



Co-funded by the
Tempus Programme
of the European Union



ҚАЗАҚСТАН РЕСПУБЛИКАСЫ БІЛІМ ЖӘНЕ ҒЫЛЫМ МИНИСТРЛІГІ
Л. Н. ГУМИЛЕВ АТЫНДАҒЫ ЕУРАЗИЯ ҰЛТТЫҚ УНИВЕРСИТЕТІ

МИНИСТЕРСТВО ОБРАЗОВАНИЯ И НАУКИ РЕСПУБЛИКИ КАЗАХСТАН
ЕВРАЗИЙСКИЙ НАЦИОНАЛЬНЫЙ УНИВЕРСИТЕТ ИМ. Л. Н. ГУМИЛЕВА

THE MINISTRY OF EDUCATION AND SCIENCES OF REPUBLIC KAZAKHSTAN
L.N. GUMILYOV EURASIAN NATIONAL UNIVERSITY

**ҚОҒАМДЫ АҚПАРАТТАНДЫРУ
VI ХАЛЫҚАРАЛЫҚ ҒЫЛЫМИ-ПРАКТИКАЛЫҚ
КОНФЕРЕНЦИЯ
ЕҢБЕКТЕРІ**

**ТРУДЫ
VI МЕЖДУНАРОДНОЙ НАУЧНО-
ПРАКТИЧЕСКОЙ КОНФЕРЕНЦИИ
ИНФОРМАТИЗАЦИЯ ОБЩЕСТВА**

**PROCEEDINGS
THE VI INTERNATIONAL SCIENTIFIC AND
PRACTICAL CONFERENCE
INFORMATIZATION OF SOCIETY**

АСТАНА 2017

УДК 002 (063)
ББК76.0
Қ 54

Редакционная коллегия: Шарипбай А.А., Казиев Г.З., Омарбекова А.С., Сауханова Ж.С., Бекманова Г.Т., Кинтонова А.Ж., Ниязова Р.С., Разахова Б.Ш., Туребаева Р.Д., Андасова Б.З.

Ученый секретарь: Омарбекова А.С.

ҚОҒАМДЫ АҚПАРАТТАНДЫРУ: VI халықар. ғыл.-практ. конф.:
Баяндамалар мен хабарламалар тезистері

ИНФОРМАТИЗАЦИЯ ОБЩЕСТВА. VI междунар. научн.-практ. конф.:
Тезисы докладов и сообщений.

ISBN 978-601-301-994-9

В сборнике трудов VI Международной научно-практической конференции «Информатизация общества» освещаются исследования в областях информатизации образования, информационной безопасности, искусственного интеллекта, латинизации казахского языка.

Сбоник издан за счет средств Международного проекта «PROMIS project № 544319-TEMPUS-1-2013-1-FR-TEMPUS-JPCR».

УДК 002 (063)
ББК76.0

ISBN 978-601-301-994-9

© Л.Н.Гумилев атындағы Еуразия ұлттық университеті, 2017
Евразийский национальный университет им. Л.Н. Гумилева, 2017

Алғысөз

Қазіргі таңда әлемдік экономикалық жүйе ақпараттық қоғамның құрылуының жаңа кезеңіне жаңа өндірістік еңбек құралымен және жаңа қоғамдық қатынастармен еніп отыр. Ақпараттық ресурс тікелей өндірістік күшке айналып, материалдық, энергетикалық және еңбек ресурстары қатарында қоғамдық процестерге белсене қатысуда. Сонымен бірге, ол материалдық өндірістің қалыптасуы мен дамуындағы маңызды факторлардың бірі бола отырып, қоғамның экономикалық потенциалын анықтайды және тауар мен капитал сипатын алып, тауар – ақша қатынасына белсене қатысады. Ақпараттық ресурсты дұрыс жасау және қолдану экономикалық жүйенің жұмысының тиімділігін және ұйымдастыру дәрежесін арттыруды қамтамасыз етеді. Бүгінгі таңда Қазақстанда ақпараттандыру басты тенденция болып саналады.

Астана қаласындағы Л.Н.Гумилев атындағы Еуразия ұлттық университетіндегі VI Халықаралық ғылыми-практикалық конференция ақпараттық қоғамның әлемдік экономиканың күшті де прогрессивті саласының бірі екендігінің дәлелі болып табылады. Байлықтың бейматериалдық түрін жасау, ақпараттық, ағартушылық, ғылыми-зерттеу және тағы басқа әлеуметтік қызметтердің жаңа мүмкіндіктеріне жол ашады. Бұл материалдар жинағы осының – айғағы.

Ұйымдастыру комитеті

Предисловие

Информатизация общества стала одной из важнейших характеристик нашего времени. Нет ни одной области человеческой деятельности, которая в той или иной степени не была связана с процессами получения и обработки информации. Информатизация - процесс, при котором создаются условия, удовлетворяющие потребностям любого человека в получении необходимой информации. Информация стала важным инструментом политики и культуры, промышленности, науки и образования. Информационный ресурс превратился в непосредственную производственную силу и наряду с материальными, энергетическими и трудовыми ресурсами активно участвует в общественных процессах и является одним из важных факторов становления и развития материального производства, определяет экономический потенциал общества, приобретает характеристики товара и капитала, активно участвует в товарно-денежных отношениях. Правильное создание и использование информационного ресурса обеспечивает повышение степени организованности и эффективность функционирования экономических систем. Информатизация становится главной тенденцией и в Казахстане.

Проведение VI Международной научно-практической конференции Информатизация общества, становится традицией в стенах Евразийского национального университета имени Л.Н.Гумилева и является доказательством того, что *информатизация общества становится одной из наиболее мощных и прогрессирующих отраслей мировой экономики* и открывает совершенно новые возможности для создания нематериальных форм богатства и оказания информационных, образовательных, научно-исследовательских и других социальных услуг. Данный сборник трудов – яркое тому свидетельство.

Оргкомитет

**МАЗМУНЫ
СОДЕРЖАНИЕ
CONTENT**

**ПЛЕНАРЛЫҚ МӘЖІЛІС БАЯНДАМАЛАРЫ
ДОКЛАДЫ ПЛЕНАРНОГО ЗАСЕДАНИЯ
PLENARY MEETING**

1	Adam Jean-Michel, Bardou Daniel. The PROMIS Project (PROfessional network of Master's degrees Informatics as a Second competence): main results.....	8
2	Шарипбай А.А. Перевод алфавита казахского языка на латиницу.....	16

СЕКЦИЯ 1

**БІЛІМ БЕРУДІ АҚПАРАТТАНДЫРУ
ИНФОРМАТИЗАЦИЯ ОБРАЗОВАНИЯ
INFORMATIZATION OF EDUCATION**

1	Jafarov A. The phases of the implementation of TEMPUS-PROMIS project at the technological university of Tajikistan.....	30
2	Merceron A., Ripphausen-Lipa H., Sauer P. Links with industry in teaching and research - experiences in the department informatics and media at the Beuth University Of Applied Sciences.....	32
3	Milosz M., Milosz E., Borys M. Efficient organization of workshops on creating teaching materials in the PROMIS project.....	36
4	Аттокуров У.Т. Сотрудничество ОШТУ с Европейскими университетами в рамках проекта TEMPUS-PROMIS.....	40
5	Бердыбек А.Е. Мобильное приложение «Студент».....	42
6	Гафарова У.А. Реализация проекта “PROMIS” В Худжандском государственном университете имени Академика Б.Гафурова.....	44
7	Жусупова Н.А., Хасенов Е.А. Создание образовательно-информационного ресурса для изучения ментальной арифметики.....	46
8	Исраилова Н.А. Результаты реализации проекта TEMPUS PROMIS “Профессиональная магистерская программа информатика как вторая компетенция” в КГТУ им. И. Разакова.....	49
9	Мухамедова Д. Реализация европейского проекта - 544319-TEMPUS-1-2013-1-FRTEMPUS-JPCR PROMIS: «Информатика как вторая компетенция» в Национальном университете Узбекистана имени Мирзо Улугбека.....	51
10	Омарбекова А.С., Бекманова Г.Т. Разработка электронных учебных пособий в рамках проекта TEMPUS PROMIS.....	56
11	Собиров Т.Р., Болтаев Т.Б. Курс переподготовки специалистов по направлению «Программирование как вторая компетенция».....	61
12	Тукеев У.А., Рахимова Д.Р., Жуманов Ж.М. Профессиональная магистратура Информатики как второй компетенции в КАЗНУ им. Аль-Фараби.....	63
13	Ускенбаева Р.К., Сатыбалдиева Р.Ж., Молдагулова А.Н., Бектемысова Г.У. О востребованности специалистов по зеленым технологиям и ИКТ.....	67
14	Ускенбаева Р.К., Ускембаева З.К. Цифровые технологии в образовании.....	70
15	Шарипбай А., Омарбекова А., Бекманова Г., Ниязова Р., Туребаева Р., Хасенов Е., Барлыбаев А. Профессиональная магистратура информатики как второй компетенции в ЕНУ имени Л.Н.Гумилева.....	73
16	Шарипбай А.А., Барлыбаев А.Б., Омарбекова А.С., Бекманова Г.Т., Разахова Б.Ш., Ниязова Р.С. Проект государственного стандарта по интеллектуальным электронным учебным изданиям (смарт-тьюторам).....	76
17	Шарипов Б.Ж. Проблемы и реальности подготовки высококвалифицированных специалистов для современной экономики.....	77
18	Шахметова Г.Б., Исайнова А.Н., Исабаева Г.Ж. Мобильное обучение.....	82

СЕКЦИЯ 2
АҚПАРАТТЫҚ ҚАУІПСІЗДІК
ИНФОРМАЦИОННАЯ БЕЗОПАСНОСТЬ
INFORMATION SECURITY

1	Барлыбаев А.Б. CYBER – онлайн-платформа для безопасного шифрования данных.....	85
2	Бекманова Г.Т., Коробов И.С., Мұқатай Г.Ж. Прямое исправление ошибок по стандарту IEEE 802.16.....	87
3	Бекманова Г.Т., Мұқатай Г.Ж. Кодер декодер Рида-Соломона.....	90
4	Зулхажав А. Автономды роботтар мен роботты жүйелерді жасаудағы навигация қауіпсіздігі.....	93
5	Оразкелдиев А.Ж. Телефон арқылы хабар алмасуды қорғау технологиялары.....	95
6	Орымбаев Б.М. Стандарты в области управления рисками: международные стандарты....	98
7	Орымбаев Б.М. Стандарты в области управления рисками: национальные стандарты.....	100
8	Өтебаева А.С., Разахова Б.Ш. Модель интегрированной системы безопасности АЗС.....	104
9	Сейтбекова Г.О., Сыдыкова М.М. Комерциялық компаниялардың киберқауіпсіздігі.....	106
10	Шағи Л.М., Абдураимова Б.К. Криптосистема на основе квантовой запутанности.....	109
11	Шарипбай А.А., Барлыбаев А.Б. Политика информационной безопасности смарт-университета.....	111
12	Ыстыбаева А. Методы оценки рисков информационной безопасности в корпоративных сетях.....	114

СЕКЦИЯ 3
ЖАСАНДЫ ЗЕРДЕ
ИСКУССТВЕННЫЙ ИНТЕЛЛЕКТ
ARTIFICIAL INTELLIGENCE

1	Altenbek G., Wu H. Research on transition-based syntactic parsing for Kazakh.....	118
2	Kantarbayev B.A., Niyazova R. Theoretical abstract model of Locomotive simulator.....	124
3	Sarsengaliyev A.A., Turebayeva R.D. Study of keyboard layout for new kazakh latin alphabet.	127
4	Zhunissova D., Omarbekova A. Visa application system.....	129
5	Актаева А., Байкенов А., Ниязова Р., Бижигитова Д. Обзор методик идентификации пользователей социальных сетей на основе онтологии.....	136
6	Әбеева И.А., Ергеш Б.Ж. Астана қаласындағы мемлекеттік ұйымдардың онтологиясын құру.....	144
7	Байбеков С.Н., Искалиева Л.С. Разработка нового метода береговой линии для секторального разделения морской территории между прибрежными государствами.....	148
8	Бекманова Г.Т., Шарипбай А.А., Омарбекова А.С., Елибаева Г.К., Ергеш Б.Ж. Автоматическая генерация отзывов пользователей на примере онтологической модели отеля.....	156
9	Бектурсынова Р.С., Ергеш Б.Ж. Уикипедияны пайдаланып қазақша мәтіндегі жалқы есімдерді тану.....	159
10	Дауренбеков Б.А. Мобильное приложение «Конвертер».....	162
11	Едигеева М.С. Использование технологии веб-дизайна в веб-приложениях.....	165
12	Ергеш Б.Ж., Бектурсынова Р.С. Білімдер базасына негізделген қазақ тіліндегі жалқы есімдерді тану.....	167
13	Инсебаева Г., Омарбекова А.С. Проектирование рекомендательной системы по планированию досуга на основе онтологической модели.....	170
14	Искакова Д.Б., Ниязова Р. Модели и методы оценки качества подсистемы мониторинга выпускников программы Болашақ.....	173
15	Искалиева Л.С., Байбеков С.Н. Разработка нового метода окружностей для установления внешних границ морских территорий.....	174
16	Исламғазиева Ә.С., Казиев Г.З. Оценка эффективности функционирования системы менеджмента качества при проведении внутреннего аудита.....	180
17	Калимолдаев М.Н., Мамырбаев О.Ж., Кейлан А. Метод проектирование системы многомодального распознавания казахской речи.....	182
18	Каманур Ү., Шәріпбай А.Ә., Бекманова Г.Т., Жеткенбай Л., Тоқтар Ө. Қазақ қытай тілдері бойынша машиналық аударманың қадамдары.....	197

19	Күзенбаев Б.А., Шарипбай А.А., Ниязова Р.С., Барлыбаев А.Б. Онтологияларды салыстыра отырып автоматтандырылған басқару жүйесінің сапасын анықтау.....	201
20	Қажымұхан Д.А., Бурибаева А.К. Facebook әлеуметтік желісіндегі жазбаларды стильдік ерекшеліктерін ескеріп сентимент талдау.....	205
21	Құлымжанова Ж.Т., Жукабаева Т.К. Нақты уақыттағы компьютерлік көру жүйесінің мәліметтерді өңдеу әдістерін және алгоритмдерін зерттеу.....	208
22	Мадрахимов Ш.Ф. Вычисления обоженных оценок по определенным наборам номинальных признаков и их интерпретация.....	211
23	Молдагулова А.Н., Ускенбаева Р.К., Сатыбалдиева Р.Ж. К вопросу обработки неструктурированных данных.....	216
24	Ниязова Р.С., Абжанова А. А. Разработка интеллектуального электронного учебника по дисциплине «Программная инженерия».....	220
25	Нургазина Г.Ш., Шарипбай А., Омарбекова А.С., Зулхажав А. Интеллектуальная система «IT-TEST.KZ» по алгебре на казахском языке	222
26	Нұрланова Ж.Н., Сауханова Ж.С. Применение классификационных методов интеллектуального анализа данных для прогнозирования сердечных заболеваний.....	225
27	Омарбекова А.С., Нургазина Г. Формирование вопросов и ответов на основе базы знаний.....	228
28	Оразбаева К.О., Өтенова Б.Е., Оразбаев Б.Б. Мұнай өңдеу өндірісінің технологиялық жүйелерін модельдеу және оптимизациялау арқылы тиімді шешім қабылдау компьютерлік жүйесі.....	231
29	Ускенбаева Р.К., Куандыков А.А. Организация и управления процессов автоматизации выполнения сложных операций совокупностью интеллектуальных роботов.....	234
30	Санкибаев А.Т., Хасенов Е.А. Применение метода самоорганизующихся карт Кохонена для кластеризации данных в задаче выявления лжепредпринимательской деятельности... ..	238
31	Сауханова М.С., Сауханова Ж.С., Қалдыбек А.Ә. RAPID MINER ортасында мәтіндерді талдау.....	243
32	Тұрмағанбетова Ш.Қ. Салық саласы бойынша қазақ тілінде интеллектуальды сұрақ-жауап жүйесін құру.....	246
33	Ускенбаева Р.К., Мукажанов Н.К. Ашық үкімет деректерін өңдеуге арналған ақпараттық-аналитикалық жүйе архитектурасын әзірлеу.....	253
34	Ускенбаева Р.К., Куандыков А.А. Методы обработки заявок в системе э-коммерции.....	256
35	Шалбаева Ж.А. Дорожная карта методологии ASAP.....	259

ПЛЕНАРЛЫҚ МӘЖІЛІС БАЯНДАМАЛАРЫ
ДОКЛАДЫ ПЛЕНАРНОГО ЗАСЕДАНИЯ
PLENARY MEETING

JEAN-MICHEL ADAM, DANIEL BARDOU

**THE PROMIS PROJECT (PROFESSIONAL NETWORK OF MASTER'S DEGREES
INFORMATICS AS A SECOND COMPETENCE): MAIN RESULTS**

(Université Grenoble Alpes (UGA), Grenoble, FR)

Abstract. The PROMIS Project is an ongoing Tempus project that started in December 2013 and is close to this end in June 2017. It involves a consortium of 10 Central Asian universities, 5 European universities and 3 European companies. This paper presents the main results of the project.

1 Introduction

The PROMIS Project (PROfessional network of Master's degrees Informatics as a Second competence) is a European Tempus project funded by Education, Audiovisual and Culture Executive Agency (EACEA). PROMIS primarily aims at extending in Central Asia (CA) an existing Master's degrees network for training "Informatics as a Second Competence" (ISC) that was created in a former Tempus project called ERAMIS [1]. The project also aims at improving the vocational features of the Master's degree network by strengthening the relationships between the universities and the local companies, and adapting the teaching process to the social needs of the students [2, 3, 4].

PROMIS project encourages a close cooperation of local IT (Information Technology) companies in the teaching process and studies organization (participation in steering committees, development of internships, setting up courses taught by companies' fellows). This gives the companies the opportunity to express their real needs in terms of required skills for their future recruitments and thus improve the employability of graduated students. For the 2nd year of study, PROMIS will develop apprenticeships: students will be able to earn money to live, and to get their first professional experience by working part-time in an IT company and completing their studies at the university.

PROMIS proposes also to improve the efficiency of the teaching process by setting up reverse teaching also called "flipped classroom" or "backward classroom" [5]. In such a pedagogical approach, the teacher does not explain new concepts in front of a silent classroom that is only listening. On the contrary, students first study the topics by themselves typically watching videos prepared by the teacher. In class, students discuss difficult aspects and apply the knowledge by solving problems. Time is allocated differently than in traditional teaching. Almost no time is spent in class in explaining new concepts. Rather time is spent to deepen critical aspects and put the knowledge into practice. A challenge in adopting reverse classroom or peer interaction is the availability of good teaching material for self-study. In particular, the teaching material needs to include simple exercises to support active reading or active watching of the students. Another challenge is to develop activities and case studies adequate for group discussion and group work during class time.

The PROMIS Project involves a large consortium of 18 partners and an external expert in the implementation of the Bologna process in Central Asia. The project was initially running from December 2013 to December 2016, EACEA granted the project an extension until the end of June 2017 for completing the unfinished tasks.

The European (EU) consortium is composed of 5 universities: Université Grenoble Alpes (Grenoble, France), Beuth Hochschule fuer Technik Berlin (Berlin, Germany), Savonia University of Applied Sciences (Kuopio, Finland), Kaunas University of Technology (Kaunas, Lithuania), and Politechnika Lubelska (Lublin, Poland), and 3 private companies: SYMETRIX (Grenoble, France), eLeDia (Berlin, Germany) and Ilmi Solutions Oy (Kuopio, Finland).

The beneficiaries universities are: Kyrgyz State Technical University named I. Razzakov (Bishkek, KG), Osh Technological University (Osh, KG), Al-Farabi Kazakh National University (Almaty, KZ), Eurasian National University (Astana, KZ), National University of Uzbekistan

(Tashkent, UZ), Bukhara State University (Bukhara, UZ), Technological University of Tajikistan (Dushanbe, TJ), Khujand State University named B.Gafurov (Khujand, TJ), Turkmen State Institute of Transport and communication (Ashgabat, TM), and Turkmen State Institute of Economics and Management (Ashgabat, TM). The first and third universities in this list were also former ERAMIS partners and have already created an ISC Master's degree, they are part of the PROMIS Project to set an example and help the other beneficiary partners.

The rest of this paper describe the main results of the PROMIS Project : section 2 describe the masters's programs creation, including professionalization, internationalization and quality assurance aspects ; section 3 focuses on pedagogical aspects ; section 4 is the conclusion of this paper.

2 Main Results

2.1 Creation of new ISC Master's Degrees in each Central Asia country

Each partner did the needs analysis after the kick-off meeting. It was composed of two surveys made in every region of Central Asia involved in the creation of the Master's degree ISC: the first survey was conducted among students and the second one among the local IT companies. These surveys clearly show: (1) the interest in learning computer development from the students trained in other disciplines, (2) the need expressed by companies in this type of training, so far non-existent in these regions. The skills and knowledge needs, expressed by companies, helped the teaching teams to define directions for their Master's programs. Some needs may require the creation of specific (elective) courses to better prepare the students for the jobs available in their region.

Each partner university drafted its accreditation file, respecting national constraints governing Master's programs, and placing the common core courses in the same semester in order to promote mobility within the network. In October 2016, six ISC Master's programs were created in the partner universities: in one Uzbek university (Tashkent), one Tajik university (Khujand) and two Turkmen institutes in Ashgabat, one in Kyrgyzstan (Osh) and one in Kazakhstan (Astana). The creation followed the regular procedure of creating a Master's program in the country. Two partner universities did not open the ISC master in October 2016: Technological University of Tajikistan who has got the accreditation, but have defined too selective criteria to be able to enrol enough students, and Bukhara State University who did not do the accreditation file on time. Table 1 shows the number of students enrolled in each master's program of the partners' universities.

For what concerns the curriculum for the Master's program, we need a common organization of the studies because we want to operate within a network of Masters. We identified a set of 10 core courses [3] that have to be part of each Master's program within the network. The core courses are organized in the same semester for all the programs in the network, so that local specific differences, from one program to another, does not prevent student mobility. The European Credit Transfer System (ECTS) is used. The number of ECTS allocated to a common core course is the same for all the programs of the network. In addition to this common courses kernel, each university proposes elective courses matching specific local requirements.

The syllabuses for the core courses, inherited from the previous ERAMIS Project, were reviewed and updated, taking into account the evolution of computer science since 2011. These revisions were made in the pedagogical workshops in European universities to jointly develop teaching materials (see section 3). Each syllabus has been specifically described in English, respecting the standard that we set in the ERAMIS project.

All the newly created Master's programs were scheduled to start in September 2015, in accordance with the work plan of the project. The existing Master's programs (KSTU Bishkek and Almaty KazNU) have also been refined so that they are closer to the other programs in the network. Each partner university has appointed a pedagogic responsible of the Master's program, and created a teaching team (6 to 15 persons) which is mainly composed of regular teachers of the university but also include sometimes teachers from local private companies.

The local teams have produced communication materials for students to inform them of this new training program, and proceeded to the recruitment of the first students. The selection procedure of the candidates for future Masters' programs was discussed. In each country, with the exception of Kyrgyzstan, the government imposes rules for the selection of the students: language tests and knowledge tests. We propose to complete this procedure by an interview with the candidate, to give him/her the opportunity to present his/her professional project and motivations for IT. In Kyrgyzstan, the rules seem to be more flexible: each university can set its own selection process, for example, it can be based on an English test, an interview and eventually a written exam.

The partners have acquired the equipment for setting up the ISC master's program. The procedure for Uzbekistan was longer because the Ministry of Education had to approve the requested equipment.

Table 1: 1st and 2nd year ISC Master's program in Central Asian Universities

			Uzbekistan		Kirghizstan		Kazakhstan		Tadjikistan		Turkmenistan		
			NUU	BSU	KSTU	OshTU	KazNU	ENU	TUT	HGU	TSIEM	TSITC	Total
2015/2016	1 st year	# of registered students	5	n/a	3	11	8	4	n/a	21	15	12	79
		# of long life learning students	0	n/a	0	9	3	4	n/a	4	0	0	20
		# of dropouts	1	n/a	0	6	3		n/a	3	0	1	14
		# of passing students	4	n/a	3	5	5	4	n/a	18	15	11	65
2016/2017	1 st year	# of registered students	6	7	11	19	5	8	10	10	n/a	n/a	78
		# of long life learning students	1	7	0	7	3	1	0	2	n/a	n/a	21
		# of dropouts	0	3	1	0	0	0	2	0	n/a	n/a	6
	2 nd year	# of registered students	4		3	5	5	4		12	n/a	n/a	33
		# of long life learning students	0		0	3	2	4		3	n/a	n/a	12

2.2 Professionalization of the Master's program

In each partner university, one person has been appointed to take charge of relations with companies. This person keeps the companies informed on the possible cooperation with the university.

Development of courses or parts of courses by company employees were encouraged. Some partners have found people from the local businesses to participate in lessons (Turkmen, Tajiks and Kazakhs partners).

We organized conferences on university-company relationships to which were invited representatives of local businesses. Our goal is to encourage university-company collaboration and to encourage these companies to enroll students for internships. Conferences were organized in Tashkent (September 2014), Bukhara (May 2015), Osh (May 2015), Bishkek (October 2015), and Dushanbe (November 2016).

The ISC Master's program is also designed for long-life learning students: the program provides a good basis of knowledge and skills in software development. Therefore, it is possible to enroll, on a full-time basis, for one year or a shorter duration, employees that can benefit from an education leave. On the other hand, as in Central Asia the majority of students work in parallel to their studies in order to support themselves, lessons are usually held in the afternoon and evening. Such an organization also can accommodate part-time employees who wish to attend the classes of the ISC Master's program. Table 1 shows, for each school year the number of long-life learning students enrolled in each partner university. We can see that this population of students represents 25% of the students in year 2015/2016, and about 30% in year 2016/2017. These figures show that the ICS Master's program corresponds to a real need for people in active employment.

2.3 Internationalization of the Master's program

In the project, we have put 25% as the target quantity of lessons delivered in English. In each teaching team, there are colleagues whose English level is still insufficient to deliver lessons in English and the project funded training in English for all partners, in accordance with the work plan. Nevertheless, most of the partners should be able to reach this target of 25% or at least to get very close to it. Anyway, the European partners have provided teaching material written in English, on the Moodle platform of the project. This material helps the Central Asian partners to prepare and deliver lessons in English.

2.4 Quality assurance

The consortium appointed the quality manager for the project, on a proposal from the Savonia University of Kuopio. Our appointed colleague is the Vice Rector in charge of the quality assurance at the University of Savonia. He is not a computer scientist and thus can have a fresh external and critical look at the project.

We also created a "Monitoring Committee". This committee is composed of one academic representative from each European university, including the quality manager of the project. The Committee visited all the partner Universities during spring 2015 to ensure that the partners were ready to start the Master's programs in September 2015. During these visits, the Committee carried out the following points with the teaching and administrative staff:

- Examination of the future Master's degree program.
- Examination of the contents of all the Master's degree program courses.
- Review of the communication material to inform students about the new Master's degree program: posters, booklets, leaflets, web pages.
- Discussion about the selection procedure for the future Master's students. Some differences were noticed among the different rules for setting up this procedure in different countries.
- Ensuring that each University has an online e-learning platform (Moodle or similar software) and is able to manage it.

The Committee gave advices and assistance to the visited partners and found solutions to almost all of the arisen problems.

Besides the advances presented above, the PROMIS project partners did a significant effort on innovative pedagogical issues during the year 2015. We present these issues in the next section.

3 Pedagogical issues

Most significant work related to pedagogy took place during the year 2015 with the organization of workshops devoted to the teaching of core courses. These workshops initiated the activity of producing reference learning and teaching materials that was continued through the year 2016. Seminars targeting students and teachers were also conducted by European teachers in Central Asia partner universities from March 2016 to May 2017. Finally student mobilities towards Europe were organized in February and April 2017. These activities are detailed in this section.

3.1 Organization of pedagogical workshops in Europe

In most of the Tempus projects "train the teacher" seminars are organized: "expert" teachers that are specialized in some course or technology are asked to train other "beginner" teachers so that they can teach the course themselves afterwards. Because this kind of seminars very often turns into

a one-directional transfer of knowledge and experience in which the emitter is the most active and the receivers are pretty passive, we preferred to organize collaborative workshops for discussing and exchanging on a specific course and actively involve all the participants [6].

Each EU partner university of the PROMIS consortium organized two workshops during the spring 2015 (see Table 1). Each workshop was devoted to one of the ten PROMIS core courses in order to cover all of them. A workshop gathered teachers from all the network that will be teaching the related core course in each Master's program. The workshops were highly motivated by our objective to implement reverse teaching, as high quality pedagogical material is required by this teaching paradigm, but they also serve in raising the overall quality of the Master's program. Indeed, a small community of the several teachers that will teach the same core course in the several programs of the Master's network was created: they met each other and are no more isolated and can continue to collaborate online.

Table 2. Summary of the pedagogical workshops (ordered by date of venue)

Host university, place and time	Workshop topic (related core course)	Number of participating CA teachers	Number of participating EU teachers (hosts included)
Beuth University of Applied Sciences, Berlin, DE (16/02-20/02/2015)	Programming	11	4
	Algorithms and Data Structures	10	4
Savonia University of Applied Sciences, Kuopio, FI (11/05-15/05/2015)	Operating Systems and Networks	10	4
	Web Development	7	2
Pierre Mendès France University, Grenoble, FR (01/06-05/06/2015)	Human Machine Interaction	9	3
	Mathematics for Computer Science	9	2
Kaunas Technical University, Kaunas, LT (15/06-19/06/2015)	Data Mining and Data Warehouses	8	2
	Databases	7	5
Lublin University of Technology, Lublin, PL (22/06-26/06/2015)	Project Management	8	3
	Software Engineering	7	3
Total		86	34

3.2 Development and sharing of learning and teaching materials and other educational resources in English, Russian and local languages

To implement reverse teaching it is necessary to have educational materials of good quality so that the students can learn the course on their own. During the pedagogical workshops that were organized in 2015 teachers shared their own teaching materials and methods, and started the production of a set of common reference educational materials in English and Russian. For each of the 10 core courses, a Central Asian colleague has been appointed for monitoring and coordinating the production of educational materials that has been continued throughout the year 2016.

3.3 Creation of an online repository for sharing educational materials

The University Grenoble-Alpes supports a Moodle platform for the project. All project teachers have access to the platform to upload or download documents. The platform is used as a virtual place for gathering and sharing all the educational materials produced in English and Russian by all the teachers involved in the workshops. The EU partners actively participate, as it's an opportunity for them to develop or improve documents for core courses in English as well. The existence of these documents will promote the internationalization of the Master's programs and foster the creation of double diploma between the partners. Each produced teaching material will be available on the online platform for sharing purposes but also as an evidence of the salary paid to the teachers that created it.

3.4 Implementation of reverse teaching

At the beginning of the project, very few teachers (European or Central Asian) had experience with reverse teaching. The availability of comprehensive teaching materials provided some help in setting up the teaching activities even though a full reverse teaching approach could not be implemented in every case. Reverse teaching was nevertheless strongly encouraged during the project and was experimented in every CA partner university. Indeed for reverse teaching to be effective, assessment tests or specific activities must be used to control that students have correctly acquired by themselves the key concepts of the course.

During the ERAMIS project tests had been developed to assess students in some core courses (approximately 800 multiple choice questions). These questions were also available for serving in the development of such tests. Furthermore, the universities of Grenoble and Lublin collaborated on the construction of a MOOC in English for teaching Human-Computer Interaction in the context of a “Leonardo da Vinci” European project. All the documents that were produced for this MOOC could straightforward be used to implement reverse teaching in this core course.

3.5 Organization of seminars in Central Asia partner universities

From May 2016 to May 2017, 31 seminars were organized in CA partner universities for training both the students and the teachers (see Table 3). EU partner universities gave training on 20 different topics most of the time in English and sometimes in Russian. In most CA partner universities, 4 seminars were organized. Only 3 seminars were organized in Turkmenistan where only TSIEM chose 3 topics. As both TUT in Dushanbe and BSU in Bukhara had no program open at the time seminars were first scheduled, 2 seminars were lately scheduled for them after the opening of their program in fall 2016.

Table 3. List of seminars organized in Central Asia

Place	Month	Seminar topic	EU Partner	Hours	Language
KazNU Almaty KZ	May 2016	Advanced Control Automation	SUAS, Kuopio	10	English
		Internet of Things	SUAS, Kuopio	10	English
	Sept. 2016	Artificial Intelligence	UGA, Grenoble	15	English
Introduction to game development for smartphones and computers		KTU, Kaunas	15	English- Russian	
ENU Astana KZ	May 2016	Advanced Control Automation	SUAS, Kuopio	10	English
		Big Data	SUAS, Kuopio	10	English
	Sept. 2016	Data Mining	BUAS, Berlin	15	English
		Storing and querying data from the web: Introduction to xml and semantic technologies	BUAS, Berlin	10	English
KSTU Bishkek KG	Sept. 2016	Software project management	LUT, Lublin	15	Russian-English
		Generic programming in C++	LUT, Lublin	15	Russian-English
	Oct. 2016	Introduction to game development for smartphones and computers	KTU, Kaunas	15	English-Russian
		Graphics in WEB and game development	KTU, Kaunas	10	English-Russian
OshTU Osh KG	Oct. 2016	Generic programming in C++	LUT, Lublin	15	Russian-English
		Software design using UML	LUT, Lublin	15	Russian-English
		Storing and querying data from the web: Introduction to xml and semantic technologies	BUAS, Berlin	10	English
		Methodology for Algorithms and data structures teaching	UGA, Grenoble	10	English
BSU	Apr. 2017	OO design and UML	BUAS, Berlin	12	English

Bukhara UZ	May 2017	Aspect-Oriented Programming with AspectJ	UGA, Grenoble	12	English
NUU Tashkent UZ	Sept. 2016	Data Mining	BUAS, Berlin	15	English
		Storing and querying data from the web: Introduction to xml and semantic technologies	BUAS, Berlin	10	English
	Nov. 2016	Design Patterns in Java	UGA, Grenoble	15	English
	Feb. 2017	Big Data	SUAS, Kuopio	10	English
HGU Khujand TJ	June 2016	Human-Computer Interaction - Theory	LUT, Lublin	15	Russian-English
		Human-Computer Interaction - Practice	LUT, Lublin	15	English
		Data Structure and Algorithms	BUAS, Berlin	10	English
		Methodology for Algorithms and data structures teaching	UGA, Grenoble	10	English
TUT Dushanbe TJ	May 2017	Software project management	LUT, Lublin	15	Russian-English
		Generic programming in C++	LUT, Lublin	15	Russian-English
TSIEM Ashgabat TM	Oct. 2016	Human Machine Interaction	UGA, Grenoble	15	English
		Business process modeling with BPMN	KTU, Kaunas	10	English-Russian
		Graphical user interface as the key to higher usability of software systems	KTU, Kaunas	10	English-Russian

3.6 Student mobilities towards Europe

In order to improve the professional skills of the students, and by this way show our colleagues from Central Asia how to better prepare the students to the professional life, 2 European Universities of the consortium organized a specific training for students: Université Grenoble-Alpes and Beuth Hochschule fuer Technik Berlin. The training were based on Java programming language, which is more and more used in Web and mobile development. This language was rarely taught in the Central Asian partner universities as the C++ language has been chosen in most of them.

Each training for 2 weeks long and consisted of a blend of fundamental classes and the development of a small software case study in which all the technology taught in classes was right away put into practice. Each Central Asian partner selected 2 students for participation to these trainings in Europe. BUAS Berlin received students from ENU Astana, KSTU Bishkek, OshTU Osh, and TUT Dushanbe in February 2017. UGA Grenoble received students from KazNU Almaty, NUU Tashkent, BSU Bukhara, and HGU Khujand end of April 2017. The Turkmen universities did not participate in these students mobilities.

3.7 Evaluation of the network's students

For each common course, we have developed tests composed of multiple-choice questions (MCQ). These questions, written in English, have been translated in Russian language. The tests have been set up on the Moodle platform of the project. A login on this platform was created for each student of the created master's programs. The realization of the tests by the students is in progress. The collected data about their results will be analysed in the next months.

4 Conclusion and perspectives

We presented in this paper the results of the PROMIS Project. In all the partner universities a ISC master's program is created and prepare students to the professional life, providing them with knowledge and skills in computer science. Their ability to combine the first competence acquired during the Bachelor program, with the application development skill is an asset to find a job. We

hope that, after their internship or master thesis defense, all the students of our master's programs will find a job.

These master's programs creation was a challenge, and not so easy to create everywhere. The training of the colleagues of the partner universities in European universities was helpful, and was completed by many seminars at the partner Universities. The notion of professional master's program is not easily understood in all Central Asian countries. For some Ministry of Education like in Uzbekistan, the masters' programs are only a preparation to research activities. The notion of professional master's program for high-level engineers' preparation, like existing programs in European Universities is difficult to be approved. Kirghizstan, Tajikistan and Kazakhstan have now approved this type of training.

The most difficulties were the communication problems with our Turkmen partners: it was difficult to have quick replies to our e-mails, and it is very difficult for us to get a visa and authorization to enter in the country, and our colleagues had the same difficulties in obtaining permission to leave the country. Always the same colleagues were present to the coordination meetings and the seminars organized in the European Universities.

The involvement of companies has proved to be complex. Only Eledia really contributed to the project by putting its management skills of Moodle platforms at the service of the project. Symetrix was involved at the beginning of the project, participating at distance to our conferences about university-company relationships. This point is important to improve the vocational aspect of the courses: all our Central Asian partners need to improve their relations with the companies in their employment area.

Agreements and partnerships between partners of PROMIS project in order to create double diplomas, make easier students mobility and promote scientific cooperation were signed or are in progress. That shows that our cooperation is not ended with this project.

ACKNOWLEDGEMENTS

We thank all colleagues and students in partner universities, who are not formally involved in the PROMIS project, but who participated in organizing or conducting all events related to this project. The work described in this paper has been carried out within the framework of the Tempus program (Ref. No. 544319-TEMPUS-1-2013-1-FR-TEMPUS-JPCR), funded by Education, Audio-visual and Culture Executive Agency (EACEA) that we thank.

References

- 1 Adam, J.-M., Merceron, A., Lujan-Mora, S., Milosz, M. and Toppinen, A. (2012). European-Russian-Central Asian Network of Master's degree Informatics as a Second Competence, in Proceedings of EDUCON 2012: The IEEE International Education Engineering Conference, Marrakesh, Morocco, April 2012, IEEE, pp. 722-729, ISBN: 978-1-4673-1456-5.
- 2 Adam, J.-M. and Bardou, D. (2014). From ERAMIS to PROMIS: Extending and Improving a Master's Degrees Network in Central Asia. In proceedings of EDULEARN14 6th International Conference on Education and New Learning Technologies, pp. 2120-2129, Barcelona, Spain, 7th – 9th of July 2014.
- 3 Bardou, D., Adam, J.M. and Goloubtsova, A. (2014). Extending an international network of professional Master's degrees Informatics as a Second Competence. In Tempus: 20 years of programme activities in Uzbekistan, prepared by the National Erasmus+ Office in Uzbekistan, pp. 186-193.
- 4 The PROMIS Project website. <http://promisproject.kaznu.kz/>
- 5 Abeysekera, L. and Dawson, P. (2015). Motivation and cognitive load in the flipped classroom: definition, rationale and a call for research. In Higher Education Research & Development 34(1). pp. 1-14.
- 6 Bardou, D., Adam, J.-M., Milosz, M., Merceron, A., Ripphausen-Lipa, H., Kapocius, K. and Toppinen, A. (2015). Pedagogy sharing in the PROMIS project. In proceedings of the IEEE Global Engineering Education Conference, EDUCON 2015, Tallinn, Estonia, March 2015. IEEE 2015, pp. 587-593, ISBN 978-1-4799-1908-6.

А. А. ШАРИПБАЙ

ПЕРЕВОД АЛФАВИТА КАЗАХСКОГО ЯЗЫКА НА ЛАТИНИЦУ

(НИИ «Искусственный интеллект» ЕНУ им. Л.Н.Гумилева, Астана, Казахстан)

1. Актуальность.

В своей программной статье «Ориентир на будущее: духовное обновление» Президент Казахстана Н.А. Назарбаев предложил *Проект перехода на латинский алфавит*: «...мы должны начинать работы по постепенному переводу казахского языка на латинский алфавит. ...» [1].

2. Исторические сведения о казахской письменности.

Казахская письменность берет свое начало с древнетюркских рунических письменностей VI–X вв. н.э. В связи с распространением ислама в тюркском мире с IX по XX век н.э. использовался арабский алфавит. В 1912 году с арабского алфавита исключили все арабские буквы, обозначающие не казахские звуки и перешли на новый алфавит казахского языка, который содержал 24 буквы на основе арабской графики и 1 апостроф. Этим алфавитом в нашей стране пользовались до 1929 года, а казахи, живущие в других странах пользуются им и по сей день, добавив еще 3 буквы для обозначения согласных (*vy*), (*fy*) и (*xy*). В 1929 году был принят алфавит на основе латиницы, а в 1940 году – алфавит на основе кириллицы. Были добавлены 10 новых букв: две буквы **Ьь** и **Ъъ** для обозначения мягкости и твердости согласных, 8 букв **Ёё, Ии, Уу, Цц, Чч, Щщ, Ээ, Юю, Яя** для обозначения не свойственных к казахскому языку русских звуков. Этот алфавит действует до сих пор [2].

3. Адаптация латинского алфавита к другим языкам.

Латинский алфавит (латиница) произошел от западного варианта греческого алфавита и возник в VII веке до н. э. и содержал 21 прописную букву, позже в него включили еще 5 букв и строчные буквы. Затем 26 буквенный латинский алфавит зафиксирован Международной организацией по стандартизации – International Organization for Standardization (ISO) как «*базовый латинский алфавит*». Он совпадает с современным английским алфавитом и является основой письменности большинства других языков [3].

В настоящее время 26 буквенный базовый латинский алфавит де-факто стал «алфавитом международного общения», так как все знакомятся с ним еще со школы. Неоднократно выдвигалась идея перевода письменностей всех языков на этот алфавит. Однако в некоторых языках количество звуков больше, чем количество букв в этом алфавите, и в зависимости от их фонетических систем некоторые буквы должны иметь другие значения. Поэтому для таких языков приходится адаптировать базовый латинский алфавит. В процессе его адаптации к другим языкам возникла проблема обозначения звуков, для которых отсутствовали буквы в нем. Для решения этой проблемы использовались следующие методы адаптации алфавита:

1. Использование цепочек букв (диграфов, триграфов, тетраграфов) для написания одного звука: например, для обозначения звуков (ш) - [ʃ] и (ч) - [tʃ] в английском языке используются диграфы *sh* и *ch*, а в немецком языке – триграф *sch* и тетраграф *tsch*, соответственно.

2. Использование положительной диакритики: например, во многих языках использовали буквы типа *á, ä, ĝ* полученные путем добавления точек, птичек и других диакритических знаков существующим буквам алфавита;

3. Использование отрицательной диакритики: например в турецком языке звук (ы) - [ɯ] обозначается буквой *ı*, полученной путем удаления точки от буквы базового латинского алфавита *i*;

4. Введение новых букв, модифицируя имеющиеся в алфавите буквы: например, модифицируя имеющиеся в алфавите буквы *c* и *s* для обозначения звуков (си) - [si:] и (эс) - [es] в английском языке получим буквы *ç* и *ş* для обозначения звуков (че) - [tʃ] и (ше) - [ʃ] в турецком языке, соответственно.

5. Введение новых букв за счет взаимственных букв из других алфавитов: например, в некоторых языках буквы *ŋ* и *ɝ* заимствованы из греческого алфавита, а буква *þ* – из рунической азбуки;

6. Использование лигатуры, полученной из сращения двух букв: например, вышеуказанное образование *W* из двойной *VV* в поздней латыни, английском, немецком и польском языках или *æ* из *a* и *e* в датском, норвежском и исландском языках.

7. Использование нескольких букв и диакритического знака для обозначения одного звука, например, в бретонском языке звук «х» обозначается диграфом *c* и *h* с апострофом ', т.е. цепочкой *c'h*.

8. Изменение значений отдельных букв в базовом латинском алфавите: например, буква *x* в английском языке обозначает звук (экс) - [eks], а в португальском языке обозначает звук (*шшш*) - [ʃiʃ].

4. Опыт тюркских стран.

Современный турецкий алфавит на основе латинской графики был принят в 1928 году и содержит 29 букв *Aa, Bb, Cc, Çç, Dd, Ee, Ff, Gg, Ğğ, Hh, Iı, İi, Jj, Kk, Ll, Mm, Nn, Oo, Öö, P, Rr, Ss, Şş, Tt, Uu, Üü, Vv, Yy, Zz*, из них 6 букв *Çç, Ğğ, İı, Öö, Şş, Üü* не входят в базовый латинский алфавит и 3 буквы базового латинского алфавита *Qq, Xx, Ww* не используются [4].

В начале 1990-х годов, в целях упрощения чтения Туркменской орфографии, туркменская письменность 3 раза была переведена на латинскую графическую основу. Последний вариант 1995 года содержит 30 букв *Aa, Bb, Çç, Dd, Ee, Ää, Ff, Gg, Hh, Iı, Jj, Žž, Kk, Ll, Mm, Nn, Ńñ, Oo, Öö, P, Rr, Ss, Şş, Tt, Uu, Üü, Ww, Yy, Ýý, Zz*, из них 8 букв *Çç, Ää, Žž, Ńñ, Öö, Şş, Üü, Ýý* используются для обозначения специфических туркменских звуков и не входят в базовый латинский алфавит, а 4 буквы базового латинского алфавита *Cc, Qq, Vv, Xx* не используются [5].

Узбекский алфавит на основе латиницы в 1990-х годах был принят дважды. Последний вариант 1995 года содержит 24 буквы базового латинского алфавита и 1 знак ' - аксантик, при этом 2 буквы базового латинского алфавита *Cc* и *Ww* не используются и ранее используемые 5 диакритические буквы *Çç, Ńñ, Şş, Ğğ, Öö* заменены на 5 диграфов *Ch ch, Ng ng, Sh sh, G' g', O' o'*, а специфическая буква *Jj* заменена на букву *Jj* [6].

Перевод азербайджанскую письменность на латинский алфавит начали в 1992 году и завершили в 2001 году. В алфавите имеются 32 буквы *Aa, Bb, Cc, Çç, Dd, Ee, Əə, Ff, Gg, Ğğ, Hh, Xx, Iı, İi, Jj, Kk, Qq, Ll, Mm, Nn, Oo, Öö, P, Rr, Ss, Şş, Tt, Uu, Üü, Vv, Yy, Zz*, из них 8 букв *Çç, Əə, Ğğ, Iı, İı, Öö, Şş, Üü* не входят в базовый латинский алфавит и 2 буквы *Cc* и *Ww* не используются [7].

По алфавитам тюркских языков можно сделать следующие выводы:

– в указанных алфавитах тюркских языков, кроме узбекского, некоторые буквы не входят в состав базового латинского алфавита и порядок следования букв в них не соответствует порядку их следования в базовом латинском алфавите, что потребует **немалых затрат на разработку программ для сортировки и поиска данных** при каждой смене поколений компьютеров и программного обеспечения, так как для базового латинского алфавита они встраиваются автоматически.

– в узбекском алфавите в качестве второй пары диграфа используется знак ' - аксантик (апостроф), который в языках программирования применяется как специальный знак для выделения в тексте отдельных его фрагментов. В случае его использования при компьютерной обработке узбекских текстов звуки [ɔ] и [ʏ], обозначающиеся как *O'o'* и *G'g'* соответственно, **не будут распознаны**, что нежелательно.

5. Критерий перевода казахского алфавита на латиницу.

Используя некоторые сочетания вышеуказанных методов адаптации латинского алфавита можно построить алфавит казахского языка на основе латинской графики. Для этого мы должны определить критерий и задачи перевода казахской письменности на новый алфавит, а также разработать правила избавления от ненужных казахскому языку русских звуков и разработать алгоритм замены кириллических букв на буквы базового латинского алфавита [8].

При реализации этого перевода обратим внимание на обстоятельства:

1. В эпоху информатизации общества во всем мире с помощью информационных коммуникационных технологий (ИКТ) создаются и обрабатываются цифровые ресурсы массового потребления очень большого объема (Big Data), которые требуют прямо пропорциональных по объему интеллектуальных, трудовых и финансовых затрат.

2. Все производимые в Мире компьютеры изначально поддерживают и будут поддерживать 26 буквенный базовый латинский алфавит, при этом адаптация этих компьютеров к национальным алфавитам, имеющим буквы, не входящие в базовый латинский алфавит, требует разработки и применения специфических шрифтов, драйверов, программ сортировки и поиска информации, которые ориентированы на эти алфавиты и требуют немалых дополнительных интеллектуальных и финансовых затрат.

3. С учетом тенденции быстрого роста объема цифровых ресурсов и расширения сферы применения ИКТ мировой общественностью предлагается использовать в глобальном информационном пространстве единый алфавит, совпадающий с базовым латинским алфавитом, в целях упрощения и снижения затрат для создания и обработки этих ресурсов, повсеместно создаются стандарты для транслитерации букв национальных алфавитов на буквы базового латинского алфавита.

4. Государственный статус казахского языка обязывает к тому, что в нашей стране цифровые ресурсы должны создаваться и обрабатываться на казахском языке.

5. Существующие информационные ресурсы на казахском языке, находящиеся в различных фондах на бумажных носителях в алфавите на основе арабицы, латиницы-жаналип и кириллицы должны оцифровываться после их отбора специальной комиссией.

6. Существующие цифровые информационные ресурсы в алфавите на основе арабицы, латиницы-жаналип и кириллицы должны конвертироваться на базовый латинский алфавит.

7. Перевод казахской письменности на базовый латинский алфавит значительно упростит изучение и улучшит развитие казахского языка, так как все грамотные люди знают этого алфавита со школьных лет;

Учитывая опыт адаптации базового латинского к другим языкам и перспективы создания и развития информационных ресурсов и технологий на казахском языке предлагаются следующий **критерий перевода казахской письменности на новый алфавит:**

1. Новый алфавит должен создаваться только на основе звуковой системы казахского языка. Это позволит построить строгие орфоэпические и орфографические нормы в соответствии с внутренними законами казахского языка. Так как в новом алфавите будут отсутствовать буквы, обозначающие звуков русского языка, которые противоречат фонетике казахского языка.

2. Новый алфавит должен образовываться только из букв базового латинского алфавита. Это позволит людям облегчить письменное общение с друг другом, создание и обработку цифровых информационных ресурсов на казахском языке с помощью электронных устройств, без разработки и установки дополнительных шрифтов и драйверов, а также программ для сортировки и поиска данных. Так как они для базового латинского алфавита автоматически встроены во всех производимых электронных устройствах.

3. Новый алфавит может менять звуковое значение некоторых букв базового латинского алфавита для обозначения звуков казахского языка. Это позволит в новом алфавите для некоторых букв определить новые звуковые значения из звуковой системы

казахского языка. Но при их подборе необходимо учитывать близость звучания казахского звука со звучанием звука – оригинала, обозначение которого включен в базовый латинский алфавит. Так как наличие близости облегчит подготовку орфоэпических и орфографических норм казахского языка на основе нового алфавита, ускорит его изучение и использование на практике.

4. Новый алфавит при отсутствии адекватного однобуквенного обозначения отдельных казахских звуков должен позволить для их обозначения использовать диграфы. Это потребует, чтобы в качестве диграфов должны быть использованы цепочки, во-первых, из двух букв, не встречающихся рядом в казахских простых словах, или во-вторых, из буквы и специального знака, не включенного в алфавит языков программирования. Так как, во-первых, если буквы, включенные в конкретный диграф будут встречаться рядом в казахском простом слове, то звук, который обозначен этим диграфом не будут распознаваться, во-вторых, если специальный знак, используемый в диграфе включен в алфавит языка программирования, с помощью которого обрабатывается казахский текст, то звук, в обозначении которого участвует этот специальный знак не будет распознаваться.

5. Новый алфавит должен иметь порядок взаиморасположения букв, полностью совпадающий с порядком взаиморасположения букв в базовом латинском алфавите. Это позволит при создании или применении информационных технологий и ресурсов на казахском языке без дополнительных затрат использовать программы для сортировки и поиска данных. Так как, для базового латинского алфавита программы для сортировки и поиска данных автоматически встроены в состав всех системных и прикладных программ. Если новый алфавит имеет порядок взаиморасположения букв, не совпадающий с порядком взаиморасположения букв в базовом латинском алфавите, то необходимо заново разработать и установить программы для сортировки и поиска данных.

6. Уточнение звуковой системы казахского языка.

Осуществление перевода письменности любого языка с одной графики на другую требует комплекса работ, связанных с уточнением его звуковой системы, с изменением его орфоэпических, орфографических, морфологических и синтаксических правил, с разработкой компьютерных программ прямого и обратного преобразования текстов в исходном алфавите на тексты в новом алфавите и наоборот, а также с созданием соответствующих словарей и подготовки многоуровневых учебных изданий [8]

Что касается казахского языка, то до сих пор состав его звуковой системы является предметом дискуссии, издаются учебные пособия и научные труды, в которых даются разные сведения о количестве и типе казахских звуков.

При уточнении системы звуков казахского языка необходимо учитывать, что среди заимствованных звуков, гласные (и), (у) и согласные (вы), (фы), (хы) часто используются для написания фамилий и международных непереводимых на казахский язык терминов. Например: *Вагнер, Форд, Хонекер, Иванов, вектор, вирус, интернет, инвестор, институт, университет, универсам, хлор, хром, хроника, факт, файл, фонетика, формула, функция*. Поэтому их следует оставить в звуковой системе казахского языка, при этом гласные (и) и (у) будут считаться аллофонами гласных фонем (i) и (y), соответственно, и будут обозначать той же буквой (графемой), что и фонема (i), а согласные (вы), (фы), (хы) останутся как фонемы и для их обозначения будут выделены отдельные буквы в новом алфавите [8].

Таким образом можно считать, что в реформе казахского языка будут участвовать **31** фонема, из них **9** гласных фонем и **22** согласных фонем: (а) - [a], (ә) - [æ], (е) - [e], (о) - [ɔ], (ө) - [ø], (ұ) - [u], (ү) - [y], (ы) - [ɣ], (і) - [i]; **22** согласных фонем: (бы) - [b], (вы) - [v], (ғы) - [ɣ], (гі) - [g], (ды) - [ki] - [k], (қы) - [q], (жы) - [ʒ], (зы) - [z], (й) - [j], (ыл) - [l], (мы) - [m], (ны) - [n], (ың) - [ŋ], (пы) - [p], (ыр) - [r], (сы) - [s], (ты) - [t], (ў) - [w], (фы) - [f], (хы) - [h], (шы) - [ʃ], которые должны обозначаться отдельными буквами в алфавите.

Во многих из них дается артикуляционная классификация 9 исконно казахских гласных на основе 3 бинарных признаков, не учитывая, что по принципу Дирихле для корректного и полного описания 9-ти объектов 3 бинарных признаков недостаточно.

В работе [9] мы добавили новый признак относительно *вертикального положения языка* с 2 значениями: *верхние, нижние*. (см. Таблицу 1), дали и впервые представили Международном фонетическом алфавите – МФА [10] новую артикуляционную классификацию гласных звуков казахского языка (см. Рисунок 1 и Таблицу 3).

Заметим, что в классификации казахских гласных звуков не будут участвовать заимствованные гласные (и) и (у), поскольку они приняты в качестве аллофонов гласных фонем фонем (і) и (ү).

Таблица 1. Четыре бинарные артикуляционные признаки казахских гласных звуков.

Гласные звуки		Артикуляционные признаки							
Наименование	Обозначение в МФА	Положение языка				Положение губ		Положение челюсти	
		Вертикальное		Горизонтальное		Неокругленные	Округленные	Открытые	Закрытые
		Нижние	Верхние	Заднеязычные	Переднеязычные				
(a)	[ɑ]	+	-	+	-	+	-	+	-
(ə)	[æ]	+	-	-	+	+	-	+	-
(ы)	[ɯ]	+	-	+	-	+	-	-	+
(і)	[ɪ]	+	-	-	+	+	-	-	+
(ү)	[ʊ]	-	+	+	-	-	+	-	+
(ү)	[y]	-	+	-	+	-	+	-	+
(o)	[ɔ]	-	+	+	-	-	+	+	-
(ө)	[ø]	-	+	-	+	-	+	+	-
(e)	[e]	-	+	-	+	+	-	+	-

Используя Булеву алгебру, мы построим математическую модель системы гласных звуков в виде алгебраического выражения (1) на основе использования значений 4-х артикуляционных бинарных признаков, показанных в таблице 1. Для этого обозначим 4 артикуляционных признака *Вертикальное положение языка* и *Горизонтальное положение языка* казахских гласных звуков через логические переменные x_1, x_2, x_3 и x_4 , соответственно. Эти переменные принимают только значения 1 для истины и 0 для лжи:

- если язык находится в верхнем положении, то $x_1 = 1$, в противном случае $x_1 = 0$;
- если язык находится в заднем положении, то $x_2 = 1$, в противном случае $x_2 = 0$;
- если губы находятся в округленном положении, то $x_3 = 1$, в противном случае $x_3 = 0$;
- если язык челюсть находится в открытом положении, то $x_4 = 1$, в противном случае $x_4 = 0$.

На основании установленных значений для переменных можно построить Булеву модель системы казахских гласных звуков, показанной в таблице 2.

Таблица 2. Булева модель системы казахских гласных.

Наименование звука	Обозначение в МФА	x_1	x_2	x_3	x_4
(a)	[a]	0	1	0	1
(ә)	[æ]	0	0	0	1
(ы)	[ɣ]	0	1	0	0
(і)	[i]	0	0	0	0
(ұ)	[u]	1	1	1	0
(ү)	[y]	1	0	1	0
(o)	[ɔ]	1	1	1	1
(ө)	[ə]	1	0	1	1
(е)	[e]	1	0	0	1

Далее для каждого из 9 гласных звуков, выписывая признаки в виде их конъюнкции и объединяя эти конъюнкции, получим следующую дизъюнктивную нормальную форму из 9 членов:

$$\left\{ \begin{array}{l} (\bar{x}_1 \wedge x_2 \wedge \bar{x}_3 \wedge x_4) \vee (\bar{x}_1 \wedge \bar{x}_2 \wedge \bar{x}_3 \wedge x_4) \vee (\bar{x}_1 \wedge x_2 \wedge \bar{x}_3 \wedge \bar{x}_4) \vee \\ (\bar{x}_1 \wedge \bar{x}_2 \wedge \bar{x}_3 \wedge \bar{x}_4) \vee (x_1 \wedge x_2 \wedge x_3 \wedge \bar{x}_4) \vee (x_1 \wedge \bar{x}_2 \wedge x_3 \wedge \bar{x}_4) \vee \\ (x_1 \wedge x_2 \wedge x_3 \wedge x_4) \vee (x_1 \wedge \bar{x}_2 \wedge x_3 \wedge x_4) \vee (x_1 \wedge \bar{x}_2 \wedge \bar{x}_3 \wedge x_4) \end{array} \right\} \quad (1)$$

Упрощая (1) получим выражение:

$$(x_1 \wedge x_3) \vee (\bar{x}_1 \wedge \bar{x}_3) \vee (x_1 \wedge \bar{x}_2 \wedge \bar{x}_3 \wedge x_4) \quad (2)$$

(2) называется *функцией принадлежности*, характеризующей систему казахских гласных в следующей теореме:

Теорема принадлежности. Гласный звук λ принадлежит системе казахских гласных тогда и только тогда, когда его артикуляционные признаки x_1, x_2, x_3 и x_4 , определенные выше, удовлетворяют дизъюнктивной нормальной форме $(x_1 \wedge x_3) \vee (\bar{x}_1 \wedge \bar{x}_3) \vee (x_1 \wedge \bar{x}_2 \wedge \bar{x}_3 \wedge x_4)$.

На основе этих четырех артикуляционных признаков можно построить геометрическую модель системы гласных звуков казахского языка, показанной на рисунке 1.

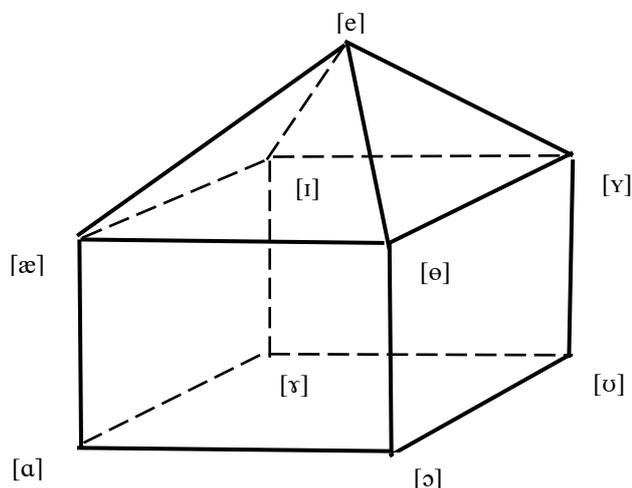
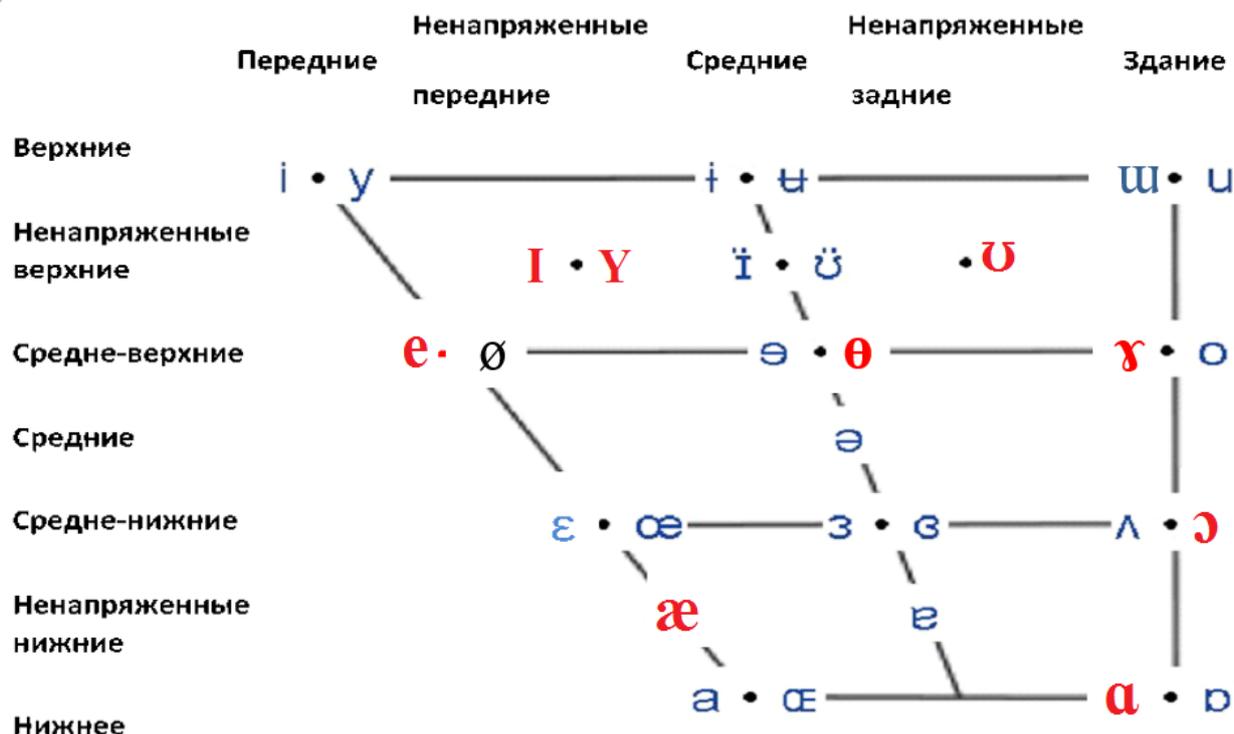


Рисунок 1. Геометрическая модель системы казахских гласных на основе 4-х бинарных артикуляционных признаков

Здесь плоскость [a] [ɔ] [u] [ɣ] представляет заднеязычные (твердые) гласные звуки (a) (o) (ұ) (ы), плоскость [æ] [ə] [y] [i] – переднеязычные (мягкие) гласные звуки (ә) (ө) (ү) (і), плоскость [a] [ɔ] [ə] [æ] – открытые гласные звуки (a) (o) (ө) (ә), плоскость [ɣ] [u] [y] [i] – закрытые гласные звуки (ы) (ұ) (ү) (і), плоскость [ɔ] [u] [y] [ə] – округленные гласные звуки (o) (ұ) (ү) (ө), плоскость [a] [ɣ] [i] [æ] – неокругленные гласные (a) (ы) (і) (ә), а вершина [e] представляет особый звук (е).

Таблица 3. Артикуляционная классификация казахских гласных звуков в МФА (выделены красным цветом).



Теперь рассмотрим артикуляционные признаки и артикуляционную классификацию казахских согласных звуков.

Во многих естественных языках характеристика согласных звуков состоит из 5 основных признаков:

- 1) место образования шума;
- 2) способ образования шума;
- 3) уровень шума (сонорность / шумность);
- 4) глухость / звонкость;
- 5) твёрдость / мягкость.

Однако в казахском языке признак твёрдость / мягкость относится к гласным звукам.

Поэтому для казахских согласных звуков можно использовать следующие 3 артикуляционные признаки;

- 1) место образования согласного звука;
- 2) способ образования согласного звука;
- 3) участие голоса в образовании согласного звука.

Согласные звуки по месту образования подразделяются на:

- губные;
- язычные.

Губные подразделяются на губно-губные и губно-зубные, а язычные - на зубно-зубные переднеязычные, среднеязычные и заднеязычные.

В свою очередь:

- переднеязычные подразделяются на зубные, альвеолярные и постальвеолярные;
- среднеязычные подразделяются на тверднённые;
- заднеязычные подразделяются на мягконённые и увулярные.

Согласные звуки по способу образования подразделяются на:

- смычные;
- щелевые.

Смычные согласные звуки подразделяются на смычно-взрывные, смычно-боковые и смычно-дрожящие.

Щелевые согласные звуки подразделяются на плоскощелевые и круглощелевые.

Согласные звуки по участию голоса подразделяются на глухие, звонкие и сонорные.

В таблице 4 показаны артикуляционные признаки всех принятых в звуковую систему казахского языка согласных звуков.

В таблице 5 приведена артикуляционная классификация согласных звуков казахского языка в МФА, где можно увидеть их относительные места среди согласных звуков других естественных языков (выделены полужирным красным цветом).

5. Задачи перевода казахского алафавита на латиницу.

Далее учитывая, что в базовом латинском алфавите имеются **26** букв, а в казахском языке - **31** фонема, для определения нового алфавита казахского языка мы должны решить следующие задачи:

- 1) определить звуки, обозначаемые только одной буквой;
- 2) определить звуки, обозначаемые диграфами;
- 3) определить значения диграфов;
- 4) обновить орфоэпические и орфографические нормы, а также морфологические и синтаксические правила казахского языка;
- 5) разработать технологии для их компьютерной обработки и подготовить интеллектуальные сетевые многоуровневые учебные пособия и учебники с речевыми интерфейсами, позволяющие обучиться казахскому языку через любые мобильные устройства.

Таблица 4. Артикуляционные признаки казахских согласных звуков.

Название согласных звуков	Обозначение согласных звуков в МФА	Место образования								Способ образования				Участие голоса				
		Губные		Язычные						Смычные				Щелевые		Глухие	Звонкие	Сонорные
		Губно-губные	Губно-зубные	Переднеязычные			Среднеязычные		Заднеязычные	Смычно-взрывные	Смычно-носовые	Смычно-боковые	Смычно-дрожящие	Плоскощелевые	Круглощелевые			
		Зубные	Альвеолярные	Постальвеолярные	Тверднёбные	Мягконёбные	Увулярные											
(бы)	[b]	+	-	-	-	-	-	-	-	+	-	-	-	-	-	-	+	-
(вы)	[v]	-	-	+	-	-	-	-	-	-	-	-	-	+	-	-	+	-
(ғы)	[ɣ]	-	-	-	-	-	-	+	-	-	-	-	-	-	+	-	+	-
(гі)	[g]	-	-	-	-	-	-	+	-	+	-	-	-	-	-	-	+	-
(ды)	[d]	-	-	-	+	-	-	-	-	+	-	-	-	-	-	-	+	-
(кі)	[k]	-	-	-	-	-	-	+	-	+	-	-	-	-	-	-	+	-
(қы)	[q]	-	-	-	-	-	-	-	+	+	-	-	-	-	-	-	+	-

(жы)	[ʒ]	-	-	-	+	+	-	-	-	-	-	-	-	+	-	+	-	
(зы)	[z]	-	-	-	+	-	-	-	-	-	-	-	-	+	-	-	+	-
(й)	[j]	-	-	-	-	-	+	-	-	-	-	-	-	+	-	-	-	+
(ыл)	[l]	-	-	-	+	-	-	-	-	-	-	+	-	-	-	-	-	+
(мы)	[m]	+	-	-	-	-	-	-	-	-	+	-	-	-	-	-	-	+
(ны)	[n]	-	-	+	-	-	-	-	-	-	+	-	-	-	-	-	-	+
(ың)	[ŋ]	-	-	-	-	-	+	-	-	-	+	-	-	-	-	-	-	+
(пы)	[p]	+	-	-	-	-	-	-	-	+	-	-	-	-	-	+	-	-
(ыр)	[r]	-	-	-	+	-	-	-	-	-	-	-	+	-	-	-	-	+
(сы)	[s]	-	-	-	+	-	-	-	-	-	-	-	-	-	+	+	-	-
(ты)	[t]	-	-	+	-	-	-	-	-	+	-	-	-	-	-	+	-	-
(ў)	[w]	+	-	-	-	-	-	-	-	-	-	-	-	-	+	-	-	+
(фы)	[f]	-	+	-	-	-	-	-	-	-	-	-	-	+	-	+	-	-
(хы)	[x]	-	-	-	-	-	-	+	-	-	-	-	-	+	-	+	-	-
(шы)	[ʃ]	-	-	-	+	+	-	-	-	-	-	-	-	+	-	+	-	-

Таблица 5. Казахские согласные звуки в МФА

	Биллабиальные		Лабiodентальные		Дентальные		Альвеольные		Палатоальвеолярные		Ретрофлексные		Палатальные		Велярные		Увулярные		Глоттальные	
Взрывные	p	b	p	b	t	d	t	d			t	ɖ	c	ɟ	k	g	q	ɢ	ʔ	
Назальные		m		m̥		n̥		n				ɳ		ɲ		ŋ		ɴ		
Фрикативные	ɸ	β	f	v	θ	ð	s	z	ʃ	ʒ	ɬ	ɮ	ç	j	x	y	χ	ʁ	ħ	ɦ
Аппроксиманты	w			v				ɹ			ɻ			j		ɥ				
Одноударные		v		v				r			ɽ								ʁ̥	
Вибранты		ʋ						r			•								ʀ	
Латеральные аппроксиманты								l			ɭ		ʎ	ʎ		ɭ				
Латеральные фрикативы							ɬ	ɮ			ɬ°		ʎ°		ɭ°	ɬ				

Решения этих задач будут основываться на:

- результатах исследования текстового корпуса казахского языка, содержащего 100 миллионов букв действующего алфавита на основе кириллицы [11], с целью определения частоты употребления букв и сочетания букв в казахском тексте;
- математической модели и геометрической модели системы гласных казахских звуков отражающие их классификацию [9];

– осцилограммы гласных звуков, позволившие установить между мягкими и твердыми гласными звуками казахского языка отношения, которые будут использоваться для обозначения диграфами мягких гласных [8]:

(ə) = (a)+(e) или [æ] = [a]+[e], (e) = (i)+(e) или [e] = [ɔ]+[e],
(y) = (y)+(e) или [y] = [u]+[e], (i) = (ы)+(e) или [i] = [ɣ]+[e].

Результаты исследования этого казахского текстового корпуса также помогли разработать правила избавления казахского текста на действующем алфавите на основе кириллицы от букв Ёё, Ии, Ээ, Юю, Яя, Цц, Чч, Щщ, Ъь, Ьь, звуковые значения которых не будут входить в звуковую систему казахского языка. На основании этих правил был создан переходной орфографический словарь казахского языка на кириллице, без этих букв [12].

Этот словарь будет служить исходной базой данных для автоматического перевода казахской письменности на латиницу, который будет состоять из двух этапов: первый этап – перевод исходного текста на промежуточный текст без букв Ёё, Ии, Ээ, Юю, Яя, Цц, Чч, Щщ, Ъь, Ьь; второй этап – перевод промежуточного текста на латиницу. Реализуя эти этапы мы разработали и внедрили конвертор для автоматического перевода любого казахского текста с кириллицы на латиницу: <http://alphabet.kz/convector>.

6. Правила перевода казахской письменности на латиницу

Осуществление перевода письменности любого языка с одной графики на другую требует комплекса работ, связанных с уточнением его звуковой системы, с изменением его орфоэпических, орфографических правил [8].

В реформе по переводу казахской письменности на латиницу будут участвовать **31** фонема, из них **9** гласных фонем, **22** согласных фонем. В действующем на основе кириллицы алфавите имеются **42** буквы, из которых **10** букв Ёё, Ии, Ээ, Юю, Яя, Цц, Чч, Щщ, Ъь, Ьь используются для обозначения звуков русского языка и в основном встречаются в заимствованных словах.

Поэтому нам необходимо построить следующие правила:

- правила удаления из слов букв Ёё, Ии, Ээ, Юю, Яя, Цц, Чч, Щщ, Ъь, Ьь;
- правила обработки буквы Уу в словах;
- правила обработки букв Аа, Ээ, Оо, Өө, Үү, Ыы, Іі в словах.

Для представления 31 казахских фонем в 26 буквенном базовом латинском алфавите необходимо применять некоторые методы его адаптации к другим языкам (см. 3.1.1).

Для обозначения казахских фонем буквами базового латинского алфавита используется схожесть их обозначения в МФА [31, 32] и предлагаются **3** группы правил:

1. Один звук и одна буква:

1.1. **5** казахские гласные фонемы (а) - [a], (е) - [e], (о) - (o), (y) - [y] и (ы) - [ɣ] будут обозначаться **5** латинскими буквами *Aa, Ee, Oo, Uu* и *Yu*, значениями которых являются созвучные с ними звуки, имеющие в МФА одинаковые обозначения с обозначениями указанных казахских фонем, соответственно;

1.2. **18** казахские согласные фонемы (бы) - [b], (вы) - [v], (гi) - [g], (ды) - [d], (фы) - [f], (й) - [j], (кi) - [k], (кы) - [q], (ыл) - [l], (мы) - [m], (ны) - [n], (пы) - [p], (ыр) - [r], (сы) - [s], (ты) - [t], (y) - [w], (хы) - [x] и (зы) - [z] будут обозначаться **18** латинскими буквами *Bb, Vv, Gg, Dd, Ff, Jj, Kk, Qq, Ll, Mm, Nn, P, Rr, Ss, Tt, Ww, Xx* и *Zz*, значениями которых являются созвучные с ними звуки, имеющие в МФА одинаковые обозначения с обозначениями указанных казахских фонем, соответственно;

1.3. Согласная фонема (ғы) - [ɣ] созвучна с согласной фонемой (гi) - [g], обозначенной буквой *Gg*. Эта буква в латинском алфавите стоит рядом с буквой *Hh*, которая обозначает звук в МФА, входящий в одну группу с обозначаемой фонемой. Поэтому целесообразно обозначить фонему (ғы) - [ɣ] латинской буквой *Hh*.

1.4. Согласная фонема (жy) - [ʒ] будет обозначаться латинской буквой *Cc* как это принято в алфавите турецкого языка. Поскольку оригинальное значение буквы *Cc* в базовом

латинском алфавите является звук (си), который также принадлежит к классу шипящих звуков, как и казахский звук (жы).

2. Один звук и диграф:

2.1. Казахские 3 гласные мягкие фонемы (ə) - [æ], (ө) - [ø], (ү) - [y] будут обозначаться диграфами, в которых первыми парами являются латинские буквы *Aa*, *Oo* и *Uu*, обозначающие близкие по звучанию твердые гласные (а) - [a], (о) - [o] и (у) - [u], соответственно, а второй парой каждого диграфа будет специальный символ ` – *аксанти грав*, который входит в состав алфавита для первого варианта алфавита (как это принято в узбекском алфавите) или латинская буква *Ee* для второго варианта алфавита (как это принято в «компьютерном» немецком алфавите).

2.2. Казахская согласная фонема (шы) - [ʃ] будет обозначаться диграфом, в котором первой парой будет латинская буква *Ss*, обозначающая близкий по звучанию согласный звук (сы) - [s], а второй парой - латинская буква *Hh*, обозначающая согласный звук (ғы) - [ɣ]. Такой диграф в новом алфавите не будет порождать орфографические противоречия в казахской письменности, так как в казахских словах не существуют сочетания согласных звуков (сы) - [s] с согласным звуком (ғы) - [ɣ].

2.3. Казахская согласная сонорная фонема (ың) - [ŋ] будет обозначаться диграфом. Первой парой диграфа будет латинская буква *Nn*, обозначающая фонему (ны) - [n], которая по своему значению близок фонеме (ың) - [ŋ]. Второй парой диграфа будет латинская буква *Gg*, обозначающая звук (гі) - [g] (как это принято в английском и узбекском языках). В записи в новом алфавите диграф «ng» должен содержаться только в одном слоге, так как в казахском слове никогда перед звуком (ың) - [ŋ] не встречается согласный звук и нет слова, которое начинается со звука (ың) - [ŋ].

Однако, некоторые разные по произношению и смыслу слова в предлагаемом алфавите будут писаться единично. Например, разные слова «күңі» и «күңгі» в предлагаемом алфавите будут писаться единично:

Для того, чтобы не допустить подобные случаи необходимо принять следующее правило: Если при записи в латинском алфавите казахского слова после буквы *Nn*, обозначающей согласный звук (ны) - [n], необходимо писать букву *Gg*, обозначающей согласный звук (гі) - [g] (такой случай возможен только в исконно казахских словах), то вместо буквы *Nn* надо писать диграф «ng», который обозначает согласный звук (ың) - [ŋ]. Это будет соответствовать закономерности произношения в казахском языке и в его письменности не порождает никакие противоречия.

3. Одна буква и два звука.

3.1. В действующем алфавите есть кириллическая буква Ии. Во-первых, она встречается в заимствованных словах и на практике носителями казахского языка произносится как исконно казахский звук (і), например, *интернет* = (*интернет*), *инвестор* = (*инвестор*), *институт* = (*институт*). Во-вторых, в исконно казахских словах во всех случаях произносится как сочетание звуков (*ій*) или (*ый*) в зависимости от мягкости или твердости контекста, например, *киік* – (*кййік*), *қыық* – (*қййық*), *тиін* – (*тййін*), *тыын* – (*тыйын*), *ғылыми* – (*ғылымй*), *ресми* – (*ресмй*), *ылғи* – (*ылғй*), *ылди* = *ылдый*, *қияли* = *қййалый*, *адами* = *адамй*, *ғылыми* = *ғылымй*, *тарихи* = *тарихый*.

Поэтому в первом случае, русский гласный звук (и) нужно принять как аллофон казахской фонемы (і). Во втором случае, просто привести в порядок орфографические правила и тогда не будет никаких орфографических противоречий, указанных во Введении настоящей работы (см. стр. 24-25). Тогда латинская буква *Ii* будет обозначать как фонему (і), так и её аллофон (и), которые в МФА будут обозначаться символами [i] и [i̞], соответственно, как это принято во многих алфавитах других языков (например, в английском языке: *bit* [bit] и *bite* [baɪt]). То есть заимствованный гласный аллофон (и) не будет обозначаться отдельной

графемой, он будет обозначаться той же латинской буквой *li*, которой обозначается казахская гласная фонема (i).

3.2. Также в действующем алфавите есть кириллическая буква Уу. Во-первых, она встречается в заимствованных словах и на практике носителями казахского языка произносится как исконно казахский звук (ұ), например, *универсам* = (ұниверсам), *университет* = (ұниверситет), *формула* = (формұла). Во-вторых, она в исконно казахских словах во всех случаях произносится как сочетание звуков (уу) или (ұу) в зависимости от мягкости или твердости контекста.

Поэтому в первом случае, русский гласный звук (y) нужно принять как аллофон казахской фонемы (ұ). Во втором случае, просто привести в порядок орфографические правила и тогда не будет никаких орфографических противоречий, указанных во Введении настоящей работы (см. стр. 24-25). Тогда латинская буква *Ui* будет обозначать как фонему (ұ), так и её аллофон (y), которые в МФА будут обозначаться символами [u] и [y], соответственно, как это принято во многих алфавитах других языков (например, в английском языке: put [put], but [bat] и tube [tu:b]). То есть гласный аллофон (y) не будет обозначаться отдельной графемой, он будет обозначаться той же латинской буквой *Ui*, которой обозначается казахская гласная фонема (ұ).

Заключение.

В заключении можно сказать, что с **26** буквами базового латинского алфавита удалось обозначить **31** фонему казахского языка, из них **26** фонем обозначены отдельными буквами, а **5** фонем обозначены диграфами. Для всех фонем, включая аллофоны, определены обозначения в МФА.

Аллофоны гласных казахских фонем (i) и (ұ) позволит сблизить написание многих заимствованных и подлежащих к заимствованию непереводаемых на казахский язык международных терминов в новом алфавите с их оригиналами, сохраняя казахское произношение.

В результате предложены 2 варианта алфавита казахского языка на основе латинской графики с указанием соответствующих обозначений в МФА: в первом варианте алфавита содержатся только 26 букв базового латинского алфавита, а во втором варианте, помимо этих букв, содержится и знак ` – *аксанти грав*, который будет использоваться в диграфах, как вторая пара каждого диграфа (см. Таблицу 6 и Таблицу 7).

Таблица 6. Казахский алфавит (на основе латиницы)

№	Латиница	Название звука	Кириллица	МФА	Пояснение
1	A a	(a)	А а	[a]	
2	B b	(бы)	Б б	[b]	
3	C c	(жы)	Ж ж	[ʒ]	Как в турецком языке
4	D d	(ды)	Д д	[d]	
5	E e	(е)	Е е	E e	
6	F f	(фы)	Ф ф	[f]	
7	G g	(gi)	Г г	[g]	
8	H h	(ғы)	Ғ ғ	[ɣ]	
9	I i	(i), (и)	І і	[i, i]	Как в английском языке
10	J j	(й)	Й й	[j]	
11	K k	(ki)	К к	[k]	
12	L l	(ыл)	Л л	[l]	
13	M m	(мы)	М м	[m]	
14	N n	(ны)	Н н	[n]	
15	O o	(o)	О о	[ɔ]	

16	P p	(пы)	П п	[p]	
17	Q q	(қы)	Қ қ	[q]	
18	R r	(ры)	Р р	[r]	
19	S s	(сы)	С с	[s]	
20	T t	(ты)	Т т	[t]	
21	U u	(ў)	Ў ў	[ʊ, u]	
22	V v	(вы)	В в	[v]	
23	W w	(ў)	У у	[w]	
24	X x	(хы)	Х х	[x]	Как в узбекском и азербайджанском языке
25	Y y	(ы)	Ы ы	[ɣ]	
26	Z z	(зы)	З з	[z]	
27	`	аксанти грав			Во втором варианте

Таблица 7. Орфографические правила для диграфов

№	Название звука	Кириллица	МФА	Латиница	Пояснение
1	(ә)	Ә ә	[æ]	Ae ae	1-вариант.
				A` a`	2-вариант.
2	(ө)	Ө ө	[ø]	Oe oe	1-вариант
				O` o`	2-вариант.
3	(ү)	Ү ү	(y)	Ue ue	1-вариант.
				U` u`	2-вариант.
4	(ың)	Ң ң	[ŋ]	Ng ng	
5	(шы)	Ш ш	[ʃ]	Sh sh	

Литература

- 1 <https://egemen.kz/article/nursultan-nazarbaev-bolashaqqa-baghdar-rukhani-zhanhgyru>.
- 2 https://ru.wikipedia.org/wiki/Казахская_письменность.
- 3 https://ru.wikipedia.org/wiki/Латинский_алфавит.
- 4 https://ru.wikipedia.org/wiki/Турецкая_письменность.
- 5 https://traditio.wiki/Туркменская_письменность.
- 6 https://ru.wikipedia.org/wiki/Узбекская_письменность
- 7 https://ru.wikipedia.org/wiki/Азербайджанская_письменность.
- 8 Шарипбай А.А. Проблемы перевода казахской письменности на латинский алфавит. Монография. ЕНУ им. Л.Н.Гумилева, Астана, 2017, -138
- 9 Yessenbayev, M. Karabalayeva and A. Sharipbayev, "Formant Analysis and Mathematical Model of Kazakh Vowels," 2012 UKSim 14th International Conference on Computer Modelling and Simulation, Cambridge, 2012, P.427-431.
- 10 https://en.wikipedia.org/wiki/International_Phonetic_Alphabet
- 11 <http://kazcorpus.kz/klcweb/en/>.

СЕКЦИЯ 1

SECTION 1

БІЛІМ БЕРУДІ АҚПАРАТТАНДЫРУ
ИНФОРМАТИЗАЦИЯ ОБРАЗОВАНИЯ
INFORMATIZATION OF EDUCATION

THE PHASES OF THE IMPLEMENTATION OF TEMPUS-PROMIS PROJECT AT THE TECHNOLOGICAL UNIVERSITY OF TAJIKISTAN

(Technological University of Tajikistan, Tajikistan)

Starting from December 2013 the Technological University of Tajikistan has been participating in the implementation of TEMPUS-PROMIS Project of the European Commission.

The main objective of the project is to exchange and transfer the best experience and knowledge of the European HEIs for the purpose of organization the similar master's study programmes at Central Asian higher educational institutions and within the implementation of this programme to attract the graduates of none IT specialties to study informatics as a second competence and as well as to strengthen their professional knowledge via the application of information and communication technologies and develop the integration of the educational process with production.

The peculiarities of the implementation of this project is that the modern trends are covering all the spheres of the people's lives with the information and communication technologies. Gradually, this process will be integrated with the technological process, production and will change its qualification requirements in relation to the preparation of the specialists. Certainly, for every higher educational institution the preparation of the competitive specialists meeting the demands of the labor market and as well as the requirements of the employers will be placed first. Therefore, the higher education institutions need to harmonize their curricula with the requirements of the employers and of course, must establish close cooperation with them.

Nowadays, for specialties of technological and economic profiles and are being taught at the Technological University of Tajikistan the acquisition and expansion of the computer knowledge is having more importance, because the technological processes and control of production needs to be improved in term of the technological level, their integration with the computer technologies in order to reduce the production loss and raise the efficiency of the production and management.

The students, of course, during their studies in the specialties with technological and economic profile will improve professional knowledge, however, the curricula cannot give them enough computer knowledge that cover all the areas of their activities. In this case, if these students after graduation continue their studies in the Master's program "Informatics as a second competence", thus they will have the opportunities to get acquainted with technological background and the logics of the technological processes. Thus, they will be able to contribute to the programming and automation of the various production processes and at the same to become competitive specialists in labor market. From this point of view Technological University of Tajikistan purposefully taking into account the best teaching conditions in the frame of implementation of the TEMPUS-PROMIS project has chosen the specialty 1-53010106 «Automation of the technological and production processes (in the food industry)».

It must be emphasized that in the implementation process of the TEMPUS-PROMIS Project there was elaborated Master's curriculum according to specialty 1-53010106 «Automation of the technological and production processes (in the food industry)». The curriculum has been analyzed by the Methodological Center under the Ministry of Education and Science of the Republic of Tajikistan. Moreover, on the basis of the Master's curriculum, the Technological University of Tajikistan for the purpose of the implementation of this program received the license of the Agency for Control of Education, where there is indicated the admission plan for the academic year 2016-

2017. During the admission period there were accepted 10 Master students for this program and at the moment teaching process is going according to the foreseen plan.

In addition, it must be noted that the project gave opportunities to 9 teaching staff of the Technological University of Tajikistan to be retrained with the financial support of the project in the partner institutions of the European Union.

№	Name and surname	Country	Receiving institution	Retraining according to course
1.	Mirzo Yusupov	Lithuania	Kaunas University of Technologies	Data mining and data warehouse
2.	Miratullo Rashidov	Lithuania	Kaunas University of Technologies	Databases
3.	Abdulnazar Gulomsafdarov	Germany	University of Applied Sciences	Programming
4.	Mirshafi Gafurov	Germany	University of Applied Sciences	Algorithms and data structure
5.	Mirzamuddin Nasriddinov	France	Pierre Mendes France University	Mathematics for computer science
6.	Ahror Jafarov	France	Pierre Mendes France University	Human machine interaction
7.	Begijon Sharipov	Poland	Lublin University of Technology	Software engineering
8.	Yusuf Nabotov	Finland	Savonia University of Applied Sciences	WEB Development
9.	Saifudin Nazarov	Finland	Savonia University of Applied Sciences	Operating Systems and Networks

The project also gave opportunities to two teaching staff of the Technological University of Tajikistan to participate in the workshops conducted by the European project partners at Khujand State University in June 2016. In addition, there is planned to conduct workshop at TUT by the European partners from Lublin University of Technology in Poland.

With the financial support of the TEMPUS-PROMIS Project and for the purpose creation of a specialized classroom for the Master students according to the signed Partnership Agreement between Technological University of Tajikistan and Grenoble Alps University there has been purchased necessary equipment on the amount of **15200 Euro** and installed in the specialized classroom. The installed equipment is being used to achieve the project aims and objectives. Generally, Technological University of Tajikistan has done all the efforts to implement the project at required level and as well to guarantee the sustainability of the project within project lifetime and beyond it.

LINKS WITH INDUSTRY IN TEACHING AND RESEARCH - EXPERIENCES IN THE DEPARTMENT INFORMATICS AND MEDIA AT THE BEUTH UNIVERSITY OF APPLIED SCIENCES

(Department Informatics and Media, Beuth University of Applied Sciences, Berlin, Germany)

1 Introduction

The aim of the PROMIS project is to introduce the Master degree *Informatics as a Second competence* in ten universities of central Asia. This degree should provide the local industry with versatile and well-trained people able to develop the software of tomorrow [1]. Therefore, a university offering this master needs to have strong links with the local industry. There are various ways to build good relationships between higher education and industry. In this contribution, we describe how this is done in our institution. Beuth University of Applied Sciences (Beuth HS) has been created in 1971 as a fusion of several schools of engineering. Belonging to the category “University of Applied Sciences”, its focus is to prepare students to work in the industry. General rules foster this focus, such as:

- Applicants to a professorship in such a university must not only have finished their PhD, they must also possess a minimum of 3-years working experience in the industry.
- About 25% of the teaching is done by industry-professionals possessing an appropriate qualification; these professionals offer courses in their specific working expertise.
- Intake of students is limited to about 40 students per study program. This way, professors are not only transmitters of knowledge but can also overtake a kind of coach-role to prepare their students better.
- Most of the bachelor programs require students to complete an internship in a company of at least 3 months. During this internship, each student is supervised by a professor who visits her/him in the company. This provides an opportunity to evaluate the content of the study program against the needs and practices of potential employers of students.
- Bachelor and Master theses can be written in cooperation with industry partners. This happens in about 40% of the cases.

Any responsible of a study program can be innovative in the design of the curriculum and creates more opportunities to strengthen further links between teaching and industry. Additionally, any professor can enrich her/his own teaching with practices and lessons learned from projects conducted with companies. In this paper, we present experiences from the department of Informatics and Media of Beuth HS to illustrate these two aspects.

This paper is organized as follows. Section 2 presents how the curriculum of the bachelor and the master “Informatics and Media” enforces stronger links with industry in the teaching. Section 3 presents projects conducted in the department and their impact on teaching. Section 4 gives further measures in place at Beuth HS to foster even closer links with industry. Finally, the last section concludes this paper.

2 Links with Industry in Teaching

In the curriculum of the bachelor and master program “Informatics and Media”, there are several courses that allow a close cooperation between the university and the industry.

In the course “Selected Topics in Informatics and Media” of the master program, representatives of the industry are invited to give a talk about important topics and questions addressed in their company. The students not only get this information, additionally, they get the opportunity to socialize with these representatives, which is helpful for getting a job or an internship in those companies. On the other hand, the companies can advertise themselves in order to find new employees.

The curriculum of the bachelor as well as of the master program contains several project modules like the “Masterprojekt”, in which companies give appropriate practical assignments to teams of students under the supervision of a university lecturer. In this way, students are able to work on practical tasks during the course of their studies, and the companies can early secure the best students as future employees.

3 Links with Industry in Research

Most of the research projects, that professors are involved in, are funded by the federal state of Berlin, by the German Ministry of Education and Research, by the German Ministry of Economics and Energy or by the European Community. Generally, these projects have to involve at least one small and medium sized enterprise that could use the research work in its business. Usually, professors involved in such projects involve their students too. There are three major ways for the involvement of students. One way is to offer topics for bachelor and master theses to investigate a well-defined area of the project. A second way is to hire students to perform specific tasks in the project. There is a well-defined status for that which is called “Studentische Hilfskraft (SHK)”, what can be translated as *student assistant*. This way allows students not only to earn extra money, but also to put in practice what they hear in their courses. Finally, such research projects provide practical case studies that can be used by professors during their lectures. The Research Day proceedings [2] give an overview of the research projects taking place in the whole university, which includes the department Informatics and Media. The Long Night of the Sciences [3] which takes place every year in June is another excellent possibility to present results of research projects in cooperation with industrial partners. Very often, the students involved in the projects enthusiastically present their research results, discuss their current work with visitors, show spectacular experiments und give science shows together with their professors. We now present succinctly projects the authors of this paper have been involved in.

Nutzerdaten und Nutzerprofile in Lernraumsystemen (NNL) and Learning Analytics Mining and Adaption (LAMA). These two two-years projects have been funded by the federal state Berlin and involved besides Beuth HS the company eLeDia [4]; this company provides complete solutions around the learning management system Moodle to schools, universities and private companies for the continuing education of their employees. The aim of these two projects is to extend learning management systems with a learning analytics module to help teachers understand how students interact with the materials they upload online for them [5, 6]. Four bachelor theses have been written in these projects and four SHK hired. These projects have proved to be a good opportunity for the company eLeDia who hired as employee a research assistant involved in the project NNL.

fMOOC. This one-year project is founded by the German Ministry of Education and Research and is a cooperation with the Charité, the Berlin medical university. Its aim is to create a massive open online course on fitness for senior citizens [7]. A unique feature of this course is the integration of wearable technology like fitness bracelets to better record the physical activity of participants and also to motivate them. The advisory board of the project includes companies like Garmin and Smartpatient. One bachelor thesis has been written and two SHK hired in this project (beside a research assistant).

Smart Learning im Handwerk (SLHw). This three-years project is funded by the German Ministry of Education and Research and involves beside Beuth HS the FOKUS-institute of the application-oriented research organization Fraunhofer [8], the institute for studies on future and evaluation of technologies IZT [9] and the chamber of crafts of Berlin [10]. Its aim is to develop an innovative learning management system called the Learning Companion App (LCA) [11] for the chamber of crafts of Berlin. LCA includes a learning analytics module and a recommender system (recommend to a student what s/he should learn next). This application allows for including mobile, online and adaptive continuing education in the programs offered by the Berlin chamber of crafts. One master thesis has been written in this project, one Bachelor thesis is beginning and 6 students are working as SHK beside the research assistants that have been especially hired. A special task

was given as project topic in the course “Masterprojekt” (master projects) in our master program Media Informatics. Results of this project illustrate a part of the course *machine learning*.

Real-time analysis and crowdsourcing for a self-organized city logistics (ExCELL). The research project ExCELL [12] is funded by the German Ministry of Economics and Energy (BMWi) and part of the SmartData campaign. It seeks to build up a modular service platform offering mobility solutions such as routing, scheduling and job pooling. ExCELL is targeting small and medium-sized enterprises to improve their productivity by saving hours waiting in traffic jams or running with empty cargos. Moreover, ExCELL is also relevant for tech start-ups by chaining their tools into the platform and using crowd sourced data from existing services. During 3 years, four students are working as SHK beside the research assistants that have been hired for the project. More than 15 Master and Bachelor thesis have been written in the project until now. Additionally, the ExCELL-project participated several times in the course “Masterprojekt” in our master program Media Informatics. Within the scope of this course, interesting subprojects are developed for ExCELL. For example, in the summer course 2016, three students of the “Masterprojekt” and the ExCELL-project participated together in the ACM SIG-SPATIAL Giscup [13], which is an annual algorithm contest and identified the 50 hotspots of taxi dropoffs between January 2009 and June 2015 in New York City. The team had to utilize a very large collection of spatio-temporal data – over one billion records.

Management of multi perspective information models to integrate Building Information Modeling (BIM) from Facility Management (FM) perspective (BIM-FM). The research project BIM-FM [14] is funded by the Institute of Applied Sciences Berlin (IFAF) [15] which was founded and is financed by the Berlin Senate Administration for Education, Science and Research. The IFAF Berlin is particularly fostering collaborative projects between the universities involved in the institute with partners from the Berlin-Brandenburg region. While most of the current BIM initiatives focus on the planning and construction phase, the research project BIM-FM addresses the Facility Management and operational phase of the building life-cycle. The results of the project will be the Generation of two BIM models for existing buildings from FM-perspective, making use of 3D-Laserscanning technology. The project part of the Beuth Team will be the **Integration** of multiple BIM models with different FM related information sources with heterogeneous data formats and a **Multi-source query engine** based on Semantic Web technologies. During two years, three students are working as SHK beside one research assistant that have been hired for the project. Some Bachelor theses have been written in the project.

Fördermitteldatenbank (FoerDaBa). The joint research and development project FoerDaBa (an acronym for German public funds database) develops and implements a web-based database solution for managing public funds handled by Friedrichshain-Kreuzberg’s facility management and urban planning departments. It ensures a consistent data management for everything related to properties, developers and development schemes. FoerDaBa has been officially launched in June 2016 by Berlin’s Friedrichshain-Kreuzberg district [16]. Prior to this, FoerDaBa was tested by a selected group for several months. All of Friedrichshain-Kreuzberg’s development related departments will be using FoerDaBa in the future. An approximated number of 100 users is planned. The project started as a case study in a lecture “database systems”. During one semester, 10 teams developed the database design and a user interface. The team with the best design was able to implement their project and was financed as SHK for two years by the Friedrichshain-Kreuzberg’s facility management and urban planning departments.

4 Other Activities with Industry at Beuth University

Since the Beuth HS is committed to practical studies, there are further links with industry.

For example, in order to promote the founding of new companies, especially through graduates, the Beuth HS has established some measures.

The so-called “*GründerScout*” [17] supports start-ups as well as founders from the initial idea to writing a business plan. Experienced persons coach the founders. They provide support in all areas, from the first critical examination of the idea to advices regarding financing. There are

several scholarships which promote start-up ideas. In addition, the founders are accompanied by experienced mentors from the university and receive scientific advices.

A second measure is the BENHU network (The Berlin Entrepreneurship Network of Universities and Business). BENHU is initiated by the Beuth HS and the Berlin School of Economics and Law in cooperation with business. Several companies as well as the federal state Berlin and IHK Berlin (chamber of Commerce and Industry of Berlin) finance the project. One of the main objectives of the network is a collaboration of start-ups and established companies in order to develop innovative business models. Furthermore, the entrepreneurship education in the universities should be strengthened as well as an international network of business incubators should be established. In workshops, actors of the following levels are brought together: universities, start-ups as well as established companies [18].

Another important link between industry and the Beuth HS is the so-called *Deutschlandstipendium* [19]. Students can apply for a scholarship of 300€ per month. 150€ is funded by the government and 150€ is funded by a business partner. Many of these companies furthermore provide non-financial support as mentoring programmes and internships. This allows the scholarship holders to establish valuable contacts to businesses and gain insight into work life. Furthermore, the contacts between universities and business are strengthened.

5 Conclusion

In this paper we have shown various forms of collaboration between our university as a University of Applied Sciences and the industry. The contacts are intensive and very fruitful for both sides. In addition to the above-mentioned relationships in teaching, research and activities like “*GründerScout*” or “*Deutschlandstipendium*”, some other ways of direct contacts between companies and students exist. Students work in companies with a special status, the so-called “*Werkstudenten*” - working students. Our computer science students for example very often work part-time as *Werkstudent* and take on design and development tasks or are involved in projects with customers. Often the companies are coming to the university to get good working students recommended. This also leads to close ties between the university and companies.

References

- 1 Bardou, D., Adam, J.-M., Merceron, A., Ripphausen-Lipa, H., Milosz, M. Kapocius, K. & Toppinen, A. (2015). Pedagogy Sharing in the PROMIS Project. Proceedings of 2015 IEEE Global Engineering Education Conference (EDUCON), 18-20 March 2015, Tallinn, Estonia, pp. 587 – 593
- 2 Research Day. <http://www.beuth-hochschule.de/researchday/> last accessed April 23, 2017.
- 3 Long Night of the Sciences. <http://www.beuth-hochschule.de/en/Indw/> , last accessed April 25, 2017
- 4 <https://eledia.de/> last accessed April 23, 2017.
- 5 A. Krueger, A. Merceron, B. Wolf: A Data Model to Ease Analysis and Mining of Educational Data. Proc. 3rd International Conference on Educational Data Mining EDM2010, Pittsburgh, USA, June 11-13, pp. 131-140.
- 6 H. Dierenfeld and A. Merceron . Learning Analytics with Excel Pivot Tables. Proceedings of the 1st Moodle Research Conference. S. Retalis & M. Dougamias pp. 115-121. Chermia, Creete, 2012.
- 7 fMOOC Project. <https://projekt.beuth-hochschule.de/fmooc/> last accessed April 25, 2017.
- 8 <https://www.fokus.fraunhofer.de/en> last accessed April 23, 2017.
- 9 <https://www.izt.de/> last accessed April 23, 2017.
- 10 <https://www.hwk-berlin.de/ueber-uns/die-handwerkskammer/> last accessed April 23, 2017.
- 11 Krauß, C., Merceron, A., An, T-S., Zwickelbauer, Arbanowski, S.: The Smart Learning Approach - A mobile Learning Companion Application. Awarded Best Paper. In Proceedings of the Eighth International Conference on Mobile, Hybrid, and On-line Learning, eLmL 2016, Venice, Italy, 24-28 April 2016, p. 13-16, ISBN: 978-1-61208-471-8.
- 12 ExCELL Project. <https://www.excell-mobility.de/>, last accessed April 25, 2017.
- 13 ACM SIGSPATIAL Giscup 2016. <http://sigspatial2016.sigspatial.org/giscup2016/home/> , last accessed April 23, 2017.
- 14 BIM-FM Project. <http://www.bim-fm.net/> last accessed April 25, 2017.
- 15 IFaf Institute Berlin. <http://www.ifaf-berlin.de/> last accessed April 25, 2017.
- 16 Project FoerDaBa - Data management project with partners from the district administration Friedrichshain-Kreuzberg’s facility management and urban planning departments. In: Beuth-Press, page 16. http://www.beuth-hochschule.de/fileadmin/oe/pressstelle/beuth_presse/2016/Beuth_Presse_2-2016.pdf last accessed April 25, 2017.
- 17 <http://www.beuth-hochschule.de/gruenderscout/>, last accessed April 25, 2017.
- 18 <http://benhu.de/english/>, last accessed April 25, 2017.
- 19 <https://www.deutschlandstipendium.de/de/2319.php>, last accessed April 25, 2017.

EFFICIENT ORGANIZATION OF WORKSHOPS ON CREATING TEACHING MATERIALS IN THE PROMIS PROJECT

(Lublin University of Technology, Lublin, Poland)

Introduction

The PROMIS project was implemented in years 2014-2017 and funded by the European Union (EU) under Tempus programme. PROMIS abbreviation derived from the name of the project: „Professional Master's Degree in Informatics as a Second Competence in Central Asia”. The project is an enlargement of the two Tempus previous projects which became implemented successfully [1]:

- “Master in Computer Science as a Second Education for Social Science Graduated”, project number: CD-JEP 26235-2005, the implementation period: 2006-2009;
- “Network Europe – Russia – Asia of Masters in Informatics as a Second Competence (ERAMIS)”, project number: 159025-TEMPUS-1-2009-1-FR-TEMPUS-JPCR, the implementation period: 2010-2013.

The objective of the PROMIS project is to extend and improve the ERAMIS academic network of partner universities in Central Asia (CA). This Master's degree called “ISC - Informatics as a Second Competence” and it is aimed for people who have other than computer science basic higher education.

The PROMIS project is coordinated by the Grenoble Alps University (Grenoble, France) and involves 18 partners including five universities from EU countries (France, Germany, Finland, Lithuania and Poland), two universities from each of Central Asia countries (Kazakhstan, Kyrgyzstan, Uzbekistan, Tajikistan and Turkmenistan) and three partners from the ICT industry.

The principle of Tempus projects is the transfer of knowledge as well as skills and pedagogical experiences from EU universities to partner universities [2]. As main components of this transfer in the PROMIS project, there were following activities realized: developing course materials and other teaching resources in the cooperation between academics from EU and CA universities, developing the pedagogical skills of their usage among academic teachers in CA. All materials were written in English as a pattern for other languages. Materials in Russian or any local language were created only in special cases. These activities were realized in the form of workshops (seminars) placed at EU universities [4].

Workshops on creating teaching materials

The workshops on creating teaching materials were held in all EU universities and were dedicated to ten different common subjects. Each workshop, beside the organizer from EU university, had its leader from CA university. The details list of workshops' organizers and leaders is presented in Table 1.

Each workshop was attended by 2 academic teachers from each of 10 CA universities as well as by several teachers from EU universities (acting as experts). The number of participants in each workshop exceeded 20 and thus it was an important organizational challenge.

To better address the organizational issues and meet expectations of all participants, the workshops were realized in five phases [4]:

- Phase I. Strategic planning,
- Phase II. Preparation ,
- Phase III. Workshop realization ,
- Phase IV. After workshop activities,
- Phase V. Results evaluation.

Table 1. List of common courses and universities [4]

<i>Course</i>	<i>CA Leader University</i>	<i>EU Host University</i>
Algorithms and Data Structure	Kyrgyz State Technical University, Bishkek, Kyrgyzstan	Beuth Hochschule fuer Technik Berlin, Germany
Programming	Bukhara State University, Bukhara, Uzbekistan	
Data Mining and Data Warehouse	National University of Uzbekistan, Tashkent, Uzbekistan	Kaunas University of Technology, Kaunas, Lithuania
Databases	Al-Farabi Kazakh National University, Almaty, Kazakhstan	
Human Machine Interaction	Kyrgyz State Technical University, Bishkek, Kyrgyzstan	Grenoble Alps University, Grenoble, France
Mathematics for Computer Science	National University of Uzbekistan, Tashkent, Uzbekistan	
Operating Systems and Networks	Khujand State University, Khujand, Tadjikistan	Savonia University of Applied Sciences, Kuopio, Finland
Web Development	Tadjik University of Technology, Dushanbe, Tadjikistan	
Project Management	Turkmen Institute of Economics and Management, Ashgabat, Turkmenistan	Lublin University of Technology, Lublin, Poland
Software Engineering	Osh Technical University, Osh, Kyrgyzstan	

The Phase III. *Workshop realization* were carried out by each EU host university in its own way. This allows other participants to become acquainted with the different organizational and educational practices, the culture of work, and participate in additional activities beyond the framework of the workshop.

Scheme of workshop organization at Lublin University of Technology

In 2015 Lublin University of Technology organized the workshops on two common subjects (see Table 1): *Project Management* and *Software Engineering*. Those subjects were completely unfamiliar as academic courses to many of academic staff from CA universities, because they had not previously taught them at a university. This forced the special organization of workshops to ensure their effectiveness and high quality of results.

In the lack of sufficient competence and experience among academic teachers from CA, a special scheme for the organization of Phase III was implemented at Lublin University of Technology (LUT). The scheme consisted of the following actions (in brackets the bodies responsible for the task were indicated):

- Task IIIa. Verification, updating and final approval of syllabus existing within ERAMIS academic network (EU universities, performed using networking).
- Task IIIb. Preparation of preliminary materials (as lectures' notes, practices' or laboratories' instructions and additional reading) in English for each course (LUT academic staff).
- Task IIIc. Distribution of preliminary materials (LUT coordinator, performed using networking).
- Task IIId. Getting acquainted with materials by CA teaching staff as well as proposing amendments (CA universities).
- Task IIIe. Realization of the workshop at LUT - CA teachers training, developing of final version of teaching materials in English as well as preparing English-Russian dictionary of keywords (all partners).

- Task IIIf. Final refining and distribution of materials among CA universities (LUT, performed using networking).
- Task IIIg. Translation of materials into Russian and national/local languages (CA leaders, performed using networking).

The schema of workshop organization at the LUT helped to achieve fruitful work. Moreover, time during the workshops was used effectively (it was 5 working days which results in 40 working hours at LUT for over 20 participants, what give in total more than 800 working hours). Additionally, some activities were carried remotely (using networking), which increased the time and financial efficiency of the task.

All activities performed using networking were held via Moodle e-learning platform [5], which was provided by project coordinator: Grenoble Alps University (France).



Fig. 1. The participants during workshops' activities

The results of workshop

The main objective of the workshop was to prepare participants to independently carry out activities as well as providing them with the necessary materials and skills. The main direct results of LUT workshops are following:

- modernized syllabuses,
- materials for lecture and practice ready to use in English,
- skills in using dedicated software tools,
- English-Russian dictionary of subject's keywords,
- knowledge and pedagogical skills.

The pedagogical skills acquired during workshops concerned not only presentation skills or guidelines how practice a concept by working on a problem set. During the workshops the academic staff from CA were introduced to reverse teaching approach (flipped classroom) [6]. The academics from EU and CA worked on ideas for class activities (e.g. brainstorming, discussions, preparing a test, individual contest), students' assessments (e.g. self- or peer-assessment) and students' preparation before classes (e.g. watching a video, reading a book chapter or tutorial).

The participants of workshops at LUT had possibility to participate in additional educational, research and cultural activities such as:

- visit the scientific and didactic laboratories of the Institute of Informatics at LUT (including the Laboratory of Motion Analysis and Interface Ergonomics),

- establish scientific and educational contacts between workshop participants,
- visit the tourist and cultural showplaces (Lublin Old Town, The Zamoyski Museum in Kozlowka) as well as get know local cuisine.

Summary

All the outcomes planned for the PROMIS project for workshops at LUT have been implemented, despite the novelty of the subject matter for many CA academics.

The realization and the atmosphere during the workshops were highly appreciated by the participants. This is expressed in the e-mail addressed to the organizers and participants by the representatives from universities of Turkmenistan with the following thanks:

"Dear organizers and participants of the seminar "Professional Master's Degree in Computer Science as a second jurisdiction in Central Asia (PROMIS)" in Lublin, Poland 22.06.2015-26.06.2015! I want to thank you for an exciting educational and cultural program, helping us time our visit to Lublin. We are impressed by your excellent and efficient way of working, creative approach to developing the concept of the event, patience and desire to do everything, even the most unexpected, the wishes of the participants. I am sure that a working meeting in Lublin, will contribute to the strengthening and further development of a trusting partnership. I would like to particularly thank the team of the Lublin University of Technology for their responsiveness, the ability to quickly respond to changing circumstances and the commitment to make of any event something special. Participants from the Turkmen side had agree at time and a lot of positive emotions from your seminar organized in Lublin in the framework of PROMIS. Give also to thank all those who participated in organizing and holding our meeting. With Respect Batyr Hezretov, Azat Atayev (Turkmen State Institute of Transport and Communications) !!!!"

Acknowledgements

The work described in this paper has been carried out within the framework of the Tempus programme (Ref. No. 544319-TEMPUS-1-2013-1-FR-TEMPUS-JPCR), partially funded by the European Commission.

References

- 1 Adam, J-M., Bardou, D. (2014). From ERAMIS to PROMIS: Extending and Improving a Master's Degrees Network in Central Asia. Proceedings of EDULEARN14: 6th International Conference on Education and New Learning Technologies Barcelona, pp. 2120-2129.
- 2 Bardou, D., Adam, J.-M., Merceron, A., Ripphausen-Lipa, H., Milosz, M., Kapocius, K., & Toppinen, A. (2015). Pedagogy Sharing in the PROMIS Project. Proceedings of 2015 IEEE Global Engineering Education Conference (EDUCON), 18-20 March 2015, Tallinn, Estonia, pp. 587-593.
- 3 Adam, J-M., Kapocius, K., Merceron, A., Milosz, M., Toppinen, A., Bardou, D. (2015). European experiences on the collective production of common master's level teaching materials for 10 Central Asian universities. Proceedings of ICERI 2015: 8th International Conference of Education, Research and Innovation, pp. 2035-2044.
- 4 Milosz, M (2015). Joint seminars for preparation of learning materials in the PROMIS Tempus Project, In: Tempus IV in Uzbekistan, Tashkent, pp. 168-172.
- 5 Moodle, open-source learning platform, <https://moodle.org/>.
- 6 Kim, M. K., Kim, S. M., Khera, O., & Getman, J. (2014). The experience of three flipped classrooms in an urban university: an exploration of design principles. The Internet and Higher Education, 22, pp. 37-50.

СОТРУДНИЧЕСТВО ОШТУ С ЕВРОПЕЙСКИМИ УНИВЕРСИТЕТАМИ В РАМКАХ ПРОЕКТА TEMPUS-PROMIS

(Ошский Технологический Университет, Ош, Кыргызстан)

Ошский технологический университет является ярким представителем инженерной школы Кыргызстана, стремящимся сохранить и приумножить достижения отечественного и мирового инженерно-технического образования. Университет включает в себя 7 факультетов, институт переподготовки кадров и технологический колледж.

ОшТУ в рамках программы TEMPUS участвует в нескольких проектах и в этой статье речь идет о проекте 544319-TEMPUS-1-2013-1-FR-TEMPUS-JPCR и о его роли в жизни университета.

Проект реализуется совместно с 5 университетами из стран Евросоюза (Германия, Литва, Польша, Финляндия, Франция), 10 университетами из стран Центральной Азии (Казахстан, Кыргызстан, Узбекистан, Таджикистан, Туркменистан) и 3 предприятиями Европы.

Основной целью проекта **TEMPUS-PROMIS** является расширение магистерской сети в Центральной Азии, и её усовершенствование, путём значительного усиления профессиональной подготовки магистров, что является возможным при укреплении связей между университетами и предприятиями.

Проект **TEMPUS-PROMIS** стартовала в ОшТУ в 2015 году и в конце 2014-2015 учебного года Министерством Образования и Науки Кыргызской Республики был утвержден экспериментальный учебный план по подготовке магистров направления “Информатика и Технология Программирования” (ИТП). А также с поддержкой проекта на базе кафедры “ПОВТАС” создана современная лаборатория для магистрантов и в 2015-2016 учебном году были приняты первые магистранты.

Развивающие эффекты проекта для кыргызских ВУЗов можно представить в следующих контекстах.

На уровне кафедры, ведущей подготовку специалистов

- возможность обмена европейским опытом подготовки специалистов в области информационных технологий;
- внедрение инновационных методов, технологий подготовки специалиста, становление теоретической, профессионально-личностной, инструмен-тальной компетентности преподавателей кафедры;
- усиление практико-ориентированной направленности университетского образования в области информационных технологий;
- внимание к организации и содержанию самостоятельной работы магистрантов (опора на проблемное обучение, требующее самостоятельного поиска решения, создание рабочих материалов, обеспечивающих сопровождение учебного процесса магистрантов в освоении ими тех или иных курсов и др.);
- подготовка пакета учебно-методических материалов для магистрантов и преподавателей.

На уровне факультета и ВУЗа

- Укрепление связей между университетами;
- Привлечение местных предприятий в учебный процесс.

В рамках проекта “Professional Master’s Degree in computer science as a second competence in Central Asia” (PROMIS) следующие 10 преподавателей факультета кибернетики и информационных технологий прошли стажировки в университетах Европы.

1. Аттокуров У., Кокозова А.- ВНТВ, Германия
2. Жолдошова Ч., Кочконбаева Б. – UPMF, Франция

3. Сайдамов Ш., Абдыраева Н.- LUT, Польша
4. Кадыркулова Н., Мусаева Р.- KTU, Литва
5. Матисаков Ж.- SUAS, Финляндия

Вышеперечисленные преподаватели читают лекции и проводят практические занятия по магистерской программе ИТП. Преподаватели Кадыркулова Н., Кочконбаева Б., Абдыраева Н. подготовили учебно-методический материал по дисциплинам “Технология программирования” и “Базы данных”.

Связи между предприятиями

Участие предприятий в проекте предполагает их участие в организации учебного процесса, и в организации производственной практики для магистрантов.

ОшТУ, в рамках проекта TEMPUS-PROMIS, сотрудничает с местными предприятиями такими, как ОшЭлектро, ОшПВЭС, ОшТЭЦ, КыргызТелеком и др.



Магистранты направления ИТП проводят предквалификационные практики в вышеперечисленных предприятиях по заключенному договору в течении 2 месяцев. После завершения магистратуры специалисты могут работать в этих организациях.

Семинары в ОшТУ

По проекту PROMIS-TEMPUS в Ошском технологическом университете прошел семинар для преподавателей и магистрантов по направлению “Информатика и Технология Программирования” с 17-октября по 29 октября 2016 года. С 17-октября по 21-октября семинар проходил по заранее планированному расписанию, преподавателями Европейских ВУЗов вошедших в данный проект. Семинар проходил по двум направлениям: “Generic programming in C++” и “Software design using UML”. А также с 24- по 28-октября семинар проходил по двум другим предметам: “Methodology for Algorithms and data structures teaching” и “Storing and querying data from the web: Introduction into xml- and semantic technologies”. В результате семинара магистранты и преподаватели получили новые методы обучения и изучали современные языки моделирования систем.

Мобильность студентов

В рамках проекта магистранты направления ИТП, Сатыбалдиева Чынара и Улугбек кызы Дарыгул, с 20 февраля по 3 марта 2017 года прошли обучающий тренинг «Программирование» в Университете прикладных наук Beuth в Берлине. По итогам тренинга магистранты успешно защитили финальный проект на тему «Hotels rating» и получили сертификаты. Такие возможности дает магистрантам расширять свои навыки.



Выводы

Международное сотрудничество высших учебных заведений представляется на сегодняшний день важным аспектом в развитии сферы образовательных услуг.

С учетом международного опыта повышается авторитет национальной системы образования. В рамках проекта TEMPUS-PROMIS многие преподаватели обменялись опытом и стали использовать методы преподавания лучших лекторов европейских университетов. Открылись научные лаборатории, преподаватели и магистранты проходили тренинги.

Литература

- 1 Материалы с сайта ОшТУ. [Электронный ресурс]/www.oshtu.kg
- 2 Материалы с сайта проекта TEMPUS-PROMIS ЕНУ им. Л.Н. Гумилева. [Электронный ресурс]/http://promis.enu.kz

МОБИЛЬНОЕ ПРИЛОЖЕНИЕ «СТУДЕНТ»

(магистрант, Евразийский национальный университет им. Л.Н.Гумилева, г. Астана, Казахстан)

Современный человек делает все, для того, чтобы достигнуть максимального комфорта. Сегодня одним из желаний большинства людей является выход в Интернет. Причем они всегда хотят оставаться онлайн. Именно поэтому огромной актуальностью пользуется такая разработка мобильных приложений под Android. Все это стало актуальным вместе с появлением мобильного Интернета. Во время поездок всегда есть возможность подключиться к сети с помощью телефона, планшета или другого устройства. Но, сразу же стоит отметить, что без специальных приложений вряд ли бы была достигнута необходимая эффективность. В наше время многие привыкли к удобству гаджетов, которые всегда находятся под рукой и позволяют постоянно оставаться в курсе всех событий. Как известно, мобильные телефоны с платформой Android появились давно, но всё ещё пользуются огромным спросом у потребителей. Этот закономерный результат сложился не столько от быстроты и удобства операционной системы, хотя это и является главной её чертой, сколько от большого разнообразия мобильных приложений, которые сотнями можно скачать в свободном доступе. Операционная система Android развивается каждый день, и десятки новых приложений поступают во всемирную паутину ежедневно. Уже сегодня современного пользователя Android нельзя представить без набора различных приложений для своего устройства. Почти у всех есть приложения социальных сетей twitter и Facebook. Путешественники закачивают себе навигаторы по городу и карты, студенты будут рады переводчику и какой-нибудь игре на Android. Автовладельцам понравятся программы по расходу топлива и выбору автомобиля, а многие другие будут пользоваться таким программами как word или excel, когда работу нужно срочно сделать, а компьютера под рукой нет.

Передача мыслей, слов и изображений на расстояние за сотую долю секунды – это такая обыденная реальность, что далеко не каждый может себе даже представить жизни без этих возможностей. Неограниченный доступ к информации за пять лет сделал то, на что не были способны сотни лет естественного развития и эволюции.

И отдельным самостоятельным явлениям стали мобильные приложения. Это те самые маленькие программки, которые способны создать целый отдельный мир в пределах одного компактного электронного устройства.

Для того, чтобы доказать необходимость, эффективность и успешность этого вида технического новшества, достаточно перечислить всего лишь несколько причин, не вдаваясь даже особенно в их своеобразную и привлекательную специфику.

Во-первых, удобство. Посещение обычных сайтов и в целом выход в интернет с компактных карманных устройств часто сопровождается неким неудобством, а особенно – потерей времени и денег. Они слишком часто уходят на загрузку тех дополнительных (но не всегда необходимых) мультимедийных элементов, которыми насыщены многие виртуальные ресурсы. Мобильные приложения загружаются значительно быстрее и могут работать даже без подключения к сети, то есть оффлайн. В этом можно выделить не один, а сразу несколько плюсов: экономия времени и средств, возможность использования приложений в отсутствии связи или средств на счету, быстрое действие (за счет фильтрации информации и задействования внутренних резервов устройства).

Во-вторых, все мобильные приложения специально создаются под конкретную модель устройства. Это дает возможность не только максимально синхронизировать их работу, но и сделать ее максимально синергетичной и результативной. В этот же подпункт

можно отнести и те обновления, которые регулярно создаются компанией-разработчиком и периодически скидываются на приложения, даже вне зависимости от того, подключены они к сети на данный конкретный момент или нет (если связи нет, обновления просто будут ожидать ее восстановления).

В данной статье рассматривается разработанное мной мобильное приложение «Студент» для гаджетов на базе операционной системы Android.

Научная новизна данной разработки очевидна, поскольку проведенные исследования в области существующих мобильных приложений показали, что для операционной системы Android приложения данной направленности и тематики разработаны не были.

Практическая значимость разработанного приложения заключается в возможности его использования студентами и преподавателями любого ВУЗа, а так же она будет интересна и для абитуриентов.

Основное назначение данного приложения заключается в оказании студенту помощи в его повседневной деятельности и учебе.

Основные функции программы:

- Учебный календарь;
- Расписание занятий;
- Задания, УМКД по дисциплинам;
- Пропуски;
- Объявления;
- Новости;
- Напоминание о важных событиях;
- Будильник;
- Текущая успеваемость;
- Текущая задолженность;
- Связь с куратором.

Данное приложение позволяет, как получать данные из сети Интернет, так и хранить их и редактировать в локальной базе данных. Это удобно, когда по какой-либо причине отсутствует подключение к Интернет, или нет доступа к сайту ВУЗа.

Интерфейс приложения «Студент» представлен на Рисунке 1.

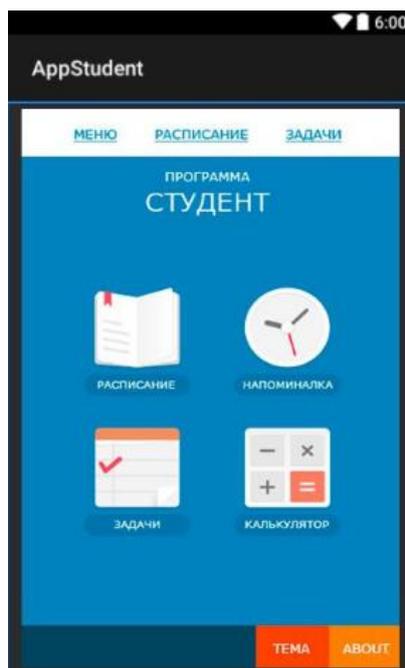


Рисунок 1. Интерфейс приложения «Студент»

Приложение выполнено в бело-голубых тонах с возможностью менять цветовые схемы по предпочтениям пользователя. На главной странице приложения находится меню программы и настраиваемая панель быстрого доступа к функциям программы. Приложение имеет простой, интуитивно понятный интерфейс, что позволяет использовать его даже неподготовленному пользователю.

В данной статье дано краткое описание основных функций и пользовательского интерфейса разработанного мобильного приложения «Студент» для устройств на базе операционной системы Android. Преимуществами приложения являются функциональность и простота в настройке и использовании.

Список использованной литературы:

- 1 Хашими С., Коматинени С., Маклин Д. Разработка приложений для Android. СПб. Питер, 2011. — 736 с.
- 2 Рето Майер. Professional Android 2: Application Development second Edition, Эксмо, 2011 – 672 с.

УДК 004

У.А.ГАФАРОВА

РЕАЛИЗАЦИЯ ПРОЕКТА “PROMIS” В ХУДЖАНДСКОМ ГОСУДАРСТВЕННОМ УНИВЕРСИТЕТЕ ИМЕНИ АКАДЕМИКА Б.ГАФУРОВА

(Худжандский государственный университет им.Академика Б.Гафурова,Таджикистан)

Системы высшего образования всех стран мира переживают эпоху быстрых перемен. Эти перемены обусловлены многими причинами: все нарастающей глобализацией экономики и политических процессов, переориентацией научной деятельности на преимущественно инновационные исследования, постоянно изменяющимися требованиями рынка труда к выпускникам университетов. Европейские страны в течение последних десятилетий взяли курс на создание Единого европейского пространства высшего образования (ЕНЕА), к которому могут подключаться и страны-соседи. Одним из основных инструментов, обеспечивающим развитие системы образования в этом направлении, является широко известный Болонский процесс. Таджикистан, вместе с соседними странами Средней Азии, также внедряет у себя современные европейские методы усовершенствования высшего образования.

Решать проблемы, связанные с адаптацией к меняющимся условиям, отдельно взятым университетам, и даже образовательным системам отдельных стран, становится невозможным. Быть успешным в наши дни можно только через расширение международного партнерства в области образования. Очень перспективным для развития таких партнёрских взаимоотношений между университетами Таджикистана и образовательными учреждениями Средней Азии и Европы являются проекты Темпус и Эрасмус+ Европейского Союза.

Проект PROMIS – Магистратура по специальности «Информатика - вторая компетенция» стартовал в октябре 2013 г. и продлился до июля 2017 г. Глобальная цель проекта состоит в способствовании дальнейшему развитию высшего образования стран Центральной Азии посредством использования лучших сторон европейской системы. Такое внедрение планируется в одно, но очень важное направление: «информатика», что соответствует национальным приоритетам этих стран. Задачи проекта:

- вклад в развитие подготовки специалистов уровня «магистр» по специальности: «информатика - вторая компетенция»;
- популяризация болонских принципов в Таджикистане;
- установление тесных контактов с организациями-работодателями, чтобы обеспечить

соответствие разрабатываемой программы требованиям рынка и нуждам социального развития на региональном и национальном уровнях.

Магистерские программы по направлению «Информатика - Вторая Компетенция» (ИВК) позволяют студентам, имеющим диплом по первой дисциплине, приобретать за 2 года основные компетенции по информатике.

В проекте принимают участие специалисты управления и преподаватели ведущих университетов стран Европейского Союза: Франции, Германии, Польши, Финляндии и Эстонии. Координатором является профессор Жан Мишель Адам и менеджером проекта является Анна Голубцова (Университет Гренобль-Альп, Франция). Из университетов Таджикистана представлены Таджикский технологический университет и Худжанский государственный университет имени академика Б.Гафурова. За время деятельности проекта достигнуты значительные успехи.

Для студентов магистратуры ХГУ имени академика Б.Гафурова Исмоиловой Дилафруз и Азамова Набиджона по разработанной программе, была возможность участвовать в стажировке в Альпийско-Гренобльском университете Франции, где они прослушали несколько курсов, подготовленных европейскими профессорами. Студенты получили кредиты по европейской системе ECTS, которые будут засчитаны в их таджикских университетах как спецкурсы. В рамках проекта на факультете телекоммуникации был открыт компьютерный класс.

Для преподавателей и студентов университета были организованы двухнедельные курсы, которые проводили европейские лектора – профессора Адам Жан-Мишель (Франция), Марек Милоч и Магдалена Борис (Польша), Хейке Марина Рипхаузен (Германия).

Были проведены тренинги для профессорско-преподавательского состава. Более десяти преподавателей университета прошли стажировку и курсы повышения квалификации в европейских университетах, в том числе: доцент Музаффаров Дилшод – в Берлинском университете и в Альпийско-Гренобльском университете Франции; доцент Нуриддинова Манзура в Люблинском университете Польши; доцент Рахимов Чамшед – в университете Куопио Финляндии и Альпийско-Гренобльском университете Франции; доцент Хамдамов Шерали - в Берлинском университете и Каунасском университете Литвы; Маликов Абдумумин в Каунасском университете Литвы.

Профессор Гафурова У. как координатор проекта в ХГУ им. Академика Б.Гафурова участвовала на всех координационных встречах, состоявшихся в Гренобле, Берлине, Ташкенте, Бишкеке, Оше, Душанбе и Астане.

В рамках проекта были разработаны и модернизированы следующие документы: учебный план этой специальности для образовательного уровня «магистр»; образовательная программа с пояснительной запиской; силлабусы и модули с учетом компетентностного подхода и европейского опыта, была обновлена библиотека.

Активное участие в проектах позволит таджикским университетам решить целый ряд задач стратегического развития университетов.

Проект призван оказать содействие реформированию и модернизации высшего образования в странах-партнерах, и предназначен для: специалистов учебно-методических управлений, преподавателей и студентов, потенциальных работодателей, представителей органов управления образованием (на республиканском и региональном уровнях), специалистов служб занятости, а также для представителей других направлений, заинтересованных в развитии и принятии международных стандартов в области образования.

Для организации работы по проекту используется сайт, где размещается актуальная информация о ходе реализации проекта, о его достижениях, электронные публикации, проводятся форумы участников. Научные и методические результаты работы участников опубликованы в сборниках трудов, в электронном и печатном виде, размещены в средствах массовой информации.

Активное участие в проектах позволит таджикским университетам решить целый ряд задач стратегического развития университетов, таких, как:

- организация научных исследований мирового уровня;
- реализация в университете основных требований Болонского процесса;
- увеличение количества реализуемых новых образовательных программ высшего профессионального образования (бакалавриат и магистратура);
- увеличение объёма и значимости на международном уровне методической и научной работы университета;
- привлечение иностранных вузов-партнеров для совместной работы;
- повышение профессионального уровня профессорско-преподавательского состава за счёт интенсивного изучения передового опыта таджикских и зарубежных коллег;
- рост мобильности студентов и профессорско-преподавательского состава в образовательных целях;
- модернизация учебных планов, разработка образовательных программ, образовательного стандарта с учетом европейского опыта.

Накопленный опыт сотрудничества с европейскими университетами показал высокую квалификацию их сотрудников и способствовал интернационализации и совершенствованию системы высшего образования.

Благодаря участию в проекте университеты имеют прекрасную возможность для интеграции в мировое образовательное пространство, повышения рейтинга среди высших образовательных заведений мира.

УДК 004

ЖУСУПОВА Н.А., ХАСЕНОВ Е.А.

СОЗДАНИЕ ОБРАЗОВАТЕЛЬНО-ИНФОРМАЦИОННОГО РЕСУРСА ДЛЯ ИЗУЧЕНИЯ МЕНТАЛЬНОЙ АРИФМЕТИКИ

(Евразийский национальный университет им. Л.Н. Гумилева, Астана, Казахстан)

На сегодняшний день, ментальная арифметика пользуется очень большой популярностью, как среди детей, так и взрослых. За последние полгода поступило больше количество запросов от клиентов по г. Астана, которые не могут посещать центры в связи с удаленностью, нехваткой времени и т.д, но хотели бы изучать ментальную арифметику. Чтобы воплотить их запросы в реальность, необходимо разработать сайт для онлайн обучения. Сайт мог бы охватить потребности клиентов не только в Астане, но и в других регионах. Создание информационно-образовательного онлайн ресурса окажет большую помощь всем желающим получить знания и навыки ментальной арифметики. Опыт работы центра интеллектуального развития “Advance” показывает, что на основе проведенного опроса сайт будет пользоваться большим спросом, как в крупных городах, так и в других населенных пунктах, имеющих доступ в интернет.

Для достижения указанной цели диссертационной работы ставятся и решаются следующие задачи:

- Разработка курсов обучения для разных возрастных категорий;
- Определение объема данных, предоставляемых обучающимся;
- Изучение способов предоставления информации;
- Разработка онлайн-методики обучения;
- Разработка методических материалов (видео, тестовые задания и т.д.);
- Разработка и реализация сайта;

Предполагается разработать информационный сайт по онлайн обучению ментальной арифметики, который будет включать в себя обучающие материалы по возрастам:

- 7-16 лет;
- 17 лет и старше.

К функциям ресурса будут относиться: обучение в формате видео, текстовые материалы, домашнее задание. Среднее время прохождения полного курса для дошкольного возраста будет составлять 2 года, школьникам- 1,5 года, взрослым- 1 год.

Если сравнивать традиционное образование с электронным образованием, то доказано, что в последнем скорость образования на 30-60% быстрее, но имеет низкое качество. Причиной низости качества электронного образования является пассивность и статичность электронных образовательных ресурсов (как правило, простой текстовый или графический материал) и отсутствие диалога с обучающимися в реальном масштабе времени. Ведь при изучении у обучающегося может появиться масса вопросов к ним, а у них просто нет возможности ответить на эти вопросы, которые с продолжением изучения могут только возрастать. Кроме того, современные методы контроля и оценки знания не всегда выдают объективную оценку знаниям. Так как сайт будет тесно сотрудничать с компанией «Advance», которая предоставляет услуги ментальной арифметики по РК и СНГ, на все возникающие вопросы смогут ответить инструктора компании, либо call center. Обучающийся будет изучать материалы на ресурсе, а уровень своих знания можно будет подтверждать только придя в центр, находящийся наиболее близко к клиенту. Там же и будут выдаваться подтверждающие свидетельства.

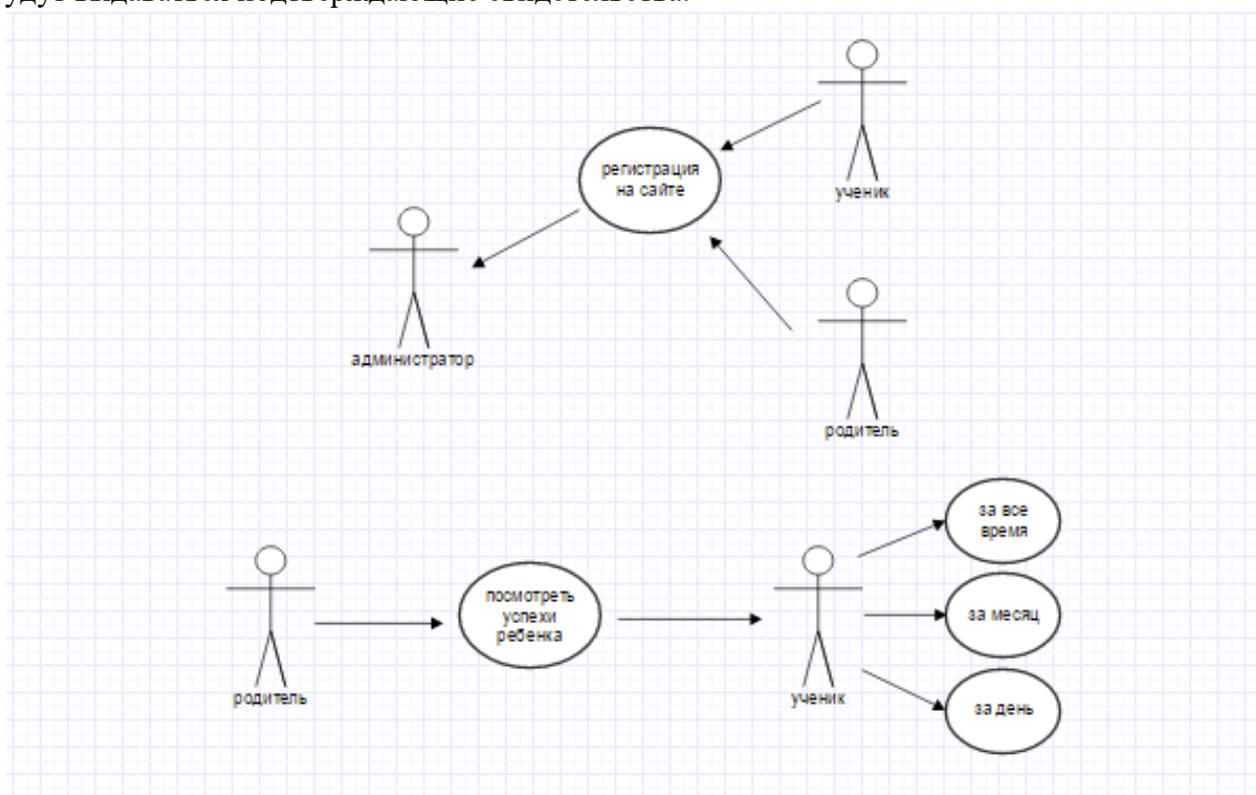


Рисунок 1

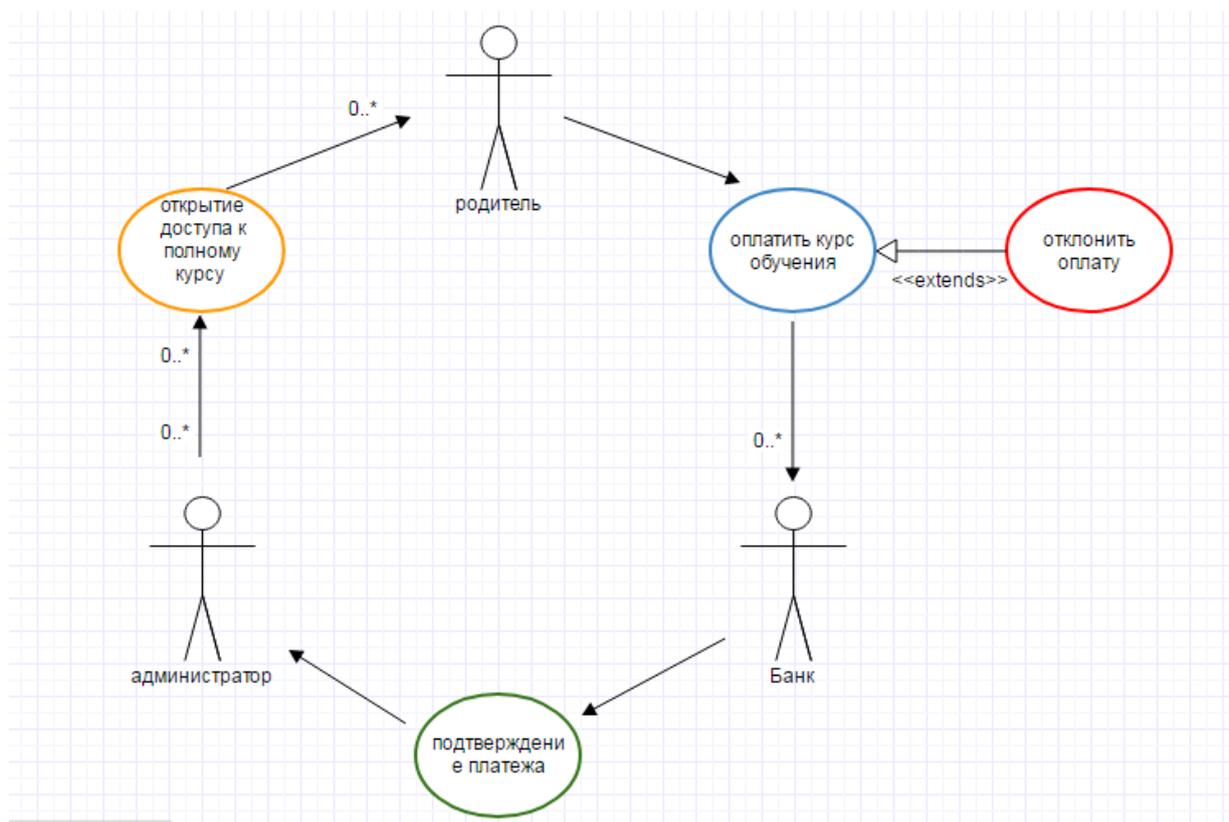


Рисунок 2.

С помощью унифицированного языка моделирования (UML), графическим языком для визуализации, построены визуальные модели образовательного сайта.

При визуальном моделировании на UML используются девять видов диаграмм, каждая из которых содержит элементы определенного типа. На диаграммах взаимодействия представлены связи между объектами; показаны, в частности, сообщения, которыми объекты обмениваются (рис.1,2).

Анализ спроса на данный вид обучения показывает, что создание электронного ресурса окажется своевременным решением запросов большого количества людей, не имеющих возможностей посещать учебные центры. К этой категории можно, отнести, например, инвалидов. Вообще, предоставление образовательных услуг для инвалидов, является актом большой гражданской и социальной значимости и способствует повышению рейтинга страны по показателям человеческого развития.

Список использованной литературы

1. Профессиональная разработка сайтов на Drupal 7; Питер - Москва, 2013. - 688 с.
2. Байков В. Интернет. Поиск информации и продвижение сайтов; Книга по Требованию - Москва, 2012. - 288 с.
3. Гарднер Л., Григсби Д. Разработка веб-сайтов для мобильных устройств; Питер - Москва, 2013. - 448 с.
4. Дронов Владимир HTML 5, CSS 3 и Web 2.0. Разработка современных Web-сайтов; БХВ-Петербург - Москва, 2011. - 416 с
5. Китинг, Джоди Flash MX. Искусство создания web-сайтов; ТИД ДС - Москва, 2012. - 848 с.

**РЕЗУЛЬТАТЫ РЕАЛИЗАЦИИ ПРОЕКТА TEMPUS PROMIS
“ПРОФЕССИОНАЛЬНАЯ МАГИСТЕРСКАЯ ПРОГРАММА ИНФОРМАТИКА КАК
ВТОРАЯ КОМПЕТЕНЦИЯ” В КГТУ ИМ. И. РАЗЗАКОВА**

(КГТУ им. И. Раззакова, Бишкек, Кыргызстан)

Идея создания магистерской сети «Информатика как вторая компетенция» была реализована в проекте ERAMIS. Благодаря результатам проекта ERAMIS многие специалисты из различных сфер обрели возможность получения второй компетенции по информатике, что позволило им применять информационные технологии в профессиональной деятельности.

Проект TEMPUS PROMIS является продолжением проекта TEMPUS ERAMIS, и ключевыми целями проекта являются расширение магистерской сети в Центральной Азии, а также её усовершенствование, путём значительного усиления профессиональной подготовки магистров.

КГТУ им. И. Раззакова является ведущим техническими университетами Кыргызстана, имеет богатые традиции и отличается высоким качеством подготовки выпускников. Университет имеет широкий опыт сотрудничества с международными организациями и участия в грантовых программах. В данном проекте КГТУ им. И. Раззакова выполняла миссию по передаче опыта запуска подобной магистерской программы в центральноазиатских университетах-партнерах.

В результате реализации проекта в КГТУ им. И. Раззакова были достигнуты следующие результаты:

В организации магистерской программы:

- Оборудован компьютерный класс для магистрантов за счёт средств проекта на общую сумму 12 487 евро (2015 г.)
- Оптимизирован и переработан учебный план магистерской программы (2016 г.)
- Продлена лицензия МОН КР на экспериментальную магистерскую программу «Информатика и технология программирования» с 2016 года бессрочно.

В повышении качества магистерской программы:

- Преподаватели академического состава прошли стажировки в европейских университетах-партнерах (6 стажировок в 2015 г.)
- Студенты магистерской программы прошли стажировки в европейских университетах-партнерах (2 мобильности в 2017 г., BUAS, Берлин)
- Организованы семинары тренинги для студентов КГТУ профессорами европейских университетов – партнеров (4 семинара-тренинга в 2016 г.)
- В ходе работы семинаров для преподавателей обновлены рабочие программы и syllabus дисциплин магистерской программы (10 дисциплин)
- Разработаны и изданы 8 учебно-методических пособий для студентов (2015-2017 г.)
- В рамках повышения уровня владения английским языком 4 преподавателя КГТУ им. И.Раззакова прошли курсы Английского языка, что позволило подготовку учебных материалов на английском языке в объёме до 25% от общего материала курсов по 4 дисциплинам.

Обучающие семинары-тренинги для преподавателей задействованных в магистерской программе были организованы в европейских вузах-партнерах проекта и включали анализ и обсуждение содержания общих курсов магистерской программы. Особое внимание было уделено реверсивным методам обучения. Европейские партнеры поделились своим опытом в методике организации и проведения практических занятий, что способствовало

эффективному достижению целей проекта. Материалы по методике преподавания и содержанию рабочих программ курсов были распространены среди преподавателей, вовлеченных в реализацию проекта. Шесть преподавателей КГТУ им. И. Раззакова участвовали в семинарах-тренингах по соответствующим курсам в ведущих европейских университетах-партнерах:

Токмергенова А.З.- Программирование (BUAS, Берлин); Шабданов М.А.- Алгоритмы и структуры данных (BUAS, Берлин); Исраилова Н.А. -Операционные системы (SUAS, Куопио); Момуналиева Н.Т.- Человеко-машинное взаимодействие (UPMF, Гренобль); Шаршеева К.Т. – Управление проектами (LUT, Люблин); Тультемирова Г.У. (КТУ, Каунас)

В результате совместной деятельности преподавателей проекта для обучения используются новые методические материалы и технологии обучения.

Студенты- магистранты 1- го года обучения Кудайберди уулу Кубат и Миррахимова Наима получили возможность участвовать в семинарах-тренингах по курсам “Программирование” и “ Базы данных” организованных в Берлинском университете прикладных наук (февраль, 2017 г.). Студенческая мобильность в рамках проекта позволила студентам получить прекрасную возможность познакомиться с культурой и системой образования другой страны, повысить уровень иноязычной компетенции, завести новых друзей, приобрести бесценный опыт. Этот опыт мотивировал студентов на активное участие в программах академической студенческой мобильности, что позволит им повысить конкурентоспособность на рынке труда.

В установлении связей с предприятиями:

- На базе КГТУ им. И. Раззакова организована и проведена международная конференция «Опыт сотрудничества университетов с предприятиями» (октябрь, 2015 г.), где обсуждались формы сотрудничества и взаимодействия университетов с предприятиями;
- Для преподавательской деятельности по магистерской программе привлечены сотрудники предприятия «IT-сервис» Дюшембеков Жоомарт и Аликулов Алишер по курсам “Web-программирование”, “Java-программирование”;
- Заключены договора с предприятиями для прохождения практики студентами;
- Участие представителей предприятий при защите магистерских диссертаций.

Для успешной реализации учебного процесса, безусловно, необходима её эффективная организация, которая в первую очередь заключается в оптимальном сочетании теории и практики. В целях достижения такого сочетания, требовалось привлечение к учебному процессу представителей предприятий и наращивание доли участия предприятий в организации учебного процесса. Для укрепления таких связей, было запланировано вовлечение местных предприятий в учебный процесс магистерской программы, по примеру европейских предприятий-партнёров проекта PROMIS. Участие предприятий в проекте предполагает их участие в организации учебного процесса, в организации производственной практики для магистрантов. Таким образом, чередование учебной нагрузки и практической работы на предприятии должно активно поддерживаться на 2-м курсе обучения: студент должен работать неполный рабочий день на предприятии и опыт его работы, накопленный за время практики на производстве на 2-м курсе обучения, будет являться основой для его выпускной квалификационной работы

В развитии международных связей с университетами-партнерами:

В результате реализации проекта PROMIS TEMPUS КГТУ им. И. Раззакова получила право участвовать в новых проектах по программе международной академической мобильности ERASMUS+ с университетами-партнерами по проекту:

- Гренобльским университетом Альп (3 студенческих мобильностей на один семестр в 2017 году)
- Люблинским технологическим университетом (2 преподавательских мобильностей в 2017 году)

С вышеуказанными университетами заключены соглашения о сотрудничестве до 2020 года.

Участие в международных проектах такого уровня стимулирует и мотивирует студентов и преподавателей активно участвовать в различных программах по международной мобильности, что обеспечивает последовательную интеграцию Кыргызстана в мировое образовательное пространство.

Заключение

Образовательная программа по подготовке магистров, разработанная в рамках проекта TEMPUS PROMIS, отвечает требованиям государственных стандартов КР третьего поколения. Разработка и реализация магистерских программ в области техники и технологии ведущим инженерным университетом республики является важным шагом в реализации Болонского процесса в Кыргызстане. Опыт, полученный КГТУ им. И. Раззакова в рамках реализации проекта, будет распространен Учебно-методическим объединением университета, которое отвечает за разработку проектов государственных образовательных стандартов для программ в области техники и технологий.

УДК 004

Д.МУХАМЕДОВА

РЕАЛИЗАЦИЯ ЕВРОПЕЙСКОГО ПРОЕКТА - 544319-TEMPUS-1-2013-1-FRTEMPUS-JPCR PROMIS: «ИНФОРМАТИКА КАК ВТОРАЯ КОМПЕТЕНЦИЯ» В НАЦИОНАЛЬНОМ УНИВЕРСИТЕТЕ УЗБЕКИСТАНА ИМЕНИ МИРЗО УЛУГБЕКА

(Национальный университет Узбекистана им. Мирзо Улугбека)

Современные потребности общества в развитии производительных сил и повышения качества жизни делают актуальным наличие знаний и навыков (специалистов различных сфер деятельности) для использования компьютерных технологий в своей работе. Наличие у специалистов опыта работы на компьютерах по большей части ограничивается использованием офисных приложений.

Очевидно, что для решения управленческих и административных задач, проведения научных исследований этого недостаточно. Ощутимый эффект даёт использование специального программного обеспечения. Разработка программного обеспечения основывается на умении формального описания сферы деятельности (постановки задачи) и выбора компьютерных (вычислительных) моделей для её реализации.

В настоящее время организации и предприятия Республики испытывают потребность в квалифицированных кадрах, умеющих эффективно применять на практике современные информационные и коммуникационные технологии (ИКТ). Удовлетворить эти потребности можно за счёт приобретения дополнительных знаний в области ИКТ специалистами различных не профильных сфер деятельности.

Свидетельством важности применения информационных технологий в жизни общества служит принятая правительством Республики Узбекистан государственная программа “Электронное правительство”.

Системы высшего образования должны оперативно удовлетворять растущий спрос на специалистов, работающих на стыке нескольких специальностей. Начинать нужно с разработки новых учебных программ. Содержание учебных программ должно учитывать особенности применения ИКТ в разных предметных областях. Следует исходить из того, что работающие в этих областях кадры изначально не готовились как специалисты по направлениям ИКТ. Дополнительная компетенция, полученная ими при обучении, позволит повысить производительность труда и эффективность принятия управленческих решений.

Подготовка специалистов с дополнительной компетенции по ИКТ реализуется на уровне магистратуры. Специальность магистратуры “Информатика как вторая компетенция”, имеет 20 летнюю историю в Европейских ВУЗах и странах ближнего зарубежья.

Национальный университет Узбекистана имени Мирза Улугбека стала участником-партнером Европейского проекта - 544319-TEMPUS-1-2013-1-FRTEMPUS-JPCR PROMIS: «Информатика как вторая компетенция ». В 2015-2016 учебном году в рамках выполнения проекта в НУУз было открыто специальность 5А330204-Информационные системы (по непрофильным отраслям) для лиц, не имеющих базового образования по ИКТ, с целью подготовки специалистов, ориентированных на использование ИКТ в своих сферах деятельности.

Данный проект включает в себя 5 европейских университетов, 3 европейских предприятия и 10 центрально-азиатских институтов, в том числе два университета Узбекистана - Национальный университет Узбекистана им. Мирзо Улугбека, Бухарский государственный университет.

В первую очередь было сформирована команда исполнителей проекта, которая состояло из двух составляющих - административная часть (2 человека) и преподавательский состав (10 человек). Были сделан соответствующий приказ ректора НУУз. по формированию команды.

Деятельность команда НУУз начала согласно календарному плану на 2014 год, в котором было необходимо изучать потребность рынка труда Узбекистана в специалистах ISC. Был произведен опрос студентов и представителей фирм об необходимости открытия магистратуры ISC и потребности их потребности в производстве. Результаты опроса приведены в таблицах 1 (студенты) и таблице 2 (представители фирм) .

Table 1. Результат опроса студентов

Specialties	Would you be interested in this master's program “Informatics as a Second Competence”?		Are you ready to apply for this master’s degree?	
	Yes	No	Yes	No
Economy	19	1	20	0
Mechanics	16	3	15	4
Mathematics	31	1	24	8
Psychology	19	4	18	5
Philosophy	25	3	23	5
TOTAL	110	12	100	22

Из 122 опрошенных 38 студентов, пожелали получить компетенцию в области сетевых технологий, 73 в области разработке программного обеспечения и 23 в других сферах IT (большинство механики и математики, они больше информированы в области IT). Данная информация может служить для определения состава дисциплин магистратуры.

Выводы на основе анализа результатов опроса представителей фирм следующие:

- средняя потребность в кадрах по IT в 2015-16 г. соответственно примерно,4-5 специалистов;
- есть необходимость в информатизации деятельности компании и в повышении квалификации сотрудников в области IT;
- относительно большая потребность в разработчиков программного обеспечения, администраторов сети и web мастерах;
- компании/организации желают повысить квалификацию по IT в среднем 20% своих сотрудников;
- все компании готовы поддержать своих работников укорченным рабочим днем, большинство оплатой обучения, отпуском во время экзаменов и т.д.;
- компании готовы сотрудничать с университетом в в трудоустройстве выпускников, а также предоставлении проведения практик.

Table 2. The result of a survey of company representatives

Questions/Company (Organization)	НИИ «UzGASHKLIT» (Кадастр)	Firm “Info-Medical technologies”	Firm “ZeroSoftService”	UCD MICROS	Firm “Hard Solutions”
A. Information about company/organization					
4. When the company/organization was created?	1964	2004	2012	1982	2001
5. Sphere of operation:	services	ICT	ICT	Sales and Manufacturing	IT
6. Total number of employees:	300	15	12	120	25
7. Number of employees in IT:	12	12	7	40	15
8. Possibility of internships:	Yes	Yes	Yes	Yes	Yes
B. Needs for specialists					
1. How many IT specialists do you plan to hire in 2015?	4	3	5	5	3
2. How many IT specialists do you plan to hire in 2016?	6	3	8	8	5
3. What kind of IT specialists do you need?	Software developers, Office workers	system administrator, web master	Software developers	Software developers	Software developers
4. Do you need IT specialists with additional competencies in the next fields:	mathematics, геология, картография, кадастр (geology, cartography, inventory)	economics	mathematics	management, engineering	management, engineering
5. Do you need to increase informatization level in the company/organization?	Yes	Yes	Yes	Yes	Yes
6. Do you have a need for IT skills improving of your employees?	Yes	Yes	Yes	Yes	Yes
7. Can you estimate a number of employees who need IT skills improving?	Yes, 30	Yes, 5	Yes, 9	Yes, 9	Yes, 10
8. Do you agree to support employees in IT skills improving during additional IT training?	tuition fee, scholarship, reduced working time, vacation during examinations	reduced working time	tuition fee, scholarship, reduced working time, vacation during examinations	tuition fee, scholarship, reduced working time, vacation during examinations, other	tuition fee, scholarship, reduced working time, vacation during examinations
9. Cooperation forms you are willing to perform with university:	Employing graduated students, Internships	Employing graduated students, Internships, Developing of new curriculums	Employing graduated students, Professional seminars made by company employees	Internships	Employing graduated students, Internships

В целом местные предприятия проявляют большой интерес к специалистам в различных областях, таких как экономика, психология, журналистика, машиностроение и другие, с дополнительными компетенциями в области информационных технологий. Что касается уровня занятости выпускников с таким профилем, у них нет трудностей найти работу сразу после учебы.

Были получены официальные письма заявки на специалистов магистратуры ISC в количестве 9 кадров на 2017 год и на основе этого был a letter was sent from the rector of NUU to the Minister HSSE of Uzbekistan on the opening of Master's Degree course 5A330204 – Information systems (second competence for professionals not specialized areas), а также были подготовлены государственный стандарт требований к специальности, учебный план, необходимых для открытия магистерской специальности.

С целью повышения квалификации преподавателей по учебным предметам входящих в учебный план магистратуры 10 преподавателей участвовали семинарах, организованных Европейскими университетами, с целью формирования и согласованию соответствующих учебных программ.

В 2016-2017 учебном году было КабМином РУз выделена квота на прием в магистратуру в количестве 6 мест. Было выпущены флаеры и буклеты для абитуриентов. Было подана 21 заявок на поступления в магистратуру. По итогам экзаменов поступила в магистратуру 6 студентов, выпускники из направления «Прикладная математика». Из них на данный момент на 2-курсе магистратуры ISC учится 4 студента, 1 отказался учиться, 1 был отчислен в 1 курсе за неуспеваемость. В 2016-2017 учебном году на 6 мест, в результате экзаменов поступила 7 студентов, где имеются выпускники бакалавров по математике, прикладные математика, педагогики преподавания математики и информатика. Два студента магистратуры Тулеев У. и Рахмонов Ж. участвовали в учебных семинарах в университете Гренобля по программе мобильности студентов. Были приобретены компьютерные оборудование и другие технический средства для организации учебного процесса на сумму 15000 евро.

В НУУз были организованы семинары, ведущих профессоров из Европейских университетов, участников проекта.

В 2016 г. в НУУз были организованы семинарские занятия профессором Агаты Мерсерон и Петра Сауре из Берлинского технологического университета, в объеме 20 часов по двум темам (14-20 сентября), профессора Даниела Бордо из университетса Гренобль Альп, в объеме 16 часов (7-11 ноября) и профессора Арто Топенана в объеме 10 часов (28.02.2017. -03.03.2017). Подготовлена и опубликована учебная пособия по предмету "Объектно-ориентированное программирования" (1 -квартал 2017г.).

В 2016 г. в НУУз были организованы семинарские занятия профессором Агаты Мерсерон и Петра Сауре из Берлинского технологического университета, в объеме 20 часов по двум темам (14-20 сентября), профессора Даниела Бордо из университетса Гренобль Альп, в объеме 16 часов (7-11 ноября) и профессора Арто Топенана в объеме 10 часов (28.02.2017. -03.03.2017). Подготовлена и опубликована учебная пособия по предмету "Объектно-ориентированное программирования" (1 -квартал 2017г.).





Деятельность команды НУУз по выполнению проекта освещалась на сайте проекта.

20.06.15 [Вторая специальность Информационные системы для непрофильных отраслей](#)

22.05.15 [Seminar in Tashkent \(19-20.may\)](#)

Moodle UPMF

Partners of the project



reform and modernization of higher education systems in the partner countries of Eastern Europe, Central Asia, the Western Balkans and the Mediterranean region.

Tempus PROMIS project is part of UPMF French projects were selected in the process of selection deemed severe by the European Commission (acceptance rate: about 15% of the projects submitted). It is also one of the most important in terms of funding and partner network Tempus projects. Worn by Jean-Michel Adam (UPF Humanities and Society, SHS) in coordination with service international relations UPMF the PROMIS project is spread over three years and has a budget of 1.2 million euros. Its launch date is scheduled for December 2013. The project is not entirely new, it follows the Tempus project ERAMIS. The latter, also led by Jean-Michel Adam between 2010 and 2013, has created a network of Masters in Computer Second Jurisdiction (ISC) in 10 universities of Russia, Kazakhstan and Kyrgyzstan. The main objectives of this new PROMIS project: "Masters extend this network to other countries in Central Asia (Kyrgyzstan, Kazakhstan, Uzbekistan, Turkmenistan and Tajikistan) and enhance the character of professionalizing Masters". It will also develop an international version of the specialty "Double Skill: Computer and Social Sciences (DCISS)" master of UPMF "Engineering, Cognition, Creation and learning" (IC2A). The network of European partners PROMIS has grown: it combines, in addition UPMF, Lublin University of Technology (Lublin, Poland), Beuth University of Applied Sciences (Berlin, Germany), Savonia University of Applied Sciences (Kuopio, Finland) and Kaunas University of Technology (Kaunas, Lithuania), as well as businesses Eledia (Berlin), Symetrix (Grenoble) and Ilimi Solutions (Kuopio). Ten universities are beneficiaries of the project. Engagement in the Tempus program is a longstanding tradition in our university. Since 1990, UPMF participated in 32 Tempus projects. She flew 19 as project manager. The Tempus program is currently in its fourth and final phase. It is likely that international cooperation projects within the scope of the Tempus program are integrated, in 2014, the "Erasmus +" program.

Распространение и передача результатов
Регулярное снабжение web сайта информацией
о результатах и осуществленной деятельности.
Функционирует отдельная страница Проект TEMPUS-
PROMIS в сайте НУУз (promis.nuu.uz)



Заключение

Система высшего образования должны оперативно удовлетворять растущий спрос на специалистов, работающих на стыке нескольких специальностей. Начинать нужно с разработки новых учебных программ. Содержание учебных программ должно учитывать особенности применения ИКТ в разных предметных областях. Следует исходить из того, что работающие в этих областях кадры изначально не готовились как специалисты по направлениям ИКТ. Дополнительная компетенция, полученная ими при обучении, позволит повысить производительность труда и эффективность принятия управленческих решений.

Подготовка специальность магистратуры "Информатика как вторая компетенция" в РУз является востребованным и выполнения проекта дает возможность внедрять новшества в учебные процесс с целью координации с зарубежными ведущими европейскими вузами и считаем целесообразным продолжении и развитие сотрудничества вузов РУз в ERASMUS проектах.

**РАЗРАБОТКА ЭЛЕКТРОННЫХ УЧЕБНЫХ ПОСОБИЙ В РАМКАХ ПРОЕКТА
TEMPUS PROMIS**

(Евразийский национальный университет им. Л.Н.Гумилева)

Характерной чертой современного этапа развития общества, является его информатизация - объективный процесс, связанный с повышением влияния интеллектуальных видов деятельности на все стороны общественной жизни и ориентированный на использование достоверного, исчерпывающего и своевременного знания во всех видах человеческой деятельности, в том числе и образования.

С 2015 года Евразийский национальный университет им.Л.Н.Гумилева стал участником международного проекта TEMPUS PROMIS (PROfessional network of Master's degrees in Informatics as a Second Competence) – "Сетевая магистерская программа «Информатика—как вторая компетенция».

Цель программы профессиональное обучение универсальных людей, которые будут сочетать навыки своей первой специальности (той, которую они приобрели во время учебы бакалавров) с теоретическими и техническими навыками в области информатики, позволяя им проектировать и управлять программными средствами.

Проект реализуется совместно с:

- 5 университетами из стран Евросоюза (Германия, Литва, Польша, Финляндия, Франция);
- 10 университетами из стран Центральной Азии (Казахстан, Кыргызстан, Узбекистан, Таджикистан, Туркменистан);
- 3 предприятиями Европы.

Была поставлена задача: разработать электронные учебные пособия (ЭУП) по обязательным дисциплинам по проекту TEMPUS PROMIS. ЭУП по дисциплинам «Программирование», «Извлечение знаний и хранилище данных», «Базы данных», «Алгоритмы и структуры данных» и др. можно скачать на сайте <http://promis.enu.kz/>.

В данной статье описано ЭУИ по дисциплине «Программирование».

После установки программной оболочки в главном меню появляется группа «Электронное учебное пособие», в ней ярлык: «Программирование» (рисунок 1).

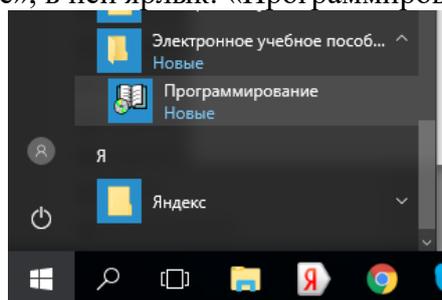


Рисунок 1

Титул представлен на рисунке 2.



Рисунок 2. Титул

Кнопки «Авторы» и «Помощь» отражают соответствующую информацию. При нажатии кнопки «Оглавление» открывается окно наглядно отражающее структуру электронного учебного пособия (рисунок 3).

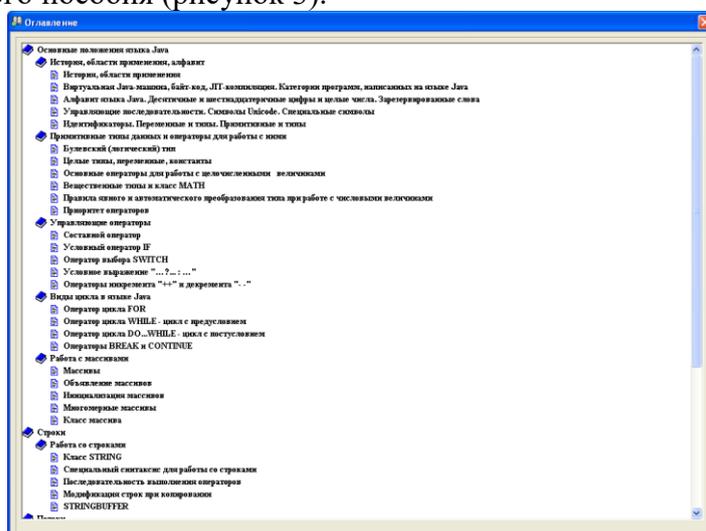


Рисунок 3

Кнопка «Содержание» позволяет обучаемому выбрать режим работы (рисунок 4).



Рисунок 4. Режим работы

Первый режим просмотра. В этом режиме обучающая программа обеспечивает просмотр только учебного материала. При этом доступа к заданиям, вопросам и файлам мультимедиа не будет (рисунок 5).

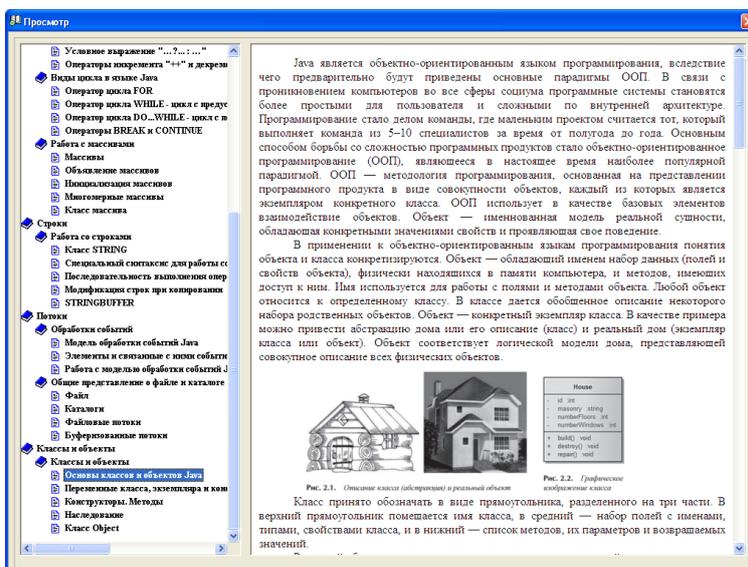


Рисунок 5. Режим просмотра.

Второй режим тестирования. В этом режиме обучающая программа обеспечивает тестирование по всему объему учебного материала. При этом после тестирования можно получить информацию о результате тестирования (рисунок 6).

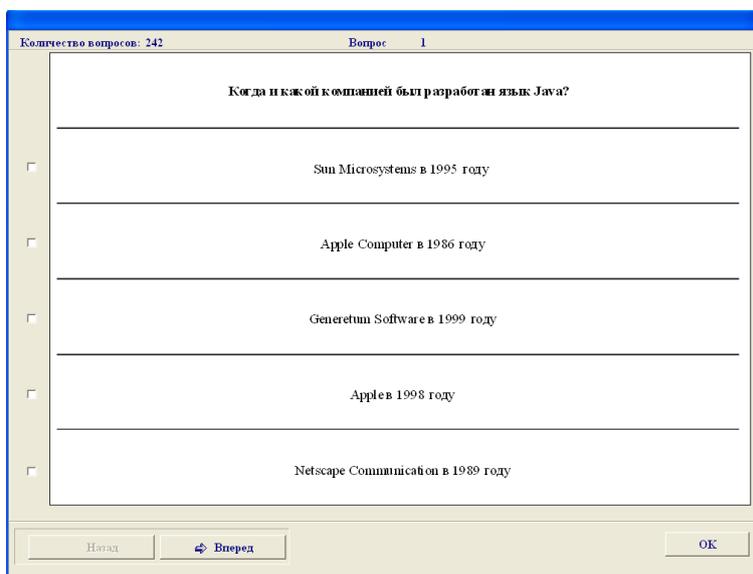


Рисунок 6

Третий режим начала обучения. Для начала обучения необходимо обучаемому зарегистрироваться (рисунок 7).

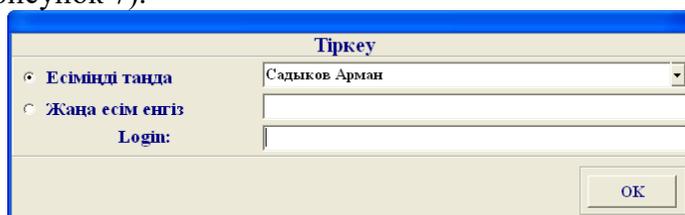


Рисунок 7. Форма регистрации

В этом режиме обучающая программа обеспечивает выбор траектории обучения. При этом после изучения теоретического материала по текущему уроку необходимо будет отвечать на тестовые вопросы. В случае недостаточного количества правильных ответов на тесты, обучаемый не сможет перейти к следующему уроку в траектории и будет продолжать изучение текущего урока. Кроме текущего тестирования предусмотрены промежуточное тестирование (при переходе к следующему блоку), рубежное (при переходе к следующему модулю) и итоговое (при завершении обучения).

Четвертый режим продолжения обучения. В этом режиме обучающая программа обеспечивает продолжение обучения по выбранной траектории. При этом процесс обучения начинается со следующего урока после прерывания.

Режим начала обучения позволяет выбрать одну из трех траекторий обучения: ручной выбор, тестовый выбор и полный выбор (рисунок 8).



Рисунок 8. Режим обучения

При ручном выборе траектория определяется обучаемым самостоятельно путем отметки номеров модулей, блоков, уроков (рисунок 9).

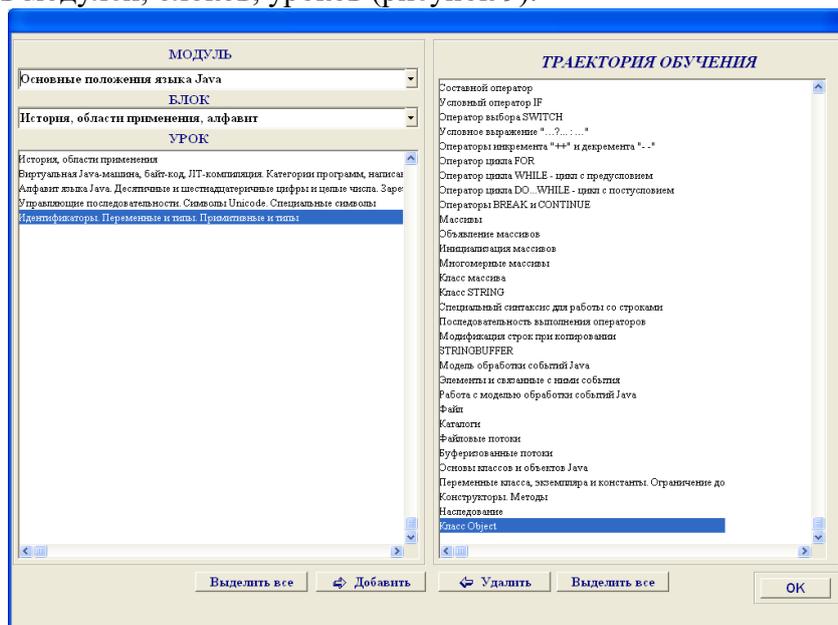


Рисунок 9. Ручной выбор траектории обучения

При тестовом выборе траектория определяется автоматически по результатам тестирования по всему объему учебного материала. В этом случае в траекторию обучения включаются только те уроки, по вопросам которых были получены недостаточное количество правильных ответов. При полном выборе в траекторию включается весь объем учебного материала данной дисциплины, включая все уроки, модули и блоки.

После определения траектории пользователь переходит непосредственно к сеансу обучения (рисунок 10).

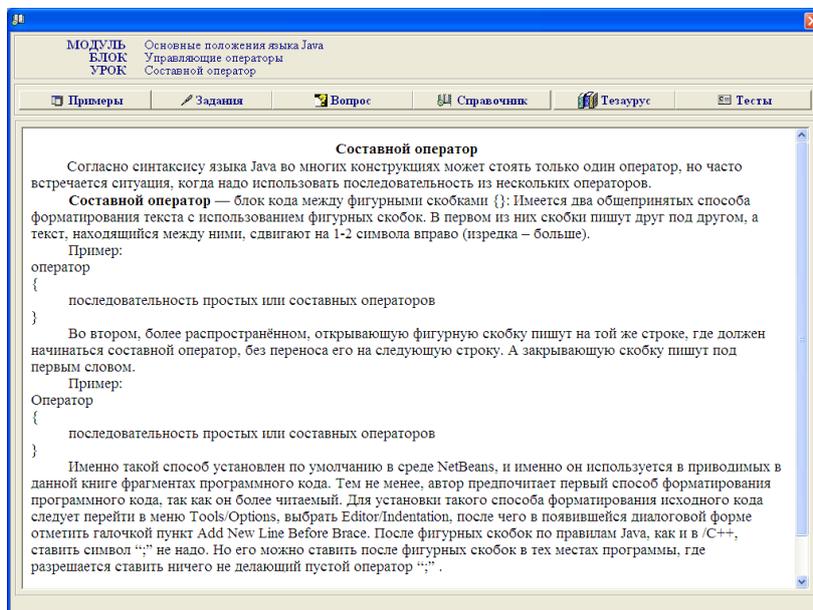


Рисунок 10. Сеанс обучения

В каждом уроке есть теоретический материал, задания для самостоятельной работы обучающегося, вопросы, тезаурус, справочник, тесты. В уроках, где предполагаются примеры и задания, соответствующие кнопки активны.

Элемент обучения «Тесты» обеспечивает доступ к тестированию, которое предназначено для осуществления самоконтроля знаний по текущей единице обучения:

- на уровне уроков – текущий контроль знаний;
- на уровне блоков – промежуточный контроль знаний;
- на уровне модулей – рубежный контроль знаний;
- на уровне всего ЭУП– итоговый контроль знаний.

Кнопка «Тесты» позволяет обучающемуся пройти текущий контроль знаний. Для перехода к следующему уроку необходимо ответить правильно на более чем 75% вопросов.

Содержание учебного пособия «Программирование» составлено в соответствии с Силабусом одноименной обязательной дисциплины для магистерской образовательной программы, реализуемой в рамках PROMIS (PROfessional network of Master's degree in Informatics as a Second Competence - Сетевая магистерская программа "Информатика как вторая компетенция).

ЭУП предназначено для изучения основ программирования на Java магистрантами, обучающимися по сетевой образовательной магистерской программы "Информатика - дополнительная специальность" по специальности 6М060200-Информатика.ЭУП будет полезно также студентам, магистрантам других специальностей, изучающих язык программирования Java самостоятельно.

ЭУП издано за счет средств Международного проекта «PROMIS project № 544319-TEMPUS-1-2013-1-FR-TEMPUS-JPCR».

**КУРС ПЕРЕПОДГОТОВКИ СПЕЦИАЛИСТОВ ПО НАПРАВЛЕНИЮ
«ПРОГРАММИРОВАНИЕ КАК ВТОРАЯ КОМПЕТЕНЦИЯ»**

. (Бухарский Госуниверситет, Узбекистан)

Основная проблема открытия магистратуры «Программирование как вторая компетенция» была связана с научным характером магистратуры в Узбекистане. Магистратура в Узбекистане – первичная стадия научной работы. Учебный план составлен тоже с учетом того, что основной период времени выделяется для научной деятельности. Цель же проекта ПРОМИС является переподготовка специалиста, у которого базовая образования является не информационная технология. Обсужденный на семинарах и тренингах проекта ПРОМИС в последние два года и в конце согласованный учебный план, распределение часов, учебные программы и учебные материалы по предметам сосредоточены для подготовки не научного сотрудника, а специалиста, который может участвовать в небольших проектах по разработке программной системы, протекающих в рамках компании как аналитик, проектировщик, разработчик, специалист по развертыванию, тестировщик, и т.д. или как специалист со стороны компании заказчика, который участвует при разработке требований к системе, в приемке, сопровождение и т.д. для больших проектов. Эти разногласия еще длится долго, но можно оказывается решить эту проблему сейчас.

Преподавательский состав Бухарского госуниверситета в рамках проекта ПРОМИС после неоднократных неуспешных попыток открыть магистратуру «Программирование как вторая компетенция» решил открыть курс переподготовки, специальный курс, после окончания которого выдается сертификат об окончании, а не диплом.

Есть конечно некоторые трудности в этом направлении. При поступлении в магистратуру состоится строгий конкурс, желающие много и выбираем лучшие из лучших. На курсы переподготовки желают поступать только хорошо осведомленные об этом курсе и не за магистрского диплома, а только за знания. Поэтому рекламу надо вести заблаговременно, в нужную аудиторию и в разных видах. Например, у нас по республике идет массовое обучение по ИТ. Почти все госучреждения нашего региона обучают своих сотрудников. Занятия проходят в университете, и ведется преподавателями нашей кафедры. Это очень удобный случай для рекламы, которого я думаю мы еще не до конца использовали. В год несколько раз состоятся ярмарки по трудоустройству, где место проведения ярмарки обычно выбирается некоторое высшее учебное заведение. Здесь можно не только раздавать рекламные листки, также можно анализировать потребность компаний к конкретным специальностям. Можно еще организовать круглые столы для обсуждения проблем внедрения новых ИТ в компаниях, где и можно обсуждать и кадровые вопросы, потребность к кадрам и т.д.

Несмотря на все наши усилия, на курс переподготовки зачислили только четыре слушателя. Мы думаем, что на следующий курс число слушателей будут больше.

Следующая проблема – заинтересованность компаний. Сегодня потребность компаний к ИТ специалисту большой, но к специалисту. Компания не хочет тратить дополнительные ресурсы (время, деньги) для дополнительного обучения. Этого и не терпит потребительский рынок. Нужен готовый специалист, а не «полуфабрикат». Конечная цель подготовки специалистов в области ИТ в наших ВУЗах, не полностью соответствует требованиям компаний, она не проблемноориентированная. При первой встрече с конкретной проблемой выпускник не всегда ее понимает до конца, с необходимым охватом и глубиной, иногда не может взять ответственность, бывает не сможет уложиться в срок или решает задачу не очень качественно. Подход к решению более кустарный (метод проб и

ошибок) чем инженерный. Риски, возникающие в процессе работы не всегда заблаговременно прогнозируются и управляются.

Быстрое развитие ИТ ставит перед ВУЗами свои условия, чтобы продукт (выпускник) соответствовал данному уровню. Учебные планы должны соответствовать реальной жизни. Они должны быть гибкими и должны существовать инструменты для их инкрементальной корректировки в зависимости от региона, конкретных потребителей выпускников.

ВУЗы разделены на опорные и типовые. При составлении учебного плана обычно не участвуют региональные компании, иногда даже региональные типовые ВУЗы, Государственный образовательный стандарт по специальностям, учебные планы, образцовые учебные программы разрабатываются в большинстве случаев опорными ВУЗами. В этом направлении начались некоторые движения в прошлом году, приглашали представителей региональных типовых ВУЗов для обсуждения и составления учебных планов. Всячески поощряется связь с развитыми университетами мира, Ташкентский Национальный университет и Технический университет получили некоторую самостоятельность в этих вопросах, и все это воодушевляет, но по всей видимости до приглашения компаний в обсуждение этих вопросов еще есть время.

Желание быстрого внедрения новых ИТ на предприятиях возникает из-за быстрого развития ИТ, повышение клиентоориентированности, создания или модернизации производственной базы, повышения эффективности в управлении и т.д. Быстрое внедрение новых ИТ диктует свои требования к специалистам по ИТ[1]. Представление проблемы, ее сложность, типизация, определение спецификации создаваемой системы, управление требованиями к системе, определение требуемых ресурсов, нахождения путей решений, представление архитектуры создаваемой системы, проектирование, программирование, управление тестированием, методы приемки системы, развертывание, сопровождение, эволюция системы, утилизация – это неполный перечень знаний и умений любого ИТ специалиста. К сожалению, более половины этих знаний, умений и навыков не приобретается в ВУЗе. Это и есть основная причина и результат не тесного сотрудничества компаний с ВУЗами.

Так как курс переподготовки не основывается на Госстандарт, мы были свободны при планировании учебного процесса. Годовой учебный план был предложен Жан-Мишель Адамом (ADAM Jean-Michel) и одобрен нашими коллегами на конференциях в Бишкеке и Оше. Мы были свободны от необходимости включения дополнительных предметов гуманитарного, политического и экономического характера и выделения часов для научной работы, как это обычно делается для магистратуры. Учебные материалы, которые были согласованы нашими коллегами на семинар-тренингах были немного переделаны с учетом сокращения часов.

Отношение к учебе у слушателей качественно отличается от студентов. Целю является не диплом, а знание, вопросы задаются по существу, хотят докопаться до той глубины, где все понятия распределяются «по полочкам». Требуют дополнительную информацию, если что-то не выясняется на достаточном уровне. Достаточно активно участвуют на занятиях.

Наши преподаватели имеют достаточного опыта преподавания соответствующих предметов, но здесь ситуация иная, мы не можем основываться на высшую математику при объяснениях. Хотя учебные материалы изначально подготавливались с этим учетом, но все равно в реальном времени преподавания пришлось «попотеть». Оказывается, можно найти объекты или процессы для объяснения не только из мира математики, но и из реальной жизни, они намного понятны и оказывается более адекватно объясняет материал. Мы создали специальную базу примеров и методов объяснений проблем не на основе математики и используем эту базу и для студентов-математиков.

Использованные источники:

1 Прикладной программист или программист прикладной области.
(http://erasmusplus.uz/images/shared/2015%20_TEMPUS%20IV%20IN%20UZBEKISTAN_3_2.pdf)

ПРОФЕССИОНАЛЬНАЯ МАГИСТРАТУРА ИНФОРМАТИКИ КАК ВТОРОЙ КОМПЕТЕНЦИИ В КАЗНУ ИМ. АЛЬ-ФАРАБИ

(Казахский Национальный Университет имени Аль-Фараби, Алматы)

В Казахском Национальном Университете им. Аль-Фараби с 2013 года стартовал международный проект Project TEMPUS - 544319-TEMPUS-1-2013-1-FR-TEMPUS-JPCR (2013-4530/001/00) "Professional Master's Degree in computer science as a second competence in Central Asia" (PROMIS)

1. Цель проекта

Целями проекта PROMIS Tempus являются:

- 1) повышение уровня профессионализации магистерских программ путем установления прочных связей с местными предприятиями в секторах ИТ и ИКТ;
- 2) развитие участия предприятий в учебном процессе Магистерская программа;
- 3) реализация магистерских программ магистратуры путем адаптации учебного графика к учащимся, работающим на предприятиях;
- 4) использование метода обучения «обратное обучение», в котором акцент делается на развитие студентов ' независимая работа.

2. Задачи проекта

1: Создание новых магистерских направлений Информатики как второй компетенции(ИБК) во всех странах Центральной Азии

1.1 Диагностика нужд в сфере занятости, связанная с внедрением магистерского направления ИБК (ISC), отчет о нуждах и ожиданиях предприятий.

1.2. Создание новых магистерских направлений: документов аккредитации, педагогических команд, силлабусов, запуск магистерских программ.

1.3 Необходимые материалы и программные обеспечения для реализации магистерских направлений и сети

2: Сеть передового опыта магистерских направлений ИБК (ISC)

2.1 Совместная разработка и распространение внутри сети учебных материалов и других педагогических ресурсов на английском, русском и местных языках

2.2 Внедрение технологии обучения «реверс обучение».

2.3 Обобщение дисциплин по выбору дистанционного обучения (e- Learning и видеоконференции)

2.4 Процесс обеспечения качества для оценивания магистерских направлений сети

3: Профессионализация магистерских направлений

3.1 Создание ячейки FARE (Образование путем Чередования и Взаимоотношения с Предприятиями)

3.2 Разработка дисциплин сотрудниками предприятий, интегрированных в педагогические команды.

3.3 Практические стажировки на предприятиях.

3.4 Второй год обучения путем чередования.

3.5 Успешная профессиональная интеграция (трудоустройство студентов)

4: Интернационализация и студенческая мобильность

4.1 Минимум 25 % обучаемых дисциплин должны быть на английском языке.

4.2 Одинаковые общие дисциплины, преподаваемые одновременно на всех магистерских направлениях.

4.3 Студенческая мобильность внутри сети.

5: Закрепление сети

5.1 Рамочные соглашения между заинтересованными партнерами.

5.2 Соглашения о возможности принятия в магистратуры студентов непрерывного обучения.

5.3 Двойные или мульти дипломы между партнерами и соответствующие министерские лицензии

6: Распространение и передача результатов

6.1 Создание Web-сайта проекта.

6.2 Публикация результатов и деятельности проекта (Интернет, СМИ).

6.3 Выпуск информационных материалов (флаеры, буклеты, Web-сайты магистерских направлений).

6.4 Информационные конференции для институтов и для заинтересованной публики.

3. Ожидаемые результаты проекта

1. Выпуск магистров Информатики как второй компетенции

2. Новая образовательная программа магистратуры Информатика как вторая компетенция по рекомендациям ведущих европейских университетов и работодателей.

3. Мобильность преподавателей и магистров.

4. Новые дисциплины, рекомендованные работодателями.

5. Учебные пособия, разработанные по образовательной программе.

4. Полученные результаты проекта (материальные и нематериальные)

1) Выпуск магистров Информатики как второй компетенции.

Выпуск магистров Информатики как второй компетенции был сделан в 2015, 2016, 2017 годах. Выпускники 100% заняты.

2) Разработана и реализована совместно с партнерами по проекту Новая образовательная программа магистратуры Информатика как вторая компетенция по рекомендациям ведущих европейских университетов и работодателей.

Контракты были подписаны с компаниями для организации производственной практики.

3) Мобильность преподавателей и магистров.

Обучение мобильных семинаров для преподавателей в европейских университетах: 10 преподавателей KAZNU

Обучение мобильным семинарам в КазНУ им. Аль-Фараби профессорами из европейских университетов-партнеров (2 профессора из Финляндии, 1 профессор из Франции, 1 профессор из Литвы): 4 семинара (слушатели магистры и молодые преподаватели)

Мобильность мастеров PROMIS путем софинансирования КазНУ: 12 магистров в Люблин, 2 магистра в Берлин, 5 магистров в Куопио.

4) Разработаны новые дисциплины, рекомендованные работодателями: Прикладные направления интеллектуальных информационных систем, Обработка естественных языков, Мультимедиа документы, Бизнес интеллидженс (Datamining)

5) Разработаны более 10 учебных пособия по образовательной программе.

6) Научно-практическая конференция, проведенная совместно с работодателями. Практический семинар-совещание с работодателями был организован совместно с проектом «Медис» - июнь 2016 года.

7) Распространение результатов: опубликовано 5 статей, представленных на 5 конференциях.

8) На базе проекта PROMIS получены два проекта мобильности Erasmus +:

- с Университетом Гренобля (Франция),

- Люблинским Технологическим университетом (Польша),

- Выигран научный грант совместно с Университетом Аликанте (Испания) на 60 миллионов тенге.

- Разработан и подан в Ерасмус+ новый совместный проект «Индустриальные информационные системы» с 5 европейскими вузами и 5 казахстанскими вузами.

9) Новые методы обучения:

- inverseteaching, позволяющее активизировать самостоятельную работу студентов.

5. Влияние проекта на развитие вуза на институциональном, национальном, региональном и международном уровнях.

Преподаватели, участники проекта ПРОМИС, получили новые знания в семинарах университетов-партнеров мобильности, повышение их индивидуальной осведомленности в методах обучения.

Мастера ПРОМИС образовательной программы «Информатика как вторая компетенция» получают современные знания и компетенции по новым дисциплинам, участвующих в данном проекте

Поступают магистры с разным базовым образованием: филологи, экономисты, математики, механики и т.д. Они получают новые компетенции в области информационных технологий. Все студенты, окончившие по данной образовательной программе, начинают работать уже во время учебы в магистратуре. Трудоустройство выпускников данной программы – 100%.

- Расширено международное сотрудничество за счет активного взаимодействия с другими проектами на факультете (проекты Медис, Пиктет)

- Расширено международное сотрудничество: со всеми университетами- партнерами Европы заключены соглашения по сотрудничеству.

- По этим соглашениям ежегодно направлялись студенты-магистранты на двухнедельные мобильности. Ежегодно несколько студентов бакалавриата направлялись на семестровую мобильность(все европейские университеты-партнеры).

- На основании этих соглашений профессора этих университетов привлечены для соруководства докторантами нашего университета (Люблин Технологический Университет, Университет Гренобля).

8. Меры, принятые для обеспечения устойчивости результатов проекта

- Активные контакты с университетами-партнерами позволили выиграть новые гранты:

-научный грант на 60 млн тенге МОН РК на 2015-2017 гг совместно с Университетом Аликанте;

-грант мобильности Erasmus+ с Люблинским Технологическим университетом 2016-2017 гг

- грант мобильности Erasmus+ с Гренобльским университетом 2016-2018 гг

- Идет совершенствование нормативно-правовой базы по сотрудничеству с предприятиями.

- Улучшение синергии происходит за счет активного сотрудничества с другими проектами и в целом с учебным процессом университета.

- Разработан и подан в Ерасмус+ новый совместный проект «Индустриальные информационные системы» с 5 европейскими вузами и 5 казахстанскими вузами.

9. Материалы, подготовленные для распространения информации о проекте и использовании его результатов

Опубликованы 5 научно-методических статей по результатам проекта.

1. Jean-Michel Adam, UalsherTukeyev.

TEMPUS project in KAZNU: from ERAMIS to PROMIS. Материалы международной научно-практической конференции «Применение информационно-коммуникационных технологий в образовании и науке», посвященной 50-летию Департамента информационно-

коммуникационных технологий и 40-летию кафедры «Информационные системы», 22 ноября 2013 года, Алматы, Қазақ университеті, 2013, с.12-16.

2. Уалшер Тукеев, Жан-Мишель Адам.

Профессионально-ориентированный ТЕМПУС проект PROMIS подготовки магистрантов по информатике как второй компетенции. Материалы 44-й научно-методической конференции «Компетентностно-ориентированная система оценки знаний» 17-18 января 2014 года. Книга 1, Алматы, Қазақ университеті, 2014, с.177-178.

3. Jean-Michel Adam, Daniel Bardou.

From ERAMIS to PROMIS: Extending and Improving a Master's Degrees Network in Central Asia. In proceedings of EDULEARN14 6th International Conference on Education and New Learning Technologies. Barcelona, Spain, 7th – 9th of July 2014.

4. Tukeyev U., Rakhimova D. On the experience of the formation of the international educational program of double-diploma training masters of computer science as a second competence in KAZNU. Proceedings of conference – 46th science-methodical conference Al-Farabi Kazakh National University. Almaty. 2016

5. Тукеев У.А., Бельгибаев Б.А., Омаров Б.С. Реализация подготовки магистрантов со специализацией "Автоматизация и управление технологическими процессами" по ГПИИР-2. Proceedings of conference – 46th science-methodical conference Al-Farabi Kazakh National University. Almaty. 2016

На сайте университета постоянно висит информация по данной международной образовательной программе магистратуры «Информатики как второй компетенции», что способствует широкому распространению информации о данном проекте. <http://welcome.kaznu.kz/ru/17126/page>,

http://www.kaznu.kz/ru/17093/page/Departments/Faculty_of_Mechanics_and_Mathematics_/Chairs/Department_of_Information_SystemsDepartment_of_Information_Systems/For_enters_masters_degree

10. Трудности и нереализованный потенциал проекта.

Основной проблемой, стоящей перед в проекте является недостаточный уровень английского языка некоторые из преподавателей и студентов: необходимо улучшить уровень владения английским языком. Недостаточный уровень английского влияния на различных участках реализации проекта негативно, в первую очередь, качество обучения, качество обучения, низкий уровень коммуникаций с партнерами, низкий уровень коммуникаций между студентами .

Литература

1 Jean-Michel Adam, UalsherTukeyev. TEMPUS project in KAZNU: from ERAMIS to PROMIS. Материалы международной научно-практической конференции «Применение информационно-коммуникационных технологий в образовании и науке», посвященной 50-летию Департамента информационно-коммуникационных технологий и 40-летию кафедры «Информационные системы», 22 ноября 2013 года, Алматы, Қазақ университеті, 2013, с.12-16.

2 Tukeyev U., Rakhimova D. On the experience of the formation of the international educational program of double-diploma training masters of computer science as a second competence in KAZNU. Proceedings of conference – 46th science-methodical conference Al-Farabi Kazakh National University. Almaty. 2016, 9-12 pp.

**Р.К.УСКЕНБАЕВА, Р.Ж.САТЫБАЛДИЕВА, А.Н.МОЛДАГУЛОВА,
Г.У.БЕКТЕМЫСОВА**

О ВОСТРЕБОВАННОСТИ СПЕЦИАЛИСТОВ ПО ЗЕЛЕНЫМ ТЕХНОЛОГИЯМ И ИКТ

(Международный университет информационных технологий, Алматы, Казахстан)

«Зеленая экономика» является одним из важных инструментов обеспечения устойчивого развития страны. Переход к «зеленой экономике» позволит Казахстану обеспечить достижение поставленной цели по вхождению в число 30-ти наиболее развитых стран мира.

При этом происходит значительное снижение рисков для окружающей среды. К основным принципам «зеленой» экономики можно отнести следующие:

1. обеспечение равенства – (принцип справедливости);
2. процветание и благополучие для всех (принцип уважения достоинства);
3. сохранение и восстановление природы (принцип предосторожности);
4. инклюзивность и широкое участие в процессе принятия решений (принцип участия);
5. подотчетность (принцип управления);
6. экономическая, социальная и экологическая устойчивость (принцип устойчивости);
7. устойчивое производство и потребление (принцип эффективности);
8. инвестирование в будущее (принцип связи разных поколений).

Одновременно, в настоящее время технологии, позволяющие улучшить коммуникаций, усовершенствовать обмен и обработку информации, затрагивают все стороны нашей жизни. Стоит отметить, что информационно-коммуникационные технологии (ИКТ) стали одной из самых быстрорастущих отраслей экономики в последние годы. Поэтому, можно отметить два аспекта влияния ИКТ. Первое - «озеленение» самой сферы ИКТ за счет снижения удельного энергопотребления, обеспечения более длительного срока службы, возможностей вторичного использования и безопасной утилизации индивидуальных устройств и компонентов сетей, а также перехода на альтернативные источники энергии для обеспечения деятельности центров хранения и обработки данных использование компьютерных технологий и средств связи с максимальным положительным эффектом и с наименьшим экологическим ущербом для окружающей среды.

Второе - дематериализация и миниатюризация потребления, а также виртуализация многих видов деятельности человека. К примеру внедрение электронного документооборота, развитие электронных СМИ и рынка электронных книг и музыки, расширение электронной коммерции и интернет-банкинга, использование электронной почты и видеоконференций для проведения деловых встреч, переход на удаленный режим работы. Дополнительно следует выделить повышение эффективности потребления энергии и ресурсов во всех других секторах за счет создания условий для интеллектуальной трансформации транспортной и энергетической инфраструктуры, оптимизации и автоматизации производственных циклов и строительства «умных» домов и офисов, что ведет к стимулирующему эффекту ИКТ.

По оценкам Климатической группы (The Climate Group), ИКТ способны к 2020 г. снизить мировые выбросы углерода на 15%. Такое снижение выбросов в экономике в целом более чем в пять раз превышает «углеродный след» самого сектора ИКТ. При этом единичные технологии, конечно, важны, но наиболее перспективным направлением являются комплексные технологические платформы, объединяющие различные подходы и инновации.

Наибольшего экологического эффекта от внедрения ИКТ следует ожидать в сфере развития интеллектуальных энергосетей. Основная идея интеллектуальных энергосетей (журнал *The Economist* ввел альтернативный термин – «энергетический интернет») – слияние информационных и энергетических технологий для оптимизации выработки, хранения, передачи, распределения и конечного потребления энергии.

Таким образом, одним из основных условий перехода к «зеленой» экономике является поддержка и распространение инноваций, как технологических, так и экономических и социокультурных. В свою очередь, процесс перехода к «зеленой» экономике приведет к созданию новых рабочих мест, включая высококвалифицированные рабочие места в новых инновационных зеленых направлениях. В этой связи, необходимо образование, которое будет готовить специалистов к решению стоящих перед обществом социальных, экономических и экологических проблем на местном, национальном и на глобальном уровнях с учетом развития ИКТ. При этом знания и умения должны подкрепляться действиями, что приведет к формированию навыков, личного опыта. Необходимо «озеленение» не только экономики и развития, но и подключение образовательного процесса, для развития технического и профессионального образования в области «зеленых» технологий в ИКТ.

Как и в случае с самой «зеленой» экономикой и «зелеными» ИКТ, существует и более узкая трактовка понятия образования в интересах «зеленой» экономики. В англоязычной литературе оно чаще всего рассматривается в связи с изменением структуры занятости и потребностью в новых кадрах или переквалификации старых. В ходе развития «зеленых» отраслей на рынке труда повышается спрос на специалистов новых профессий, а часто возникает и заметный дефицит специалистов конкретных квалификаций. В частности, нехватка кадров отмечается в секторе производства биотоплива в Бразилии, в возобновляемой энергетике и производстве экологических товаров и технологий в Германии, США и Бангладеш, в строительном секторе в Австралии, Китае, Европе и ЮАР.

Кроме специфических квалификаций и компетенций, связанных с занятостью в секторе экологических товаров и услуг, не менее важна задача «озеленения» всего трудового капитала во всех секторах и на всех уровнях. Если отойти от конкретных видов трудовой деятельности, то в целом по содержанию, подходам и методам образование для «зеленой» экономики – по сути то же самое, что и образование для устойчивого развития. Это в первую очередь образование для перемен, поскольку переход к устойчивому развитию и «зеленой» экономике требует отхода от устаревших форм хозяйствования, изменения правил не только формальных (государственная политика, законодательство), но и неформальных (бизнес-этика, социальные нормы и ценности). В новых условиях от образования требуется эффективность в подготовке творческих инициативных личностей, способных решать сложные проблемы инновационными и гибкими способами. А для этого, в первую очередь, требуется переход от репродуктивного к креативному подходу в организации образовательной системы и образовательного процесса, а также в содержании и методах преподавания.

В настоящее время во всем мире наблюдается дефицит инвестиций в образование, несмотря на то, что развитие человеческого капитала признается в качестве одной из ключевых целей и условий успешного развития. На уровне национальных систем озабоченность вызывают процессы коммерциализации и стандартизации (формализации) образования. К примеру, Джон Сили Браун (бывший главный научный сотрудник корпорации Херох и директор его исследовательского центра Palo Alto Research Center (PARC) придерживается идеи «элегантного минимализма» в формальном образовании: обязательный компонент содержания обучения должен обеспечивать только базовые компетенции – грамота, арифметика и критическое мышление. Остальное содержание («открытый» компонент) должно определяться самими учащимися исходя из обилия вариантов и возможностей, которые предлагают (или будут в будущем предлагать) так называемые «образовательные социальные сети».

Так, ИКТ стали жизненно необходимыми инструментами для разработки инновационных решений в вопросах развития. Тем не менее, в целях стимулирования полноценного развития важно, чтобы молодежь понимала, как надо использовать ИКТ и преимущества, которые могут предоставить ИКТ, в своей стране и регионе. Концепция «зеленого» роста предусматривает дальнейшее развитие человечества с учетом необходимости восстановления и бережного использования природных богатств, борьбы с бедностью, отказа от бездумного увеличения потребления, устранения социального и гендерного неравенства. Знание концепции «зеленого» роста и роли ИКТ позволит студентам теоретически освоить и гораздо лучше понять значимость ИКТ в наблюдении и мониторинге окружающей среды, обмене информацией, мобилизации действий, обеспечении экологической устойчивости и смягчении последствий изменения климата.

Обучение и формирование экологической культуры в бизнесе и среди населения: необходимо совершенствовать действующие и разработать новые образовательные программы по рациональному использованию ресурсов и охране окружающей среды в системе образования и подготовки кадров.

Так, в целях обучения и повышения квалификации ИКТ-специалистов на данный момент зарубежные институты проводят курсы в области «зеленых» технологий в ИКТ. К примеру, Athabasca University (Канада) проводит курс «зеленая стратегия в области ИКТ» предназначенный для магистрантов университета. Курс обучает «зеленой» стратегии в сфере ИКТ с профессиональной точки зрения с предоставлением инструкций о том, как лучше всего применять «зеленые» стратегии в области ИКТ в производственной сфере.

Singapore Infocomm Technology Federation (SiTF) (Сингапур) проводит курс «Зеленые ИТ» для специалистов имеющих опыт в сфере ИКТ. Курс затрагивает различные аспекты не только ИКТ-технологий, но и практика стандартов, которые имеют международное качество.

Australian National University (Австралия) проводит курс «ИКТ устойчивость» в рамках чего исследуется, как ИКТ влияют на выбросы углекислого газа и как технологии могут повлиять на сокращение этих выбросов.

Таким образом, подготовка кадров для «зеленой» экономики является одной из главных вопросов, от которой зависит успех «зеленой» экономики.

Список использованных источников:

1 ИКТ - информационные и коммуникационные технологии», Рабочая программа конкурса Европейской Комиссии ,2010г.

2 Уша Рани Вьясулу Редди. Серия учебников по ИКТР для молодежи. Учебник 1: Введение в ИКТ для развития. UN-APCICT/ESCAP 2011

3 Primer Series on ICTD for Youth. Primer 4: An Introduction to ICT, Climate Change and Green Growth, 2013

4 Замлелый А.Ю. Формирование smart (интеллектуальной) экономики: теория и практика // Современные проблемы науки и образования. – 2012. – № 4; URL: <http://www.science-education.ru/104-6684>

5 Максимова В.Ф. Smart есоному (умная, интеллектуальная экономика)– веление времени //МЭСИ

6 Казахстан: Такая «зеленая» экономика URL: <http://kief.com.ua/rus/news/kazakhstan-takaya-zelenaya-ekonomika>

7 Основы перехода к «зеленой» экономике: международный опыт и инструменты / Под ред. проф., д.э.н. Б.К. Есекиной. – Астана, 2013.

8 Стейн Рамсли. Актуальность "зеленых технологий". URL: <http://tssonline.ru/articles2/fix-corp/aktualnost-zelenyh-tehnologiy>

9 Н.А. Сидоров. Зеленые информационные системы и технологии. Национальный авиационный университет, Инженерия программного забезпечення/№3(7) 2011.

ЦИФРОВЫЕ ТЕХНОЛОГИИ В ОБРАЗОВАНИИ

*(Международный Университет Информационных Технологий, Алматы, Казахстан,
Казахский Национальный исследовательский технический университета
им.К.Сатпаева, Алматы, Казахстан)*

Цифровизация быстро проникает в повседневную жизнь людей по всему миру, предоставляя новые возможности для компаний в области разработки новых продуктов и услуг и модернизации методов работы, меняя производственные процессы, каналы дистрибуции и способы взаимодействия с поставщиками и клиентами.

Цифровизация это постепенная миграция многих видов человеческой деятельности в сеть интернет, которая предоставляет массу возможностей не только для создания, хранения и доставки контента, но и для коммуникации и проектной деятельности между людьми.

К 2020 году следует ожидать наступления тотальной цифровизации – момента, когда вся человеческая культура окажется оцифрованной. На передний план выходит необходимость целенаправленной фильтрации информации и, как следствие, усиливается роль образования в формировании соответствующих компетенций. Навык эффективного поиска в сети интернет предполагает способность к рефлексии, способность построить обобщенный образ результата поиска и, исходя из него, сформулировать запрос. Развитие цифровых технологий меняет способы, которыми фиксируется, передается и создается знание, а также формируются навыки. Кроме этого, цифровые технологии меняют процесс оценки и фиксации достижений, процесс формирования собственной образовательной траектории, процессы в управлении учебными заведениями. Цифровые технологии позволяют индивидуализировать образование без увеличения издержек на него.

Благодаря развитию инфраструктуры и удешевлению процессов обработки, хранения и передачи человечество оказалось на пороге самого значительного этапа цифровой революции – мы говорим об офлайн-онлайн-конвергенции и появлении киберфизического мира.

Это стало возможным благодаря нескольким фундаментальным факторам – всеобщей подключенности и быстрому распространению сенсорных устройств и данных.

Степень влияния цифровых технологий в разных отраслях неоднородна. Тем не менее не вызывает сомнений, что все отрасли и игроки в них будут рано или поздно вынуждены пройти через цифровую трансформацию.

Четвертая промышленная революция, о чем президент Назарбаев Н.А. отметил в своем послании народу, первым приоритетом он отметил: «Ускоренная технологическая модернизация экономики. Мы должны культивировать новые индустрии, которые создаются с применением цифровых технологий. Это важная комплексная задача.

Необходимо развивать в стране такие перспективные отрасли, как 3D-принтинг, онлайн-торговля, мобильный банкинг, цифровые сервисы, в том числе в здравоохранении и образовании, и другие.

Также важно обеспечить развитие коммуникаций, повсеместный доступ к оптоволоконной инфраструктуре. Развитие цифровой индустрии обеспечит импульс всем другим отраслям. Поэтому вопрос развития IT-сферы Правительство должно держать на особом контроле.»

Цифровая трансформация предусматривает цифровизацию и интеграцию процессов по вертикали в рамках каждой отрасли, начиная от разработки продуктов и закупок и заканчивая производством и обслуживанием. Все данные об операционных процессах, эффективности процессов, управлении качеством и операционном планировании доступны в режиме реального времени в интегральной сети.

В связи с вышесказанным предлагается выделить ряд направлений первоочередного развития. Это следующие направления:

1. Здравоохранение (телемедицина, превентивная медицина, медицинская генетика)
2. Образование (онлайн образование)
3. Наука (открытая среда для хранения, обмена и использования научных данных)
4. Промышленность (Internet of Things)
5. Сельское хозяйство (цифровые системы мониторинга и контроля (точное земледелие))
6. Ритейл (отслеживание качества изделий и их происхождения)
7. Транспорт (беспилотный транспорт и услуги на его основе)
8. Электроэнергетика (Smart Grid)
9. ЖКХ
10. Рынок финансовых услуг (в том числе, финтех-инновации)

Для достижения высоких результатов в цифровизации экономики необходимо достижение ряда поставленных целей, одной из которых является создание инфраструктуры, обеспечивающей взаимодействие субъектов в цифровом пространстве.

Развитие цифровой инфраструктуры подразумевает:

1. Обеспечение всеобщего и безлимитного доступа в интернет
2. Электронное государство (правительство)
3. Обеспечение хранения, обработки и конвертации баз данных
4. Работа с big data и системы поддержки принятия решений
5. Идентификация пользователей
6. Идентификация технических устройств
7. Обеспечение безопасности и защиты данных
8. Восстановление после злонамеренных действий
9. Доступ к базовым приложениям
10. Алгоритмы и сервисы на основе искусственного интеллекта
11. Сертификация алгоритмов и систем, используемых в цифровой экономике
12. Обеспечение раскрытия информации (раскрытие которой целесообразно) и программного доступа к базам данных органов власти, государственных и частных компаний и других организаций (где применимо) по всем направлениям деятельности
13. Единые технологические стандарты («облака», «большие» данные, биометрия, пространственные данные и пр.)

Цифровизация в образовании

Повсеместная цифровизация не могла не коснуться сферы образования. Скорее она идет совместно с ней, а в идеале, должна идти на опережение. Сфера образования должна развиваться гораздо активнее, чтобы успевать за быстро меняющимся миром.

В Послании народу, темой которого стала «Третья модернизация Казахстана», Президент Назарбаев Н.А. ставит самый высокий приоритет на инвестициях в образование, он всегда подчеркивает важность человеческого капитала, который стал самым важным фактором, благодаря которому последние 25 лет были успешными. «Прежде всего должна измениться роль системы образования. Наша задача – сделать образование центральным звеном новой модели экономического роста. Учебные программы необходимо нацелить на развитие способностей критического мышления и навыков самостоятельного поиска информации. Наряду с этим нужно уделить большое внимание формированию IT-знаний.»

Традиционная образовательная система уже не способна обеспечить выпускникам долговременную гарантию занятости, поскольку стремительные темпы обновления знаний, объем которых удваивается в среднем каждые полтора года, требуют постоянной переподготовки.

Современный этап развития образования в мире характеризуется особой интенсивностью преобразований, в равной мере затрагивающих организационные и

управленческие структуры образования, его целевые установки и содержание, методы и технологии обучения, источники и механизмы финансирования, а также условия и формы образовательного, научного и производственного сотрудничества. То есть развитие человеческого капитала является решающим фактором в реализации поставленных задач по пути построения цифровой экономики

В электронном обществе пересматривается само представление об обучении, связях обучения с работой и повседневной жизнью. Поскольку информационное общество основывается на умственном труде, работа все теснее переплетается с учебной, которая превращается в пожизненное занятие. Сам процесс обучения, однако, совсем не требует обязательного посещения школы или университета: интернет позволяет учиться дистанционно у самых лучших специалистов в своих областях. Это при том, что объем знаний можно увеличить на 30%, затратив на 30% меньше средств и на 40% меньше времени, чем при традиционной форме. Именно поэтому система просвещения, интегрированная с информационной магистралью, близка к тому, чтобы уже в ближайшем будущем может стать основой образовательной системы.

Исходя из выше изложенного, задачи Вуза на современном этапе - обеспечивать актуальность образования, регулярно менять программы в связи с изменениями конъюнктуры рынка, а именно: ориентация на глобальный рынок; образовательные и научные процессы должны опережать текущее развитие; деятельность должна быть ориентирована на практическое использование знаний и научных исследований в целях развития региона.

Новое время диктует новую систему подачи знаний. Образовательный процесс должен быть более технологичным, нужно отходить от чисто вербально-репродуктивного подхода в обучении, получение информации происходит из разнообразных источников. Обучение должно быть все более технологичным. Студенту нужно больше самостоятельности в добывании знаний. Роль преподавателя больше становится координирующей, направляющей. Современный студент и преподаватель значительно отличаются в навыках получения и усвоения материала.

Требуется проработка вопросов внедрения в образовательный процесс обучения курсов практического использования новейших программных решений и приложений для конкретной профессиональной деятельности по специализациям. Инновационные технологии в образовании позволяют получать ценные кадры.

Таким образом, в результате эффективного применения современных передовых образовательных технологий в практике преподавания отечественных образовательных учреждений в долгосрочной перспективе можно ожидать повышение качества образовательных услуг, увеличение количества высококвалифицированных кадров в экономике страны.

Для ускоренного развития и модернизации казахстанского образования, в свете последних разработок по цифровизации и внедрению новых технологий предлагаем проведение следующих мероприятий:

- увеличить критическую массу обучающихся на специальностях ИКТ;
- пересмотреть контентное содержание специальностей, включить дисциплины для освоения современных областей науки и технологий, формате виртуальных игр и дополненной реальности можно изучать любую образовательную дисциплину.
- реализация образовательных программ и инновационных проектов с использованием технологий сетевого обучения. инновационные практикумы и сетевое обучение, мастер-классы, интенсивные мини-курсы.
- государственно-частное финансирование образовательных программ и инновационных проектов. Проекты по линии правительственных и общественных организаций. Проекты по линии министерств и госкомпаний. Проекты по линии частных промышленных предприятий.
- провести системный анализ подготовки PhD и магистратуры, с целью приближения их потребностям Цифровой экономики.

- разработка мероприятий и программ модернизации экономики через ИКТ.
- разработать Аналитическое Содержание Киберщита.
- мотивировать отечественного разработчика. использовать не только зарубежные платформы для создания отечественных продуктов, а создавать свои отечественные платформы. Тем самым увеличить отечественные разработки – разработка платформ Big Data, платформы IoT.
- сделать трендом образования online-обучение. персонализация обучения. обучение в любое удобное и комфортное время, non-stop. обучение через мобильный телефон.
- ввести в образование обязательное полиязычие.
- для эффективного развития стартап движения создать университетские консорциумы по стартап движениям.
- университетам с международной аккредитацией дать право ввести новые специальности в области ИКТ с грантовым обеспечением от государства.

Заключение. Для системного внедрения новшеств в образовательных учреждениях всех уровней необходимы инициативные менеджеры, профессиональные IT-специалисты и надежная инфраструктура. Необходима разработка новых образовательных программ, проведение совместных конференций, мастер-классов, тренингов и других форм эффективных коммуникаций для активного внедрения образовательных стандартов XXI века.

Литература.

1. Послание Президента Республики Казахстан Н.Назарбаева народу Казахстана. 31 января 2017 г. «Третья модернизация Казахстана: глобальная конкурентоспособность».
2. Макроэкономические перспективы Казахстана и новые инвестиционные горизонты в цифровизации. Новое исследование Фонда национального благосостояния "Самрук-Казына" и The Boston Consulting Group., декабрь, 2016г.

УДК 004

ШАРИПБАЙ А., ОМАРБЕКОВА А., БЕКМАНОВА Г., НИЯЗОВА Р., ТУРЕБАЕВА Р., ХАСЕНОВ Е., БАРЛЫБАЕВ А.

ПРОФЕССИОНАЛЬНАЯ МАГИСТРАТУРА ИНФОРМАТИКИ КАК ВТОРОЙ КОМПЕТЕНЦИИ В ЕНУ ИМЕНИ Л.Н. ГУМИЛЕВА

(Евразийский национальный университет имени Л.Н. Гумилева, Астана, Казахстан)

В Евразийском национальном университете имени Л.Н. Гумилева (далее - ЕНУ) с 2015 года стартовал международный проект Project TEMPUS - 544319-TEMPUS-1-2013-1-FR-TEMPUS-JPCR (2013-4530/001/00) “Professional Master’s Degree in computer science as a second competence in Central Asia” (PROMIS)

В целях укрепления связей между университетами и предприятиями, местные предприятия были вовлечены в организацию магистратуры с начала проекта, по примеру европейских предприятий-партнёров проекта (участвовали в организации учебного процесса, работники предприятий были задействованы в преподавании, в организации исследовательской практики).

1 Разработана и реализована совместно с партнерами по проекту новая образовательная программа магистратуры «Информатика как вторая компетенция» по рекомендациям ведущих европейских университетов и работодателей.

2 Первый выпуск магистров Информатики как второй компетенции были выпущены в 2017 году, количество выпускников 4 и все они трудоустроены.

Контракты были подписаны с компаниями для организации исследовательской практики (АО «Национальный инфокоммуникационный холдинг «Зерде», АО «Национальные информационные технологии», АО «Казахтелеком», АО «КазИнжиниринг», АО «Казахстан Темір жолы», АО «КЕГОГ» и др).

Обучение мобильных семинаров для преподавателей ЕНУ им.Л.Н.Гумилева в европейских университетах прошли 8 преподавателей: Шарипбай А., Омарбекова А., Бекманова Г., Ниязова Р., Туребаева Р., Хасенов Е., Барлыбаев А., Казиев Г.

Обучение мобильным семинарам в ЕНУ им.Л.Н. Гумилева профессорами из европейских университетов-партнеров (2 профессора из Финляндии, 2 профессора из Германии, 1 профессор из Литвы) провели 4 семинара, слушателями которых были магистранты и молодые преподаватели.

Мобильность магистрантов PROMIS путем софинансирования ЕНУ: 3 магистранта в Берлине.

3 Разработаны новые дисциплины, рекомендованные работодателями: «Сентимент анализ в социальных сетях», «Язык веб онтологий (OWL)», «Структурный анализ в социальных сетях».

4 Разработаны 7 учебных пособий и 8 электронных учебных изданий по образовательной программе.

Таблица 1. Наименование учебных пособий и электронных учебных изданий

п/н	Учебные пособия на русском языке	Учебные пособия на английском языке	Электронные учебные издания на русском языке	Электронные учебные издания на английском языке
1.	Алгоритмы и структуры данных	Algorithms and data structures	Алгоритмы и структуры данных	Algorithms and data structures
2.	Математика для компьютерных наук	Mathematics for computer science	Математика для компьютерных наук	Mathematics for computer science
3.	Программирование	Programming	Программирование	Programming
4.	Программная инженерия		Программная инженерия	
5.			Извлечение знаний и ХД	

5 Практический семинар-совещание с работодателями был организован совместно с проектом «Интеллектуальная система обучения, контроля и оценки знаний» в 2016 году. Участники проекта участвовали с докладом на конференции в Анталии и в Барселоне.

6 В магистратуру поступают с разным базовым образованием. Они получают новые компетенции в области информатики. Все магистранты, окончившие по данной образовательной программе, начинают работать уже во время учебы в магистратуре. Трудоустройство выпускников данной программы – 100%.

Расширено международное сотрудничество за счет активного взаимодействия с другими проектами кафедры (проекты ИСОКОЗ);

Расширено международное сотрудничество: со всеми университетами- партнерами Европы заключены соглашения по сотрудничеству;

По этим соглашениям ежегодно направлялись магистранты на двухнедельные мобильности.

7 Опубликовано 7 научно-методических статей по результатам проекта:

– Sharipbay A.A., Niyazova R. «The organization and conduct of the teaching process by educational program of Informatics as a second competence within project PROMIS in the ENU named L.N. Gumilyov». Collection of materials «The experience of cooperation between universities and industry» within the project Tempus-Promis. October 15, 2015 KSTU named after Razzakov I.

– Абжанова А., Ниязова Р.С. Разработка интеллектуального электронного учебника по дисциплине «Программная инженерия». Труды IV международной научно – практической конференции: Информатизация общества. Астана: 2016. – С. 245-247.

– Шарипбай А.А., Омарбекова А.С., Бекманова Г.Т., Ниязова Р.С. Разработка базы знаний по дисциплинам проекта Tempus Promis. Тезисы докладов. Труды III Международной научно-практической конференции «Интеллектуальные информационные и коммуникационные технологии» - средство осуществления третьей индустриальной революции в свете стратегии «Казахстан-2050» -Астана, 2016.- С.181-185.

– Абжанова А. Онтологическая модель интеллектуального электронного учебника по дисциплине «Программная инженерия». XII Международная научная конференция студентов и молодых ученых «Наука и образование - 2017» г. Астана 14 апреля 2017 г.

– Omarbekova A., Zakirova A., Abduraimova B., Duzelbayev S. Automated generation of test questions from knowledge-base. 3rd International conference on “Multidisciplinary innovation in business engineering science & technology” (MI-BEST April 01-04, 2017). Bangkok, Thailand.

– Шарипбай А.А., Бекманова Г.Т., Омарбекова А.С., Разахова Б.Ш., Туребаева Р.Д., Хасенов Е.А. Онтологическая модель образовательной программы магистратуры информатика – вторая компетенция в рамках проекта Темпус Промис. Информатизация общества: V межд. научно-практ. конф. - Астана, 2016.-С. 269-272

– Sharipbay A., Razakhova B., Bekmanova G., Omarbekova A., Khassenov Ye., and Turebayeva R. Ontological Model of Educational Programs in Computer Science (Bachelor and Master Degrees). International Journal of Computing, Communication and Instrumentation Engineering. Vol. 2, Issue 2 (2015) ISSN 2349-1469 EISSN 2349-1477. http://ijccie.ieng.org/journals_pdf/22.

На сайте университета размещена информация по данной международной образовательной программе магистратуры «Информатики как второй компетенции», что способствует широкому распространению информации о данном проекте (<http://promis.enu.kz/>)

Литература

1 Jean-Michel Adam, Ualsher Tukeyev. TEMPUS project in KAZNU: from ERAMIS to PROMIS. Материалы международной научно-практической конференции «Применение информационно-коммуникационных технологий в образовании и науке», посвященной 50-летию Департамента информационно-коммуникационных технологий и 40-летию кафедры «Информационные системы», 22 ноября 2013 года, Алматы, Казак университеті, 2013, с.12-16.

**ШАРИПБАЙ А.А., БАРЛЫБАЕВ А.Б., ОМАРБЕКОВА А.С., БЕКМАНОВА Г.Т.,
РАЗАХОВА Б.Ш., НИЯЗОВА Р.С.**

**ПРОЕКТ ГОСУДАРСТВЕННОГО СТАНДАРТА ПО ИНТЕЛЛЕКТУАЛЬНЫМ
ЭЛЕКТРОННЫМ УЧЕБНЫМ ИЗДАНИЯМ (СМАРТ-ТЮТОРАМ)**

*(НИИ «Искусственный интеллект» Евразийский национальный университет им.
Л.Н.Гумилева, Астана, Казахстан)*

Достижения в области технологий и рост электронного обучения в целях обеспечения преподавателей и преподавателей с уникальными возможностями для расширения обучения и преподавания в области корпоративного управления, государственного управления, здравоохранения и высшего образования приводят к вопросу пересмотра государственного стандарта по электронному учебному изданию. Для начала мы соберем и проанализируем библиографические материалы в области электронных учебных изданиях.

SCORM. В 1997 году департамент обороны США установил инициативу по созданию ADL (Расширенное распределенное обучение) по стандартизации и модернизации обучения. В 2001 году была выпущена первая производственная версия нового стандарта, которая была известна как Sharable Content Object Reference Model (SCORM) 1.1. Позднее в том же году, это было оно заменено SCORM 1.2, который является, пожалуй, самым распространенным стандартом электронного обучения, используемым сегодня. Значительная часть стандарта SCORM была взята из фундамента, заложенной AICC, и обновлена до использования фреймворка JavaScript. Это потребовало, чтобы LMS обернула курс во фрейм и использовала JavaScript API для связи между курсом и LMS. Этот стандарт с тех пор превратился в более новую версию, известную как SCORM 2004, а его последний выпуск известен как SCORM 2004 4th издание, выпущенный в 2009 году. В SCORM 2004 добавлена концепция последовательности объектов курса, а также способность LMS управлять навигацией контента. Хотя SCORM был более современным, чем AICC, он был разработан для работы в среде настольных компьютеров и ноутбуков, требуя от пользователя входа в сеть и входа в LMS для записи информации. В сегодняшнем мобильном мире, прежде всего, мире BYOD, это не всегда так. SCORM также очень ограничен в информации, которую он может записывать из курса, его язык поддерживает только оценки, завершение/прохождение/неудачу и ответы на вопросы. В 2010 году ADL начала исследования нового стандарта eLearning, который позволил бы намного больше гибкости, и ему было дано кодовое название Tin Can API.

xAPI. В апреле 2013 года была официально выпущена спецификация Tin Can API 1.0, а название проекта было изменено на «Experience API» или xAPI. Имя Tin Can API, или просто Tin Can, по-прежнему является широко используемым именем для xAPI. xAPI – это простой, гибкий стандарт, который позволяет собирать данные по широкому спектру информации, а не только строго информацию о курсе. Он также позволяет отправлять данные в полностью независимом от платформы виде, так что операторы xAPI могут быть отправлены из чего угодно - от действий на курсе, запущенном в веб-браузере, до физической кнопки, нажатой на устройстве. Операторы xAPI формулируются как предлог объекта-глагола, подобно «я сделал это», и допускают большую гибкость при составлении оператора, чтобы можно было передавать практически любую информацию. xAPI действительно просто определяет язык для передачи данных, и как эти данные должны быть сохранены. В стандарте электронного обучения, который заменит SCORM, необходимо будет не только определить это, но также указать, какие данные необходимо передать для того, чтобы курс был пройден, оценен и завершен. Сейчас разрабатывается новое поколение стандартов электронного обучения, smi5.

smi5. До того, как AICC была распущена, началась работа по созданию нового стандарта электронного обучения, который должен был заменить SCORM. Этот стандарт известен как smi5 и фактически является «профилем» xAPI, что означает, что он представляет собой стандартный набор операторов xAPI, которые должны использовать курс для связи smi5 с совместимым LMS. Работа над этим стандартом была передана ADL, и сегодня она активно развивается. В дополнение к стандартным наборам операторов, которые должны отправлять курс, smi5 также позволяет LMS захватывать все операторы на основе xAPI, отправленные из курса, которые необходимо сохранить и сообщить, с предоставлением стандартной структуры, которую SCORM всегда предоставлял, и гибкости, которую обеспечивает xAPI. Сегодня интересное время для стандартов eLearning, поскольку достижения в xAPI и smi5 позволят менеджерам по обучению все больше и больше узнавать о ходе обучения и продолжать совершенствовать свои учебные программы.

Open eBook (или ОЕВ), или формально, Open eBook Publication Structure (OEBPS), является устаревшим форматом электронных книг, который был заменен форматом EPUB. Он был «основан главным образом на технологии, разработанной SoftBook Press» и на XML. ОЕВ был выпущен с бесплатной версией, общего пользования, и полной версией, которая будет использоваться с или без DRM в издательской промышленности. Открытый eBook является файлом ZIP плюс файл манифеста. Внутри пакета может использоваться определенное подмножество XHTML наряду с CSS и метаданными Dublin Core. По умолчанию расширение файла .opf (формат пакета ОЕВ).

В Казахстане действует ГОСУДАРСТВЕННЫЙ СТАНДАРТ РЕСПУБЛИКИ КАЗАХСТАН ИНФОРМАЦИОННЫЕ ТЕХНОЛОГИИ. ЭЛЕКТРОННОЕ ИЗДАНИЕ. Электронное учебное издание. СТ РК 34.017-2005. На сегодняшний этот стандарт устарел, его нужно актуализировать. Предлагается добавить в новые блоки для интеллектуализации электронных учебников. В частности:

- Модель базы знаний.
- Модель студента.
- Модель тьютора.
- Модель пользовательского интерфейса.
- Интеллектуальное обучение.
- Интеллектуальный контроль знаний.

УДК 004

Б.Ж.ШАРИПОВ

ПРОБЛЕМЫ И РЕАЛЬНОСТИ ПОДГОТОВКИ ВЫСОККВАЛИФИЦИРОВАННЫХ СПЕЦИАЛИСТОВ ДЛЯ СОВРЕМЕННОЙ ЭКОНОМИКИ

(проректор АО «Финансовая академия», д.п.н, к.т.н., академик МАИИ)

Новые высокотехнологические технологии, непрерывно меняются и совершенствуются каждые 3–5 года. Адаптироваться к непрерывно изменяющемуся миру, быть в нем активным действующим лицом, главным преобразователем, творцом и двигателем прогресса, оставаясь высококультурным человеком – вот задача специалиста будущего.

Проблемы подготовки специалистов будущего трансформируются вместе с политическими и социально-экономическими изменениями в стране и обществе, а их содержание и средства решения отражают позицию государства и общества в отношении системы технического образования, а также возможности конкретных ВУЗов и колледжей в новых условиях развития.

В экономике Казахстана и России да и во всех стран ЕАЭС сегодня наблюдается дефицит квалифицированных рабочих, техников, инженеров и это один из факторов, сдерживающих развитие целых отраслей в регионах и странах в целом. Остроту и серьезность проблемы отмечают крупные компании, лидеры экономики. Некоторые из них сами начинают активно влиять на систему профессионального образования. Почти 25% компаний считают ключевым ограничением для инновационного развития отсутствие необходимых кадров, 70% компаний полагают, что столкнутся с нехваткой кадров в течение ближайших двух-трех лет. Нехватка кадров нужной квалификации ограничивает текущее развитие 60% компаний.

В тоже время наблюдается такой парадокс в колледжах Казахстана обучается более чем полмиллиона человек. Ежегодное число выпускников колледжей составляет около 200 тыс. человек. Из числа выпускников трудоустраиваются либо поступают в вузы лишь 50 тыс., а остальные 150 тыс. остаются в большинстве своем нетрудоустроенными. Такие показатели носят угрожающий характер, и положение лишь ухудшается

Министерство образования и науки РК утверждает, что 75% выпускников колледжей и вузов трудоустраиваются после окончания учебы. Однако пока нет четкого взаимодействия между бизнесом и учебными заведениями, сложно говорить об эффективности высшего и среднего специального образования [1].

Никто точно не знает, на какую конкретно работу устроился математик, окончивший Казахский государственный университет, работают ли педагогами те, кто окончил педагогический вуз, в какой компании и на какой конкретно позиции работает программист, окончивший, например, технологический колледж. И ведь дело не в том, сколько человек после окончания колледжа или вуза куда устроилось работать. Дело в том, что и, как 10-20 лет назад, в Казахстане учебные заведения оторваны от реальной жизни и потому просто штампуют специалистов и раздают дипломы. И пока ничего не изменилось [1].

Реализация программы индустриализации в Казахстане предъявляет новые требования к подготовке кадров. Формально в образование вводятся инновации: многоуровневая подготовка, кредитная система, дистанционное обучение. Главная же проблема в том, что годами не меняется содержание обучения, хотя требования на порядок выросли.

Инновации в большинстве случаев декларируются, а в реальности продолжается чтение лекций по старым конспектам. Студенты занимаются по устаревшим учебникам, на оборудовании 70-80-х годов. Как в такой ситуации можно обеспечить должное качество?

В ближайшие годы в стране будут созданы новые производства: автомобилестроение, авиастроение, морской транспорт, атомная энергетика. Подготовка кадров для этих сфер еще не начата или только-только начинается.

Система образования слишком неповоротлива. Продолжается подготовка по невостребованным профессиям и специальностям. Осмысленно прогнозами развития рынка труда Правительство занялось только сейчас [2].

Сведения по потребности рабочей силы в Карте занятости обоснованы недостаточно, прогноз рынка труда сведен к простой экстраполяции. Ни о каких перспективных долгосрочных планах нет и речи на сегодня. Официальная информация о прогнозах по трудовым ресурсам носит общий характер, не структурирована по отраслям и профессиям.

В настоящее время нет реальной картины о состоянии рынка труда, нет полной статистики о профессионально-квалификационном уровне работающих, о количестве рабочих по разрядам, нет качественных данных о безработных, самозанятых, нет информации о количестве обучающихся по разным специальностям.

В вопросах определения потребностей в кадрах и их компетенции нет четкой координации между государственными органами. Формирование

госзаказа идет по текущим вакансиям, которые уже в ходе учебы становятся неактуальными. С таким положением вещей мириться решительно невозможно. Современный мир быстро меняется, и в системе образования обязательно должен быть

заложен институт изучения, мониторинга этих изменений и механизм оперативного реагирования на них [2].

По данным Всемирного экономического форума, в 2013 году казахстанский рынок труда занимал достаточно высокую 21-ю позицию, однако по ряду показателей наша страна имеет низкий рейтинг, профессионализм управления - 116-ое место, доступность переподготовки штатного персонала - 96-ое место, утечка мозгов -88-ое место, кооперация труда и работодателей -78-ое место.

Сегодня отечественный рынок труда нуждается не в кадрах вообще, а в кадрах высококвалифицированных.

Это подготовка и переподготовка кадров для успешной реализации Государственной программы форсированного индустриально-инновационного развития. Важно развивать дуальное профессиональное образование, создавать современные центры прикладных квалификаций, не говоря об отраслях инновационных. Сегодня есть проблемы с кадрами промышленности даже в сельском хозяйстве.

Основные шаги в направлении повышения качества подготовки технических специалистов, это создание и обновление профессиональных стандартов. Профессиональные стандарты - каналы взаимосвязи между образованием и бизнесом, которые определяют минимальные требования к знаниям, навыкам и компетенциям на каждом профессиональном уровне. На его основе должны быть обновлены или разработаны новые образовательные программы.

Профессиональные стандарты станут также основой для проведения независимой оценки и присвоения квалификации при непосредственном участии предприятий.

Другим шагом будет являться внедрение системы определения и прогнозирования потребности в кадрах в разрезе конкретных профессий что позволит ориентироваться на реальные потребности рынка труда.

Следующий шаг - повышение качества образовательного процесса. Система образования должна готовить кадры, которые вместе с теоретическими знаниями получают необходимые навыки для быстрой адаптации в реальных производственных условиях.

В техническом и профессиональном образовании необходимо укреплять тесные партнерские отношения с бизнесом. Вузы должны быть ориентированы на потребности работодателей. Программы обучения должны быть направлены на формирование определенных компетенций через передачу студенту интегрированных знаний, умений и навыков.

Для условий Казахстана была предложена относительно новая система подготовки специалистов – бакалавриат производственного направления (БПН). Совместно с Министерством образования и науки Республики Казахстан начат пилотный проект БПН в 4 ведущих вузах страны: Евразийском национальном университете (г. Астана), Международном университете информационных технологий (г. Алматы), Восточно-Казахстанском государственном техническом университете (г. Усть-Каменогорск), Карагандинском государственном техническом университете (г. Караганда) по специальностям сферы информационно-коммуникационных технологий (ИКТ).

Характеристика квалификации выпускника БПН. Квалификация выпускника БПН соответствует шестому квалификационному уровню отраслевой рамки квалификации (ОРК). Введение программ БПН не требует увеличения числа уровней ОРК или включения дополнительных подуровней.

Результаты образования выпускника БПН. Выпускник программы БПН наряду с академической степенью бакалавра получает профессиональную квалификацию подтвержденную сертификатами и трудовую книжку.

Профессиональная квалификация – результат освоения программы БПН и ее отличительная особенность – достигается за счет особого содержательного наполнения программы.

Определение требований к квалификации выпускников по программам БПН производится в зависимости от потребностей рынка труда в работниках соответствующей квалификации.

Квалификация всегда является результатом освоения определенной образовательной программы, в том числе, практического опыта (как результата неформального образования). Достичь определенного квалификационного уровня возможно разными путями.

Работодатели в основном предпочитают принимать на работу людей с высшим профессиональным образованием. Выпускники среднего профессионального образования часто превосходят их в области практической подготовки, однако проигрывают в плане общей компетенции работника (фундаментальности подготовки), связанной с масштабом деятельности, ценой возможной ошибки, способностью действовать в нестандартной рабочей ситуации, а также с полнотой реализации в профессиональной деятельности основных функций руководства (целеполагание, организация, контроль, мотивация исполнителей). Выпускники высшего профессионального образования отличаются объёмом и степенью сложности используемой информации, инновационностью применяемых знаний и степенью их абстрактности.

Принципиальным отличием содержания программ БПН от программ углубленной подготовки в технических колледжах является наличие теоретической подготовки, соответствующей уровню высшего образования при сохранении методологического подхода к формированию профессиональной квалификации, существующего в колледжах - введение в структуру программы профессиональных модулей.

Принципиальным отличием БПН от академического является наличие профессиональной (прикладной) квалификации как результата освоения программы БПН.

Выпускники программ БПН отличаются практикоориентированным характером полученных умений и знаний. На этом основании можно прогнозировать их большую востребованность на рынке труда, в тех областях, где требуется профессиональная квалификация наряду с теоретической подготовкой, соответствующей высшему профессиональному образованию.

На рисунке 1 приведена технология подготовки специалистов в БПН.

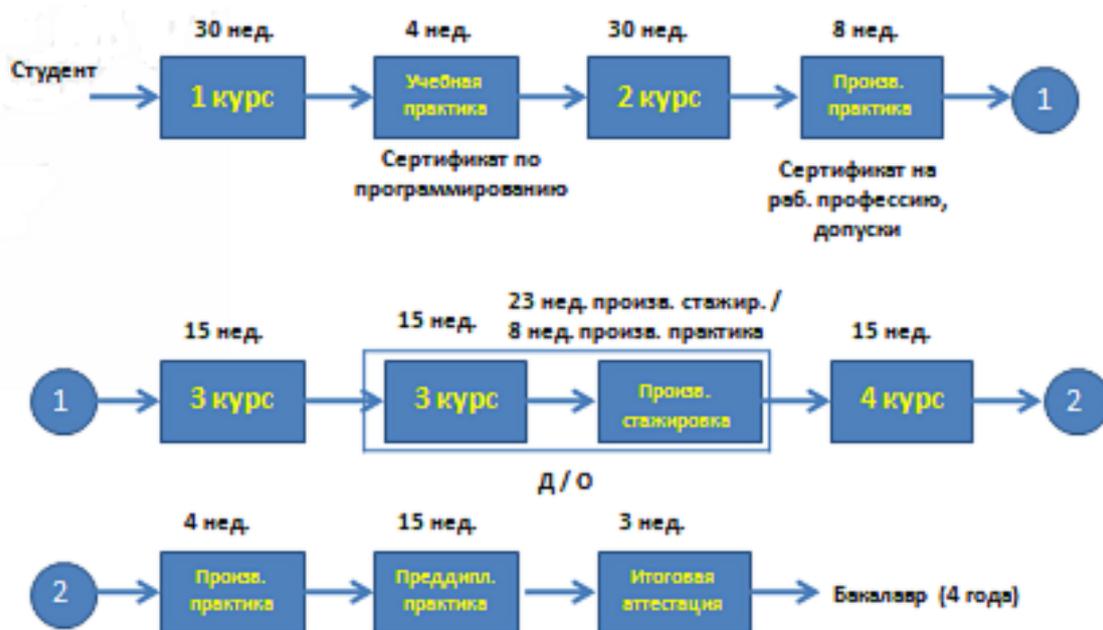


Рисунок 1 – Технология подготовки специалистов в БПН для ИКТ- специальностей

Как видно по схеме (рисунок 1) после окончания 1 курса (30 недель обучения) студент проходит учебную практику (4 недели) и получает сертификат по программированию (для

любой специальности сферы ИКТ), затем после 2 курса обучения он проходит 4 недельную производственную практику с целью получения рабочей специальности (сертификат). Вид рабочей специальности определяется направлением подготовки данного студента. Например, студент специальности «Информационные системы», «Автоматика и управление» помимо рабочей специальности должен получить и допуск к оборудованию до 1000 V (удостоверение).

На 3 курсе после зимней сессии и каникул студент направляется на производственную стажировку на 23 недели, обучение в период весеннего семестра производится по дистанционной форме. Следует подчеркнуть, что в рабочих планах по возможности произведено смещение дисциплин, часть специальных дисциплин читается на более ранней стадии, чем на традиционных курсах обучения. Общественно-политические, экономические дисциплины, охрана труда и т.д. переведены на дистанционную форму.

Студенты, заболевшие в период производственной стажировки или которые не смогли устроиться на производственную в обязательном порядке проходят 8 недельную производственную практику.

Наличие у студентов сертификатов на рабочую профессию, наличие допуска к работе на электрооборудовании и увеличенный срок производственной стажировки делают шансы студентов на получение места практики очень высокими. К тому же, отдельные работодатели высказались о возможности оплаты труда студентов согласно тарифной сетки. После окончания практики работодатель оформляет трудовую книжку с записью выполняемой работы в период стажировки.

Следует подчеркнуть, что в период производственной стажировки за студентом закрепляется руководитель от производства (наставник), который одновременно является одним из руководителей дипломного проекта студента. Тема дипломного проекта выбирается по проблемным вопросам производства.

Возвращаясь на очное обучение в 4 курсе, студент при выборе спецдисциплин, уже сознательно ориентируется в курсах по выбору и знает свои пробелы в знаниях, выявленные в ходе производственной стажировки.

После зимних каникул студент может выполнять дипломный проект или в стенах вуза, или на производстве. Защита дипломного проекта может быть предусмотрена также или в стенах вуза, или на производстве.

В новом учебном году 2016/2017 года планируется первый прием студентов на БПН во всех 4 пилотных вузах по различным специальностям сферы ИКТ. Это начало пути, работы предстоит еще много и как по организации дистанционного обучения, так и по разработке пакета нормативно-правовой документации по пилотному проекту "Бакалавриат производственного направления":

- разработка проекта предложений по гарантиям и преференциям для работодателей по государственно-частному финансированию подготовки, и переподготовки ИКТ-специалистов;
- проект предложений по введению в законодательство РК (в Закон "Об Образовании") понятий и механизмов сетевого взаимодействия учебных заведений, механизмов взаимодействия сферы производства и сферы образования (в Кодекс о труде) для реализации проекта БПН.

Не смотря все сложности, мы полны оптимизма и уверенности в реальности нашего проекта и его значимости для удовлетворения требований современной экономики Казахстана.

Литература

1 Почему сложно трудоустроить выпускников вузов и колледжей? - <https://kapital.kz/gosudarstvo/36952/pochemu-slozhno-trudoustroit-vypusknikov-vuzov-i-kolledzhej.html>.

2 Материалы по Парламентским слушаниям «О законодательном обеспечении реализации государственной программы развития технического и профессионального образования в Республике Казахстан» от 16 мая, 2014 г.

Г.Б.ШАХМЕТОВА¹, А.Н. ИСАЙНОВА¹, Г.Ж. ИСАБАЕВА²**МОБИЛЬНОЕ ОБУЧЕНИЕ***(¹ЕНУ им. Л.Н.Гумилева, г.Астана, ²АО «Финансовая академия», г.Астана, Казахстан)*

Термин М-обучение или "мобильное обучение", имеет различное значение для разных сфер, которое является подвидом электронного обучения, образовательных технологий и дистанционного обучения, и которое фокусируется на обучении посредством мобильных устройств. Мобильное обучение имеет множество разных определений, и оно известно так же как М-обучение, индивидуальное обучение, обучение в любом месте в любое время и портативное обучение. Одно из определений мобильного обучения звучит следующим образом: «это любой вид обучения, который происходит, когда учащиеся находятся в любом удобном для обучения месте, используя мобильные технологии». (MOBLearn., 2003). Другими словами можно сказать, что с использованием мобильных устройств, учащиеся могут обучаться в любом месте и в любое время (Crescente и Lee, 2011). Мобильное обучение рассматривает возможность использования мобильных устройств для поддержки преподавания и обучения.

Мобильное обучение это определено не просто объединение двух терминов «мобильное» и «обучение»; оно всегда неявно подразумевает «мобильное электронное обучение» и его история и развитие должны восприниматься как продолжение «обычного» электронного обучения и реагировать на это «простое» электронное обучение с его преимуществами и недостатками. Именно аспект мобильности делает мобильное обучение отличным от других форм обучения, а именно специфическое проектирование теоретического и практического материала с учетом использования возможностей мобильных устройств. М-обучение сфокусировано на мобильности обучающегося, взаимодействующего с портативными технологиями, и на обучении что отражает акцент на том, как общество и его институты могут взаимодействовать и поддерживать все более мобильное население. Это потому, что мобильные устройства имеют возможность и функциональность для поддержки учащихся. Например, подкасты лекций могут быть доступны для скачивания, и обучающийся может изучить материал, находясь далеко от традиционного места учебы.

За последние десять лет мобильное обучение выросло из мелких исследовательских исследований до более значимых проектах в школах, рабочих мест, музеев, городах и сельских районах по всему миру. М-образовательное сообщество по-прежнему фрагментировано, с разными национальными перспективами, еще имеет различия между научными учреждениями и промышленностью, а также между школой, высшим образованием и сектором самообразования.

В мобильном обучении использую мобильные и портативные IT-устройства, такие как смартфоны, нетбуки, планшеты и другие. Так как, на сегодняшний день, компьютеры и Интернет стали неотъемлемой частью учебного процесса, информационные технологии стали более доступными, эффективными и легкими в использовании, все эти возможности открыли широкие перспективы для расширения участия ИКТ в процессе обучения. Мобильные девайсы, такие как смартфоны и телефоны более доступны по ценовой характеристике, чем настольные компьютеры, поэтому представляют собой менее дорогостоящий доступ выхода в Интернет.

Большинство мобильных устройств могут быть полезны в области образования, управления и преподавания, а так же техническими средствами поддержки обучения для учащихся.

Можно сказать, что существуют такие преимущества, как:

- Обучающиеся могут взаимодействовать между собой и с преподавателем;
- Намного проще разместить в классе несколько мобильных девайсов, чем дорогостоящие ПК;
- Карманные и планшетные ПК намного легче и занимают меньше места, чем учебники, тетради и различные бумаги.
- У учащихся и преподавателей появляется возможность быстрого обмена заданиями и совместной работы, т.е. работы могут передаваться друг другу внутри группы, производится совместные обсуждения и исправления.
- Мобильные устройства могут быть использованы в любом месте, в любое время, в том числе дома, в поезде, в гостиницах - это неограниченное значение для обучения по месту работы.
- Обучение становится действительно индивидуальным. Учащиеся могут сами выбрать подходящую для них траекторию обучения с учетом их интересов, отсюда можно сказать, что М-обучение является ориентированным на студента.
- Немедленный доступ к любой информации, необходимой для конкретной работы, посредством мобильного устройства повышает производительность обучающегося.

Так же можно отметить, что в последнее время появилась тенденция смешанного обучения. Эта форма обучения включает в себя различные виды преподавания, что делает процесс более эффективным и интересным. Смешанное обучение сочетает в себе преимущества различных форм обучения, и лучше всего подходит к контексту обучения в интерактивной учебной среде. Мобильное обучение можно комбинировать и с другими видами обучения, данный метод обеспечит интерактивные условия обучения для студентов[3].

Говоря о новом направлении в учебном процессе, нельзя не затронуть и недостатки мобильного обучения. Существуют следующие возможные проблемы:

- Мобильные устройства имеют небольшого размера экраны, что в свою очередь ограничивают количество и тип доносимой информации.
- Существуют ограниченные возможности для хранения мобильных телефонов.
- Батареи имеют небольшое время работы, что может затруднить процесс обучения, к примеру, могут быть потеряны данные, если студент не успел вовремя их сохранить.
- У мобильных телефонов гораздо менее надежны, чем настольные компьютеры.
- Можно столкнуться с трудностями воспроизведения графики, больших объемов видео материала.
- Рынок мобильных устройств весьма динамичен и быстро изменяется, так что устройства могут устареть очень быстро.
- Пропускная способность может снизиться при большом количестве пользователей, использующих беспроводные сети.

В связи с вышесказанным, можно выделить тот факт, что будущее сферы мобильного обучения требует совместных усилий мобильных производителей, поставщиков мобильных услуг, а также экспертов индустрии обучения.

Литература.

1. Mobiledu and Widsets for China. <http://mobiled.uiah.fi/?p=67> (дата обращения 17.04.2017)
2. Бабичев Н.В., Водостоева Е.Н., Масленикова О.Н., Соколова Н.Ю. Роль и значение интерактивных наглядных пособий в системе современного биологического образования. <http://e-drofa.ru/aboutnavigator/40> (дата обращения 17.04.2017)
3. Голицина И.Н., Половникова Н.Л. Мобильное обучение как новая технология в образовании. http://ifets.ieee.org/russian/depository/v14_i1/html/1.htm (дата обращения 17.04.2017)

СЕКЦИЯ 2

SECTION 2

АҚПАРАТТЫҚ ҚАУІПСІЗДІК
ИНФОРМАЦИОННАЯ БЕЗОПАСНОСТЬ
INFORMATION SECURITY

А.Б. БАРЛЫБАЕВ

СУВЕР – ОНЛАЙН-ПЛАТФОРМА ДЛЯ БЕЗОПАСНОГО ШИФРОВАНИЯ ДАННЫХ

*(НИИ «Искусственный интеллект» Евразийский национальный университет
им. Л.Н.Гумилева, Астана, Казахстан)*

Цель проекта СУВЕР – разработка онлайн-платформы для безопасного шифрования данных.

Проект направлен на исследование и создание технологий повышенной конфиденциальности (ТПК). ТПК должны быть доступны в широком спектре продуктов и услуг, с пригодными для использования, удобными и доступными безопасными опциями. Основными задачами проекта являются: Анализ методов и сервисов шифрования, в том числе анализ таких сервисов на доступность в сети, удобство работы пользователя и дружелюбность интерфейса; Разработка платформы сувер для бесплатного онлайн сервиса шифрования (дешифрования) различных файлов пользователей на основе симметричных алгоритмов с использованием различных режимов шифрования (современные, современные сложные, современные ультра сложные); Разработка сервиса оплаты за шифрование, при условии что данные «большие», (с подключением сервиса «Donate») по очень умеренной цене; Разработка сервиса учета пользователей через мобильный номер, передача ID шифрования на мобильный номер для безопасного дешифрования файла; Разработка сервиса risk management framework для выгрузки пользователю отчета и пояснения шифрования.

Рабочий план проекта:

1. Анализ методов и сервисов шифрования, в том числе анализ таких сервисов на доступность в сети, удобство работы пользователя и дружелюбность интерфейса.
 - a. Анализ методов и сервисов шифрования.
 - b. Анализ доступности сервисов шифрования в сети.
 - c. Анализ сервисов шифрования на удобство работы пользователя и дружелюбность интерфейса.
 - d. Разработать онтологию методов шифрования.
 - e. Разработать онтологию сервисов шифрования.
2. Определить основные методы и сервисы, которые будут разрабатываться в рамках проекта.
 - a. Определить основные методы и сервисы шифрования, которые необходимо разработать в рамках Проекта.
 - b. Подготовка спецификаций требований для функций подсистем и пользователей.
 - c. Подготовка спецификаций требований к инструментам и технологиям.
3. Разработка платформы сувер для бесплатного онлайн сервиса шифрования (дешифрования) различных файлов пользователей на основе симметричных алгоритмов с использованием различных режимов шифрования (современные, современные сложные, современные ультра сложные).
 - a. Проектирование архитектуры сувер.
 - b. Проектирование базы данных и интерфейсов.
 - c. Разработка базы данных.
 - d. Разработка интерфейса платформы сувер.

е. Разработка бесплатного онлайн сервиса шифрования различных файлов с использованием различных режимов шифрования (современные, современные сложные, современные ультра сложные).

ф. Разработка сервиса дешифрования.

4. Разработка библиотек процедур и функций, реализующих шифрование и дешифрование файлов разных пользователей на основе симметричных алгоритмов с использованием различных методов и режимов шифрования.

а. Разработка бесплатного онлайн-сервиса библиотек процедур и функций для шифрования и дешифрования симметричных алгоритмов на основе использования современных методов и режимов шифрования.

б. Разработка бесплатного онлайн-сервиса библиотек процедур и функций для шифрования и дешифрования симметричных алгоритмов на основе использования современных сложных режимов и методов шифрования.

с. Разработка бесплатного онлайн-сервиса библиотек процедур и функций для шифрования и дешифрования симметричных алгоритмов на основе использования современных сверхсложных режимов и методов шифрования.

5. Разработка сервиса оплаты за шифрование, при условии что данные «большие», (с подключением сервиса «Donate») по очень умеренной цене.

а. Анализ цен онлайн сервисов шифрования и определение оптимально низкой цены за шифрование больших файлов.

б. Открытие аккаунтов cyber в популярных сервисах интернет оплаты и интеграция их с cyber,

с. Разработка интерфейса оплаты за шифрование «больших» файлов.

6. Разработка сервиса учета пользователей через мобильный номер, передача ID шифрования на мобильный номер для безопасного дешифрования файла.

а. Разработка регистрации пользователей.

б. Разработка функции изменения пароля.

с. Разработка авторизации.

д. Разработка ведения учетной записи пользователя.

е. Разработка привязки телефонного номера пользователя и отправки специального кода на него для дешифрования файла.

7. Разработка сервиса журнализации шифрования и дешифрования, блокировки дешифрования.

а. Разработка сервиса журнализации шифрования.

б. Разработка сервиса журнализации дешифрования.

с. Разработка сервиса блокировки дешифрования.

8. Разработка сервиса risk management framework для выгрузки пользователю отчета и пояснения шифрования.

а. Разработка сервиса для выгрузки пользователю отчета шифрования.

б. Разработка сервиса для выгрузки пользователю пояснения процесса шифрования.

9. Разработка сервиса feedback обратной связи с пользователями в виде форума для улучшения работы сервисов cyber в целом.

а. Разработка форума.

б. Разработка обратной связи пользователей с разработчиками.

10. Разработка сервиса тонкого клиента для шифрования и дешифрования файлов на стороне клиента не загружая файлы на сервер.
 - a. Разработка сервиса тонкого клиента для шифрования файлов на стороне клиента не загружая файлы на сервер.
 - b. Разработка сервиса тонкого клиента для дешифрования файлов на стороне клиента не загружая файлы на сервер.
11. Анализ юзабилити системы для ее улучшения.
 - a. Обзор методов анализа юзабилити.
 - b. Анализ юзабилити.
 - c. Построение требований юзабилити.
12. Улучшение интерфейса для достижения требований юзабилити.
13. Пилотное тестирование.
 - a. Разработка программ и методик тестирования.
 - b. Тестирование и модификация.
 - c. Создание руководства администратора и пользователя.
14. Управление проектом и координация.
 - a. Координация и управление проектом.
 - b. Стартовое совещание и отчетность;
 - c. Партнерская деятельность и финансовая отчетность;
 - d. Совещание по статусу и отчетность;
 - e. Завершение и проектная документация.
15. Распространение и эксплуатация.
 - a. Продвижение результатов проекта.
 - b. Подготовка ежемесячных отчетов по использованию cyber.
 - c. Запуск сервера cyber и его периодичное обновление.

УДК 003.26.09

Г.Т. БЕКМАНОВА, И.С. КОРОБОВ, Г.Ж.МУҚАТАЙ

ПРЯМОЕ ИСПРАВЛЕНИЕ ОШИБОК ПО СТАНДАРТУ IEEE 802.16

(Евразийский национальный университет им.Л.Н.Гумилева, Астана, Казахстан)

Широкополосные сигналы - это сигналы, у которых произведения активной ширины спектра F на длительность T много больше единицы. Это произведение называется базой сигнала B для широкополосных сигналов

$$B = FT, (FT \gg 1)$$

Повышение базы в широкополосных сигналах достигается путем дополнительной манипуляции (или модуляции) на времени длительности сигнала по фазе или частоте. В результате, спектр сигнала F (если его длительность T остается неизменным) значительно расширяется. Дополнительная внутри сигнальная манипуляция по амплитуде используется редко. [1]

Одна из наиболее острых проблем в информационных технологиях – это защита данных от разрушения. Как каналы передачи данных, так и носители информации на сегодняшний день остаются далекими от совершенства, несмотря на все усилия производителей современных аппаратных средств. Кабельные и беспроводные линии

передачи информации подвержены воздействию внешних помех, искажающих форм передаваемых сигналов и тем самым делающих невозможным однозначное распознавание информации на стороне приемника, магнитные и оптические носители информации чувствительны к физическим повреждениям, делающим невозможным чтение информации из отдельных участков на поверхности носителя. В такой ситуации особенно актуальным становится применение специальных технологий информационного резервирования, позволяющих, как минимум, обнаруживать искажения, а как максимум, не только обнаруживать, но и исправлять искажения.

Для этого согласно стандарту IEEE 802.16-2012 для радиointерфейса систем широкополосного беспроводного доступа используются коды Рида Соломона.[2]

Канальное кодирование состоит из трех этапов: рандомизатор, прямое исправление ошибок(FEC) и перемежение.

Кодирование Рида-Соломона должно быть получено из систематического кода RS ($N = 255, K = 239, T = 8$) с использованием GF (2^8), где

N - количество общих байтов после кодирования

K - это количество байтов данных перед кодированием

T - количество байтов данных, которое может быть скорректировано

Для систематического кода используются следующие многочлены:

$$\text{Генератор} \quad \text{кода } g(x) = (x + \lambda^0)(x + \lambda^1)(x + \lambda^2)\dots(x + \lambda^{2T-1}), \quad (1)$$

$$\text{Полином:} \quad X = 02_{HEX}$$

$$(2)$$

$$\text{Полевой полином поля } p(x) = x^8 + x^4 + x^3 + x^2 + 1$$

Этот код укорачивается и прокалывается, чтобы обеспечить возможность изменения размеров блоков и возможности исправления ошибок. Когда блок сокращен до K' байтов данных, добавьте $239-K'$ нулевые байты в качестве префикса. После кодирования отбросьте эти $239-K'$ нулевые байты. Когда кодовое слово проколото, чтобы можно было исправить T байт, должны использоваться только первые $2T$ из всех 16 байтов четности. Преобразование бит / байта должно быть сначала MSB.

Каждый RS-блок кодируется двоичным сверточным кодером, который должен иметь собственную скорость $1/2$, длину ограничения, равную 7, и должен использовать коды генераторных многочленов, показанные в уравнении (3), для получения его двух кодовых битов.

$$\begin{aligned} G_1 &= 171_{OCT} \quad \text{FOR } X & (3) \\ G_2 &= 133_{OCT} \quad \text{FOR } Y \end{aligned}$$

Кодирование выполняется, сначала передавая данные в блочном формате через RS-кодер, а затем передавая их через сверточный кодер. Один конец байта 0x00 добавляется в конец каждого пакета. Этот хвостовой байт должен быть добавлен после рандомизации. В RS-коде избыточные биты отправляются перед входными битами, сохраняя 0x00 хвостовой байт в конце выделения. Для обеспечения того, чтобы количество битов после сверточного кодера делилось на N_{cbps} , нулевые (0b0) биты подушки добавляются после нулевых хвостовых битов перед кодировщиком. Биты нулевой подушки не рандомизированы. Обратите внимание, что такая ситуация может возникнуть только при подканализации. В этом случае кодирование RS не используется.

Когда применяется подканализация, FEC должен обходить RS-кодер и использовать общую скорость кодирования, как скорость CC-кода. Размер некодированного размера блока и закодированного блока можно вычислить, умножив значения, на количество выделенных

КОДЕР ДЕКОДЕР РИДА-СОЛОМОНА

(Евразийский национальный университет им.Л.Н.Гумилева, Астана, Казахстан)

Коды Рида-Соломона (РС) нашли широкое применение в системах хранения данных, в системах радиосвязи и дальней космической связи, в цифровом телевидении. Стандартное описание кодов Рида-Соломона основывается на понятии минимального кодового расстояния и гарантированно исправляемой кратности ошибок, что по сути своей способствует выбору кодов с заданными параметрами[2].

В общем случае коды РС - это недвоичные блочные циклические коды, которые позволяют исправлять ошибки в блоках данных. На (рис.1) представлена классификация корректирующих кодов [1].

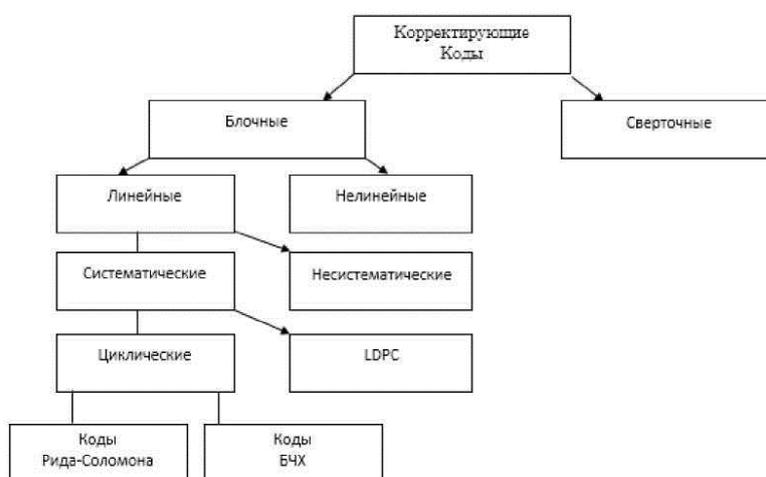


Рис.1. Классификация корректирующих кодов

Недвоичными коды РС называют потому, что элементами кодового слова выступают, не биты, а группы битов (байты). Для восьмибитовых символов полная длина кодового слова составит $2^8 - 1 = 255$ символов.

В настоящее время широко используется в системах восстановления данных с компакт-дисков, при создании архивов с информацией для восстановления в случае повреждений, в помехоустойчивом кодировании.

Типовая система кодера и декодера Рида-Соломона (рис.2) представлена ниже



Рис. 2. Схема коррекции ошибок Рида-Соломона

Кодировщик Рида-Соломона берет блок цифровых данных и добавляет дополнительные «избыточные» биты. Ошибки происходят при передаче по каналам связи или по разным причинам при запоминании (например, из-за шума или наводок и т.д.). Декодер Рида-Соломона обрабатывает каждый блок, пытается исправить ошибки и

восстановить исходные данные. Число и типы ошибок, которые могут быть исправлены, зависят от характеристик кода Рида-Соломона.

Исходные данные количество символов кодируемого сообщения (называемого также информационным словом) по общепринятому соглашению обозначается буквой k ; полная длина кодового слова, включающего в себя кодируемые данные и символы четности, n . Отсюда, количество символов четности равно: $n - k$. За максимальным количеством исправляемых ошибок «закреплена» буква t . Поскольку для исправления одной ошибки требуется два символа четности, общее количество символов четности равно $2t$. Выражение $RS(n, k)$ описывает определенную разновидность корректирующих кодов Рида-Соломона, оперирующую с n -символьными блоками, k -символов, из которых представляют полезные данные, а все остальные задействованы под символы четности.

Алгоритм действий Кодировщика Рида-Соломона(рис.3).

При передаче данных происходят ошибки. Современный радиозфир сильно зашумлен антропогенными и природными шумами, и при передачи данных могут быть искажения от входящих данных в приемнике. Чтобы обнаружить какие данные были зашумлены, и получить исходное сообщение используем код Рида-Соломона. В программной реализаций добавляем шум в виде вектора.



Рис.3. Алгоритм кодирования информационного сообщения с помощью кодов РС

Операция декодирования состоит из следующих этапов [3]:

1. Вычисления синдрома ошибки (синдромный декодер);
2. Построения полинома ошибки, осуществляемое (в нашем случае) посредством высокоэффективного, алгоритма Берлекэмп-Мессис;
3. Нахождение корней данного полинома, решается лобовым перебором (использование алгоритма Ченя);

4. Определение характера ошибки, сводящееся к построению битовой маски, вычисляемой на основе обращения алгоритма матрицы;

5. Исправление ошибочных символов путем наложения битовой маски на информационное слово и последовательного инвертирования всех искаженных бит через операцию XOR.

В цифровых каналах связи данные передаются в виде последовательного потока бит, соответственно шум в канале определяется вероятностью искажения одиночного бита. При этом количество исправленных ошибок определяется числом байт и не зависит от того сколько искаженных бит находится в каждом байте, отсюда следует, что результат коррекции зависит от характера распределения ошибочных бит в информационном потоке.

Общая схема архитектуры декодера кодов Рида-Соломона показана на рисунке 4.

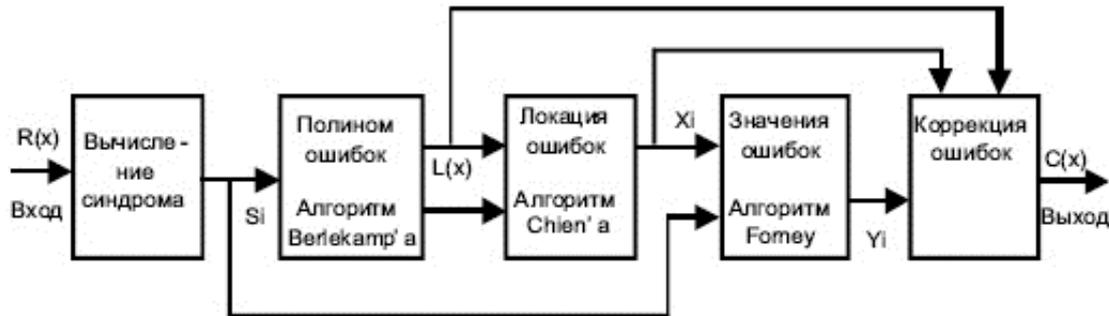


Рис. 4. Схема работы с кодами Рида-Соломона

Обозначения:

- $r(x)$ – Полученное кодовое слово
- S_i – Синдромы
- $L(x)$ – Полином локации ошибок
- X_i – Положения ошибок
- Y_i – Значения ошибок
- $c(x)$ – Восстановленное кодовое слово
- v – Число ошибок

Полученное кодовое слово $r(x)$ представляет собой исходное (переданное) кодовое слово $c(x)$ плюс ошибки:

$$r(x) = c(x) + e(x)$$

Декодер Рида-Соломона пытается определить позицию и значение ошибки для t ошибок (или $2t$ потерь) и исправить ошибки и потери.

Если говорить упрощенно, то основная идея помехозащитного кодирования Рида-Соломона заключается в умножении информационного слова, представленного в виде полинома D , на неприводимый полином G , известный обеим сторонам, в результате чего получается кодовое слово C , опять-таки представленное в виде полинома.

Декодирование осуществляется с точностью до наоборот: если при делении кодового слова C на полином G декодер внезапно получает остаток, то он может рапортовать вверх об ошибке. Соответственно, если кодовое слово разделилось нацело, его передача завершилась успешно

Список литературы

1 Владимирова С.С. Анализ эффективности декодирования циклических кодов Рида-Соломона с использованием двойственного базиса // Материалы диссертации на соискание ученой степени кандидата технических наук. - СПб. - 2013

2 Быков В.В., Меньшиков К.В. Помехоустойчивые коды цифрового телевидения // Т-Comm. - №9 - 2013. - С. 30-32.

3 Ничипорук Н.Е., Сай С.В. Оценка корректирующей способности кодов Рида-Соломона при передаче подводных изображений через зашумленный канал связи // Вестник ТОГУ. - №3(22). - 2011. - С. 31-34.

АВТОНОМДЫ РОБОТТАР МЕН РОБОТТЫ ЖҮЙЕЛЕРДІ ЖАСАУДАҒЫ НАВИГАЦИЯ ҚАУПСІЗДІГІ

(Л.Н.Гумилев ат.Еуразия ұлттық университеті, Қазақстан Республикасы. Астана қ.)

Қазіргі таңда күзет саласында, барлау жұмыстарында, төтенше жағдайлар зардабының алдын алуда, т.б. арнайы салаларда роботтарды қолдану кеңінен қолданыс табууда. Жер үсті роботтарын, ұшқышсыз ұшу аппараттарын жасау барысындағы басты жұмыс автономды жүйелерді басқару жүйесін жасау. Барлау, күзету және іздеу жұмыстарымен айналысатын автономды роботтар бейтаныс аумақтарда қозғалысты жүзеге асыра білуі қажет. Өз бетімен қозғала алатын барлық мобильді роботтардың негізгі проблемасы навигация болып табылады. Кеңістікте дұрыс бағытты таба білу үшін автономды роботтың борттық жүйесі маршрут құру, қозғалыс параметрлерін басқару (бұрылу бұрыштары мен жылдамдықты дұрыс беру), қоршаған орта жайлы ақпаратты дұрыс тану және тұрақты түрде өз координатасын бақылау қызметін атқара алуы тиіс. Осыған байланыст азаматтық салада болсын, әскери салада болсын навигациялық құрылғыларды қолдану кеңінен белең алуда. Мысалы, АҚШ полиция басқармасы қылмыспен күрес жүргізу үшін ұшқышсыз ұшу жүйелерін қолдануды бастаса, мексика полициясы құрамында есірткі бар өсімдіктер плантацияларын анықтау мақсатында ұшқышсыз ұшу аппараттарын қолдануда. Қытай елі де көштен қалмай көпфункционалы ұшқышсыз авиациялық полиция қызметі кешенін пайдалана бастаған [1].

Ал Израиль елі Guardian автономды жерүсті барлаушы машинасын шекара күзетінде, патруль қызметінде қолдануда [2].

Қазақстан Республикасының Қорғаныс Министрлігі де 2017 жылдан бастап елімізде Израильмен бірге ұшқышсыз ұшу аппараттарын шығаруды қолға алғанын айтса, ал Тараз қаласының әуежайының әскери секторында ұшқышсыз ұшу аппараттарын ұрысқа қолдану мен дайындау орталығы құрылуда екенін хабарлады[3].

Дегенмен, автономды құрылғылардың навигациялық жүйесіне шабуыл жасау фактілер бүгінгі күннің өзінде тіркеліп үлгерді. Нақты дәлел келтірмесе де Ирандықтар Америкалық ұшқышсыз ұшу аппаратын қондырып алғанын хабарлады. Олар мұны жүзеге асыруға GPS-спуфинг тәсілін (1-сурет), яғни спутниктік сигналды ауыстыруды қолдану арқылы қол жеткізгенін айтады [4, 5].



1-сурет. GPS-спуфинг

Бұған күмән келтіргендер де болды, дегенмен америкалық мамандардың тобы глобалды позициялау жүйесі спутниктерін ауыстыру мүмкін екенін дәлелдеді[6]. Зерттеуді

жүргізу барысында навигациялық құрылғыларға сырттан әсер етудің келесі түрлері анықталды:

- деректер шегінде шабуылдау;
- навигациялық қабылдағыштың программалық жабдықтамасына шабуылдау;
- тәуелді жүйелерге шабуылдау.

Келесі әдістер қолданылды:

- эфемеридтерді, яғни спутниктерді ағымдағы орбиталдық параметрлері жайлы тұрақты жаңартылып отыратын деректер жиынын фльсификациялау;

- ағымдағы уақыт жайлы теріс ақпарат беру;
- уақыт шкалаларын сәйкессіздендіру;
- құрылғының жүйелік программалық жабдықтамасына шабуыл;
- сигналды бағыттан жаңылдыру мақсатында өзгерту – spoofing.

Осы уақытқа дейін тек сигналдың тікелей өзіне шабуыл жасау әдістері айтылып келсе, зерттеуден кейін сигналды қабылдауыштарға да шабуыл жасауға болатындығына көз жеткізілді.

Атап айтсақ, зерттеулер барысында келесі мүмкіндіктер дәлелденді:

- навигациялық сигналдарды қабылдауыштарды берілетін ақпараттарда басқара отырып істен шығару мүмкіндігі;
- спутниктік навигациялық жүйелер негізінде салынған басқару инфрақұрылымдарына шабуыл жасау (нақты уақыт жүйелері, маңызды инфрақұрылым объектілері, аэродромдар, ұшақтар, ұшқышсыз роботтандырылған кешендер және т.б. жүйелер).

Сонымен қатар, мұндай шабуылдарға қарсы амалдардың айтарлықтай ойластырылмағандығы анықталды. Сигналды қабылдағыштарға сырттай программалы-аппараттық әсер етуді анықтаудың түрлі әдістерін қарастыра отырып, олардың роботтандырылған кешендер мен т.б. навигациялық жүйелерге қатысты өте нашар дамығанын көруге болады. Глобалды позициялау жүйелерінің қымбаттығы мен спутниктерге сәйкес қорғаныс құралдарын орнату мүмкін еместігін ескеретін болсақ, онда қорғаныс жүйелері қабылдағыштардың тікелей өздері үшін жасалуы тиіс.

Аталмыш ғылыми есепті шешудің, яғни шабуылды анықтаудың ең қолайлы әдісі - орта мән және орта квадраттық ауытқу моделі. Яғни белгілі бір x шамасы жайлы алдыңғы бақылаулар жайлы білетініміз оның орта шамасы

$$\mu = \sum x_i / n$$

және орта квадраттық ауытқуы:

$$\sigma = \sqrt{((x_1^2 + \dots + x_n^2) / (n - 1)) - \mu^2}$$

Олай болса жаңа бақылау кезіндегі шама сенімділік интервалы шегіне сыйыспаса, онда жүйеге белгілі дәрежеде сырттан әсер бар екенін анықтай аламыз[7]. Навигация қауіпсіздігін қамтамасыз ету үшін шабуылды анықтаумен қатар, инерциалды навигациялық жүйелерді пайдалану мүмкіндігін қарастыру қажет. Мәселен, Northrop Grumman компаниясының мамандары DARPA қорғаныстық-ғылыми агенттігінің тапсырысымен микрогироскоп жасауға кіріскені туралы хабарланды.

Қолданылған әдебиеттер тізімі:

1 <http://bp-la.ru/bespilotniki-specialnogo-naznacheniya/>

2 <https://ru.wikipedia.org/wiki/Guardium>

3 https://www.kt.kz/rus/state/v_voennom_sektore_aeroporta_taraza_sozdaetsja_centra_podgotovki_i_boevogo_primeneniya_bespilotnikov_1153626195.html

4 <https://ru.wikipedia.org/wiki/Спуфинг>

5 <https://xakep.ru/2012/07/02/58927/>

6 <http://www.gps-forum.ru/content/view/128/38/>

7 <http://sibac.info/conf/innovation/lv/49489>

8 <https://www.imena.ua/blog/gps-steal-yaht/>

ТЕЛЕФОН АРҚЫЛЫ ХАБАР АЛМАСУДЫ ҚОРҒАУ ТЕХНОЛОГИЯЛАРЫ

(Л.Н.Гумилев ат.Еуразия ұлттық университеті, Қазақстан Республикасы. Астана қ.)

Бүгінгі таңда телефон арқылы хабар алмасу бұл барлық адам үшін үйреншікті жайт. Дегенмен, дәл осы хабар алмасу түрінің де өзіндік кемшіліктері мен үшінші тараптың рұқсатсыз ену мүмкіндігі бар. Жалпы, телефон желілерін қорғау әдістері мен технологияларын айтпас бұрын, бұл желіде ақпараттарды қалай рұқсатсыз алуға болатындығына тоқталамыз.

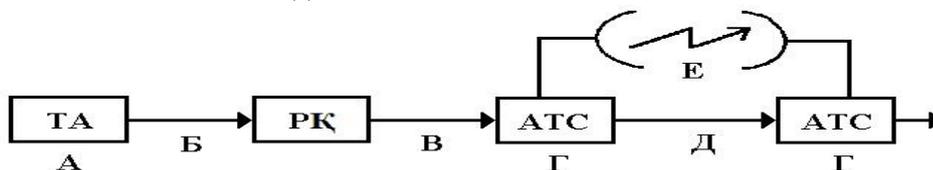
Кез –келген ақпаратты қорғау жүйесін қорғау жоспарының негізі, қауіптің қайдан келетіндігін білу және оның қауіптілік деңгейін анықтау болып табылады. Қазіргі кезде көп жағдайда ақпаратты жіберу үшін тропосфералық, космостық арналар желісі, радиожелілік, сонымен қоса кабельдік, волоконды- оптикалық желі байланыстарын қолданады. Байланысу арналарының түріне байланысты ақпаратты ұрлау техникалық арналарын электромагниттік, электрикалық және индукциялық деп бөледі.

Телефон желілері арқылы жіберілетін ауызекі ақпаратты рұқсатсыз ұрлаудың мүмкін деген жағдайларын қарастырайық. Телефондық жүйені бірнеше шартты аймақтарға бөліп қарастыруға болады. А аймағына қолданушының телефон аппаратының (ТА) өзі жатады. Сигнал аппараттан телефон сымы арқылы реттеуші қорапқа(РҚ)(Б аймағы) келіп, осы жерден магистралдық кабельге(В аймағы) барады. Сигнал автоматтандырылған телефон станциясында(АТС)(Г аймағы) коммутацияланғаннан кейін көпжелілі кабельдер(Д аймағы) арқылы не радиожелі(Е аймағы) арқылы келесі АТС дейін таралады. Бұл аталған әр аймақта ақпаратты ұрлаудың өзіндік ерекше тұстары бар, дегенмен барлығындағы да рұқсатсыз қосылу техникаларының жұмыс істеу барысы бір – бірінен өзгешеленбейді.

Тыңдаушы аппараттардың орналасуы мүмкін жағдайлары көбірек болатын А,Б және В аймақтары болып табылады. А аймағының сипаттамалары кең танымал болғандықтан, тек қалалық телефон желілерінің(ҚТЖ) сызықтық құрылымының құрамын қарастырайық. ҚТЖ құрамына абоненттік желілер, тексеруші құрылғыларынан тұратын телефондық кәрізқұбырды және терминалдык жабдықтаушы құрылғылар.

Телефондық желілер қолданушы аппараттарын қалалық АТС -ға немесе телефондық қосалқы станцияларға қосуға арналады және үш бөліктен тұрады: магистральдық(АТС -дан реттеуші шкафқа дейін(РШ)), реттеуші(РШ- тан реттеуші қорапқа РҚ дейін), абоненттік(РҚ -тан телефондық аппаратқа дейін).

Соңғы екі бөлім(реттеуші және абоненттік) салыстырмалы түрде ақпараттың ұрлануы мүмкін ең осал бөлім болып табылады.



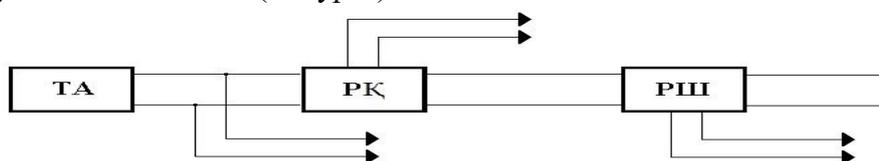
1- Сурет. Телефон желілері арналарындағы ақпаратты ұрлаудың негізгі аймақтары.



2-сурет. АТС- абонент желісінің негізгі элементтері

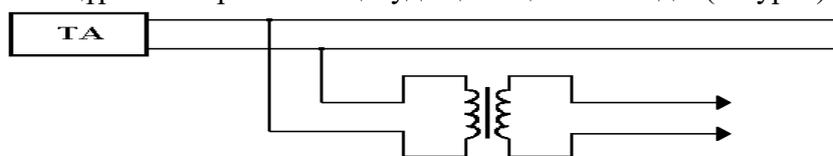
Өндірістік тыңшылықпен айналысатын мемлекеттік емес ұйымдар үшін ақпаратты ұрлау мақсатымен қосылуға ең ыңғайлы орындар “А”, “Б” және “В” аймақтары болып табылады.

Тыңдаушы құрылғының шунты “А”, “Б” және “В” аймақтарының, телефондық сымдарға немесе телефондық аппаратқа қолжетімділігі бар: телефондық ток көзі не реттеуші қорапқа дейінгі кез -келген жерінде орналасуы мүмкін. В аймағында магистраль кабелін қолдану барысында тыңдаушы құрылғының орналасу мүмкіндігі өте аз. Оның бірден -бір себебі, ол құрылғыны орнату үшін телефондық желілердің кәрізқұбырына өту қажет, яғни бір не бірнеше трубаларға біріккен және кабелдерді төсеу, жөндеу және уақытылы жөндеу жұмыстарын өткізуге арналған байқаушы құрылғылардан(құдықтар) тұратын жерасты құрылысынан тұрады. Сонымен, қандай да бір тыңшылық құрылғыны орнатпас бұрын, тек қана жерастының коммуникациялық әккі әдістермен біріктірілген құрылғыларын білу жеткіліксіз, сонымен қоса жүздеген және де соған ұқсас тағы да жүздеген кабелдер ішінен өзіне қажеттісін табуы қажет. Ал, әуелік желіні таңдаған сәтте барлығы да әлдеқайда жеңілдейді. Сондықтан, телефон арқылы арзан байланыс жолдарын, жерасты не арзан “әуелік” кабелді, таңдамас бұрын қауіпсіздік тұсын да жіті қарастырған жөн. Техникалық жағынан қарастырғанда Б және В аймақтарында заңсыз қосылудың ең жеңілі қатынасу арқылы қосылу болып табылады.(3- сурет)



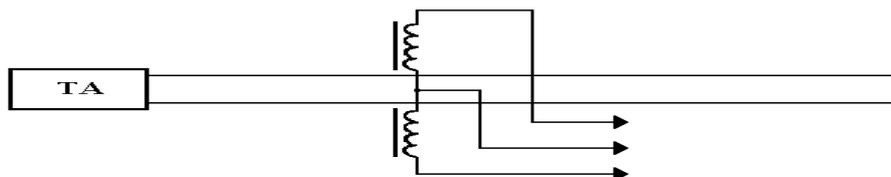
3-сурет. Телефондық желіге қатынасу арқылы қосылу

Әуелік желілерге қосылу әлдеқайда жеңілірек, әрі қауіпсіз және ол келесідей ретте жүргізіледі: бірнеше жұп, лакпен көмкерілген өте жұқа(адам шашындай) өткізгіштері арқылы телефон сымдарынан төменге қарай ағаш сырығы арқылы алдын -ала жалданып алынған бөлмеге әкеледі. Онда, ақпаратты ұрлайтын дайындалған оператор отырады. Бірақ, бұл әдістің өзіндік кемшіліктері бар. Себебі, қосымша қосылған құрылғы кернеудің өте қатты төмендеуіне әкеліп, байланысушылар арасындағы дыбыстың дұрыс естілмеуіне әкеледі. Сәйкесінше, ортадағы адамды анықтау өте жеңіл болады. Осы себепті, тиімдірек тәсіл ретінде сәйкестендіруші құрылғы арқылы қосылуы айтады. Себебі, ол ортадағы кернеудің аз ғана өзгеруіне әсер етеді, сондықтан ортада тыңдаушы бар екенін арнайы құрылғылар арқылы құрылғылармен анықтаудың өзі қиын болады.(4-сурет)



4-сурет. Телефондық желіге сәйкестендіруші құрылғы арқылы қосылу.

Қосылған құрылғы әсерінен желі байланысының сипаттамаларын өзгертуден туындайтын кемшіліктерді жою мақсатында, қатынас орнатпай байланысу әдісін қолданады. Әдетте, бұл мақсатта трансформатор ретінде жасалынған индуктивті датчиктер қолданылады.



5-сурет. Индуктивті датчик көмегімен телефон желісіне қосылу әдісі

Телефондық желілерді үшінші жақтың рұқсатсыз тыңдауынан қорғау үшін жасалатын заманауи қорғау әдістері мен құралдары, олардың іске асырылу жолдарының көптігі мен

әртүрлілігі төнетін қауіптер түрінің көптігін дәлелдейді. Кең мағынасында қарсы тұрудың 2 әдісін қарастыруға болады:

- Филтрлар және ақпараттың ұрланып жатқан бөлігін физикалық түрде анықтауға мүмкіндік беретін физикалық қорғау құралдары.
- Мағыналық ақпаратты қорғау құрылғылар(жеке жағдайларда, криптографиялық)

Кедергі келтіруші бөгеттердің көпшілігі телефондық байланыстардың өткізгіш желілері бар тұсына арналған.(«АТС – телефондық аппарат» бөлігін қорғау). Сәйкесінше, бөгет сөздік сигнал сызығынан тыс құрылады және оның номиналды деңгейінен біршама артық болады. Өте күшті бөгеттің болуы сызықтық режимнен , телефондық желіге қатынас ету арқылы және қатынамай-ақ байланысқан барлық қарапайым құрылғыларды істен шығарады.(дауыс диапазонында шу пайда болып, айтылған сөздер адам айырғысыз болады). Ал, телефондық аппараттың өзінде алдын-ала кіріс сигналына жасалынған жоғары жиілікті фильтрация жасалынуы себепті шулау жоқ.

Басқа құрылғылар, сөйлеу барысында болған сигналдарды өзгерту арқылы жұмыс жасайды. Көбіне, телефон аппараттары түрінде және оған қосымша ретінде болады. Барлық желіні өзгерген түрде(қорғалған) өткеннен кейін,адам еститін қарапайым сигнал ретінде келесі абонентке жетеді. Бұл жағдайда, хабарларды тыңдау тиімсіз болып табылады. Дегенмен, аталған тәсілде сұхбат жүргізу үшін әр қолданушы үшін арнайы аппараттарды сатып алу қажет. Қазіргі кезде мамандар телефон арқылы сұхбат жүргізуде қауіпсіздікті қамтамасыз ететін ең сенімді тәсіл ретінде, криптографиялық қорғауды айтады.

Ал, енді нақты қорғау құрылғыларының ішінен скремблер мен криптофонды қарастыруға болады.

Скремблер – арна арқылы жіберілетін сөздік хабарларды шифрлауға арналған құрылғы. Скремблеу барысында хабарды үш параметрі бойынша шифрлауға болады: амплитуда, жиілік, уақыт. Заманауи радиобайланыста кең қолданыс тапқан – сигналды уақыты, жиілігі бойынша өзгертулер. Ал, радиоарналардағы бөгеттер сөздік хабар амплитудаларының түп –тура қайта қалпына келтірілуіне кедергі келтіретіндіктен , бұл әдіс бүгінде өзіндік қолданысын таппаған.

Скремблер артықшылығы: желінің барлық аймағында қорғау жүргізіледі. Ал, кемшілігі хабар алмасу үшін қажетті барлық абоненттерді сәйкес келетін құрылғылармен қамтамасыз ету.

Аралас әдістер арқылы сигналды өзгерту бір мезетте жиілікті және уақытты өзгерту арқылы жасалады. Скремблер телефонға қосылады.(өткізгіш не Bluetooth арқылы) және өшірілген жағдайда оны анықтау мүлде мүмкін емес. Аппарат иесі оны қосқан сәтте, ол микрофон арқылы келетін барлық сигналдарды қабылдап, шифрлайды және тек осыдан кейін ғана шығысқа жібереді. Ал, хабарды декодтау кері тәртіпте іске асады. Сигнал антеннадан скремблерге жіберіледі, ал содан кейін – динамикке. Осы әдіспен, скремблер жіберілетін хабарды шифрлап, қабылданған сигналды дешифрлайды.

Криптофон – телефондық сұхбатты қорғаудың жаңа әдісі. Криптофон бұл арнайы бағдарламалық қамтама орнатылған қарапайым смартфон. Криптофонның жұмыс жасау тәртібі скремблерге ұқсас: сигнал микрофоннан цифрланады, содан кейін кодталып абонентке жіберіледі. Ал, скремблерден айырмасы – шифрлауында. Криптофон арнайы криптографиялық қорғауды қолданады. Заманауи криптофондар AES, Twofish және т.с.с шифрлау алгоритмдарын қолданады. Кемшіліктері :

- Қос хабар алмасушыларда да құрылғының болуы.
- Дауыстың кішкене уақытқа кідіріп жетуі
- Сөйлесу барысында эханың пайда болуы.

Қорытындылай келе, комплексті түрде барлық техникалық құрылғыларды қолдану арқылы толық түрде телефон желілерінде қауіпсіз байланыс орнатуға қол жеткізуге болады.

СТАНДАРТЫ В ОБЛАСТИ УПРАВЛЕНИЯ РИСКАМИ: МЕЖДУНАРОДНЫЕ СТАНДАРТЫ

(ЕНУ им. Гумилева, г. Астана, Республика Казахстан)

На сегодняшний день все лучшие компании/организации в мире признали необходимость управления рисками и осуществления результативного риск-менеджмента для эффективного достижения целей и выполнения задач компаний/организаций. Безусловным является необходимость интегрирования процесса риск-менеджмента в общее управление организацией, стратегию, планирование, менеджмент, процессы отчетности, политики, ценности и культуру организации.

Классификация стандартов

В целом классификация стандартов в области управления рисками выглядит следующим образом:

- международные;
- разработанные профессиональными объединениями и стандарты;
- разработанные национальными органами стандартизации (рисунок 1).

В ноябре 2009 года опубликован стандарт **ISO 31000 «Риск-менеджмент. Принципы и рекомендации»**, который призван помочь организациям любых размеров, эффективно управлять рисками [1].

Стандарт ISO 31000 описывает общие подходы, предоставляет принципы и руководства по систематическому, понятному и надежному менеджменту любого вида рисков в рамках любой области применения и контекста.

Одновременно был опубликован ISO Guide 73:2009 «Словарь управления рисками», как дополнение к ISO 31000, дающий определения понятий, относящихся к управлению рисками.

Другой заслуживающий внимания документ в этой серии – ISO 31010 «Риск-менеджмент. Методы риск менеджмента». Данный международный стандарт поддерживает стандарт ISO 31000 и обеспечивает руководство по выбору и применению систематических методов оценки риска. Применение ряда методов представляется в форме специальных ссылок на другие международные стандарты, где концепция и применение методов описываются более подробно.

Тогда как все организации, в той или иной мере, занимаются управлением рисков, стандарт ISO 31000 устанавливает ряд принципов, необходимых для осуществления результативного риск-менеджмента.

ISO 31000 разработан, чтобы помочь организациям:

- повышать вероятность достижения целей;
- стимулировать упреждающий менеджмент;
- осознавать потребность в идентификации и обработке рисков во всей организации;
- улучшать идентификацию возможностей и угроз;
- соответствовать соответствующим организационным, законодательным и нормативным требованиям и международным нормам;
- улучшать обязательную и добровольную отчетность;
- улучшать управление;
- улучшать доверие заинтересованных лиц;
- улучшать меры управления;
- эффективно распределять и использовать ресурсы для обработки риска;
- минимизировать убытки.

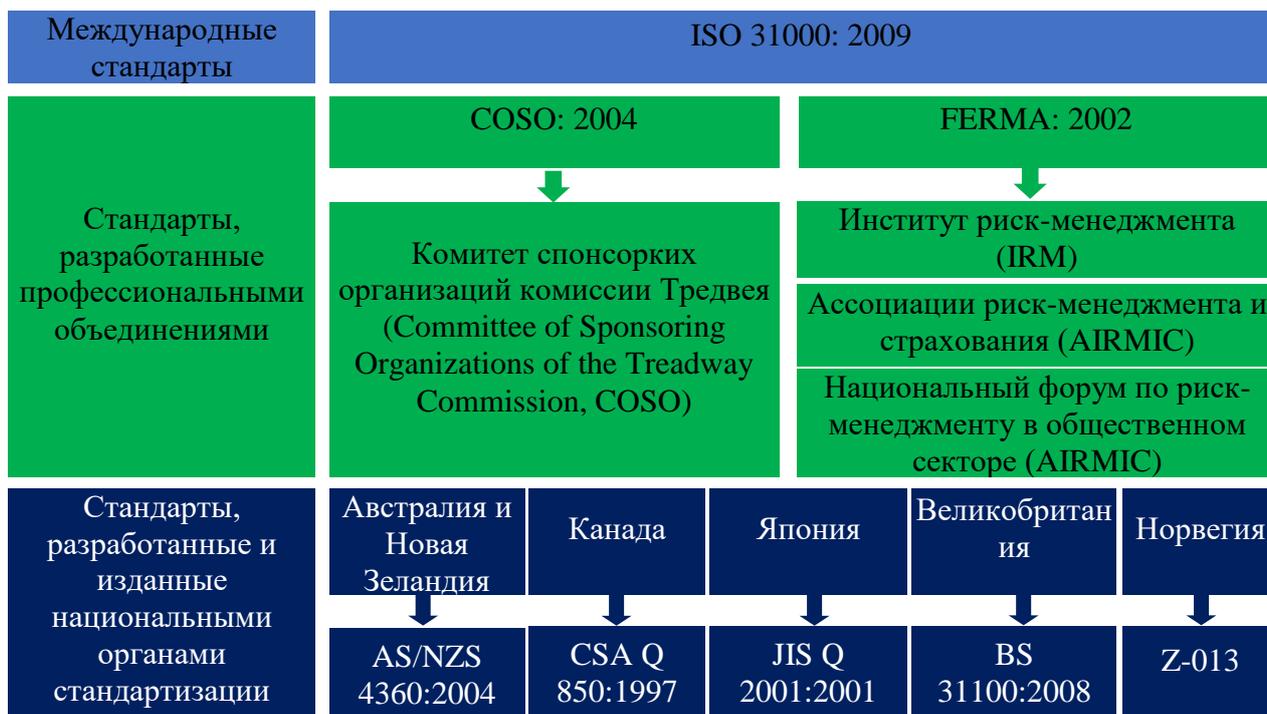


Рисунок 1. Существующие стандарты управления рисками

Согласно рассматриваемому стандарту система риск-менеджмента включает: принципы, концепцию и процесс риск-менеджмента.

Необходимо отметить, что стандарт ИСО 31000 предоставляет общий подход к любым отраслям, любым видам рисков, любым потребностям, любому контексту и любым критериям для оценки риска. Это обеспечивает необходимую результативность и эффективность, последовательность и системность, прозрачность и согласованность. В целом ИСО 31000 гибко касательно разнообразных подходов к системе риск-менеджмента, включая возможность относительно безболезненной адаптации существующих в организации наработок в данном направлении.

За основу при подготовке проекта стандарта ИСО 31000 разработчиками был принят рассмотренный выше стандарт Австралии и Новой Зеландии, о чем, в частности, свидетельствует схожесть использованного подхода к определению и описанию процесса риск-менеджмента и его отдельных составляющих. Вместе с тем существуют некоторые отличия. Например, проектом стандарта ИСО 31000 стадии «Идентификация риска», «Анализа риска», а также «Оценивание риска» рассматриваются не самостоятельно, а в качестве составляющих стадии «Оценка риска». Кроме того, отдельным образом подчеркивается необходимость документирования процесса управления риском с учетом выгод повторного использования накопленной информации для целей управления, оценки затрат на создание и хранение документов и ряда других факторов.

Вместе с тем характеристика принципов риск-менеджмента и описание модели управления риском отделены в рамках стандарта от характеристики собственно процесса риск-менеджмента. Признается, что процесс управления риском не существует сам по себе, а должен стать составным элементом управления в организации, должен внедряться в организационную культуру, настраиваться под действующие в рамках организации бизнес-процессы.

Список литературы

- 1 <http://www.iso.org/iso/home/standards/iso31000.htm>
- 2 Источник: AS/NZS 4360:2004 - Risk Management, issued by Standards Australia. P. 13
- 3 ISO/IEC 31010:2009 Risk management — Risk assessment techniques (IDT)
- 4 В. Rechel, J. Erskine, S. Wright, B. Dowdeswell, M. McKee Capital investment for health . Case studies from Europe, European observatory on Health System and Policies, p.123-139

СТАНДАРТЫ В ОБЛАСТИ УПРАВЛЕНИЯ РИСКАМИ: НАЦИОНАЛЬНЫЕ СТАНДАРТЫ

(ЕНУ им. Гумилева, г. Астана, Республика Казахстан)

В наше время необходимость интегрирования процесса риск-менеджмента в общее управление организацией, стратегию, планирование, менеджмент, процессы отчетности, политики, ценности и культуру организации является обязательным условием на пути к эффективности достижения целей и выполнения задач организаций. Лучшие организации в мире признали необходимость управления рисками и осуществления результативного риск-менеджмента.

В данной работе рассмотрены существующие национальные стандарты определения/идентификации и оценки рисков.

В целом классификация стандартов в области управления рисками выглядит следующим образом:

- международные;
- разработанные профессиональными объединениями и стандарты;
- разработанные национальными органами стандартизации (рисунок 1).

Национальные стандарты разработаны и действуют в следующих странах:

- Австралии и Новой Зеландии (AS/NZS 4360, принят в 1995 г., дорабатывался в 1999 г. и 2004 г.);
- Канаде (CAN/CSA-Q850 – 97, принят в 1997 г.);
- Японии (JIS Q 2001, принят 2001 г.),
- Великобритании в сфере управления проектными рисками (BS-6079-3:2000, принят в 2000 г.).
- Норвегии (Z-013 «Анализ готовности к риску и возникновению аварийных ситуаций»).

Одним из наиболее полных и проработанных национальных стандартов в области управления риском эксперты признают Стандарт по риск-менеджменту Австралии и Новой Зеландии. Стандарт AS/NZS 4360 имеет общий (внеотраслевой) характер, его основные положения адаптированы для построения систем управления риском рядом транснациональных компаний.

Согласно стандарту AS/NZS 4360 управление риском на уровне компании представляет собой совокупность пяти последовательно сменяющих друг друга стадий и двух процессов сквозного характера (рисунок 3) [2]. При этом под управлением риском в стандарте понимается «совокупность культуры, процессов и структур, ориентированная на использование потенциальных возможностей при одновременном управлении негативными воздействиями».

Стадия 1. Определение окружения (среды).

Среди факторов, определяющих необходимость анализа и идентификации внутренней среды компании, выделяются следующие:

- управление риском должно осуществляться в контексте определенных целей и задач организации;
- одним из основных рисков компании является возникновение препятствий в процессе достижения поставленных стратегических, операционных, проектных и прочих целей;
- четкая формулировка принципов организационной политики и целей компании будет способствовать определению основных направлений корпоративной политики в области управления рисками;

– цели и задачи компании по сегментам деятельности, а также целевые ориентиры, формируемые при реализации отдельных корпоративных проектов, должны рассматриваться в соответствии с целями компании как единого целого.

Стадия 2. Идентификация рисков.

На данной стадии должны быть идентифицированы риски, обусловленные особенностями внешней и внутренней среды, проанализированной на предыдущем этапе: рассматриваются все возможные источники риска, а также имеющаяся информация о восприятии риска (осознании риска) причастными сторонами как внутренними по отношению к организации, так и внешними.

Стадия 3. Анализ рисков.

Результатом прохождения рассматриваемой стадии является определение уровня риска, отражающего оценки последствий и вероятности рисков событий. Выделяют количественный и качественный анализ. Ценность и значение качественного анализа существенным образом повышаются в случае, если определение риска формируется широким кругом причастных сторон.

Стадия 4. Оценивание рисков.

Задачей данной стадии является принятие решения о допустимости/недопустимости риска (в отношении допустимого риска не применяются процедуры обработки риска, предусмотренные стадией 5 рассматриваемого процесса управления риском).

Оценивание риска предполагает исследование уровней подконтрольности рисков события, затрат на осуществление воздействия, потенциальных издержек и выгод, связанных с рисковым событием.

Стадия 5. Обработка риска.

На данной стадии осуществляется работа с оцененными и ранжированными рисками, в отношении которых принято решение об их неприемлемости/недопустимости для компании в соответствии с критериями, определенными на начальных стадиях рассматриваемого процесса управления рисками.

Альтернативные варианты обработки рисков:

Избежание риска, осуществляемое либо посредством прекращения деятельности, сопряженной с недопустимым для компании уровнем риска, и выбором других, более приемлемых направлений деятельности, отвечающих задачам организации, либо в ходе избрания альтернативной, менее рискованной методологии в отношении организации рассматриваемого процесса или направления деятельности;

Снижение вероятности реализации рисков события и (или) возможных последствий реализации; важно учитывать, что должен быть найден баланс между уровнем риска и издержками, сопряженными со снижением риска до заданного уровня. Когда разработанные подходы к снижению риска отнесены к категории оправданных, одновременно имея высокие издержки реализации, необходимые затраты требуют бюджетирования. Рекомендованные в рамках данной альтернативы процедуры: контроль; улучшение процессов; тренинги и повышение квалификации персонала; аудит и определение соответствия установленным правилам;

Разделение риска с третьими сторонами. Необходимо учитывать, что передающая сторона сталкивается с новым риском, связанным с неспособностью принявшей риск организации эффективно управлять им;

Удержание риска. Данная альтернатива применяется в отношении остаточных, а также невыявленных рисков.

Practice Standard for Project Risk Management (Практический стандарт по управлению рисками проектов) разработан Project Management Institute (PMI) -США в 2009 году.

Практические стандарты PMI описывают процесс, деятельность, входы и выходы определенной области знаний. Предоставляет информацию о значимых процессах, инструментах или методах и их применении.

Принципы управления рисками проекта, лежащие в основе шести главных процессов Руководства РМВОК это такие процессы, как:

- Планирование управления рисками
- Идентификация рисков
- Количественный анализ рисков
- Качественный анализ рисков
- Планирование противодействия рискам
- Мониторинг и контроль рисков.

Немецкий стандарт DGQ – 2007 – Deutsche Gesellschaft für Qualität (Немецкое Общество по Качеству), согласно которому управление рисками включает: концепцию и определение риска, номенклатуру рисков, требования к управлению рисками, процесс управления рисками, методы управления рисками, резюме.

Определяет основные 3 элемента риска: Событие, Вероятность, Количественный ущерб.

Требования к управлению рисками:

-Закон о контроле и прозрачности в корпоративном секторе от 1998 года (Руководство должно создать систему раннего выявления рисков. Риск-менеджер должен постоянно и обширно контролировать процесс и бить тревогу, если обнаружены риски);

- База II (Критерии оценки Базы II: качество управления, бизнес стратегии, человеческие ресурсы и организация, управление и контроль качества, структура управления и т.д);

- Жизнеспособность компании.

Процесс управления рисками состоит из следующих процессов:

-определение контекста

- идентификация рисков;

- анализ рисков;

- оценка рисков;

- нивелирование рисков (снижение риска, передача риска, принятие риска).

ON – 2008 – Австрия (Austrian Standards Institute) - Risikomanagement für Organisationen und Systeme – Begriffe und Grundlagen (Risk Management for Organisations and Systems) - ONR 49000

Первое издание стандарта ONR 49000 «Управление рисками для организаций и систем» было выпущено Австрийским институтом стандартов в 2004 году. Первая часть стандарта ONR 49000:2004 интегрирована со стандартом ISO 31000 «Управление рисками – принципы и Руководство к применению».

Управление рисками объединяет и координирует сверху вниз и снизу вверх. Такое Управление рисками является мощным инструментом руководству организации для того чтобы справиться с возрастающими сложностями.

Взаимодействие. Для деятельности руководителей и сотрудников по достижению общих целей, разработать Политику организации. Она включает в себя цели, стратегии, инструменты и процедуры управления ресурсами и поведение сотрудников.

Система управления организует механизмы управления и контроля. Они касаются отдельных субрегионов и/или системы управления качеством. Все аспекты управления, учитывающие различные требования должны рассматриваться как интегрированная система менеджмента.

Для оценки рисков используется внутренняя система контроля. Она состоит из систематически разработанных мер и средств контроля организации и предотвращения рисков. Задачей системы управления рисками является определение и уменьшение рисков во всех областях организации. Процесс управления рисками включает в себя мероприятия, которые направлены на организацию управления и контроля рисков. Процесс управления рисками состоит из основного процесса (определение рисков, анализ и оценка рисков, управление рисками) и мониторинга.

Ряд стран не разрабатывают принципиально новые национальные стандарты. Руководствуясь международным стандартом ИСО/МЭК 31000 Казахстан и Россия гармонизировали его и внедрили на территории своей страны в качестве руководства по управлению рисками предприятий.

В Российской Федерации в настоящее время действует стандарт **ГОСТ Р ИСО 31000–2010«Менеджмент риска, принципы и руководства»**, который идентичен международному стандарту ISO 31000: 2009.

Настоящий стандарт разработан в дополнение к ИСО 31000 и содержит рекомендации по выбору и применению методов оценки риска. Оценка риска является основным элементом процесса менеджмента риска, включающего в соответствии с ИСО 31000 следующие элементы:

- обмен информацией и консультации;
- установление области применения менеджмента риска;
- оценку риска (включая идентификацию риска, анализ риска и сравнительную оценку риска);
- обработку риска;
- мониторинг и анализ риска.

Оценка риска — процесс, объединяющий идентификацию, анализ и сравнительную оценку риска.

Риск может быть оценен для всей организации, ее подразделений, отдельных проектов, деятельности или конкретного опасного события. Поэтому в различных ситуациях могут быть применены различные методы оценки риска.

Обработка риска

После завершения оценки риска принимают и выполняют одно или несколько решений об обработке риска, позволяющих изменить вероятность возникновения опасного события и/или его воздействие.

Обработка риска обычно является адаптивным процессом проверки риска на его приемлемость и соответствие ранее установленным критериям для определения необходимости дальнейшей обработки риска.

Мониторинг и анализ

Мониторинг и анализ риска являются составной частью процесса менеджмента риска. Регулярное проведение мониторинга, анализа и управления риском направлены на проверку:

- достоверности предположений о риске;
- достоверности предположений, на которых основана оценка риска, включая внешние и внутренние области применения;
- достижимости ожидаемых результатов;
- соответствия результатов оценки риска фактической информации о риске;
- правильности применения методов оценки риска;
- эффективности обработки риска.

Процессы мониторинга и анализа риска должны быть документированы, а результаты мониторинга и анализа риска зафиксированы в отчете.

На территории Республики Казахстан с 1 июля 2011 года действуют стандарт СТ РК ИСО 31000-2010 Менеджмент риска. Принципы и руководящие указания, СТ РК ИСО/МЭК 31010-2010 Менеджмент риска. Методы оценки риска, СТ РК ИСО 73-2010 Менеджмент риска. Словарь.

Список литературы

- 1 <http://www.iso.org/iso/home/standards/iso31000.htm>
- 2 Источник: AS/NZS 4360:2004 - Risk Management, issued by Standards Australia. P. 13
- 3 ISO/IEC 31010:2009 Risk management — Risk assessment techniques (IDT)
- 4 В. Rechel, J. Erskine, S. Wright, B. Dowdeswell, M. McKee Capital investment for health . Case studies from Europe, European observatory on Health System and Policies, p.123-139
- 5 Журнал «GlobalTenant», июль 2012 года
- 6 <http://adilet.zan.kz/rus/docs/P1400000728>

А.С.ӨТЕБАЕВА, Б.Ш.РАЗАХОВА

МОДЕЛЬ ИНТЕГРИРОВАННОЙ СИСТЕМЫ БЕЗОПАСНОСТИ АЗС

(Евразийский национальный университет имени Л.Н.Гумилева, Казахстан г.Астана)

Современные АЗС предлагают широкий перечень услуг. В связи с данным появляется необходимость применения комплексного подхода к решению вопросов безопасности. АЗС является объектом повышенной опасности, в связи с этим, к безопасности АЗС предъявляются повышенные требования. В целях обеспечения безопасности на АЗС целенаправленно использовать инженерно технические средства охраны, а для повышения качества безопасности и контроля выполнения мер безопасности как сотрудников, так и клиентов АЗС - целенаправленно использовать комплексные(интегрированные) системы безопасности.

Интегрированной системой безопасности (ИСБ) называют совокупность технических средств из специфических систем безопасности, которые реализованы на единой аппаратной или программной платформе. Данная система обеспечивает выполнение заранее определенных алгоритмов систем безопасности в автоматическом режиме. Автоматизация работы оператора снижает риск принятия неверных решений, а также ускоряет реагирование сотрудников при возникновении форс-мажорной ситуации на охраняемом объекте.

Из этого определения видно, что интегрированная система безопасности вполне определенный и конкретный комплекс программ и аппаратного обеспечения. Интегрированная система безопасности предназначена для решения утилитарных задач, а основная ее цель - это повышение безопасности на охраняемом объекте.

Целями интегрирования являются:

- снижение рисков принятия ошибочных решений и уменьшение времени реакции при появлении внештатной ситуации на объекте;
- получение новых функций, связанных с возможностью обеспечения оперативного взаимодействия подсистем и компонентов СБ при сохранении в полном объеме возможностей ее составных частей;
- экономия необходимых для реализации данных функций средств;
- максимальная автоматизация действий по всем направлениям защиты объекта.

Среди функций, обязательных для исполнения в рамках ИСБ, следует считать следующие :

- контроль состояния объекта с созданием многорубежной защиты;
- иерархический доступ персонала и посетителей в помещения с четким разграничением полномочий по праву доступа в зависимости от времени суток и дней недели;
- идентификацию и аутентификацию личности человека, пересекающего рубеж контроля;
- предупреждение утечки информации;
- предупреждение попадания на объект запрещенных материалов и оборудования;
- документирование и архивирование информации для ее использования при расследовании происшествий и анализе действий служб охраны;

–оперативный инструктаж службы охраны о порядке действий в различных штатных и нештатных ситуациях путем автоматического вывода на экран монитора инструкций в нужный момент;

–обеспечение полной интеграции систем видеонаблюдения, сигнализации, мониторинга доступа, оповещения;

–связи между персоналами службы охраны, службы пожарной безопасности, служб жизнеобеспечения объекта и т.д.;

–обеспечение взаимодействия постов охраны и органов правопорядка;

–контроль исполнения персоналом охраны своих служебных обязанностей.[1]

Учитывая особенности организации интегрированных систем безопасности, а также специфику рассматриваемого объекта, в качестве комплексного решения вопросов безопасности на АЗС предлагается использовать следующую модель.(Рисунок 1)



Рисунок 1

Используем модульный принцип построения системы безопасности на АЗС,при котором каждый ее компонент сам является полнофункциональной системой и способен работать независимо от других. В состав охранного комплекса входит следующее: телевизионная система охраны и наблюдения, система охранной и тревожной сигнализации, система пожарной сигнализации, система контроля и управления доступом, электронная система защиты от краж, автоматическая установка пожаротушения, система оповещения и управления эвакуацией, а также (при необходимости) оборудование, имеющее открытые способы интеграции со специализированным оборудованием, применяемым на АЗС.

Список литературы

1.Рыжова В.А. Проектирование и исследование комплексных систем безопасности. – СПб: НИУ ИТМО, 2013. – 156 с.

2.ГОСТ Р 53704-2009 Системы безопасности комплексные и интегрированные. Общие технические требования. – Москва: Стандартинформ, 2010.

КОМЕРЦИЯЛЫҚ КОМПАНИЯЛАРДЫҢ КИБЕРҚАУІПСІЗДІГІ

(Алматы технологиялық университет, Алматы қаласы)

Жаңа ақпараттық телекоммуникациялық технологиялардың арқасында ақпараттық қоғам қарқынды дамып келеді. Адамзат қызметінің барлық салаларында электронды түрде сақталған ақпарат көлемі мың еселеп өсті. Компьютерлік желілердің пайда болуымен қатар ақпарат қауіпсіздігіне сенім азайды. Бізді қоршаған ақпараттық орта әр күн өткен сайын қауіптірек болып, компаниялардың ақпараттық жүйелеріне төнетін қатерлер коммерциялық сипат алуда. Кез келген комерциялық бизнестегі компания басшысы, немесе сенімді адамы өздерінің ақпараттық объектілерін ұйымдастырады. Оны әдейі бұзу ақпаратқа шабуыл болып табылады.

Ақпаратқа шабуылдың зардаптары:

1. Комерциялық ақпаратты бұзу нарықта үлкен шығындарға әкеліп соқтырады;
2. Ақпаратты үлкен көлемде ұрлау фирма репутациясына зардабын тигізеді;
3. Бәсекелес фирмалар ұрланған ақпаратты қолданып, басқа фирманы банкротқа ұшыратады;
4. Ақпараттық қызмет көрсететін фирмаға көп шабуыл жасау фирма клиенттерінің сенімділігін азайтады, сол арақылы оның кіріс көлемі азаяды.

Компьютерлер мен желілер арқылы ақпаратқа шабуылдар компанияларға өте үлкен зардабын тигізуде. Осыған байланысты комерциялық компаниялардың ақпараттық қауіпсіздігін қарастырайық.

Ақпараттық қауіпсіздік ақпарат конфиденциалдылығын, ақпарат тұтастығын, ақпаратқа қол жеткізу мүмкіндігін қаматамасыз ететін процес

Құпиялылығы – белгілі бір ақпараттың белгілі бір адамдарға ғана тиесілі екеніне сенімділік.

Тұтастығы – ақпараттың сақталуы немесе таралуы кезінде санкцияланбаған өзгерістердің болмағанына сенімділік.

Ақпаратқа қол жеткізу- әртүрлі адамдардың объектілерге әртүрлі жету жолдарының болуына сенімділік.

Техникада ақпарат қауіпсіздігін сақтай отырып ақпаратты тарату соңғы 40 жылда екі кезеңнен өтті.

1-кезең. Ертеректе ақпаратты шифрлап, құпия түрде таратуды барлаушылар, саясаткерлер, дипломаттар қолданса, қазіргі кезде интернеттің пайда болуына байланыты шифрлау арқылы конфиденциалды ақпаратты тарататын адамдар саны миллиондаған есеге артты.

2-кезең. Әлемге белгілі Австралиялық интернет-журналист, WikiLeaks негізін салушы, ұлы державалардың дипломтиялық үлкен көлемдегі шпиондық құпия материалдары мен әскери қылмыстарын халыққа жариялаушы, 1971 жылы туылған Джулиан Ассанж бен әлемге әйгілі ұлттық қауіпсіздік агенттігінің қызметкері болған, 1,7 млн құпия файлдарды құрықтаған, Америкалық техникалық маман, 1983 жылы туылған Эдвард Джозеф Сноуденнің іс-әрекетіне байланысты болды. Олардың пайымдауынша заманауи техникалық құралдар бір жағынан өте үлкен көлемдегі конфиденциалды ақпараттарды сақтаса, екінші жағынан осы техникалық құралдар көмегімен көптеген мемлекеттердің құпия қызмет (секретная служба) көрсету орталықтары барлық адамдардың телефон арқылы, мобильді телефон арқылы тілдесулерін құпия жағдайда тыңдайды және миллиондаған адамдардың, сондай-ақ әлемдік деңгейдегі саясаткерлердің СМС хабарламалары мен электронды пошталарындағы хаттарын оқи алатындығын Джулиан Ассанж бен Эдвард Сноуден өз қызметтерінде көз жеткізді.

Сноуденнің айтуынша ғаламдық сипатта жасырын тілдесетін кеңістігі жоқ шындық әлемінде өмір сүріп отырғанын әрбір адам сезіне білуі керек.

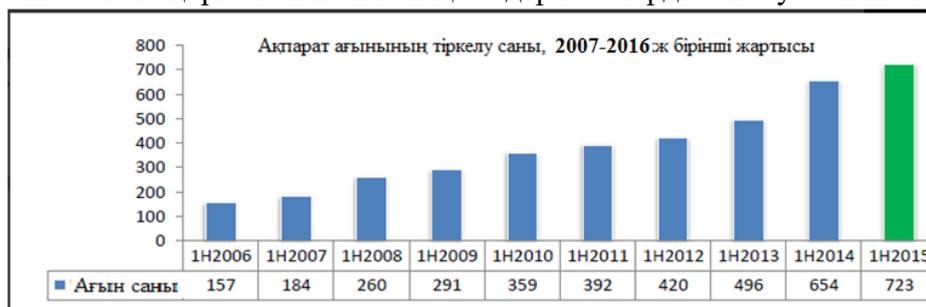
Киберқылмыс қазіргі таңда эпидемия сияқты кең таралып, шағын және орта бизнес компаниялары хакерлердің негізгі нысанасына айналып отыр. Зиянкестер оларды өздерінің желілерін кеңейту мақсатында пайдалануда, сондықтан осы тектес шабуылдардан нарықтың бір де бір секторы сақтанбаған. Киберқауіпсіздік ірі компаниялар үшін қаншалықты қажет болса, шағын компанияларға да қажет болып отыр.

Жақында «Wall Street» журналы шағын компаниялардың кибершабуылдардан көп жағдайда толық арыла алмайтынын атап өткен болатын. Алайда бірнеше қадам жасау арқылы да өз бизнесіңізді қауіпсіздендіруге болады. Ол үшін бизнес ақпараттарын қорғау жүйелерін қарастырамыз.



1-сурет. Бизнес ақпараттарын қорғау жүйелері

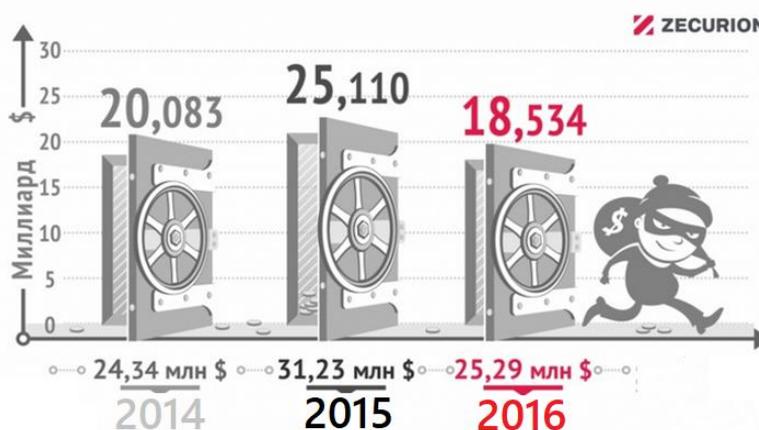
Қазіргі «Ақпараттық технологиялар» жүйесі біршама күрделі. Ол жүйелік администратор арқылы бірнеше күрделі командалармен басқарылады. Сондықтан, ақпараттық жүйедегі жүйелік администратордың рөлі өте зор. Оны сол ақпараттық жүйенің патшасы десек те болады. Кейде барлық адамдар сияқты оның да қателесуі мүмкін. Егер ол өз жұмысын дұрыс атқармаса компания үшін өте қауіпті. Оның жұмысының дұрыс-дұрыс емес екеніне компания басшысы баға бере алмайды. Сондықтан оған көз жеткізу үшін IT-департаментінен мамандарды шақыртып, Аудит жасату керек. Олар Wallix AdminBastion көмегімен тексеріс жүргізеді. Егер администратор көп болса, жұмыс күрделеніп, бірнеше топ параллель жұмыс жасайды. Бұл жүйелік администратор үшін де пайдалы. Өзінің параллель әрекеті арқылы жұмысына талдау жасай алады. Мысалы, манипуляция кезінде қызметкерлер туралы мәліметтер бірін бірі жауып қалмау керек. Олай болған жағдайда АЖ апатты жағдайға түседі. Бұдан кейін күрделі ақпараттық жүйені қалпына келтіру қиынға соғады. Сонымен қатар, жүйелік администратордың басты мақсаты ақпараттардың сыртқа шықпай ұрланбауын қамтамасыз етеді. Ақпаратқа тек рұқсат бойынша кіруді ұйымдастырады. Ақпараттардың қорғалу жүйесі әлсіз болған жағдайда хакерлер мен бұзық ойлылардың іс-әрекеті оңай орындалып, ақпарат ағыны жүзеге асады. Хакерлердің шабуылдары мен Locky, KeRanger, Cryptolocker, CryptoWall и TeslaCrypt сияқты бопсалаушы программаларымен күресудің ең тиімді әдісі тұйық программалау ортасын құру үшін виртуальданған, оқшауланған және қорғалған пошталық бағдарламаларды енгізу болып табылады.



2-сурет. Әлемдік ақпарат ағынының көрсеткіші

Хакерлерді ешқандай заң бағындыра алмайды. Олардың жасаған ақпараттық шабуылдарынан 2016 жылдың бірінші жартысында дүние жүзі бойынша құпия ақпараттар ағыны, яғни ұрлануы 10% ға өсті. Ал Ресей желі арқылы құпия ақпараттардың ұрлану жағынан екінші орынға шықты. InfoWatch компаниясының әлемдік зерттеуі бойынша 2015 жылы 723 ақпаратты ұрлау оқиғасы тіркелген. Бұл 2014 жылмен салыстырғанда 10% көтерілген. InfoWatch зерттеуі бойынша әлемдік рейтингте ақпарат ағынының саны жағынан, Ресей 2013 жылдан бастап екінші орында (2015 жарты жылдығында 59 ақпарат ағыны). Бірінші орын әдеттегідей АҚШ еншісінде (430 немесе 59%). Ал үшінші орында — Канада (39 ақпарат ағыны)

Қауіпсіздіктің үлкен көлемдегі дербес ақпарат ағыны ішкі шабуылдарға байланысты. Gemalto зерттеуі бойынша дербес және сәйкестендіру мәліметтерінің ұрлануы 2016 жылы 53% болған. Бұл дербес ақпараттардың да ұрланғанын көрсетіп отыр. Жалпы әлем бойынша 888 дербес ақпарат ұрланған.



3-сурет. Ақпарат ағынының орташа залал келтіру шығыны

Хакерлер Банк желісіндегі ұрланған ақпараттардан құпия мәліметтер алып қана қоймай, сол арқылы ақша табады. Zecurion Analytics зерттеуінде 2015 жылы дүние жүзі бойынша бір ғана ақпарат ағынының орташа залал келтіру шығыны 25 миллион долларды құрағаны көрсетілген. Ал Ресей компанияларында максимум 30 миллион долларға жеткен.

Хакерлер шабуылынан қорғанудың алғы шарттары:

- Заңды тұлға уақтылы программалық жабдықтарын жаңалап, трафиктерді филтрлеп, құпия ақпараттарды өңдеп, қызметкерлердің жұмыс регламентін реттеу керек. Ақпарат ағынының жартысынан көбі компания қызметкерлерінің кінасынан болатыны белгілі. Сол үшін олардың ақпараттық білім деңгейін үнемі жетілдіріп отыру қажет.

- Қазіргі кезде негізгі ақпарат алмасу каналдарын бақылау үшін DLP (Data Loss/Leakage Prevention) класындағы жүйені қолданған тиімді. Ал ақпараттық ресурстарға қол жеткізу құқығын басқару IRM (Information Right Management) класындағы қосымшаларды қолданған сенімдірек.

Қорыта келгенде, жоғарыда зерттелген әдістер бойынша бірнеше қадам жасау арқылы да өз бизнесіңізді қауіпсіздендіруге болады.

Әдебиеттер

1 Симионс Г.Дж. Обзор методов аутентификации информации // ТИИЭР, 1998. – Т.6. – С. 105-125.

2 Грушо А.А., Тимонина Е.Е. Теоретические основы защиты информации. – М.: Изд. агентства «Яхтсмен», 1996. – 71 с.

3 Пшенин Е.С. Теоретические основы защиты информации. Учебное пособие. – Алматы: КазНТУ, 2000. – 125 с.

Л.М.ШАҒИ, Б.К.АБДУРАИМОВА

КРИПТОСИСТЕМА НА ОСНОВЕ КВАНТОВОЙ ЗАПУТАННОСТИ

(Евразийский национальный университет им. Л.Н.Гумилева, г. Астана, Казахстан)

Целью исследования является создание криптосистемы на основе квантовой запутанности, которая будет иметь абсолютную криптостойкость, то есть не подвергаемая методу полного перебора.

Современные алгоритмы шифрования можно взломать, но их защищенность обусловлена тем, что взлом займет настолько много времени, что в этом не будет смысла. А что если вовсе откинуть возможность перебора всех вариантов. Тогда предполагаемый алгоритм будет иметь высший уровень криптостойкости. Если такой алгоритм можно будет создать, то он провернёт революцию в криптографии. Если же такой алгоритм пока что невозможно создать, то это тоже ожидаемый результат данного исследования. К примеру недавно были совершены успешные атаки на алгоритм SHA-1. Также актуальность исследования основывается на угрозе квантовых вычислений, которые безусловно, рано или поздно, сделают современные алгоритмы неэффективными.

В поисках материала по реализации подобной системы, мы не нашли исчерпывающей информации. В связи с чем мы взялся за самостоятельный поиск, и начали исследование. Статей со схожей темой мы не встретили, и считаем, что внесём свой вклад в развитие криптографии и информационной безопасности, в частности в развитие квантовой криптографии.

В криптографии на вычислительной сложности полного перебора основывается оценка криптостойкости шифров. В частности, шифр считается криптостойким, если не существует метода «взлома» существенно более быстрого чем полный перебор всех ключей.

Криптографические атаки, основанные на методе полного перебора, являются самыми универсальными, но и самыми долгими.

Абсолютная криптостойкость определяется следующими требованиями к системе. Требуется, чтобы апостериорные вероятности различных сообщений, полученные после перехвата противником данной криптограммы, были бы в точности равны априорным вероятностям тех же сообщений до перехвата[4].

Полный перебор — метод решения математических задач. Относится к классу методов поиска решения исчерпыванием всевозможных вариантов. Сложность полного перебора зависит от количества всех возможных решений задачи. Если пространство решений очень велико, то полный перебор может не дать результатов в течение нескольких лет или даже столетий.

Квантовая запутанность — квантовомеханическое явление, при котором квантовые состояния двух или большего числа объектов оказываются взаимозависимыми (например, можно получить пару фотонов, находящихся в запутанном состоянии, и тогда если при измерении спина первой частицы спиральность оказывается положительной, то спиральность второй всегда оказывается отрицательной, и наоборот). Такая взаимозависимость сохраняется, даже если эти объекты разнесены в пространстве за пределы любых известных взаимодействий, что находится в логическом противоречии с принципом локальности[1].

Целью информационной безопасности является обеспечение трех наиболее важных сервисов безопасности: конфиденциальность, целостность и доступность. Первые два не имеют прямого отношения к нашему исследованию. А вот доступность играет ключевую роль в построении предполагаемой криптосистемы. Дело в том, теоретически можно создать

абсолютно криптостойкий алгоритм, если не брать в расчёт сервис доступности. То есть зашифровать информацию так, чтобы невозможно было её расшифровать вовсе, даже методом полного перебора. Например наложить абсолютно случайную последовательность(шум) на исходную.

Таблица 1.Таблица кодировки

A	00000	G	00111	M	01101	S	10011	Y	11001
B	00001	H	01000	N	01110	T	10100	Z	11011
C	00011	I	01001	O	01111	U	10101	.	11100
D	00100	J	01010	P	10000	V	10110	?	11101
E	00101	K	01011	Q	10001	W	10111	_	11110
F	00110	L	01100	R	10010	X	11000	,	11111

Выше представлена таблица кодировки. Зашируем сообщение "WHY_ME?". Переведём в кодировку: 10111010001100111110011010010111101. А теперь наложим случайную последовательность (Рис. 1):

$$\oplus \begin{array}{r} 10111010001100111110011010010111101 \\ 10101110110010111101001010111010110 \\ \hline 00010100111110000011010000101101011 \end{array}$$

Рисунок 1 - Наложение случайной последовательности

Зашифрованное сообщение потеряло свойство доступности, если нет возможности воссоздать ключ, следовательно потеряло исходную информацию. Методом полного перебора из этого сообщения можно извлечь любое сообщение этой длины. А если доступность присутствует, то предоставив возможность хоть один раз дешифровать шифртекст, мы уже даём возможность для перебора всех вариантов ключей.

Наложение шума, или шифр Вернама, действительно абсолютно криптостойкое, как доказал К.Э.Шеннон, но при условии истинной случайности шума. Он доказал, что шифртекст не содержит информацию о исходном сообщении, а только лишь информацию о случайной последовательности. Для работы шифра Вернама необходима истинно случайная последовательность (ключ). По определению, последовательность, полученная с использованием любого алгоритма, является не истинно случайной, а псевдослучайной. То есть, нужно получить случайную последовательность не алгоритмически. Остаётся найти способ создания двух синхронизированных источников случайных чисел. Именно случайных, а не псевдослучайных. И в этом нам поможет явление квантовой запутанности. Ещё сам А.Эйнштейн был противником этой теории, но это явление экспериментально доказано. Теоретически можно воспользоваться свойством запутанных частиц, то есть можно производить замеры одной частицы, и инвертировать результаты замеров второй, что даёт нам два синхронизированных генератора случайных чисел. И вероятности того, что частица окажется в одном из двух состояний равновероятно. Причём эти пары запутанных частиц могут находиться в разных частях вселенной. Далее необходимо вести запись результатов. В итоге отправляется зашифрованное сообщение и в виде ключа отправляется порядковый номер бита с которого начинается последовательность ключа для дешифрования в генерируемой последовательности. А длина ключа для повторного сложения равна длине сообщения. Злоумышленник не имея доступа к сгенерированной последовательности не имеет возможность расшифровать сообщение.

Остаётся ещё одна задача, связанная с доставкой запутанных частиц. Частица передаётся получателю по квантовому каналу. Примером запутанной пары можно взять бифотоны[4]. После через классический канал информации мы передаём шифр. При перехвате информации любого из каналов, попытки расшифровать будут безуспешны из-за абсолютной стойкости шифра Вернама для классического канала и теоремы о запрете

клонирования для квантового канала. То есть мы узнаем, что сообщение было перехвачено, помимо этого сообщение без ключа расшифровать невозможно.

В заключении всё опирается на доступность. Используя явление квантовой запутанности и шифр Вернема можно теоритический посторить криптосистему с иммунитетом к методу полного перебора, которая еще даёт информацию о перехвате сообщения. Остаётся обеспечить защиту от других возможных видов криптоаналитических атак. На сегодня не решенной технологической проблемой является доставка оптического сигнала на большие расстояния.

Список литературы:

1 Нильсен М., Чанг И. Квантовые вычисления и квантовая информация = Quantum Computation and Quantum Information. — М.: Мир, 2006. — ISBN 5-03-003524-9.

2 Paar, Christof et al.,. Understanding Cryptography: A Textbook for Students and Practitioners. — Springer, 2010. — P. 1292. — ISBN 3-642-04100-0.

3 Shannon C. E. Communication Theory of Secrecy Systems/ C. E. Shannon //Bell System Technical Journa. October 1949. Vol. 28. Issue 4. P. 656-715.

4 Бурлаков А. В., Клышко Д. Н. Поляризованные бифотоны как "оптические кварки" // Письма в ЖЭТФ : журнал. — 1999. — Т. 69, вып. 11.

W.K. Wootters and W.H. Zurek, A Single Quantum Cannot be Cloned, Nature 299 (1982), pp. 802-803.

УДК 004.056.53

А.А. ШАРИПБАЙ, А.Б. БАРЛЫБАЕВ

ПОЛИТИКА ИНФОРМАЦИОННОЙ БЕЗОПАСНОСТИ СМАРТ-УНИВЕРСИТЕТА

*(НИИ «Искусственный интеллект» Евразийский национальный университет
им. Л.Н.Гумилева, Астана, Казахстан)*

Основной характерной чертой современного мира является переход к информационному обществу, где информация считается первичной субстанцией всего мировоззрения. Главным механизмом формирования информационного общества является информатизация, которая предполагает создание индустрии производства и обработки информационных ресурсов на основе эффективного использования информационных технологий во всех сферах жизнедеятельности общества. С одной стороны, информатизация выступает одной из основ экономического развития общества, преобразует устои жизни общества, ускоряет общественное развитие и дает мощный импульс структурной перестройке экономики, появлению наукоемких отраслей и высоких технологий, развитию сферы услуг, реорганизации традиционных базовых отраслей. С другой стороны, информатизация становится одной из наиболее мощных и прогрессирующих отраслей мировой экономики и открывает совершенно новые возможности для создания нематериальных форм богатства и оказания информационных, образовательных, научно-исследовательских и других социальных услуг. Обработка информационных ресурсов, наряду с материальными, энергетическими и трудовыми ресурсами включаются в общественное производство и становятся компонентами любого производственного, экономического и социального процесса. Поэтому информационная безопасность в информационных системах (системах обработки информации) является очень важной и сложной проблемой.

Интеллектуальный электронный университет (Смарт-университет) является сложной, большой и специальной информационной системой, разработанной с применением объектно-ориентированных технологий (программирования и базы данных), методов

искусственного интеллекта (онтологического моделирования знаний, обработки естественных языков) и нечеткой логики (нечетких описаний и логических выводов).

Интеллектуальный электронный университет (далее – ИЭУ) состоит из:

подсистемы «Электронная библиотека» позволяет создавать, хранить и управлять интеллектуальными электронными учебными изданиями;

подсистемы «Учебный процесс» позволяет организовать и управлять учебным процессом по всем уровням (бакалавриат, магистратура, докторантура) высшего образования.

В подсистеме «Электронная библиотека» имеется Генератор, который позволяет специалистам по предметным областям создавать Смарт-тьюторы в соответствии учебной программой (Силлабусом) выбранной дисциплины из входящих в учебные планы конкретных образовательных программ определенного уровня образования и должен удовлетворять требованиям государственного стандарта.

Смарт-тьюторы без вмешательства человека позволяет пользователю осуществить следующие задачи:

обучиться учебному материалу (получить свободный доступ к учебному материалу в удобной для своего восприятия форме в виде текстовой, графической, анимационной, аудио или видео информации);

консультироваться (получить смысловые ответы на свои заданные вопросы) по учебному материалу;

осуществить самоконтроль (дать возможность сравнить ответы пользователя и эталонных ответов на контрольные вопросы) по учебному материалу;

оценить свои знания (получить обоснованные оценки на свои ответы на тестовые вопросы) по учебному материалу.

Бесперебойное функционирование Смарт-университета достигается на основе обеспечения информационной безопасности всех процессов, происходящих в нём, и идентификацией их пользователей, которое включает в себя любую деятельность, направленную на защиту информации и/или поддерживающей инфраструктуры. Неотъемлемой частью организации защиты информации является непрерывный контроль эффективности предпринимаемых мер, определение для пользователей Смарт-университета перечня недопустимых действий, возможных последствий и ответственности. Все это будет описано в политике информационной безопасности Смарт-университета, являющейся документом первого уровня, ее должны дополнять другие документы (положения и инструкции), которые будут регламентировать исполнение ее положений.

Политика информационной безопасности является основой для обеспечения режима информационной безопасности всех процессов в Смарт-университета, служит руководством при разработке соответствующих Положений, Правил, Инструкций. Её положения относятся ко всем пользователям Смарт-университета, имеющим доступ к его подсистемам и данным.

Нормативно-правовую основу Политики информационной безопасности составляет законодательство Республики Казахстан по вопросам использования информационных систем, нормативные правовые акты Республики Казахстан, требования международных стандартов управления информационной безопасностью.

В качестве архитектуры Смарт-университета была выбрана трехуровневая модель.

Трёхуровневая архитектура – архитектурная модель программного комплекса, предполагающая наличие в нём трёх компонентов: клиентского приложения (обычно называемого «тонким клиентом» или терминалом), сервера приложений, к которому подключено клиентское приложение, и сервера базы данных, с которым работает сервер приложений.

На рисунке 1 описана трехуровневая архитектура смарт-университета.

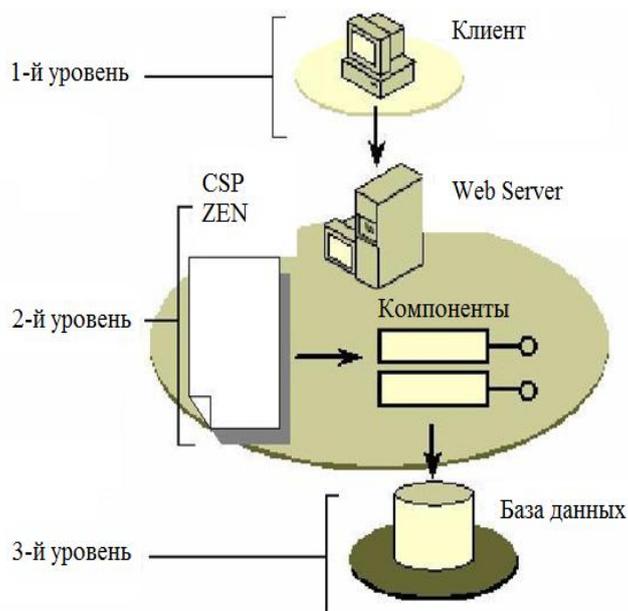


Рисунок 1. Трехуровневая архитектура смарт-университета

В Смарт-университете для личных данных пользователей (дата рождения, индивидуальный идентификационный номер, транскрипт и др.) должна быть организована информационная безопасность.

Политика информационной безопасности организации - совокупность руководящих принципов, правил, процедур и практических приёмов в области безопасности, которые регулируют управление, защиту и распределение конфиденциальной информации.

Рассмотрим характеристики, далее построим архитектуру описывающую целостность модели политики информационной безопасности. Модель политики информационной безопасности Смарт-университета изображена в таблице 1.

Реализация Политики информационной безопасности Смарт-университета должна исходить из предпосылки, что его отдельные подсистемы должны рассматриваться как часть единой информационной системы в защищенном исполнении при оптимальном соотношении технических (аппаратных, программных) средств и организационных мероприятий. Разработать инструкцию по предоставлению доступа к сети интернет, инструкцию по предоставлению доступа в помещения с ограниченным доступом, инструкции по работе с подсистемами Смарт-университета и т.п.). Не реже, чем раз в квартал пересматривать Политику и прочие документы по информационной безопасности с целью их актуализации.

Таблица 1. Модель политики информационной безопасности Смарт-университета.

Настройки безопасности времени		Параметры безопасности после установки		
Настройки безопасности ОС		Управление пользователями	Политика паролей	
C S P	Частный, закодированный CSP		Z E N	
				Уровень приложения
				Уровень страницы
%Сессия		Уровень компонента		
Отключение		Ведение журнала		
Безопасность обновлений				
Шифрование данных	баз	CSP шлюз	Резервное копирование	

МЕТОДЫ ОЦЕНКИ РИСКОВ ИНФОРМАЦИОННОЙ БЕЗОПАСНОСТИ В КОРПОРАТИВНЫХ СЕТЯХ

В настоящее время вопросы информационной безопасности не теряют своей актуальности, количество кибер-преступлений не снижается.

По данным аналитического портала в области информационной безопасности InfoWatch: за I полугодие 2016 года Аналитическим центром InfoWatch зарегистрировано 840 случаев утечки конфиденциальной информации. Это на 16% больше, чем за аналогичный период 2015 года.

Зарегистрировано 506 (67%) утечек информации, причиной которых стал внутренний нарушитель. В 250 (33%) случаях утечка информации произошла из-за внешнего воздействия.

Примечательно, что среднее количество записей, которые скомпрометированы в ходе одной утечки, мало отличается от отрасли к отрасли.

В организациях среднего размера зафиксировано существенно больше утечек, чем в крупных компаниях. В ряде случаев совокупный объем скомпрометированных записей в средних компаниях равен совокупному объему скомпрометированных записей в крупных компаниях в пределах одной отрасли.

Совокупная доля утечек персональных и платежных данных¹² выросла на 4 п. п. и составила 94%. При этом на персональные данные пришлось 88% утечек. В 6% случаев утекала платежная информация. Как правило, при утечке платежной информации речь идет о компрометации реквизитов платежных карт. По сравнению с I полугодием 2015 года, наблюдается незначительное (на 1 п. п.) уменьшение долей утечек информации, составляющей государственную и коммерческую тайну.

В 2016-м году доля утечек данных, сопряженных с последующим использованием скомпрометированной информации в целях мошенничества (банковский фрод) снизилась и составила 8%. Доля утечек данных, сопряженных с неправомерным доступом к информации (злоупотребление правами доступа, внутренний шпионаж), составила 11%. 81% инцидентов, сопряженных с потерей контроля над информацией, относится к типу «классических» утечек, не сопряженных с дополнительными нарушениями.

Во многом неожиданный трехкратный рост объема скомпрометированных данных свидетельствует о растущей день ото дня ценности данных в цифровом виде.

Злоумышленники поняли это даже раньше, чем владельцы информации, которые до сих пор не всегда готовы оценить в деньгах свои информационные активы. Между тем, очевидно, что дальнейшее развитие подходов к обеспечению информационной безопасности данных неизбежно потребует оценки стоимости активов, ясного представления, прежде всего, от владельцев информации, относительно того, какие данные для них важны, каковы финансовые потери в случае утечки этих данных.

Оценка стоимости активов и потенциального ущерба является вопросом аудита информационной безопасности организации.

В настоящие дни вопросы аудита информационной безопасности корпоративных систем предприятия приобретают новую актуальность. Представляют собой динамично развивающееся направление менеджмента в области информационной безопасности.

Основной задачей аудита является объективная оценка текущего состояния информационной безопасности объекта, эффективности существующей системы информационной безопасности и экономической целесообразности системы.

Аудитом информационной безопасности является систематизированный процесс оценки качественных и количественных существующей системы информационной безопасности и соответствие ее определенным критериям и показателям безопасности.

На основе результатов аудита информационной безопасности возможно построение оптимальной по эффективности и затратам корпоративной системы защиты.

В настоящее время разработано большое количество методик и рекомендаций в области оценки рисков информационной безопасности.

При этом данный вопрос не теряет актуальности, так как постоянно появляются все новые и новые угрозы информационной безопасности.

Объектом исследования являются вопросы оценки рисков информационной безопасности.

Предметом исследования являются методы оценки рисков информационной безопасности в корпоративных сетях.

Целью является разработка программной реализации системы оценки рисков информационной безопасности в корпоративных сетях.

Для достижения поставленной цели необходимо выполнение следующих задач:

- Рассмотреть существующие угрозы информационной безопасности в корпоративных сетях;
- Рассмотреть понятие рисков информационной безопасности;
- Рассмотреть вопросы аудита состояния информационной безопасности;
- Рассмотреть существующие методы оценки рисков информационной безопасности;
- Разработать техническое задание на разработку программной реализации системы оценки рисков информационной безопасности;
- Разработать проект системы оценки рисков информационной безопасности;
- Разработать программную реализацию системы оценки рисков информационной безопасности.

Практическая значимость заключается в разработке программной реализации системы оценки рисков информационной безопасности в корпоративных сетях.

Перед началом программной реализации системы была разработана структура программных модулей с описанием.

Диаграмма структуры программных модулей приведена на рис. 1.

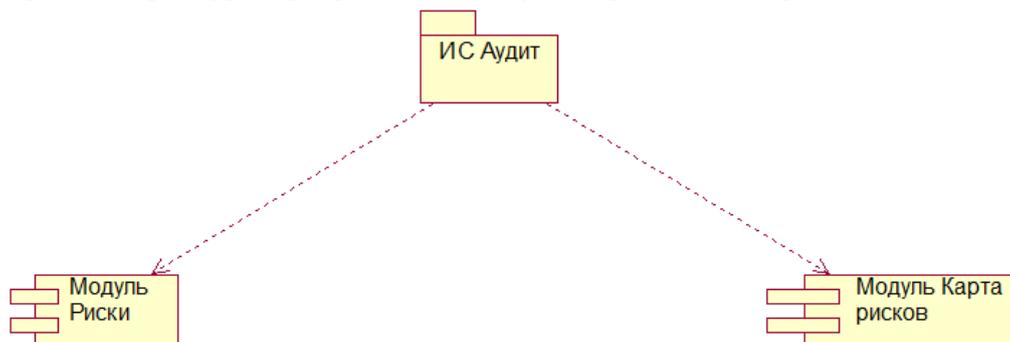


Рисунок 1. Диаграмма структуры программных модулей

На главном окне программы представлен перечень рисков информационной безопасности для корпоративной сети предприятия. На рис. 2 приведен пример главного окна программы.

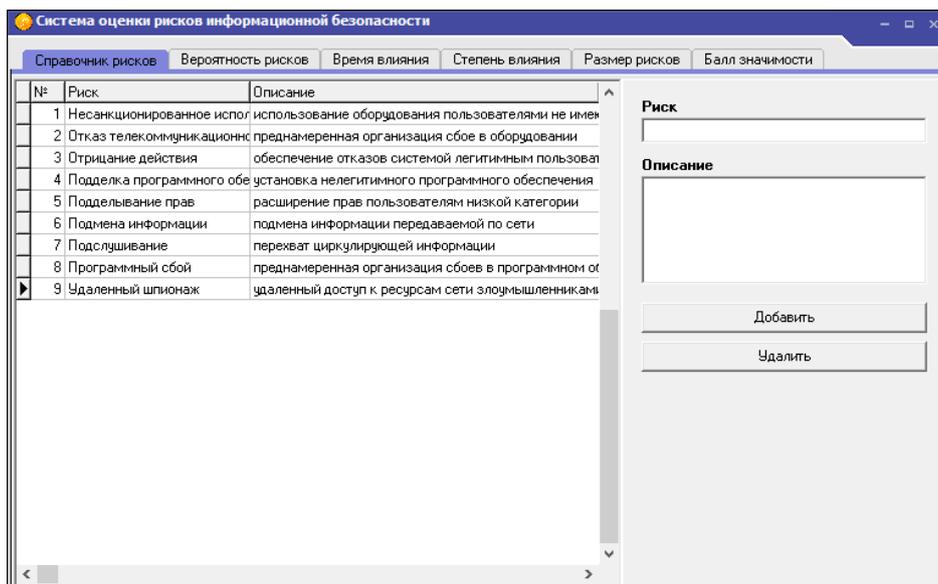


Рисунок 2. Главное окно программы

Также можно вывести на печать отчет – Карту рисков.
На рис. 3 приведен пример отчета системы «Карта рисков».

КАРТА РИСКОВ					
№	Риск	Вероятность	Время влияния	Размер рисков	Балл значимости
1	Несанкционированное использование оборудования	3	1	2	16
2	Отказ телекоммуникационного оборудования	1	1	1	2
3	Отрицание действия	3	3	3	12
4	Подделка программного обеспечения	4	2	5	24
5	Поддельвание прав	5	2	1	21
6	Подмена информации	4	3	5	21
7	Подслушивание	2	1	1	12
8	Программный сбой	4	3	5	14
9	Удаленный шпионаж	3	1	3	12

Рисунок 3. Отчет системы «Карта рисков»

Назначение разрабатываемой системы: повышение эффективности управления рисками организации. Пользователь входит в ИС Аудитор, вводит перечень возможных рисков, дает оценку характеристик для каждого вида рисков:

- Вероятность рисков;
- Время влияния риска;
- Степень влияния риска;
- Размер риска.

Далее система на основе выставленных оценок рассчитывает балл значимости риска. На выходе системы – отчет «карта рисков».

СЕКЦИЯ 3

SECTION 3

ЖАСАНДЫ ЗЕРДЕ
ИСКУССТВЕННЫЙ ИНТЕЛЛЕКТ
ARTIFICIAL INTELLIGENCE

RESEARCH ON TRANSITION-BASED SYNTACTIC PARSING FOR KAZAKH

*(College of Information Science and Engineering, XINJIANG UNIVERSITY, P.R. China.
The Base of Kazakh and Kirghiz Language of National Language Resource.
Monitoring and Research Center on Minority Languages, P.R. China)*

Abstract. This paper focuses on the Syntactic Parsing of the Arabic scripts in modern Kazakh based on Transition method by Perceptron algorithm. Traditional Kazakh Syntactic Parsing techniques generally adopt hierarchical and sequential analysis of individual tasks, such as Part-of-Speech (POS), Chunks and Syntactic Parse, and analysis of each task by their local state best models independently. This paper describes the Kazakh parsing system that using the perceptron model to do training and the shift-reduce to parsing with Beam-Search algorithm to decode. We use the average perceptron instead of the traditional perceptron model to solve the problem of over-fitting. Improve the Beam-Search decoding algorithm, since the search space of the joint model is the product of the search space of each task, that is, the fixed column value B to the dynamic column value, Improve the accuracy of the search space. Also, design the reward function to reduce the beam-Search decoding algorithm in the search when the pruning of the best results of the risk so that the accuracy of the results improved. As Syntactic Parsing are important parts of Kazakh language processing, they are the basis tasks of semantic analysis, machine translation, and many other applications in Kazakh language.

Key words : Transition-Based parsing; Perceptron Algorithm ; Beam-Search.

1 Introduction

Syntactic analysis is important tasks in natural language processing research for agglutinative languages. Syntactic Parsing have been developed for different languages with different methods, including the rule-based, statistical model or other model. The perceptron model is a classification model, which uses the perceptron algorithm to train the corpus and obtain the parameters, then use the perceptron to classify the phrases or words to achieve the effect of parsing. Collins^[1] explores an alternative approach to parsing, based on the perceptron training algorithm introduced in 2002. In this approach the training and decoding problems are very closely related – the training method decodes training examples in sequence, and makes simple corrective updates to the parameters when errors are made. Traditional perceptron algorithms sometimes perform very well in training, but it have the problem of parameter over-fitting, and the average perceptron can solve this problem well.

Briscoe and Carroll^[2] associates scores with each decision in the parsing process, selecting the parse which is built by the highest scoring sequence of decisions. In addition Yue Zhang^[3] describe a global discriminative model for Chinese shift-reduce parsing with Beam-Search algorithm to decode. They apply beam search to decoding instead of greedy search. The parser still operates in linear time. In the tradition beam-search algorithm, choose the B highest scoring state items in an agenda during the shift-reduce process. For uniform B, we select the appropriate value of B is difficult. If we choose the value of B is too large, while decoding accuracy has improved, but the decoding speed is greatly reduced. The value of B is too small, the opposite effect.

In this paper, we describe the parsing system that using the perceptron model to do training and the shift-reduce to parsing with Beam-Search algorithm to decode. In the aspect of parameter training, we use the mean perceptron instead of the traditional perceptron algorithm to avoid over-fitting. In addition, we use the Beam-Search decoding algorithm and design the dynamic B-value to construct the search space more rationally, so that the accuracy and efficiency While achieving the best results.

2. Learning Algorithm

In this section, we describe the overall framework of this lesson - linear model^[4,5], which has been widely used in many fields. It gets good performance not only in the part-of-speech tagging, but also in the syntax analysis. We will deal with linear models by taking syntactic analysis as an example.

2.1 Linear Models in NLP

In this part, we use syntactic analysis as an example to explain the basic principles of linear models. The main task of the linear model is mapping $y \in \mathcal{Y}$ for the input $x \in \mathcal{X}$. In the syntactic analysis, \mathcal{X} is a set of untreated sentences, and \mathcal{Y} is the set of the most probable syntactic trees for a χ sentence. Here are four assumptions:

- (1) Training samples (x_i, y_i) , where the range of i from 1 to n
- (2) $\text{GEN}(x)$, a function which collects all possible sets of syntactic trees of input x
- (3) Φ denotes the mapping from each group $(x, y) \in \mathcal{X} \times \mathcal{Y}$ to the feature vector $\Phi(x, y) \in \mathbb{R}^d$
- (4) $\bar{\alpha}$ is a parameter vector which belongs to \mathbb{R}^d

The above four hypothetical contents, like GEN , Φ , $\bar{\alpha}$ define a mapping from input x to output $F(x)$:

$$F(x) = \arg \max_{y \in \text{GEN}(x)} \Phi(x, y) \cdot \bar{\alpha} \quad 2 - 1$$

Where $\Phi(x, y) \cdot \bar{\alpha}$ can be expressed as $\sum_s \alpha_s \Phi_s(x, y)$, and $\bar{\alpha}$ can be obtained by training features. If we want to get the results of the formula, we need to find the appropriate decoding algorithm. We will give the training method about $\bar{\alpha}$ in next section.

2.2 Parameter Estimation

This section focuses on the use of training samples, which are used for training $\bar{\alpha}$. Its convergence properties— see Collins for a full description. The algorithm and theorems are based on the approach to classification problems described in Freund and Schapire. We use the perceptron algorithm^[1,6], and its description is specified as follows:

Algorithm 1 Tradition perceptron algorithm Algorithm 2 Averaged perceptron algorithm

```

Input:  $D = \{(x, y)\}_N$ 
 $w \leftarrow 0$ 
for  $t = 1 \dots T$  do
  for  $(x_i, y_i) \in D$  do
     $z = \arg \max_{y' \in \text{GEN}(x_i)} (\varphi(x_i, y') \cdot w)$ 
    if  $z \neq y_i$  then
       $w \leftarrow w + \varphi(x_i, y_i) - \varphi(x_i, z)$ 
    end if
  end for
end for
return  $\bar{w}$ 

```

```

Input:  $D = \{(x, y)\}_N$ 
 $w \leftarrow 0$ 
for  $t = 1 \dots T$  do
  for  $(x_i, y_i) \in D$  do
     $z = \arg \max_{y' \in \text{GEN}(x_i)} (\varphi(x_i, y') \cdot w)$ 
    if  $z \neq y_i$  then
       $w \leftarrow w + \varphi(x_i, y_i) - \varphi(x_i, z)$ 
    end if
  end for
end for
 $\bar{w} = \frac{1}{NT} \sum_{n=1 \dots N, t=1 \dots T} w_{n,t}$ 
return  $\bar{w}$ 

```

Algorithm 1 is the tradition perceptron algorithm with two drawbacks: 1) When the data is inseparable, how to guarantee the correctness of the algorithm? 2) When the performance is very good in the training corpus, whether or not it appears over fitting in the test corpus? In order to solve the above problems, we use the averaged perceptron algorithm to average the training parameters, and the averaged perceptron algorithm is shown in Algorithm 2.

3. Decoding algorithm

The most complex step in the algorithm is to find the value of $z_i = \arg \max_{y \in \text{GEN}(x)} \Phi(x, y) \cdot \bar{\alpha}$ in the same way that we have introduced the basic principle of the perceptron algorithm in Section 2. This step is the decoding of the syntax tree, so the main problem of the perceptron algorithm is to decode the syntax tree. Once the decoding algorithm is implemented, the training of the parameters is relatively simple. The Beam-Search algorithm^[3,7]

belongs to the heuristic search algorithm, compared with the dynamic programming algorithm, Beam - Search algorithm with the expansion of every step, not looking for all possible situations, but rather through to lose some small weight and probability of node to control the size of the Search space. Yue Zhang use the Beam-Search algorithm to decode the syntax analysis, which combines the transfer model to introduce the basic principles of Beam-Search as shown in Algorithm 3:

Algorithm 3 Beam-Search decoding algorithm

Variables: state item $item = (S, Q)$, where
 S is stack and Q is incoming queue;
 the agenda agenda;
 list of state items next;

Algorithm:

for $item \in agenda$:
 if $item.score = agenda.bestScore$ and
 $item.isFinished$:
 $rval = item$
 break
 $next = []$
 for $move \in item.legalMoves$:
 $next.push(item.TakeAction(move))$
 $agenda = next.getBBest()$

Outputs: rval

Algorithm 4 Improved Beam-Search decoding

Variables: state item $item = (S, Q)$, where
 S is stack and Q is incoming queue;
 the agenda agenda;
 list of state items next;

Algorithm:

for $item \in agenda$:
 if $item.score = agenda.bestScore$ and
 $item.isFinished$:
 $rval = item$
 break
 $next = []$
 $B = 0$
 $threshold = threshold * \alpha$ ($0 < \alpha < 1$)
 for $move \in item.legalMoves$:
 $next.push(item.TakeAction(move))$
 for elect $\in next$:
 if $next.bestScore/elect.Score$
 $\leq threshold$
 $B = B + 1$
 $agenda = next.getBBest()$

Outputs: rval

First, the beam-search algorithm is used to select the highest B candidate sets during the Shift-Reduce process. The agenda is initialized to contain all the initial states. Such as: an empty stack and a queue, at each step of the decoding process, the current items from the agenda are progressed through the transfer reduction step, and after each action, a series of new candidate sets are generated, and the Beam-Search decoding The algorithm still chooses the highest score of the B candidate sets into the agenda until the Q state space is empty to indicate that the decoding algorithm terminates. In this paper, the static B value is changed to the dynamic B value. The choice of the dynamic B value is based on the score of the candidate set and the highest value in the candidate set, and the threshold value is selected. The size of the threshold decreases linearly with the delay of the transfer step. Prune. The specific algorithm description is described in algorithm 4.

4 The Shift-Reduce Parsing Process

Assume that the Shift-Reduce parsing process is a process of constructing a syntactic tree. A sample example is shown in Figure 1 below to show the Shift-Reduce process in detail. In the Shift-Reduce^[6,8,9] syntax analysis process, we first assume that the input is the result of the tagging of the sentence's parts of speech. In the initial state of Shift-Reduce, the word is stored in the queue which is already tagged; and stack is empty. The parse state is defined as $\langle queue, stack \rangle$. The analysis action consists mainly of a Shift operation and two Reduce operations. Shift-Reduce process is the elements of the queue in turn Shift to the stack, then make the stack elements do Reduce processing. The Shift-Reduce action used in this article is defined as follows:

初始状态	栈	队列
		$\overset{n}{\text{لادلىق}}$ $\overset{n}{\text{يدى تاماش}}$ $\overset{v}{\text{عبىز مۇراج}}$
SHIFT步后:	$\overset{n}{\text{لادلىق}}$	$\overset{n}{\text{يدى تاماش}}$ $\overset{v}{\text{عبىز مۇراج}}$
REDUCE – unary – (X,Y)步后:	$\overset{n}{\text{لادلىق}}$ $\text{S}(b)$	$\overset{n}{\text{يدى تاماش}}$ $\overset{v}{\text{عبىز مۇراج}}$
SHIFT步后:	$\text{S}(b)$ $\overset{n}{\text{لادلىق}}$ $\overset{n}{\text{يدى تاماش}}$	$\overset{v}{\text{عبىز مۇراج}}$
REDUCE – unary – (X,Y)步后:	$\text{S}(b)$ $\text{VP}(b)$ $\overset{n}{\text{لادلىق}}$ $\overset{n}{\text{يدى تاماش}}$	
SHIFT步后:	$\text{S}(b)$ $\text{VP}(b)$ $\overset{n}{\text{لادلىق}}$ $\overset{n}{\text{يدى تاماش}}$ $\overset{v}{\text{عبىز مۇراج}}$	
REDUCE – unary – (X,Y)步后:	$\text{S}(b)$ $\text{VP}(b)$ $\text{VP}(b)$ $\overset{n}{\text{لادلىق}}$ $\overset{n}{\text{يدى تاماش}}$ $\overset{v}{\text{عبىز مۇراج}}$	
REDUCE – binary – {L/R} – Z步后:	$\text{S}(e)$ $\text{S}(b)$ VP $\overset{n}{\text{لادلىق}}$ $\overset{n}{\text{يدى تاماش}}$ $\overset{v}{\text{عبىز مۇراج}}$	
REDUCE – binary – {L/R} – Z步后:	S $\text{S}(b)$ $\text{S}(e)$ $\overset{n}{\text{لادلىق}}$ $\overset{n}{\text{يدى تاماش}}$ $\overset{v}{\text{عبىز مۇراج}}$	

Figure 1 The Shift-Reduce Parsing Process

(1) SHIFT: Remove the first word-POS in the queue to the stack.

(2) REDUCE–unary–(X,Y): the categories to which the newly moved nodes are classified and labeled X, Y and X include VP, NP and other phrase structures. Y represents b, i, e (eg: b, i, e, combine the previous phrase structure (VP, b / i / e), (VP, b) represents an element belonging to the current parent node VP leftmost child node (VP, i) represents the middle child node of the parent node VP, and (VP, e) represents the rightmost child node of the parent node VP).

(3) REDUCE–binary–{b/i/e }–(Z,Y): when the elements in the stack are marked with (X, Y), check whether the classification Y in the top of the stack is i, If i, the system will merge the top two elements of the stack to form a new sub-tree, the sub-tree is a Z phrase, and a Y flag for Z.

(4) TERMINATE: when the elements in the queue is empty, there is only one element in the stack, that is, the parse tree to stop the Shift-Reduce operation.

5 Features

Feature Selection (Feature Selection) is to select all the features from all the characteristics of a subset, the use of partial feature subset to establish statistical models, so that to achieve the best results. In this paper, we choose two sets of feature to compare, in which the first group does not add lexical information, the second group adds lexical information, the accuracy rate is 1.2% higher. The two feature templates are shown in Figures 2 (a) and (b).

		Unigrams	$P_{0w}, P_{0t}, P_{bt}, P_{bw}, P_{fcw}, P_{fct}, P_{cw}, P_{ct}, P_{clw}, P_{clt}, P_{crw}, P_{crt}$.
Unigrams	$P_{0w}, P_{0t}, P_{bt}, P_{bw}, P_{fcw}, P_{fct}, P_{cw}, P_{ct}$.	Bigrams	$P_{0w}P_{0t}, P_{0w}P_{bt}, P_{0t}P_{bt}, P_{0t}P_{fct}, P_{bt}P_{fct}, P_{0t}P_{ct}, P_{0t}P_{clw}, P_{0t}P_{clt}, P_{0t}P_{crw}, P_{0t}P_{crt}, P_{clt}P_{crt}$.
Bigrams	$P_{0w}P_{0t}, P_{0w}P_{bt}, P_{0t}P_{bt}, P_{0t}P_{fct}, P_{bt}P_{fct}, P_{0t}P_{ct}$.	Trigrams	$P_{0w}P_{0t}P_{bt}, P_{0w}P_{0t}P_{fct}, P_{0t}P_{bt}P_{bt}, P_{0t}P_{bt}P_{fct}, P_{0t}P_{bt}P_{ct}, P_{0t}P_{clt}P_{clw}, P_{0t}P_{clt}P_{crt}$.
Trigrams	$P_{0w}P_{0t}P_{bt}, P_{0w}P_{0t}P_{fct}, P_{0t}P_{bt}P_{bt}, P_{0t}P_{bt}P_{fct}, P_{0t}P_{bt}P_{ct}$.		

(a) (b)
Figure 2 Feature Template

In the feature template, P_{0w} represents the current phrase, P_{0t} represents the mark of the current phrase, P_{bw} represents the brother node, P_{bt} represents the brother node of the current phrase, P_{fcw} represents the parent node of the parent node, P_{fct} represents the parent node's parent node of the mark, P_{cw} , P_{ct} , respectively, said the current node behind the cousin node and its mark.

6 Empirical Result

6.1 Experimental Settings

In the experiment, we select 120 sentences from 1200 sentences of manually annotated in the Xinjiang Daily (Kazakh) for small-scale testing. Precision, recall and F-measure which is used to identify the results of the evaluation. For the result of Kazak parsing, the evaluation function is defined as follows:

a = Identify the correct number of phrases ;

b = The number of phrases is not recognized ;

c = Identify the number of error phrases ;

$$\text{Precision} : P = \frac{a}{a+c} \times 100\%$$

$$\text{Recall} : R = \left(\frac{a}{a+b} \right) \times 100\%$$

$$\text{F-measure} : F\text{-measure} = \left(\frac{2 \times P \times R}{P+R} \right) \times 100\%$$

6.2 Experimental Analysis

In this paper, we have analyzed the improved algorithm and the original algorithm section. In this section, we focus on the results of the improved algorithm and the original algorithm. First, the averaged perceptron algorithm is compared with the traditional perceptron algorithm. In contrast to Table 1, we used two algorithms to train five sets of data, and the averaged perceptron algorithm was 1.4% higher than the average of the five sets of data accuracy of the traditional perceptron algorithm. There are two sets of data in the five sets of data. The accuracy of the traditional perceptron model is higher than the averaged perceptron model.

Table 1 Comparison of Averaged Perceptron and Traditional Perceptron Results

	Averaged perceptron model			Tradition perceptron model		
	Precision(%)	Recall(%)	F(%)	Precision(%)	Recall(%)	F(%)
First group	70.1	66.3	68.1	65.8	62.6	64.2
Second group	69.3	65.8	67.5	73.6	67.8	70.6
Third group	69.8	66.1	67.9	63.7	61.7	62.7
Fourth group	69.7	66.4	68.0	64.3	61.4	62.8
Fifth group	70.6	67.2	68.9	75.1	70.2	72.6
Averager	69.9	66.4	68.1	68.5	64.7	66.6

In the decoding algorithm, this paper makes a corresponding improvement to the Beam-Search algorithm, and uses the dynamic B value instead of the original fixed B value to make the search space more accurate. When we get the threshold and the parameter α are 7 and 0.92 respectively acquire best results. In Table 2, the decoding accuracy and decoding space contrast of the dynamic B value and the fixed B value (B value selection is 16) are shown. As can be seen from table 2, although the accuracy of using the dynamic B-value decoding is only 0.4 higher than that of B = 16, the average search number of sentences in the sentence is greatly reduced.

Table 2 Comparison of Dynamic Beam and Fixed Beam Results

Dynamic Beam (Threshold = 7, $\alpha = 0.92$)				Fixation Beam(B = 16)			
Precision(%)	Recall(%)	F(%)	The Averager number of searches	Precision(%)	Recall(%)	F(%)	The Averager number of searches
69.9	66.4	68.1	106	69.5	66.3	67.9	138

7 Conclusion

Traditional Kazakh Syntactic Parsing techniques generally adopt hierarchical and sequential analysis of individual tasks, such as Part-of-Speech (POS) , Chunks and Syntactic Parse, and analysis of each task by their local state best models independently. However, there are two problems that the low-level task analysis of the error will have an impact on the high-level analysis, and each layer of the task using local optimal. For this problem, the Transition-Based joint model is a good solution to solve the problem, its decoding efficiency has obvious advantages. In this paper, we use the averaged perceptron model instead of the traditional perceptron model for corpus training, the overall training results increased by 1.3%. In addition, in the decoding process we use the improved Beam-Search decoding algorithm, the use of dynamic B value instead of fixed B value decoding efficiency and accuracy have a corresponding increase.

Acknowledgements

This work is supported by National Natural Science Foundation of China, under grant (NSFC), under grant No. 61363062 and other No. NMLR 201601).

References

- 1 COLLINS M. Discriminative training methods for hidden markov models: Theory and experiments with perceptron algorithms[C]//Proceedings of the ACL-02 conference on Empirical methods in natural language processing-Volume 10. Association for Computational Linguistics, 2002: 1–8.
- 2 BRISCOE T, CARROLL J. Generalized probabilistic LR parsing of natural language (corpora) with unification-based grammars[J]. Computational linguistics, 1993, 19(1): 25–59.
- 3 ZHANG Y, CLARK S. Transition-based parsing of the Chinese treebank using a global discriminative model[C]//Proceedings of the 11th International Conference on Parsing Technologies. Association for Computational Linguistics, 2009: 162–171.
- 4 JIANG W, HUANG L, LIU Q等. A cascaded linear model for joint chinese word segmentation and part-of-speech tagging[C]//In Proceedings of the 46th Annual Meeting of the Association for Computational Linguistics. Citeseer, 2008.
- 5 RATNAPARKHI A. A linear observed time statistical parser based on maximum entropy models[J]. arXiv preprint cmp-lg/9706014, 1997.
- 6 ZHANG Y, CLARK S. Chinese segmentation with a word-based perceptron algorithm[C]//Annual Meeting-Association for Computational Linguistics. 2007, 45: 840.
- 7 ZHANG Y, CLARK S. Syntactic processing using the generalized perceptron and beam search[J]. Computational linguistics, 2011, 37(1): 105–151.
- 8 SAGAE K, LAVIE A. A classifier-based parser with linear run-time complexity[C]//Proceedings of the Ninth International Workshop on Parsing Technology. Association for Computational Linguistics, 2005: 125–132.
- 9 WANG M, SAGAE K, MITAMURA T. A fast, accurate deterministic parser for Chinese[C]//Proceedings of the 21st International Conference on Computational Linguistics and the 44th annual meeting of the Association for Computational Linguistics. Association for Computational Linguistics, 2006: 425–432.

KANTARBAYEV B.A., NIYAZOVA R.

THEORETICAL ABSTRACT MODEL OF LOCOMOTIVE SIMULATOR

Introduction

Theoretical novelty of this work lies in the development and further use of an ontological system. The system, which is going to be based on “ТЭ33А Evolution series” locomotive model, will ultimately lead to creation of an operational simulator with various interfaces. Thus, the distinctive characteristic of this research lies in creation of a unique simulation based system and its elements that could be readjusted based of requiring needs and different locomotive model types for better control of railroad systems and education of locomotive operators.

Brisk development of information technologies worldwide has led to an increased need of IT development and adoption in all possible areas of our everyday life – engineering facilities, industrial bases, production operation facilities, economies, coffee shops, and so on. Nevertheless, the pace of IT incorporation into our life causes new problems to appear. Such problems may involve low quality means of introduction of these technologies. This specific problem is particularly noticeable when introducing technically complex mechanisms and machines in the production and industry of some countries

Since the beginning of the 20th century, great minds have been actively using simulation methods and techniques in order to achieve better results in many areas of science and engineering. One of these endeavors can be traced back to the times of the First World War, when modified versions of a normal horse powered carts or wagons were used for pilot training. These carts and wagons would be operated manually by human force in order to imitate “dog fights” (aircraft combat engagements). In the modern era, to ensure higher chances of success, different types of simulators (anything from a driving simulator to a space aircraft simulator) are being implemented.

Thus, simulators have had and still have a huge impact on the way modern society exists, its security and continuous progress.

Ontology of the locomotive simulator.

A basic ontology model was created as a first step in the simulator development process. In this basic model the most important main and supportive components were included, such as control, model, load and atmospheric conditions. Hierarchy and relations of these elements are shown in Figures 1.1 ad 1.2.

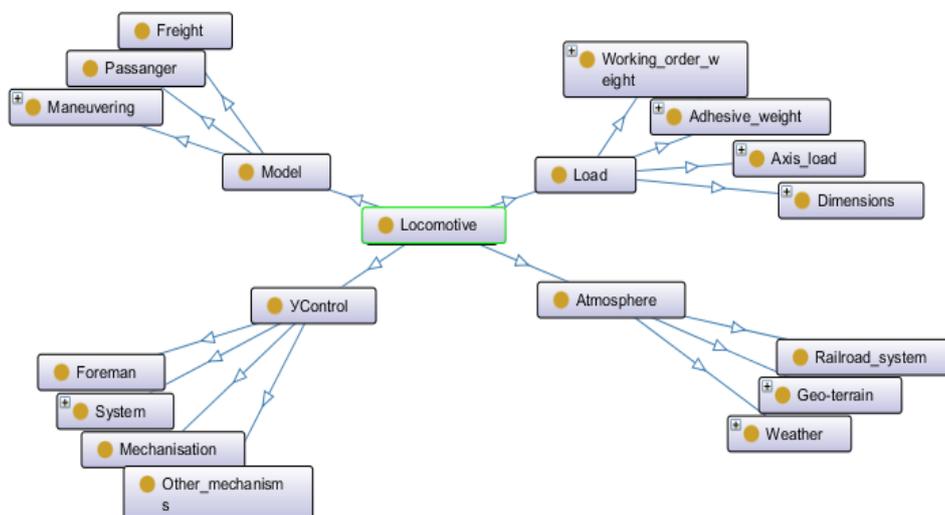


Figure 1.1 - Ontology of the locomotive simulator

industry without use of computer based tools. Hence, information technology keeps developing at an enormous rate.

Table 1.1. Thesaurus of the ontological model

Word	Definition	Synonym	Antonym	Hyponym	Meronym
Locomotive	a self-propelled, vehicular engine, powered by steam, a diesel, or electricity, for pulling or, sometimes, pushing a train or individual railroad cars.	Тепловоз	--	--	wheel, engine ...
Frame (model)	Locomotive's base (usually made of metal)	Body	Interior	--	--
Freight	A train for carrying goods rather than people.	Shipping	--	--	--
Maneuvering	Used for maneuvering operations on railroad stations and approach rails	Manipulative	--	--	--
Passenger (carrying)	Public transport provided by a line of railway cars coupled together and drawn by a locomotive; "express trains don't stop at Princeton Junction	--	--	--	--

While being an advantage, this development also causes different issues in the areas of industrial sector in Kazakhstan. Without a doubt, creation, approbation and incorporation of simulator based tools and technologies will allow having a better control and levels of proficiency in the railroad system among locomotive operators and tool developers.

Prospects of the research

This research is not the first or a unique in the area of IT, yet, it will have a humble but considerable influence and prospects for future development. The idea of creation of a multifunctional locomotive simulator is both hard and fascinating, has a great scientific interest for specialists in the area of railroad operations and development.

Ultimately, creating of such simulator will facilitate promoting the idea of simulation tool development in many neighboring fields of science, computer science and engineering. Also, the success of this research can lead to an improvement not only in the model described in this paper, but also in subsequent systems in the future.

References:

1 V. V. Golenkov, L. S. Globa, N. A. Gulyakhina, I. V. Efimenko, O. P. Kuznetsov, B. M. Lobanov, D. S. Suleimanov, A. A. Kharlamov, V. F. Khoroshevskii, "Open semantic technologies in the design of intelligent systems "OSTIS-2014"" («Открытые семантические технологии проектирования интеллектуальных систем "OSTIS-2014"»), Belorussian state university of computer science and radio electronics (Белорусский государственный университет информатики и радиоэлектроники), 2014.

2 A. I. Belousov, N. M. Bogrest ... Sharipbayev Altynbek, journal "Ontology of design" (журнал «Онтология проектирования»), 2013.

3 Natl Academy, «Simulated Voyages: Using Simulation Technology to Train and License Mariners», June 1996

4 <http://www.thesaurus.com/> (multiple definitions)

5 <http://www.dictionary.com/> (multiple definitions)

6 B. A. Kantarbayev «ONTOLOGICAL MODEL OF LOCOMOTIVE SIMULATOR»

STUDY OF KEYBOARD LAYOUT FOR NEW KAZAKH LATIN ALPHABET

(L. N. Gumilyov Eurasian National University)

The Kazakh alphabet has faced several changes and it has an ancient history from ancient Turkic writing (enharmonically), using letter based on the Arabic script almost for a thousand years, and from 1940 switching to Cyrillic graphics, which is used up to present years. In addition, it is remarkable to note that there was a short period of time when Kazakh alphabet has been written in Latin alphabet.

As we can see during the twentieth century Kazakh letter was changed three times. Moreover, in this kind Turkic-speaking countries such as Azerbaijan, Turkmenistan, Uzbekistan and the Republic of Karakalpakstan have faced fourth change of graphics, as they again moved to the Latin alphabet at the end of the XX century.

In Kazakhstan most people are bilingual who can fluently speak in Kazakh and Russian. Today, as English language is getting popular among the youth and with the promotion and encouragement of Government to make country trilingual various state programmes have been introduced. It is claimed that English language is language of successful integration into the global economy.

Nowadays, besides English language, Internet is growing rapidly and widely used in Kazakhstan. Development of Kazakh language Internet space is directly related to the development of KazNet. In Kazakhstan and beyond there is a large number of potential users of the Kazakh language Internet, who speak in Kazakh. Given the importance of the status of Kazakh language as the state language, the number of potential users of Kazakh language Internet, will grow it can be observed from the number of registered domains and websites created.

However, it is been actively discussed for the long time now the problem of representation of the Kazakh alphabet characters in computer programmes and the Internet. The crux of the problem lay in the fact that in the past when many software products were developed the use of symbols, different from those presented in the English language alphabet was not taken into account. Thus, if there are additional characters as in many European languages, or when using, for example, Cyrillic or Chinese characters, there is a need to create and add new sets of fonts.

As one of the solutions of the problem with the encodings, it was proposed to transfer the alphabet to Latin characters. The main arguments of supporters of this idea was that the modern computer keyboard created for the Latin script, so the transmission of the Kazakh language will allow solving all problems.

Moreover, there is a difficulty when typing Kazakh on standard keyboard. While English use Latin script, Kazakh and Russian share Cyrillic and Kazakh alphabet includes all Russian symbols. It might look that two keyboard layouts would be sufficient to type in all three languages, however people's preference differs from the statement. A survey has been conducted for this paper, where 87 out of 91 stated that they use at least three layouts at their home or on work computers. The main reason for that is current Kazakh Cyrillic alphabet has too many letters than a standard QWERTY keyboard can fit without compromise by replacing numbers with specific Kazakh letters.

Many supporters of the transition of the Kazakh letters from the Cyrillic to the Latin alphabet say that Latin is very easy that it will be easier to write than the Cyrillic alphabet. Murat Auezov says "...the Latin alphabet is very simple". Actually it is not so.

Internal capabilities of Latin letters are very limited. The reason is that the Latin alphabet was not seriously transformed in the languages that adopt it as a national letter. The Latin script and alphabet was adapted to these languages, in fact, with no change or solely minor changes. Therefore, due to the importance of adapting Latin script to Kazakh language, number of alphabets

proposed by various scientific institutions and amateurs. Some argue that no additional letters should be added to the current English alphabet, to make it easier to adapt for native speakers. This approach was successfully adopted in Uzbekistan. Since there are more sounds in Kazakh language than letters in standard Latin script, use of apostrophe or dual letters is being recommended. If this method is chosen as the most appropriate, there will be no need to invent a new Kazakh layout. On the other hand, a strong group of scientists is in favour of adding some letters with diacritics to standard Latin alphabet in order to distinctly express different sounds. Universal Turkic alphabet was created in 1991 to promote panturkistic ideas. Azerbaijan was one of the first countries shifting to that alphabet. In proposed Kazakh Latin alphabets, which promote adding extra letters, the number of extra letters varies from 3 to 9. In case of confirmation of an alphabet with diacritic letters, invention of a new keyboard layout will be necessity. Instead of considering how to align letters of the confirmed alphabet to the keyboard, this paper will analyse how many letters can be added to the keyboard. For this study standard QazAqparat alphabet is used that is shown below in Table 1

Cyrillic	а	ә	б	в	г	ғ	д	е	ё	ж	з	и	й	к	қ	л	м	н	ң	о	ө
Qazaqparat latin	А	Ä	В	У	Г	Ğ	Д	Е	Yo	Ж	З	İ	У	К	Q	L	M	N	Ñ	О	Ö
	а	ä	в	у	г	ğ	д	е	yo	ж	з	ı	у	к	q	l	m	n	ñ	о	ö
Cyrillic	п	р	с	т	у	ұ	ү	ф	х	һ	ц	ч	ш	щ	ъ	ы	і	ь	э	ю	я
Qazaqparat latin	P	R	S	T	W	U	Ü	F	X	H	C	Ç	Ş	ŞŞ	"	ı	ı	'	É	Yw	Ya
	p	r	s	t	w	u	ü	f	x	h	c	ç	ş	şş		ı	ı		é	yw	ya

Table 1. Standard QazAqparat alphabet

From the table we can see that all symbols from English alphabet is used and there are 9 extra letters added. In standard keyboard there are only 7 keys left for these 9 letters and punctuation marks. Accommodating all of these symbols might lead to the mentioned problem, when users have to change layout to reach some missing symbols on the keyboard. Moreover, having four different layouts at the same time where same symbols located on various places depending on the layout might influence users to neglect the orthography and punctuation. The entire point of switching the alphabet will be lost and this will be following footsteps of Uzbekistan and Azerbaijan. Both countries changed latin alphabets within years after first adopting it.

To solve the problem with key constraint on the keyboard and instead of adding extra letters to the existing alphabet it can be proposed to create a new alphabet based on native Kazakh sounds not direct transliteration of the current cyrillic alphabet. Combining 'x' and 'h' sounds into one and removing 'u' would give three more available keys. It is also good to leave punctuation marks in their places in QWERTY layout to avoid confusions.

One of the attempts of an efficient keyboard layout was standardized DVORAK layout. The design puts majority of regularly utilized letterpress in the home row, where it is simple to reach, likewise, rarely used letters are placed in the bottom row. DVORAK layout is designed for the right hand users, while QWERTY is more suitable for the left hand users. QWERTY itself was invented in order to place letters that come together more frequently far apart from each other to avoid jams.

These techniques can be used to place extra letters to the available keys. First, we must count frequency rate of extra letters and punctuation symbols, so more frequent symbols can be placed on the home row, where they are easy to reach. Similarly, digraph frequency table must be

1.3 Tangible Benefits:

- Cost saving, user does not need to visit the embassy and pay in order to get some information.
- Paperless system, user does not need to fill up the form and photocopy individual documents in case of registration.
- User does not need to spend any money for transportation in order to visit a certain embassy.

1.4 Intangible Benefits:

- User will save their time by getting all the information and requirements online, rather than spending time for the way to go to the embassy and take a queue.
- The user can access to the system at any place and anytime, whereas consular department of the embassy can consult only at the working time.
- User's information will be more secure since only authorized personnel can view personal details.
- User can create a forum or discussion board where she/he can share the knowledge and experience with others.
- Much convenience for users to get a fast answer or respond online in case of any question rather than making an appointment and wait for the appointment day.
- Each individual's problem will be kept private.

1.5 Nature of challenge

While gathering the information it was hard to identify the group of users for the portal. On the other hand how many people to survey and interview. Moreover was difficult to decide if the **WPFVA** needs to have a payment session for the application or no. Other difficulty is making the system web based or mobile application. Which programming language will be more an appropriate for the web based system, Html that is standard or PHP that is more professional. In order to overcome these, need to do some additional research on programming languages and gather more information on visa application portals.

1.6 Scope and objectives

Web Portal For receiving and analyzing Visa Applications allows users to control and manage their visa application on web by connecting to the internet. By using this system user does not need to physically visit the Consular department or the Embassy. User can easily visit the website start to manage their documents. This can definitely save users' time and money.

The core functions of the system to be achieved before delivering as follow:

- Allow users to register.
- Allow users to get information online.
- Allow users to check the requirements
- Allow users to visit other pages as well as guest users.
- Allow users to create a forum on visa issues.
- Allow users to get an email notification as a reminder.
- Allow users to pay online for a visa.
- Allow users to contact a consular department

1.7 Functionalities

Besides that, the additional features of the system are as follows.

- Allow users to open, edit and delete a profile.
- Allow users to watch TV online.
- Allow users to watch a documentary videos and pictures of the country

Apart from that, the special features of the system will be added if there is any extra time.

- Send SMS notification to the users as a reminder.
- Allow users to create an event such as conference, live chat.
- Develop a mobile application.
- Change language (russian, kazakh, arabic)

2 METHODOLOGY

As for a methodology was chosen Waterfall to implement the “**Web Portal For receiving and analyzing Visa Applications**”. Waterfall model is a sequential software development model where development is seen as flowing downwards like waterfall. There is no overlap or going backward in waterfall. The system components are designed by coders and integrated together.

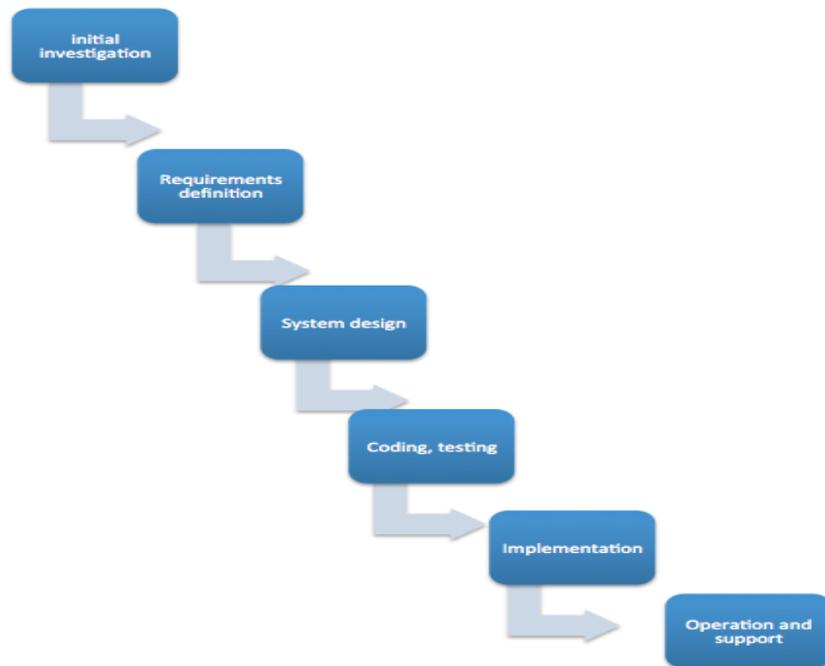


Figure 1. Waterfall Methodology (Self-drawn)

Another acceptable system development methodology for **WPFVA** will be Prototyping methodology.

It is a sample implementation of the system that shows limited and main functionalities of the proposed system. It helps to overcome the limitations that were found in previous Waterfall methodology. After prototype has been done it will be delivered to the customer to evaluate it later. There are two types of prototype such as throwaway prototype and evolutionary prototype. As for implementing system screen mock-ups and storyboards must be used. And for evolutionary prototyping must get feedback on the system from the users. In this system storyboards will be used as for prototyping.

So combination of two methodologies, which are Waterfall and Prototyping, will help to define the software concept, requirement analysis, design of architecture and system core and then will follow to culminate the installation of the final prototype.

In conclusion, **Web Portal For receiving and analyzing Visa Applications** would be a requested system for almost all people who will travel to the Middle East. This is because the current system needs to be improved and add some more functionality. Similar system will help to develop the system in more efficient and interactive way. In order to develop the system researcher will follow the selected methodology and proceed with storyboarding before delivering the system to the users.

3 DESIGN

Introduction

Under this chapter developer will describe the design of the system, such as use case diagram, activity diagram, UML class diagram, entity relationship diagram, data dictionary and storyboards for all the pages.

System design : Storyboards

Home page:

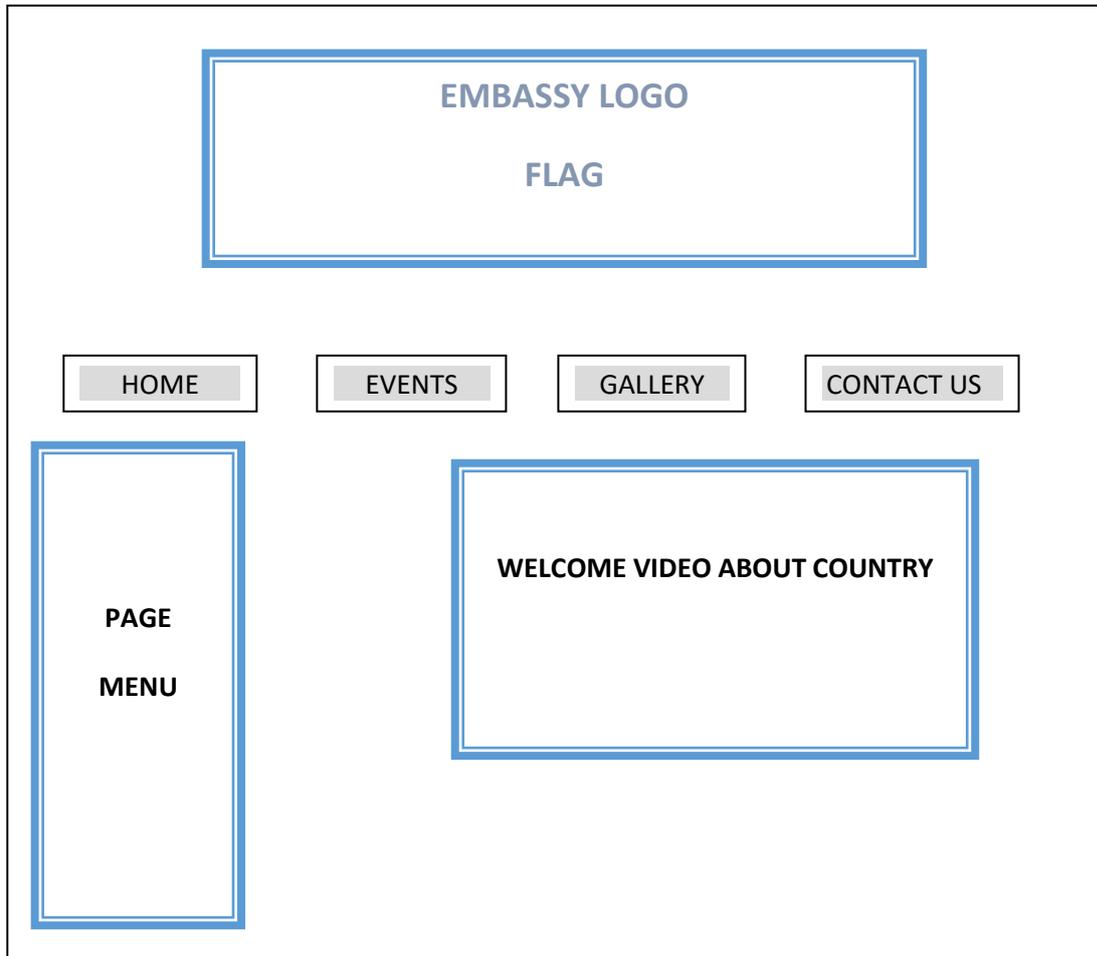


Figure 2. Home page.

Description: This is a Home page. Once user will open the website, he/she will see this page first.

WELCOME VIDEO: It allows the users to take a look on COUNTRY location, culture and traditions.

EVENTS: This is about the events that embassy holds (updated).

GALLERY: It allows the users to see all pictures that was taken during events and receptions.

Contact us: User can get numbers of the consular department and also see working information, besides that they can give suggestions to the system admin. In order to improve the features and services and also user can just email admin for some request.

Page menu: there is some information about the country. Information has been divided into certain parts that can be useful for the users.

Visa application submission:

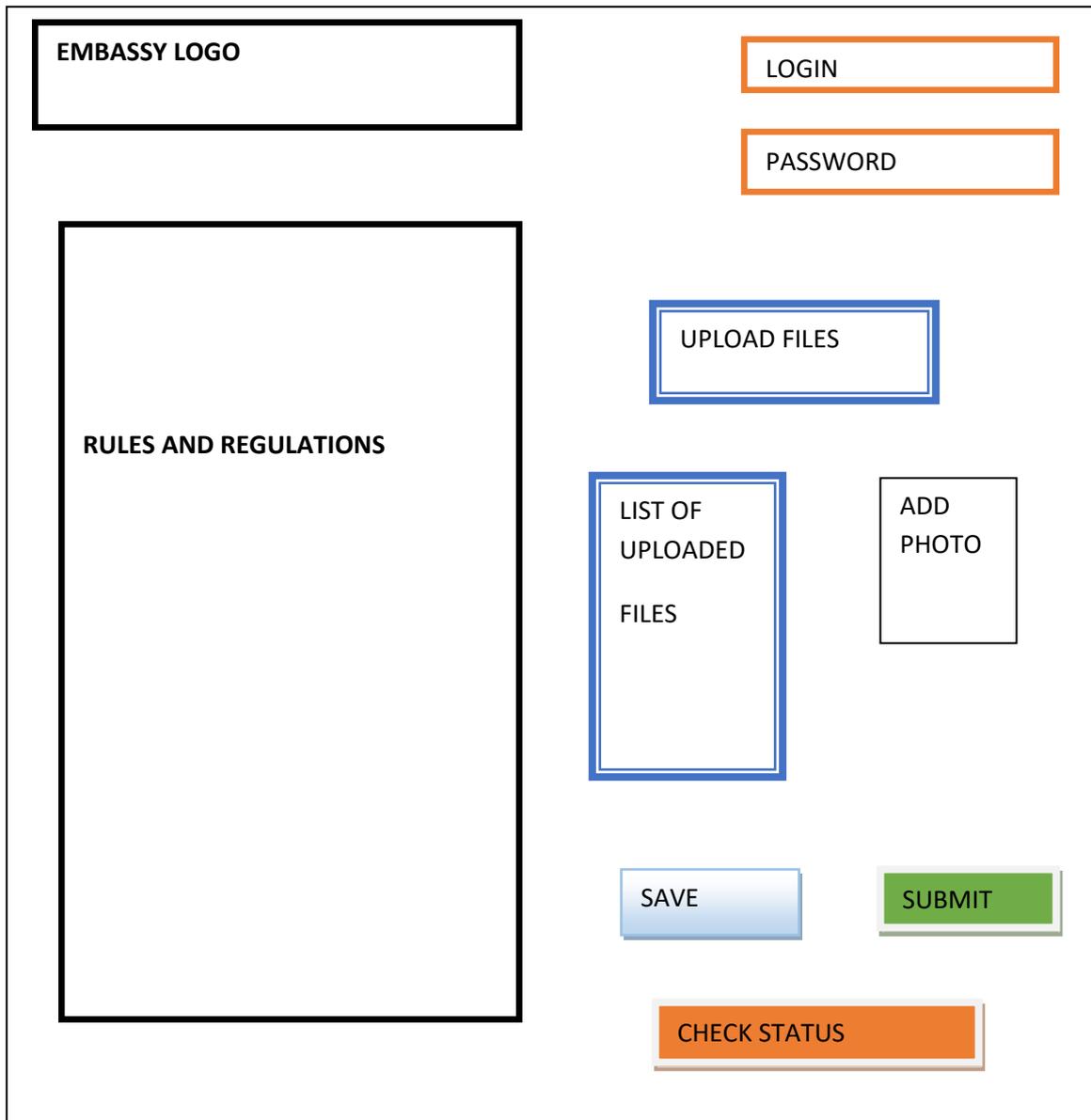


Figure 3. Visa application submission

Description: This page allows users to login, upload their documents, read the rules and regulations concerning visa for each nationality, they can save their documents that has been already uploaded. User has to upload a photo for a visa. After completion user can submit all the documents in order to proceed. By the end they can check their visa status.

Current system

EMBASSY OF THE STATE OF PALESTINE

Astana, Kazakhstan

HOME
EVENTS
GALLERY
CONTACT US

Palestine

PLO

Jerusalem

Refugees

The Wall

Peace Process

Declaration

Embassy of the State of Palestine in the Republic of Kazakhstan, Astana

Ambassador of the State of Palestine to the Republic of Kazakhstan

Dr. Montaser Abu Zaid

Tel: +7 (7172) 28 75 57
+7 (7172) 28 75 62

Fax: +7 (7172) 28 75 55

e-Mail: pal.embkz@gmail.com

Consular section : Head of Consular section

Mr. Hazem Zweid

EMBASSY OF THE STATE OF PALESTINE

Astana, Kazakhstan

HOME
EVENTS
GALLERY
CONTACT US

Palestine

PLO

Jerusalem

Refugees

The Wall

Peace Process

Declaration

Embassy of the State of Palestine in the Republic of Kazakhstan, Astana

Ambassador of the State of Palestine to the Republic of Kazakhstan

Dr. Montaser Abu Zaid

Tel: +7 (7172) 28 75 57
+7 (7172) 28 75 62

Fax: +7 (7172) 28 75 55

e-Mail: pal.embkz@gmail.com

Consular section : Head of Consular section

Mr. Hazem Zweid

Mob: +7 (7017) 94 32 36
+7 (7071) 19 42 36

Working days: Monday to Friday
Working time: 9:30AM to 5:30PM

Address:
Kabanbay-Batyr Str., 28, Block 4, entrance 1, 010000,
Astana, Kazakhstan

Copyright Embassy of the State of Palestine © 2015 - 2016
Website by [Zhunissova](#)

4 CONCLUSION AND RECOMMENDATIONS

Critical Success Factor

Web Portal For receiving and analyzing Visa Applications has met all the objectives of the project and has been delivered within the time frame.

Waterfall methodology has ensured that user requirements were absorbed during the development phase and integrated into the project with no complication or disruption to other components of the system.

The system gives the ability to the users to submit their documents for a visa without visiting the embassy.

Future enhancement

As for a future enhancement researcher can develop **Web Portal For receiving and analyzing Visa Applications** application for Androids. This will allow users to use the system anywhere and anytime. Besides that can add live chat in order to communicate with the consular department online and live. There will be one consultant where he/she to answer any questions concerning visa issues.

Apart from that can implement SMS notification for the users, where each registered user will receive a notification message to their mobile number. Notification will contain status and the date to collect a passport. Last but not least researcher would like to add few languages to allow the users change the language that they are comfortable with. English will be set as a main language, Russian, Kazakh and Arabic languages as optional languages.

Lessons learnt/ Problems faced

During project researcher has gained a lot of skills and experiences. There were some problems while conducting a primary research. While gathering the information it was hard to identify the group of users for the system. On the other hand how many people to survey and interview. Moreover was difficult to decide if the **WPFVA** needs to have a payment session for a visa or no. Other difficulty is making the system web based or mobile application. Which programming language will be more an appropriate for the web based system, Html that is standard or PHP that is more professional. In order to overcome these, need to do some additional research on programming languages and gather more information about weight loss and obesity.

Besides above stated, there was this problem where every student faces - time management. Due to time constraint researcher could not implement some of the additional functions such as live chat where user can talk to a consultant live. It was challenging to the researcher to develop this feature. And last but not least is was hard to identify the language of the system, since this system implemented for researcher's country, this is because in Kazakhstan there are not many similar systems as **Web Portal For receiving and analyzing Visa Applications**. The researcher will use the skills that were gained from this project to avoid such problems in future projects.

Conclusion

In conclusion, the researcher had some findings during investigation report. It helped the researcher to gain and widen the knowledge. The researcher has learned how to control and manage the time during project, types of methodologies, tools and techniques, programming languages and system design techniques.

As a nature of challenge researcher faced several issues during project. There was confusion on which group of people can use Web portal. It was also not easy to find a surveyors. Overall result of survey was as expected.

Therefore, the researcher will make all the improvement on the project. Due to time constraint the researched might not meet all the requirements, but will be added in future final documentation. At the end of the project the researcher hopes to achieve all project scope as stated in the project.

References

- 1 Learn with the smart method. 2011. The waterfall development methodology. [Online]. Available at http://www.learnaccessvba.com/application_development/waterfall_method.htm . [Accessed on 7th of May 2013].
- 2 Marzuki B. Khalid. Research methodology. [book]
- 3 Moser C.S., Kalton G. Questionnaires. [Book].
- 4 Shafiee. 2012. Waterfall Methodology. [Online]. Available at <http://comp.utm.my/shafiee92/waterfall-methodology/> . [Accessed on 19th of May 2013].
- 5 Selecting a development approach. 2005. System Development Methodologies. [Online]. Available at <http://www.cms.gov/Research-Statistics-Data-and-Systems/CMS-Information-Technology/XLC/Downloads/SelectingDevelopmentApproach.pdf> . [Accessed on 7th of May 2013].

АКТАЕВА А.¹, БАЙКЕНОВ А.², НИЯЗОВА Р.³, БИЖИГИТОВА Д.⁴

ОБЗОР МЕТОДИК ИДЕНТИФИКАЦИИ ПОЛЬЗОВАТЕЛЕЙ СОЦИАЛЬНЫХ СЕТЕЙ НА ОСНОВЕ ОНТОЛОГИИ

¹ *Алматинский Технологический Университет, г.Алматы, Казахстан*

² *Алматинский Университет Энергетики И Связи, г.Алматы, Казахстан*

³ *Евразийский Национальный Университет им. Л.Гумилева, Астана, Казахстан*

⁴ *Алматинский Технологический Университет, г.Алматы, Казахстан)*

Проблемам обеспечения национальной безопасности всегда была характерна особая актуальность, поскольку здесь затрагиваются вопросы формирования потенциала развития страны, а также стабильности и благополучия общества.

В связи с глобализацией и информатизацией и небывалым НТП, а также привлечением всего человечества обострились экономические, сырьевые и др. проблемы на мировом уровне. В современном многоликом и динамичном мире проблемы безопасности приобретают принципиально новые черты, теперь они выходят далеко за рамки предотвращения войн и конфликтов [2,5,6,8,9,15].

Знание динамики и особенностей процессов развития процессов глобализации общества необходимое условие для принятия мер по наиболее полной информатизации общества. Интенсивное развитие и использование современных информационных технологий уже в настоящее время привели к серьезным качественным изменениям во всех сферах общественной жизни. Одним из приоритетов модернизации экономики страны является развитие стратегических информационных технологий, являющихся наиболее важной составляющей процесса информатизации общества.

Новейшие Интернет-технологии теперь широко применяются в различных областях науки, техники и бизнеса и стремительно перешли в все сферы человеческого бытия: бытовая сфера – умный дом, бытовые приборы с помощью средств Интернет-технологий; средства связи – IP-телефония, IP-телевидение; гаджеты – сотовые телефоны, и др. Это, в свою очередь, обуславливает необходимость контроля субъектов информационных процессов для идентификации возможных направлений информационного воздействия на пользователей сети Интернет. В связи с тем, что порядка 80% информации в сетях Интернет представлено в текстовом виде, возникает необходимость в разработке средств идентификации пользователей на основе методов математической лингвистики и семантики, предназначенных для обеспечения информационной безопасности важных объектов политической, социально-экономической, оборонной, культурной и других сфер деятельности от внешних и внутренних угроз модификации информации [2, 5,8,9,15,14].

ИКТ-компоненты, как элементы системы информатизации общества непосредственно и активно влияют на состояние экономической, экологической, энергетической, транспортной, продовольственной, криминогенной, информационной и других составляющих комплексной структуры безопасности страны. Таким образом, ИКТ компоненты являются одним из системообразующих факторов жизни современного общества, и влияние информационной безопасности и на все стороны жизни общества с течением времени будет только возрастать и оказывают информационно - манипуляторное воздействие на пользователей. Для мониторинга и контроля комплексной структуры безопасности используется комплексный критерий (см. рис.1)



Рис.3 - Критерий параметров манипуляторного воздействие на пользователей

С учетом данного комплексного критерия можно выделить группы манипуляторных приемов, имеющих наиболее универсальный характер, с высокой частотой встречаемости в различных технологиях информационных войн, имеющих сферу применения в таких информационно-коммуникативных ситуациях, как публичные дискуссии и групповые обсуждения, выступления на митингах и демонстрациях, в средствах массовой коммуникации, в межгрупповых и межличностных конфликтах, характеризующиеся достаточно высокой действенностью и влиянием на психику человека. Данные приемы характеризуются высокой степенью выраженности по всем параметрам манипуляторного воздействия на пользователей социальных сетей [2, 5,8,9,15].

Основные методы идентификации вычисления сетевых данных пользователей, использующие технические характеристики для применения технологии обработки сообщений для защиты от модификации информации перечислены на рисунке 2:



Рис.2 - Методы идентификации вычисления сетевых данных пользователей

Большое количество Интернет ресурсов и сервисов, таких как форумы, порталы, интернет-магазины, сталкиваются с различными проявлениями проблемы манипуляции и искусственного формирования общественного мнения, путем «организации» целенаправленных тематических диалогов, в которых ряд пользователей имеют несколько учетных записей. В таблице 1 приведены основные приемы информационно-психологического воздействия на массовое сознание [2, 5,8,9,15].

Возможность использования порталов и сайтов для распространения информации и недостаточная функциональность механизмов идентификации и аутентификации пользователей, оставляющих сообщения, определяет ряд направлений совершенствования систем защиты и систем мониторинга информационной безопасности ИТКС.

Задача выделения информации (information extraction) отличается от задачи поиска информации (information retrieval). Системы выделения информации можно разделить на следующие четыре типа по степени участия эксперта в создании и адаптации системы:

1. *Настраиваемые вручную*, где Пользователь задает на некотором языке правила выделения информации из конкретных сайтов.
2. *С обучением*. Пользователь вручную размечает обучающую выборку документов, которая используется для построения модуля выделения информации.

3. *С частичным обучением.* Пользователь не размечает всю обучающую выборку, а лишь предоставляет некоторую дополнительную информацию, например, выбирает шаблон из вариантов, представленных системой, и отмечает данные, которые нужно выделить.

4. *Без обучения.* Система автоматически размечает обучающую выборку и создает модуль выделения информации полностью без участия пользователя [2, 5,8,9,10,11,13].

Таблица 1 - Основные приемы информационно-психологического воздействия на массовое сознание пользователей

№ пп	Основные приемы информационно-психологического воздействия	Пояснение
1	name calling	Приклеивание или навешивание ярлыков»
2	glittering generality	Сияющие обобщения» или «блистательная неопределенность»
3	transfer	Перенос» или «трансфер
4	testimonial	Ссылка на авторитеты, по рекомендации, свидетельствование
5	plain folks	Свои ребята или игра в простонародность
6	card stacking	Перетасовка или подтасовка
7	band wagon	Общая платформа или «фургон с оркестром»

В последнее время все большее внимание исследователей обращается к задаче выделения информации из Всемирной паутины. Согласно исследованию Рассела и Норвига, выделение информации находится «посередине» между поиском информации (information retrieval), который заключается в подборе документов, отвечающих запросу пользователя, и пониманием смысла текста, подразумевающим глубокий анализ текста с целью выявления его семантики [1, 3, 4, 9,10,11,13].

Другой особенностью Web-документов является тот факт, что информация в них обычно генерируется автоматически из некоторых баз данных, что превращает процесс выделения информации из таких текстов в процесс «декодирования». Наконец, нельзя не отметить огромное количество данных, содержащихся в Интернет и их неоднородность.

Перечисленные особенности обуславливают набор методов, обычно применяемых при выделении данных из Всемирной паутины. Если в традиционных системах выделения информации из неструктурированных текстов обычно применяются методы обработки естественного языка, такие как словари и грамматики, то в системах выделения информации из Интернет чаще используются алгоритмы машинного обучения и выделения шаблонов, которые опираются на синтаксические свойства и визуальную структуру Web-страниц (см.рис.3).

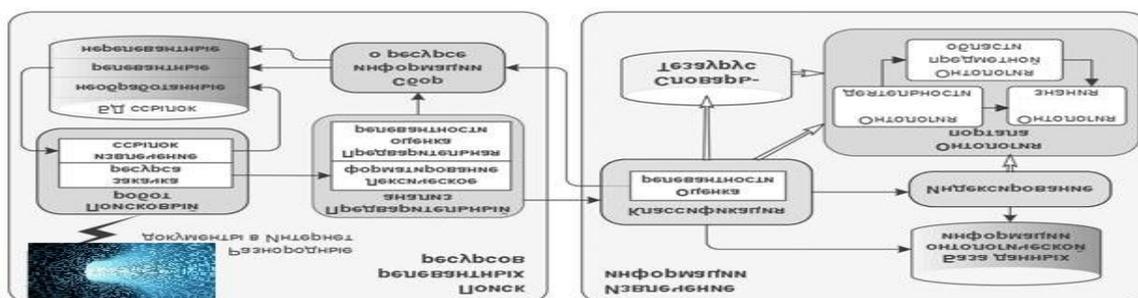


Рис. 3 - Схема функционирования информационного объекта [5,8,9]

Для выделения данных из Всемирной паутины может использоваться обучение на подготовленной выборке источников, и/или эвристические модели, которые могут быть

основаны, например, на использовании заданных лексико-синтаксических шаблонов или же на онтологиях.

Онтологии используются для формального описания области знания. Согласно классическому определению, онтология – формальная, явная спецификация общей концептуализации. Это способ представления знаний с помощью конечного множества понятий и отношений между ними. Связи между понятиями бывают следующих типов (см.рис.4). Понятие, или сущность - это класс индивидуальных объектов, или экземпляров [5,8,9].

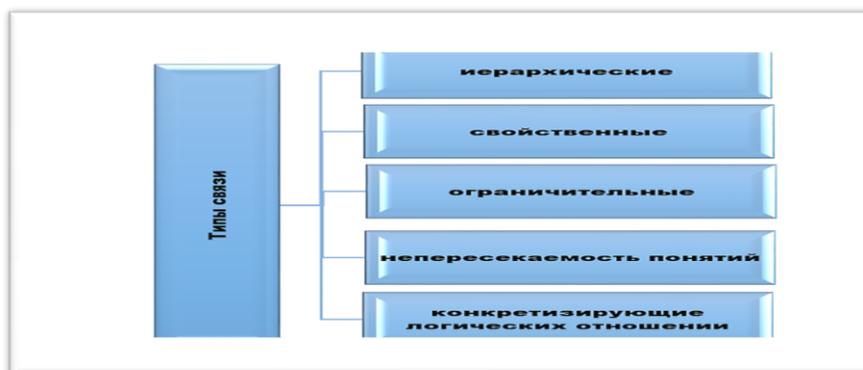


Рис.4 – Типы связи понятий онтологии

А направление извлечения информации с помощью онтологий выделилось из общей задачи выделения информации сравнительно недавно, но уже отмечено как перспективное направление развития систем выделения информации.

Представление знаний с помощью онтологий в информационно-поисковых системах позволяет выполнять сложные объектно-ориентированные запросы, которые можно представить в виде графа. Практически выполнение такого запроса сводится к поиску подграфа в графе.

На основе онтологии информационно-поисковая система может предложить конкретизировать или, наоборот, расширить запрос, если ему удовлетворяет слишком /мало объектов. Более того, онтология позволяет предложить пользователю самому осуществить обзор области знания с помощью навигации по понятиям, переходя от одного понятия к другому по связям между ними. Во многих системах, основанных на онтологиях, используются следующие методы извлечения информации:

- Языковые правила, сформулированные с помощью регулярных выражений;
- Справочники (gazetteers);
- Методы классификации;
- Построение частичных деревьев синтаксического разбора предложений с семантической аннотацией;
- Анализ HTML/XML-тегов;
- Поиск в Интернет;
- Описание используемых алгоритмов.

Отметим еще несколько преимуществ использования онтологий для представления знаний:

- Гибкость модели данных, которая позволяет сравнительно легко ее изменять и расширять.
- Возможность повторного применения существующих онтологий [5,8,9].

Для задания онтологии предметной области используются классы языков, основанных на математических моделях (см.рис.5).

Использование онтологий для представления знаний имеет ряд преимуществ:

1. Онтология содержит общее видение области знания, что позволяет устранить различия в терминологии, проблемы синонимии и многоязычности [5,8,9].

2. Математический аппарат дескриптивной логики предоставляет средства логического вывода новых фактов на основе существующих.

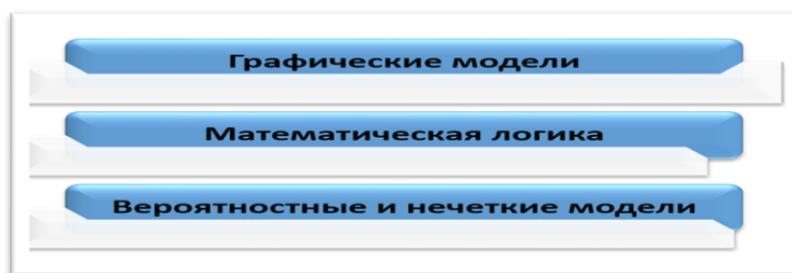


Рис.5 - Классы языков для задания предметной области онтологии

В качестве примера рассмотрим, *математическую модель алгоритма Brainstern*:

пусть W – множество всех слов, которые встречаются во всех документах заданной коллекции Doc , включая ϵ – пустое слово, а PW – множество всех упорядоченных пар слов, то есть $PW = W \times W$. Определим документ d как отображение $d: N \rightarrow W$, которое сопоставляет каждому натуральному числу n слово, стоящее на n -той позиции в данном документе коллекции. Номера позиций, на которых нет слов (после конца документа), отображаются в пустое слово. Аналогично определим абзац p как отображение $p: N \rightarrow W$, которое сопоставляет каждому натуральному числу n слово, стоящее на n -той позиции в данном абзаце. Номера позиций, на которых нет слов, отображаются в пустое слово. Обозначим множество всех абзацев в коллекции через P . Определим рубрику r как произвольное подмножество множества документов, а именно – $r \in 2^{doc}$. Мощность рубрики, как количество документов в ней, будем обозначать через $|r|$. Обозначим множество всех заданных рубрик через R [5,8,9].

Выделение информации традиционно имеет целью найти сведения, которые описывают некоторую область знаний, заданную структурой данных. Онтологии же как раз и представляют собой формальную модель предметной области, выраженную, например, в виде графа понятий и связей, что обобщает иерархическую структуру данных, обычно используемую для заполнения в задаче выделения информации.

Логический вывод новых утверждений позволяет превращать скрытое знание в явное, а также находить противоречия в онтологиях. Следует отметить, что неявное знание, как правило, и содержит смысл явного. Именно в этом контексте принято утверждать, что в рамках семантического подхода компьютер «понимает» смысл информации.

С помощью логического вывода можно автоматически классифицировать экземпляры по классам на основе их атрибутов. Большим достоинством логического вывода является возможность не вводить в онтологию всю информацию, а ввести лишь ее базовую часть, из которой можно вывести все необходимые утверждения.

Также онтологии позволяют вынести процесс согласования на семантический уровень. Это означает, что если разные системы используют одну и ту же онтологию, то есть модель данных, то конкретная реализация системы не имеет значения.

А построение онтологии включает выделение множеств понятий и имен отношений, а также экземпляров этих понятий и связей между ними. Заполнение онтологии подразумевает наличие понятий, связей и имеет целью поиск экземпляров понятий и отношений между ними. Следует отметить, что задачи построения и заполнения онтологий в системах информационной безопасности социальных сетей в настоящее время являются очень актуальными по следующим причинам:

1. Построение и заполнение онтологий требует разработки алгоритмов автоматического выделения информации из текстов на естественном языке. Большая часть информации в Интернет содержится именно в таком виде. Ручная обработка таких данных требует очень больших перманентно увеличивающихся человеческих ресурсов в связи с

колоссальными объемами накапливаемой информации. Именно поэтому интеллектуальные алгоритмы, которые автоматизируют этот процесс, приобретают огромную важность;

2. Заполненная онтология является готовым информационным ресурсом для семантической паутины. Семантическая аннотация позволяет в дальнейшем обрабатывать эту информацию машинами, воплощая концепцию Семантической паутины. Для реализации видения паутины необходимы автоматические средства генерации метаданных;

3. Заполнение онтологии может использоваться для повышения ее качества. Основная идея состоит в том, что в случае, если онтология помогает эффективно выделять необходимую информацию из текстов, можно сделать вывод, что эта онтология адекватно описывает область знания [2, 5,8,9,15].

Построение и заполнение онтологий тесно связано с выделением информации с использованием онтологий. В онтологии должно содержаться как можно больше информации об области знания, в частности, не только иерархия понятий и направлений, но и ассоциативные связи. Алгоритма построения онтологии:

1. Извлекать иерархические и ассоциативные отношения между терминами;
2. Отражать актуальное состояние заданной области научного знания;
3. Точность и полнота извлекаемых терминов и отношений не должна уступать показателям существующих и прошедших апробацию алгоритмов построения онтологии;
4. Существовать возможность получать необходимые обучающие выборки из открытых источников, не затрачивая больших усилий по их обработке;
5. Не требовать большого объема ручного труда экспертов по настройке на заданную предметную область;
6. Источники данных должны находиться в открытом доступе и регулярно обновляться;
7. Иметь модульную архитектуру;
8. Возможность автоматизированной настройки алгоритма на конкретную область знания [5,8,9,15].

Показатели эффективности алгоритмов выделения информации делятся на два класса: *точность-корректность* выделенной информации, полнота - количество выделенной информации по отношению к объему всей доступной информации и *мера избыточности*, а также оценки вычислительных ресурсов (время и память).

Онтологический подход к представлению знаний позволяет применять существующие и прошедшие апробацию алгоритмы выполнения аналитических запросов. Выполнение аналитических запросов к данным обеспечивается в процессе взаимодействия конечного пользователя системы с программной реализацией модели, описывающей область знания. Запрос при использовании онтологий может выполняться автоматически с помощью механизмов логического вывода. В качестве языка запросов к онтологиям можно использовать язык SPARQL. Выбор именно этого языка обусловлен высоким уровнем его развития, зрелости и хорошим потенциалом, что подтверждают следующие факты:

- в 2008 году язык SPARQL получил статус официальной рекомендации консорциума W3C2;

- язык SPARQL не привязан к конкретному программному комплексу, в отличие от других языков запросов к онтологиям;

- для языка SPARQL существует большое число программных реализаций и приложений [2, 5,8,9,15].

Ниже приведены примеры использования языка SPARQL:

*Пример 1. Перечень направлений, которые активно исследуются в рамках интересующей области знания. Интерпретация запроса: «выдать все термины, сопоставленные результатам ... деятельности за последний (2015) год и отсортировать их по убыванию встречаемости в этих результатах». Формализуем его на языке SPARQL. Сначала сформируем множество **Terms**, содержащее все термины (возможно, с повторениями), сопоставленные результатам деятельности за последний год.*

```

SELECT ?term WHERE {
  ?term a cs:term .
  ?res a swrc:Result .
  ?res swrc:isAbout ?term . ?res swrc:year 2015 .
}

```

Полученное множество терминов *Terms*, необходимо отсортировать по убыванию количества повторений каждого уникального элемента. Термины, расположенные в начале отсортированного списка, и определяют направления, которые активно исследуются в рамках интересующей области знания.

Пример 2. Перечень направлений интересов пользователя. Интерпретация запроса: «выдать список всех терминов, связанных с результатами деятельности, в авторах которых есть заданный пользователь User X». Далее представлена формальная запись этого запроса на языке SPARQL.

```

SELECT DISTINCT ?term WHERE { ?term a cs:term .
  ?res a swrc:Result .
  ?res swrc:isAbout ?term. ?res dc:creator User X .
}

```

Пример 3. Перечень форумов, посвященных интересующему направлению.

Перепишем этот запрос следующим образом: «выдать список форумов, связанных с терминами заданного направления $T = \{t_1, \dots, t_n\}$ ». Интерпретация запроса на языке SPARQL:

```

SELECT DISTINCT ?forum
WHERE {
  ? forum a swrc: Forums .
  { ? forum swrc:isAbout t_1 }
  UNION { ?forum swrc:isAbout t_2 } ...
  UNION { ?forum swrc:isAbout t_n } . }

```

Связь между запросами, формальной моделью разрабатываемой системы и кодом запросов на языке SPARQL позволяет контролировать влияние:

- модификаций множества принятых в системе запросов и используемых онтологий на программный код системы;
- модификаций программного кода системы на используемые онтологии и рассматриваемые запросы [5,8,9].

В таких случаях для идентификации пользователя возможно применять методы математической лингвистики. При этом возникает необходимость в разработке:

- модели идентификатора пользователя портала сети Интернет на основе кортежа лингвистических признаков короткого сообщения;
- метода создания компонентного профиля пользователя портала сети Интернет, базирующегося на модели идентификатора, содержащего кортеж лингвистических признаков;
- методики идентификации пользователя портала сети Интернет на основе компонентного профиля.

Метод создания компонентного профиля пользователя сети Интернет предполагает реализацию ряда шагов:

- обработка пользовательских сообщений в рамках Интернет портала;
- разбор сообщений по частям речи с последующим применением шаблонов (синтаксических паттернов) для выделения наиболее распространенных конструкций;
- лексикографический анализ сообщения и выделение конструкций в соответствии с описанными шаблонами и сбор статистики об использовании знаков препинания и специальных символов;

• выделение лексических конструкций на основе слов и словоформ языка, а также выявление тематических специальных слов и выражений, характерных для аудитории конкретного форума.

На каждом шаге используются специализированные словарные базы данных и анализаторы текстовой информации, описанные в предыдущих разделах.

Заключение. Онтологии же как раз и представляют собой формальную модель предметной области, выраженную, например, в виде графа понятий и связей, что обобщает иерархическую структуру данных, обычно используемую для заполнения в задаче выделения информации в сетях Интернет.

На основе исследования предметной области информационной безопасности могут быть построены математические модели и алгоритмы, разработаны опирающиеся на онтологии архитектурные и технологические решения для создания системы пополнения и хранения, анализа и выдачи по запросу информации, характеризующей результаты деятельности *Пользователя* информация на Web-страницах в Интернет.

С использованием онтологий возможно формальное описание запросов к системе, создающее гарантии их вычисления и дополнительные возможности для эффективной верификации кода системы на всех этапах ее жизненного цикла, а также позволяет использовать для учета и анализа информации при моделировании и распознавании профиля *Пользователя* в социальных сетях в условиях информационной войны.

Литература

1. Jowett G. S., O'Donnel V. Propaganda and Persuasion. — Newbury Park, 1992
2. Кучумов Д.О. Проблема региональной информационной безопасности // Безопасность Евразии. – Москва, 2007, № 4. – С.53-56.
3. Russell Stuart J. Artificial Intelligence: A Modern Approach / Stuart J Russell, Peter Norvig. — 2 ed. — Pearson Education, 2003
4. Cowie, Jim. Information Extraction / Jim Cowie, Yorick Wilks // A Handbook of Natural Language Processing: Techniques and Applications for the Processing of Language as Text / Ed. by Robert Dale, Hermann Moisl, Harold Somers. -New York, USA: Marcel Dekker, New York, 2000.
5. Голомазов Д. Д. Методы и средства управления научной информацией с использованием онтологий //05.13.17-теоретические основы информатики диссертация на соискание ученой степени кандидата физико-математических наук.- Moscow– 2012
6. Урсул А.Д. Природа информации: философский очерк. – 2-е изд. – Челябинск, 2010, 231 с
7. Горбунов-Посадов М.М. Интернет-активность как обязанность ученого / М.М. Горбунов-Посадов // Информационные технологии и вычислительные системы. — 2007, № 3, 88–93с.
8. Сухопаров М.Е. Методика идентификации пользователей порталов сети интернет на основе методов математической лингвистики // 05.13.19 – Методы и системы защиты информации, информационная безопасность - Диссертация на соискание ученой степени кандидата технических наук. -Санкт-Петербург – 2015
9. Aktayeva A., Galyieva N., Baiman G & etc Cognitive Ontology of information security priorities in social networks // Proceedings of the International Scientific and Practical Conference Open Semantic Technologies for Intelligent Systems – OSTIS-2017, 16-18 Feb., ISSN 2415-7074, <http://proc.ostis.net/eng/main.html>
10. Голомазов Д.Д. Выделение терминов из коллекции текстов с заданным тематическим делением / Д.Д. Голомазов // Информационные технологии. — 2010. — № 2. — С. 8–13.
11. Васенин, В.А. и др. Использование семантических технологий для обнаружения грид-ресурсов // Программная инженерия. — 2011, № 7, 2–8 с.
12. Yeong W. Lightweight Directory Access Protocol// Search. -1995, Vol. 2251, №1777, 1–11pp.
13. Боровикова О.И. Онтологический подход к построению систем информационной поддержки научной и производственной деятельности // Материалы Всероссийской конференции с международным участием «Знания – Онтологии – Теории» (ЗОНТ–09). — Т. 2. — Новосибирск: Институт математики им. С.Л. Соболева СО РАН, 2009, 93–102 с.
14. Манжуева О.М. Философская парадигма информационной безопасности //09.00.08 – Философия науки и техники диссертация на соискание ученой степени доктора философских наук . - Moscow– 2016
15. Aktaeva A., etc. Development of a mathematical model of information warfare // International Journal of Open Information Technologies, vol. 2, №11, 2014, 28-33 pp., www.injoit.org, <https://doaj.org/toc/2307-8162>

И.А. ӘБЕЕВА, Б.Ж. ЕРГЕШ

АСТАНА ҚАЛАСЫНДАҒЫ МЕМЛЕКЕТТІК ҰЙЫМДАРДЫҢ ОНТОЛОГИЯСЫН ҚҰРУ

(Л.Н. Гумилев атындағы Еуразия ұлттық университеті, Астана, Қазақстан)

Онтология деп әр түрлі ережелер мен қарым-қатынас арқылы бір-бірімен байланысқан класстардан тұратын сызбаны айтады. Онтология кез келген білім ортасын формалды түрде ерекше формада көрсетеді. Онтологияның ерекшелігі - білім түсінігінің формалды құрылымын жасауы, себебі ол компьютерлік өндеуді едәуір жеңілдетеді [2].

Білімге негізделген ұсынымдар жүйесі бойынша кәдімгі мазасыздық инженерлі білім аясындағы сипаттамаларға негізделеді, ол жүйенің үлестірудегі құнын арттырады. Білім базасын жасау үшін алдымен таңдалған пәндік облыс бойынша онтология құру қажет етіледі.

Білім базасы кез келген жүйеде орынды әрі тиімді, олардың қолданылу саласы шектелмеген. Адамға көмек керек жерде жасанды интеллект қажеттілікті толықтырып отыруы қажет. Жасанды интеллект арқасында ғаламтор қолданылу кезіндегі білім шекарасының мүмкіндігі мың есеге кеңейтіліп жатыр [1].

Білім базасы әртүрлі мақсаттағы интелектуалды жүйелердің кілттік компоненттерінің бірі. Бұл компонентті дайындау қиындықты келтіретін және ұзақ уақытты процесс [2, 5].

Білім базасын жасау кезінде білім сақтауды қамтамасыз ету мүмкіндігі ғана емес, сонымен қоса әзірлеуші ұжым арасында база құру мен оны өзгерту мүмкіндігін үлестіру керек.

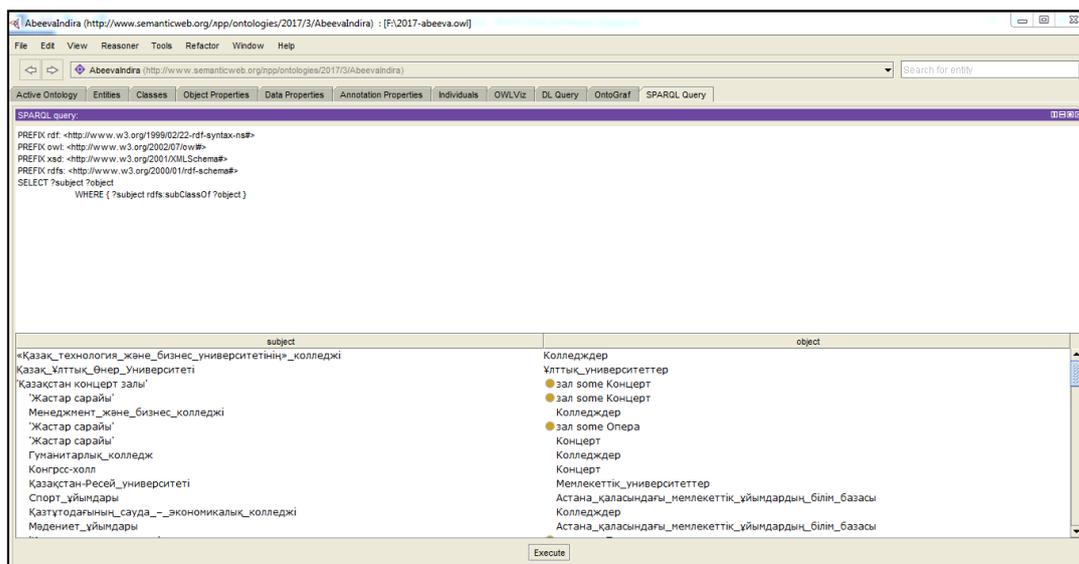
Білім базасының мәліметтер базасынан негізгі айырмашылығы – үлкен шығармашылық мүмкіндіктердің болуы. Мәліметтер базасындағы деректер әдетте пассивті болып келеді : олар бар болады немесе жоқ болады. Білім базасы болса белсенді түрде жетпеген ақпаратты толықтыруға тырысады. Білім базасы логикалық шешім механизмі үшін кіру ағыны болып табылады[4].

Мемлекеттік ұйымдардың онтологиясын Protégé редакторында құру

Protege редакторында[3] мемлекеттік ұйымдардың онтологиясы құру процесі сипатталды. программасын іске қосып, жаңа жобасын құруды бастаймыз. Параметрлер терезесінде, Protege файлдарды таңдаймыз. Экранда бізге қажет жұмыс терезесі пайда болады. Онтология құруға алғашқы қадамында класстар құру қажет. Біздің барлық класстар Class Browser терезесінде көрсетіледі. Жаңа класс құру үшін «Create Class» батырмасын қолданамыз.

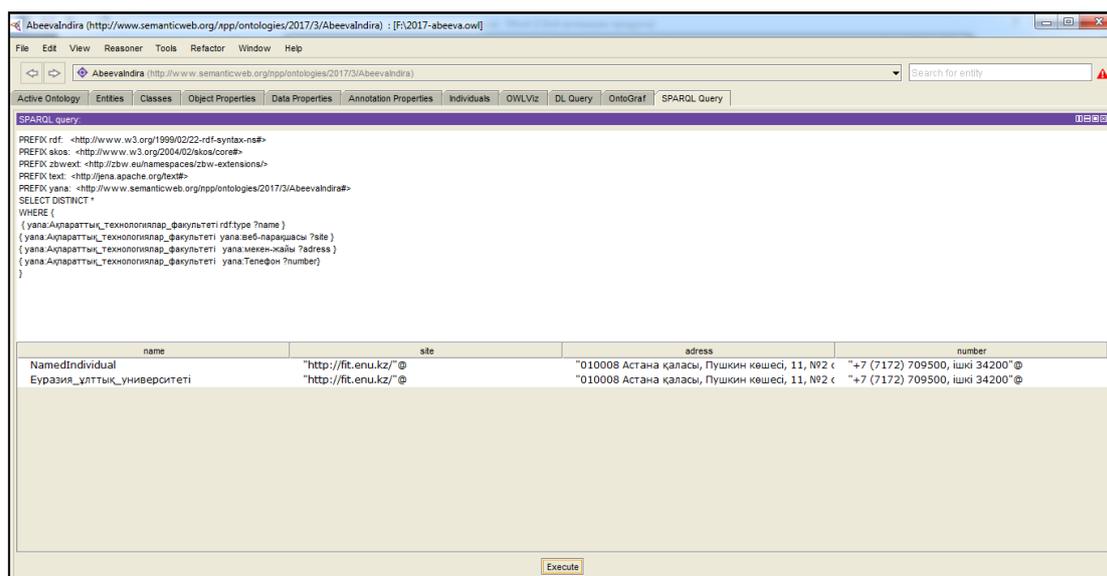
Онтологияның ерекшелігі- білім түсінігінің формалды құрылымын жасауы, себебі ол компьютерлік өндіруді едәуір жеңілдетеді.

«Астана қаласындағы мемлекеттік ұйымдардың білім базасы» классының: «Білім ұйымдары», «Мәдениет ұйымдары», «Спорт ұйымдары» подкласстары таралады. Солардың ішінде «Білім ұйымдары» 5 астыңғы классқа бөлінеді «Балабақша», «Жоғарғы оқу орындары», «Колледж», «Лицей» «Мектеп» болып. Және олар индивидке жеткенше өз іштерінен классқа тарала береді. Оны (Сурет 1) көруге болады.



Сурет 3. SparQl query де сұраныс жасау

Сонымен ары қарай «Еуразия ұлттық университетінің» индивиді «Ақпараттық технологиялар факультетінің» веб-парақшасын, мекен-жайын, байланыс телефон номерін іздеуге болады (Сурет 4).



Сурет 4. SparQl query де «Ақпараттық технологиялар факультетіне» сұраныс жасау

Және тағы бір сұраныста «Астана қаласындағы мемлекеттік ұйымдардың білім базасының» типтерін, класстарын және индивиттерін шығардым (Сурет 5).

ind	type	class
БАЗАЛЫҚ ДАЙЫНДЫҚ ФАКУЛЬТЕТІ	Қазақ экономика_қаржы_және_халықаралық_сауда_университ	Мемлекеттік_университеттер
КАШЫҚТҚТАН ОҚЫТУ ФАКУЛЬТЕТІ	Қазақ экономика_қаржы_және_халықаралық_сауда_университ	Мемлекеттік_университеттер
КОЛДАНАЛАЙ ҒЫЛЫМДАР ФАКУЛЬТЕТІ	Қазақ экономика_қаржы_және_халықаралық_сауда_университ	Мемлекеттік_университеттер
Технология_факультеті	Қазақ технология_және_бизнес_университеті	Мемлекеттік_университеттер
ЭКОНОМИКА_ ФАКУЛЬТЕТІ	Қазақ экономика_қаржы_және_халықаралық_сауда_университ	Мемлекеттік_университеттер
Сәулет-құрылыс_факультеті	Еуразия_ұлттық_университеті	Ұлттық_университеттер
Әскери_кафедра	Еуразия_ұлттық_университеті	Ұлттық_университеттер
Журналистика_және_саясаттану_факультеті	Еуразия_ұлттық_университеті	Ұлттық_университеттер
Халықаралық_қатынастар_факультеті	Еуразия_ұлттық_университеті	Ұлттық_университеттер
Әлеуметтік_ғылымдар_факультеті	Еуразия_ұлттық_университеті	Ұлттық_университеттер
Механика-математика_факультеті	Еуразия_ұлттық_университеті	Ұлттық_университеттер
Заң_факультеті	Еуразия_ұлттық_университеті	Ұлттық_университеттер
Келік-энергетика_факультеті	Еуразия_ұлттық_университеті	Ұлттық_университеттер

Сурет 5. Сұраныста индивид, тип және класстары көрсетілген

Бұл жұмыста Астана қаласындағы мемлекеттік ұйымдардың жоғарғы деңгейде онтология құрылды. Бағдарлама барысында Астана қаласында орналасқан мемлекеттік ұйымдар енгізілді. Ұйымдар 3 ке бөлінді: білім, мәдениет және спорт. Құралған онтологияға SparQl query тілінде сұраныстар жасау жолдары сипатталды.

Қолданылған әдебиеттер тізімі

- 1 Искусственный интеллект. Кн. 2: Модели и методы: Справочник / Под ред. Э.В. Попова. М.: Ра-дио и связь, 1990. 303 с.
- 2 Методы представления знаний: Метод. указ. / Сост. И. Л. Коробова. Тамбов: Изд-во Тамб. гос. техн. ун-та, 2003. 24 с.
- 3 Руководство по Protege 4.2 beta (разделы 4.1-4.7) <https://docs.google.com/document/d/1kKOom1gBV5uvji62Z0fhQpbhcINBNAPZIPZmjk-n-zU/edit>
- 4 <http://www.tstu.ru/book/elib/pdf/2003/korobova.pdf>
- 5 Горбань А.Н., Дунин-Барковский В.Л., Кирдин Н., Миркес Е.М., Новоходько А.Ю., Россиев Д.А., Терехов С.А., Сенашова М.Ю. Нейроинформатика. <http://www.bmstu.ru/facult/iu/iu4/rus/stst/book2/ann.htm>
- 6 Вопросы приближения функций: Метод. указ. / Авт.-сост.: Ю.В. Литовка, А.В. Романенко, И.Л. Коробова. Тамбов: Изд-во Тамб. гос. техн. ун-та, 2001. 20 с.

С.Н.БАЙБЕКОВ, Л.С.ИСКАЛИЕВА

**РАЗРАБОТКА НОВОГО МЕТОДА БЕРЕГОВОЙ ЛИНИИ ДЛЯ
СЕКТОРАЛЬНОГО РАЗДЕЛЕНИЯ МОРСКОЙ ТЕРРИТОРИИ МЕЖДУ
ПРИБРЕЖНЫМИ ГОСУДАРСТВАМИ**

(Институт химии угля и технологии, г.Астана, Казахстан)

Конвенция ООН по морскому праву 1982 года [1] (в дальнейшем – Конвенция) состоит из 17 частей, включающих 320 статей и 9 приложений, и регламентирует, по существу, все политические, юридические, экономические, экологические, технологические и другие аспекты, связанные с использованием морских пространств.

В ней однозначно классифицированы все основные критерии морских зон прибрежного государства, которыми являются:

-**территориальное море**, ширина которого устанавливается до предела, не превышающего двенадцати морских миль, отмеряемых от исходных линий (ст.5). Из ст.5 и ст.7 настоящей Конвенции следует, что *исходной линией* является линия наибольшего отлива вдоль берега;

-**прилежащая зона**, которая не может распространяться за пределы двадцати четырех морских миль от исходных линий (ст.33);

-**исключительная экономическая зона**, ширина которой не должна превышать 200 морских миль, отчитываемых от исходных линий (ст.57);

-**континентальный шельф**, который должен находиться не далее 350 морских миль от исходных линий или не далее 100 морских миль от 2500-метровой изобаты и т.д. (ст.76).

Следует сказать, что всем островам (за исключением искусственных островов и сооружений, также скал, которые не пригодны для поддержания жизни человека или для самостоятельной хозяйственной деятельности) предоставляются территориальное море, прилежащая зона, исключительная экономическая зона и континентальный шельф (ст.121). Причем они определяются на тех же основаниях, которые применимы и к другим сухопутным территориям.

В Части 1-6 настоящей Конвенции регламентированы все права и юрисдикции прибрежного государства, права и свободы других государств на эти зоны, так же и их обязанности. Следует сказать, что Конвенция достаточно полно устанавливает внешнюю границу вышеуказанных зон в целом. Что касается определения границ их секторального разделения, т.е. делимитации, между соседними государствами с противоположными или смежными побережьями, то методика делимитации их границ указаны в статьях 15, 74 и 83.

Статья 15 этой Конвенции имеет следующую формулировку:

*- если берега двух государств расположены один против другого или примыкают друг к другу, то, ни то, ни другое государство **не имеет права**, если только между ними не заключено соглашение об ином, распространять свое территориальное море за **срединную линию**, проведенную таким образом, что каждая ее точка является равноотстоящей от **ближайших точек** исходных линий, от которых отмеряется ширина территориального моря каждого из этих двух государств.*

Ст.ст. 74 и 83 этой Конвенции 1982 года посвящены проблемам делимитации соответственно исключительной экономической зоны и континентального шельфа, и имеют почти идентичную формулировку. Эти статьи имеют чисто рекомендательный характер и призывают решить проблему делимитации континентального шельфа и исключительной экономической зоны путем соглашения на основе международного права в целях достижения справедливого решения. С другой стороны если в течение разумного срока между ними не может быть принято разумное соглашение, то сторонам рекомендуется предпринимать все

усилий для того, чтобы достигнуть временной договоренности практического характера и в течение этого переходного периода не ставить под угрозу достижение окончательного соглашения [1].

Проблема делимитации прилегающей зоны между государствами с противоположными и смежными побережьями осталась открытой, она не нашла отражение в Конвенции ООН по морскому праву 1982 года.

Заметим, что методика делимитации территориального моря, приведенная в ст.15, имеет обязательный характер (см. слова «не имеет права»).

Следует сказать, что в Конвенции отсутствует конкретный механизм делимитации прилегающей зоны, исключительной экономической зоны и континентального шельфа между соседними государствами. Быть может, поэтому для этих случаев иногда использует методику, приведенную в статье 15, в которой срединная линия играет роль границы делимитации морских зон между прибрежными государствами. Этот метод, который обычно называется *методом срединной линии*, является предметом исследования настоящей работы.

Из формулировки метода срединной линии (ст.15) видно, что для определения точек границ секторального разделения морских зон (например, территориального моря) между прибрежными соседними государствами рассматриваются только ближайшие точки исходных линий берегов каждой страны, а остальные точки в расчет не берутся. Почему? В принципе все точки исходной линии должны иметь определенную значимость. Каждая точка исходной линии, независимо от того является ли она самой ближайшей или отдаленной точкой, должна внести свой вклад в процесс разделения морских зон между прибрежными государствами. Если рассмотреть этот вопрос с «юридической» точностью, то соседние точки должны быть не менее «весомыми», чем самая ближайшая точка. Мы убеждены, что рассмотрение только ближайших точек без рассмотрения других соседних к ним точек, также удаленных точек, совокупность которых определяет конфигурацию исходной (береговой) линии, неизбежно должно привести к искаженным результатам. Это можно проиллюстрировать на следующих простейших примерах секторального разделения морских зон.

Для удобства рассмотрим случай островных государств.

Пример 1. Пусть принадлежащие разным государствам два острова прямоугольной формы одинаковой ширины и разной длины, как это показано на рис.1, находятся рядом. Предположим, что левый остров принадлежит государству *A*, а правый остров - государству *B*. Согласно п.2 ст.121 Конвенции, эти острова должны иметь свои территориальное море, прилегающую зону, исключительную экономическую зону и континентальный шельф. Если расстояние между этими островами меньше 24 мили, то нужно определить границу разделения территориального моря между государствами *A* и *B*. Если расстояние между этими островами порядка 200 мили, тогда предметом делимитации является исключительная экономическая зона и т.д.

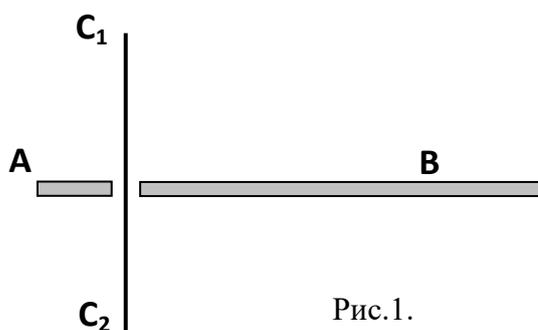


Рис.1.

Если применяя для этого случая метод, приведенный в статье 15 Конвенции, построим среднюю линию, то она имеет вид прямой C_1C_2 . Эта прямая проходит ровно по середине островов A и B . Причем, морская зона, лежащая с левой стороны прямой C_1C_2 , должна принадлежать государству A . Морская зона, находящаяся с правой стороны этой прямой, должна принадлежать государству B . В зависимости от расстояния до островов, эти зоны могут быть территориальным морем или прилегающей зоной и т.д.

Результат такого разделения приведет к следующим двум случаям.

Первый случай. Такое деление вызывает разумное недовольство второго острова B , который по площади на нескольких порядков больше чем первый остров A , так как при этом получаемая им территория, лежащая с правой стороны срединной линии, равна территории, отводимой острову A .

Второй случай. Если обязать эти островные государства проводить требующее значительную финансовую сумму мероприятие по предотвращению, например, экологического загрязнения морской территории, то в этом случае справедливое недовольство будет выражаться страной A . Так как она должна проводить данное мероприятие по такой же площади морской территории, что страна B , хотя площадь ее островной территории на нескольких порядков меньше территории острова B . В результате не исключено, что получаемая граница разделения не удовлетворяет ни ту, ни другую сторону.

Пример 2. Теперь рассмотрим этот же случай, показанный на рис.1, с незначительным изменением. Пусть в этом случае эти два острова, будут расположены перпендикулярно друг к другу, как это показано на рис.2. Остальные параметры островов (расстояние между ними, длина и ширина островов, также и другие характеристики) пусть остаются без изменения.

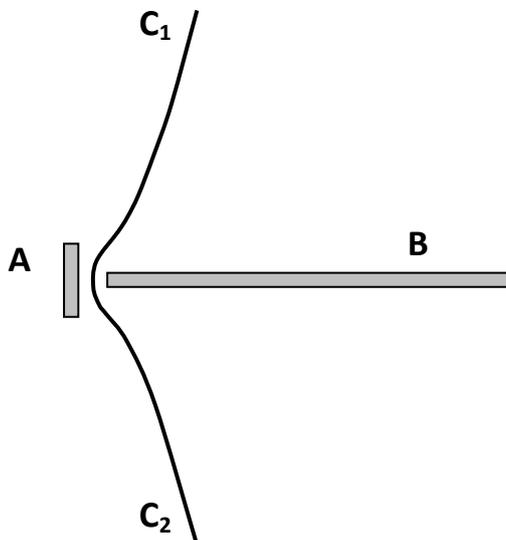


Рис.2.

Для определения границу разделения морской территории между островами A и B применим метод срединной линии. Полученная граница имеет вид кривой C_1C_2 (рис.2). В отличие от 1-го примера, в этом случае граница разделения морских зон (т.е. срединная линия) имеет совершенно другой вид, она значительно изогнута в сторону острова B . Т.е. отводимое морское пространство острову B становится значительно меньше, чем даже в первом случае. Такое разделение, очевидно, является не справедливым. При этом то, что государство B будет выражать справедливое недовольство, не вызывает сомнения.

Следует сказать, что эти простые примеры являются убедительным и наглядным доказательством несовершенства метода срединной линии.

Анализируя множество подобных примеров, также можно убедиться в том, что применение метода срединной линии для реальных случаев, когда береговые линии имеют чрезвычайно сложную конфигурацию, может привести к картинам, искаженным до неузнаваемости.

Мы убеждены, что главной причиной несовершенства метода срединной линии является то, что в этом методе протяженность и конфигурация исходных (береговых) линий и их взаимное расположение учитываются не полностью.

Полагаем, что устранение этих недостатков, которое является целью настоящей работы, приведет к более точному разделению морских зон между соседними прибрежными странами и будет способствовать укреплению сотрудничества и дружественных отношений между ними.

* * *

В предлагаемом ниже методе, прежде всего, в процессе моделирования секторального разделения морской территории всем прибрежным государствам создаются одинаковые и равные условия, которые достигаются путем исключения таких факторов как численность населения, общая площадь, политический или военный потенциал и другие характеристики той или иной страны, ибо они неизбежно приведут к искаженным картинам. Также полностью исключаются наличие и отсутствие природных богатств того или иного региона разделяемых морских зон, так как последующие открытия новых месторождений могут провоцировать новое разделение. В результате единственным не исключенным фактором является наличие исходных (береговых) линий. Следовательно, процесс разделения морской территории должен моделироваться через исходные (береговые) линии с полным учетом протяженности, конфигурации и их взаимного расположения.

Естественно, чем больше граничит разделяемая морская зона с данной страной, тем больше должна быть территория, отводимой этой стране и наоборот. Т.е. исходная (береговая) линия является источником, создающий интересы той или иной страны в том или ином регионе морской зоны. Эти интересы являются более значительными, если рассматриваемый локальный морской регион находится в непосредственной близости к берегу данной страны, они также могут быть незначительными, если рассматриваемый морской регион находится вдали от нее. Следовательно, величина, направление и другие характеристики *поля заинтересованности* каждой прибрежной страны в том или ином регионе морского пространства должны определяться протяженностью и конфигурацией исходной (береговой) линии, а также их взаимным расположением.

Предположим, что береговая линия $A_1 - A_1$ (рис.3), принадлежащая стране A , создает поле заинтересованности на поверхности морской зоны. Обозначим через dE силовую характеристику этого поля в произвольной точке P , создаваемой бесконечно малым отрезком dL береговой линии $A_1 - A_1$. В данном случае вектор $d\vec{E}$ направлен к центру отрезка dL , его величина прямо пропорциональна dL , т.е.

$$dE \sim \lambda dL, \quad (1)$$

где λ - коэффициент, выражающий способность береговой линии $A_1 - A_1$ создавать поле заинтересованности.

Как было сказано выше, данная силовая характеристика также зависит от R - расстояния от элемента dL до рассматриваемой точки P . Их зависимость имеет следующий вид

$$dE \sim \frac{1}{R^\alpha}. \quad (2)$$

Определим показатель α . Так как поток поля, создаваемого элементом dL , через любую окружающую этот элемент сферу имеет одно и то же значение [2,3], то

$$dE \cdot 4\pi R^2 = const \quad \text{или} \quad dE = \frac{const}{R^2}$$

Откуда получаем, что $\alpha = 2$. Используя это значение, из (1) и (2) получим

$$dE = \frac{\mu\lambda dL}{R^2}, \quad (3-1)$$

или в векторном виде:

$$d\vec{E} = -\mu \frac{\lambda dL}{R^3} \vec{R}. \quad (3-2)$$

Здесь μ - коэффициент пропорциональности. Для определения силовой характеристики общего поля, создаваемого береговой линией $A_1 - A_1$ в целом, сложим все вклады от всех элементов dL

$$\vec{E} = \int_L d\vec{E} = -\mu\lambda \int_L \frac{dL}{R^3} \vec{R} \quad (4)$$

Для определения силовой характеристики поля заинтересованности, создаваемого другими прибрежными странами, перепишем выражение (4). Так как для них данная морская зона является общей, то коэффициенты пропорциональностей для любой исходной (береговой) линии, независимо от того какой стране она принадлежит, а также независимо от их протяженности и конфигурации, должны быть постоянными, т.е.

$$\lambda_1 = \lambda_2 = \dots = \lambda_i = \lambda = const, \quad \mu_1 = \mu_2 = \dots = \mu_i = \mu = const.$$

Следовательно, из (4) получим

$$\vec{E}_i = \int_{L_i} d\vec{E}_i = -\mu\lambda \int_{L_i} \frac{dL_i}{R_i^3} \vec{R}_i, \quad (5)$$

где i - индекс, показывающий порядковый номер прибрежной страны. Интеграл берется по всей длине L_i береговой линии $A_i - A_i$.

Используя эту систему уравнений (5), можно получить границу секторального разделение данного морского пространства между прибрежными государствами с прилежащими и противоположащими береговыми линиями. При этом особо отметим, что система уравнений (5) учитывает протяженность и конфигурацию исходной (береговой) линии каждой страны (а также их взаимное расположение) в полном объеме, без какого либо исключения.

Из-за сложности конфигурации исходных (береговых) линий $A_i - A_i$ систему уравнений (5) решить аналитически невозможно. Поэтому используем ЭВМ.

Алгоритм расчета на ЭВМ. Первоначально определим координаты всех точек исходной (береговой) линии $A_i - A_i$, отстоящих друг от друга на одинаково равном малом интервале ΔL . При этом $\Delta L = const$. Следует сказать, что чем меньше расстояние ΔL , тем больше точность расчета. Координаты этих точек фиксируются в

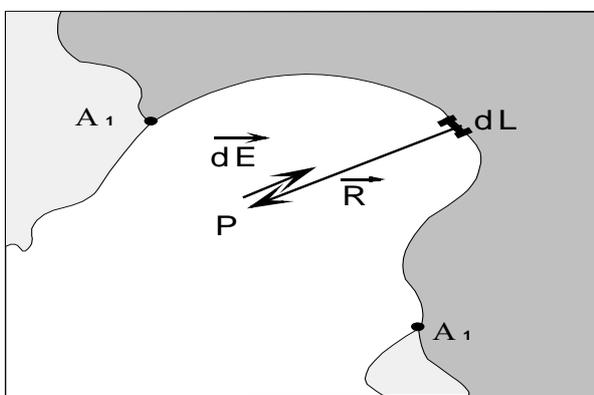


Рис 3

памяти ЭВМ. Затем берется произвольная точка P на поверхности водоема. Используя уравнение (5) найдем силовые характеристики \vec{E}_i полей, создаваемые исходными (береговыми) линиями $A_i - A_i$ этих стран в этой точке в отдельности.

После этого сравниваем их между собой. Если силовая характеристика поля заинтересованности одной из стран в этой точке, окажется больше, чем силовые характеристики полей других стран в отдельности, то эта точка считается принадлежащей этой стране.

После этого берем другую произвольную точку P и повторяем расчет. Если в этой точке силовые характеристики двух стран окажутся равными, то она считается одновременно принадлежащей этим двум странам, т.е. она находится на границе секторального разделения. А если в какой-то точке силовые характеристики трех стран окажутся равными, то в этой точке находится пересечение границ трех стран и т.д.

Продолжая расчет, получим множество всех точек, совокупность которых дает границу секторального разделения морской зоны между соседними прибрежными странами.

Одним словом, каждая точка границы разделения морского пространства между прибрежными государствами определяется по принципу равенства силовых характеристик полей заинтересованности двух или нескольких стран в данной точке.

*Это - то же самое, что математическое описание того, что было сказано многими известными политологами и геополитиками. Например, Н.Спикман писал, что «...границы являются теми **силовыми линиями**, где происходит нейтрализация взаимного напряжения между государствами», «...с точки зрения международной политики граница есть выражение соответствующих **силовых взаимоотношений** как линий, где уравнивается конфликтное напряжение», «... качество границы более выражается не в категориях стратегической ценности пограничной линии, а в категориях **силового потенциала** прилегающей к границе территории». Также по Р.Страус-Хюпе: «...границы определяют разделение политических образований и выражают **силовые отношения** между государствами...». Аналогичное определение было дано Ю.В.Тихонравовым: «... именно **баланс сил** и **динамическое равновесие интересов** определяют жизнеспособность тех или иных границ ...» [4, 5, 6].*

Из (5) следует, что преимуществом данного метода по сравнению с методом срединной линии является то, что в этом случае при определении границы секторального разделения морской территории между прибрежными государствами протяженность и конфигурация всех исходных (береговых) линий, а также их взаимное расположение учитываются целиком и полностью. Поэтому предлагаемый этот метод назовем **методом береговой линии**.

* * *

Теперь рассмотрим практическое применение предлагаемого нами метода береговой линии.

Для наглядности рассмотрим случай островов, приведенный на рис.2 (пример 2), так как при этом получим более убедительный результат. Для упрощения вычисления расстояние между островами берем равным 7,5 милям, ширину островов равной 1 мили, а длину первого и второго острова соответственно берем равной 3 и 10 милям.

Применяя предлагаемый нами метод береговой линии, найдем новую границу разделения морского пространства между островами A и B , принадлежащих разным государствам. Вычисленная новая граница имеет совершенно другой вид - вид замкнутой кривой, показанной на рис.4 в точных масштабах, где размеры каждой координатной клетки равны 4 милям. Отметим, что эта замкнутая линия представляет собой совокупность точек, в каждой из которых силовая характеристика поля заинтересованности двух островов имеет равную величину. За пределами заштрихованной области силовая характеристика поля заинтересованности, создаваемого островом B , больше силовой характеристики поля острова A . Внутри заштрихованной области, т.е. внутри замкнутой кривой – наоборот, т.е. силовая

характеристика острова *A* больше силовой характеристики острова *B*. Следовательно, заштрихованная область целиком принадлежит острову *A*. В зависимости удаленности от острова *A* часть этой области может быть территориальным морем или прилегающей зоной и т.д.

Область морского пространства за пределами заштрихованной области принадлежит острову *B*. Часть этой области является территориальным морем, если она находится на расстоянии 12 миль от острова *B*. Если часть этой области находится на расстоянии от 12 до 24 миль, то она является прилегающей зоной острова *B*. Аналогично можно найти исключительную экономическую зону и т.д.

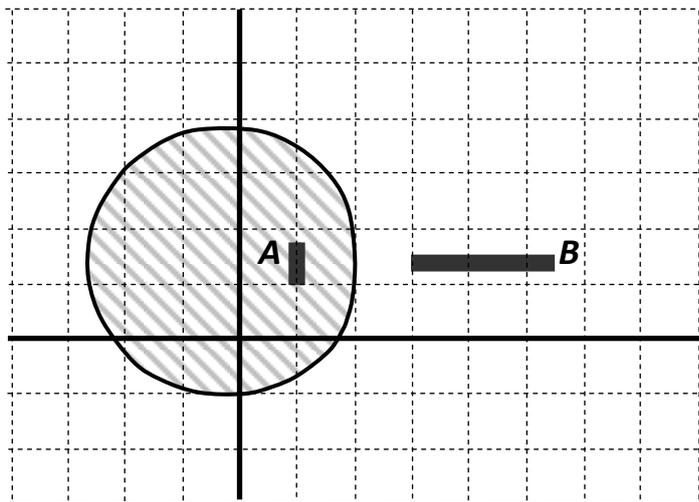


Рис.4.

Отметим, полученная картина разделения морского пространства между островами *A* и *B* является более объективной, чем картина, полученная в вышеприведенных примерах, где роль границы играла срединная линия.

На рис.5 в точных масштабах показаны вычисленные нами границы разделения территориального моря и прилегающей зоны между островами *A* и *B*. Здесь также, как в случае рис.4, размеры каждой координатной клетки равны 4 милям. Области морского пространства, обозначенные цифрами 1 и 2, являются территориальным морем соответственно островов *B* и *A*. Линия соприкосновения этих зон в точности совпадает с линией границ разделения морского пространства, показанной на рис.4. При этом заметим, что самые отдаленные точки указанных территориальных морей (т.е. точки их внешних границ) находятся на расстоянии 12 миль от соответствующих островов.

Морские пространства, обозначенные цифрами 3 и 4, которые соответственно находятся за областями территориальных морей, в свою очередь являются прилегающими зонами островов *B* и *A*.

Обратим внимание, что обозначенное цифрой 5 на рис.5 морское пространство, находящееся за пределами прилегающей зоны 3 острова *B*, является его исключительной экономической зоной. Как видно, остров *A* в данном случае не имеет своей исключительной экономической зоны, так как согласно ст. 55 Конвенции она должна прилегать к территориальному морю этого острова. В данном случае между территориальным морем острова *A* и его предполагаемой исключительной экономической зоной находится прилегающая зона острова *B*. Аналогично следует отметить, что в данном случае остров *A* также не имеет своего континентального шельфа.

Сравнивая результаты, показанные на рисунках 2, 4 и 5, можно сказать, что метод береговой линии, в отличие от метода срединной линии, дает более объективную и правдивую картину границ разделения морского пространства между прибрежными государствами.

Г.Т. БЕКМАНОВА, А.А. ШАРИПБАЙ, А.С. ОМАРБЕКОВА,
Г.К. ЕЛИБАЕВА, Б.Ж. ЕРГЕШ

АВТОМАТИЧЕСКАЯ ГЕНЕРАЦИЯ ОТЗЫВОВ ПОЛЬЗОВАТЕЛЕЙ НА ПРИМЕРЕ ОНТОЛОГИЧЕСКОЙ МОДЕЛИ ОТЕЛЯ

(НИИ «Искусственный интеллект» ЕНУ имени Л.Н.Гумилева, Астана, Казахстан)

1 Введение

Авторами ведется работа по сентимент анализу отзывов об отелях на казахском языке. Во время данной работы авторы столкнулись с проблемой фейковых отзывов об отелях на казахском языке, которые искусственно искажают результаты сентимент анализа. Поэтому возникла задача определения фейковых отзывов.

Есть несколько примитивных способов отличить правдивый отзыв от фейка: во-первых, настоящие отзывы, как правило, всегда «живые». В них присутствуют грамматические ошибки или сленг, они часто содержат больше эмоций, чем описаний технических достоинств товара. Если вы видите, что у товара 10 отзывов, и все они являются положительными, к этому нужно относиться с осторожностью[1].

Также стоит обращать внимание на дату публикации отзывов: если все они написаны практически в один и тот же день — это явный признак фейка. Конечно, ни один из вышеперечисленных способов на 100% не даст гарантии достоверности отзывов, но как минимум позволит взглянуть на читаемые мнения немного под другим углом.

Анализ отзывов показал, что как правило, фейковые отзывы являются однотипными, имеют правильную структуру и грамматику[2-4].

Для автоматической генерации отзывов об отелях нами была построена онтология для аспектов отелей и их категорий, по которым как правило, дается оценка в системе www.tripadvisor.com.

2 Онтологическая модель сервисов отеля

Для построения онтологической модели нами были выбраны аспекты и их категории для отелей. Для выявления аспектов и их категорий были исследованы различные системы для путешественников. В качестве основной системы выбрана система для путешественников, расположенная по адресу www.tripadvisor.com.

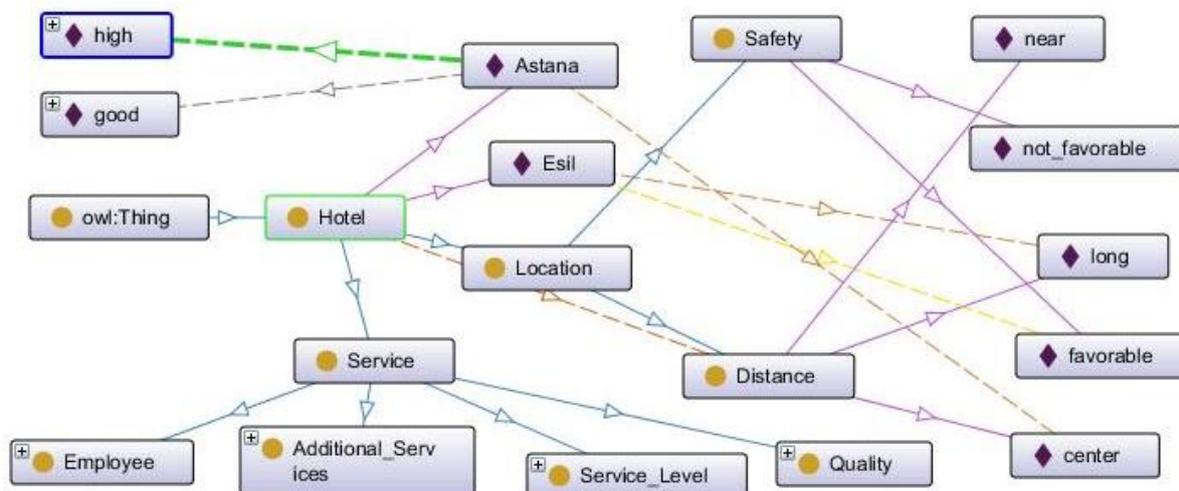


Рисунок 1. Онтологическая модель сервисов отеля

Онтологическая модель гостиничных услуг позволяет нам описывать аспекты и категории услуг и их взаимоотношения. Кроме того, представление онтологии позволяет трансформировать модель в схему RDF, которая будет использоваться для генерации эталонных фейковых отзывов.

Для генерации эталонных фейковых отзывов необходимо создать онтологии всех простых предложений казахского языка. В казахском языке существует 20 видов простых предложений, формальные модели этих предложений представлены в работе [5]. В данной работе для формализации предложений используется семантическая нейронная сеть. В казахском языке существуют следующие члены предложений: подлежащее, сказуемое, дополнение, обстоятельство, определение. В качестве частей речи казахского языка, которые могут быть тем или иным членом предложения выступают существительное, прилагательное, числительное, наречие, местоимение, глагол.

Связь между вышеперечисленными частями речи и членами предложения представлены в онтологиях. Онтологии доступны для скачивания (<http://e-zerde.kz/kazlang/>).

3 Генерация отзывов

Для генерации отзывов используются онтологическая модель. Нами используется три базы знаний. Первая база знаний содержит просто имена прилагательные с указанием их позитивности, негативности или нейтральности. Данная база знаний содержит слов.

Вторая и третьи базы знаний разработаны с использованием математической модели с нечеткой оценкой (простой и расширенной соответственно).

Вторая база знаний использует **Simple fuzzy model** оценки отзывов. В базе знаний хранятся значения каждого аспекта с уже заданными показателями «Позитивный», «Нейтральный», «Негативный». Тогда множества термов для этих лингвистических переменных определяются так:

$$T(x) = \{\text{«Позитивный»}, \text{«Нейтральный»}, \text{«Негативный»}\},$$

где

$$x ::= \langle \text{Service} \rangle \mid \langle \text{Location} \rangle \mid \langle \text{Room} \rangle \mid \langle \text{Sleep} \rangle \mid \langle \text{Cleanliness} \rangle \mid \langle \text{Food} \rangle \mid \langle \text{Price} \rangle$$

Для генерации отзывов с заданной эмоциональной оценкой используются термы из T , которые можно представить в виде функций принадлежности лингвистической переменной.

$$M(x) = \begin{cases} \text{positive}, & [0;0,33] \\ \text{neutral}, & [0,34;0,66] \\ \text{negative}, & [0,67;1] \end{cases}$$

Третья база знаний использует **Extended fuzzy model** оценки отзывов. Структура Extended Fuzzy Model представлена аспектами более чем одного уровня. Тогда как, Simple Fuzzy Model оценивает отзывы, содержащие аспект только первого уровня.

Чтобы расширить количества показателей оценок рассмотрим два уровня аспектов. Для этого аспекты первого и второго уровней обозначим через буквы A и B с нижним индексами, состоящими из цепочек букв и цифр.

Аспекты первого уровня:

$$A_S = \text{Service}; A_L = \text{Location}; A_R = \text{Room}; A_{Sq} = \text{Sleep_quality};$$

$$A_C = \text{Cleanliness}; A_F = \text{Food}; A_P = \text{Price}.$$

Аспекты второго уровня (или аспектные категории):

$$B_{S1} = \text{general}; B_{S2} = \text{staff}; B_{S3} = \text{quality}; B_{S4} = \text{service level};$$

$$B_{S5} = \text{additional_service}; B_{I1} = \text{safety and etc}.$$

Аспект сервис оценивается с учетом отношении работников к гостю, уровнем и качеством оказания услуг, дополнительных услуг оказываемой в гостинице. При оценивании расположения учитываются расстояние от интересующего пункта и безопасности района.

В качестве примера взяли аспект *Service* и его аспектные категории (аспект второго уровня).

Для порождения новых термов (имен новых значений) лингвистической переменной *Service* используем квантификаторов "очень".

Например, математическое описание аспекта первого уровня будет иметь вид:

$$A_S = \textit{Service};$$

$$T(\textit{service}) = \{\textit{very bad}, \textit{bad}, \textit{satisfactory}, \textit{good}, \textit{very good}\};$$

$$U = [0, 1];$$

$$G = \{\langle \text{новый терм} \rangle ::= \textit{very} \langle \text{базовый терм} \rangle, \\ \langle \text{базовый терм} \rangle ::= \textit{bad} \mid \textit{good}\};$$

$$M(\textit{service}) = \frac{\sum_1^n M_n}{n}$$

$$M(\textit{service}) = \begin{cases} \textit{very negative}, [0; 0,2], \\ \textit{negative}, [0,21; 0,45], \\ \textit{neutral}, [0,46; 0,75], \\ \textit{positive}, [0,76; 0,89], \\ \textit{very positive}, [0,90; 1] \end{cases}$$

$$A_S = \{B_{s1}, B_{s2}, B_{s3}, B_{s4}, B_{s5}\}$$

Подобные описания используются для каждого из 7 выделенных нами аспектов первого уровня и 17 аспектов второго уровня.

Таким образом, аспект онтологии включает в себя возможные сентимент значения того или иного аспекта отеля, а также включает числовую оценку негативности, позитивности или нейтральности. Это позволяет нам генерировать отзывы с заданным порогом позитивности, негативности или нейтральности, по всем правилам генерации простых предложений казахского языка, описанными их онтологическими моделями.

4 Выводы

В результате данной работы нами было сгенерировано в общей сложности 50242 фейковых отзывов. Так из первой базы знаний путем использования прилагательных нами было сгенерировано 49008 отзывов, из них положительных 42192, негативных 6680, нейтральных 136. Процент отзывов, явно определяющихся как фейковые, составил 6371. Данные отзывы семантически неточны. Из второй базы знаний, нами было сгенерировано 674 отзывов. Из третьей базы знаний 432. Поскольку во втором и третьем случае использовалась онтология, отличить на первый взгляд, что отзывы являются фейковыми невозможно. Однако нами сгенерирована база данных объемом 50242 записей, позволяющая в будущем использовать данную базу данных для анализа реальных отзывов пользователей и в случаях обнаружения большого количества отзывов, написанных в один день, неравномерного заполнения отзывами относительно времени, либо большого количества однотипных отзывов, имеющих правильную структуру и грамматику.

Список литературы

1 <http://mneniya.pro/>

2 <https://www.cs.uic.edu/~liub/FBS/fake-reviews.html>

3 Samsung probed in Taiwan over 'fake web reviews', April 16, 2013 [BBC News].

4 Woman Paid To Post Five-Star Google Feedback, September 15, 2012 [ABC7 News] (our algorithm in WWW-2012 paper can catch them).

5 Sharipbayev A.A., Razakhova B. Sh. Mathematical models of syntactical rules of Kazakh language subject to semantics of parts of the sentence./ The 4th Congress of the Turkic World Mathematical Society (TWMS). – Baku, 2011. –p. 463;

Р.С. БЕКТУРСЫНОВА, Б.Ж. ЕРГЕШ

УИКИПЕДИЯНЫ ПАЙДАЛАНЫП ҚАЗАҚША МӘТІНДЕГІ ЖАЛҚЫ ЕСІМДЕРДІ ТАНУ

(Л.Н. Гумилев атындағы Еуразия ұлттық университеті, Астана, Қазақстан)

Бұл мақалада Уикипедиядан алынған ақпараттар пайдаланылған мәтіннен жалқы есімдерді іздеп табу әдісі талқыланды. Ол үшін қажет:

- Жалқы есімдерді іздеп табуға арналған қолданыстағы әдістерді зерттеу;
- Уикипедиядан алынған ақпараттар пайдаланылған мәтіннен жалқы есімдерді іздеп табу әдісін дайындау;
- Дайындалған әдісті Texterra сөз өңдеу жүйесінде интеграцияланатын, бағдарлама түрінде жүзеге асыру;
- Қазақ тіліндегі жаңалық мақалалардан тұратын сынақ (тестілік) денесін дайындау;
- Дайындалған әдістің сапасына тестілеу жүргізу.

Мақала қазақ тіліне арналған жалқы есімдерді іздеп табу жүйесі ретінде әзірленуін қарастыру қажет. Ақпарат алу үшін қазақ тіліндегі Уикипедия [1] бөлімін пайдаланылды. Уикипедия — бұл көптілді қолжетімді еркін таратылатын Интернет-энциклопедия. Уикипедияда да 250 тілдегі 10 млн-нан астам мақала бар. Уикипедияда жиі қарастырылатын жалқы есімдердің негізгі түрлерінің ауқымды санына арналған мақалалар бар, бұл оның жалқы есімдерді іздеу мәселесі кезінде пайдалылығымен шартталады.

Уикипедиядан құрылымданған ақпараттарды іздеу кезінде пайдалануға болады:

- Уикипедияның санаттар беттері
- Лексикалық көпмағыналық шешу беттері
- Қайта бағыттау беттері
- Уикипедияның басқа да мақалаларына сілтеме
- Уикипедияның басқа тілдердегі мақалаларына сілтеме
- Уикипедия шаблондары (соның ішінде, Infobox-шаблоны).
- Қойылған міндет

Әдетте жалқы есімдерді іздеу жүйесін бағалау белгіленген корпуста қолмен жүзеге асырылады. Жалқы есімдер өзінің шекарасы (оған кіретін сөздермен) мен түрлері бойынша анықталады. Өлшеу мен бағалаудың түрлі тәсілдері қолданылады. CoNLL конференциясында [2] келесідей бағалау тәсілі ұсынылған болатын. Егер есімнің шекарасы мен түрі корпустағы сарапшы белгісімен сәйкес келетін болса, онда жалқы есім дұрыс анықталған болып есептеледі, әйтпесе жүйе қате бар. Мұндай бағалау тәсілі дәлме-дәл сәйкестік әдісімен бағалау деп аталады. Бұл тәсіл кеңінен қолданылады, алайда оның өзіндік кемшіліктері де бар: ол белгілеу кезінде сарапшы тарапынан кетуі мүмкін қателіктерді есепке алмайды. Мысалы, қосымша қателіктер түрін көрсетуге болады: типі дұрыс анықталған, ал шекарасы қате, шекарасы дұрыс белгіленген, бірақ типі қате, және есім типі де және шекарасы да қате анықталған. MUC конференцияларында [3] жүйелер екі бағыт бойынша: жалқы есімдер шекарасын дұрыс белгілеу және типтерді дұрыс анықтау бойынша параллельді түрде бағаланды. Жүйе жұмысының сапа көрсеткіштеріне толықтық (Recall, R), дәлдік (Precision, P) және F-өлшем (F) жатады, олар келесідей анықталады:

$$R = \frac{\text{дұрыс анықталған есімдер саны}}{\text{барлық анықталған есімдер саны}}$$

$$P = \frac{\text{дұрыс анықталған есімдер саны}}{\text{корпустың есімдер саны}},$$

$$F = \frac{2+P+R}{P+R}.$$

Жалқы есімдерді іздеу көптеген конференцияларда, атап айтқанда CoNLL және MUC конференцияларда шешімі табылмаған мәселе еді. Осыған байланысты жалқы есімдерді іздеуге арналған кейбір дайын корпустар бар. Яғни, мысалы, CoNLL-2003 конференциясында мынадай қолжетімді дайын корпустар [5] пайдаланылды:

- - оқыту үшін 946 мақаладан, тестілеу үшін 231 мақаладан тұратын, деректері Reuters Corpus-тан алынған ағылшын тіліне арналған корпус.
- - оқыту үшін 553 мақаладан, тестілеу үшін 155 мақаладан тұратын, деректері ECI Multilingual Text Corpus -тан алынған неміс тіліне арналған корпус.

Осы жұмыс барысында қолданылатын Уикипедияның жіктелуін жүзеге асыратын жүйені қарастырайық. Жүйе логистикалық регрессияны пайдаланатын машиналық оқыту әдісіне негізделген. Жіктелу объектісі ретінде Уикипедия мақалалары алынады. Классификатор оларды жалқы есім түрлері мен арнайы типтер деп бөлінетін кластарға (сыныптарға) жіктейді. Арнайы типтер бұл мақаланың қарастырылып отырған жалқы есім түрлерінің ешқайсысына жатпайтындығын білдіреді. Жіктеу кезінде объекті үшін келесідей белгілер айшықталады: Мақала тақырыбы; Категория атауы; Бөлімдер тақырыбы; Мақаланың бірінші сөйлемі; Мақаланың бірінші абзацы; Шаблондар тақырыбы және олардың ішіндегісі («кілт-мән» жұбын көрсететін параметрлер).

Жалқы есімдердің, BBN класына[4-7] жататын тапсырмаларды шешуде ұсынылатын, жалқы есімдер түрлерінің ішкі жиынтығы болып табылатын келесідей түрлерін қарастыратын боламыз.

Тұлғалар; Ұйымдар; Саяси/мемлекеттік; Білім беру; Коммерциялық; Коммерциялық емес; Орналасқан жері; Мемлекеттер; Қалалар; Субъектілер; Басқалары; Географиялық объектілер; Өзендер; Көлдер, теңіздер, мұхиттар; Континенттер; Аймақтар; Басқалары; Әртүрлі; Ақша; Күн; Уақыт; Азық-түлік; Оқиғалар; Өнер туындылары; Атаулы ғимараттар, инфрақұрылымдық объектілер.

Мұндай жиынтықтардың таңдалуы қазақ тілінің ерекшеліктеріне (жоғарыда аталғандардың көпшілігі зат есімнің өзіндік атауы болып табылады, бұл ең бірінші кезекте жалқы есім мағынасы келеді) және қолданушылар қызығушылығына негізделеді: жүйені жаңалық мақалаларда сынақтан өткіземіз, ал жаңалықтарда осындай жалқы есімдер түрлерін іздеу әлдеқайда тиімді.

Жалқы есімдерді іздеп табу базалық жүйесі ретінде жіктеудің әртүрлі келесідей мүмкіндіктерін пайдаланатын жүйені қарастыратын боламыз:

1. Локалды:

- Сөздің префиксі
- Сөздің суффиксі
- Сөз тұлғасы
- Сөйлемдегі сөздің орны

Сөздің жазылу ерекшеліктері (барлық символдар – бас әріппен; барлық символдар – цифрлар немесе әріптер; барлық символдар – цифрлар немесе тыныс белгілері; барлық символдар – әріптер емес; символдар арасында цифрлар бар; бас әріппен басталады) Аталған сөздің жазылу нұсқасы емлеге сай цифрлардың кіру

2. Локалды емес: Контекстік агрегация белгілері

- Болжамдардың локалды тарихының белгілері (сөздің шағын сол жағында орналасқан жалқы есім түрлерін есепке алатын, мысалы, берілген сөздің алдындағы 2 сөз)
- Болжамдардың кеңейтілген тарихының белгілері (сөздің үлкен сол жағында орналасқан шағын жалқы есім түрлерін есепке алатын, мысалы, берілген сөздің алдындағы 1000 сөз).

Төмендегі 1,2-суреттерде жоғарыда сипатталған әдістің программалық жүзеге асуының мысалы келтірілген.

Мәзір Анықтама

Мәтінді еңгізіңіз:

Аслан мен Иван Көкшетау табиғатында 8-нші наурыздан бері бірге жазғы демалысты өткізді.



Сурет 1 Талдауға арналған қазақша мәтінді

Анализ нәтижесі: 

Сөз	Анықтамасы	Сұрағы
[Date: 8-нші наурыздан]		Қашан?
[Name: Аслан]	(Википедиядан алынған)	Кім?
[Name: Иван]	(Википедиядан алынған)	Кім?
[Location: Көкшетау]	- 140 685 адам (2014)	Қайда?

Сурет 2 Талдау нәтижесі

Пайдаланылған әдебиеттер

- 1 <https://kk.wikipedia.org>
- 2 <http://www.conll.org/>
- 3 https://en.wikipedia.org/wiki/Message_Understanding_Conference
- 4 Ratinov L., Roth D. Design challenges and misconceptions in named entity recognition. // Proceedings of the Thirteenth Conference on Computational Natural Language Learning / Association for Computational Linguistics. 2009. P. 147–155.
- 5 Tjong Kim Sang E.F., De Meulder F. Introduction to the CoNLL-2003 Shared Task: Language-independent Named Entity Recognition // Proceedings of the seventh conference on Natural language learning at HLT-NAACL 2003-Volume 4 / Association for Computational Linguistics. 2003. P. 142-147.
- 6 Kim J.-D., Ohta T., Tsuruoka Y., Tateisi Y. Introduction to the Bio-Entity Recognition Task at JNLPBA // Proceedings of the international joint workshop on natural language processing in biomedicine and its applications / Association for Computational Linguistics. 2004. P. 70-75.
- 7 Nadeau D., Sekine S. A survey of named entity recognition and classification // Lingvisticae Investigationes. 2007. Vol. 30, no. 1. P. 3–26.

Б.А.ДАУРЕНБЕКОВ

МОБИЛЬНОЕ ПРИЛОЖЕНИЕ «КОНВЕРТЕР»

(Магистрант, ЕНУ им. Л.Н. Гумилева, Астана, Казахстан)

Введение

Актуальность. Все мобильные приложения условно можно поделить на программы для рабочих целей и на развлекательные программы. Первые позволяют бизнесменом и офисным работникам контролировать бизнес-процессы, составлять аналитическую отчетность, выполнять такие задачи, разработка дизайна фирменного стиля. Вторые включают в себя разнообразные игры, софт для просмотра фильмов и прослушивания музыки, средства для общения и т.д. каждое из мобильных приложений находит своего потребителя, однако, как отмечают специалисты из этой области, наибольшей популярностью пользуется специализированный софт, каким и является разработанное мной приложение «Конвертер».

Новизна данной разработки в том, что Казахстан с 2025 года полностью перейдет на латинскую графику. Об этом сообщил населению президент Нурсултан Назарбаев. По мнению экспертов, это будет способствовать более плотному вхождению страны в тюркский мир. Приложение позволит переводить казахский текст с кириллицы в латиницу. Приложение будет полезно как для студентов, так и для обычных пользователей.

Описание приложения «Конвертер»

Мобильное приложение «Конвертер» предназначено для перевода слов и фраз с кириллицы на латиницу. Отличительной особенностью программы является возможность перевода с кириллической раскладки на латинскую. Т.е. пользователь выбирает направления перевода, вводит нужную фразу и программа отображает перевод на экране гаджета.

Важной особенностью программы «Конвертер» является возможность перевода без использования подключения к сети Интернет. Редактируемые словари находятся в памяти устройства, или на флэш-накопителе.

Таблица соответствия кириллического казахского алфавита латинскому представлена в табл. 1.

Таблица 1. Таблица соответствия кириллического казахского алфавита латинскому

а	ә	б	в	г	ғ	д	е	ж	з	й	к	қ	л	м	н	ң	о	ө	п	р	с	т	у	ұ	ү	ф	х	ш	ы	і
a	ae	b	v	g	gh	d	e	zh	z	j	k	q	l	m	n	ng	o	oe	p	r	s	t	w	u	ue	f	kh	sh	y	i

После получения данных в поле ввода, программа в цикле перебирает строку посимвольно. Каждый символ проверяется на наличие в таблице соответствия. Если соответствие установлено – символ заменяется. Если нет, пропускается и осуществляется переход в начало цикла. При достижении окончания строки – алгоритм прекращает действие.

Интерфейс и основные функции приложения показаны на Рисунках 2-3.

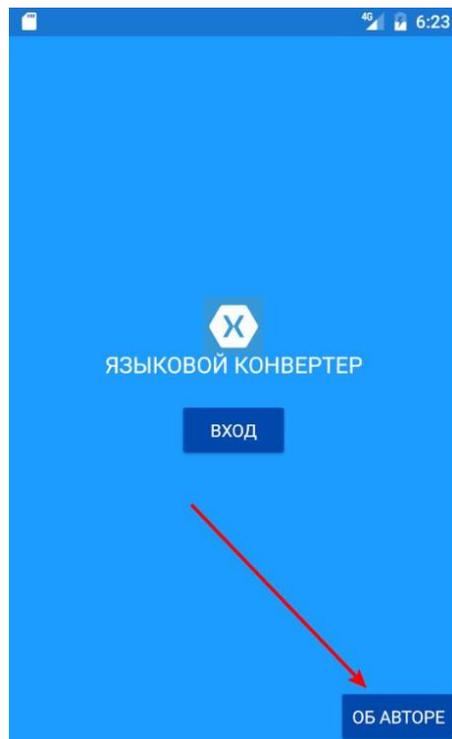


Рисунок 2. Главная страница приложения «Конвертер»

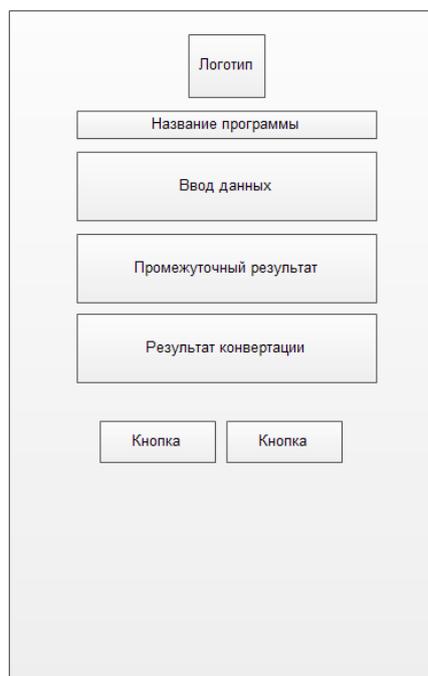


Рисунок 3. Макет страницы ввода данных и отображения результата конвертации

На странице расположены следующие блоки:

- ввод исходных данных;
- отображение промежуточного текста;
- отображение результата конвертации.

Созданная страница представлена на рис. 4.

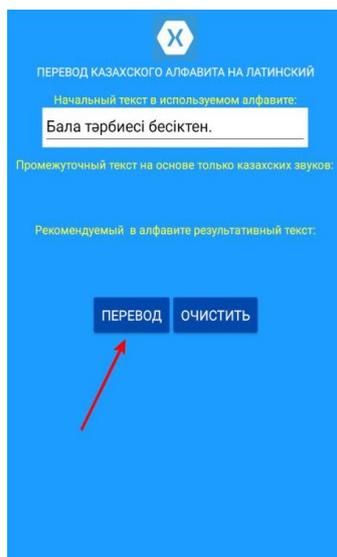


Рисунок 4 – Страница ввода данных и отображения результата конвертации
 Для конвертации текста необходимо кликнуть на кнопку «Перевод». Данные результата выполненной конвертации отобразятся в соответствующих полях (рис. 5).

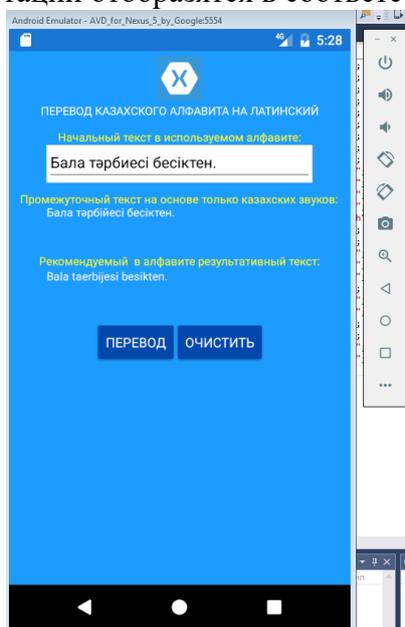


Рисунок 5 – Отображение результата конвертации

Для очистки результата выполнения необходимо кликнуть на кнопку «Очистить». Для выхода из приложения необходимо кликнуть на кнопку внизу экрана устройства.

Заключение

В результате проделанной работы была выполнена поставленная цель, а именно, разработан эволюционный прототип приложения «Конвертер» для осуществления конвертации казахского текста с кириллицы на латиницу.

Список использованной литературы и источников

- 1 Хашими С., Коматинени С., Маклин Д. Разработка приложений для Android. СПб. Питер, 2011. — 736 с.
- 2 Рето Майер. Professional Android 2: Application Developmentecond Edition, Эксмо, 2011 – 672 с.
- 3 www.alphabet.kz

М.С. ЕДИГЕЕВА

ИСПОЛЬЗОВАНИЕ ТЕХНОЛОГИИ ВЕБ-ДИЗАЙНА В ВЕБ-ПРИЛОЖЕНИЯХ

(ЕНУ имени Л.Н. Гумилева, г. Астана, Казахстан)

Создание веб-приложений сегодня считается одним из перспективных направлений не только в Казахстане, но и во всем мире. Потребителями информационных продуктов и услуг могут быть юридические и физические лица, решающие различные задачи.

Веб-приложения широко используются в глобальной сети Интернет для различных целей, таких как реклама товаров и услуг, виртуальные представительства фирм, общение в тематических сообществах и др. Для любого веб-приложения, необходимо использовать дизайн, который отражает его назначение и цели.

Целью данной работы является проектирование и реализация веб-приложения – конструктора сайтов с использованием технологий веб-дизайна и дизайн стилей.

Веб-дизайн представляет собой процесс, который включает в себя целый ряд различных дисциплин, таких как графический дизайн, информационный дизайн, дизайн интерфейса, изготовление документов, создание сценариев, программирование и мультимедиа[1].

Данное приложение – конструктор построен как распределенная система с применением клиент-серверной архитектуры. База данных (БД) системы является централизованной. В основе разработанной системы лежит хранилище разработанных пользователями системы сайтов и механизмов их разработки и редактирования. На рисунке 1 представлена инфологическая модель системы.

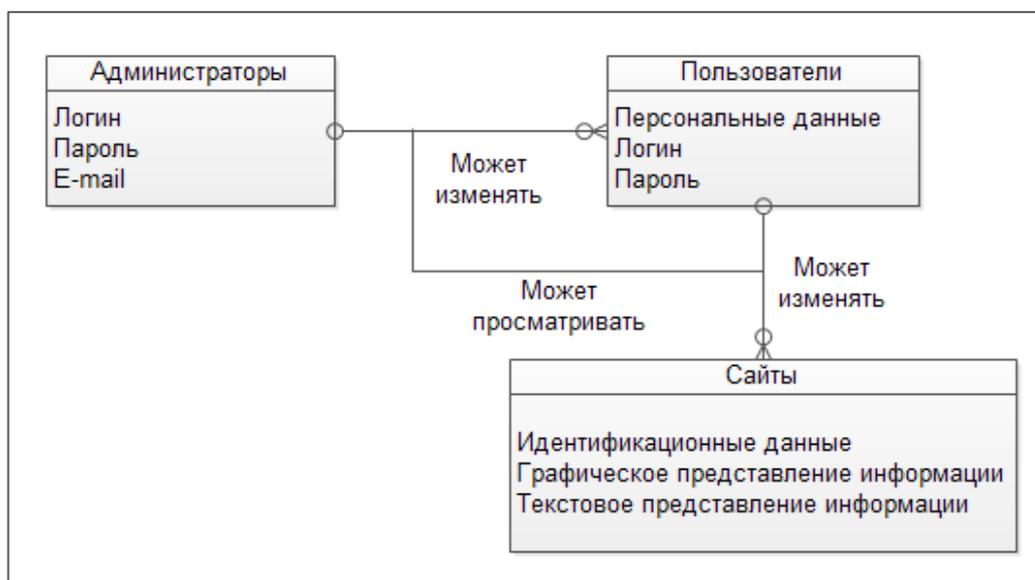


Рис. 1 – Инфологическая модель

Каждый сайт характеризуется следующей информацией:

- Идентификационные данные (название, логотип);
- Графическое представление информации (шаблон сайта, фотогалерея, видеогалерея и т.д.);
- Текстовое представление информации (общая информация, справка, контактные данные, отзывы и т.д.).

С системой будут взаимодействовать следующие группы лиц:

- Администраторы системы;
- Пользователи системы.

Администраторы системы имеют права на выполнение следующих действий:

- Добавление, удаление и редактирование пользователей;
- Добавление, удаление и редактирование контентных страниц;
- Просмотр отзывов, оставленных на сайте с помощью формы обратной связи.

Пользователи системы могут выполнять следующие действия:

- Создавать персональные сайты;
- Просматривать контентные страницы сайта;
- Редактировать персональные сайты;
- Удалять персональные сайты;
- Оставлять отзывы о сайте в форме обратной связи.

Для аутентификации пользователей в системе используется имя пользователя и пароль при входе в систему, доступ к функциям системы определяется процедурами авторизации пользователей.

Готовый графический интерфейс изображен на рисунке 2.

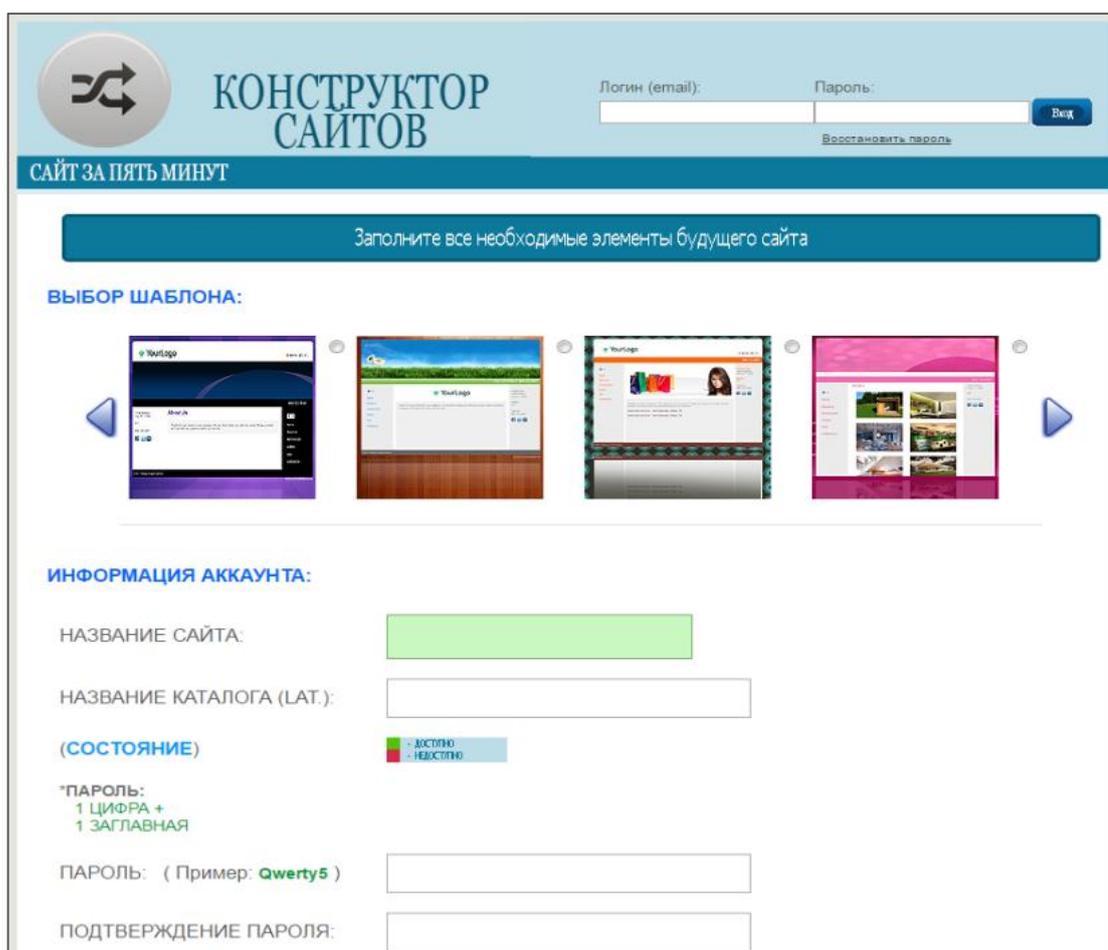


Рис. 2 – Фрагмент графического интерфейса системы

Система была успешно протестирована, ошибок в разработке и коде выявлено не было.

Список использованной литературы:

1 Niederst Robbins, J. (2007). Learning Web Design. Sebastopol, California, USA: O'Reilly Media.

Б.Ж. ЕРГЕШ, Р.С. БЕКТУРСЫНОВА

БІЛІМДЕР БАЗАСЫНА НЕГІЗДЕЛГЕН ҚАЗАҚ ТІЛІНДЕГІ ЖАЛҚЫ ЕСІМДЕРДІ ТАНУ

(Л.Н. Гумилев атындағы Еуразия ұлттық университеті, Астана, Қазақстан)

Жалқы есімдерді анықтау (Named Entity Recognition, NER) тапсырмасы мәтіннен мән атауларын (сөздер мен сөз тіркестерін) анықтаудан және алдын ала белгіленген категориялар, мысалға, тұлғалар (PERSON), ұйымдар (ORGANIZATION), географиялық объектілер (LOCATION) және басқалар бойынша олардың классификациясын танып білуден тұрады. Бұл тапсырма құрылымданбаған немесе нашар құрылымданған ақпарат көздерінен, соның ішінде Интернеттен құрылымданған ақпаратты автоматты түрде іздейтін және, өз кезегінде, ақпаратты іздеу және оны табиғи тілде өңдеумен байланысты ақпаратты іздеудің (Information extraction, IE) ішкі тапсырмасы болып табылады [1].

Жалқы есімдерін іздеуші компонент көптеген қосымшаларда қолданылады, әсіресе болашағы өте зор бағыттарының бірі құрылымданбаған ақпарат көздерінен түскен мәліметтермен семантикалық веб (мысалы, LOD- онтология) деректерін толтыру болып табылады.

Жалқы есімдерге адам есімдерін, компаниялардың атауларын, географиялық атаулар, мекен жайлар, күндер, сонымен қатар электронды адресстер, телефон нөмірлері және олардың өзара байланысын жатқызуға болады. Қазіргі кезде жалқы есімдерді анықтау бизнестің әр түрлі салаларында маркетингтік зерттеулерден бастап, интеллектуалды жартылай және автоматты шешім қабылдау жүйелеріне дейін пайдаланылады. Жалқы есімдерді анықтау қайда және қалай қолдануға болады? Бірінші айқын пайдалану бұл - кейбір нарықтық зерттеулер, ұйымдар немесе олардың брендтері болып табылады. Осы зерттеулердің мәні кейбір брендтер немесе ұйымдар жайлы әртүрлі көздерден (әлеуметтік желілер, блогтар, сайтта түсіндірмелер, әріптес сайттар және т.б.) автоматты түрде пайдаланушыларының пікірлері жайлы ақпарат алу болып табылады. Жалқы есімдерді анықтау қалай жұмыс жасайды? Тұлғалар мен мәтіндерден бастап олардың арасындағы қарым-қатынастардың өндіру мәселесі оңай жұмыс емес. Сонымен қатар, ол белгілі бір тілге байланысты болады [2]. Мысалға, қазақ тілі үшін жазылған программа украин немесе орыс тілі үшін жұмыс істей бермейді.

Білім базасы болса белсенді түрде жетпеген ақпаратты толықтыруға тырысады. Егер ... , онда ... форматындағы ережелер кең таралған, бірақ білімді көрсетудің жалғыз амалы емес. Осы мақсатта сараптамалық жүйелерде семантикалық желілер, фреймдар, нейрондық желілер және басқа да амалдар қолданылады. Кейбір терең деңгейлерде эквиваленттер анық көрінеді. Таңдалған әдіс мәселе түрімен және пәннің спецификасымен анықталады. Сараптамалық жүйедегі көптеген ережелер эвристикалық, берілген пәндегі сарапшылардың эксперименттеріне негізделеді. Егер алгоритмдік әдіс нақты әрі оптималды шешімге кепілдік берсе, эвристикалық әдіс тек мүмкін шешімді қарастырады. Білім базасы логикалық шешім механизмі үшін кіру ағыны болып табылады [3, 4]. Қазақша мәтіндерден жалқы есімдерді тану үшін ережеге және білімдер базасына негізделген әдіс таңдалды.

Қазақ тіліндегі мәтіннен жалқы есімдерді анықтау есебін шешу үшін келесі категорияларды мәтін ішінде анықтауды жөн көрдік: name(есім), geo(жер-су), org (өндіріс, мекеме, ұйым), animals (үй жануарлары).

Осы категориялар бойынша білімдер базасы құрастырылды[5]. Адам есімдері білімдер базасында 2230 сөз. Жер-су атаулары білімдер базасында 3099 сөз. Өндіріс, мекеме, ұйым атаулары базасында 900 сөз. Үй жануарлары білімдер базасында 132 сөз.

Келесі суретте (сурет 1) сөздердің білімдер базасында көрсетілуі келтірілді.

	A	B	C	D	E	F	G
1	Қазақ қыздары қазақша есімдері мен есімдері мағынасы						
2	Сөз	Анықтамасы	Сұрағы				
3	АБИРА	- тәжік тілінен алынған абера — шөбере, шөбере қыз.	Кім?				
4	АГИЛА	- араб және иран тілдерінен алынған агиля — ақылды, парасатты сөзінен жасалған.	Кім?				
5	АҒЛИЯ	- грек тілінен — сұлулық, көркемдік, салтанат.	Кім?				
6	АДЛИЯ	- тәжік тілінен алынған адл — әділдік, ақиқаттық + ия қосымша.	Кім?				
7	АЖАР	- көрікті, келбет, нұрлы; түр, өң.	Кім?				
8	АЖАРГҮЛ	- гүл сияқты әдемі, сұлу, көрікті.	Кім?				
9	АЖАРХАН	- көрікті, ажарлы, әрлі, нұрлы.	Кім?				

Сурет 1. Адам есімдері білімдер базасы

Қазақ тіліндегі атау сөздерге байланысты ережелерді толықтай қарастыру нәтижесінде келесі жалқы есімдерді анықтау ережелері туындады:

– Егер сөз бас әріптен басталса және кім? сұрағына жауап беретін сөзді деректер қорынан қарастырамыз, ол деректер қорында бар болса ол жалқы, жанды зат есім болса онда ол адамның есімі(name) болып табылады.

Мысалы төменгі суретте көрсетілген деректер қорында кім? сұрағына жауап беретін жалқы, жанды зат есімдер.

– Егер сөз бас әріптен басталса, сөзді деректер қорынан қарастырамыз, ол деректер қорында бар болса, ол жалқы, жансыз зат есім болса, онда ол жер-су, мекеме, қала, көше, аң-құс, газет, журнал атаулары болып табылады.

Жалқы есімдерді анықтау алгоритмі:

- Берілген сөйлемдегі сөздерді қарастырады;
- Бас әріптен басталған сөздерді табады;
- Табылған сөздерді деректер базасынан іздейді;
- Деректер базасында табылған сөздерді мәліметтер шығару терезесіне шығарады.

Жоғарыдағы алгоритм бойынша жалқы есімдерді анықтаудың автоматтандырылған жүйесін жүзеге асырылады.

Қазақ тіліндегі мәтіннен жалқы есімдерді анықтау есебін шешу үшін мысал ретінде келесі категорияны қарастырайық: егер сөз бас әріптен басталса, кім? сұрағына жауап беретін сөзді деректер қорынан қарастырамыз, ол деректер қорында бар болса ол жалқы, жанды зат есім ол адамның есімі(name) болып табылады.

Программа адам есімін интерфейстегі кестеге шығару үшін бірінші программаға адам есімін енгіземіз, егер де ол есім базада бар болса, программа төмендегідей интерфейсте адам есімін және оның мағынасын бізге көрсетеді (сурет 2).

Мәзір Анықтама

Мәтінді еңгізіңіз:

Айбала үйге келді

Анализ нәтижесі:

Сөз	Анықтамасы
[Name: Айбала]	- ай сияқты сұлу, көркем болсын деген тілекпей қойылған есім.

Сурет 2. Адам есімін енгізген кездегі интерфейстегі көрінісі

Жұмыс барысында қазақ тіліндегі адам есімдері, жер-су аттары, мекеме атаулары, жан-жануарлар атаулары, ақша атауларының білімдер базасы құрылды. Жалқы есімдерді анықтау ережелері сипатталып, жалқы есімдерді тану алгоритмі құрылды. Әдістерді зерттей келе ережеге негізделген жүйе жүзеге асырылды.

Болашақта басқа да осы жұмысты әрі қарай зерттеп, омонимия мәселесін шешу мақсаты қойылды. Бұл істелген қазақша мәтіндердегі жалқы есімдерді тану жүйесі қазақ тілін өңдеу жүйелерінің бір бөлігі болады.

Пайдаланылған әдебиеттер

- 1 Kim J.-D., Ohta T., Tsuruoka Y., Tateisi Y. Introduction to the Bio-Entity Recognition Task at JNLPBA // Proceedings of the international joint workshop on natural language processing in biomedicine and its applications / Association for Computational Linguistics. 2004. P. 70-75.
- 2 http://www.kazntu.kz/sites/default/files/kazntu.kz.SUMMARY_Nugumanoa
- 3 Лорьер Ж. Системы искусственного интеллекта. – М.: Мир, 1991. Нильсон Н. Принципы искусственного интеллекта. - М.: Радио и связь, 1995.
- 4 Krishnan V., Manning C. D. An effective two-stage model for exploiting non-local dependencies in named entity recognition // Proceedings of the 21st International Conference on Computational Linguistics and the 44th annual meeting of the Association for Computational Linguistics / Association for Computational Linguistics. 2006. P. 1121–1128.
- 5 Бектұрсынова Р.С. Қазақша мәтіндегі жалқы есімдерді анықтау ережелері. «Ғылым және білім - 2017» студенттер мен жас ғалымдардың XII Халықаралық ғылыми конференциясының баяндамалар жинағы, Астана, 2017. 675-678 б., ISBN 978-9965-31-827-6.

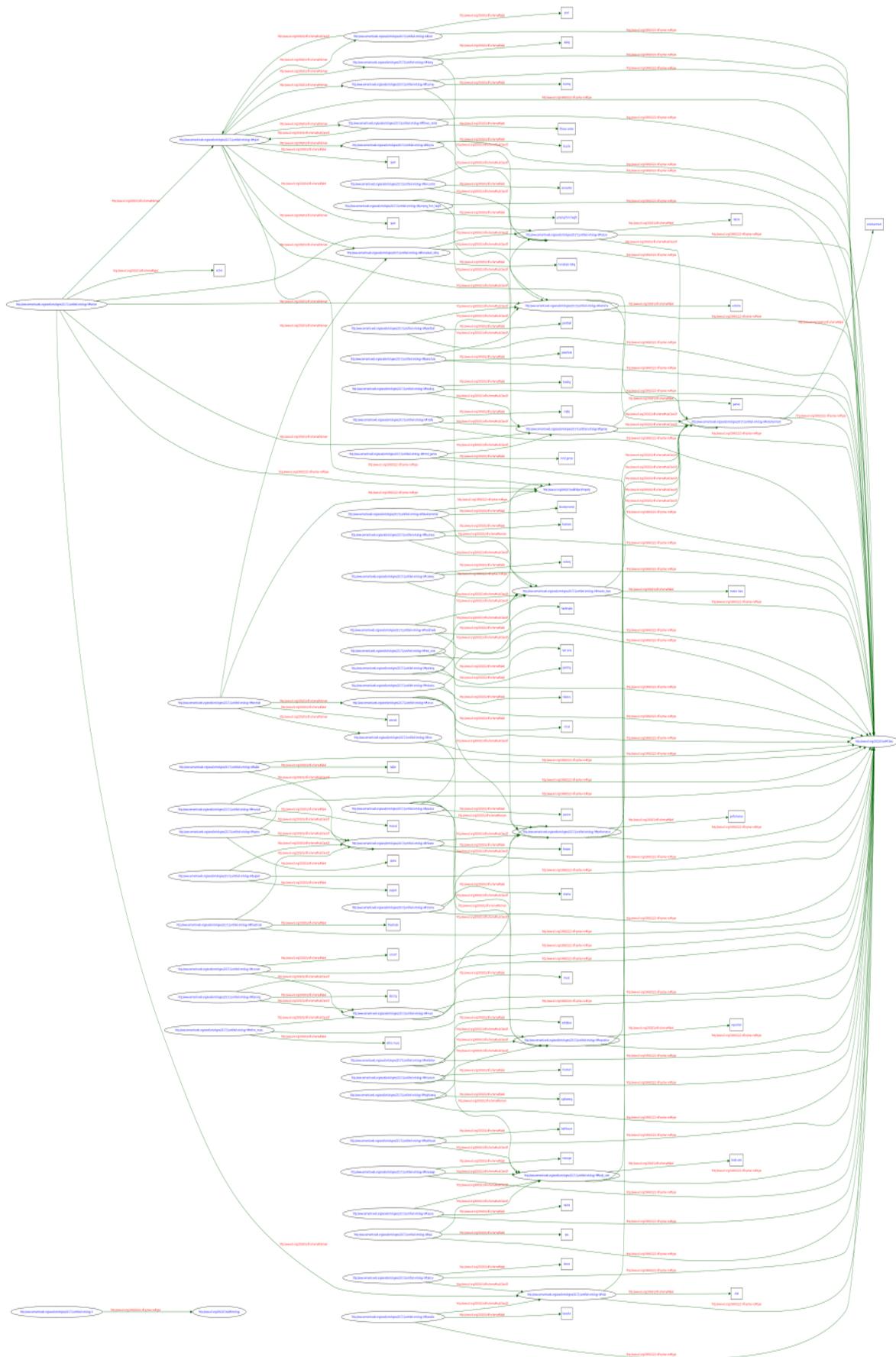


Рисунок 2. Граф rdf-документа

На основе данной базы знаний будет реализован портал рекомендательной системы по планированию досуга. На рисунке 3 отражена карта сайта.

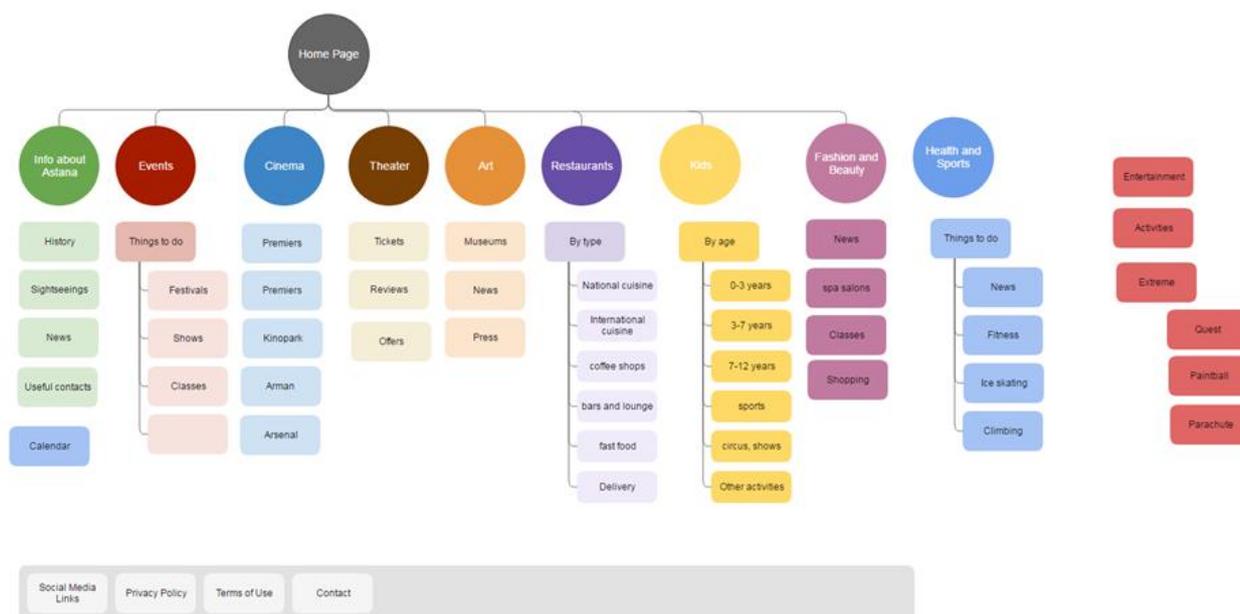


Рисунок 3. Карта сайта

Таким образом, в настоящее время изучена предметная область, проведен анализ существующих аналогичных систем, разработана онтологическая модель, онтология проверена на валидность, построена карта сайта. Новизна разрабатываемой системы заключается в проведении анализа данных предпочтений клиентов, использовании методов коллаборативной фильтрации, используемых в образовательных процессах (обучении, контроле и оценке знаний); представлении материала по объектам досуга на языке представления знаний, основанного на семантических гиперграфах.

Использованные источники:

1. Гаврилова Т.А., Хорошевский В.Ф. «Базы знаний интеллектуальных систем» - С.-Петербург.:ПИТЕР, 2000 с.81–112
2. OWL Web Ontology Language. W3C Recommendation [Electronic Resource] / 2004 – Mode of access: <http://www.w3.org/TR/owl-features/>
3. Auer S., Bizer C., Kobilarov G., Lehmann J., Cyganiak R., Ives Z. DBpedia: A nucleus for a web of open data. ISWC, 2007.
4. Cambridge Mathematical Thesaurus. URL: <http://thesaurus.maths.org>
5. Aberer K., Boyarsky A., Cudré-Mauroux P., Demartini G., Ruchayskiy O. ScienceWISE: A Web-based Interactive Semantic Platform for Scientific Collaboration // 10th International Semantic Web Conference (ISWC 2011 - Demo), Bonn, Germany, 2011.
6. Добров Б.В., Лукашевич Н.В. Онтология по естественным наукам и технологиям ОЕНТ: структура, состав и современное состояние, Электронные библиотеки. – 2008. – Т.11. – В. 1.
7. Аникин А. В. Формальная модель описания учебных объектов / А. В. Аникин, И. Г. Жукова, М. Б. Сипливая // Сборник научных трудов международной конференции "Информационные технологии в образовании, технике и медицине", Волгоград, 2004.
8. Е.А. Жыжырий, С.С. Щербак Применение web-онтологий в задачах дистанционного обучения [Электронный ресурс]. – Режим доступа: <http://shcherbak.net/dist/>
9. Брейкин Е. А. Рекомендательная система на основе коллаборативной фильтрации // Молодой ученый. — 2015. — №13. — С. 31-33.

Д.Б. ИСКАКОВА, Р. НИЯЗОВА

**МОДЕЛИ И МЕТОДЫ ОЦЕНКИ КАЧЕСТВА ПОДСИСТЕМЫ МОНИТОРИНГА
ВЫПУСКНИКОВ ПРОГРАММЫ БОЛАШАК**

(Магистрант, Евразийский национальный университет им. Л.Н.Гумилева)

В соответствии с идеологией стандартов ИСО серии 9000 можно выделить четыре направления деятельности в области качества, посредством которых система качества воздействует на процесс формирования качества продукции и услуг на различных этапах их жизненного цикла:

1. планирование;
2. управление;
3. обеспечение;
4. улучшение.

Каждое направление деятельности имеет особенности, и представляют собой четыре функциональных подсистемы системы качества.

Планирование качества включает в себя деятельность по установлению целей и требований к качеству и применению элементов системы качества. Планирование качества продукции и услуг охватывает идентификацию, классификацию и оценку качества, установление целей и нормирование требований к качеству продукции или процессов. Планирование качества процессов охватывает подготовку программы качества, выработку предложений по улучшению качества, подготовку применения системы качества, включая составление стандартных графиков ее ввода в действие и применения.

Управление качеством включает в себя методы и виды деятельности оперативного характера, используемые для выполнения требований к качеству. В состав процедур управления качеством входят контроль качества, разработка и реализация мер по корректировке процессов. Основное назначение управления качеством в организации - выявлять каждое отклонение от установленных требований к качеству продукции и услуг, применять решения по дальнейшему использованию продукции, имеющей отклонения или дефекты, не допускать появления повторных отклонений или дефектов за счет своевременной разработки и реализации корректирующих мер.

Обеспечение качества включает в себя все планируемые и систематически осуществляемые виды деятельности в рамках системы качества, необходимые для создания и подтверждения достаточной уверенности в том, что продукт, процесс или услуга удовлетворяют установленным требованиям к качеству.

Различают внутреннее обеспечение качества (деятельность по созданию уверенности в выполнении требований к качеству продукции или услуг у руководства организации) и внешнее обеспечение качества (деятельность по созданию такой уверенности у потребителя или других лиц, например, экспертов-аудиторов систем качества, государственных инспекторов по качеству и т.д.).

В состав процедур по обеспечению качества входят внутренние проверки системы качества и другие предупреждающие действия, предусмотренные стандартами ИСО серии 9000 или разработанные по инициативе самой организации исходя из специфики продукции и производственных процессов.

Улучшение качества включает в себя все мероприятия, осуществляемые в организации в целях повышения эффективности и результативности деятельности и процессов для получения прибыли организации и выгоды потребителей[1].

Закключение. В процессе оценивания систем менеджмента качества должны быть получены ответы на следующие четыре основных вопроса в отношении каждого оцениваемого процесса:

- а) Выявлен и определен ли соответствующим образом процесс?
- б) Внедрены и поддерживаются ли в рабочем состоянии процедуры?
- в) Обеспечивает ли процесс достижение требуемых результатов?

Оценивание системы менеджмента качества может различаться по области применения и включать в себя такие виды деятельности как аудит (проверку) и анализ системы менеджмента качества, а также самооценку[2].

Далее будут рассматриваться соответствующие стандарты качества.

Использованная литература

1 <http://quality.eur.ru/MATERIALY3/q-subsys.htm>

2 ГОСТ Р ИСО 9000-2008 Системы менеджмента качества. Основные положения и словарь

3 Шарипбай А.А., Ниязова Р.С., Кузенбаев Б.А. Жоғары оқу орнындағы сапа менеджмент жүйесіне талдау жасау. Сборник научных Актуальные научные исследования в современном мире. iScience ISSN 2524-0986.

УДК 35.071.51+341.225.5

Л.С. ИСКАЛИЕВА, С.Н. БАЙБЕКОВ

РАЗРАБОТКА НОВОГО МЕТОДА ОКРУЖНОСТЕЙ ДЛЯ УСТАНОВЛЕНИЯ ВНЕШНИХ ГРАНИЦ МОРСКИХ ТЕРРИТОРИЙ

(Институт химии угля и технологии, г. Астана, Казахстан)

Установление внешней границы морской территории является довольно сложной задачей. Например, согласно ст. 3 и ст.4 Конвенции ООН по морскому праву 1982 года внешней границей территориального моря является линия, каждая точка которой находится от ближайшей точки исходной береговой линии на расстоянии, равном 12 морских миль (*далее – морск.м*). Согласно ст.33 и ст.55 этой Конвенции внешняя граница прилегающей зоны и исключительно экономической зоны не может распространяться соответственно за пределы 24 и 200 морск.м от исходных береговых линий, от которых отмеряется ширина территориального моря. Согласно ст.76 этой Конвенции, внешняя граница континентального шельфа должна находиться не далее 350 морск.м от исходных линий, от которых отмеряется ширина территориального моря [1].

Из этих определений следует, что внешняя граница указанных морских территорий может устанавливаться одинаковой методикой. С другой стороны, нахождение каждой точки внешней границы, например, территориального моря, порождает новую проблему – как определить и найти самую ближайшую точку исходной линии, чтобы она была на 12 морск.м расстоянии от рассматриваемой точки внешней границы.

Очевидно, что аналитического решения этой задачи не существует, т.к. нам неизвестно уравнение исходной береговой линии, от которой отмеряется ширина территориального моря.

Естественно, моделирование построения внешней границы морских территорий на компьютере сильно облегчает решение этой проблемы, но, с другой стороны составление компьютерной программы является достаточно объемной и трудоемкой, также сложной задачей. Поэтому предлагаем новый довольно простой метод решения указанной проблемы. Назовем этот метод - **методом окружностей**.

Суть этого метода заключается в следующем.

Сначала берем окружность определенного диаметра. Затем устанавливаем ее на произвольную точку исходной линии, так чтобы она прикоснулась к исходной линии. После этого берем другую окружность такого же диаметра и устанавливаем ее рядом с первой, чтобы и она тоже прикоснулась к исходной линии. Аналогично устанавливаем третью, четвертую и других окружностей, чтобы они все прикоснулись к исходной линии. Очевидно, чем ближе друг другу и теснее будут расположены окружности на исходной линии, тем точнее будет решена указанная задача.

После этого с внешней стороны окружностей (т.е. со стороны водного простора) проводим огибающую их линию. Полученная линия будет внешней границей рассматриваемой морской территории. Эта граница на рис.1 показана пунктирной линией. Заметим, если диаметр окружностей равен 12 морск.м, то огибающая линия будет внешней границей территориального моря. Если диаметр окружностей берется равным 24 морск.м, то устанавливается внешняя граница прилегающей зоны. Если имеем дело с исключительной экономической зоной, то диаметр окружностей не должен превышать 200 морск.м. Если диаметр равен 350 морск.м, то в определенных случаях согласно ст. 76 вышеназванной Конвенции определяется внешняя граница континентального шельфа.

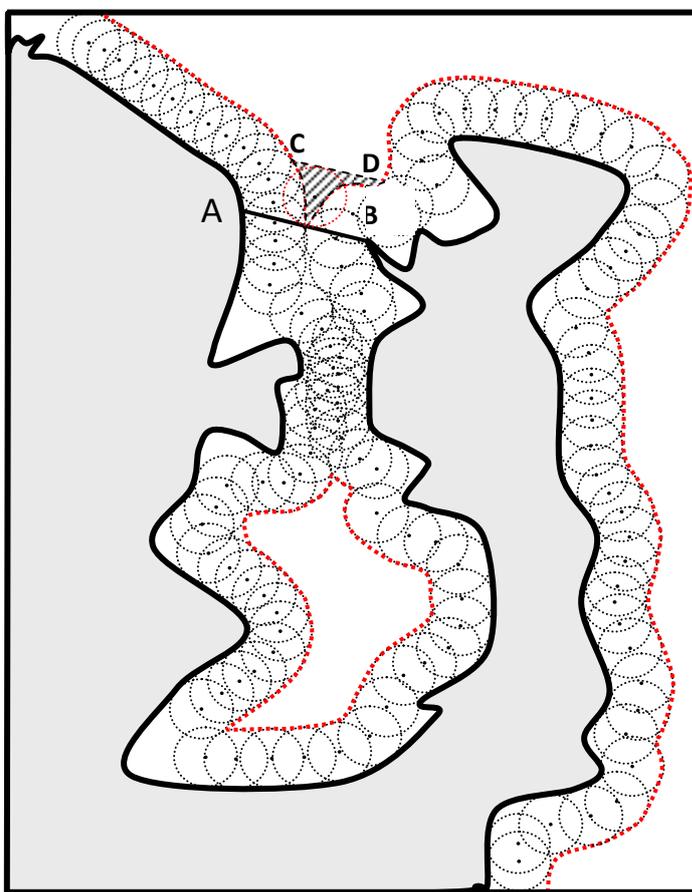


Рис.1. Применение метода окружностей для определения внешних границ морского пространства, заключенного внутри залива

Теперь более подробно рассмотрим разные случаи построения внешней границы территориального моря.

Сначала рассмотрим случай залива и построим для него внешнюю границу территориального моря. Полученная картина показана на рис.1, где диаметр всех окружностей равен 12 морск.м. Из этого рисунка видно, что водное пространство,

находящееся между исходной береговой линией и линией, огибающей все окружности (на рис.1 эта линия указана как пунктирная линия), является территориальным морем, причем огибающая линия является его внешней границей. При детальном рассмотрении этой картины напрашивается следующий резонный вопрос, который касается водному пространству, которое находится во внутренней и нижней части залива. Это пространство с одной стороны находится за территориальным морем, с другой стороны оно со всех сторон окружено территориальным морем. Поэтому возникает вопрос, каким статусом обладает это водное пространство – статусом территориального моря или прилежащей зоны или чего-то другого?

Мы полагаем, что:

1. Если указанная на рисунке суша, т.е. «материк и полуостров», полностью принадлежит одному государству, то, согласно п.3 ст.7, п.1 ст.8 и п.п.4 и 5 ст. 10 Конвенции 1982 года, рассматриваемое водное пространство является не только территориальным морем, но и внутренней водой этого государства.

2. Если указанные на рисунке 1 полуостров и материковая суша принадлежат разным государствам, то рассматриваемое водное пространство необходимо делить между этими государствами, следовательно, это пространство не является ни внутренней водой и не территориальным морем. Она в зависимости от размера становится прилежащей зоной или исключительной экономической зоной и т.д.

Подобные примеры часто встречается в международном праве по проблемам морских территорий [2].

Рассмотрим еще один интересный случай, проблема которого не решается тривиально. Если ширина естественного входа в залив не превышает 24 морск.м, воды, заключенные внутри залива, согласно п.4 ст.10 Конвенции 1982 года считаются внутренними водами. А если ширина естественного входа в залив превышает 24 морск.м, то согласно п.5 ст. 10 Конвенции 1982 года прямая исходная линия в 24 морск.м проводится внутри залива таким образом, чтобы линией такого протяжения было ограничено возможно большее водное пространство.

Как видно из рис.1, ширина естественного входа рассматриваемого залива превышает 24 морск.м. Затем по мере приближения во внутреннее пространство залива ширина его входа сужается. При этом возникает следующий вопрос – как надо проводить эту прямую длиной в 24 морск.м, чтобы охватить максимальное водное пространство?

Естественно, в узком месте входа залива прямую длиной в 24 морск.м можно проводить разными способами. Но нас интересует единственный случай, когда эта прямая охватывает самое максимальное водное пространство внутри залива.

Для решения этой проблемы предлагаем следующий алгоритм.

Рассмотрим случай залива, ширина естественного входа которого больше 24 морск.м (рис.2). На этом рисунке для удобства показана часть входа в залив, при этом надо подразумевать, что внутренняя часть залива находится в нижней части рисунка.

Пусть кривая линия I-I является левым берегом, а кривая линия II-II - правым берегом залива.

Попробуем построить прямой отрезок длиной 24 морск.м между левым и правым берегами входа в залив, чтобы эта прямая охватила максимально большее водное пространство указанного залива. Используем метод окружности. Первоначально рассмотрим левый берег залива. Сначала построим окружность радиуса $R=24$ морск.м так, чтобы ее центр находился на верхней части исходной линии I-I левого берега залива, а сама окружность прикоснулась к правому берегу II-II. Если прикосновения не будет (как например, в случае окружности с центром в точке O), то центр этой окружности сдвинем вниз по линии левого берега. Если и при этом не будет прикосновения, то центр окружности еще раз сдвинем вниз по линии левого берега. Процесс продолжим до тех пор, чтобы искомое прикосновение имело место. Это показано на рис.2, где точка A является центром

окружности, точка B есть точка прикосновения окружности с правым берегом. Длина отрезок AB равна радиусу окружности.

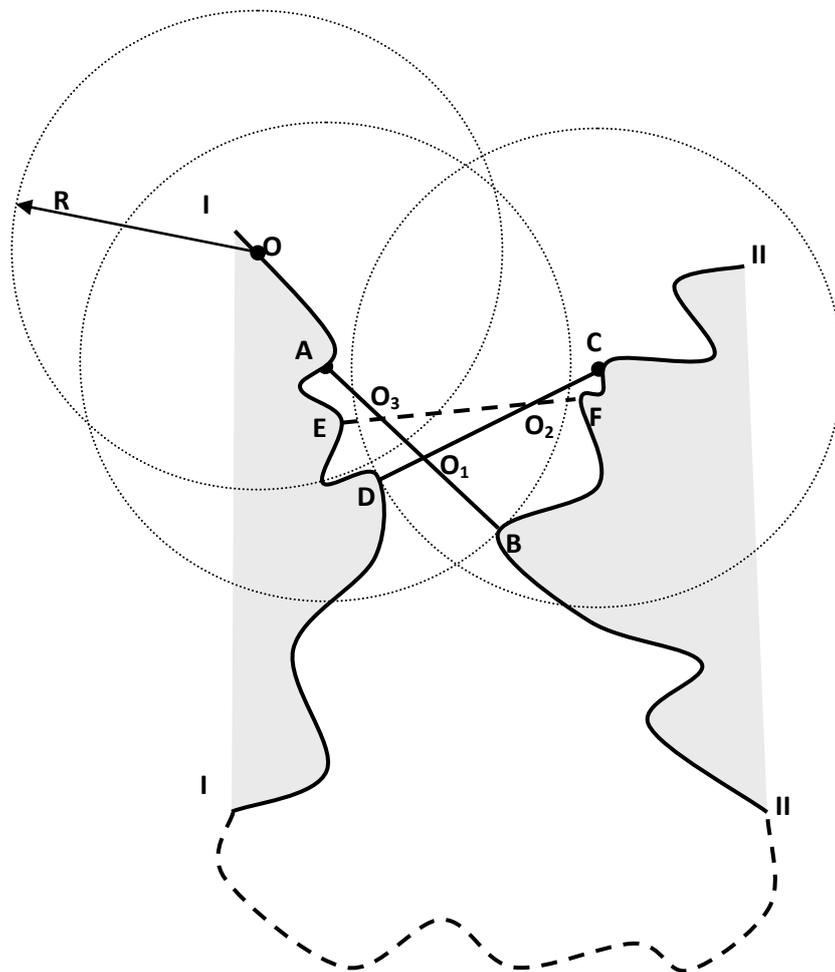


Рис.2. Применение метода окружностей для случая залива, ширина естественного входа которого больше 24 морск.м.

После этого эту процедуру повторим для правого берега залива так, чтобы центр окружности находился на исходной линии II-II правого берега, а сама окружность прикоснулась к левому берегу I-I. Полученная картина показана на рис.2, где точка C является центром окружности, а точка D есть точка прикосновения окружности с левым берегом. Прямая CD есть радиус этой окружности равный 24 морск.м.

Теперь рассмотрим, какой из этих отрезков (AB и CD) длиной 24 морск.м ограничивает большее водное пространство залива. В данном случае прямая CD охватывает большее пространство, чем прямая AB , так как площадь пространства CO_1BFC больше чем площадь пространства AO_1DEA .

В данном случае мы рассмотрели две крайние случаи. Не исключено, что прямую длиной в 24 морск.м можно построить между береговыми кривыми AD и CB так, чтобы с одной стороны один конец этой прямой находился на кривой AD , а другой конец – на кривой CB и с другой стороны - эта прямая могла охватить еще большее пространство, чем прямые AB и CD . Заметим, что это зависит от конфигурации кривых AD и CB . Для данного случая, показанного на рис.2, из всевозможного семейства прямых длиной 24 морск.м, прямая EF

охватывает максимальное пространство залива. В этом случае площадь EO_2D больше чем площадь CO_2F , также площадь FO_3B больше площади AO_3E , поэтому площадь водного пространства залива, ограниченная прямой EF больше чем площади ограниченные прямыми AB и CD в отдельности. O_2

Следовательно, прямая EF является искомой прямой длиной в 24 морск.м, ограничивающая максимальное большее водное пространство залива.

Теперь вернемся п.4 ст. 10 Конвенции ООН по морскому праву 1982 года, согласно которому если расстояние между отметками естественного входа в залив не превышает 24 морск.м, то *замыкающая линия* может быть проведена между этими двумя отметками, и ограниченные таким образом воды считаются внутренними водами.

Обратим внимание, что, так как здесь речь идет о внутренней воде, то каждая заинтересованная сторона будет стремиться охватить как можно большее водное пространство. При этом возникает следующая проблема, как определить эти отметки естественного входа в залив и как соответственно проводить указанную замыкающую линию, хотя ее длина меньше 24 морск.м?

Как видно, суть этой проблемы и вышеприведенного на рис.2 примера идентичны. Поэтому решение этой проблемы при помощи вышеприведенного метода окружностей не составляет особого труда.

Теперь рассмотрим еще один интересный случай, вытекающий из примера, показанного на рис.1, где ширина естественного входа в залив превышает 24 морск.м. На этом рисунке сплошной прямой линией показана *прямая исходная линия* AB длиной в 24 морск.м, построенная для данного случая вышеприведенным методом окружностей, показанного на рис.2. Прямая AB была построена, так чтобы она ограничивала максимальное большее водное пространство залива. Так как эта линия является исходной линией, то согласно ст. 3, ст.4 и других статей Конвенции 1982 года каждое государство имеет право устанавливать свое территориальное море, отмеряемого и из этой прямой исходной линии. Определенное нами для этого случая территориальное море приведено на этом рисунке в виде заштрихованной площади. Как видно из этого рисунка, данное территориальное море ограничивается прямой CD . Причем прямые AB и CD между собой являются параллельными и расстояние между ними равно 12 морск.м.

Обратим внимание, что часть морского пространства, находящаяся за пределом территориального моря благодаря вышесказанной прямой исходной линии AB , переходит в категорию территориального моря.

Теперь рассмотрим случай «материк-остров» (рис.3). Согласно ст.121 Конвенции ООН по морскому праву 1982 года любой остров, за исключением искусственных островов и скал, не пригодных для поддержания жизни человека и для самостоятельной хозяйственной деятельности, имеет свои территориальное море, прилежащую зону, исключительную экономическую зону и континентальный шельф.

Для определения внешних границ этих водных пространств используем предложенный выше метод окружности. Полученная картина приведена на рис.3. Пунктирная линия, огибающая окружностей со стороны водного пространства, является внешней границей водных зон. При этом если диаметр окружностей равен 12 морск.м, то огибающая линия будет внешней границей территориального моря. Если диаметр окружностей берется равным 24 морск.м, то огибающая линия устанавливает внешнюю границу прилежащей зоны. Если имеем дело с исключительной экономической зоной, то диаметр окружностей не должен превышать 200 морск.м. Если диаметр равен 350 морск.м, то в определенных случаях согласно ст. 76 этой Конвенции определяется внешняя граница континентального шельфа.

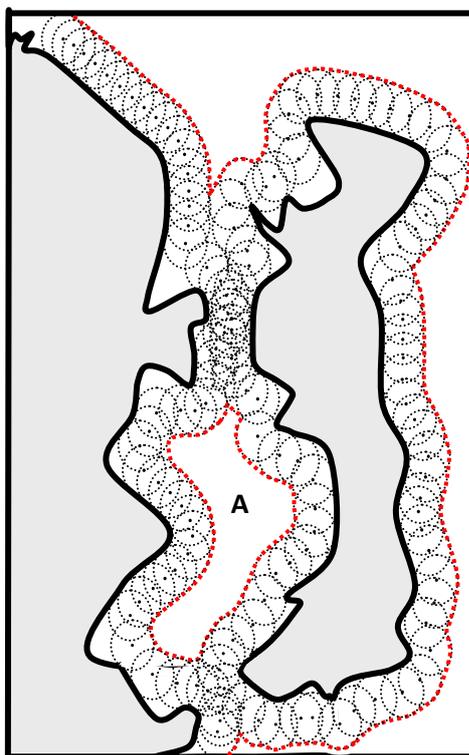


Рис.3. Применение метода окружностей для определения внешних

Из этих случаев более подробно рассмотрим случай территориального моря, при котором диаметр всех окружностей равен 12 морск.м.

Из рисунка 3 видно, что водное пространство, находящееся между исходной береговой линией и линией, огибающей все окружности (на рисунке она показана как пунктирная линия), является территориальным морем, причем огибающая линия является его внешней границей.

При детальном рассмотрении этой картины напрашивается следующий резонный вопрос, который касается водному пространству *A* (рис.3). Водное пространство *A* с одной стороны находится за территориальным морем, с другой стороны оно со всех сторон окружено территориальным морем. (На рис.3 пространство *A* заключено внутри пунктирной линией, которая является огибающей линией, т.е. внешней границей территориальных морей «материка» и острова.) Поэтому возникает вопрос, каким статусом обладает это водное пространство – статусом территориального моря или прилежащей зоны или другое чего-то другого?

Мы полагаем, что:

1. Если указанные на рисунке 3 «материк» и остров полностью принадлежат одному государству, то рассматриваемое водное пространство является территориальным морем данного государства, так как к этому водному пространству нет доступа других государств.

2. Если указанные на этом рисунке «материк» и остров принадлежат разным государствам, то рассматриваемое водное пространство необходимо делить между ними, следовательно, в этом случае это пространство не является территориальным морем. Она в зависимости от размера становится прилежащей зоной или исключительной экономической зоной и т.п.

Литература

1 Конвенция ООН по морскому праву 1982 года.

2 Тихонравов Ю.В. Геополитика. М., ЗАО "Интел-Синтез", 1998 г.

ИСЛАМҒАЗИЕВА Ә.С., КАЗИЕВ Г.З.

ОЦЕНКА ЭФФЕКТИВНОСТИ ФУНКЦИОНИРОВАНИЯ СИСТЕМЫ МЕНЕДЖМЕНТА КАЧЕСТВА ПРИ ПРОВЕДЕНИИ ВНУТРЕННЕГО АУДИТА

(Евразийский национальный университет им. Л.Н. Гумилева, Астана, Казахстан)

Внутренний аудит является высшей формой контроля руководством системы менеджмента качества организации.

Целью внутреннего аудита является экспертная оценка соответствия деятельности и результатов в области качества запланированным мероприятиям, требованиям ISO 9001, а также требованиям, разработанным самой организацией.

Необходимость проведения внутреннего аудита обуславливается тем, что руководители организации хотят знать о реальном состоянии организации. Для этого в организациях создается группа аудиторов, работники которых занимаются анализом всех звеньев управления, участков производства, каждого технологического цикла. Анализируется рациональность функционирования как системы в целом, так и каждого элемента, уровень взаимосвязи между ними, соответствие должностных инструкций, контролируют их реальное выполнение, выявляются слабые места в работе организации и даются рекомендации по устранению имеющихся недостатков.

Актуальность и недостаточная разработанность отдельных проблем внутреннего аудита и оценка его эффективности, всевозрастающая практическая значимость и потребность в использовании внутреннего аудита предприятиями обусловили выбор ее структуры и основные исследования.

Внутренние аудиторы могут также решать другие организационно-управленческие, правовые, технологические, технико-экономические и иные задачи, связанные со спецификой видов деятельности предприятия.

Внутренний аудит СМК является одним из ключевых процессов СМК. По мнению специалистов, если такие процессы СМК, как внутренний аудит, корректирующие и предупреждающие действия и анализ со стороны руководства отлажены и функционируют в соответствии с требованиями СТ РК ИСО 9001–2009, то и все другие процессы СМК будут не только работать, но и совершенствоваться с каждым годом.

Аудит СМК, представляя собой аналитическую по своему характеру работу, должен осуществляться на основе рекомендаций ГОСТ ISO 19011-2013 «Руководящие указания по аудиту систем менеджмента». В этой работе участвуют, как правило, многие работники предприятия, в первую очередь аудиторы, проверяемый персонал, руководители разных уровней.

Для осуществления процесса внутреннего аудита необходимо реализовать следующие функции:

- планирование аудитов;
- разработка заданий;
- оформление результатов аудита;
- обеспечение контроля над устранением несоответствий.

Процесс внутреннего аудита представляет собой деятельность аудиторов по получению объективных данных о функционировании СМК и оцениванию на их основе степени соответствия установленным критериям аудита.

Организация процесса проведения внутренних аудитов СМК начинается с план-графика проведения внутреннего аудита. На основании утвержденного графика, до начала проверки определенного подразделения организации назначается группа аудиторов. Назначенный аудитор оформляет задание представителю подразделения. В процессе

проведения внутренней проверки аудиторы проводят на проверяемых участках сбор информации путем наблюдений, опросов и изучения документов. При обнаружении несоответствия аудитор фиксирует в протоколе несоответствий и доводит сведения до руководителя проверяемого подразделения после завершения проверки. После завершения проверки аудиторы оформляют отчет об итогах проверки.

С помощью унифицированного языка моделирования (UML), графическим языком для визуализации, построены визуальные модели организации проведения внутреннего аудита.

При визуальном моделировании на UML используются девять видов диаграмм, каждая из которых содержит элементы определенного типа. На диаграммах взаимодействия представлены связи между объектами; показаны, в частности, сообщения, которыми объекты обмениваются (рис.1, 2).

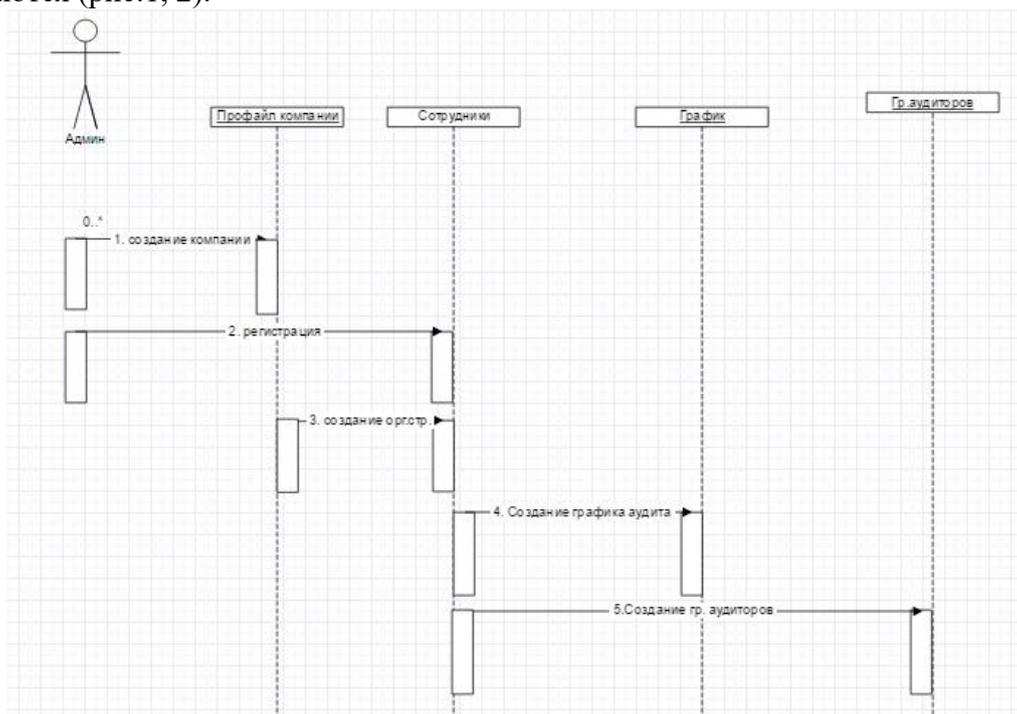


Рисунок 1. Диаграмма последовательности.

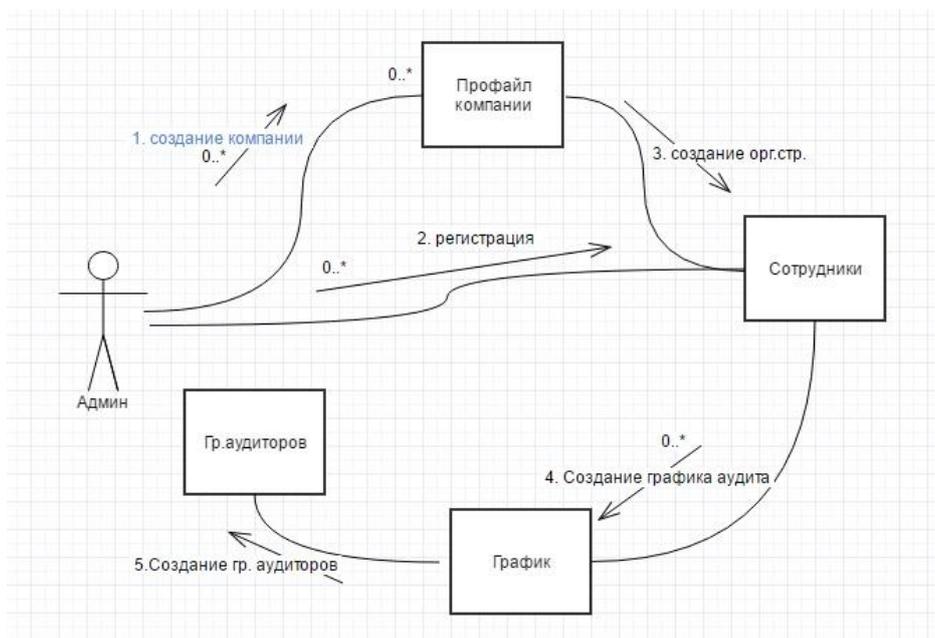


Рисунок 2. Диаграмма кооперации.

В условиях работы существует достаточно серьезная проблема поиска средств оптимизации труда аудитора, снижения трудоемкости с одновременным соблюдением высокого качества работы. Одним из путей решения этой задачи является развитие экспертных систем аудита (ЭСА). Экспертные системы можно характеризовать как комплекс тематических вопросов, сравнительных показателей и ситуаций, алгоритмов, позволяющий аудитору получить одновременно информацию и результат обработки такой информации

Экспертные системы можно определить как компьютерные программы, включающие оригинальную технику программирования для обеспечения обработанной информацией пользователей ЭВМ. На её основе с помощью обработки информации можно осуществить анализ.

Базовыми компонентами экспертной системы является информационная база, которая содержит правила и информацию, пункты стандарта, используемые при формировании заданий, отчета о проведенном аудите. Экспертные системы используются для оказания помощи аудитору в определении производственного потенциала компании и оценки эффективности системы менеджмента качества путём уменьшения документооборота.

Список использованной литературы

1. Крылова, Т.Д. Зарубежный опыт управления качеством Текст. / Т.Д. Крылова. М.: Изд-во стандартов, 1992. - 140 с.
2. Подлипаев, Л.А. Технология внедрения и постоянного улучшения системы менеджмента качества на предприятии Текст. / Л.А. Подлипаев. -М.: Гелиос АРВ, 2004. 408 с.
3. Адлер, Ю.П. Опыт внедрения современной системы качества на АО «Автоагрегат». Первые итоги Текст. / Ю.П. Адлер [и др.] // Стандарты и качество. 1998. - № 10. - С. 67 - 70.
4. Белобрагин, В. Системы управления и их эволюция Текст. / В. Белобрагин // Стандарты и качество. 2007. - № 4 - С. 56 - 59.

¹КАЛИМОЛДАЕВ М.Н., ²МАМЫРБАЕВ О.Ж., ³КЕЙЛАН А.

МЕТОД ПРОЕКТИРОВАНИЕ СИСТЕМЫ МНОГОМОДАЛЬНОГО РАСПОЗНАВАНИЯ КАЗАХСКОЙ РЕЧИ

^{1,2}*Институт информационных и вычислительных технологий, Казахстан, Алматы,*
³*School of Science and Engineering, Tokyo Denky University, Hatoyama-Machi, Hiki-Gun, Japan)*

Введение. Разработка системы эффективного взаимодействия на естественном языке человека с компьютером сегодня является одним из приоритетных направлений развития области распознавания речи. Это связано с тем, что ресурсы вычислительной техники не используются в полной мере из-за отсутствия полноценного, привычного для человека, интерфейса для взаимодействия пользователя с компьютером.

С развитием современных речевых технологий появилась принципиальная возможность перехода от формальных языков-посредников между человеком и машиной к естественному языку в устной форме, как универсальному средству выражения целей и желаний человека. Речевая форма диалога обладает рядом преимуществ, таких, как естественность, оперативность, смысловая точность ввода, освобождение рук и зрения пользователя, возможность управления и обработки в экстремальных условиях. Считая, что речь является наиболее естественной формой взаимодействия, многие разработчики пытались заменить ряд интерфейсов многих систем речевым интерфейсами и заставить компьютеров и пользователей выполнять все функции с помощью голоса.

Каждая такая система имеет некоторые задачи, которые она призвана решать и комплекс подходов, которые применяются для решения поставленных задач. Рассмотрим основные признаки, по которым можно классифицировать системы распознавания человеческой речи и то, как этот признак может влиять на работу системы.

Размер словаря. Очевидно, что чем больше размер словаря, который заложен в систему распознавания, тем больше частота ошибок при распознавании слов системой. Например, словарь из 10 цифр может быть распознан практически безошибочно, тогда как частота ошибок при распознавании словаря в 100000 слов может достигать 45%. С другой стороны, даже распознавание небольшого словаря может давать большое количество ошибок распознавания, если слова в этом словаре очень похожи друг на друга.

Дикторозависимость или дикторонезависимость системы. По определению, дикторозависимая система предназначена для использования одним пользователем, в то время как дикторонезависимая система предназначена для работы с любым диктором. Дикторонезависимость – труднодостижимая цель, так как при обучении системы, она настраивается на параметры того диктора, на примере которого обучается. Частота ошибок распознавания таких систем обычно в 3-5 раз больше, чем частота ошибок дикторозависимых систем.

Раздельная или слитная речь. Если в речи каждое слово разделяется от другого участком тишины, то говорят, что эта речь – раздельная. Слитная речь – это естественно произнесенные предложения. Распознавание слитной речи намного труднее в связи с тем, что границы отдельных слов не четко определены и их произношение сильно искажено смазыванием произносимых звуков.

Назначение. Назначение системы определяет требуемый уровень абстракции, на котором будет происходить распознавание произнесенной речи. В командной системе (например, голосовой набор в сотовом телефоне) скорее всего, распознавание слова или фразы будет происходить как распознавание единого речевого элемента. А система диктовки текста потребует большей точности распознавания и, скорее всего, при интерпретации произнесенной фразы будет полагаться не только на то, что было произнесено в текущий момент, но и на то, как оно соотносится с тем, что было произнесено до этого. Также, в системе должен быть встроен набор грамматических правил, которым должен удовлетворять произносимый и распознаваемый текст. Чем строже эти правила, тем проще реализовать систему распознавания и тем ограниченной будет набор предложений, которые она сможет распознать.

Основные подходы к решению задачи распознавания речи. Первый подход, который используется для улучшения показателей распознавания речи, основывается на выделении векторов свойств из сигнала с учетом особенностей восприятия звука человеческим ухом. Он включает в себя анализ несущих частот и выравнивание сигнала по громкости. Наиболее распространенными технологиями, использующими такой подход, являются метод кепстральных коэффициентов тоновой частоты (Mel Frequency Cepstral Coefficients, MFCC, Davis & Mermelstein, 1980) и метод коэффициентов линейного предсказания (Perceptual Linear Prediction, PLP, Hermansky, 1990). Одновременное и опережающее сопоставление с шаблоном (маскирование) (Paliwal & Lilly, 1997), характерное для человеческого восприятия, может быть смоделировано и использовано для выделения свойств, обеспечивающих большую устойчивость от шумов. С этой целью был создан метод варьирования размерностей кадров (Variable Frame Rate analysis, VFR, Zhu & Alwan, 2000). Учитывая специфику работы нервных клеток, отвечающих за слуховые рецепторы, был предложен метод диапазонной автокорреляции (Subband-Autocorrelation, SBCOR, Kajita & Itakura, 1994).

Другой подход основан на анализе звуковых сигналов. Различие поступающих в систему зашумленных сигналов от шаблонов, полученных в ходе обучения «чистыми» сигналами, является основной причиной неустойчивости работы систем распознавания. Целью подхода является уменьшение этого различия. Предполагается, что шум в звуковых сигналах аддитивный и стационарный. Оценки среднего значения усредненного шума

вычитаются из кепстра (Cepstral Mean Subtraction, CMS, Furui, 1981) или спектра (Spectral Subtraction, SS, Virag, 1999), вычисленного по зашумленным данным. Некоторые модификации таких методов включают в себя нелинейное спектральное вычитание (Non-linear Spectral Subtraction, NSS, Lockwood & Boudy, 1992), которые используют спектральные огибающие. Такие техники требуют хорошей оценки шума, которую на практике бывает сложно получить, особенно в случае нестационарного фонового шума.

Еще одним способом борьбы с разницей между полученными свойствами из зашумленных и чистых сигналов является использование высокочастотного фильтра. Предполагается, что шум в сигнале не стационарный, а медленно изменяющийся во времени. Метод RASTA (Relative Spectral Analysis, Hermansky & Morgan, 1994) представлен таким образом, что относительные спектральные изменения фиксируются. И те медленные изменения, которые были вызваны шумом, удаляются. В этом случае отпадает необходимость в явном оценивании шума.

Третий подход основан на использовании многомерных пространств (Ephraim & Trees, 1994). Основной идеей этого подхода является нахождение линейного отображения, которое минимизирует функцию стоимости. Часто в качестве такого отображения берется умножение вектора свойств на матрицу преобразования. Примерами данного подхода могут служить основной компонентный анализ (Principal Component Analysis, PCA) и независимый компонентный анализ (Independent Component Analysis, ICA, Koscor, 2000), а также проектирование на многомерные подпространства (Gales, 2002).

Соответствующие системы автоматического распознавания речи пока еще значительно уступают речевым способностям человека, по причине их недостаточной адекватности к естественной речи, что ограничивает применение речевых технологий в промышленности и быту. Для решения этих проблем человеко-машинного взаимодействия многие исследователи стали использовать дополнительные виды каналов передачи информации (аудио, видео и т.д.). В результате для улучшения качества распознавания начали разрабатывать так называемые многомодальные методы распознавания речи. Такие системы позволяют более качественно обеспечить эффективное взаимодействие с различными автоматизированными средствами управления и коммуникации.

В моровой науке, в настоящее время, использование многомодальных методов распознавания речи актуально в смартфонах (умный телефон), в которых возможен отдельный ввод с помощью голоса, незргономичной клавиатуры или сенсорного экрана. Совместное использование этих коммуникативных каналов позволит пользователю более оперативно и надежно обмениваться информацией с такими устройствами.

Методы. При создании системы распознавания речи требуется выбрать, какой уровень абстракции адекватен поставленной задаче, какие параметры звуковой волны будут использоваться для распознавания и методы распознавания этих параметров. Рассмотрим основные различия в структуре и процессе работы различных систем распознавания речи.

По типу структурной единицы. При анализе речи, в качестве базовой единицы могут быть выбраны отдельные слова или части произнесенных слов, такие как фонемы, ди- или трифоны, аллофоны. В зависимости от того, какая структурная часть выбрана, изменяется структура, универсальность и сложность словаря распознаваемых элементов.

По выделению признаков. Сама последовательность отсчетов давления звуковой волны – чрезмерно избыточна для систем распознавания звуков и содержит много лишней информации, которая при распознавании не нужна, либо даже вредна. Таким образом, для представления речевого сигнала из него требуется выделить какие-либо параметры, адекватно представляющие этот сигнал для распознавания.

По механизму функционирования. В современных системах широко используются различные подходы к механизму функционирования распознающих систем. Вероятностно-сетевой подход состоит в том, что речевой сигнал разбивается на определенные части (кадры, либо по фонетическому признаку), после чего происходит вероятностная оценка того, к какому именно элементу словаря распознаваемого словаря имеет отношение данная часть и

(или) весь входной сигнал. Подход, основанный на решении обратной задачи синтеза звука, состоит в том, что по входному сигналу определяется характер движения артикуляторов речевого тракта и, по специальному словарю происходит определение произнесенных фонем.

В современных системах распознавания речи задача понимания смысла чаще всего решается методом «снизу-вверх», т.е. сначала происходит распознавание речевых сегментов, а затем все распознанное поступает на семантический модуль. Как правило, сигнал на входе семантического блока представляет собой матрицу, составленную из векторов вероятности распознавания каждого сегмента потока речи, который соответствует при удачной сегментации какому-либо слову или словоформе. Дальнейшая работа семантического блока предполагает построение из этих векторов вероятности списка осмысленных предложений, ограниченных заданным порогом минимальной вероятности. Естественная речь зачастую аграмматична и практически сложно применить грамматику для построения высказывания, учитывая еще и то, что падежные окончания во флексивных языках чаще всего «заглатываются», т.е. не проговариваются достаточно четко. Поэтому используют другие разнообразные «улучшители» понимания как, например, учет предыстории, выявление контекста и падежно-ролевых отношений или использование различных статистически-вероятностных методов (частотности, ассоциативности и пр.). Как правило, на данном этапе используется обратная связь семантического модуля с модулем распознавания: список поиска вероятных слов при распознавании пополняется ассоциативной лексикой с последующим пересчетом векторов вероятности. Повторяя цикл можно достичь более высокий процент правильного понимания смысла.

Особыми проблемами при таком подходе является, как уже указывалось, омонимия и так называемый «мусор» – слова, которых нет в словаре распознавания, а так же различного рода помехи как речевого, так и неречевого типа. Если степень омонимии можно уменьшить, выявляя и запоминая контекст сообщения, то проблема «мусора» не имеет простого решения, поскольку здесь помимо внешних помех необходимо выявлять и учитывать индивидуальные характеристики говорящего (хезитации, употребление эмотивных лексических элементов).

Языковые модели. С ростом словаря увеличивается количество слов, схожих или даже одинаковых по звучанию. При слитном произнесении акустическая схожесть отдельных фрагментов речи проявляется настолько, что часто и человек, прослушивая запись вне контекста, не может в точности распознать то, что было произнесено. Поэтому значительную роль в распознавании речи играют так называемые языковые модели. Они позволяют определить наиболее вероятные словные последовательности. Сложность построения языковой модели во многом зависит от конкретного языка. Так, для английского языка достаточно использовать статистические модели (так называемые N-граммы). Для высокофлективных языков (языков, в которых существует много форм одного и того же слова), к которым относится и русский, языковые модели, построенные только с использованием статистики, уже не дают такого эффекта – слишком много нужно данных, чтобы достоверно оценить статистические связи между словами.

Качество распознавания зависит от двух факторов – структуры каркаса системы распознавания речи (набора программных модулей и алгоритмов, использующихся при распознавании) и качества моделей – акустических, языковых, тематических.

Все модели обучаются с использованием большого объема материала. Так, для акустических моделей используются сотни часов записей речи тысяч дикторов. Для повышения устойчивости распознавания к помехам и искажениям, при обучении используются записи в различных каналах и различных условиях. Для обучения языковых моделей и моделей тематик используются текстовые корпуса объемом от сотен миллионов словоформ до нескольких миллиардов. Подготовка такого объема обучающего материала – это сложная и кропотливая работа. Центр речевых технологий в течение нескольких десятилетий накапливал обучающий материал и на данный момент обладает уникальным по

своим объемам, разнообразию и качеству набором записей и текстов, способных обеспечить высочайшее качество распознавания речи.

Созданная система рассматривает двухмодальный (речь и губы) метод распознавания. По первому каналу поступает речевой сигнал от микрофона, а из второго канала поступает сигнал от видео камеры Kinect, описывающий движение губ. Разработанная многомодальная система, получая информацию из двух каналов совместно обрабатывая их, позволяет более качественно распознавать речь. В результате, установлено, что качество распознавания речи на основе двухмодального подхода выше, чем качество, полученное от отдельно взятой речи в системах учета.

В предварительной обработке речевого сигнала важна, точность определения моментов начала и окончания слов. Возможность обнаружения моментов начала и окончания фразы, позволяющие обрабатывать только те сегменты, в которых имеются речевые сигналы, позволяет существенно уменьшить вычислительные процедуры. Вследствие этого, скорость и эффективность обработки будут увеличиваться. Как показали многочисленные исследования, в речи могут содержаться до 50 % пауз, а в диалоге их объем может достигать до 70%. В связи с этим, было создано множество различных алгоритмов, которые в свою очередь устраняют избыточность речи, пытаясь выделить только значимые параметры речи .

Voice activity detector (VAD) – является технология сжатия речевого сигнала, за счет поиска речи и пауз и их кодирования. Алгоритм VAD работает в процессе кодирования речевого сигнала перед распознаванием речи. Наличие пауз в речевом сигнале определяется на основе анализа и синтеза речевых данных. Предположим что, речь содержит паузу, которую необходимо предсказывать, определения паузы является самым сложным аспектом в алгоритме VAD (Рис. 1). Алгоритм VAD работает на основе вычисления значений

кратковременной энергии $E_n = \sum_{m=-\infty}^{\infty} [x(m)w(n-m)]^2$ (или кратковременное значение модуля энергии).

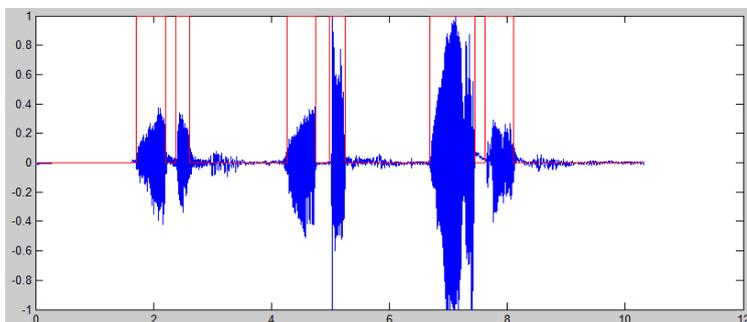


Рис. 1. Результаты определения паузы в речевом сигнале с помощью VAD

В наиболее простых реализациях наличие паузы в наборе цифровых отсчетов определяется на основе сравнения суммарной энергии пакета речевых данных с некоторым пороговым значением. В таком случае необходимо подобрать порог так, чтобы не допустить слишком частое устранение ошибочных пауз, т.к. это может привести к ухудшению качества и потере полезных данных, что может послужить снижению эффективности алгоритма VAD. Обычно, для определения пауз, применяется сложный алгоритм, учитывающий не только энергию пакета, но и энергию спектральных составляющих отрезка сигнала. После алгоритма VAD применяется алгоритм сегментации речевого сигнала. Сегментация речевого сигнала необходима для выделения характерных признаков голоса диктора на определённых участках речевого сигнала и восстановления формы речевого тракта по акустическому признаку, которая может быть использована в синтезе речи по входному тексту и распознавании речи [10, 11].

Литературный обзор. С развитием речевых технологий и все большим внедрением мобильных устройств, возникла идея применения речевого управления при построении

сетевых приложений. Для этого было необходимо разработать унифицированный стандарт для интеграции речевых технологий.

Один из открытых стандартов на основе XML-языка – VoiceXML (Voice eXtensible Markup Language), первая версия опубликована в мае 2000 г. международным консорциумом World Wide Web (W3 Consortium) – предназначен для разработки интерактивных голосовых приложений (Interactive Voice Response, IVR) управления медиаресурсами. Цель создания стандарта – привнесение всех преимуществ web-программирования в разработку IVR-приложений.

Однако интерес к многомодальным приложениям, сочетающим распознавание речи с другими формами ввода информации (при помощи клавиатуры, пера или набора цифровых кнопок) побудил ряд компаний, в том числе Microsoft, поддержать проект SALT Forum (Speech Application Language Tags – теги языка речевых приложений). И теперь вокруг SALT и VoiceXML консорциума W3C формируются два разных лагеря. До сих пор компании не могут прийти к единому мнению о выборе главного стандарта, и сейчас оба направления развиваются в равной степени.

В настоящее время за рубежом многомодальные методы распознавания речи уже используются в некоторых прикладных областях: картографических системах, системах виртуальной реальности, медицинских системах, робототехнике, web-приложениях, и т.д. [1]. Помимо этого, многомодальные методы распознавания речи могут быть полезны в мобильных устройствах, где имеются ограничения для ввода информации с помощью обычной клавиатуры. Сейчас в карманных персональных компьютерах и современных мобильных телефонах используются системы распознавания рукописного текста и речи на естественном языке. Комбинирование таких систем с голосовым вводом позволит обмениваться информацией с пользователем более эффективно. Исследования, посвященные распознаванию речи, мимики лица, положения человека в окружающем пространстве ведутся уже более полувека. Однако, системы объединяющие различные способы ввода информации в единой форме стали разрабатываться совсем недавно. Такие распознающие системы используют многомодальных (мультимодальных) методов распознавания речи. Многомодальные методы распознавания речи обрабатывают данные полученных из двух и более каналов ввода информации например, как речь и движения губ. Китлер (Kittler) и другие ученые [2, 3] объединили вероятностные классификаторы Байеса в одну систему для распознавания речи. В такой системе учитываются данные полученные из трех каналов – речь, фронтальная и профильная части лица. Чибулуши (Chibelushi) и другие [4, 5, 6] разработали методы комбинированного распознавания речи, состоящие из речи и изображений лица.

Звуковой строй казахского языка стал предметом изучения еще с XIX века. Казахский язык, как один из тюркских языков, относится к кыпчакской группе. Тюркология в последнее время сделала заметный поворот к изучению закономерностей, лежащих в сфере речи. Обращение к фактам разговорной речи обнаруживает ряд синтаксических примеров, которые не только интересны по организации и функции, но и расширяют наши представления о соотношении между языком и речью, о возможностях речевой реализации общезыковых моделей, схем.

Фонетическую систему языка образуют не только сегментные средства, но и суперсегментные, которые накладываются на фонемную (линейную) структуру речи. Минимальными единицами звуковой системы языка являются фонемы (гласные, согласные), являющиеся элементами звуковой оболочки слов и морфем. Фонемы составляют сегментный уровень речевого звучания и определяют линейности речи, вторую линию речевого звучания составляет суперсегментный уровень, т. е. интонация. Интонационный (просодический) уровень имеет достаточно сложное строение и выполняет весьма многообразные функции. Просодические средства в каждом языке функционируют в виде тональных, динамических и темпоральных модификаций и изменений одновременно с сегментами речевого потока.

Звуки казахского языка по их артикуляционной близости разделены на две категории: восемь звуков отнесены к группе гласных (а -э, е, ы, і, о, ө, ұ, ү), девятнадцать звуков – к согласным (п, б, м, у, н, д, т, ж, з, ш, с, р, л, ј, к, г, қ, ғ, ң). Переднеязычный широкий э не выделяется как самостоятельная фонема. В составе согласных отсутствуют аффрикаты дж, ч и спирант һ. Приведем отдельные сведения о некоторых артикуляционных особенностях казахских звуков, комбинаторных чередованиях их (б~м, н~б~т, н~ң, к~г и т. п.) и ударении выделяются твердые и мягкие варианты согласных. В составе казахского вокализма отмечают девять гласных (а, э, е, о, ө, ұ, ү, ы, і), в консонантизме – 20 согласных звуков. Согласные классифицируются им по четырем группам: глухие (к, қ, т, п, с, ш), звонкие (ғ, г, д, б, з, ж), сонорные (н, ң, м, р, л, л') и промежуточные (у, й). По мнению ученых, не совсем обоснованно выделены у, й из состава других сонорных. Представляется неверным также выделение мягкой и твердой разновидности сонорного л в качестве самостоятельных фонем [7]. Казахский ученый, языковед С. К. Кенесбаев подробно анализирует фонемный состав, закон сингармонизма, звуковые изменения, слоговую структуру слов и категорию акцентуации казахского языка [8]. При анализе фонологической структуры казахских слов, определено, что казахский вокализм состоит из одиннадцати фонем: девяти монофтонгов (а, э, е, ө, о, ұ, ұ, ы, і) и двух дифтонгоидов (и, у). С. К. Кенесбаев подчёркивает восточное происхождение фонемы э в казахском языке и наличие нескольких согласных (в, х, ф, ч, щ) заимствованных из русского языка. Классификация согласных проводится в трех направлениях: по участию голоса и шума, по способу и месту образования. По месту образования согласные делятся на восемь групп: губно-губные (билабиальные) – п, б, м (у); зубно-губные (дентолабиальные) – ф, в; зубные (дентальные) – т, с, з, д, ц; альвеолярные – н, л, ч; переднеязычные (палатальные) – р, ш, ж, й; среднеязычные (препалатальные) – к, г; заднеязычные (велярные) – қ, ғ, ң, х; гортанные (фарингальные) – һ.

Состояния губ, соответствующие фонемам устной речи, называются виземами. Предварительный анализ возможности автоматической классификации образов такого алфавита показал необходимость их существенного сокращения в направлении использования базовых или опорных визем [9]. Поэтому, для дальнейших исследований в направлении разработки системы автоматического чтения с губ можно принять следующий рабочий алфавит визем, за основу которого приняты опорные виземы, показанные в рисунке 2.

1	2	3	4	5	6	7
						
а, э /a/, /æ/	ж, ш /jo/, /ʃ/	ы, ң /ɯ/, /ŋ/	д, т, н, р, л /d/, /t/, /n/, /r/, /l/	е /e/	с, з /s/, /z/	і /i/
8	9	10	11	12	13	
						
о, ө /o/, /ø/	у, ұ, ү /w/, /ɣ/, /y/	б, п, м /b/, /p/, /m/	й /j/	в, ф /v/, /f/	г, ғ, қ, к /g/, /ɣ/, /q/, /k/	

Рис. 2. Состав визем казахского языка.

Результаты. В настоящее время известно довольно много одномодальных методов распознавания речи: на основе методов динамического программирования (DTW) [16, 17], скрытых марковских моделей (НММ) [18], с использованием нейросетевых методов и др. Однако, в любом из методов распознавания речи присутствуют ошибки, и задача улучшения качества распознавания является актуальной. Один из основных вариантов решения данной проблемы – комплексное распознавание по нескольким биометрическим признакам.

В данной работе исследуется метод двухмодального распознавания речи, основанный на анализе речевого сигнала и изображения лица говорящего. Интеграция данных подходов рассматривается с точки зрения комбинирования классификаторов, когда каждый метод обучается отдельно по своей модальности. При этом под модальностью подразумевается множество признаков, относящихся к одной группе измерений (например, параметры речевого сигнала или параметры изображения лица говорящего). Распознавание по речевому сигналу обладает относительно высоким, но недостаточным уровнем точности распознавания. По результатам теоретических оценок и проведенных экспериментов установлено, что распознавание речи в комбинации с менее точным методом распознавания, как распознавание по изображению лица говорящего, качество распознавания может быть улучшено.

Целью комбинирования методов распознавания является преодоление указанных недостатков каждого из описанных методов.

Для определения точки-границы между двумя слогами применяется следующий алгоритм [12]:

1. Определения слоговых пиков.
2. Определение точки наименьшей энергией между слоговыми пиками.

В большинстве случаев эта точка является границей между двух слов. Но есть случаи, когда эта точка была конечной точкой шипящего согласного следующего слова. Так что, надо ответить на вопрос – слева от этой точки был ли согласный шипящий или голосовой. Определим шипящих, реализованных по числу переходов речевого сигнала через нуль. Подсчитаем долю числа переходов через нуль для участка длиной в N отсчетов, который находится слева от точки с минимумом энергии и заканчивается отсчетом m :

$$Z_x(m) = \frac{1}{N} \sum_{n=m-N+1}^m \frac{|\text{sgn}(x_n - \text{sgn}(x_{n-1}))|}{2}, \text{ где } N = 256;$$

$$\text{sgn}(x) = \begin{cases} 1, & \text{если } x > 0 \\ 0, & \text{если } x = 0 \\ -1, & \text{если } x < 0 \end{cases}$$

При принятой частоте дискретизации 11025 отсчетов в секунду число переходов сигнала через нуль у щелевого звука всегда больше 14 на 100 отсчетов. Соответственно переходов через нуль больше 0,14, а у голосового сигнала меньше этого числа. Интервал наблюдения равен $100 \cdot \Delta t = \frac{100}{11025}$ [сек], число переходов через нуль 14, т.к. в среднем на 1

период приходится 2 перехода, то в интервале $\frac{100}{11025}$ [сек] укладывается $\frac{14}{2}$ периодов. Если

пропорция $\frac{(Z_x(m) \cdot 100)}{256}$ больше 14, то считаем участок щелевым и считаем переходов на предыдущем участке. Выполнение этого условия определяет отправную начальную точку шипящего согласного.

3. Проверим числа переходов через нуль слева от точки с минимальной энергией, чтобы точно определить точку границы между слогами и это будет точное разделение двух слогов, соответствующих двум словам (рис. 3).

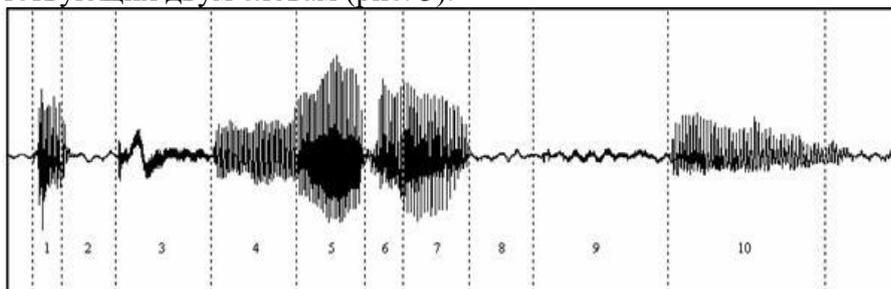


Рис. 3. Сегментация речевого сигнала.

Для распознавания движения губ используются активные модели внешнего вида (Active appearance model, ААМ) [13, 14]. В первую очередь распознаётся вся область лица диктора, а потом выделяет контур губ для распознавания движения губ, что бы улучшить распознавания речи с чтением движения губ (Рис. 4).

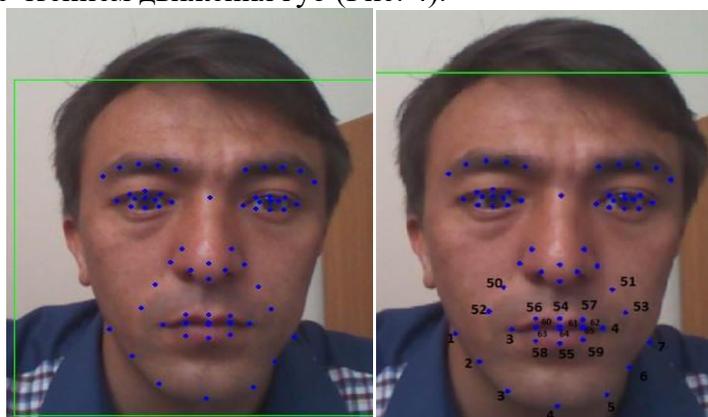


Рис. 4. Распознавания лица и губ.

При выделении контура губ и их составляющие точки по оси x и y формируют матрицу $F=[f_1, f_2, \dots, f_n]$, где $f_n=[x_1, x_2, \dots, x_n, y_1, y_2, \dots, y_n]$ и форма губ имеет следующий вид: $f=f_0+M_n p_n$, где f_0 – нормальная форма губ, M_n – матрица главных векторов, p_n – параметры формы. Для быстрого и надежного отслеживания губ, мы предлагается метод слежения Lukas-Kande feature tracer [15]. Изменяя вектор параметров p_n можно получить разного рода деформации формы для сопоставления ее с виземами. После сопоставления получаем множество Q_{lip} . Распознавание лица и чтение движения губ осуществляется с помощью алгоритма FaceSDK и для видеозаписи была использована камера Kinect model 1517.

Для распознавания осуществляется запись изображений лица диктора, говорящего на цифровую видеокамеру и синхронно записывается речевой сигнал на микрофон. Отдельно по каждой из этих модальностей происходит распознавание, и формируется, на основе алгоритмов распознавания речи и мимики, два множества предполагаемых к распознаванию фонем Q_{lip} и Q_{voice} . Далее ищется пересечение этих множеств $Q=Q_{lip} \cap Q_{voice}$, при этом возможны следующие варианты:

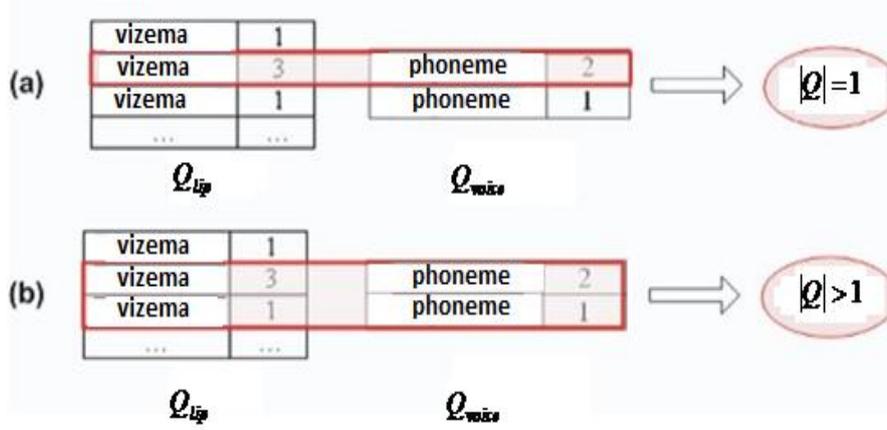


Рис. 5. Комбинирование классификаторов: а) $|Q|=1$; б) $|Q|>1$

1. $Q_{lip} = \emptyset$, и в этом случае фонема не распознаётся.

2. Ситуация $|Q|=1$ представлена на рисунке 5 а. Фонема распознаётся, как попавшая в пересечение двух множеств.

3. В случае $|Q|>1$ считается общее число попаданий в группу "ближайших" по изображению и по голосу одновременно. Например, для ситуации на рисунке 5 б для Фонемы 2 эта величина равна $n_1 = 3 + 2 = 5$, а для Фонемы 3 - $n_2 = 1 + 1 = 2$. Так как $n_1 > n_2$, то фонема идентифицируется как Фонема 2. В случае, когда для нескольких фонем эти величины совпадают, система идентифицирует фонему, но распознавать отказывается. Обозначим через FRR_{lip} и FAR_{lip} соответственно FalseRejectionRate и FalseAcceptanceRate для неполного метода распознавания последовательности фонем по изображению лица, а через FRR_{voice} и FAR_{voice} - по речевому сигналу. Пусть u_o - идентифицированная фонема, а u_s - не идентифицированная. Для рассмотренных методов распознавания эти величины определяются следующим образом:

1. $FRR_{lip} = P(\{u_o \notin Q_{lip}\})$;
2. $FRR_{lip} = P(\{u_{lip} \neq \emptyset\} | u_s)$;
3. $FRR_{voice} = P(\{u_o \notin Q_{voice}\})$;
4. $FRR_{voice} = P(\{u_{voice} \neq \emptyset\} | u_s) = 1$, так как множество всегда непустое.

Оценим FAR и FRR для рассматриваемой двухмодальной системы. Как отмечалось выше, фонема не идентифицируется в случае, если она не попала в пересечение списков $Q = Q_{lip} \cap Q_{voice}$ для обоих методов. Поэтому:

$$\begin{aligned}
 FRR &= P(\{u_o \notin Q\}) = P(\{u_o \notin Q_{lip} \cap Q_{voice}\}) = \\
 &= 1 - P(\{u \in Q_{lip}\}) \cdot P(\{u_o \in Q_{voice}\}) = \\
 &= FRR_{lip} + FRR_{voice} - FRR_{lip} \cdot FRR_{voice}
 \end{aligned}$$

Фонема распознаётся в случае, когда множество оказывается непустым, оценка для FAR имеет следующий вид: $FAR = FAR_{lip} \cdot \frac{k}{n}$. В этой формуле k – число отбираемых эталонных записей фонем, а n - число фонем.

Проведённые эксперименты и полученные теоретические оценки показывают, что предложенный в работе метод комбинирования классификаторов действительно позволяет повысить надёжность распознавания. Для анализа результатов системы, рассмотрено 10 дикторов, показанных в таблице 1, среди которых 5 дикторов мужчины и 5 дикторов женщины. Каждый диктор говорит по 200 слов с разными сложностями.

Таблица 1. Анализ показателей системы распознавания.

Дикторы	Система распознавания речи	Многомодальная система распознавания речи
Мужчина-1	90,1%	93,4%
Мужчина-2	91%	93,6%
Мужчина-3	90,4%	94%
Мужчина-4	89,8%	91,2%
Мужчина-5	90,7%	93,6%
Женщина-1	91,2%	93,6%
Женщина-2	91,3%	93,6%
Женщина-3	91%	93,1%
Женщина-4	92,1%	93,8%
Женщина-5	90%	92,7%

По результатам анализа показателей системы распознавания речи получено следующий график (Рис. 6.):

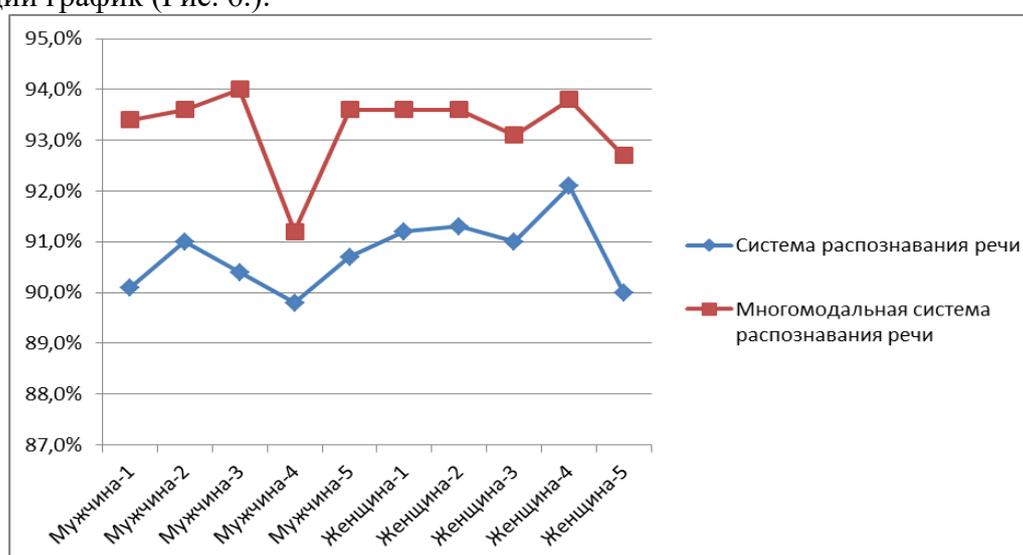


Рис. 6. Результаты сравнительного анализа системы распознавания речи.

Обсуждение. Таким образом, в работе предложен новый метод многомодального распознавания, который отслеживает и читает движения губ, совместно сопоставляя речь для распознавания казахских слов (Рис. 7). Использование изображения визем помогает обучить систему онлайн режиме и для сопоставления фонемы с виземами для распознавания. К базам данных визем доступ 97% со скоростью 15 кадров в секунду. Время тестирования зависит от человека, от произношения и классификации визем или фонем. Результаты анализа и сравнения с другими существующими методами и моделями для распознавания, превзошли качество ХМ моделей и KNN (Nearest Neighbor) классификаторов слов на 21,30% и 11% соответственно. Представленная многомодальная система распознавания позволяет улучшить функциональность адаптивной классификацией визем, добавив к звуковым данным данные о мимике в режиме реального времени. При этом правильность распознавания речи улучшается на 37% в среднем. Разработанная система позволяет улучшить распознавание, совместив звуковые сигналы с данными о положении губ. При этом улучшается распознавания на 39,6% в среднем.

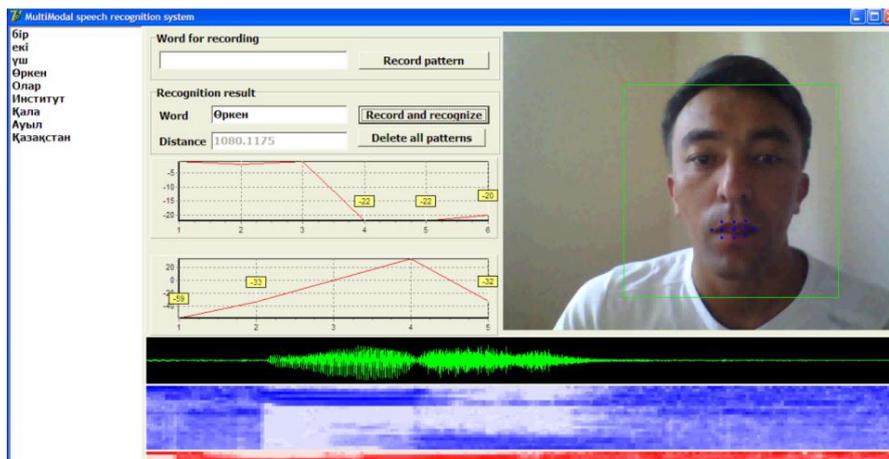


Рис. 7. Интерфейс многомодальной системы распознавания речи

Сравним результаты данных исследований с исследованиями классической «Обобщенной моделью распознавания речи», в которой рассмотрение процесса анализа речевого сигнала показало необходимость введения обратных связей [26]. Очевидно, обобщенная модель распознавания речевого сигнала подобна синтезу речи с той лишь разницей, что движение информации идет в обратном направлении. Для реализации базовых алгоритмов необходимо сформировать прогнозируемое получение смысла и цели принимаемого речевого сигнала.

Таким образом, необходимо отобразить схему получения смысла речевого сообщения. Данная схема может быть использована в качестве методологических основ систем синтеза речи. Приведенная модель описывает получение печатного текста, однако в этой схеме отсутствует собственно получение семантики и прагматики сообщения и диалога в частности. При этом целесообразно рассматривать полную систему диалога при человеко-машинном взаимодействии. Эта модель используют только прямой канал преобразования, показанные непрерывными линиями на рис. 8. Прерывистыми стрелками показаны предполагаемые для наполнения связи для согласования формирования и разбора сообщения. Для повышения эффективности работы системы преобразования информации предложена следующая система синтеза и распознавания речи на рис. 9.

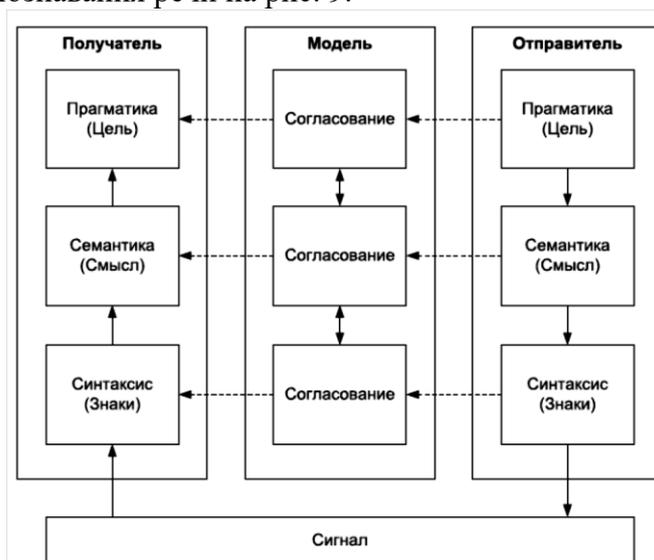


Рис. 8. Упрощенная схема преобразования информации

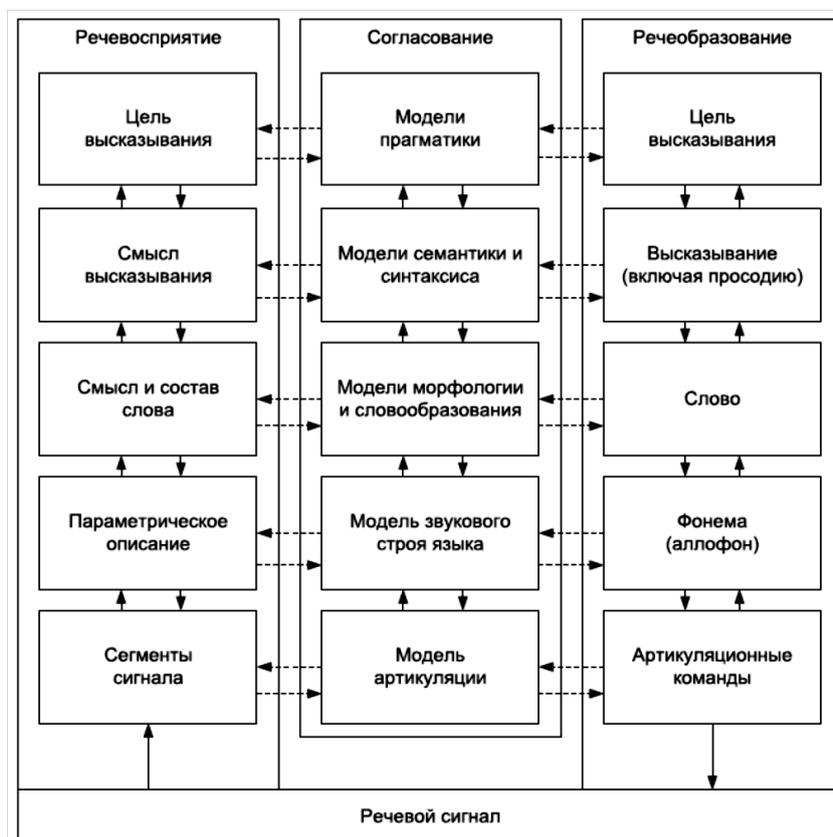


Рис. 9. Обобщенная схема речеобразования и речевосприятия

Таким образом, обобщенная схема учитывает различные априорные знания: языковые (фонетика, лексика, синтаксис, семантика, просодия и т. д.) и неязыковые, т. е. знания предметной области диалога. Внешними входными и выходными данными речевых систем являются: смысловое пространство слов и фраз заданного языка и предметной области, речевой сигнал, а также для систем синтеза – параметры речеобразующей системы.

Но так, как она является одномодальной и не имеет возможности анализа мимики (чтения с губ), а имеет лишь речевой аудиоканал, то она малоэффективна для людей с особыми потребностями.

На данный момент в мире есть 2 ветки развития одномодальных систем распознавания речи (объем рынка с \$1.09 по \$2.42 миллиарда с 2012 по 2016 гг., темп роста +22.07%)

Преобразование речи в текст (объем рынка с \$860млн. (2012г.) до \$1727млн. (2016г.) — общая доля 79%-71% с 2012 по 2016 гг.)

Верификация и идентификация голоса человека (объем рынка с \$229млн. (2012г.) до \$697млн. — общая доля 21%-28,8% с 2012 по 2016 гг.)

В конкурентной борьбе будут более активно развиваться компании, которые существуют на грани эти двух направлений — с одной стороны улучшая точность программ распознавания речи и перевода его в текста, с другой стороны решая эту задачу посредством идентификации диктора и верификации его речи, используя дополнительный канал (например видео) в качестве источника информации.

Согласно исследованию [Technavio](#) — основная проблема существующих программ распознавания речи — это их подверженность в подавлении окружающего шума;

— Основная тенденция — распространение речевых технологий за счет увеличения количества и качества мобильных устройств и развития решений мобильного банкинга;

— Большую погоду в развитии технологий распознавания речи на данный момент играет государственные организации, военная сфера, медицина и финансовый сектор.

Однако наметился большой спрос на такого рода технологии в виде мобильных приложений и задач голосовой навигации, а также биометрии;

— Основной рынок систем распознавания речи находится в США, однако самая быстрая и платежеспособная аудитория проживает в странах юго-восточной Азии, особенно в Японии (за счет полной голосовой автоматизации работы call-центров). Предполагается, что именно в данном регионе должен появиться сильный игрок, который станет серьезным подспорьем для мирового могущества [Nuance Communications](#) (текущая доля общемирового рынка — 70%);

— Наиболее распространенная политика на рынке систем распознавания речи — это слияния и поглощения (M&A) — компании -лидеры рынка часто скупают небольшие технологические лаборатории или фирмы по всему миру, чтобы сохранить гегемонию.

— Стоимость приложений стремительно падает, точность растет, фильтрация посторонних шумов улучшается, безопасность возрастает — предполагаемая дата реализации сверхточной технологии распознавания речи — 2018 гг.

Выводы. Таким образом, по прогнозам [Technavio](#) в период 2014-2018 гг. ожидается увеличение рынка одномодальных систем распознавания речи более чем в 2,5 раза. Большую долю на одном из самых динамичных и быстрых рынков IT технологии получают игроки, которые смогут в своем продукте решить 2 задачи одновременно: научиться качественно распознавать речь и переводить ее в текст, а также хорошо уметь идентифицировать голос диктора, верифицировать его из общего потока. Большим преимуществом в конкурентной борьбе можно назвать демпинг (искусственное снижение стоимости подобных технологий), создание программ с дружелюбным интерфейсом и быстрым процессом адаптации — при высоком качестве работы. Предполагается, что в течение ближайших 5 лет — появятся новые игроки на рынке, которые могут поставить под сомнение менее поворотливых крупных корпораций типа [Nuance Communications](#).

На мировом рынке распознавания голоса единой проблемой является невысокие показатели точности распознавания, не смотря на то, что в настоящее время системы распознавания голоса способны распознавать различные языки и определять подлинность голоса. Так как система включает в себя сложный процесс согласования баз данных с произносимыми командами и интегрированной технологией распознавания речи и голосовой верификации, даже незначительная ошибка в любой часть процесса может привести к неверному результату.

Погрешность в распознавании речи является одним из основных ограничений в приложениях распознавания голоса. Однако некоторые производители начали разработку систем с очень низким уровнем погрешности в распознавании голоса. Они разработали системы с менее чем 4% неточных результатов (например, измерения голосовой биометрии неверно идентифицируют и отвергают голос человека, у которого есть доступ).

При таких направлениях развития рынка одномодальных систем, многомодальные системы (как в данной работе) имеют еще большие перспективы развития. Так, как анализируя не только голос но и мимику, можно значительно увеличить точность распознавания.

Созданную многомодальную систему можно применять в автомобильных навигационных приборах и для удобного использования людьми с ограниченными возможностями. До сих пор методы чтения с губ не были эффективно использованы в режиме реального времени из-за трудности отслеживание мимики. В этой работе предлагается система чтения с губ, что может эффективно применяться в качестве опоры для существующей системы распознавания речи. Предлагается многомодальная система распознавания речи в качестве эффективного инструмента для распознавания казахской речи.

Список использованных источников:

1. Jain A.K., Ross A., Prabhakar S., (2014) An introduction to biometric recognition, IEEE Trans. Circuits Syst. Video Technol. Vol. 14, 4–20.

2. Jain A.K., Ross A., Prabhakar S., An introduction to biometric recognition, *IEEE Trans. Circuits Syst. Video Technol.* 14 (2014) 4–20.
3. Raghavendra R. (2010). *Robust Algorithms for Person Verification using Multimodal Biometrics*, Ph.D Thesis, Feb.
4. Raghavendra R., Ashok Rao, Hemantha Kumar (2013). *Multisensor Biometric Evidence Fusion of Face and Palmprint for Person Authentication using Particle Swarm Optimization (PSO)*, *International Journal of Biometrics*, Vol.2, No.1, 19–33.
5. Sanderson C., & Paliwal K. (2013). *Noise compensation in a person verification system using face and multiple speech features*, *Pattern Recognition* 36 (2), 293-302.
6. Bishop C. M. (2014). *Pattern Recognition and Machine Learning*, Springer-Verlag Edition.
7. Andabayeva K. (2015) *Comparative analysis of consonants Kazakh and English language*. *Bulletin KazNPU*. Vol.4, 133-136.
8. Aralbayeva J.A. (2005). *Vowels Kazakh language*. *The Science*, (pp.178-181).
9. Ben-Yacoub S., Abdeljaoued Y., & Mayoraz. *Fusion of face and speech data for person identity verification*, *IEEE Transactions on Neural Networks* 10.
10. Kalimoldayev M. N., Amirgaliyev Ye., Mussabayev R. R., Mamyrbayev O. J. (2014). *Methods of forming a dictionary vizeem for multimodal speech recognition*. X- *International Conference*, Vol.35, No. 1, 60-65.
11. Rabiner L. R., & Schafer R.V. (2013). *Digital processing of speech signals*. *Radio and Communication* (pp. 495-515).
12. Kalimoldayev Maxat N., Keylan Alimhan & Mamyrbayev Orken J. (2014). *Methods for applying VAD in Kazakh speech recognition systems*. *International Journal of Speech Technology*, 17(2), 199-204.
13. Kalimoldayev M. N., Mamyrbayev O. J., Mussabayev R. R., Orazbekov J. N. (2014). *Segmentation and speech signal processing algorithm using the average frequency level crossing*. *Problems of Informatics*, 1(22), 73-82.
14. Lanitis A., Taylor C., & Cootes T. (2013) *Automatic interpretation and coding of face images using flexible models*. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 19(7), 743-756.
15. Cootes T., Edwards G., & Taylor C. (2015). *Active appearance models*. In *Proceedings of the European Conference on Computer Vision*, volume 2, 484–498.
16. Baker S., Gross R., & Matthews I. (2014). *Lucas-Kanade 20 years on: A unifying framework: Part 1*. Technical Report CMU-RI-TR-02-16, Carnegie Mellon University Robotics Institute.
17. Erzin E., Yemez Y., Tekalp A. (2015). *Multimodal speaker identification using an adaptive classifier cascade based on modality reliability*, *IEEE Trans. Multimedia* 7(5), 840–852.
18. Nilsson M., & Ejarsson M. (2013). *Speech recognition using hidden Markov model*. Department of Telecommunications and Speech Processing. Blekinge Institute of Technology, Blekinge.
19. Мещеряков Р.В., Бондаренко В.П., Организация баз знаний в системе синтеза речи // Теория и практика речевых исследований (АРСО_99). Матер. конф. – М.: МГУ им. М.В. Ломоносова, 2014. – С. 37–38.
20. Златоустова Л.В., Потапова Р.К., Трунин_Донской В.Н. *Общая прикладная фонетика*. – М.: Изд_во МГУ, 1986. – 304 с.
21. *Математическая лингвистика. Сб. переводов / под ред. Ю.А. Шрейдера, И.И. Ревзина, Д.Г. Лахути, В.К. Финна*. – М.: Мир, 1964. – 144 с.
22. Филичева Т.Б., Чевелева Н.А., Чиркина Г.В. *Основы логопедии*. – М.: Просвещение, 1989. – 223 с.
23. Потапова Р.К. *Речь: коммуникация, информация, кибернетика*. – М.: Радио и связь, 2014. – 528 с.
24. Сорокин В.Н. *Синтез речи*. – М.: Наука, 1992. – 392 с.
25. Потапова Р.К. *Речевое управление роботом*. – М.: Радио и связь, 1989. – 246 с.
26. Дмитриев Л.Б., Теляева Л.М., Таптапова С.Л., Ермакова И.И. *Фониатрия и фонопедия*. – М.: Медицина, 1990. – 272 с.
27. Бондаренко В.П., Мещеряков Р.В. *Диалог как основа построения речевых систем // Кибернетика и системный анализ*. – 2013. – № 2. – С. 30–41.
28. Бондаренко В.П., Квасов А.Н., Конев А.А., Мещеряков Р.В., Чойнзонов Е.Л., Чижевская С.Ю. *Программные средства комплекса исследования речевого сигнала при злокачественных заболеваниях гортани // Медицинская техника*. – 2014. – № 4. – С. 33–37.

Ү. КАМАНУР, А.Ә. ШӘРІПБАЙ, Г.Т. БЕКМАНОВА, Л. ЖЕТКЕНБАЙ, Ө.ТОҚТАР

ҚАЗАҚ ҚЫТАЙ ТІЛДЕРІ БОЙЫНША МАШИНАЛЫҚ АУДАРМАНЫҢ ҚАДАМДАРЫ

(Л.Н.Гумилев атындағы Еуразия ұлттық университеті)

Қазақ тілі роман-герман, славян, семит, қытай тілдері тобынан айтарлықтай ерекшеленетін түркі тілдер тобына жатады. Жоғарыда аталған тілдер топтарына машиналық аударудың сапалы жүйелері бар. Дегенмен қытай және түрік тілдері топтарының арасында машиналық аудармашы жоқ. Жұмыс барысында ең алғаш рет қытай және қазақ тілдері арасында машиналық аудару технологиясы жасалады. Ал қазақ тілінен басқа туыс түркі тілдеріне (түрікше, әзірбайжан, өзбек, түркімен, ұйғыр, татар және т.б. шамамен 50 тіл) оңай аударуға болады. Жоғарыда аталған жайттар қазіргі заманда бар жүйелерден ерекшелігін көрсетеді.

1 машиналық аудармаға шолу

Қазақ және қытай тілдерінің машиналық аудармашысын құру қазақ тілі үшін машиналық аудару жүйелері, семантикаға негізделген ақпараттық іздеу жүйелері, сұрақ-жауап жүйелерін және білімдерді алу жүйелерінің дамуына жаңа серпіліс береді. Мұндай семантикалық аудармашының негізі ретінде бір біріне метатіл арқылы сәйкестендірілген қазақ тілінің формалды семантикалық моделі және қытай тілінің формалды семантикалық моделі болғандықтан, бұл аударманың жоғары деңгейдегі дәлдігін қамтамасыз етеді. Бұл жағдай компьютерлік лингвистика саласында әлемдік ғылымның дамуына әсер етеді.

Қоғамда қазақ тілінен қытай тіліне және керісінше машиналық аудару технологиясына, деңгейі ағылшын, француз, испан және басқа да дамыған елдердің тілдерінің машиналық аударуымен бірдей әлеуметтік сұраныс бар.

Қазақша-қытайша және қытайша-қазақша машиналық аударма екі ел арасындағы экономикалық, білім саласындағы, әлеуметтік байланыстардың күшеюіне алып келеді, ел және мәдениет арасындағы тілдік тосқауыл азаяды.

Қазіргі таңда бүкіл дүние жүзінде дамыған елдердің тілдерін машиналық аудару технологиялары қарқынды дамуда. Мысалы, мәзірінде 80-не жуық тілдер қолжетімді Google Translate сол сияқты, бірақ мүмкіншіліктері азырақ Яндекс.Перевод, Translate.ru, Bing Translator, Apertium, Babel Fish, Dicto, PROMT, GoldenDict, StarDict, WikiBhasha, SYSTRANet, Free Translation, Babel Fish, Worldlingo, InterTran, ImTransator, Translate Online.ua, Windows Life Translator [1-3] сияқты машиналық аударудың тиімді технологиялары бар. Бұл барлық танымал дүниежүзілік машиналық аудару саласындағы лидерлер сапалы түрде қазақ тілін қолдамайды. Осы кезде Қазақстан әлемдік қауымдастыққа белсенді түрде кірістіріліп жатыр және біздің елімізге, мәдениетімізге және тілімізге қызығушылық қарқынды өсуде.

Қазіргі таңда машиналық аудару жүйелерінің көптеген түрлері бар. Машиналық аудару жүйелерінің қолдануы табиғи тілді формалдау күрделілігіне және табиғи тілдің тілдік корпусының бар не жоқ болуына байланысты. Грамматикалық ережелерге негізделген жүйелер (Rule-Based Machine Translation, RBMT) формалдауға жақсы келетін тілдерге қолдануға қолайлы. Осындай жүйелердің венгер-ағылшын машиналық аударуға арналғаны бар [6]. Көптеген ғалымдар машиналық аударудың статистикалық әдістерін [7-8] зерттейді және қолданады, неміс, испан, француз және ағылшын машиналық аудару жұмысында [9] фразалар-аналогтарға негізделген N-gram-based models co-exist with their phrase-based counterparts as an alternative SMT framework N-граммды модел жүйесі қолданылды. Жұмыстардың біразында ережелер мен статистикалық әдістерге негізделген әдістерді қолданатын гибриді әдістер қолданылады [10]. Сонымен қатар мәтіннің семантикасына тәуелділікке негізделген әдістер бар.

Мысалы, [11] жұмысында мәтін құрылымын құруды семантикалық гиперграф қолданылады, ал [12] жұмысында тіл семантикалық түрде тілдің сембанкте (семантикалық банкте), яғни семантикалық құрылым болып табылатын ағылшын тіліндегі мыңдаған сөйлемдер сақталады.

Интернетте аударманың статистикалық әдісіне негізделген қазақ және орыс тілдері арасындағы машиналық аудару жүйелері [13-16] бар. Өкінішке орай, бұл жұмыстар бойынша ғылыми басылымдар жоқ, сондықтан олардың ғылыми құрамын және қолданылған технологияларды бағалау мүмкін емес.

Қоғамда қазақ тілінен қытай тіліне және керісінше машиналық аудару технологиясына, деңгейі ағылшын, француз, испан және басқа да дамыған елдердің тілдерінің машиналық аударуымен бірдей әлеуметтік сұраныс бар.

2 Қазақ және қытай тілдерінің машиналық аударманың есептері

I. Таңдалған пәндік салалар бойынша деректер базасын және білімдер базасын құру:

I.1. Таңдалған пәндік салалардың қазақша және қытайша терминдердің деректер базасын құру.

Нәтижесі: Таңдалған пәндік салалардың қазақша және қытайша терминдердің деректер базасы.

I.2. Пәндік салалар бойынша тезаурус және онтологиялық моделдер құру.

Нәтижесі: Пәндік салалардың тезаурусы және онтологиялық моделдері;

I.3. Таңдалған пәндік салалар бойынша сөз тіркестер мен фразалардан тұратын деректер базасын құру.

Нәтижесі: Таңдалған пәндік салалар бойынша сөз тіркестер мен фразалардан тұратын деректер базасы.

I.4. Таңдалған пәндік салалар бойынша білімдер базасын құру.

Нәтижесі: Таңдалған пәндік салалар бойынша білімдер базасы.

II. Қазақ және қытай тілдерінің лингвистикалық моделдерін құру:

II.1. Қазақ және қытай тілдерінің графематикалық моделдерін құру.

Нәтижесі: Параграфтар, абзацтар, сөйлемдер, қазақ және қытай тілдеріндегі сөздер мен мәтін құрылымының басқа да элементтері.

II.2. Қазақ және қытай тілдерінің морфологиялық ережелерінің семантикалық моделдерін құру.

Нәтижесі: Формалды метатілде қазақ және қытай тілдерінің морфологиялық ережелерінің семантикалық моделдері;

II.3. Қазақ және қытай тілдерінің синтаксистік ережелерінің семантикалық моделдерін құру.

Нәтижесі: Қазақ және қытай тілдерінің синтаксистік ережелерінің семантикалық моделдері;

II.4. Қазақ –қытай трансляторының бірегей семантикалық моделдерін құру.

Нәтижесі: Қазақ-қытай трансляторының бірегей семантикалық моделдері.

III. Қазақ тілінен қытай тіліне және керісінше машиналық аудару алгоритмдерін және программаларын жасау:

III.1. Қазақ тіліндегі сөздер мен сөз тіркестерін қытай тіліне және керісінше машиналық аудару алгоритмдері мен программаларын құру.

Нәтижесі: Қазақ тіліндегі сөздер мен сөз тіркестерін қытай тіліне және керісінше машиналық аудару алгоритмдері мен программалары;

III.2. Қазақ тіліндегі сөйлемдерді қытай тіліне машиналық аудару алгоритмдері мен программаларын құру.

Нәтижесі: Қазақ тіліндегі сөйлемдерді қытай тіліне машиналық аудару алгоритмдері мен программалары;

III.3. Қытай тіліндегі сөйлемдерді қазақ тіліне машиналық аудару алгоритмдері мен программаларын құру.

Нәтижесі: Қытай тіліндегі сөйлемдерді қазақ тіліне машиналық аудару алгоритмдері мен программалары;

III.4. Машиналық аудару сапасын бағалау әдістерін құру және оны әзірленген моделдер, алгоритмдер мен программаларды бағалау үшін қолдану және пайдаланушы интерфейсін жасау.

Нәтижесі: Машиналық аудару сапасын бағалау әдістері және моделдер, алгоритмдер мен программаларды бағалау әдістері, пайдаланушы интерфейсі.

3 Қазақ және қытай тілдерінің машиналық аударманың ғылыми жаңашылдығы және маңыздылығы

1990-2000 жылдар арасында қазақ тілінен машиналық аударға байланысты ғылыми жұмыстар жоқтың қасы. 2000 жылдың басынан тек Қазақстандық ғылыми-зерттеу институттары ғана емес, Jonathan North Washington (2006) "A Novel Approach to Delineating Kazakh's Five Present Tenses:Lexical Aspect" [17]; A.Dzhubanov, B.Khasauov (2007) "Computational description of the kazakh language"; B.J. Bayachorova, P.Pankov (2009) "Independent Computer presentation of a natural language": G. Altenbek and Wang Xiao-long (2010) "Kazakh Segmentation System of Inflectional Affixes" [18] сияқты бірнеше ғылыми еңбектер қазақ тілінің автоматтандырудың дамуын көрсетеді. Барлық аталған жұмыстар лексикалық және синтаксистік деңгейде қазақ тілі сегменттеу және автоматтандыру саласында қызықты. Бірақ олар табиғи тілді өңдеудің кейбір аспектілерін ғана қарастырады. Машиналық аудару қолданбалы лингвистикадағы қиын әрі көпфункционалды жүйе болып табылады. Сонымен қатар, Шарипбаев А.А., Бекманова Г.Т., Ергеш Б.Ж., Бурибаева А. К., Карабалаева М. Х. (2012) " Семантикалық желіге негізделген интеллектуалды морфологиялық талдағыш "[19] атты ғылыми жұмысты атап өтуге болады. Авторлар семантикалық желі арқылы қазақ тілінің морфологиялық талдағышын кеңейткендігін көрсетеді. Жұмыста қазақ тілінің морфологиялық ережелерін және морфологиялық алгоритмін формалдау берілген. Сонымен қоса, сөз түберлерінің семантикалық қорын құру көрсетілген.

Сонымен қатар соңғы жылдары қазақ тілі негізінде табиғи тілдегі мәтінді автоматты өңдеуді іске асыру тәжірибесімен тығыз байланыстағы зерттеулер жүргізілуде. Атап айтқанда, бүгінгі күні, қазақ тіліндегі мәтіндердің орфографиялық ерекшеліктері қарастырылды және қатені автоматты анықтаудың негізгі кезеңдері белгіленді. Бұл жағдайда, қазақ мәтіндерін автоматты өңдеу процессіне мәтін дұрыс сөздермен қамтамасыз ету, сөз формаларын дұрыс талдауын бақылау және т.б. сияқты деңгейлері енгізілген. Осы жұмыстың логикалық қорытындысы ретінде қазақ тіліндегі мәтінді дұрыстығына тексеретін жүйе (Spell Check) және өндірісте қолданылуға берілген қазақ тілінің электронды сөздігі жасалды. Сонымен қатар ҚР БҒМ-нің ақпараттық және есептеу технологиясы институтында және Л.Н.Гумилев атындағы ЕҰУ, «Жаснды интеллект» ҒЗИ-да қазақ тіліндегі сөйлеуді тану және синтездеу аумағында белсенді жұмыстар жүргізілуде[20-23].

Табиғи тілдегі мәтіндерді аударудағы негізгі мәселе ретінде келесілерді санауға болады: сөздің көпмағыналылығы және синтаксистік құрылымы, тіпті пәндік аумақтың шектеуінде де ғаламдық семантикалық құрылымды сипаттаудың тәжірибелік қиындығы, лингвистикалық заңдылықтарды сипаттайтын тиімді формальды әдістердің болмауы және т.б. [24]. Арнайы пәндік облысқа арналған терминдер сөздігін қолданғанда, сонымен қоса өңделетін құжаттың типі қайталанған кезде ғана МА тиімді болады.

Алынған нәтижелерді мақсаттық тұтынушылары жынысына, ұлтына, жасына, азаматтығына, тұрғылықты жеріне, социалдық статусына, жұмыс орнына және т.б. тәуелсіз қазақ және/немесе қытай тілдерін білетін және интернетте жұмыс істеу дағдысы бар физикалық тұлға, сонымен қатар меншік түріне, ұйымдастыру формасына, қызмет пәніне, юрисдикцияның орналасқан орнына тәуелсіз заңды тұлғалар, сонымен қатар халықаралық ұйымдар және дипломатиялық корпус болады.

Жұмысны жүзеге асыру нәтижелерінің ұлттық және халықаралық ауқым бойынша ғылыми жаңалығы мыналарда: біріншіден, қазақ және қытай тілдерінде сөздер мен

сөйлемдер мағыналық қасиеттерін ескере отырып, морфологиялық және синтаксистік ережелер математикалық модельдерін жасау; екіншіден формалды ережелерге негізделген қазақ тілінен қытайшаға машиналық аудару алгоритмдері мен программаларын жасау; үшіншіден, құрылатын қазақ және қытай тілдерінің параллель корпусында статистикалық әдістерін қолдану; төртіншіден, қазақ-қытай трансляторының семантикалық моделінің формалды ережелері мен статистикалық әдістерге негізделген әдістерді тиімді үйлестіру технологиясын құру болып табылады.

Жұмыстың практикалық маңыздылығы: біріншіден, әлемнің ең дамыған 30 елінің қатарына кіру жолында Қазақстанның мемлекеттік тілінің рөлін және мәртебесін арттыру болып табылады. Екіншіден, қазақ және қытай тілдерін оқытудың мүмкіндіктерін арттыру. Үшіншіден, тиімді қазақ-қытай машиналық аударманың болуы Қазақстан мен Қытай арасындағы барлық салалардағы халықаралық қарым-қатынастарды дамытуға және «Жібек жолы» жобасын тиімді жүзеге асыруға ықпал етеді.

Пайдаланған әдебиеттер тізімі

- 1 <https://ru.wikipedia.org/wiki/>.
- 2 glosbe.com
- 3 <http://www.primavista.ru>
- 4 Rakhimova D.R. “Research models and semantics methods of machine translation from Russian into Kazakh language” PhD thesis, al-Farabi Kazakhstan National University, Almaty, Kazakhstan, 2014. – pp.8–9.
- 5 Tukeev U., Rakhimova D.R. Augmented attribute grammar in meaning of natural language sentences The 6th International Conference on Soft Computing and Intelligent Systems / The 13th International Symposium on Advanced Intelligent Systems. – Kobe, Japan, 2012, P.1080-1084.
- 6 Attila Novák, László Tihanyi, Gábor Prószéky. The MetaMorpho translation system. Proceedings of the Third Workshop on Statistical Machine Translation at ACL 2008. Columbus, Ohio. 2008. – pp. 111–114,
- 7 Final Report of the 2006 Language Engineering Workshop Open Source Toolkit for Statistical Machine Translation: Factored Translation Models and Confusion Network Decoding
- 8 Barry Haddow. Applying Pairwise Ranked Optimisation to Improve the Interpolation of Translation Models. Proceedings of NAACL-HLT, Atlanta, Georgia, 2013. – pp. 342–347.
- 9 Nadir Durrani, Alexander Fraser, Helmut Schmid. Model With Minimal Translation Units, But Decode With Phrases. Proceedings of NAACL-HLT, Atlanta, Georgia, 2013. – pp.1–11.
- 10 Andreas Eisele and etc. Using Moses to Integrate Multiple Rule-Based Machine Translation Engines into a Hybrid System. Proceedings of the Third Workshop on Statistical Machine Translation, Columbus, Ohio, USA, 2008. – pp.179–182.
- 11 Kenneth Heafield, Philipp Koehn, Alon Lavie. Grouping Language Model Boundary Words to Speed K-Best Extraction from Hypergraphs. Proceedings of NAACL-HLT 2013, Atlanta, Georgia, 2013. – pp. 958–968.
- 12 Laura Banarescu and etc. Abstract Meaning Representation for Sembanking// Proceedings of the 7th Linguistic Annotation Workshop & Interoperability with Discourse, Sofia, Bulgaria, 2013 – pp.178–186.
- 13 <http://sozdik.kz>
- 14 <http://www.soylem.kz>
- 15 <https://translate.google.kz>
- 16 <http://translate.meta.ua/ru>
- 17 Washington J. N., Salimzyanov L, Tyers F.M. Finite-state morphological transducers for three Kypchak languages // Proceedings of the 9th Conference on Language Resources and Evaluation, LREC. - Granada.-2014.- P 123.
- 18 Altenbek G., Wang X . Kazakh Segmentation System of Inflectional Affixes // Proceedings of CIPS -SIGHAN Joint Conference on Chinese Language Processing.- China: Beijing, 2010.- P. 183—190.
- 19 Шарипбаев А.А., Бекманова Г.Т., Ергеш Б.Ж., Бурибаева А. К., Карабалаева М. Х. Интеллектуальный морфологический анализатор, основанный на семантических сетях // Материалы 2-й международной научно-технической конференции «Открытые семантические технологии проектирования интеллектуальных систем», - Минск: 2012.- С. 397-400.
- 20 Шарипбаев А. А., Бекманова Г. Т. Построение логической семантики слов казахского языка // Материалы Всероссийской конференции с международным участием «Знания - Онтологии -Теории (ЗОНТ-09)». - Новосибирск: 2009. - Т. 2. -С. 246-249.
- 21 Амиргалиев Б.Е., Амиргалиев Е.Н. Методы анализа речевого сигнала в системах синтез речи и распознавания // Материалы международной конференции «Современные проблемы математики, информатики и управления», посвященной 60-летию академика МАИ, д.ф.-м.н., профессора М.Б. Айдарханова. - Алматы, 2008. - С. 50-53.
- 22 Бурибаева А.К. Распознавание казахских слов на основе дифонной базы // Труды 1-й международной конференции «Компьютерная обработка тюркских языков».— Астана: ЕНУ им. ЛИ. Гумилева, 2013. - С. 230-238.
- 23 Амиргалиев Б.Е., Мусабаев Р.Р. и др. Унифицированный язык фонетического представления для систем и распознавания речевого сигнала // Материалы международной научно-практической конференции «Актуальные проблемы математики, информатики, механики и теории управления». - Алматы: ИПИУ-КБТУ, 2009. - Ч. 1. - С. 55-57.

Б.А.КҮЗЕНБАЕВ, А.А. ШАРИПБАЙ, Р.С. НИЯЗОВА, А.Б. БАРЛЫБАЕВ

**ОНТОЛОГИЯЛАРДЫ САЛЫСТЫРА ОТЫРЫП АВТОМАТТАНДЫРЫЛҒАН
БАСҚАРУ ЖҮЙЕСІНІҢ САПАСЫН АНЫҚТАУ**

(Л. Гумилев атындағы Еуразия ұлттық университеті, Астана қаласы)

ЖОО –ның бөлімшелеріндегі сапа менеджмент жүйесі негізгі екі бөлімнен тұрады:

- бөлімше туралы құжатнама
- қызметкерлердің қызметтік нұсқаулары

Бөлімше туралы құжатнамада бөлімшелер арасындағы қарым-қатынас, ал қызметкерлердің қызметтік нұсқауларында қызметкерлер арасындағы қарым – қатынастары көрсетілген. Осыған байланысты office, docs, function, user объектітерін арнайы фигуралармен белгілеу қажет. Негізгі онтографтарда фигуралық шыңдар жасауға мүмкіндіктер қарастырылмаған. Олар шыңдарды қарапайым нүкте ретінде қарастырады.

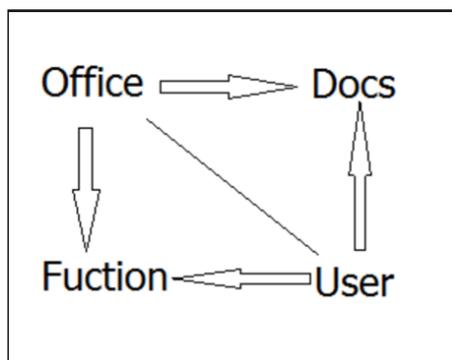
Осы арнайы фигуралар деректер қорын немесе бағдарламалық қамтаманы жобалауда қолданылатын UML және mockups методологиясына ұқсайды, мысалы use case diagrams немесе business models.

Осыны негізге ала отырып жаңадан онтограф жасау қажеттілігі туындады.

Нәтижесінде бір форматтағы екі онтологияны аламыз, біріншісі қолданушының қызметтік құжаттарына негізделген, екіншісі осы қолданушы жұмыс атқаратын автоматтандырылған басқару жүйесінің бизне үрдістеріне негізделеді.

ЖОО оқу үрдісін автоматтандырылған басқарудың сапасын бағалайтын эксперттік жүйені құрастыру мақсатында онтология конструкторы құрастырылды. Құрастырылған онтология екі режимде орындалады. Бірінші режим - сапа менеджмент жүйесінің expertтінің автоматтандырылған жұмыс орнын қарастырады. Екінші режим - басқарудың автоматтандырылған жүйесінің expertтінің автоматтандырылған жұмыс орнын қарастырады.

Басқарудың автоматтандырылған жүйесінің expertтінің автоматтандырылған жұмыс орны екі бөлімді қамтиды: онтологияны құрастыру және құрастырылған онтологияны бағалау. Онтологияны құрастыру үшін төрт объект қарастырылады: Office, User, Docs, Function.



Сурет 1. Конструктор онтологиясының құрылымы

Конструктор бізге онтологияны және оның тезаурусын жасауға және құрастырған онтологияларды салыстыра отырып бағалауға мүмкіндік береді. Объекттердің арасында қарым-қатынасты орнату мақсатында Қабырға элементті қолданылады.

Кесте 2 – Бағаны шығару ережесі.

Variable	Rule	Conclusion
Unsatisfactorily	If $0*n \leq l \leq 0,49*n$	Then “ACS deviates fully from the quality the QMS requirements”
Satisfactorily	If $0,5*n \leq l \leq 0,74*n$	Then “ACS deviates significantly from the quality the QMS requirements”
Good	If $0,75*n \leq l \leq 0,89*n$	Then “ACS deviates slightly from the quality the QMS requirements”
Excellent	If $0,9*n \leq l \leq n$	Then “ACS satisfy fully the quality the QMS requirements”

Екі онтологияны салыстыра отырып төмендегі нәтижелерге ие боламыз: қанша пайызға олар ұқсас, дұрыс байланыстарды анықтау, дұрыс емес байланыстар және қандай элементтер қосу қажет.

Көрсетілген параметрлер бойынша екі онтология салытырылып, автоматтандырылған басқару жүйесінің сапа менеджмент жүйесінің талаптарынан аса ірі көлемде ауытқитыны анықталды. Екі онтология бір-біріне 61,82%-ға ұқсас екені айқындалды. Сондай-ақ, ауытқуды жою үшін дұрыс емес байланыстар, артық элементтер, қандай элементтерді қосу керек екені төмендегі суретте анық белгіленген.

Мәні	Түрі	Мәні	Түрі	Мәні	Түрі	Мәні	Түрі
Кафедра	OFFICE	Мұғалым	USER	Оқу аудиториясы	FUNCTION	Тестілеу	FUNCTION
Журнал	DOCS						
Толтыралы	1) Мәні	Тіркеу офісі	2) Мәні	Басқаралы	Тіркеу офісі		
Түрі	USER	Түрі	USER				
Академиялық календарь	DOCS	Бұйрықтар	DOCS				
Факультет	OFFICE						
Оқу процесін жоспарлау жән...	USER	Оқушылардың қозғалысын ...	USER	ОЖТ	DOCS	Декан	USER
				Өкілдік	DOCS		

Сурет 5. Екі онтологияны салыстырудың нәтижесінің формасы

Осы онтологиялық моделдерді салыстыра отырып, функционалдық талаптарына автоматтандырылған басқару жүйесінің қаншалықты сәйкес келетінін анықтауға мүмкіндік аламыз. Мұнда онтологияларды салыстыру әдісі қолданылады. Өйткені, жасалынған онтографтың көмегімен ЖОО оқу үрдісінің жұмысын басқаратын нормативті құжаттарға және автоматтандырылған басқару жүйесін құрастырудың негізгі идеяларына сәйкес фигуралық шындар арқылы белгілеуге мүмкіндік аламыз.

Құрастырылған онтографты басқа салаларда қолданылатын кез келген автоматтандырылған басқару жүйесін бағалау үшін қолдануға болады.

Әдебиеттер

1 ГОСТ Р ИСО 9001-2008. Системы менеджмента качества. Рекомендации по улучшению деятельности.–М.:Изд-во стандартов, 2008

2 Абакумова, О.Г. Управление качеством: Конспект лекций / О.Г. Абакумова. - М.: А-Приор, 2011. - 128 с.

3 Шемякина, Т.Ю. Производственный менеджмент: управление качеством: Учебное пособие / Т.Ю. Шемякина, М.Ю. Селивохин. - М.: Альфа-М, НИЦ ИНФРА-М, 2013. - 272 с.

FACEBOOK ӘЛЕУМЕТТІК ЖЕЛІСІНДЕГІ ЖАЗБАЛАРДЫ СТИЛЬДІК ЕРЕКШЕЛІКТЕРІН ЕСКЕРІП СЕНТИМЕНТ ТАЛДАУ

(Л.Н.Гумилев атындағы Еуразия ұлттық университеті, Ақпараттық технологиялар факультеті, Астана, Қазақстан)

Сентимент талдау – бұл Text Mining-тың бір бөлігі, мәтіннен субъективті пікірді шығарып алу жүйесі, мәтіннің мазмұны мен реңкін зерттейтін есептік лингвистика және ақпаратты іздеу пәні [1].

Сентимент талдау бірнеше синонимдері мен жақын терминдері бар: сентимент талдау (sentiment analysis), сентиментометрия (sentiment metrics), брендомониторинг (brand monitoring), әлеуметтік медиа талдау (social media analysis), пікірді шолу (opinion mining), пікірді «сыртынан тыңдау» (opinion listening), пікірдің эмоционалды реңкін анықтау және т.с.с.

Сентимент талдаудың танымал әрі жиі қолданылатын сөздіктер негізіндегі әдістері, эмоционалды реңкі бар сөздер және эмоцияны білдіретін символдар сөздігін қолданады. Сөздіктер негізіндегі әдістерде әр сөз эмоционалды реңкін анықтайтын салмаққа ие. Әдетте сөздіктер WordNet сияқты құрылғыларды қолдану арқылы құралады, одан кейін терминдерге салмақтары беріледі [8].

Сөздіктерге негізделген зерттеулерде мәтіннің полярлылығы әртүрлі тәсілдермен анықталады. Үш тәуелсіз бағалау бар: жағымды, жағымсыз және нейтралды [2]. Құжаттың полярлылығы сөйлемдердің полярлылығы анықталғаннан кейін құрылады. Сөйлемдердің полярлылығын анықтау терминдердің салмағының жалпы қосындысын есептеумен анықталады. Сөйлемде әр терминнің жағымды, жағымсыз немесе субъективті салмағы 0 және 1 арасындағы шкалада беріледі.

2016 жылдың ақпан айына дейінгі көрсеткіш бойынша, SocialBakers атты facebook әлеуметтік желісіне сараптама жасайтын ұйымның хабарлауынша (<http://www.socialbakers.com/facebook-statistics/kazakhstan>) Қазақстан facebook әлеуметтік желісін қолданатын елдердің ішінде 100 орынға көтерілді. Соңғы мәліметтер бойынша Қазақстанда 125,8 мың адам аталған әлеуметтік желінің қызметін қолданады. 2016 жылдың 6 айында 51 480 қазақстандық facebook әлеуметтік желісіне қосылған. Әлеуметтік желіні Қазақстанда еркектерге (47%) қарағанда әйелдер (53%) көбірек қолданады.

Қазіргі кезде Facebook-ке тіркелгенде басқа әлеуметтік желілерден достарыңды шақыру қызметі қосылған. Осы қызмет Facebook-те қазақстандық интернет қолданушылар көбеюіне септесіп отыр. Қазақстанда 18-24 жас аралығындағы азаматтар facebook желісін қолданатындардың ең үлкен тобын құрайды. Қазақстан Орталық Азия елдері ішінде бірінші орынды алып отыр.

Facebook әлеуметтік желісінде сентимент әдістерді қолданып, ағылшынын тіліндегі бірнеше қолданушының көңіл-күйін анықтау жобалары бар. Мысалы, SAS Sentiment Analysis, Lithium, OpinionEQ. Бірақ, қазақ тіліндегі жазбаларды талдайтын программалар жоқтың қасы.

Facebook әлеуметтік желісінің қолданушыларының жазбаларында жиі кездесетін стильдік ерекшеліктері:

- Қазақша әріптерді орысша әріптермен алмастырып жазу;
- Эмоционалды реңкі бар сленгтерді қолдану;
- Эмотикондар;
- Латын қарпімен жазу.

Жазбалардың жоғарыда аталған стильдік ерекшеліктерінің болуына себепші факторларды қарастыра отыра сентимент талдау жүргізуге ережелер құруды қарастырсақ.

Facebook әлеуметтік желі қолданушыларының барлығы бірдей сауатты жаза бермейді, диалект сөздерді қолданады, ұялы смартфонда қазақ қарпінің жоқ болуына байланысты оларды орыс әріптерімен алмастырып жазып жатады. Сол себепті, талдау жүргізудің тиімділігін қазақ тілінің фонетикалық ережелеріне сүйене отыра арттыруға болады. Қазақ тілінде 9 дауысты дыбыс бар: а, ә, о, ө, е, ұ, ү, ы, і. Олардың ішінде 5 әріп: ә, і, ө, ұ, ү ұялы смартфондардың пернетақтасында кездеспейді. 25 дауыссыз дыбыс бар: б, г, ғ, д, ж, з, й, к, л, м, н, ң, қ, п, р, с, т, у, ш, в, ф, х, ц, ч, һ. Ұялы смартфондардың пернетақтасында ғ, қ, ң әріптері ғана кездеспейді. Әлеуметтік желі қолданушылары көбіне 1- кестегідей алмастырулар жасайды:

1-кесте Қазақша әріптерді орысша әріптермен алмастыру

Дыбыс	Алмастыратын дыбыс
ә	а
і	ы, и
ө	о
ұ	у
ү	у
ғ	г
қ	к
ң	н
и	й

Бұл алмастыруларды да талдау үшін келесідей алгоритм құрамыз. Талданатын сөйлемде егер деректер қорында жоқ сөздер кездесе оларды қате сөздер кестесіне енгіземіз. Қате сөздер кестесіндегі сөздер түзетуден кейін қайтадан талдаудан өтеді. Түзетулерді іске асыруда жоғарыдағыдай алмастыру кездесе, оларды фонетикалық ережелерге сүйене отыра түзетеміз. Оны көру үшін төмендегі жарияланымды қарастырсақ:

-Даригер,мумкин укол салмай-ак коярсыз,а?! Мен осем гой али, сизди тауып алам дегендей...))))

Даригер сөзіні түзетуді енгізу: А- жуан, ашық дауыстыны ә - жіңішке, ашық дауыстымен, и – жіңішке, қысаң дауыстыны і – жіңішке, қысаң дауыстымен алмастыру керек, сонда дәрігер деген дұрыс сөзді аламыз. Оны деректер қорынан тексеріп, эмоционалды реңкін анықтай аламыз.

Халықтың ауыз екі тілі - әр өлкенің жергілікті қазағының тіліне тән диалектілік белгілер мен құбылыстардың жиынтығын, басқаша айтқанда, халық диалектілері мен говорларын білдіретін, ауызекі тілді білдіретін, бірақ ауызекі тіл болғанда да, оның тек жергілікті формасын ғана білдіретін ұғым деп түсінген дұрыс [5]. Қазіргі қазақ тілінің дыбыс жүйесін, әдеби айту нормаларын дұрыс пайымдау үшін[5] еңбектің маңызы зор. Халық ауыз екі тілінде қолданатын сөздерді зерттей отыра а мен ә дауыстыларын бірінің орнына бірін қолдану, мысалы, кәте –хато - қата – қате, әсілі – асил - асыл, кәтер – хатар – хатар –кәтер – қатер – қатар және т.б., е,ы,і әріптері сөздің барлық буындарында жазылады, бірақ олар ылғи да е,ы,і деп дыбыстала бермейтіндігін, олардың айтылу кезінде ө,ұ,ү әріптерімен алмастырылады[6]. Бұл есептің шешуі, оларды жаңа сөз ретінде деректер қорына қосып, сипаттамаларын беру.

Сонымен қатар, қолданушылардың сленг сөздерді қолданатындығында ескеру керек. Ол халық арасында жиі қолданылатын басқа тілден кірген сөздер. Оларда қосымша деректер қорына енгізіліп олардың эмоционалдық реңкі беріледі. Мысалы, Жарайсың! Сөзін көп қолданушылар Молодец! –деп орысша дұрыс жазса, кейбірі бұрмалап, Мәледес!, Мәледессің! Деп те жазуы мүмкін. Осы тәріздес сленгтер тізбегі де құрылды. Мысалы, *ОК, Норм, Кайф, Реальный, Класс, Релакс, Лайк, Лайф, Селфи, Комментар.*

Елбасы Н.Ә.Назарбаевтың «Болашаққа бағдар: рухани жаңғыру» мақаласының маңызды бағытының бірі қазақ жазуын латын әліпбиіне көшу, жыл соңына дейін оның нақты жобасы жасалуы тиіс деген тапсырамасы берілгеннен кейін әлеуметтік желілерді қоғам белсенділері латын қаріптерімен жаза бастады, халық пікірін білу кезінде барлық қазақша пікірлерді талдау қажеттігі туындады. Бұл есепті шешу үшін Л.Н.Гумилев атындағы ЕҰУ, «Жасанды интеллект» ғылыми-зерттеу орталығы директоры, т.ғ.д., профессор А.Ә. Шәріпбайдың еңбектері негізінде латын әріптерін танып, оларды талдауда қолдану алгоритмі құрылды.

Aa, Bb, Cc, Dd, Ee, Ff, Gg, Hh, Ii, Jj, Kk, Ll, Mm, Nn, Oo, Pp, Qq, Rr, Ss, Tt, Uu, Vv, Ww, Xx, Yy, Zz. Бұл 26 әріптік латын әліпбиі Халықаралық стандарттау ұйымында (International Organization for Standardization) «базалық латын әліпбиі» болып тіркелді және ол замануи ағылшын әліпбиімен бірдей. Осы күнде латын әліпбиі негізіндегі жазуды роман, герман, кельт және балтика топтарының барлық тілдері, сонымен қатар, кейбір славян, фин-угор, түркі, семит және иран тобындағы тілдер, албан, баскі, Индоқытай, Мьянма, Филиппин, Африка, Америка, Австралия тілдері және басқалары пайдаланады. Кейбір тілдерде дыбыстардың саны латын әріптерінің санынан көп, сондықтан базалық латын әліпбиі осы тілдерге бейімделеді [6].

Сентимент талдау жүргізу программасында латын әріпімен жазылған қазақша мәтінді де талдау жүргізіп нәтиже алу үшін келесі алгоритм құрылды: Латынша жазылған сөздерді [6] берілген әріптер негізінде автоматты анықтап, деректер қорына жеке латын әріптерімен арналған сөздерді енгізетін кестеге қосылады. Ол жерден оның сипаттамаларын беру қажет, сөз табын, эмоционалды реңкін (баға).

Қазақша мәтінге сентимент талдау жасайтын зерттеу программа жүйесін әлеуметтік желіге интеграция жасау арқылы қолданушылардың (қазақстан азаматтары) көңіл күйлерін анықтап, алдағы террористік актілер, көтерілістерді алдын ала болжауға септігін тигізуі мүмкін. Қазақ тіліндегі мәтінге сентимент-талдау жасаудың тиімді әдісі сөздіктер мен ережелер негізіндегі талдау нәтижесінде құрылған программа арқылы әлеуметтік көзқарасты анықтауда айтарлықтай. Сонымен қатар елімізде қазақ тілінің қолданылу аясын кеңейтеді.

Болашақ жұмыстар ретінде келесі бағыттарды атап өтуге болады:

- Қате сөздерді автоматты түрде анықтап, оның дұрыс нұсқасын ұсыну алгоритмін дайындау;

- Қазақ тіліндегі эмоционалды реңкі бар сөздер қатарын толықтыру;
- #хэштегтер арқылы біріккен жазбаларды талдау жұмысын жүйелеу;
- Эмотикондарды анықтап, оларға эмоционалды реңк беруді автоматтандыру.

Қорытындылап айтқанда бұл жұмыс қазақ тілді Facebook әлеуметтік желісіндегі жазбаларды сентимент талдау жасаудың бастапқы нүктесі болып табылады.

Пайдаланылған әдебиеттер:

1 Bing Liu. Sentiment Analysis Tutorial // AAAI-2011, San Francisco, USA.

2 Thelwall, M., Buckley, K., Paltoglou, G. Cai, D., & Kappas, A. Sentiment strength detection in short informal text. // Journal of the American Society for Information Science and Technology, Vol., 2544–2558. 2010.

3 <https://developers.facebook.com/>

4 С.Мырзабеков, Қазіргі қазақ тілінің фонетикасы. Оқу құралы – Алматы: «Дәуір-Кітап», 2013 ж.-220 б.

5 Қалиев Б.А., Қазақ тілінің фонетикасы. Оқу құралы. – Алматы: «Эверо», 2014 -168 бет.

6 Шәріпбай А.Ә., Қазақ жазуын Латын әліпбиіне ауыстыру негіздемесі, Астана, 2014

Ж.Т.ҚҰЛЫМЖАНОВА, Т.К.ЖУКАБАЕВА

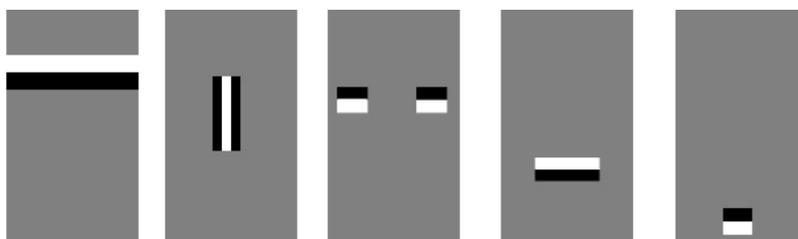
НАҚТЫ УАҚЫТТАҒЫ КОМПЬЮТЕРЛІК КӨРУ ЖҮЙЕСІНІҢ МӘЛІМЕТТЕРДІ ӨҢДЕУ ӘДІСТЕРІН ЖӘНЕ АЛГОРИТМДЕРІН ЗЕРТТЕУ

(Л.Н.Гумилев атындағы Еуразия Ұлттық Университеті, Астана, Қазақстан)

Қазіргі уақытта адам қызметінің түрлі салаларында компьютерлік көру технологиялары кеңінен қолданылуда. Компьютерлік көру –объектілердің классификациясы және қадағалау, объектілерді табуға мүмкіндік беретін машиналарды жасау технологиясы және теориясы болып табылады.

Компьютерлік көру жүйесі жасанды интеллекттің маңызды бөлігін, роботты белгілі бір орта арқылы жүргізу секілді іс-әрекеттерді орындай алатын жүйелерді құрайды. Өңдеудің бұл түрі әдетте видеосенсор ретінде әрекет жасайтын және робот және орта жөнінде жоғары деңгейдегі ақпарат беретін компьютерлік көру жүйелері кіріс мәліметтерін сұрайды. Жасанды интеллектке қатысты деп сипатталатын және компьютерлік көру жүйесі қолданылатын басқада салалар ол бейнелерді және үйретуші әдістер. Нәтижесінде, компьютерлік көру кейде жасанды интеллекттің немесе жалып компьютерлік ғылымдардың бөлігі ретінде қарастырылады.

Компьютерлік көру жүйесін іске асыру үшін, бейнелерді танып, алынған мәліметтерді талдап, танылған бейнелерді санауды жүзеге асырады. Алдымен, суреттегі адамның бетін қалай танып білу керектігі анықталынды. Біріншіден, адамның бетін бөлмедегі басқада объектілермен шатастырмай қай жерде орналасқанын дұрыс табуы керек. Оңай тапсырма болғанымен, компьютер үшін өте күрделі болып табылады. Адамның бетін табу үшін оның негізгі компоненттерін, яғни, мұрын, маңдай, көз, ерін және т.б. табуы керек. Ол үшін шаблондарды, яғни Хаар примитиві қолданылады (Сурет 1).



Сурет 1. Хаар примитиві

Егер шаблондар суреттегі нақты ауданға сәйкес келсе, онда суретте адамның беті бар деп саналады. Дегенмен осындай шаблондар өте көп. Әрқайсысы үшін суреттегі ақ және қара аудандар арасындағы айырмашылық есептеледі. Бұл мән үлгімен салыстырылып, содан кейін адамның беті бар жоғы жөнінде шешім қабылданады. Бұл әдіс Виола-Джонс әдісі деп аталады. Қазіргі таңда Виола-Джонс әдісі объектіні суреттерде іздеу үшін өзінің жылдамдығы және тиімділігі жөнінен танымал болып табылады. Виола-Джонс әдісі негізінде: Хаар белгілері бойынша суреттің интегралды ұсынылуы, адаптивті бустинг алгоритмі негізінде классификаторды құру және классификаторды каскадты құрылымға орналастыру. Бұл идеялар объектіні нақты уақытта іздеуді жүзеге асыруға мүмкіндік береді. Оларды толығырақ қарастырайық.

Суреттің интегралды бейнесі – алғашқы суреттің өлшемімен бірдей матрица. Матрицаның әрбір элементінде барлық пикселдерінің қарқындылығының суммасы сақталады. L матрицасының элементтерін формула бойынша есептеуге болады:

мұндағы, $I(i, j)$ – алғашқы сурет пикселінің жарықтығы. Матрица элементтері мәндерінің есебі, алғашқы суреттегі үйлесімді пикселдер саны уақытта өтеді, сондықтан бір

өткенде интегралды сурет есептелінеді. Матрица элементтері формула бойынша есептеледі. Суретті интегралды ұсыну көмегімен суреттің кез келген тік бұрышты аумағында жарықтың қосындысын тез есептеуге болады. Есептеу мысала А қосымшасында көрсетілген. Виола-Джонс әдісінде объектіні табу сатысында сурет бойынша қозғалатын берілген өлшемдегі терезе қолданылады. Терезе өтетін суреттің әрбір аумағында Хаар белгісі есептеледі. Көрсету белгісі, Df – белгі мәндерінің көп мүшесі. Егер f_1, \dots, f_n белгілері берілсе, белгілер векторы $x = (f_1(x), \dots, f_n(x))$, x объектісінің белгілік сипаттамасы болып табылады. Белгілік сипаттаманы объектінің өзімен сәйкестендіру керек. $X = Df_1 \times \dots \times Df_n$ көпмүшесін белгілік кеңістік деп атайды.

Белгілер Df көпмүшесіне байланысты келесі түрлерге бөлінеді:

- бинарлы белгісі, $Df = \{0, 1\}$;
- номиналды белгісі: Df – реттелген соңғы көпмүше;
- реттік белгісі: Df – нақты сандар көпмүшесі.

Хаар белгісі тікбұрышы аймақтары бойынша есептеледі. Виола-Джонс стандартты әдісінде сурет 1-де көрсетілген тік бұрышты примитивтер қолданылады. Хаар белгісінің F есептелетін мәні – примитивтің жарық бөлігінің жабылатын, X - нүкте жарығы мәнінің қосындысы, Y – қараңғы бөлігімен жабылатын жарық нүктелері мәнінің қосындысы болып табылады. Есептеу үшін интегралды сурет түсіні, жоғарыда қарастырылған, және де Хаар белгілері тез есептеледі. Хаар белгісін қолдану сәйкесінше X және Y остері бойынша жарық төмендеуінің нүктелік мәнін береді. Саны көп белгілерді қажетті және жеткілікті дәлдікпен сипаттау үшін, үйрету немесе классификация үшін Хаар белгісі сәйкес келмейді. Сондықтан Хаар белгісі каскадты классификаторға келеді. Классификатор суреттің берілген аумағынадағы ізделінетін объекті үшін сәйкес белгілерді таңдау үшін бустинг алгоритмі негізінде құрылады (boost - ағылшын тілінен аударғанда жақсару, күшею). Жалпы жағдайда бустинг – аналитикалық моделдердің дәлдігін күшейтетін әдстер кешені. Классификацияда аз қате жіберетін тиімді модель «күшті» модель деп аталады. «Әлсіз» класстарды бөлуге мүмкіндік бермейді, және де қателер саны өте көп. Сондықтанда бустинг «әлсіз» модельдерді қуаттандыру болып табылады.

Бустинг алгоритмі жұмысының нәтижесінде әрбір итерацияда классификатордың жай түрі қалыптасады:

- теңсіздік белгісінің бағыты, - шек мәні, - белгінің есептелген мәні, -24×24 пикселді өлшемді сурет терезесі.

Алынған классификаторда қателерді анықтау үшін үйрету процедурасында қосылған ағымдағы мәндердің салмақтылығына қатысты минималды қате болады. Егерде бізде суретте бір бейне емес, бірнеше адамның беті болса, және де бүкіл суретке шаблонды қолданатын болсақ, біз барлық бейнелерді таба алмаймыз, өйткені, олардың өлшемі шаблоннан кішкентай. Суреттегі әр түрлі өлшемдегі адамның бейнесін табу үшін жылжымалы терезе әдісі қолданылады. Дәл сол терезенің ішінде примитивтер есептелінеді. Аты айтып тұрғандай бұл терезе суретте барлық аймақтары бойынша жылжып өтеді. Үлкен масштабтағы бейнені табу үшін терезе суретті өтіп шыққан сайын терезе ұлғайып отырады.

Компьютерлік көру жүйесінде бейнелер, сурет, жалпы мәліметтер жұмыс жасап, оларды өңдеу үшін кітапханалар қолданылды. Зерттеу жұмысында OpenCV кітапханасын және Python бағдарламалау тілін қолданылды. OpenCV - ашық кодты жалпы мақсаттағы санаулы алгоритмдер және суреттерді өңдейтін компьютерлік жүйенің алгоритмдер кітапханасы. C/C++ тілінде жүзеге асырылған, сондай-ақ Python, Java, Ruby, Matlab, Lua тілдерінде де өңделеді. Бұл кітапхананың негізгі артықшылары:

- Компьютерлік көрудің мындаған алгоритмдері және функциясы
- Кросс-платформалы
- Өнімділік
- Либералды BSD лицензиясы

Тапсырманы шешу үшін жазылған бағдарлама коды төменде келтірілген.

```

import cv2
import sys
import os
#cascPath = sys.argv[0]
cascPath = "C://Python27/haarcascade_frontal_default.xml". format (
base_path = os.path.abspath(os.path.dirname(_file_)))
faceCascade = cv2.CascadeClassifier(cascPath)

```

CascadePath параметрі адам бейнесін анықтау үшін қолднылатын дайын мәндері бар файл атынан тұрады. Бұл файлды OpenCV директориясынан алсақ болады (opencv\build\etc\haarcascades).

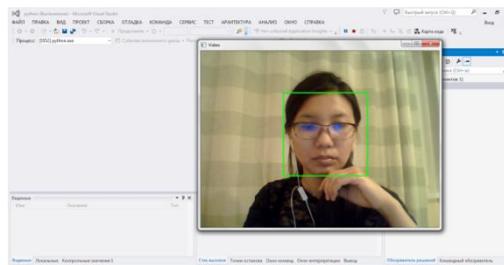
video_capture = cv2.VideoCapture(1) – бұл жол веб-камерадан видеосигналды қабылдауға мүмкіндік береді.

```

while True:
    #Capture frame-by-frame
    ret, frame = video_capture.read()

```

Бұл жерде видео бекітіледі. Read() функциясы видеодан алынған кадрларды оқиды. Ол әрбір циклдегі бір кадрды және қайтару кодын береді. Әрі қарай белгіленген кадрлардан бейнелер ізделінеді. Және де жазылған бағдарлама адам бейнесін тауып қана қоймай, табылған адам бейнесін нақты уақытта санайды. Бағдарлама нәтижесіне төменде келтірілген (Сурет 2).



Сурет 2. Бейнені нақты уақытта іздеп табу

Сурет 3-те көрсетілгендей, ол жерде адамның бейнесін тауып қана қоймай, қанша бейне табылғанын санап көрсетіп тұр.



Сурет 3. Табылған бейнені санау

Қорытындылай келе, зерттеу кезінде анықтау әдістері және сурет фрагменттері классификациясы, Виола-Джонс әдісі зерттелді. Іздеу алгоритмін байланысқан аймақтарды белгілеуін және нормализациялауын және табылған символдар анықталып, бейнелерді нақты уақытта іздеп, оларды санау жүзеге асырылды.

Ш.Ф.МАДРАХИМОВ**ВЫЧИСЛЕНИЯ ОБОБЩЕННЫХ ОЦЕНОК ПО ОПРЕДЕЛЕННЫМ НАБОРАМ
НОМИНАЛЬНЫХ ПРИЗНАКОВ И ИХ ИНТЕРПРЕТАЦИЯ**

(Национальный университет Узбекистана им.Мирзо Улугбека)

Рассматривается задача поиска скрытых закономерностей по результатам отображения объектов на числовую ось по определенным наборам номинальных признаков. Предложен метод разбиения признаков на наборы через построения графов связанностей признаков. Обобщенные оценки по каждому набору признаков интерпретируется как латентный количественный признак для описания объектов в двухклассовой задаче распознавания. Выявленные значения характеризующие закономерности по латентному признаку, рассматривается как значения лингвистической переменной и интерпретируется.

Ключевые слова: латентные признаки, графы, обобщенные оценки, скрытые закономерности.

Введение

Поиск скрытых закономерностей в базах данных одно из важнейших направлений интеллектуального анализа данных.

Выявленные закономерности являются источником новых знаний, которые используются для объяснения процесса принятия решений в различных предметных областях.

С помощью интеллектуального анализа данных эксперты имеют возможность выдвигать и проверять свои гипотезы относительно рассматриваемой задачи.

По результатам проверки добытое (извлечённое) знание могут быть представлены в виде лингвистических правил, а в ряде случаев записываться в виде чётких формул.

Рассматривается задача сокращения размерности признакового пространства путём отображения описаний объектов на числовую ось по определяемым наборам номинальных признаков.

Результат отображения каждого набора признаков интерпретируется как латентный (явно неизмеримый) количественный признак для описания объектов в двухклассовой задаче распознавания.

Множества значений латентного признака разбивается на непересекающиеся интервалы, границы которых вычисляются по специальному критерию.

Множество значений критерия в $[0,1]$ интерпретируются в терминах нечеткой логики.

Для данных описываемых номинальными признаками потенциальная возможность обнаружения закономерностей существенно ограничена.

Это объясняется тем, что в рамках теории измерений свойства отношений по значениям номинального признака определены как слабая шкала.

Принадлежность латентных количественных признаков к сильным шкалам измерений значительно увеличивает возможности по обнаружению закономерностей по их значениям.

Предлагается процесс добычи новых знаний проводить в форме последовательного выполнения следующих этапов:

- отображение описаний объектов по указанным наборам номинальных признаков на числовую ось и интерпретация результатов отображения как латентных признаков;
- разбиение значений латентных признаков на непересекающиеся интервалы по специальному критерию принимающего оптимальное значение в $[0,1]$;
- значений критерия интерпретируется в термины предметной области с использованием аппарата нечёткой логики.

Для реализации целей исследования требуется решить две задачи оптимизации:

– отображение описаний объектов на числовую ось по указанным наборам номинальных признаков;

– разбиение результатов отображение на непересекающиеся интервалы.

Проблема извлечения знаний по результатам последовательного решения описанных выше задач рассматривается впервые.

1. Постановка задачи

Считается, что определён алгоритм A для отображения описаний объектов E_0 по указанному набору $X(k) \subset X(n)$, $1 < k \leq n$ на числовую ось и критерий R для разбиения результата отображения на два непересекающихся интервала, множество оптимальных значений которого принадлежит $[0,1]$.

Требуется:

– алгоритмом A отобразить описание объектов E_0 по набору признаков $X(k)$, $1 < k \leq n$ на числовую ось;

– получить оптимальное по критерию разбиение результата отображения на два интервала.

2. Отображение наборов номинальных признаков на числовую ось

Обозначим через I – множество индексов в наборе номинальных признаков.

Отображение указанного подмножества номинальных признаков $X(k)$ на числовую ось осуществим через вычисления обобщенной оценки принадлежности объектов классов к классу K_d , $d \in \{1, 2\}$ [2,3].

Обозначим через p число градаций признака $r \in I$, g_{dr}^t – количество значений t -й ($1 \leq t \leq p$) градации r -го признака в описании объектов класса K_d , l_{dr} – число градаций r -го признака в K_d .

Различие по r -му признаку между классами K_1 и K_2 определяется как величина

$$\lambda_r = 1 - \frac{\sum_{t=1}^p g_{1r}^t g_{2r}^t}{|K_1| |K_2|} \quad (1)$$

Степень однородности (мера внутриклассового сходства) β_r значений градаций r -го признака по классам K_1, K_2 вычисляется по формулам:

$$D_{dr} = \begin{cases} (|K_d| - l_{dr} + 1)(|K_d| - l_{dr}), & p > 2, \\ |K_d|(|K_d| - 1), & p \leq 2; \end{cases}$$

$$\beta_r = \begin{cases} \frac{\sum_{t=1}^p g_{1r}^t (g_{1r}^t - 1) + g_{2r}^t (g_{2r}^t - 1)}{D_{1r} + D_{2r}}, & D_{1r} + D_{2r} > 0, \\ 0, & D_{1r} + D_{2r} = 0. \end{cases} \quad (2)$$

С помощью (1), (2) вес номинального признака $r \in I$ определяется как

$$v_r = \lambda_r \beta_r. \quad (3)$$

Легко проверить, что множество значений весов номинальных признаков, вычисляемых по (3), принадлежит интервалу $[0,1][0,1]$.

Очевидно, что множество чисел, идентифицирующих p градаций номинального признака, всегда можно взаимно однозначно отобразить в множество $\{1, \dots, p\}$. С учётом такого отображения для объекта $S = (x_1, \dots, x_n)$ вклад признака $x_i = j, j \in \{1, \dots, p\}$ в обобщённую оценку определяется величиной

$$\mu_i(j) = v_i \left(\frac{\alpha_{ij}^1}{|K_1|} - \frac{\alpha_{ij}^2}{|K_2|} \right),$$

где $\alpha_{ij}^1, \alpha_{ij}^2$ – количество значений j -й градации i -го признака соответственно в классах K_1 и K_2 , v_i – вес i -го признака, вычисляемый по (3).

Обобщённая оценка объекта $S_i \in E_0, S_i = \{x_{ij}\}_{j \in I}$ по набору $X(k)$ определяемым множеством индексов $I, |I| = k$, вычисляется как

$$R(S_i) = \sum_{j \in I} \mu_i(x_{ij}) \quad (4)$$

Результаты вычисления обобщённых оценок $R(S_1), R(S_2), \dots, R(S_m)$ рассматриваются как значения нового латентного, количественного признака u в описании объектов выборки E_0 .

Теоретически число различных латентных признаков ограничено $C_n^2 + C_n^3 + \dots + C_n^{n-1}$.

Степень делимости (перемешанности) – $w(u)$ значений признака u в описании объектов классов K_1 и K_2 вычислим следующим образом.

Упорядоченное множество значений признака u разбивается на два интервала $[c_0, c_1](c_1, c_2]$.

Критерий для определения границы c_1 основывается на проверке гипотезы (утверждения) о том, что каждый из двух интервалов содержит значения количественного признака объектов только одного класса.

Пусть $u_1^1, u_1^2(u_2^1, u_2^2)$ – количество значений признака u класса $K_i, i = 1, 2$ соответственно в интервалах $[c_0, c_1]$ и $(c_1, c_2]$, p – порядковый номер элемента упорядоченной по возрастанию последовательности $r_{j1}, \dots, r_{jp}, \dots, r_{jm}$ значений u из E_0 , определяющий границы интервалов как $c_0 = r_{j1}, c_1 = r_{jp}, c_2 = r_{jm}$.

Критерия

$$\left(\frac{\sum_{p=1}^2 \sum_{i=1}^2 u_i^p (u_i^p - 1)}{m_1(m_1 - 1) + m_2(m_2 - 1)} \right) \left(\frac{\sum_{p=1}^2 u_1^p (m_2 - u_2^p) + u_2^p (m_1 - u_1^p)}{2m_1 m_2} \right) \rightarrow \max \quad (5)$$

позволяет вычислять оптимальное значение границы между интервалами $[c_0, c_1][c_1, c_2]$.

Пусть $w(y)$ – оптимальное значение критерия (5) по признаку y , c_0, c_1, c_2 – соответствующие этому значению концы интервалов разбиения и по K_1 и K_2 .

Значение $w(y) \in [0, 1]$ в терминах нечеткой логики можно использовать для описания закономерностей на естественном языке с помощью лингвистических переменных.

Например, принадлежность $w(y)$ одному из интервалов $[0, 0.3], (0.3, 0.6], (0.6, 1]$ может интерпретироваться как “слабое”, “удовлетворительное”, “сильное”.

3. Вычислительный эксперимент

В вычислительном эксперименте были использованы данные опроса врачей по применению лекарственных препаратов для лечения больных бронхиальной астмой [4].

Классификация данных опроса производилась исходя из ответов врачей по специализации – аллергологи, пульмонологи, терапевты.

Использовались три варианта градаций ответа по каждому препарату (VEN-оценка): 3 – жизненно важный и улучшающий качество жизни (Vital – V); 2 – необходимый (Essential – E); 1 – второстепенный (Non-essential – N).

Было сформирована выборка из 91 объекта (экспертов-врачей), описываемых 102 признаками со значениями градаций лекарственных препаратов.

На базе этой выборки проверялось ряд гипотез по критерию (5), для описания результатов которых использовалась лингвистическая переменная [3] “Мнения экспертов аллергологов, пульмонологов, терапевтов по VEN оценке набора лекарственных препаратов”.

Изучалось мнение двух групп врачей K_1 и K_2 , представленных соответственно аллергологами ($|K_1| = 23$), пульмонологами и терапевтами ($|K_2| = 68$), по 3-м указанным наборам лекарственных препаратов:

- подмножества χ^1 лекарств группы «Глюкокортикоидные средства» - {Дексаметазон, Беклометазон, Преднизолон, Флутиказон, Метилпреднизолон, Триамцинолон, Будесонид};

- подмножества χ^2 лекарств группы «Бронхолитические средства» - {Аминофиллин, Сальбутамол, Флутиказон пропионат, Аминофиллин, Теофиллин, Орципреналин, Фенотерол-ипратропия бромид, Фенотерол, Сальметерол};

- подмножества χ^3 лекарств группы «Антибиотики» - {Мидекамицин, Цефтазидим, Цефотаксим, Цефоперазон, Цефазолин, Цефтриаксон, Азитромицин}.

Множество значений, вычисленных по (4) и являющимися обобщенными оценками (признаками y^1, y^2, y^3) по наборам χ^1, χ^2, χ^3 , анализировались по критерию (5).

Для интерпретации результатов анализа использовались значения лингвистической переменной, представленные в табл.1.

Таблица 1. Значения лингвистической переменной

Interval	The value of the variable
[1..1]	<i>Absolutely do not match</i>
[0,7..1)	<i>Virtually do not match</i>
[0,5..0,7)	<i>Few match</i>
[0,2..0,5)	<i>Substantially coincide</i>
(0..0,2)	<i>Practically the same</i>
[0..0]	<i>Fully matched</i>

В ходе вычислительного эксперимента проверялось истинность гипотезы "Мнения врачей из группы K_1 и K_2 по использованию наборов лекарственных препаратов χ^1, χ^2, χ^3 для лечения бронхиальной астмы существенно совпадают".

Результаты вычислительного эксперимента приведены в табл. 2.

Таблица 2. Результаты вычислительного эксперимента по наборам признаков

Набор признаков	Значение (S) по обобщенным оценкам	Значение лингвистической переменной
χ^1	0,55	<i>Few match</i>
χ^2	0,51	<i>Few match</i>
χ^3	0,36	<i>Substantially coincide</i>

Как видно из табл.2 истинность гипотезы подтверждается только для набора χ^3 .

Следует отметить, что наборы номинальных признаков могут быть сформированы (а не указаны как в предыдущем случае) по специальным критериям.

Например, критериям кластерного анализа. Изучение других методов формирования наборов признаков и использование различных способов задания значений лингвистической переменной является предметом дальнейших исследований.

Заключение

Поиск закономерностей по выборке осуществляется путём отображения описаний объектов на числовую ось по определяемым наборам номинальных признаков.

Результат отображения каждого набора интерпретируется как латентный количественный признак для описания объектов в двухклассовой задаче распознавания.

На множестве значений латентного признака производится разбиение на непересекающиеся интервалы, границы которых вычисляются по специальному критерию.

Множество значений критерия в $[0,1]$ интерпретируются в терминах нечеткой логики.

Результаты исследования могут применены для объяснения процесса принятия решения в задачах распознавания образов из различных предметных областей.

Литература

1. Ignatev N.A. Vchislenie obobshyonnix pokazateley i intellektualniy analiz dannix // Avtomatika i telemexanika. 2011. № 5. S.183-190.
2. Ignatev N.A. Intellektualniy analiz dannix na baze neparаметricheskix metodov klassifikatsii i razdeleniya viborok ob'ektov poverxnostyami. – Tashkent: Natsionalny universitet Uzbekistana im. Mirzo Ulugbeka, 2009. – 120 s.
3. Kruglov V. V., Dli M. I., Golunov R. Yu. Nechetkaya logika i iskusstvennie neyronnie seti: Ucheb.posobie. – M.: Izdatelstvo Fiziko-matematicheskoy literatur, 2001. – 224 s. – ISBN 5-94052027-8.
4. Suyunov N. D., Madraximov Sh.F., Ikramova G. M. Analisis of drug provision of patients with bronchial asthma // News of Science and Education, NR 3 (27) 2015, Sheffield, Science And Education Ltd, p. 27-35.

А.Н. МОЛДАГУЛОВА, Р.К. УСКЕНБАЕВА, Р.Ж. САТЫБАЛДИЕВА

К ВОПРОСУ ОБРАБОТКИ НЕСТРУКТУРИРОВАННЫХ ДАННЫХ

(Международный Университет Информационных Технологий, Алматы, Казахстан)

В настоящее время анализ неструктурированных данных стал одним из важных задач в связи с необходимостью обработки таких данных, как электронная почта, новости, чаты, сообщения, комментарии, маркетинговые материалы, презентации и другие документы, которые не могут быть обработаны с помощью реляционных баз данных. Согласно опросу IDC, более 80% всех хранимых данных предприятий являются неструктурированными [1], и их количество растет гораздо быстрее по сравнению со структурированными данными. Сотрудники офисов проводят свое рабочее время над рутинной, неавтоматизированной работой, связанной с необходимостью обработки документов, которые могут храниться в виде текстовых файлов в различных форматах, и эти файлы могут иметь свою внутреннюю структуру. Хотя интерес к работе с неструктурированными данными возник еще в 1950-х годах [8], тем не менее методы обработки данных улучшались не так быстро как рост объема этих данных, и это не столько технологии, сколько результаты научных исследований. Однако технологии Больших Данных (Big Data) значительно ускорили этот процесс.

Анализ неструктурированных данных является одной из нишевых областей Больших Данных. Проблема обработки неструктурированных данных появилась в рамках дисциплины Text Mining.

1. Обзор современного состояния обработки неструктурированных данных

Представление неструктурированных данных определяет тип алгоритмов, применяемых для их интеллектуального анализа. Одним из типов неструктурированных данных являются текстовые документы. Существует несколько различных представлений текстовых данных: мешок слов (bag of words), существительные фразы (noun phrases - NP) и именованные сущности (named entities). Schumaker и Chen [2] использовали эти методы для представления финансовых новостей для определения важных терминов.

Мешок слов представляет собой модель представления текста, где текст или документ представляется в виде набора слов, без учета грамматики и порядка слов, но сохраняя при этом множественность [2]. Эта модель обычно используется в задаче текстовой классификации, где частота встречаемости слов используется как атрибут для обучения классификатора.

Например, рассмотрим два простых текстовых документа:

«Аскар любит играть в шахматы. Дина тоже любит шахматы».

«Аскар также любит играть в футбол».

Из этих двух текстовых документов можно построить мешок слов следующим образом:

```
{  
  "Аскар" : 1,  
  "любит" : 2,  
  "играть" : 3,  
  "в": 4,  
  "шахматы" : 5,  
  "Дина" : 6,  
  "тоже" : 7,  
  "также" : 8,  
  "футбол" : 9  
}
```

Словарь содержит 9 различных слов. И используя индексы словаря, каждый документ может быть представлен вектором из 9 элементов, где каждый элемент векторов относится к соответствующему элементу в словаре, как показано ниже:

[1, 2, 1, 1, 2, 1, 1, 0, 0]

[1, 1, 1, 1, 0, 0, 0, 1, 1]

Существительные фразы определяются как единицы, основным словом которых является существительное. Они являются основными носителями содержания документа. Поэтому идентификация NP является ключевой во многих алгоритмах интеллектуального анализа текста. В задачах анализа текстов **именованный объект** представляет собой фразу, которая четко идентифицирует один элемент из набора других элементов, имеющих аналогичные атрибуты. Примерами именованных объектов являются имя и фамилия, географическое расположение, возраст, адрес, номер телефона, название компании и т.п. Именованные объекты часто используются для целей извлечения информации. Известно, что алгоритмы интеллектуального анализа текста работают с представлениями документов, основанными на их признаках [5]. Фельдман и Зангер [5] разделили наиболее часто используемые представления, основанных на признаках на четыре типа: основанные на символах, основанные на словах, основанные на терминах, основанные на концепции.

Представление документов на основе символа использует набор символов в документе и содержит позиционную информацию, такую как биграммы или триграммы.

А в представлении, основанном на словах, представляется полным набором слов в этом документе.

Представление документа на основе терминов обязательно состоит из подмножества терминов, извлеченных из этого документа. Существует различные методологии извлечения терминов.

В концептуальных представлениях характеристики документа, так называемые концепции, генерируются с помощью ручных, статистических, основанных на правилах или гибридных методологиях категоризации.

Многие исследователи доказали, что текстовое представление имеет много преимуществ благодаря простоте в алгоритмах проектирования, которые эффективно декомпозируют текст на слова [5]. Кроме того, текстовые представления считаются очень продуктивными в задачах классификации текста и категоризации текстового поиска [9]. Фельдман и Зангер [5] утверждают, что «конкретные слова, выбранные непосредственно из «исходного» документа, находятся на основном уровне семантического множества» [5].

Одной из задач интеллектуального анализа текста является категоризация текста. Цель задачи категоризации текста - классифицировать документы по заданным темам. Основные методы классификации текста, такие как деревья решений, максимальное моделирование энтропии, перцептроны и классификация ближайших соседей, описаны в книге «Основы статистической обработки естественного языка», написанной Мэннинг и Шютце [6]. Классификация текста (Text Classification - TC) или определение темы - это процесс маркировки текстов естественного языка по тематическим категориям из предопределенного набора. Термин «Text mining» все чаще используется для обозначения всех задач, которые пытаются извлечь полезную информацию из огромного объема текстовых данных. Согласно этой точке зрения, TC является экземпляром текстового анализа [4].

Поддержка векторной машины (Support Vector Machine - SVM) является одним из популярных и мощных алгоритмов машинного обучения для классификации текста. Фан и Ватанабэ [3] использовали подход выбора терминов зависимых слов в качестве признаков статей и использования Support Vector Classification для классификации новостных статей. Наши исследовательские задачи соответствуют вышеупомянутой области научного направления Text Mining.

Hadoop является доминирующей платформой для неструктурированной аналитики данных для крупных веб-компаний. Наряду с Hadoop, модель программирования MapReduce

позволяет намного быстрее анализировать данные, а также значительно упростить их сбор и отладку.

2. Методология исследования

Существуют различные методы обработки неструктурированных данных, связанных с парадигмой больших данных. В основном, эти подходы основаны на научных направлениях информатики, таких как обработка естественного языка (Natural Language Processing - NLP), обнаружение знаний, интеллектуальный анализ данных, извлечение информации и статистика.

В таблице 1 приводится краткое изложение методологий.

Таблица 1. Обобщенные методологии неструктурированного интеллектуального анализа данных (Источник: [7])

	Представление данных	Выход	Инструменты	Проблемы и будущие исследования
Извлечение информации	Длинный или короткий текст, семантика и онтологии, ключевые слова	Документы (структурированные или полуструктурированные)	Общий поиск данных, Тегирование, Нормализация данных, Токены, Обнаружение сущностей	Отсутствие исследований по этимологии данных
Правила ассоциации	Ключевые слова, длинные или короткие предложения	Визуализация, сводный отчет	Онтологии, семантика, лингвистика	Применение различных подходов к различным источникам данных. Например, реляционные, пространственные и транзакционные хранилища данных
Темы	Ключевые слова, Лексическая цепочка	Тематические связи	Публикация / Подписка, Темы-Ассоциации - События	Идентификация интересующих тем может быть сложной
Термы	Ключевые слова	Визуализация, глубина краулера	Лингвистическая предварительная обработка, генерация термов	Применение к междоменному подходу инструмента, основанного на сообществе является сложным
Кластеризация документов	Данные, документы	Визуализация	K-means кластеризация, иерархическая кластеризация	Может быть ресурсоемким, поэтому необходимо провести исследования по распараллеливанию
Сводные документы	Термы, темы, документы	Визуализация, сводный отчет	Анализ структуры документов, Классификация документов	Нуждается в исследовании этимологии данных
Многоразовые словари	Слова, предложения	Сводный отчет	Аннотации, маркировка	Отсутствие исследований по адаптации словаря к новым словам

Методы классификации текстовых документов находятся на стыке двух областей - поиск информации и машинное обучение. Общие части двух из этих подходов - способы подачи документов и способы оценки качества классификации текстов, а различия состоят только в способах собственно поиска.

Несмотря на то, что проблемы классификации текстовых документов находятся в центре внимания ряда исследовательских групп, по многим вопросам до сих пор не найдены удовлетворительные ответы. Точность различных методов существенно зависит от реализации априорных предположений и допущений, а также от структуры текстовых данных. На рисунке 1 показана общая модель процесса для приложения классификации текста с использованием алгоритмов машинного обучения.

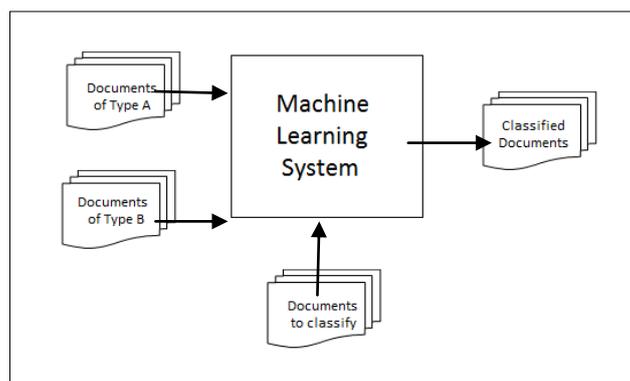


Рис1. Пример классификации текста (R. Feldman, J. Sanger [5]).

3. Результаты

Широкий круг задач, которые могут использовать методы классификации, а также наличие больших коллекций документов в электронной форме, вызвал живой интерес к области классификации текста со стороны многих прикладных дисциплин. Классификация текста используется во многих сферах: от автоматической индексации документов до фильтрации документов, автоматической генерации метаданных, расширения иерархического каталога веб-ресурсов и структурирования документов. Например, классификация текста может быть использована для классификации статей в онлайн-журналах, предложений по трудоустройству через Интернет и т. д. Офисные работники проводят свое рабочее время в рутинной неавтоматизированной работе, связанной с необходимостью обработки неструктурированных данных: электронных писем, заметок, новостей, чатов, отчетов, маркетинговых материалов, презентаций и других документов, которые не могут быть обработаны должным образом в реляционных базах данных, т.к. они хранятся в виде текстовых файлов в различных форматах, и эти файлы могут иметь свою характерную внутреннюю структуру. Более того, методы текстовой классификации могут быть адресованы вопросам анализа текстовых документов на предмет определения критериев с дальнейшей идентификацией дублирующих документов, документов, задерживающих выполнение правительственных поручений, документов, содержащих запросы для исходной информации и отчетов, которые требуют большого количества рабочего времени, отвлекая сотрудников от основной работы на ответы на письма с незначительным содержанием.

Выводы

Хотя обработка неструктурированных данных является не новой концепцией, но тем не менее ее находят очень сложной. Она требует масштабируемого индекса хранения и распределенного подхода для получения требуемых результатов в режиме реального времени. Основопологающим фактом является то, что данные слишком велики для традиционной обработки. Таким образом, требуется особый метод для анализа большого количества текстовых документов, поступающих из разных источников. Для этого мы выбрали методологию исследования, состоящую из комбинации методов обработки естественного языка и подхода машинного обучения.

Список литературы

- 1 Nadkarni, A., & Yezhkova, N. (2014). Structured Versus Unstructured Data: The Balance of Power Continues to Shift. Doc # 247106. Retrieved from <http://www.idc.com/getdoc.jsp?containerId=247106>
- 2 Schumaker, R. P., & Chen, H. (2009). Textual analysis of stock market prediction using breaking financial news: The AZFin text system. *ACM Trans. Inform. Syst.* 27, 2, Article 12, 19 pages.
- 3 Fan, W., & Watanabe, T., (2012). Dynamic Prediction of Forthcoming Trends in Stock Prices from News Articles. *WIMS '12*, Craiova, Romania.
- 4 Sebastiani, F. (2002). Machine Learning in Automated Text Categorization. *ACM Computing Surveys*, Vol. 34, No. 1, pp. 1–47.
- 5 Feldman, R., & Sanger, J. (2007). *The Text Mining Handbook. Advanced Approaches in Analyzing Unstructured Data.* Cambridge University Press.
- 6 Manning C. D., & Schütze, H. (2000). *Foundations of Statistical Natural Language Processing.* The MIT Press.

РАЗРАБОТКА ИНТЕЛЛЕКТУАЛЬНОГО ЭЛЕКТРОННОГО УЧЕБНИКА ПО ДИСЦИПЛИНЕ «ПРОГРАММНАЯ ИНЖЕНЕРИЯ»

(Евразийский национальный университет имени Л.Н.Гумилева, Астана, Казахстан)

Одним из главных качеств информатизации образования является разработка электронных средств поддержки процесса обучения. В этом спектре программных продуктов наиболее востребованными являются электронные учебные издания [1]

Содержание интеллектуального электронного учебника по дисциплине «Программная инженерия» составлено в соответствии с Силабусом одноименной обязательной дисциплины для магистерской образовательной программы, реализуемой в рамках PROMIS (PROfessional network of Master's degree in Informatics as a Second Competence - Сетевая магистерская программа «Информатика как вторая компетенция»).

Целью исследования является разработка интеллектуального электронного учебника по дисциплине «Программная инженерия». В ходе были определены цели, задачи обучения, изучаемые темы. Интеллектуальный электронный учебник - это система искусственного интеллекта, имеющая базу знаний предметной области.

В ходе работы был составлен тезаурус определений (рис.1), тезаурус это словарь, где все значения слов связаны семантическими отношениями друг с другом и отражают ключевые отношения определений в конкретной предметной области. Тезаурус является: источником специальных знаний в широкой либо узкой предметной области.

	В	С	D	Е	F	G
1	Субъект	Предикат		Объект(значение свойства)	гипероним (выражающее более общую сущность)	гипоним(обозначающее видо понятие)
2		прилагательное(глагол			
3	Программная инженерия	систематический, дисциплинированный, измеримый	есть	подход к (разработке, функционированию, сопровождению) программного обеспечения систематическим, дисциплинированным, измеримым подходом к разработке, функционированию и сопровождению программного обеспечения		
4	Программная инженерия		является	программное средство и документации к ним для решения классов задач в конкретных предметных областях	Программное обеспечения	Программирование
5	Программное обеспечение	комплексное	является		Программное обеспечения	

Рисунок 1 Тезаурус определений

Также была составлена онтология предметной области. Под онтологией будем понимать как точную спецификацию концептуализации. Концептуализация предполагает описание множества объектов и понятий некоторой предметной области, знаний об этих объектах и понятиях и связей между ними.[2] Онтология описывает основные концепции

НУРГАЗИНОВА Г.Ш., ШАРИПБАЙ А., ОМАРБЕКОВА А.С., ЗУЛХАЖАВ А.

ИНТЕЛЛЕКТУАЛЬНАЯ СИСТЕМА «IT-TEST.KZ» ПО АЛГЕБРЕ НА КАЗАХСКОМ ЯЗЫКЕ

(Евразийский национальный университет имени Л.Н. Гумилёва)

На сегодняшний день в цифровизации общества особо важную роль приобретают исследования и разработки, направленные на повышение эффективности средств электронного обучения и контроля знаний, а также сокращение затрат на их разработку. Одним из возможных путей совершенствования инфраструктуры электронного обучения и повышения его эффективности является создание интеллектуальных средств обучения путем интеграции новых программных сценариев и онтологий предметной области изучения предмета и формирования основ интеллектуальной оценки знаний.

Наиболее удачно и успешно вопросы формализации знаний решаются в таких предметных областях обучения, как математика. Сейчас более чем когда-либо, математические навыки имеют основополагающее значение для успешной работы в любой сфере деятельности, в том числе и в образовании.

Для представления знаний в формализованном виде предметной области могут быть использованы следующие часто используемые и популярные на сегодняшний день модели представления знаний (логические модели, продукционные модели, сетевые модели, фреймовые модели, математические модели), которые относятся к прагматическому направлению (основано на предположении о том, что мыслительная деятельность человека «черный ящик») в исследованиях по искусственному интеллекту.

Наиболее в большей мере организации долговременной памяти человека соответствует сетевые модели представления знаний, но тем самым являются наиболее сложными в создании и модификации. Это так называемые семантические модели представления знаний.

На практике онтологии часто разрабатываются с использованием интегрированных графических средств разработки онтологий, таких как Protégé, OIled и OntoEdit. Они используются для разработки новых онтологий и модификации существующих.

В данный момент в научно-исследовательском институте «Искусственный интеллект» ЕНУ им. Л.Н. Гумилева разрабатываются интеллектуальные средства электронного обучения и контроля знаний.

Данное направление разработок порождает новую парадигму образования - Смарт-образование, в которой статические электронные ресурсы будут заменены на динамические смарт-тьюторы.

Разрабатываемая интеллектуальная система «It-test.kz» удовлетворяет следующим требованиям:

- представлять формализованные знания, так что их семантика должна быть полностью сохранена;
- быть недорогой, или менее дорогой с точки зрения вычислительной стоимости разработки, чем существующие механизмы, которые обычно используются при разработке интеллектуальных обучающих систем (ИОС);
- быть в соответствии с действующими современными стандартами и Web-технологиями обучения.

Для создания базы знаний интеллектуального средства обучения мы структурировали базу знаний по элементарной алгебре, представленную в виде семантической сети.

В результате работ [1-3] разработана база знаний интеллектуального средства обучения и контроля знаний по элементарной алгебре. В среде Protege построена онтология

«Алгебра», подробнее которую можно рассмотреть в Интернете по следующей ссылке http://e-zerde.kz/zerde/Algebra_ru/OntologAlgebra.jpg.

При проектировании интеллектуальной системы были использованы следующие языки программирования и технологии: java, apache jena, owl, gae, json COS.

Обучение реализовано в виде вопросно-ответной системы. Система выдает ответ на поставленный вопрос ученика путем нахождения его в базе знаний.

Интеллектуальные средства обучения в основном состоят из следующих компонентов:

- База знаний, которая основана на онтологии и тезаурус;
- Машина обработки знаний;
- Пользовательский интерфейс, необходимый для установления интерфейса между компьютером и пользователем на естественном языке.

В нашем случае, в качестве пользовательского интерфейса мы рассматриваем интеллектуальную вопросно-ответную систему.

Вопросно-ответная система – информационная система, способная принимать вопросы и отвечать на них на естественном языке, другими словами, это система с естественно-языковым интерфейсом. Для концептуального структурирования знаний предметной области вопросно-ответной системы применяется онтология. В общих чертах, под онтологией понимается система понятий некоторой предметной области, которая представляется как набор сущностей, соединенных различными отношениями. Онтологии используются для формальной спецификации понятий и отношений, которые характеризуют определенную область знаний. Преимуществом онтологий в качестве способа представления знаний является их формальная структура, которая упрощает их компьютерную обработку.

Таким образом, вопросно-ответную систему можно отнести к классу интеллектуальных средств обучения.

В вопросно-ответной системе используются технологии обработки естественного языка. Используя данные технологии разработанная вопросно-ответная система обрабатывает полученный от пользователя ответ на естественном языке в зависимости от смысла и дает ответ на заданный вопрос.

Любая вопросно-ответная система изначально получает текст вопроса на естественном языке. Затем данный текст обрабатывается, осуществляется поиск и на выходе выдается ответ [4].

Вопросно-ответная система состоит из следующих приложений:

- Формулировка и задание вопроса
- Словарь
- Онтология
- Схема вопросов

Задать вопрос системе можно на естественном языке и на языке SparQl запросов. Словарь необходим для введения новых понятий, не имеющих в базе, зная класс понятия и указав его значение на английском языке. В приложении Онтология имеется возможность изменения существующей онтологии или введения новой онтологии другой предметной области.

Предположим, что ученику задан вопрос и получен ответ на него. Вопросы и ответы приведены на казахском языке.

Вопрос: Рационал сандар жиыны қалай бөлінеді?

Эталонный ответ: *Рационал сандар жиыны бүтін және бөлшек сандар жиынына бөлінеді.*

Термин / Ұғым

Рационал сандар жиыны

Анықтамасы / Түсініктемесі

Рационал сандар жиыны бүтін және бөлшек сандар жиынына бөлінеді

Сұрақ қою түрі

қалай бөлінеді

Түрі

Анықтама

Рационал - сөз табы: Сын есім
сан - сөз табы: Зат есім Қосымшалары: -дар(көптік)
жиын - сөз табы: Зат есім Қосымшалары: -ы(III жақ тәуелдік)
Рационал - сөз табы: Сын есім
сан - сөз табы: Зат есім Қосымшалары: -дар(көптік)
жиын - сөз табы: Зат есім Қосымшалары: -ы(III жақ тәуелдік)
бүтін - сөз табы: Сын есім
және - сөз табы: Шылау
бөлшек - сөз табы: Зат есім
сан - сөз табы: Зат есім Қосымшалары: -дар(көптік)
жиын - сөз табы: Зат есім Қосымшалары: -ы(III жақ тәуелдік) -на(барыс септігі)
бөл - сөз табы: Етістік Қосымшалары: -н(VerbsToVerbs) -е(VerbsToVerbs) -д(III-жақ жіктік)
бөлшек → сан(бөлшек,сандар);

Талданған сөйлемдер:1

Рисунок 1. Фрагмент интеллектуальной системы в обработке

Пәнді таңдау:

Сұрақ

Рационал сандар жиыны қалай бөлінеді?

Жауабы

Рационал сандар жиыны бүтін және бөлшек сандар жиынына бөлінбейді

Сіздің жауабыңыз сұраққа сәйкес емес. Себебі: рационал сандар жиыны бүтін және бөлшек сандар жиынына бөлінеді, ал сіздің жауабыңыз: Рационал сандар жиыны бүтін және бөлшек сандар жиынына бөлінбейді

Рисунок 2. Фрагмент вопросно-ответной системы

Сұрақ

Рационал сандар жиыны қалай бөлінеді?

Жауабы

Рационал сандар жиыны бүтін және бөлшек сандар жиынына бөлінеді

Сіздің жауабыңыз 100%

Рисунок 3. Фрагмент оценивания результата

Ответ ученика: бүтін мен бөлшек сандар жиыны рационал сандар жиынын құрайды.

Оценивание результатов обучения основано на сравнении моделей знаний обучаемого и эталонной модели знаний, которые были формально определены согласно работе [5].

После определения соответствия двух предложений, найденный коэффициент умножается на 100%, для определения процентного соотношения оценки.

Создание и развитие интеллектуальных средств обучения и контроля знаний позволяет свободный доступ к образованию и существенно повысить его качество, поэтому они будут находиться в центре внимания многих современных исследователей, как зарубежных, так и отечественных.

Список использованных источников:

1 Шарипбай А.А., Омарбекова А.С., Нургазиева Г.Ш. Проектирование базы знаний интеллектуальной справочной системы по алгебре. //Материалы международной научно-технической конференции OSTIS (г. Минск, 19-21 февраля 2015) Минск: БГУИР, 2015. – с. 157-160.

2 On creation of the intelligent help system Journal of Theoretical and Applied Information Technology: Pakistan, 10th April 2016. Vol.86. No.1, p. 26-33 ISBN 1992-8645, входящий в базу данных Scopus (Elsevier, Нидерланды) SJR 0.151 www.jatit.org

3 Designing of intelligent reference system for algebra based on the knowledge database. 2015 International Conference on Control, Automation and Artificial Intelligence (CAAI2015): Phuket, Thailand, August 23-24, 2015, p. 230-235 <http://www.caaai2015.org>

4 Муқанова А.С., Асқарова С.А. Қазақ тілі бойынша интеллектуалды анықтамалық жүйенің білімдер базасы // Вестник.ЕНУ.- Астана: ЕНУ, 2015.-С.91-95

Ж.Н. НҮРЛАНОВА, Ж.С. САУХАНОВА

ПРИМЕНЕНИЕ КЛАССИФИКАЦИОННЫХ МЕТОДОВ ИНТЕЛЛЕКТУАЛЬНОГО АНАЛИЗА ДАННЫХ ДЛЯ ПРОГНОЗИРОВАНИЯ СЕРДЕЧНЫХ ЗАБОЛЕВАНИЙ

(ЕНУ им. Л.Н. Гумилева, Астана, Казахстан)

При возросшей частоте сердечных заболеваний классификация и прогнозирование сердечных заболеваний значительно важны. Автоматизация классификации сердечных заболеваний более полезна для врачей для быстрого диагноза. Точно предсказание сердечной болезни может помочь в спасении жизни пациентов. В данной статье описывается применение классификационных методов интеллектуального анализа данных для прогнозирования сердечных заболеваний. Используемый набор данных - это база данных Cleveland Heart, взятая из репозитория наборов данных обучения UCI [1].

Симптомы сердечного заболевания обычно возникают без предварительного предупреждения и часто приводят к летальному исходу, прежде чем можно будет предложить медицинское лечение. Важно диагностировать пациента на ранней стадии, чтобы действия против заболевания могли быть предприняты заранее.

Основными симптомами сердечных заболеваний являются дискомфорт в груди, одышка, усталость, отек, сердцебиение и обмороки. Другие более мелкие признаки включают кашель, кашель крови, хромоту, боль в конечностях, расстройство желудка, отрывка и обесцвечивание кожи. В большинстве случаев внимание к характеристикам дискомфорта в грудной клетке может сузить диагноз [2].

Эхокардиография является одной из наиболее широко используемых диагностических средств в сердечно-сосудистых заболеваниях. Он обеспечивает полной клинически значимой информацией. Помимо наиболее распространенных методов диагностики и эхокардиографии, существуют другие методы, такие как электрокардиография, рентгенография грудной клетки, магнитно-резонансная томография, ангиография и интракоронаризм и компьютерная томография.

В экспериментальной работе будет использоваться среда для решения задач машинного обучения и интеллектуального анализа данных RapidMiner, версия 7.4. RapidMiner - это инструмент с открытым исходным кодом, предназначенный для выполнения интеллектуального анализа данных и машинного обучения, включая: загрузку и преобразование данных, предварительную обработку и визуализацию данных, моделирование и оценку.

Представим информацию об атрибутах: 1. **Age**: возраст в годах; 2. **Sex**: пол (1 = мужчина; 0 = женский); 3. **Ср**: тип грудной боли: значение 1: типичная стенокардия; значение 2: атипичная стенокардия; значение 3: боль, не относящая к стенокардии; значение 4: бессимптомная боль; 4. **Trestbps**: кровяное давление (в мм ртутного столба, при поступлении в больницу); 5. **Chol**: холестерин в мг/дл; 6. **Fbs**: (уровень сахара в крови > 120 мг / дл, 1 = истина; 0 = ложь); 7. **Restecg**: электрокардиографические результаты: значение 0: нормальный; значение 1: аномальные изменения ST-T волны (волновые инверсий T и / или ST повышение или понижение > 0,05 мВ); значение 2: показывает вероятную или явную гипертрофию левого желудочка по критериям Эстеса; 8. **Thalach**: максимальная частота сердечных сокращений; 9. **Exang**: нагрузка, стимулирующая стенокардию (1 = да; 0 = нет); 10. **Oldpeak**: ST депрессии, стимулируемые нагрузкой; 11. **Slope**: подъем (элевация) сегмента ST; 12. **Ca**: число крупных сосудов (0-3) окрашенных по рентгеноскопии; 13. **Thal**: стресс тест с таллием. 3 = норма; 6 = фиксированный дефект; 7 = обратимый дефект; 14. **Num** – прогнозируемый атрибут (наличие/отсутствие болезнь сердца).

В случае если миокард испытывает значительный или критический дефицит кислорода возникает каскад биохимических изменений, вследствие чего появляются определенные изменения на ЭКГ — элевация или депрессия сегмента ST.

На ЭКГ можно выделить 5 зубцов: P, Q, R, S, T. Иногда можно увидеть малозаметную волну U. Зубец P отображает процесс деполяризации миокарда предсердий, комплекс QRS — деполяризации желудочков, сегмент ST и зубец T отражают процессы реполяризации миокарда желудочков.

Стресс-тест с таллием используется для того, чтобы определить, какой объем крови поступает в сердце и как он изменяется в зависимости от физической нагрузки. Также он используется при контроле уровня напряжения у пациентов, перенесших сердечный приступ и при определении причин таких симптомов, как боль в груди и одышка. Иногда этот тест проводят после операции для оценки ее эффективности. Он поможет установить, насколько заблокирован кровоток в коронарных артериях.

Во время этого теста, пациент ходит по беговой дорожке до тех пор, пока нагрузка не станет максимальной. После этого, пациенту в вену вводят таллий и, с помощью гамма-камеры, наблюдают за движением крови к сердцу. Если имеется нарушение кровотока, (что происходит при ишемической болезни), скintiграмма (изображение сердца) покажет очаги, в которых снижено накопление таллия. Это и будет признаком болезни [3].

Для выявления степени связности, и отбора факторов, оказывающих наиболее существенное влияние на целевой признак, построена корреляционная матрица. На основе полученных результатов, сделан вывод, что между атрибутами нет корреляционной связи, таким образом, можно предположить, что исследования медицинских данных носят комплексный характер. В этом случае изучается не одна характеристика исследуемого объекта, а целая совокупность показателей.

Реализован алгоритм классификации – дерева принятия решений. Идея метода состоит в том, чтобы создать набор правил, которые могут предсказать конкретную переменную – класс признака болезни сердца на основе данных.

Оценка эффективности модели классификации основана на сумме правильно и неправильно предсказанных моделей [4]. Эти суммы подсчитываются в таблице, известной как матрица ошибок (Таблица 1).

Таблица 1. Матрица ошибок для задачи бинарной классификации

		Predicted Class	
		Class=1	Class=0
Actual Class	Class=1	f_{11}	f_{10}
	Class=0	f_{01}	f_{00}

Каждая запись f_{ij} в этой таблице обозначает количество записей класса i , были распознаны как объекты класса j . Например, f_{01} - это число записей из класса 0, некорректно прогнозируемых как класс 1. На основе записей в матрице ошибок общее число правильных предсказаний, сделанных моделью, равно $(f_{11}+f_{00})$, а общее количество неправильных предсказаний равно $(f_{10}+f_{01})$. Хотя матрица ошибок содержит информацию, необходимую для определения насколько хорошо работает модель классификация, обобщая эту информацию одним числом будет удобнее сравнивать производительность различных моделей. Это можно сделать, используя показатель Точность, которая определяется следующим образом:

$$\text{Accuracy} = \frac{\text{Количество правильных прогнозов}}{\text{Общее количество прогнозов}} = \frac{f_{11}+f_{00}}{f_{11}+f_{10}+f_{01}+f_{00}}$$

Эквивалентно, производительность модели может быть выражена в коэффициенте ошибок, которая определяется следующим уравнением:

$$\text{Classification error} = \frac{\text{Количество неправильных прогнозов}}{\text{Общее количество прогнозов}} = \frac{f_{10}+f_{01}}{f_{11}+f_{10}+f_{01}+f_{00}}$$

Большинство алгоритмов классификации ищут модели с максимальной точностью, или с наименьшим коэффициентом ошибки.

С помощью оператора дерева решений Decision Tree сформирована схема принятия решений для классификации данных. Цель данного оператора – создать модель классификации, которая будет прогнозировать значения целевого атрибута, на основании значений входных атрибутов. На входе оператору подается учебный набор данных, который будет использоваться для обучения. На выходе оператора получаем обученную модель. В процессе тестирования и оценки эффективности алгоритма классификации использованы операторы X-Validation и Performance Classification.

Существует множество критериев, которые могут быть использованы, чтобы определить наилучший способ разбиения множества данных [5]:

- Information gain. Выбор атрибута основан на подсчете энтропии каждого входного атрибута. Затем энтропия добавляется пропорционально, чтобы получить полную энтропию для разбиения. Результирующая энтропия вычитается из энтропии до разбиения. Разделение происходит по атрибуту с минимальной энтропией.

$$Gain(T,X)=Entropy(T)-Entropy(T,X)$$

- Индекс Gini. При использовании критерия атрибут выбирается на основании расстояний между распределениями классов. Если дано множество T , включающее примеры из n классов, индекс Gini определяется по формуле:

$$Gini(T)=1-(p_1^2+p_2^2+\dots+p_n^2)$$

где T - текущий узел, p_j - вероятность класса j в узле T , n - количество классов.

- Assurasy. При выборе данного критерия, выбирается такой атрибут, который максимизирует точность для всего генерируемого дерева.

- Gain ratio. Критерий использует идею энтропии и использует значения прироста информации. Прирост информации — это разница между энтропией множества до разбиения и после разбиения относительно параметра.

$$Gain\ ratio=\Delta_{info}/Split\ info$$

Выбран критерий Information gain, так как при данном критерии была достигнута наибольшая точность модели.

При построении дерева решений наиболее информативными атрибутами оказались Стресс-тест с таллием, Число крупных сосудов, окрашенных по рентгеноскопии, Тип грудной боли.

В процессе тестирования и оценки эффективности алгоритма классификации использованы операторы X-Validation и Performance Classification. Одним из выходных данных операторов является Assurasy - количество правильно классифицированных объектов, т.е. процент правильных прогнозов. Assurasy построенной модели составляет 72,5%.

В ходе проделанной работы были сделаны выводы о том, что в наборе данных между атрибутами нет корреляционной связи, и исследуемые медицинские данные должны быть проанализированы методами интеллектуального анализа данных. При построении дерева решений наиболее информативными атрибутами оказались Стресс-тест с таллием, Число крупных сосудов, окрашенных по рентгеноскопии, Тип грудной боли. Для диагностирования признака болезни сердца по входным данным разработано Java приложение, интегрированное с обученной RapidMiner моделью.

Список использованных источников

1 Machine learning repository. <https://archive.ics.uci.edu/ml/datasets/Heart+Disease>

2 Основы кардиологии. Принципы и практика. Под редакцией проф.Клива Розендорффа. Humana Press Inc., 2005. - 113 с.

3 Медицинские обследования, анализы и тесты. Типы кардиологических стресс-тестов. Источник - <http://doctor-test.ru/kardiologicheskie-stress-testy>.

4 Introduction to Data Mining.P.Tan, M.Steinbach, V. Kumar., 2006. – p 145-165.

5 Анализ данных и процессов / А.А.Барсегян, М.С. Куприянов, И.И. Холод, М.Д. Тесс, С.И. Елизаров. – 3-е изд. перераб. и доп. – СПб.: БХВ-Петербург, 2009. - 512 с.

ФОРМИРОВАНИЕ ВОПРОСОВ И ОТВЕТОВ НА ОСНОВЕ БАЗЫ ЗНАНИЙ

(ЕНУ им. Л.Н. Гумилева, Астана, Казахстан)

Введение

В настоящее время во всем мире интенсивно применяется электронное обучение. Количество студентов электронного образования увеличивается в каждом году. Что касается Казахстана, то сектор электронного образования почти отсутствует. Имеются только системы, автоматизирующие некоторые сопутствующие процессы традиционного образования, которые никак не относятся к электронному образованию. Казахстан, интегрируясь в мировое сообщество, должен обеспечивать достаточный (по отношению к мировому уровню) уровень развития электронного образования.

Одним из основных направлений Государственной программы развития образования Республики Казахстан на 2011-2020 годы является электронное обучение «e-learning». В программе реализации данного направления сформулирована следующая задача: по каждому предмету, изучаемому в основной и профильной школе, будут разработаны интерактивные и интеллектуальные цифровые образовательные ресурсы [1].

Разработка базы знаний

В связи с этим, в настоящее время в научно-исследовательском институте «Искусственный интеллект» Евразийского национального университета им.Л.Н.Гумилева ведутся исследования в данной области, разрабатываются базы знаний по различным предметным областям, а также системы контроля и обучения, основанные на базе знаний [2].

В ходе исследований разработана онтология по Алгебре в Protege. Онтограф показан на рисунке 1 (<http://e-zerde.kz/zerde/Algebra/OntologAlgebra.jpg>).

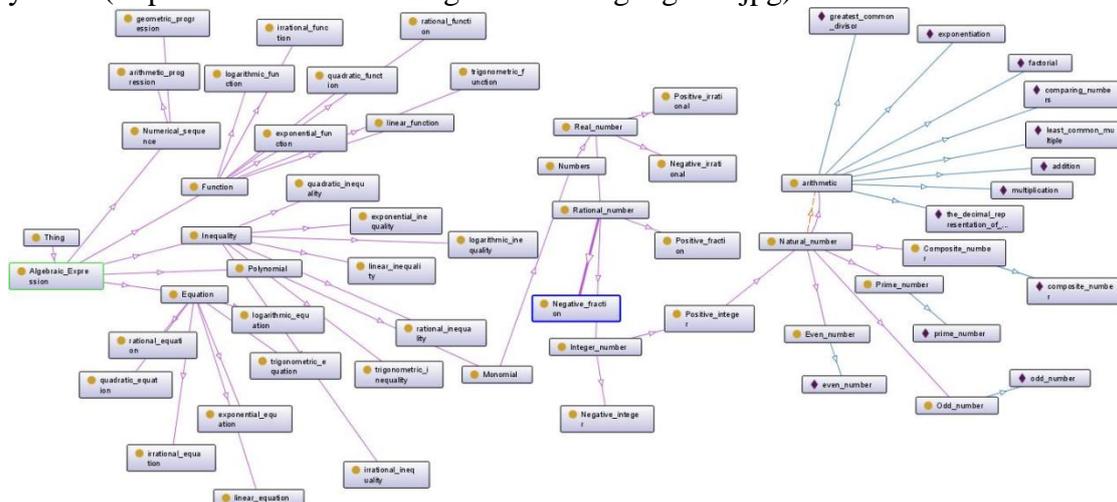


Рисунок 1 – Онтограф базы знаний по алгебре

После проектирования онтологии, файл была проверена на валидность с помощью веб-службы «RDF Validator» на сайте «www.w3.org». Система выдает список триплов и граф модели данных. Сгенерированный граф после успешной валидации документа можно просмотреть по ссылке <http://e-zerde.kz/zerde/Algebra/ValidAlgebra.png>.

Для демонстрации возьмем часть базы знаний Действительные числа, граф которой показан на рисунке 2 (<http://e-zerde.kz/zerde/Algebra/ValidNumbers.png>)

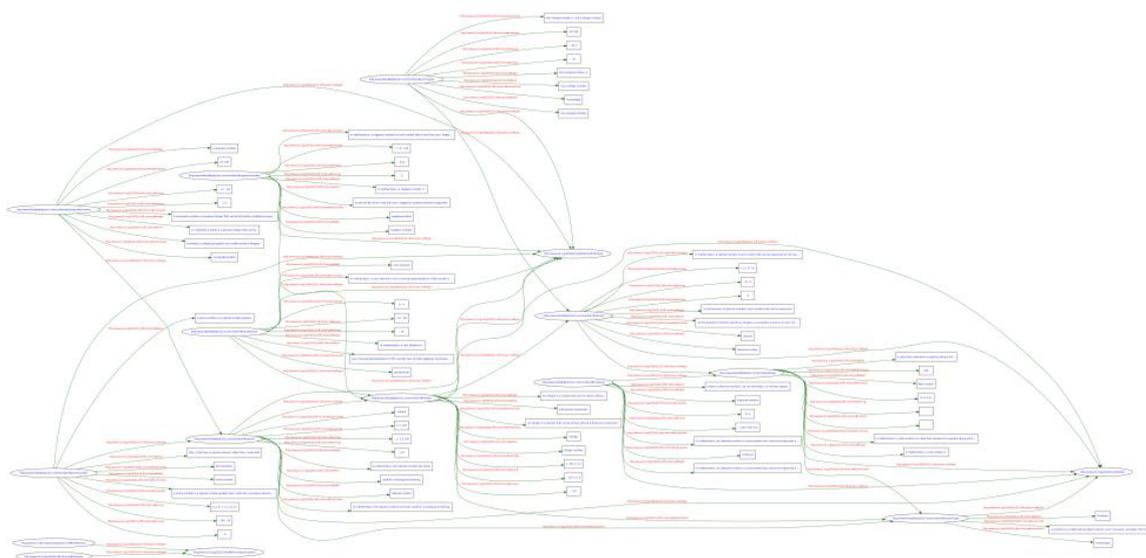


Рисунок 2 – Онтограф базы знаний Действительные числа

Для применения разработанных баз знаний по различным предметам в учебном процессе нами разработана технология для автоматической генерации и экспорта в Excel контрольных вопросов и ответов.

Разработка контрольных вопросов для оценивания знания учащихся требует серьезного подхода. Вопросы не должны быть двусмысленными или неясными для обучающихся. При разработке контрольных вопросов были исследованы различные методики. Особую известность в мире получила таксономия, разработанная коллективом американских ученых под руководством Б.С. Блума. В процессе приобретения знаний таксономия Б.С.Блума выделяет шесть уровней, на которых осуществляются отдельные познавательные процессы. К таксономическим категориям относятся следующие: знания (информация); понимание; применение; анализ; синтез; оценка материалов и методов с учетом принятых целей. Самые примитивные из навыков являются знание и понимание, а самые продвинутые: применение; анализ; синтез; оценка. В процессе разработки вопросов было решено генерировать задания на оценку знаний и понимания, так как для формирования вопросов на применение, синтез, анализ и оценку необходима система умеющая делать рассуждения в полной мере. Ниже приведены примеры вопросов разработанной системы для уровня «Знание» и «Понимание» [3,4].

Познавательный процесс	Примеры вопросов разработанной системы
Знание	
Воспроизведение	– Дайте определение термину.
Припоминание, узнавание	– Какие подклассы содержит класс. – К какому классу относятся подклассы.
Понимание	
Интерпретация	– Продолжите определение.
Приведение примеров	– Приведите примеры.
Классификация	– Найдите лишний элемент из ряда.
Сравнение	– К какому классу относятся элементы. – Найдите элементы, которые не относятся к классу.

Реализация программы начинается с установки IDLE Python 3.5 и библиотеки «rdflib», для разработки графического интерфейса импортируется в код библиотека «tkinter», для экспорта в Excel - «xlwt».

На рисунке 3 показана стартовая форма программы.

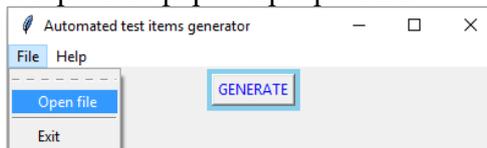


Рисунок 3- Стартовая страница программы

Система позволяет открывать любые файлы типа rdf и owl. Необходимо выбрать файл и нажать кнопку “Generate”. В текущей директории создается файл example.xls со сгенерированными вопросами. На основе 10 классов сформировано 66 корректных тестовых заданий с правильными ответами. Загрузить файл можно по ссылке <http://e-zerde.kz/zerde/Algebra/example.xls>.

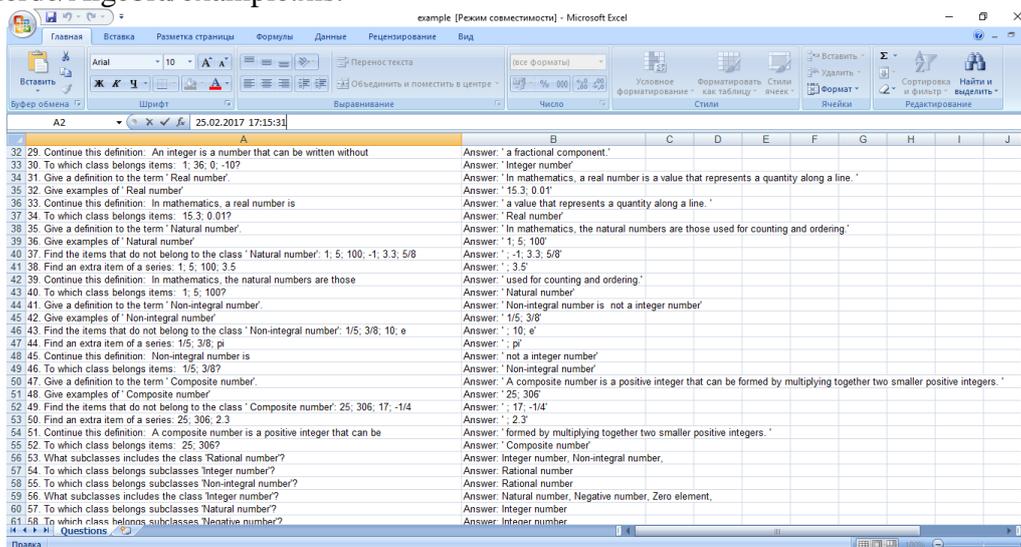


Рисунок 4 – Сгенерированные тестовые вопросы с ответами

Заключение

Подводя итоги можно сделать выводы относительно актуальности и тенденции развития данного направления. На сегодняшний день во всем мире ведется большая работа в направлении искусственного интеллекта и интеллектуальных систем на основе знания. Проведенные исследования в данной работе, дают новые идеи и открывают новые возможности интеллектуальных систем на основе баз знания для дальнейшей работы и изучения.

Таким образом, нами разработана база знаний по алгебре в редакторе онтологий Protege. А также разработано программное приложение с использованием IDLE Python для автоматической генерации тестовых вопросов открытого типа на проверку учащегося на знание и понимание предметной области из базы. Данная технология позволит автоматически генерировать тестовые вопросы из любой базы знаний по различным предметным областям. Данная система позволит преподавателям автоматизировать процесс формирования корректных тестовых вопросов и ответов по различным дисциплинам. Система для генерации заданий из базы знаний является одним из маленьких шагов автоматизации процесса оценивания учащихся. Дальнейшие разработки направлены на обработку ответа обучающегося и оценки правильности ответа.

Использованные ссылки:

1 Omarbekova, A; Sharipbay, A; Barlybaev, B. Generation of Test Questions from RDF Files Using PYTHON and SPARQL. Journal of Physics: Conference Series.-2017. –Vol.806, 1.

2 Liu, BL; Hu, B; Wang, XY; Zheng, L. A RDF-based knowledgebase implementation for E-learning. New Horizon in Web-based Learning. 3rd International Conference on Web-Based Learning (ICWL 2004). P.: 45-50. 2004

К.О. ОРАЗБАЕВА, Б.Е. ӨТЕНОВА, Б.Б. ОРАЗБАЕВ

**МҰНАЙ ӨНДЕУ ӨНДІРІСІНІҢ ТЕХНОЛОГИЯЛЫҚ ЖҮЙЕЛЕРІН МОДЕЛЬДЕУ
ЖӘНЕ ОПТИМИЗАЦИЯЛАУ АРҚЫЛЫ ТИІМДІ ШЕШІМ ҚАБЫЛДАУ
КОМПЬЮТЕРЛІК ЖҮЙЕСІ**

(Атырау мұнай және газ институты, Атырау қ.)

Өндірісте шешім қабылдайтын тұлға (технолог) дұрыс шешім қабылдау үшін ақпараттың мол көлемін өңдеу, көптеген баламаларды қарастыру, түрлі (экономикалық, технологиялық, экологиялық) факторлардың әсерін ескеру, айқынсыздық жағдайда шешімнің салдарын бағалау қажет болатын жағдайға жиі ұшырайды. Мұндай жағдай мұнай өңдеу технологиялық кешендері сияқты көпкритерийлі жүйелерді тиімді басқару т.с.с өндірістік есептерін шешу қажет болғанда туындайды.

Мұндай есептерді шешу үшін компьютерлік модельдеу және оптимизациялау арқылы тиімділеу жүйелері (КТЖ) өте пайдалы. Мұндай жүйелер модельдеу, оптимизациялау тәсілдерін және қазіргі заманғы компьютерлік техниканың мүмкіндігін біріктіреді, бұл оптимизациялау, тиімділеу процедура-сын анағұрлым жақсартып, жеделдетеді. Компьютерлік тиімділеу жүйесі құрамына мынадай негізгі блоктар кіреді: оптимизациялау алгоритмдерінің кешені, технологиялық объектілердің математикалық модельдер кешені, білім және мәліметтер базасы, шешімді түсіндіру блогы, модельді идентификациялау блогы, пайдаланушы интерфейсі. Бұл блоктар ақпараттық ағымдармен байланысты, олардың әрбіреуі белгілі бір қызмет атқарады [1].

Мұнай өңдеу және мұнай химиясы өндірістік объектілерінің көпшілігінің негізгі ерекшелігі бастапқы ақпараттың айқынсыздығы болып табылады. Бұл жағдайда, атап өткеніміздей, сапалы сипаты бар ШҚТ, сарапшы мамандардың білімі мен пікірін негізге алуға тура келеді. Оптимизациялаудың мұндай айқын емес есебін шешу және шешім қабылдау үшін КТЖ құрамына, онымен табиғи немесе кәсіби тілде байланысуға мүмкіндік беретін, зерделеу (интеллектуал) элементтерін енгізу қажет. Бұл мүмкіндіктерге компьютерлік жүйе құрамына білім базасын, логикалық қорыту блогы және нәтижелерді түсіндіру, көпкритерийлі айқын емес оптимизациялау (оптималды басқару) есебін шешу алгоритмдері және зерде интерфейсін қосу арқылы жасанды зерде тәсілдері [2] негізінде қол жетеді.

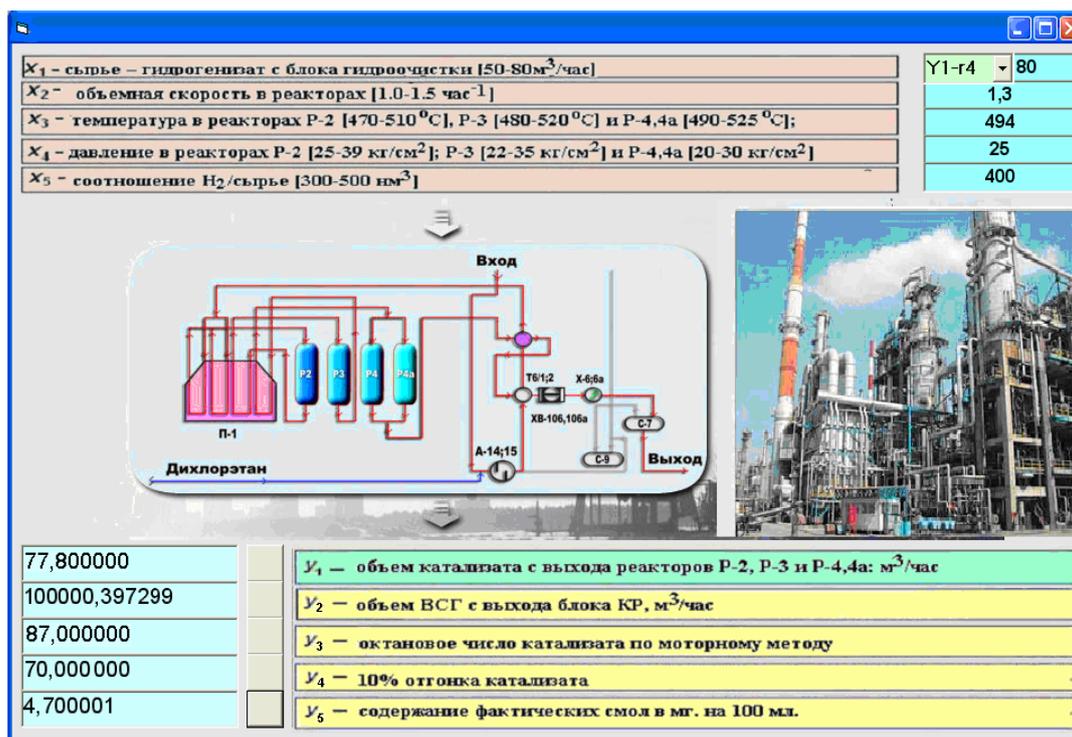
Мұнай өңдеу өндірісінің технологиялық кешендерінің КТЖ-ін құруды мынадай негізгі кезеңдер бойынша жүргізуге болады:

1. Проблемалық аумақты және шешілетін есепті идентификациялау, тиімділеу, оптимизациялау есебінің мазмұнды қойылымы;
2. ШҚТ және сарапшы мамандардың объект, процесс және мәселе жөніндегі білімін қалыптастыру (формализациялау);
3. Білім және мәліметтер базасын құру;
4. Объект моделдер кешенін жасақтау;
5. Оптимизациялау мен тиімді шешім қабылдау есебін алгоритмдеу;
6. Пайдаланушының интеллектуалды интерфейсін жасақтау;
7. Жасақталған модельдер мен алгоритмдерді программалық жүзеге асыру.

Интеллектуалды КТЖ тиімділігі білімнің қалыптасу сапасымен, жасақталған математикалық модельдер мен оптимизациялау, тиімді шешім қабылдау есептерін шешу алгоритмдерінің сапасымен, сондай-ақ, пайдаланушы интерфейсін қолайлылығымен анықталады.

Өндірісті модельдеу және оптимизациялау арқылы тиімді шешім қабылдау есептері көп жағдайларда айқынсыздық жағдайда қалыптасып, шешіледі. Мұндай есептерді шешу үшін негізінен адамнан айқын емес пікірлері, ой-тұжырымдары және айтылған сөздері түріндегі ақпараттар пайдаланылады. Осыған байланысты қарастырылған есептерді шешу тиімділігі қомақты деңгейде айқын емес модельдерді құру тәсіліне және айқын емес ортада көпкритерийлі оптимизациялау есебін шешу алгоритмдеріне байланысты болады [3, 4].

Құрылған модельдерді өндірістік кешеннің тиімді режимін табу үшін қолдану мәселелерін, яғни программа түріндегі жүзеге асыру нәтижелері мен риформинг блогы мен бензол өндіру кешендері үшін жасақталған интерфейсті сипаттайық. «Риформинг процесін жүйелі модельдеу» менюін таңдағанда (1-сурет) процессті тікелей модельдеу жүргізіледі.



1-сурет – Риформинг процесін жүйелі модельдеу режимі

1 - суреттен келтірілгендей Интерфейсте модельдеу режимінде экранның жоғары жағында риформинг процесінің негізгі кіріс-режимдік параметрлері (x_1, x_2, x_3, x_4, x_5) келтірілген, оларды өзгерте отырып, модельдеу арқылы процесстің шығыс параметрлерін риформинг ректорлары шығысындағы өнім – катализаттың көлемін (y_1), шығыстағы СКГ көлемін (y_3), және катализаттың сапа көрсеткіштерін (y_4 – октандық саны, y_5 – 10% - айдау, y_8 – шайырдың құрамы) бақылай аламыз, яғни олардың оптималды мәндерін іздеуге болады. Менюде әр кіріс - режимдік параметрлердің өзгеру интервалдары мен өлшем бірліктері келтірілген. Модельденетін риформинг реакторларын (P-2, P-3, P-4,4a) таңдау үшін арнаулы терезе орналасқан. Әр x_1, x_2, x_3, x_4, x_5 параметрлерін өзгерту үшін, олардың оң жақтарында арнаулы терезелер келтірілген.

Экранның төменгі жағында модельдеу нәтижелері – процесстің шығыс параметрлері, жоғарыда аталған y векторы, яғни өнімдердің көлемі мен сапа көрсеткіштері шығады. Модельдеу кезінде, яғни кіріс параметрлері (x) өзгертілген соң, шығыс параметрлерінің жаңа мәндерін (y) көру үшін арнаулы әр y параметрінің тұсында орналасқан арнаулы пернені () басу қажет.

Келтірілген суретте модельдеу арқылы риформинг блогының оптималды мәндерін анықтау нәтижесі келтірілген. Бұл процессті тиімділеу үшін режимді оптимизациялау алгоритмдерін (оптимсизациялау режимін) қолдану қажет.

Жұмыста алынған зерттеу нәтижелерінің тәжірибелік-өнеркәсіптік сынақ қорытындылары құрылған математикалық модельдер мен оптимизациялау алгоритмдердің дұрыстығын растады. Мәселен, каталитикалық риформинг қондырғысының оптималды жұмыс режимдерін ұсынылған модельдер мен алгоритмдер көмегімен анықтау арқылы қабылданатын тиімді шешімнен түсетін экономикалық тиімділік 5 миллионнан астам теңге құрайды. Бұл көрсеткіштерге талап етілген сапалы мақсаттық өнімнің шығуын арттыру жолымен және жоспарлау мен басқару есептерін өндірістегі қалыптасқан жағдайға байланысты жедел шешумен қол жеткізіледі.

Жасақталған КТЖ құрылымы болашақта математикалық қамтамасыз-дандыруды, жаңа функциялар мен жүйе мүмкіндігін қоса отырып, оңай кеңейтуге мүмкіндік береді. Бұдан басқа ұсынылып отырған КТЖ құрылымы тиісті аппаратты-техникалық (объектімен байланыс құралы) және қосымша арнаулы программалары болғанда басқару шеңберін тұйықтауға мүмкіндік береді, яғни, компьютер объектіні тікелей басқаратын болады. Айқынсыздықта модельдерді құрудың және көпкритерийлі оптимизациялау есептерін шешудің ұсынылған әдістемелері мен алгоритмдері өнеркәсіптің түрлі салаларында (мұнай химиясында, мұнай-газ, металлургия және т.б.) өндірістік мәселелерді шешуде тиімді қолданыла алады.

Таяу жылдарда пайдаланушы мен компьютер арасындағы қатынас жүйесі жаңа сапалы деңгейге ауысуы тиіс. Клавиатурамен жұмыс істеуді талап ететін мәтіндік қатынастың орнына дыбыстық қатынасты орнату жобалары зерттеліп, іске асырылуда. Мұнда пайдаланушы өзіне керекті ақпаратты дауысымен енгізіп, хабарлама жүйесінен де осындай түрде ақпарат ала алатын болады. Адам үшін үйреншікті сөйлеу формасы КТЖ-мен байланысудың анағұрлым қолайлылығын арттырып, осындай жүйе тиімділігін әжептәуір көтереді [5].

Осылайша, бұл жұмыста каталитикалық риформинг технологиялық қондырғысы мысалында мұнай өңдеу жүйелерінің математикалық моделдеу және оптимизациялау арқылы оларды басқару үшін тиімді шешім қабылдаудың жаңа перспективті тәсілдері құрылған. Осы нәтижелер базасында интеллектуалды компьютерлік модельдеу мен оптимизациялау жүйесінің математикалық қамтамасыздандыру және басқа компоненттері жасақталады. Атырау МӨЗ-ның жағдайында технологиялық кешендердің тиімділігін арттыру мәселелері шешіліп, өндірісті экологизациялау мақсатында тиімді техникалық және технологиялық шешімдер зерттеліп, ұсынылған.

Қорытынды: Мұнай өңдеу технологиялық жүйелерін модельдеу және оптимизациялау арқылы тиімді шешім қабылдау компьютерлік жүйесінің құрылымы мен функционалды блоктары қарастырылған. Ұсынылған КТЖ-нің осыған ұқсас жүйелерден артықшылығы: оның құрамына өзара байланысқан мұнай өңдеу және мұнай химиясы өндірістерінің технологиялық объектілерін математикалық модельдеу және көпкритерийлі оптимизациялау мәселелерін айқын емес ортада да тиімді шешетін алгоритмдер кешені енгізілген және зерделі интерфейс бар.

Қолданылған ақпарат көздері

1 Бигалиева М.Ж., Оразбаев Б.Б., Сериков Ф.Т. Информационные системы и компьютерное моделирование. -Алматы: Ғылым: 2003. -182 с.

2 Поспелов Г.С., Поспелов Д.А. Искусственный интеллект – прикладные системы. -М.: Знание, 1985. - 195 с.

3 Сериков Ф.Т., Оразбаева К.Н., Гайсина С.Б. Применение математических методов и компьютерной технологии для эффективного решения экономико-экологических задач нефтегазового производства //Труды 5 Казахстанско-Российской международной научно-практической конференции «Мат. моделирование научно-технолог. и экологических проблем в нефтедобывающей промышленности». -Алматы-Атырау: 2005. -С. 257-263.

4 Оразбаев Б.Б. Методы моделирования и принятия решений для управления производством в нечеткой среде. ЕНУ им. Л.Н.Гумилева, -Астана: 2016, 398 с.

5 Оразбаев Б.Б. Интеллектуальные системы принятия решений для управления технологическими объектами при дефиците информации //ИТЖ Автоматизация, телемеханизация, и связь в нефтяной промышленности. -М.: 1994, №6-7. -С.12-13.

Р.К. УСКЕНБАЕВА, А.А. КУАНДЫКОВ

**ОРГАНИЗАЦИЯ И УПРАВЛЕНИЯ ПРОЦЕССОВ АВТОМАТИЗАЦИИ
ВЫПОЛНЕНИЯ СЛОЖНЫХ ОПЕРАЦИЙ СОВОКУПНОСТЬЮ
ИНТЕЛЛЕКТУАЛЬНЫХ РОБОТОВ**

(Международный Университет Информационных Технологий, Алматы, Казахстан)

1. Введение

Процесс решения и выполнения некоторых сложных задач сводится к совокупностям сложных операций или подпроцессов. Эти операций или подпроцессы:

- во-первых обладают такими свойствами, как стационарность и подвижность;
- во-вторых еще могут обладать такими свойствами, как когнитивность и физичность (причем механический-подвижными свойствами, как в микропространстве(рабочее пространство с одной позицией платформы и манипулятора) или в макропространстве(в разных позициях платформы и манипулятора)), умственно-рутинность и физическо-рутинность.

Кроме того, пространство задач состоит из таких компонентов (в том числе систем) инфраструктуры, что для решения и выполнения таких сложных задач функциональность одиноких систем могут быть недостаточным. Поэтому решения и выполнения таких задач могут быть достигнуты при объединении функциональность нескольких систем [1, 2]. Причем эти системы могут быть как стационарными, так и подвижными, а архитектура сильно-связанными и/или слабосвязанными. Функциональность слабосвязанных систем формируется из объединение функциональности сильно-связанных систем так, чтобы было достаточно для решения и выполнения сложных задач заданного пространство.

Объединение функциональности различных систем можно увеличить путем специальной организации (т.е. координации, согласования) по определенному методу. От метода организации сильно-связанных систем, зависит уровень объединенной функциональность слабосвязанных систем, которой должна быть достаточной для автоматизированного выполнения задач и операций.

Отмети, что в качестве сильно-связанных систем выступают интеллектуальные роботы (ИР) на уровне модулей.

С другой стороны уменьшение уровня сложности задач достигается специальным организационным методом – декомпозицией и созданием бизнес-процессов и еще раз декомпозицией операций бизнес-процесса.

В целом, работа посвящена исследования вопросов теории ИР, которые представлены из модулей программных или интеллектуальных агентов и модулей механотроники, в том числе из механических устройств.

Изучение вопросов участие в процессе автоматизации сложных задач, исходя из психофизиологических особенности человека, в работе сознательно опускается, т.к. эти вопросы требует отдельных исследований.

2. Постановка класса задач и процессов выполнения операций

Формулировка задачи, которое формализует выполняемые действия в задании с указанием ограничения, осуществляется исходя из условий, поставленных перед системой, а также исходя из инфраструктуры.

Например, задания ИР могут быть следующего утверждения (которое состоит из совокупности признаков):

«Используя предметы, произвольным образом, находящиеся в районе «А» построит объект по заданному проекту в районе «В»».

Для данного задания, постановка задачи осуществим при заданном пространстве, по которой составляется инфраструктура: <задание, инфраструктура> → <задача>.

Инфраструктура. Инфраструктура для задачи, которую формулируем, имеет следующий вид.

Пусть задано пространство, трехмерное Евклидовое (горизонталь и вертикаль) пространство Ω . В зависимости от характера решаемой задачи, пространство представляется различным образом, например, нужны только горизонтальные размеры предметов или еще вертикальные размеры предметов.

Пространство Ω может разбиваться со стороны роботов таким образом. Пространство Ω для i -го робота, состоит из Ω^{Mi} – микропространство или рабочем пространстве робота (пространстве вытянутой манипулятор-руки робота) i -го робота, а также Ω^{MAi} – макропространство для i -го робота – это пространство деятельности данного робота. Полное пространство Ω^i для i -го робота. Для робота R_i , которые выполняют задания Z_d пространство выполнения Ω^i может совпадать с полным пространством Ω всего робота [3].

Если роботы в пространстве Ω являются подвижными, то для его микро-пространство Ω^{Mi} задается окрестностью предметов, с которыми робот работает.

В частности, для нашей задачи пространство разбиты на ячейки в зависимости от необходимости для решения задачи. Причем ячейки могут объединяться (группироваться) в районы А, В, С, и т.д.

Пусть, координаты ячейки (например, в ячейках $Esh(X, Y)$) района А пространство Ω (т.е. $Esh(X, Y) \in A \in \Omega$), задается/характеризуется значением координат, где X_0 и Y_0 декартовы координаты центра ячейки, района. Любой район (например, А) может включать от одной ячейки до несколько.

В ячейке (например, в районе А) Пространство Ω может иметь конечное множество предметов или элементов $Pr = \{Pr_i\}$, $i = 1, I$, где Pr_i - i -ой предмет I ограничен ($Pr = \{Pr_i\}$, $i = 1, I$, где Pr_i - i -ой предмет), Pr – множество объектов или предметов.

Каждый предмет (объект или элемент) имеет известные характеристики и характеризуется признаками объекта $Pr_i \in Pr: \chi(Pr_i) = \{\chi_{ij}\}$, $i = 1, I, j = 1, J_i$.

Характеристики для предметов $Pr_i \in Pr: \chi(Pr_i) = \{\chi_{ij}\}$, $i = 1, I; j = 1, J_i$, и для предмета $Pr_k \in Pr: \chi(Pr_k) = \{\chi_{kh}\}$, $k = 1, K = I; h = 1, H_k$ таковы, что выполняется: $\exists \chi_{ij} \neq \chi_{kh}$, где χ_{ij} - j -ая характеристика i – предмета Pr_i и χ_{kh} - h -ая характеристика k – предмета Pr_k .

Эти объекты или предметы могут быть как в микропространстве $\forall \omega_i \subseteq \Omega$ (т.е. рабочем пространстве робота или $\Omega = \cup \omega_i$) робота, так и макропространстве пространства деятельности роботов Ω .

Причем, надо отметить, что значения характеристик этих предметов может оставаться постоянным на длительное время $t > \Delta T$.

Отметим, что формализация предметов, элементов или участки пространство могут задаваться пассфреймами, посфреймами и ортофреймами, унифреймами, описания которых в работе сознательно опускаем.

Если Роботы в пространстве Ω является подвижными, то для них микро-пространство Ω^{Mi} задается окрестностью предметов.

В задаче требуется выполнить следующее.

Пусть, из заданных предметов с определенными признаками Pr , $Pr = \{Pr_i\}$, $i = 1, I'$; $I' \leq I$ следует построить роботам проект (чертеж) объекта PrO из имеющихся предметов.

Требуется в ячейках (Esh) района В, пространства Ω построит объект согласно проекту PrO из предметов $Pr' = \{Pr_i\}$, $i = 1, I'$; $I' \leq I$.

Требуется взять предмет $A(x, y)$ с адреса (x_1, y_1) , перенести по адресу $B(x, y)$ с адресом (x_2, y_2) с помощью одной или несколькими рук $\{P_1, P_2, P_3, \dots, P_i, \dots, P_n\}$. Характеристики рук, например, $P_i: \chi(P_i) = \{\chi_{ij}\}$, $J = 1, J_i$ зададим позже.

Требуется сначала построить алгоритм построения объекта из исходных объектов или предметов заданных районах A и других в Ω .

3. Основные особенности процесса решения задачи такого класса

Анализ показывает, что особенностями исследуемого класса задач следующие:

Процессе решения задач состоит из операций (независимо от субъект и/или системы выполнения) имеющие определенные сложности. Эти операции могут быть и стационарными и подвижными. Причем эти операции могут быть: когнитивными, механическими и рутинно-умственными и рутинно-физическими операций.

Кроме того имеется такие задачи или их операций, в которых свойства когнитивность и физичность (микроопераций) в их составе не делимыми при существующей инфраструктуре. Причем, для автоматизаций и управления процессов выполнения этих операций функциональность одиночных систем недостаточны. Поэтому для автоматизации процессов выполнения таких операций возникает необходимость привлечение нескольких систем (т.е. ИР).

В связи с этим, предлагаем общий подход к решению этих проблем, для которых требуется метод и средства, позволяющие:

- уменьшить сложности автоматизируемых операции и
- координировать и согласовать поведение и работы ИР, привлеченных для автоматизации выполнения операций задач между собой для того, чтобы максимизировать эффективность их суммарной функциональности.

4. Принципы автоматизации сложных задач с помощью ИР

Принцип упрощения и приспособление к процессам автоматизации выполнения этих классов задач, должна осуществляться с двух сторон: как со стороны автоматизируемых операций; так и со стороны исполнителей.

Со стороны задач или операций, которые должны быть автоматизированы:

- 1) сложные задачи для автоматизации выполнения могут быть представлены (для автоматического выполнения) в виде ряда простых задач при имеющей инфраструктуре;
- 2) простые задачи должны быть организованы между собой определенным образом, так чтобы сложность этих простых задач стали упрощеннее и приспособлены к автоматизированному выполнению.

Со стороны исполнителей автоматизация выполнения задач или операций описываются:

- 1) Некоторые операции после упрощения и организации могут быть сложными, что функциональность одиночных сильно-связанных систем недостаточна (в частности ИР) для выполнения сложных операций. Поэтому для автоматизации процессов выполнения этих операций, требуется привлечение нескольких систем (сильно-связанных), суммарная функциональность которых была достаточна для автоматизации процессов выполнения этих операций.

- 2) Привлечение нескольких систем (сильно-связанных), обеспечивающие автоматизацию задач необходимо организовать определенным методом, а именно, обеспечить совместную работу и взаимодействие сильно-связанных систем исполнителей в слабо связанной системы. Данный метод должен обеспечить повышение суммарной функциональности слабосвязанной системы из сильно-связанных систем исполнителей, т.е. ИР.

- 3) Расширить круг сильно-связанных систем, создавая из них сеть слабосвязанной системы. Повышение функциональность слабосвязанных систем достигается путем организационного метода из совокупности слабосвязанных систем, какими являются ИР.

Тем самым, для исполнения сложных операций с совокупностью сильно-связанных систем (это ИР), требуется сильно-связанные системы организовать в виде слабосвязанные системы.

4. Технологические основы процессов упрощения сложных задач

Процесс упрощения автоматизации состоит из следующих этапов.

1 этап. Сложные задачи могут быть в первую очередь декомпозированы (процесс декомпозиции могут несколько раз повторяться) на ряд более простых подзадач, которые в сумме наследуют свойства исходной задачи: когнитивные, физические (механические, подвижные), умственно-рутинные и физическо-рутинные операций.

2 этап. Процесс решения и выполнения декомпозированной задачи, в свою очередь, состоит из операций: когнитивных, физических (механические, подвижные), умственно-рутинных и физическо-рутинных. Поэтому нужны такие исполнители, которые обладают функциями, которые выполняют эти операций.

3 этап. Деление сложных операций на простые зависит от инфраструктуры их выполнения, и инфраструктуры или средств декомпозиций сложных операций на простые.

Таким образом, сложные задачи или операций сводятся на ряд декомпозированных, более простых, чем первоначальной постановки, т.е. «простым» задачам и их организации между собой. Эти простые задачи далее не декомпозируются при наличии существующих средств декомпозиции или инфраструктур, поэтому их назовем «конечными» задачами.

Причем, суммарная сложность этих декомпозированных «конечных» задач таковы, что их решения могут быть достигнуты при совместном (применений) выполнении как когнитивных, так и механических операций, при существующей инфраструктуре. Таким образом, не только первоначальные сложные задачи, но и конечные задачи в сумме содержат: когнитивные, механические и рутинные (умственные и физические) компоненты - операций.

Таким образом, в свою очередь эти «конечные» задачи являются трудной и сложной для автоматизации при наличии имеющихся средств (инфраструктуры) автоматизации.

Подытоживая можно утверждать, что для решения вопросов автоматизации сложных задач или операций следует упростить сложность операций, особенно механических операций, т.к. для выполнения механических операций требуется контакт систем-исполнителей. Для выполнения сложных когнитивных и механических задач (операции) применяют как стационарные, так и подвижные системы, но для выполнения сложных механических задач (операции) применяют только подвижные системы.

Литература

1. Kuandykov A.A. "The formalization of problem area, implementation and maintenance of and service of business-process by group of unmanned vehicles," IJCTA 2013, pp. 79-82.
2. Uskenbayeva R.K., Kuandykov A.A., Cho Young Im, Kalpeyeva Zh.B., Kozhamzharova D.K. "Organization of computing processes in the large heterogeneous distributed systems," Robotics (ISR) 2013, 44th International Symposium on Digital Object Identifier: 10.1109/ISR.2013.6695697, Publication Year: 2013, pp. 1- 4.
3. Uskenbayeva R.K., Kuandykov A.A, Cho Young, Kozhamzharova D.K., Kalpeyeva Zh.B. "Models and Methods of Joint Work Management of Group of Unmanned Vehicles," 2013 13th International Conference on Control, Automation and Systems (ICCAS 2013), Oct. 20-23, 2013 in Kimdaejung Convention Center, Gwangju, Korea.

А.Т. САНКИБАЕВ, Е.А ХАСЕНОВ

**ПРИМЕНЕНИЕ МЕТОДА САМООРГАНИЗУЮЩИХСЯ КАРТ КОХОНЕНА ДЛЯ
КЛАСТЕРИЗАЦИИ ДАННЫХ В ЗАДАЧЕ ВЫЯВЛЕНИЯ
ЛЖЕПРЕДПРИНИМАТЕЛЬСКОЙ ДЕЯТЕЛЬНОСТИ**

(ЕНУ им. Л.Н. Гумилева, Астана, Казахстан)

Задача выявления незаконной предпринимательской деятельности всегда была в поле зрения налоговых ведомств, силовых структур и органов власти всех уровней во многих странах мира. Научно-исследовательские учреждения также уделяют внимание этой задаче: ряд публикаций по этой теме дает представление об основных подходах к ее решению. М.Gupta и V.Nagadevara [1] использовали методы интеллектуального анализа данных для поиска потенциальных неплательщиков налогов в выборке 2003-2004 хозяйствующих субъектов в целом, включающей большие объемы налоговых деклараций в системе НДС в Индии. J.Velasques и др. [4] реализовали алгоритмы кластеризации для повышения эффективности налоговой проверки в Чили. Сегментация налогоплательщиков была получена с использованием методов интеллектуального анализа данных: самоорганизующихся карт Кохонена и K-means.

В данной работе представлены некоторые результаты решения задачи выявления лжепредприятий среди казахстанских юридических лиц на основе применения методов интеллектуального анализа заданного набора данных о деятельности предприятий за определенный промежуток времени. Ключевым для анализа задачи является предположение о том, все предприятия, по которым представлены данные, можно сгруппировать по классам, используя известные методы кластеризации и классификации.

Решением задачи классификации является отнесение каждого из объектов данных к одному (или нескольким) из заранее определенных классов и построение в конечном счете одним из методов классификации модели данных, определяющей разбиение множества объектов данных на классы.

В задаче кластеризации отнесение каждого из объектов данных осуществляется к одному (или нескольким) из заранее неопределенных классов. Разбиение объектов данных по кластерам осуществляется при одновременном их формировании. Определение кластеров и разбиение по ним объектов данных выражается в итоговой модели данных, которая является решением задачи кластеризации [3].

Если выходная переменная одна и она является дискретной (метка класса), то речь идет о задаче классификации.

Одним из широко применяемых методов кластеризации является метод самоорганизующихся карт Кохонена. Начиная с выбора признаков мы выполняем предварительный анализ данных, это позволяет нам сделать основные предположения о природе компаний, соблюдающих правила налогообложения. Затем мы выбираем статистическую модель и предлагаем алгоритм оценки его параметров в обучении с учителем [1].

Затем мы предлагаем определить статистическое распределение для описания поведения налоговых хозяйствующих субъектов и предлагаем алгоритм для оценки параметров такого распределения в обучении с учителем. Данный метод было производит обучение на реальных налоговых данных одного года, а затем были проведены расчеты на реальных данных следующего года.

Для того чтобы иметь возможность сравнить наш подход с некоторой базовой составляющей, мы использовали следующие атрибуты:

1. размерность предприятия;

2. суммы платежей, которые производило предприятие;
3. коды бюджетной классификации по платежам;
4. коды типа платежа;
5. срок уплаты;
6. сумма;
7. дата начала регистрации;
8. дата окончания регистрации;
9. дата реорганизации;
10. тип реорганизации.

Самоорганизующиеся карты и кластеризация

Самоорганизующиеся карты (SOM) являются одним из видов искусственной нейронной сети, который обучен используя обучение без учителя для получения двумерного, дискретного представления входного пространства образцов обучающих данных, называемого карта. Самоорганизующиеся карты отличаются от других искусственных нейронных сетей в том смысле, что использовать функцию двумерного пространства, чтобы сохранить топологические свойства пространства исходных данных. Это делает SOM полезным для визуализации с помощью низкоразмерных срезов многомерных данных, похожих на многомерное шкалирование. Самоорганизующаяся карта состоит из компонентов, называемых узлами или нейронами. Каждому узлу ставится в соответствие положение в пространстве карты K , и вектор прототипа m_k , той же размерности, что и вектор входных данных. Обычно узлы размещаются в виде двумерной гексагональной или прямоугольной сетки (рис 1). В гексагональной сетке, каждый нейрон соединен с шестью соседними, а в прямоугольной сетке каждый нейрон соединен с четырьмя соседями. В пространстве карты, соседи находятся на одинаковом расстоянии [2].

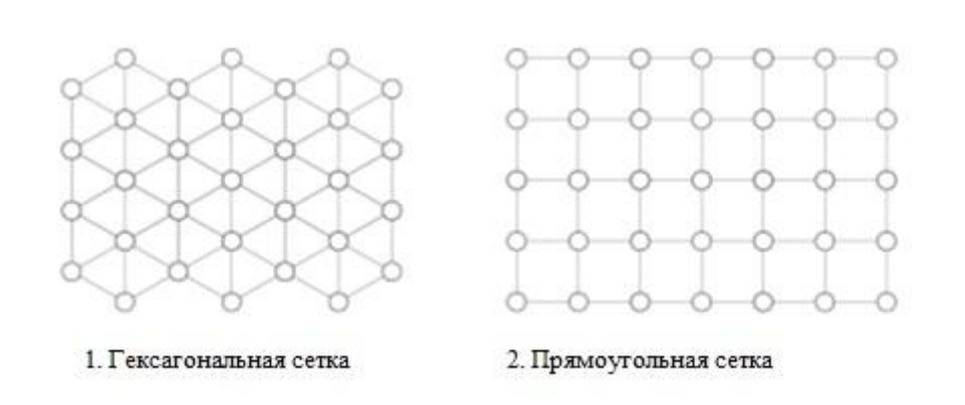


Рисунок 1. Виды узлов

Карты обучаются путем представления векторов данных на карте и корректировки прототипов векторов соответственно. Эти прототипы векторов инициализируются случайными значениями. Есть два подхода к обучению SOM: последовательное обучение и пакетное обучение. В последовательном обучении, учебные векторы могут быть взяты из набора данных в случайном порядке или циклически.

Мы использовали пакет машинного обучения Weka и базу данных Microsoft SQL Server 2014 как СУБД, а также среду разработки IntelliJ Idea. Вычисления производились на основе различных наблюдений (коммерческих предприятий). Алгоритм метода представлен в виде Java кода из файла SelfOrganizingMap.java, входящего в пакет Weka.

Алгоритм состоит из двух этапов:

- Обучение модели на заданном наборе данных;
- Проверка модели на том же наборе данных;

- Решение задачи кластеризации с обученной моделью на новом наборе данных
В результате работы алгоритма создается матрица из априори заданного количества узлов (нейронов) с вычисленными весовыми векторами. Эта матрица затем используется для кластеризации новых наборов данных, которые должны иметь такую же структуру, что и тренировочный набор данных, использованный для обучения модели.

Этапы работы с алгоритмом.

Данный алгоритм работает только с числовыми или списочными типами данных. Данные должны быть подготовлены как набор записей. Каждая запись состоит из набора атрибутов числового или списочного типа. Физически набор записей может представлять собой файл, оформленный в соответствии с форматом Weka arff, либо он может быть таблицей базы данных с аналогичной структурой. Настройка соответствия типов данных MSSQL к типам Weka осуществляется путем редактирования файла DatabaseUtils.props из пакета Weka.

Для работы алгоритма нужны следующие входные параметры:

- L : скорость обучения (значение по умолчанию 1)
- O : количество проходов на этапе упорядочивания (значение по умолчанию 2000)
- C : количество проходов на этапе схождения (значение по умолчанию 1000)
- H : высота кластерной решетки
- W : ширина кластерной решетки
- I : выполнять нормализацию атрибутов (по умолчанию да)

Имя файла или таблицы базы данных, содержащей тренировочные данные

Имя файла или таблицы базы данных, содержащей тестовые данные

По окончании работы выдается текст с результатами построения модели, тестирования и цветная диаграмма кластеризации тестовых данных

Описание алгоритма

Сначала инициализируется объект класса SelfOrganizingMap (строки 307-319 листинга программы).

Затем загружаются тренировочные данные в объект m_instances (строки 99-100 листинга).

Вычисляются постоянные коэффициенты для расчета эффективного радиуса решетки и скорости обучения (строки 685-687 листинга).

Инициализируется кластерная решетка, когда весовые вектора каждого узла принимают одинаковые начальные значения, т.е. значения атрибутов равны нулю, если записи нормализованы, или равны среднему арифметическому от диапазона изменения каждого атрибута (строки 983-998 листинга).

Начинается основной цикл обучения, при котором общее число проходов равно сумме количеств проходов этапов упорядочивания и схождения (строки 714-724 листинга).

Внутри цикла происходит следующее.

Вычисляется скорость обучения, как функция текущего номера прохода.

Вычисляется эффективный радиус решетки, как функция текущего номера прохода.

Во внутреннем цикле прогоняются все записи тренировочного набора данных с целью обновления весовых векторов узлов (нейронов) кластерной решетки. Внутри этого цикла выполняются следующие шаги:

- для текущей записи из тренировочного набора данных выявляется самый близкий к ней узел решетки в евклидовой метрике (строки 718 листинга). Этот узел называется победившим нейроном.

- обновляются весовые векторы всех узлов решетки на основе победившего нейрона. Этот цикл по всем узлам решетки и для каждого узла решетки выполняет следующее: по сквозному номеру победившего нейрона вычисляется его целочисленные

координаты в прямоугольной кластерной решетке (строки 830-831 листинга), аналогично для текущего узла (строки 832-838 листинга).

- По этим данным вычисляется евклидово расстояние между текущим нейроном и победившим нейроном [5]. Это расстояние затем используется для расчета поправочного коэффициента h . Затем для каждого атрибута текущей записи тренировочного набора данных вычисляется отклонение его от значения соответствующего атрибута весового вектора текущего узла с учетом поправки h и скорости обучения. Затем значения атрибутов весового вектора обновляются на величину этого отклонения.

В процессе прохождения всех этих циклов весовые вектора непрерывно обновляются, при этом величины обновлений постепенно затухают, т. е. Значения весовых векторов кластерной решетки стабилизируются.

Результатом обучения модели является набор весовых векторов. Обученная модель как алгоритм с вычисленными весовыми векторами сохраняется в виде файла `weka.clusterers.SelfOrganizingMap.cls`. Обученная модель используется для кластеризации нового набора данных. В нашем случае новый набор данных означает другой временной период для того же множества предприятий. В результате работы обученной модели на новом наборе данных заданное множество предприятий делится на группы. По результатам численного эксперимента множество из 5780 предприятий разделилось на три группы: 45 попали в первую подгруппу, 5632 попали во вторую подгруппу, 103 – в третью подгруппу.

Таким образом, предположительно 0,08% всех предприятий были отнесены к лжепредприятиям. На самом деле результаты эксперимента не указывают явно – является ли найденная группа предприятий лжепредприятиями. Для выяснения этого вопроса требуется дальнейший анализ данных с привлечением методов классификации. На данном этапе, было выяснено, что действительно по представленному набору данных можно эффективно разделить предприятия на группы.

Листинг программы на языке JAVA

```
99      /** The training instances. */
100     private Instances m_instances;
...
307     public SelfOrganizingMap() {
308         m_clusters = null;
309         m_width = 2;
310         m_height = 2;
311         m_convergenceEpochs = 1000;
312         m_orderingEpochs = 2000;
313         m_learningRate = 1.0;
314         m_normalizeAttributes = true;
315         m_calcStats = true;
316     }
...
686     // normalize instances
687     m_instances = normalize(m_instances);
...
714     // update the weights
715     for (int neuron = 0; neuron < m_clusters.numInstances(); neuron++) {
716         updateWeights(neuron, winningNeuron, instance, learningRate,
effectiveWidth);
717     }
718 }
719 }
```

```

720
721     if (m_calcStats) {
722         calcStatistics();
723     }
724 }
...
830     double distance = Math.pow((winnerWidth - width), 2) + Math.pow((winnerHeight
- height), 2);
831     double h = Math.exp(-distance / (2 * Math.pow(effWidth, 2)));
832     for (int j = 0; j < m_clusters.numAttributes(); j++) {
833         diff = learningRate * h * (m_instances.get(instance).value(j) -
m_clusters.get(updateCluster).value(j));
834         if (!Double.isNaN(diff)) {
835             m_clusters.get(updateCluster).setValue(j, m_clusters.get(updateCluster).value(j) +
diff);
836         }
837     }
838 }
...
983     Instances weights = new Instances(m_instances, m_width * m_height);
984     for (int i = 0; i < m_width * m_height; i++) {
985         double[] instValues = new double[m_instances.numAttributes()];
986         for (int j = 0; j < m_instances.numAttributes(); j++) {
987             if (m_normalizeAttributes) {
988                 instValues[j] = 0;
989             } else {
990                 instValues[j] = (m_attributeMax[j] + m_attributeMin[j]) / 2;
991             }
992         }
993         Instance inst = new DenseInstance(1, instValues);
994         weights.add(i, inst);
995     }
996
997     return weights;
998 }

```

Список литературы:

- 1 Gupta M., Nagadevara V. Audit Selection Strategy for Improving Tax Compliance – Application of Data Mining Techniques – Computer Society of India p.378-379;
- 2 Горбаченко В.И. Сети и карты Кохонена – стр.18;
- 3 Барсегян А.А., Куприянов М.С., Холод И.И. Технологии анализа данных – ВHV строки145;
- 4 P.Gonzalez, J.Velasques Characterization and detection of taxpayers with false invoices using data – Expert Systems with Applications p.1427-1430.
- 5 Г.Шилдт Java 8. Полное руководство, 9-е издание, стр. 102-105

М.С. САУХАНОВА, Ж.С. САУХАНОВА, А.Ә. ҚАЛДЫБЕК

RAPID MINER ОРТАСЫНДА МӘТІНДЕРДІ ТАЛДАУ

(Л.Н. Гумилев атындағы ЕҰУ, Астана, Қазақстан)

Data Mining – адамзат қызметінің әр түрлі салаларында шешім қабылдау үшін қажетті бұрын белгісіз болған, пайдалы және қолжетімді білімді анықтау әдістерінің жиынын белгілеуде қолданылатын жалпы атау.

Data Mining қарастыратын негізгі есептердің бірі - кластерлеу есебі. Кластерлеу есебін шешу барысында зерттелетін объекттер жиынын кластер (cluster) деп аталатын бір-біріне ұқсас объекттер тобына бөлу қарастырылады. Cluster сөзі ағылшын тілінен аударғанда топ, түйін, байлам дегенді білдіреді. Әдетте элементтер жиынын кластерлерге бөлу есебін кластерлік талдау деп атайды.

Ғылыми зерттеулер үшін кластерлеу нәтижелерін, атап айтқанда қандай жағдайларға байланысты объектілер топтарға бірігетінін зерттеу көптеген жаңа бағыттарды ашуға мүмкіндік туғызады.

Кластерлеу есебінде бір-біріне ұқсас объектілер ізделінеді. Кластерлерді анықтаған соң мұндай бөлініс не мағына білдіретінін, қандай критерийлер бойынша бөлінгенін анықтау үшін Data Mining-тің басқа әдістері қолданылады.

Кластерлеу өте көлемді ақпараттарды кластерлерге бөлу арқылы талдау есебінің көлемін азайтуға мүмкіндік береді.

Кластерлік талдаудың ең басты артықшылығы объектілерді тек бір ғана параметр негізінде емес, параметрлер жиынтығы негізінде объектілерді топтау болып табылады. Кластерлік талдаудың басқа көптеген математикалық статистика әдістерінен өзгешелігі - қарастырылатын объектілердің түрлеріне ешқандай шектеулер қойылмауы.

Мәтінде әдетте көп ақпарат жасырын болады, бірақ оның құрылымсыз түрде болуы Data Mining әдістерін қолдануға мүмкіндік бермейді. Бұл мәселені шешумен құрылымсыз мәтінді талдаудың әдістері айналысады. Батыс әдебиеттерінде мұндай талдауды Text Mining деп атайды. Құрылымсыз мәтінді талдаудың әдістері бірнеше салалардың тоғысында жатыр: Data Mining, табиғи тілде мәтінді өңдеу, ақпарат іздеу, ақпаратты алу және білімді басқару.

Мәтінде білімді анықтау – бұл құрылымсыз мәтінді деректерде жаңа, пайдалы және түсінікті шаблондарды анықтау.

Тақырыптық санаттар бойынша мәтіндерді кластерлеу әр түрлі ақпараттық жүйелерде қажетті маңызды есеп болып табылады.

Мәтіндерді кластерлеу үшін құжатты терминдер векторы түрінде ұсыну қолданылады.

T - n -өлшемді кеңістікті анықтайтын барлық бірегей жүйе терминдерінің жиыны болсын:

$$T = \{ t_1, t_2, \dots, t_n \},$$

Мұнда әрбір t_i термині бір кеңістік өлшеміне сәйкес келеді.

Берілген n -өлшемді кеңістікте K топтаманың әрбір D құжаты v_D векторымен анықталатын нүктеге сәйкес келеді:

$$v_D = \langle (t_1, \omega_1), (t_2, \omega_2), \dots, (t_n, \omega_n) \rangle,$$

мұндағы ω_i - t_i терминінің салмақ коэффициенті.

Салмақ коэффициентін есептеудің тиімді тәсілдерінің бірі $tf.idf$ әдісі болып табылады. Ол бойынша ω_i мәні келесі түрмен анықталады:

$$\omega_i = \begin{cases} tf.idf_{D,K}(t_i), & \text{егер } t_i \in D \\ 0, & \text{басқа жағдайларда} \end{cases}$$

$tf.idf$ салмақ коэффициентін келесі түрде есептейміз.

Терминнің салмақ коэффициентіндегі tf терм жиілігі (term frequency) және idf құжат жиілігінің кері мәні (inverted document frequency) дегенді білдіреді.

D құжатындағы t терминінің tf коэффициенті келесі формуламен анықталады:

$$tf_D(t) = \frac{freq_D(t)}{|D|} \quad (1)$$

мұндағы $freq_D(t)$ - D құжатындағы t терминінің кездесу жиілігі, $|D|$ - D құжатындағы барлық терминдер саны (D құжатының өлшемі).

K топтамасындағы t терминіне қатысты df (document frequency) құжатының жиілігі келесі формуламен есептеледі:

$$df_K(t) = \frac{|\{D_i | D_i \in K, t \in D_i\}|}{|K|} \quad (2)$$

мұндағы $|K|$ - K топтамасындағы құжаттар саны. Құжаттар жиілігінің кері мәні келесі формуламен есептеледі:

$$idf_K(t) = \log\left(\frac{1}{df_K(t)}\right) \quad (3)$$

K топтамасындағы D құжаттың t_i терминінің ω_i коэффициентін келесі формуламен есептейміз:

$$tf * idf_{D,K}(t) = \frac{freq_D(t)}{|D|} * \log\left(\frac{|K|}{|\{D_i | D_i \in K, t \in D_i\}|}\right) \quad (4)$$

Мәтіндерді талдау барысында жолдары әрбір құжатқа сәйкес $v_D = \langle (t_1, \omega_1), t_2, \omega_2, \dots, t_n, \omega_n \rangle$ векторларының t_i терминіне сәйкес салмақ коэффициенттерінен тұратын терминдер-құжаттар матрицасы қарастырылады. Құжаттардың ұқсастығы – семантикалық жақындығы n өлшемді кеңістіктегі векторлардың жақындығы негізінде анықталады. Арақашықтық өлшемі ретінде косинустық арақашықтық қарастырылады:

$$cosine(a, b) = \frac{\sum_i a_i b_i}{\sqrt{\sum_i a_i^2 \sum_i b_i^2}} \quad (5)$$

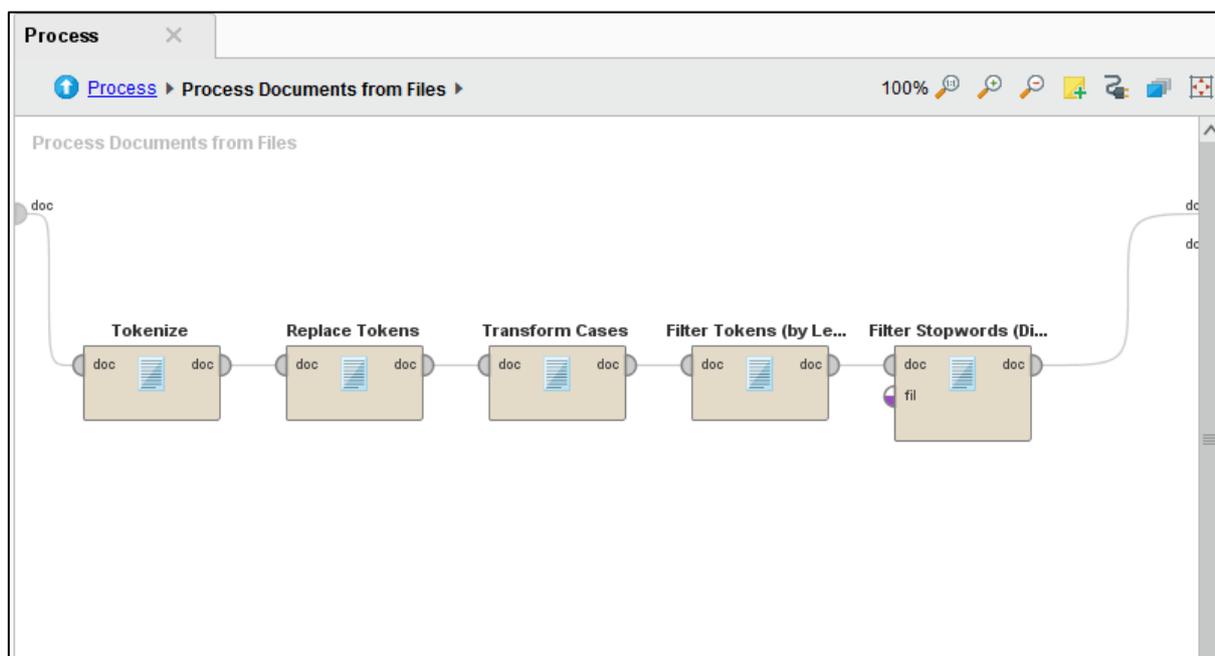
Терминдер-құжаттар матрицасы Rapid Miner ортасында «Text Processing» операторлар жиынтығы көмегімен құрылады. Rapid Miner ортасында қазақ тіліндегі мәтіндерді алдын ала notepad++ қосымшасы көмегімен UTF-8 кодировкасын қолданып .txt кеңеймесімен сақтаймыз.

Tokenize операторының көмегімен мәтіндердегі терминдер анықталады. Ол құжат мәтінін лексемаларға бөледі. Барлық дерлік құжаттарда кездесетін, сондықтан құжаттарды талдау барысында бағалы ақпарат бермейтін stop-сөздер жиынтығын Filter stopwords (by dictionary) операторы көмегімен құрамыз. Оператордың encoding параметрінің мәні UTF-8 кодировкасына тең болу керек.

Түбірлес сөздердің қайталануын болдырмау үшін Stem(by dictionary) операторын қолданылуы қазақша ә,ң,ғ,ү,ұ,қ әріптерінің дұрыс бейнеленбеуіне алып келеді, сондықтан Replace Tokens операторының көмегімен стемминг сөздігін құрамыз. Қосымша

Терминдер кеңістігінің өлшемін төмендету мақсатында Filter stopwords (by dictionary), Replace Tokens операторларымен қоса Transformer cases, Filter tokens(by length) операторлары қолданылды.

Text processing операторлар жиынтығындағы құжаттарды талдау Process documents from files күрделі операторының құрамдас бөліктері келесі түрде анықталады (Сурет 1):



Сурет 1. Process documents from files күрделі операторының құрамдас бөліктері

Құжаттарды кластерлеу үшін k-means алгоритмі қолданылды. Бұл алгоритм барлық кластер үшін кластер центрінен кластер объектілеріне дейінгі ортақвадраттық ауытқулар қосындысын барынша азайтуға негізделген. Тиімді кластерлер санын анықтау үшін Loop Parameters операторын қолдануға болады.

Қолданылған әдебиеттер:

1 Анализ данных и процессов / А.А.Барсегян, М.С. Куприянов, И.И. Холод, М.Д. Тесс, С.И. Елизаров. – 3-е изд. перераб. и доп. – СПб.: БХВ-Петербург, 2009. - 512 с.

2 Нгуен Ба Нгок, Тузовский А.Ф. Классификация текстов на основе оценки семантической близости терминов// Управление, вычислительная техника и информатика, 2012. URL: <http://cyberleninka.ru/article/n/klassifikatsiya-tekstov-na-osnove-otsenki-semanticheskoy-blizosti-terminov> .pdf

САЛЫҚ САЛАСЫ БОЙЫНША ҚАЗАҚ ТІЛІНДЕ ИНТЕЛЛЕКТУАЛЬДЫ СҰРАҚ-ЖАУАП ЖҮЙЕСІН ҚҰРУ

(Л.Н.Гумилев атындағы Еуразия ұлттық университеті, Астана, Қазақстан)

Кіріспе

Компьютерлерді алғашқы құру дәуірінен бастап білімдерді формальды бейнелеу және білімді автоматты өңдеу әдістері саласында зерттеудің түрлі бағыттары пайда болды. Бұл зерттеулер жыл сайын артып, 1980 жылдардың соңында информатиканың жеке бір саласын құрады, бүгінгі күні мұны білім инженериясы деп атауға болады. Аталмыш зерттеулерге қызығушылықтың артуының шешуші факторы ол – қолжетімді цифрлы ақпарат көлемінің артуы және осы ақпаратқа қолжеткізе алатын қолданушылар санының өсуі. Осыған байланысты кең көлемдегі мәлімет ауқымынан өз қажетіне жарар ақпаратты алуға мүмкіндік беретін ақпараттық жүйелерді жасау қажеттілігі туындайды. Мұндай бағытта ең тиімді шешім ол - сұрақ-жауап жүйелерін құру.

Сұрақ-жауап жүйелері саласындағы зерттеулер екі бағытта жүреді: ақпараттық-іздеу жүйелері(*information retrieval systems*) және интеллектуальды жүйелер(*intelligence support systems*)[1].

Ақпараттық-іздеу саласында дамып келе жатқан сұрақ-жауап жүйелері қойылатын сұраққа Интернет желісінде қолжетімді, қысқаша мәтін түрінде жауап беруге бағытталған. Жалпы сұрақ-жауап жүйелері екі түрге бөлінеді: ашық доменді(*open domain*) және шектеулі доменді(*restricted domain*) жүйелер[2;3].

Ашық доменді жүйелер кез келген пәндік саланың сұрақтарын өңдеуге бағытталған жалпы сұрақ-жауап жүйелеріне жатады. Ал, шектеулі доменді жүйелер белгілі бір пәндік сала, яғни шектелген тақырыптардың сұрақтарын өңдеуге арналған және олар арнайымамандандырылған сұрақ-жауап жүйелері болып табылады.

Бірақта олардың барлығы мәтінді талдау(*text mining*) және табиғи тілдік өңдеу(*natural language prossing*) тәсілдерін жасауға назар аударады. Олар тек онтологияны құру дәрежесімен ғана ерекшеленеді[1].

Ашық доменді сұрақ-жауап жүйелері табиғи тілдің жалпы онтологиясын қолданады(мысалы, WordNet онтологиясы). Осының салдарынан олар табиғи тілдің синономия, омонимия, полиморфизм және о.ұ. секілді мәселелерін шешуге мәжбүр. Қазіргі заманғы ашық доменді сұрақ-жауап жүйелері тек бірнеше алдын ала анықталған сұрақтар класстарын-фактографиялық сұрақтарын кезекті өңдеу үшін формалдай алады. Фактографиялық сұрақтарды былай бөлуге болады: адам туралы сұрақтар, уақыт туралы сұрақтар, анықтама туралы сұрақтар және о.ұ.

Шектеулі доменді сұрақ-жауап жүйелері тармамандандырылған онтологияларды пайдаланады. Бұл біржағынан табиғи тіл сөздерінің көпмағыналық мәселесін шешеді, ал екінші жағынан берілген пәндік саланың арнайы сұрақтарына жауаптар алу мүмкіндігін береді.

Жасанды интеллект саласындағы зерттеулер білім көздері ретінде түрлі білімдер қорларын пайдаланатын білімдерге негізделген(*knowledge based QA-systems*) сұрақ-жауап жүйелерін құруға әкелді. Әрине мұндай сұрақ-жауап жүйелерінің шектеулі доменді жүйелер болатыны айқын. Бұл жүйелердің алдын ала білімдер қорында кодталған ақпаратпен байланысты, яғни сұрақтарды қалыптастыруда неғұрлым жайлы жауаптар беретінін атап өту керек.

Бұл қадамның басты ерекшелігі білімдер қоры құрылымында көрсетілген логикалық қорытынды, ұқсастық дәлелі және о.ұ. секілді құрылымдық ақпараттарды өңдеудің озық технологияларын қолдануға мүмкіндік беретін концептуальды моделдің болуы. Негізінде

мұндай жүйелер сұралатын ақпаратты іздеу және оқшаулауға емес, олар берілген пәндік саланың жасырын заңдылығын, сыни жағдайларын талдауды, тәуекелділікті бейнелеуге бағытталған.

Медициналық салада құрылған L&C[5] жүйесі нақты емделушілер туралы ақпараттарымен қатар беделді медициналық білім интеграциясының міндеттерін орындайды. Демир-Фушманның[6] сұрақ-жауап жүйесі клиникалық медицинаның статистикалық тәсілдерін қолдану негізінде құрылған. WEBCOOP[7] жүйесі туризм саласы бойынша жауаптарды генерациялау үшін логикалық қорытынды процедурасын қолданады.

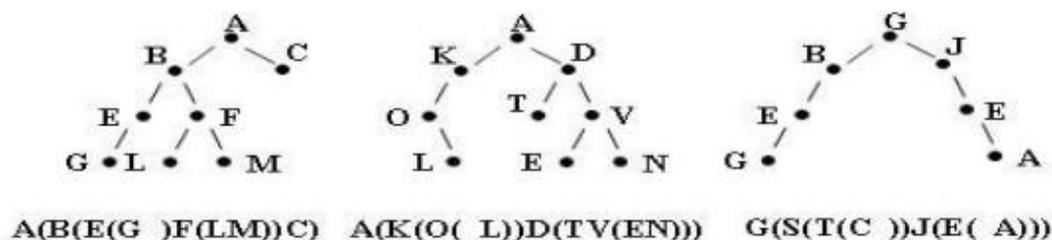
Бұл жұмыста біз салық саласы пәндік облысы үшін құрылған «*Gaztaxexpert*» сұрақ-жауап жүйесін сипаттаймыз. Аталмыш жүйе пәнді саланы моделдеуге прецеденттік жолтабу негізінде құрылып және қолданушының салық салуға қатысты сұрақтарына жауап алуға мүмкіндік береді.

Салық жүйесі және салық түрлерінің онтологиялары

Онтология түсінігі грек тілінен аударғанда онтос – жан, мән,мағына, логос - білім дегенді білдіреді, яғни болмысты зерттейтін философиялық ілім. Техникалық ғылымдарда онтологияның қолданылуының негізгі мәні белгілі бір білім облысы бойынша мәліметтер жиынының барлығын қамтитын және бөліктік формализацияны концептуальді сызбамен көрсетуі. Концептуальды сызба негізінде түсініктер жиыны + түсініктер жайлы мәліметтер (қасиет, қатынас, шектеу, аксиомалар және түсініктердің бекітілуі, бұл ақпараттардың барлығы таңдалынған пәндік облыс бойынша есептің шешілу процесін сипаттау үшін қажет) беріледі.

Бұл формализм онтологияны семантикалық гиперграф түрінде бейнелеуге мүмкіндік береді: $O(X, R, I)$, мұндағы X – пәндік облыс түсініктерінің жиыны (гиперграф төбесі), R – түсініктер арасындағы қарым- қатынастар жиыны (гиперграф доғалары мен қабырғалары), ал I – осы пәндік облыстағы түсініктер мен қарым- қатынастарының жиыны.

Семантикалық гиперграф арқылы түбір сөздерге жалғаулар мен жұрнақтарды қосу ережелері формалданады. Гиперграфтар көп өлшемді құрылым болып табылады. Бірақ көптеген жағдайда бір өлшемді құрылымдарды пайдаланған тиімді. Сондықтан да көп өлшемді графтар құрылымындағы ақпараттың бір өлшемді құрылымға өткені біз үшін маңызды. Графтарды бірөлшемді құрылым ретінде бейнелеу үшін сызықтық жақшалық жазбаларды қолдануға болады, яғни, гиперграфтарды жолдар түрінде бейнелеу [6]. Онда гиперграфтың төбелері ашылған және жабылған дөңгелек жақша түрінде бейнеленеді. Гиперграфтар мен олардың сызықтық жақшалық жазбалары арасында өзара тығыз біртепті байланыс бар. Сызықтық жақшалық жазба қандай да бір гиперграфты толықтай тексеріп, айналып өткеннен кейін жасалынады. Графтар мысалы және оларға сәйкес келетін сызықтық жақшалық жазбалардың мысалдары сурет 1-де көрсетілген.



Сурет 1. Сызықтық жақшалық жазба түрлері

Төменде салық жүйесі және салық түрлерінің онтологиялық модельдері, сызықтық жақшалық жазба арқылы алынған формалды ережелері келтіріледі.

Салық жүйесінің онтологиясы салықтың жалпы құрылымын, зерттеу объектілері мен субъектілерін, жалпы ұғымдарын бейнелейді. Салық жүйесі және салық түрлерінің онтологиялық моделдері онтология құрудың негізгі ережелеріне сай құрылады және білімді ұсыну тілі ретінде семантикалық гиперграф алынады[7].

Салық жүйесінің онтологиясы келесі түрде болуы мүмкін:

$O_{СЖ} = \langle X, R, I \rangle$, мұнда

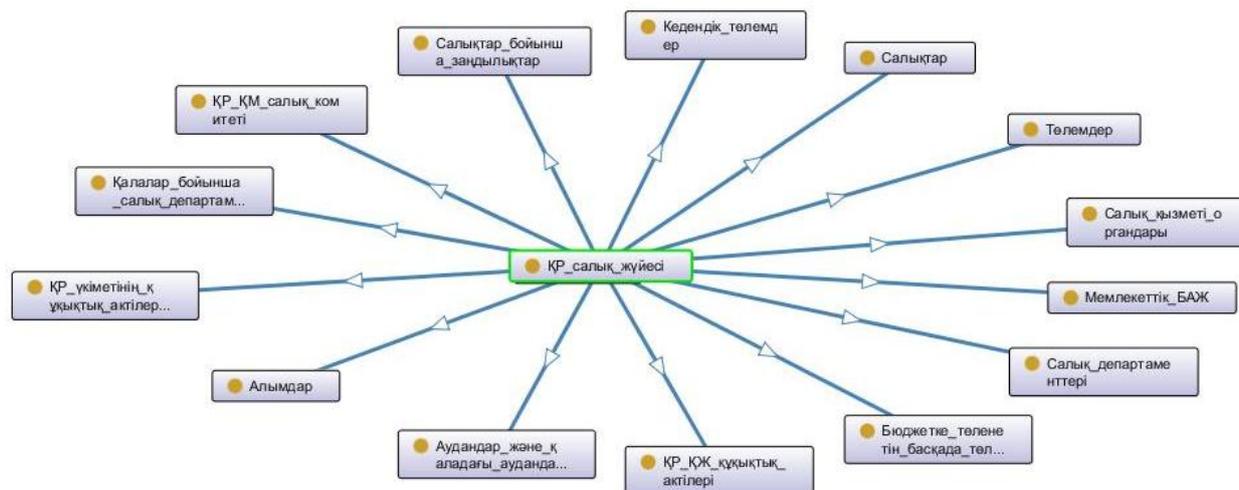
X - нақты пәндік сала ұғымдарының жиыны;

R - ұғымдар анықтамасының жиыны;

I – ұғымдар арасындағы қатынастар жиыны.

Салық жүйесі онтологиясы құрамына $O_{СЖ}$. бюджетке төленетін міндетті төлемдер мен салықтар, ҚР ҚМ құқықтық актілері, ҚР үкіметінің құқықтық актілері, ҚМ салық комитеті, салық департаменттері, қалалар бойынша салық департаменттері, аудандар бойынша салық департаменттері, аудандағы және қаладағы аудандар бойынша салық департаменттері, салықтар бойынша заңдылықтар, кедендік төлемдер, салық қызметі органдары, мемлекеттік БАЖ, бюджетке төленетін басқа да төлемдер, алымдар, салықтар кіреді. Ал, салықтар құрамы салық түрлерінен тұрады. Сурет 2 – салық жүйесінің онтологиялық моделі көрсетілген.

Салық жүйесі пәндік облысы онтологиясын графикалық жобалау және формалды сипаттамасын құру. Алдыңғы деңгейлер бойынша анықталған жиындар негізінде салық жүйесі облыстың концептуалды моделі құрылды. Ол үшін Protégé онтологиялық редакторы қолданылды. 3 суретте салық жүйесі онтологиясының бөлігі бейнеленген.



Сурет 2 – салық жүйесінің онтологиялық моделі

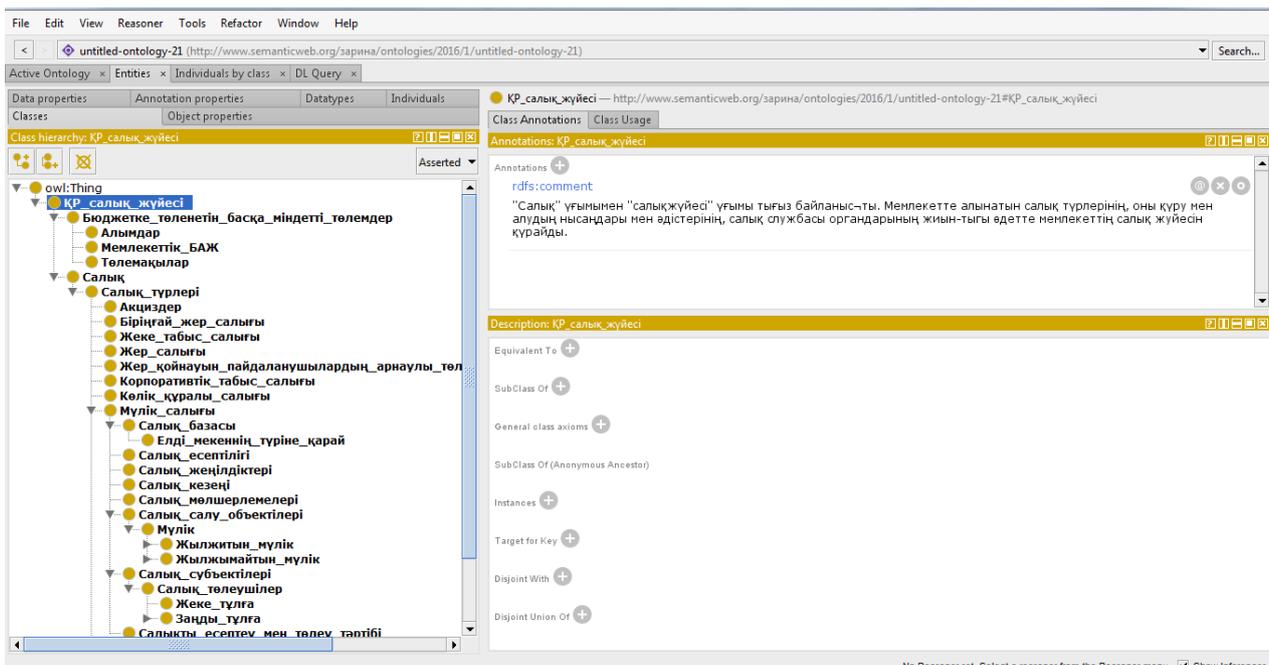
Салық түрлерінің онтологиясы келесі түрде болуы мүмкін:

$O_{СТ} = \langle X, R, I \rangle$, мұнда

X - нақты пәндік сала ұғымдарының жиыны;

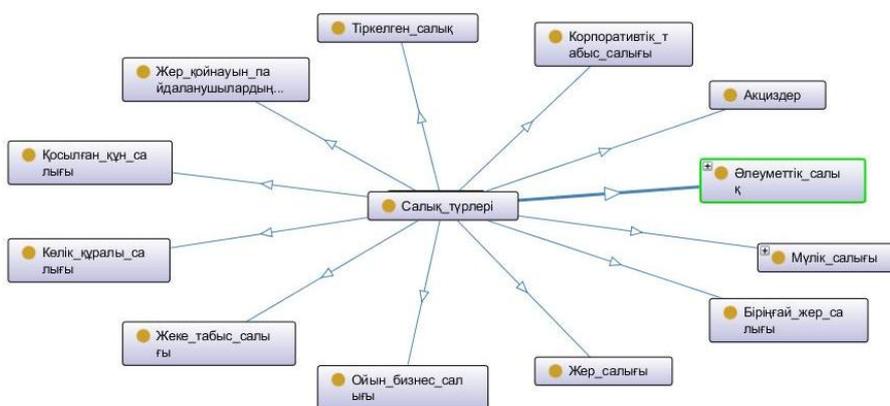
R - ұғымдар анықтамасының жиыны;

I – ұғымдар арасындағы қатынастар жиыны.



Сурет 3 – салық жүйесі пәндік облысы онтологиясы бір бөлігі

Салық түрлері онтологиясы құрамына *ОСТ* - тіркелген салық, корпоративтік табыс салығы, акциздер, әлеуметтік салық, мүлік салығы, бірыңғай жер салығы, жер салығы, ойын-бизнес салығы, жеке табыс салығы, көлік құралы салығы, қосылған құн салығы және жер қойнауын пайдаланушылар салығы кіреді. Ал, салықтар құрамы салық түрлерінен тұрады. Сурет 4 – салық жүйесінің онтологиялық моделі көрсетілген.



Сурет 4 – салық түрлерінің онтологиялық моделі

Gaztaxexpert жүйесінің қолданушы интерфейсі

Қолданушылар сұрақ-жауап жүйелеріндегі сұрақтарды неғұрлым табиғи тілге жақын болуы қажет деп санайды. Ал, жүйе білім қорынан жауапты іздеу бірнеше бекітілген логикалық қойылымда қолданушы сұрағын бізмәнді аударуда ғана жүреді дейді. Осылайша программалық жүйеде сұрақты қалыптастыру екі қарама-қайшы мақсатты көздейді, яғни программалық интерфейсті құру міндеті табиғи тіл және сұраудың маңыздылығы арасындағы шешуді іздеуге айналады.

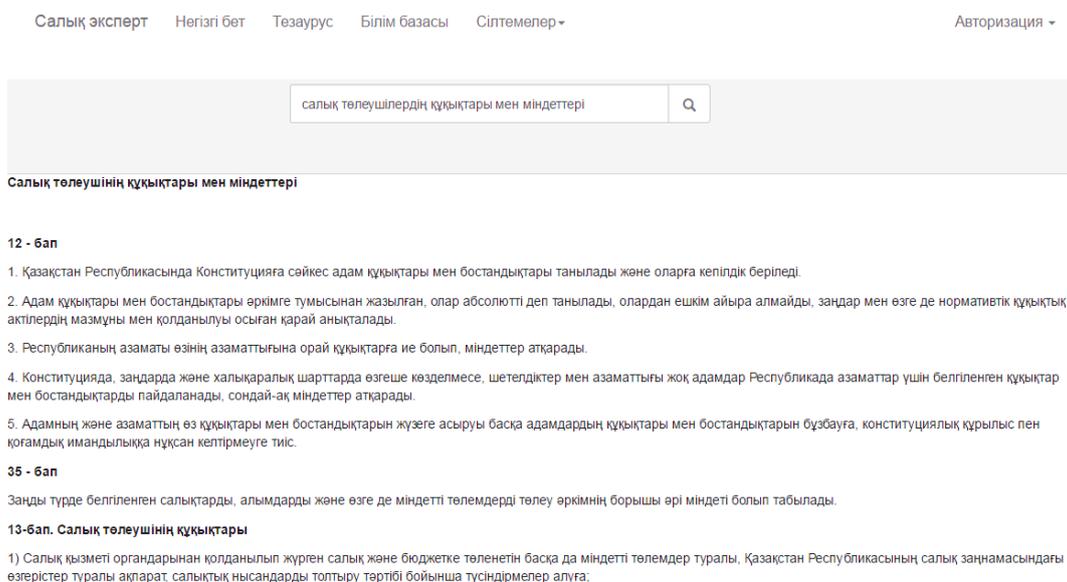
Құрылатын сұрақ-жауап жүйесі пәнді сала бойынша түрлі типтегі сұрақтарға жауап бере алады.

Салық салуға қатысты қолданушы сұрақтарына табиғи тілде жауап беретін *Gaztaxexpert* сұрақ-жауап жүйесін іске қосу интернет-браузер арқылы жасалады. Жүйеге келесі сілтеме арқылы өтуге болады: qaztaxexpert.appspot.com (5 сурет).

Gaztaxexpert табиғи тіліндегі сұрақ-жауап жүйесі келесі қосымшалардан тұрады:

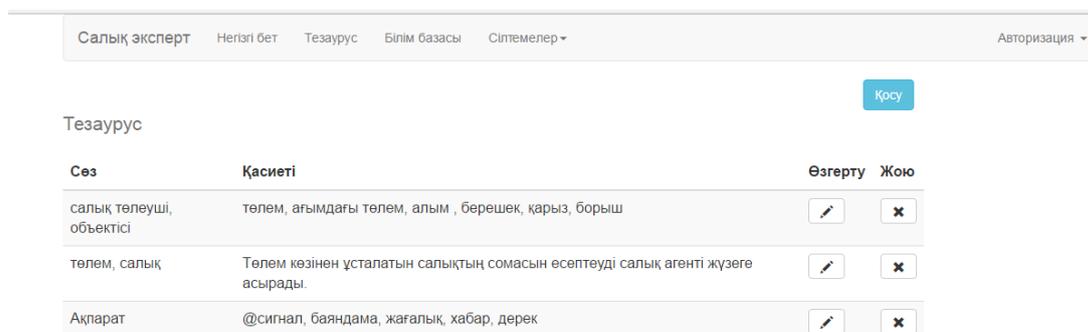
- салық эксперт;
- негізгі бет;
- білім базасы;
- сілтемелер;
- авторизация.

Салық эксперт пен негізгі бет табиғи тілде білім базасына сұрақ қойылады, табиғи тілдегі сұрақ SPARQL сұратымға аударылады. SPARQL сұратым арқылы алынған жауап қолданушыға табиғи тілге аударылып көрсетіледі(5 сурет).



5 сурет - сұрақ жауап жүйесі басты беті

Тезаурус білім базасында және пәндік салада кездесетін барлық түсініктер мен ұғымдар, яғни сөздер қоры жүргізіледі. Тезаурустағы сөздер лингвистикалық талдауға негіз болады(6 сурет).



6 сурет – тезаурус беті

Білім базасы нормалды қазақ тілінде білім базасын толтыруға арналған мәтіндік редактор мен json форматындағы мәліметтерді көрсетуге арналған jstree компонентінен тұрады.

Нормалды қазақ тілінде білім базасын толтыруға арналған мәтіндік редактор - бұл білімді енгізудің күрделі жүйелерін меңгермеген қарапайым қолданушыларға білім қорын кәдімгі ауызекі тілде толтыруға мүкіндік береді. Нормалды қазақ тілі дегеніміз қазақ тілінің бір қалыпқа келтірілген, белгілі бір анықталған грамматикалық және лексикалық ережелер шеңберіндегі ішкі жиыны. Бұл мәндік редактор қолданушының бейберкет еркін стильде жазу жазуын шектейді және нақты тұжырымдар енгізуге көмектесіп отырады. Нормалды қазақ тіліндегі жазу OWL, RDF синтаксисіне автоматты түрде оңай аударылады(7 сурет).

Зат есім	Анықтамасы	Гипероним	Гипоним	Холоним	Синоним	Антоним	Омоним	Мероним	Өзгерту	Жою
салық, құн	Салық базасы салық салу объектісінің құн, физикалық немесе өзге де сипаттамаларын білдіреді, олардың негізінде салықтың және бюджетке төленуге жататын басқа да міндетті төлемдердің сомасы айқындалады.	төлем	ағымдағы төлем, алым, берешек, қарыз, борыш	салық	тиісті	тиісті емес, міндетті емес		мөлшері, кезеңі, түрі, төлемі		
Ақпарат	бұл ақпараттық процесс барысында түзілетін қозғалыстағы объект	мәлімет	@сигнал, баяндама, жағалық, хабар, дерек	деректер базасы	мәлімет, дерек, мағұмат, мәлімдеме, хабарлама,			ұзындығы, энтропия, биттер		

7 сурет – білім базасы беті

JSON форматындағы білім қоры jsontree ағаш түріндегі иерархиялық мәліметтерді көрсетуге арналған компонент. json форматындағы мәліметтерді тармақтарын тарқатып көрсете алады. **JSON (JavaScript Object Notation ағылшынша, әдетте /'dʒeɪs'əp/)** — деп айтылады. JavaScript негізінде қаланған мәлімет алмасудың мәтіндік форматы(8 сурет). Басқа да көптеген мәтіндік форматтар секілді JSON адамдармен жеңіл оқылады. JSON форматын Дугласом Крокфорд жасаған.

Көптеген тілдер үшін JSON форматында мәліметтер өңдеу мен құру үшін код дайын.

The screenshot shows a SPARQL query interface. The query endpoint is set to `http://localhost:3030/taxontology/query`. The content type is set to `JSON`. The results are displayed in a table view, with a 'Raw Response' tab selected, showing the following JSON output:

```

1 {
2   "head": {
3     "vars": [ "subject", "predicate", "object" ]
4   },
5   "results": {
6     "bindings": [
7     {
8       "subject": { "type": "uri", "value": "http://www.alphabet.kz/ontologies/tax" },
9       "predicate": { "type": "uri", "value": "http://www.w3.org/1999/02/22-rdf-syntax-ns#type" },
10      "object": { "type": "uri", "value": "http://www.w3.org/2002/07/owl#Ontology" }
11    },
12    {
13      "subject": { "type": "uri", "value": "http://www.alphabet.kz/ontologies/tax#салық" },
14      "predicate": { "type": "uri", "value": "http://www.w3.org/1999/02/22-rdf-syntax-ns#type" },
15      "object": { "type": "uri", "value": "http://www.w3.org/2002/07/owl#Class" }
16    },
17    {
18      "subject": { "type": "uri", "value": "http://www.alphabet.kz/ontologies/tax#салық" },
19      "predicate": { "type": "uri", "value": "http://www.w3.org/2002/07/owl#Class" }
20    }
21  ]
22 }

```

8 сурет – JSON форматындағы білім қоры

Сілтемелер салық жүйесіне қатысты интернет ресурстарға сілтемелер келтірілген(9 сурет).

Зат есім	Анықтамасы	Гипероним	Гипоним	Холоним	Синоним	Онтоним	Омоним	Мероним	Өзгерту	Жою
Ақпарат	бұл ақпараттық процесс барысында түзілетін қозғалыстағы объект	мәлімет	@сигнал, баяндама, жағалық, хабар, дерек	деректер базасы	мәлімет, дерек, мағұлмат, мәлімдеме, хабарлама,			ұзындығы, энтропия, биттер		
Ақпарат	бұл ақпараттық процесс барысында түзілетін қозғалыстағы объект	мәлімет	@сигнал, баяндама, жағалық, хабар, дерек	деректер базасы	мәлімет, дерек, мағұлмат, мәлімдеме, хабарлама,			ұзындығы, энтропия, биттер		
Ақпарат	бұл ақпараттық процесс барысында түзілетін қозғалыстағы объект	мәлімет	@сигнал, баяндама, жағалық, хабар, дерек	деректер базасы	мәлімет, дерек, мағұлмат, мәлімдеме, хабарлама,			ұзындығы, энтропия, биттер		

9 сурет – Сілтемелер беті

Қорытынды

Қазіргі кезде мәтіндік ақпараттың күн санап өсуі және олардың құрылымдарының күрделі болуы табиғи тілдегі мәтіндерді талдайтын, олардың мағынасын түсіне алатын және олардың негізінде жаңа білімдер алатын, адам мен машина арасында табиғи тілде қарым-қатынас жасауға мүмкіндік беретін интеллектуалды жүйелерді қажет етеді. Ондай жүйелерді жобалау және құру үшін білімдерді бейнелеу мен өңдеудің моделдері мен әдістерін қарастыру керек.

Мақалада сұрақ-жауап жүйелері, салық жүйесі және салық түрлерінің онтологиялық моделдері берілді. **Gaztaxexpert** сұрақ-жауап жүйесінің жұмыс жасалу интерфейстерінің әр бетіне түсініктеме берілді. Бұл жүйе әлі толық емес, келешекте білім қорлары толықтырылатын болады.

Әдебиеттер

- 1 Базы знаний интеллектуальных систем / Т.А. Гаврилова, В.Ф. Хорошевский. – СПб.: Питер, 2000. – 384 с.
- 2 Ефимов Е.И. Решатели интеллектуальных задач. – М.: Наука, 1982. – 320 с.
- 3 Искусственный интеллект: в 3-х кн.: справочник / под ред. Э.В. Попова. – М.: Радио и связь, 1990. – Кн. 1. - 464 с.
- 4 Искусственный интеллект: в 3-х кн.: справочник / под ред. В.Н. Захарова, В.Ф. Хорошевского. – М.: Радио и связь, 1990. – Кн. 3. - 368 с.
- 5 Попов Э.В. Экспертные системы – состояние, проблемы, перспективы // Техническая кибернетика. – М.: Наука, 1989. – №5. – С. 152-161.
- 6 Построение экспертных систем / пер. с англ.; под ред. Ф. Хейеса-Рота, Д. Уотермана, Д. Лената. – М.: Мир, 1987. – 441 с.
- 7 Гаврилова Т.А. Об одном подходе к онтологическому инжинирингу // Новости искусств. интеллекта. - 2005. – №3. – С. 25-30.
- 8 Гуревич И.Б., Труслова Ю.О. Тезаурус и онтология предметной области «Анализ изображений» // Всерос. конф. с междунар. участием «Знания – Онтологии – Теории» (ЗОНТ–09). – Новосибирск: Институт математики им. С.Л. Соболева СО РАН, 2009. – 10 с.
- 9 SPARQL query language for RDF // [http:// www.w3.org/TR/rdf-sparql-query](http://www.w3.org/TR/rdf-sparql-query).
- 10 Хоанг В.К., Тузовский А.Ф. Выбор языка запросов к семантическим данным // Матер. X Всерос. науч.-практ. конф. студентов, аспирантов и молодых учёных «Технологии Microsoft в теории и практике программирования». – Томск, 2013. – С. 117-119.
- 11 SPARQL protocol for RDF // <http://www.w3.org/TR/rdf-sparql-protocol/>.
- 12 Turing A.M. Computing machinery and intelligence // Mind. – 1950. – №59. - P. 433-460.

Р.К.УСКЕНБАЕВА, Н.К.МУКАЖАНОВ

АШЫҚ ҮКІМЕТ ДЕРЕКТЕРІН ӨНДЕУГЕ АРНАЛҒАН АҚПАРАТТЫҚ-АНАЛИТИКАЛЫҚ ЖҮЙЕ АРХИТЕКТУРАСЫН ӨЗІРЛЕУ

(Халықаралық ақпараттық технологиялар университеті, Алматы, Қазақстан)

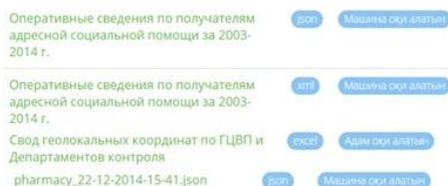
1 Кіріспе

Бұл мақала шеңберінде ашық үкімет порталындағы ашық деректерді жедел аналитикалық өңдеу және деректерді аналитикалық өңдеуге арналған ақпараттық-аналитикалық жүйесінің архитектурасы құрылымы қарастырылады.

Бүгінгі күнде мемлекеттік секторда жинақталған ауқымды деректерді жедел аналитикалық өңдеу негізгі трендтердің бірі болуда. Онда ағымдағы жағдайларды бақылау және талдау нақты уақытта орындалады. Мемлекеттік органдарда басқару шешімдерін қабылдауда деректерді аналитикалық өңдеудің көп жағдайларда қолданылмауының басты себептерінің бірі ондағы ақпараттың басым бөлігі біріктірілмеген жеке-жеке және автоматтандырып өңдеуге жарамсыз түрде болуында. Бұдан басқа, ақпарат сапасына, дұрыстығына және мағұлыматтарды жедел алудың мәселелері кездеседі. Көптеген мемлекеттік мекемелерде деректерді талдау түрлі есеп берулерді және анықтамалық деректерді дайындаудан тұрады. Бұл тапсырманы тиімді шешу үшін арнайы дайындалған ақпараттан қажетті кестелер мен сұлбаларды қолайлы және жылдам алу мүмкіндігі болу қажет. Аталған тапсырманы шешудің бірден бір жолы ретінде мақала ашық деректерді жедел аналитикалық өңдеу жүйесін құру тапсырмасы қарастаралып, жүйе архитектурасы ұсынылады.

2 Деректерді жедел аналитикалық өңдеудің мемлекеттік секторда қолданылуы

Еліміздің дамуындағы басты ұстанымдарының бірі мемлекеттік басқарудың ашықтығы. Осы ұстанымға ақпарат иемденушілердің, соның ішінде мемлекеттік органдардың қызметтерінің ашықтығы мен мөлдірлігін қамтамасыз ету үшін "Электронды үкімет" пен қатар "Ашық үкімет" құру арқылы қол жеткізу. Қазіргі уақытта мемлекеттік органдармен аса үлкен көлемді деректер туындап, сақталынууда. Олардың халыққа маңызды және қол жетімділері ашық деректер порталында жариялануда. Бұл ашық деректер азаматтардың өзбетінше келесі өңдеулер, талдаулар жасауына ұсынылып отыр.



Сурет 1 – Ашық үкімет порталындағы деректердің ұсынылу форматтары

Соңғы жылдарда әлемде ақпарат көлемінің артуына, жаңа сақтау құрылымдараның әзірленуіне байланысты пайда болған жаңа тренд – Big Data-технологияларын мемлекеттік секторларда, электронды үкіметте (e-Government), электронды басқаруларда (e-Governance) қолдану белсенді жүргізілуде. Біздің елемізде де электронды үкімет жүйесінің аппараттық-программалық әзірлемелерін жетілдіріуде заманауи Big Data-технологияларын қолдану арқылы шешу жолдары қарастырылуда [1].

Big Data-технологиялары мүмкіндіктерін пайдаланудың бірден бір жолы ретінде барлық мемлекеттік ведомствалардан жинақталған ашық ақпаратты көлденең масштабталатын таратылған ДБ-ларында сақтау (NoSQL-ДБ, Hadoop). Ары қарай пайдаланушылардың қажеттігіне байланысты өңдеу, қолайлы формаларда ұсыну, талдау және .т.б. тапсырмаларды жүзеге асыру. Ашық деректерден басқа әр бір мемлекеттік

мекемелердің тек өз ішінде ғана пайдаланылатын жабық деректерде бар. Ашық деректерді аналитикалық өңдеудің нәтижелерін тек мемлекеттік қызметтің жетекшілері ғана емес, сонымен қатар республика азаматтарының пайдалануына ұсыну.

3 Ашық деректерді жедел аналитикалық өңдеу жүйесі архитектурасы.

Қазіргі уақытта ақпараттық-аналитикалық жүйелерді құруға арналған Aster MapReduce appliance (Teradata корпорациясы), Oracle Big Data appliance, Greenplum appliance (EMC корпорациясы), SAP HANA appliance және т.б. көптеген платформалар құрылған. Сонымен қатар SAS OLAP Server, Hyperion, Oracle Exadata, Microsoft Analysis Services және т.б. коммерциялық программалық өнімдері мен Mondrian, Pentaho Community, Jedox, olap4j және т.б. ашық кодтағы программалық өнімдері ұсынылуда [1-3].

Жоғарыда келтірілген коммерциялық бағыттағы программалық өнімдердің барлығы дерлік деректерді аналитикалық өңдеудің базалық талаптарын қанағаттандырады. Бірақ, олардың программалық кодтары жабық болуына байланысты әр пайдаланушы, бағдарламалаушы өзінің әзірлемесін ендіре алмайды.

Бұл мақалада қарастырылатын аналитикалық жүйе архитектурасына ашық кодтағы өнімдер алынады. Ашық кодтағы толыққанды программалық өнімдердің бірі Pentaho Community. Pentaho Community-программалық қамтамасы java-ортасында әзірленген.

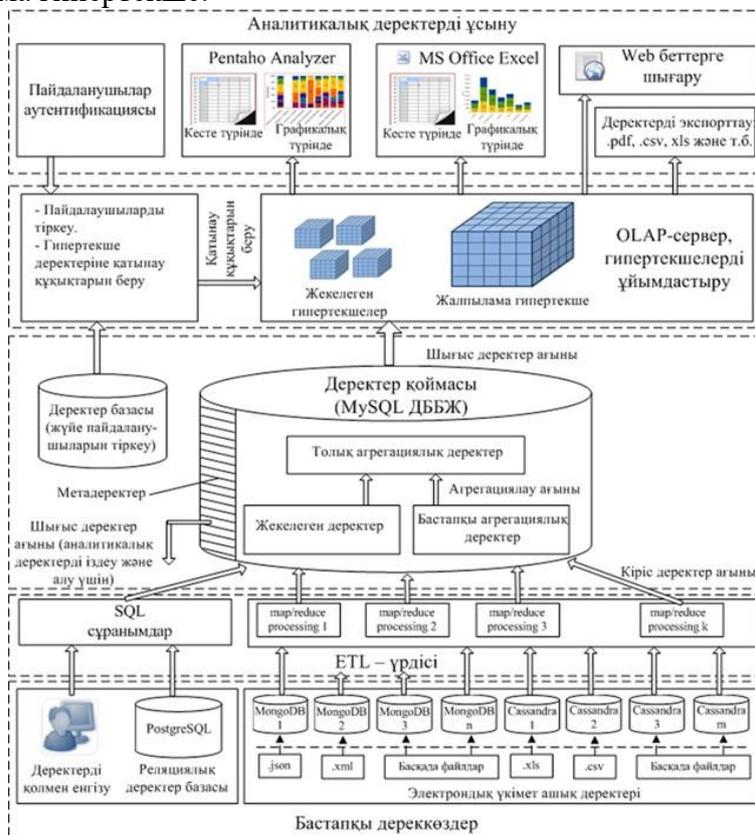
Ашық деректерді жедел аналитикалық өңдеу жүйесі архитектурасы келісдегдей компоненттерден тұрады (2-суретте берілген):

- Аналитикалық деректерді ұсыну құралдары:

1. Pentaho Analyzer;
2. MS Office Excel;
3. .pdf, .csv, .xls және басқа да форматтарына ауыстыру қосымшалары.

- Аналитикалық гипертекшелерді ұйымдастыру модулі:

1. OLAP-сервер;
- а) Жекелеген гипертекшелер;
- б) Жалпылама гипертекше.



Сурет 2 – Ашық деректерді аналитикалық өңдеудің ақпараттық-аналитикалық жүйесі архитектурасы

2. Пайдаланушыларды тіркеу және қатынау құқықтарын беру қосымшасы.

- Деректерді сақтау жүйесі:

1. MySQL ДББЖ құрылған деректер қоймасы.

2. Пайдаланушыларды тіркеуге арналған ДБ.

- Деректерді бірыңғай құрылымға келтіру құралдары: SQL-сұранымдар, Map Reduce - моделінде әзірленген, деректерді бірыңғай құрылымға келтіру қосымшасы.

- Бастапқы дереккөздер: деректерді қолмен енгізуге арналған қосымшалар, реляциялық үлгіде құрылған ДБ (PostgreSQL және т.б.), MongoDB-ДБ, Cassandra -ДБ.

Ведомостваларда деректер реляциялық ДБ-да, excel кестелер, машина оқи алатын форматтарда, түрлі файлдарда берілген. Реляциялық деректерді және кесте түрінде берілген деректерді деректер қоймасына бірден жүктеп аналитикалық талдау жасауға болады. Ал, түрлі файлдар, құжаттарды бірден реляциялық деректер базасына көшіру күрделі тапсырма және көлемді деректерді реляциялық базаға көшіру көп уақыт қажет етеді, сондықтан жартылай құрылымдандырылған және құрылымға келтірілмеген деректерді сақтауға арналған NoSQL ДБ-лары пайдаланылды.

Қорытынды

Бұл мақала ашық үкіметтің "ашық деректерін" аналитикалық өңдеу қажеттігі және аналитикалық өңдеу жүйесі архитектурасын ұсынуға арналды. Мақаланың бастапқы бөлігі деректерді аналитикалық өңдеу жүйелерінің мемлекеттік секторларда пайдаланудың қажеттігі және аналитикалық жүйелерде Big Data технологияларын қолдану айтылды. Мақаланың үшінші бөлімінде ашық үкіметтің деректерін жедел аналитикалық өңдеуге арналған жүйе архитектурасы келтіліді.

Пайдаланылған әдебиеттер тізімі

1 Codd E.F., Codd S.B., and Salley C.T. Providing OLAP (On-line Analytical Processing) to User-Analists: An IT Mandate. - 1993.

2 Vatika Sharma and Meenu Dave. SQL and NoSQL Databases // International Journal of Advanced Research in Computer Science and Software Engineering. – 2012, august. – Vol. 2, issue 8. - P. 21-27.

3 Ralph Kimball and Richard Merz. The Data Webhouse Toolkit: Building the Web-Enabled Data Warehouse. - New York: John Wiley & Sons, 2000. - 417 p.

4 Uskenbayeva R.K., Kurmangaliyeva B., Mukazhanov N.K. Development of multidimensional model of data for information and analytical decision-making support system // Computer modelling and new technologies. - 2014. – №2(18). – P. 170-175.

5 Barbara Ubaldi Open Government Data. Open Government Data: Towards Empirical Analysis of Open Government Data Initiatives // OECD Working Papers on Public Governance. - OECD Publishing, 2013. – №22. – P. 61.

МЕТОДЫ ОБРАБОТКИ ЗАЯВОК В СИСТЕМЕ Э-КОММЕРЦИИ

(Международный Университет Информационных Технологий, Алматы, Казахстан)

Введение

Э-коммерция состоит из множества процессов, начиная с процессов этапа закупки и завоза товаров на склад, заканчивая процессами доставкой товаров клиентам и т.д. Поэтому, для правильного и эффективного усовершенствования и развития системы э-коммерции, в первую очередь требуется, исходя из системных позиций, формализация всех процессов э-коммерции, формулировка задач оптимального выполнения этих процессов, а также поиск эффективного алгоритма их решения с учетом взаимосвязи.

Для того, чтобы все эти этапы выполнялись (работали) оптимально и бесперебойно, необходимо рассматривать все процессы электронной коммерции комплексно, охватывая все специализированные процессы, которые входят в состав общего бизнес-процесса.

В данной работе, исходя из выше отмеченных подходов исследуются процессы первичной обработки заявок: прием, регистрация, проверка надежности подаваемых заявок, формализация содержания.

1. Общая задача процессов выполнения заявок

В работе [1] проведен анализ проблемной области э-коммерции и сформулирована общая задача оптимального ведения э-коммерции, а именно процессы выполнения заявок при заданном варианте инфраструктуры э-коммерции. Далее, определены этапы решения данной общей задачи. Здесь рассматривается этап подготовки заявок для формализованной обработки, которая изложена в [2].

Формализация процессов приема и содержания заявок.

В текущий момент t подается (или прием) проект-заявка на приобретение товаров покупателем (заявитель). Подача заявки осуществляется путем заполнения заявителем формы заявки э-магазина. Если в тексте, при заполнении, имеются орфографические ошибки и неточности, то их устраняют вместе с заявителем в диалоговом режиме.

Прежде чем принимать и регистрировать, проект-заявки проходят проверку на надежность заявителя. Один из способов оценки надежности покупателя является анализ поведения покупателя по историческим данным, как раньше заказывал данный покупатель и как он себя вел.

Надежность покупателя или заявителя также можно определить по историческим данным из других ресурсов или Интернета путем поиска данных о корректности и надежности данного заявителя. Данный процесс напоминает банковский процесс проверки надежности будущего заемщика в процессе оформления.

Отсекаем те проект-заявки, которые невозможно выполнить или, если заявитель ненадежный, например, если есть сомнения, что заявитель дает ложные реквизиты или неплатежеспособен. Выполняемые проект-заявки назовем перспективными и принимаем как заявки.

После приема текущей проект-заявки $Or_{тек}$, следующие данные регистрируются путем занесения входных данных в базу регистрации заявок (Рисунок 1):

- дата заказа (с идентификатором или кодом D),
- время подачи заявки (с идентификатором или кодом T),
- порядковый номер цикла обслуживания покупателя (с идентификатором или кодом L),
- ФИО заявителя (с идентификатором или кодом F),
- порядковый номер заявителя (с идентификатором или кодом J),

- порядковый номер подачи заявки (с идентификатором или кодом I),
- адреса доставки товаров заявки (с идентификатором или кодом A, если товар доставить по нескольким адресам, то $\{A_1, A_2, A_3, \dots\}$),
- заказываемые товары в заявке (с идентификатором или кодом Tv, если заявлено несколько видов товаров, то $\{Tv_1, Tv_2, Tv_3, \dots\}$).

Таблица 1. Таблица 1 базы регистрации заявок

п/п ID код заявки	дата заказа	время подачи	порядковый номер цикла обслуживания	ФИО заявителя	порядковый номер заявителя	порядковый номер подачи заявки	адреса доставки товаров в заявке	заказываемые товары в заявке (код пакета заказанных товаров)	Код заявок D.T.L.J.I
	26.07.20016	15 ³⁰							D.T.L.J.I
	26.07.20016	15 ⁴⁵							D.T.L.J.I

Таблица 2. Таблицы 1 базы регистрации заявок

п/п	ID код заявки	код пакета заказанных товаров	Товар №1	Адрес доставки Товар №1	Товар 2	Адрес доставки Товар №2	Товар 3	Адрес доставки Товар №3				Товар k	Адрес доставки Товар №k

Отметим, что в базе регистрации хранится еще E-mail и телефон сотовый клиентов.

Базы регистрации заявок таблица 1 сохраняет основные атрибуты, принятых заявок.

Таблица 2 - Базы регистрации заявок. В ней запоминают характеристики заказанных товаров в количестве из k видов товаров. По каждому товару в заявке указывается вид товара, адреса их доставки.

Отметим, что может даваться максимальное количество видов товаров в одной заявке, например, значение $k=10$ в одной заявке.

После приема и регистрации текущая проект-заявка $Or_{тек}$ приобретает статус заявки и присваивается регистрационный номер Or_i , состоящий из следующего кода:

$ID(Or_i) = [\text{код дата}] [\text{код время}] [\text{код номер цикла обслуживания}] [\text{код порядковый номер заявителя}] [\text{код порядковый номер заявки в цикле}] = DTLJI$ (или D.T.L.J.I),

где D – код дата заказа, T – код время подачи заявки, L – код порядковый номер цикла обслуживания покупателя, J – код порядковый номер заявителя, I – код порядковый номер подачи заявки. Код признака занесем код заявок в базу (БДЗ) согласно порядку поступления (Рисунок 1). В данную базу данных заявок происходит заполнение по мере обработки заявок.

п/п	ID код заявки	Код заявок D.T.L.J.I	
	ID1	D1.T1.L1.J1.I1	
	ID2	D2.T2.L2.J2.I2	

Рисунок 1. База данных код заявок.

Приняты следующие правила или допущения подачи заявки:

- 1) Один и тот же человек (покупатель) может подавать несколько заявок.
- 2) В заявке несколько видов товаров в неограниченном количестве (но предполагаем в пределе разумности, все таки не официально существует верхний предел).
- 3) Можно на каждый товар оформить заявку.
- 4) Прием заявки осуществляется 24 часа без выходных.

Прием заявок производится в течение времени $T = \Delta t$, в частности $\Delta t = 1$ час.

Накопленные заявки за интервал времени Δt будет таков:

$$Or = \{Or_i\}, i=1, I, \quad (1)$$

где Or_i – заявка с i -ым номером, поданная в текущем цикле обслуживания заявителя (покупателя) или порядковый номер поступления в текущем цикле обслуживания покупателей.

Если считать один покупатель подает одну заявку, то порядковый номер покупателя и номер заявки совпадает. Тогда, один покупатель в одной заявке заявляет несколько видов товаров с неограниченным объемом (весом, количеством).

Совокупность заявок (1) составляет пакет заявок для одного цикла обслуживания заявок соответственно заявителей или покупателей.

Далее, содержание каждой заявки переводится в признаковое пространство, т.е. каждая заявка представляется в виде набора признаков. Таким образом, атрибуты и характеристики каждой заявки в виде вектора признаков, а всех заявок в пакете текущего цикла обслуживания заявителя в виде:

$$\chi(Or_i) = \{\chi_{ij}\}, i=1, I, j=1, J_i, \quad (2)$$

где $\chi(Or_i) = \{\chi(Or_{ij})\}$ или $\chi(Or_i) = \{\chi_{ij}\}, j=1, J$ – атрибуты и характеристики заявок Or_i .

Признаки по содержанию поданной заявки $\chi(Or_i) = \{\chi_{ij}\}$ таковы:

- χ_{i1} – имя (ФИО) или название организации/предприятия покупателя.
- χ_{i2} – статус клиента: юридический, физический $St(Kl_i)$ – клиент товаров i -ой заявки. Статус клиента: $St(Kl_i) = \{1 - \text{если юридический}, 0 - \text{если физический}\}$.
- χ_{i3} – категория покупателя: частное или юридическое лицо.
- χ_{i4} – признаки заказанных товаров. В связи с тем, что покупатель заказывает несколько видов товаров $\chi_{i4} = \{\chi_{i4k}\}, k=1, K_i$.
- χ_{i5} – признаки количества или веса и объема товаров. Для каждого вида следует показать количество или вес товаров $\chi_{i5} = \{\chi_{i5k}\}, k=1, K_i$.
- χ_{i6} – дата и время заказа.
- χ_{i7} – форма оплаты: с предоплатой, после получения клиентом товаров, частичная предплата, частично после получения клиентом товаров.
- χ_{i8} – форма доставки: само вывоз, с доставкой.
- χ_{i9} – признаки адресов доставки товаров клиенту $Adr(g_i)$. Возможно доставить различные товары по разным адресам: $\chi_{i9} = \{\chi_{i9l}\}, l=1, L_i$.
- χ_{i10} – признаки времени доставки товаров клиенту $Tim(g_i)$. Возможно доставить различные товары в различное время: $\chi_{i10} = \{\chi_{i10\theta}\}, \theta=1, \Theta_i$.

Таким образом, формализованное представление заявки или модель заявки (заявки с выше установленными данными)

$$Or = \{Or_i\}, i=1, I$$

$$Or_{tek} = Or_i$$

$$ID(Or_i) = DTLJI,$$

$$\chi(Or_i) = \{\chi_{ij}\}, i=1, I, j=1, J_i, \quad (3)$$

$$A = \{A_1, A_2, A_3, \dots\},$$

$$Tv = \{Tv_1, Tv_2, Tv_3, \dots\},$$

поступает на обработку, т.е. дальнейшая обработка заявок проводится по признакам.

Литература

1 Касымова А.Б., Куандыков А.А. Системно-объектный анализ проблемной области э-коммерции // Журнал Промышленность Казахстана. – 2016. - №3(96). – С. 29-31.

2 Касымова А.Б. Формализация заявок для обработки в системе э-коммерции // Вестник ВКГТУ им.Серикбаева. – 2016. - №4.

ДОРОЖНАЯ КАРТА МЕТОДОЛОГИИ ASAP

(Магистрант, Евразийский национальный университет им. Л.Н. Гумилева, г. Астана)

Методология Accelerated SAP 8 (далее ASAP) - используется для планирования, управления и сопровождение проектов по установке систем SAP. Цель методологии - оптимизация времени, ресурсов и процессов для наиболее эффективного внедрения SAP решений. Методическая структура ASAP предоставляется конечным пользователям в четырех вариантах для поддержки различных стратегий развертывания.

- ASAP 8 for Simplified Rapid Deployment Solution Experience - это упрощенное решение для быстрого развертывания. Он сочетает в себе способность изучать возможности решения SAP, выбирать подходящую предварительно собранную RDS в качестве отправной точки для проекта и реализовывать решение в бизнесе с помощью подхода структурированной реализации - методологии ASAP 8.

- ASAP 8 for Assemble to Order Project (проект единичного производства, A2O) – эта методология основана на опыте быстрого развертывания решений, реализации гибких проектов и возможности комбинирования нескольких служб RDS и инженерных услуг в одном проекте. Эта методология предоставляет контент, инструменты и передовую практику, которые помогают консультантам обеспечивать последовательные и успешные результаты в различных отраслях промышленности и в среде клиентов. ASAP for Assemble-to-Order строится на прочной основе ASAP 8, используя подход итеративной реализации, представленный в методологии SCRUM, широко признанную гибкую структуру и способность использовать готовые решения на основе RDS или SAP Best Practices.

- Методология Agile ASAP 8 (Agile ASAP 8 Methodology) - новая, практическая методология реализации, которая позволяет реализовать функциональные возможности в коротких итерационных циклах. В каждом цикле команда реализует наиболее ценные и важные функции. Это позволяет быстрее генерировать результаты, получить мгновенное представление о ценности, повысить гибкость внедрения и улучшить контроль выполнения.

- Стандартная методология ASAP 8 (Standard ASAP 8 Methodology) - использует дисциплинированный подход к управлению проектами, управлению организационными изменениями, управлению решениями и другим дисциплинам, применяемым при внедрении решений SAP. ASAP 8 построен на основе модели SAP Advanced Delivery Management и поддерживает проектные команды с шаблонами, инструментами, вопросниками и контрольными перечнями, включая путеводители и ускорители. ASAP 8 позволяет компаниям использовать возможности ускоренных функций и инструментов, уже встроенных в решения SAP.

В таблице 1 предоставлена дорожная карта методологии ASAP с описаниями этапов четырех вариантов развертывания.

Методология ASAP предоставляет следующие преимущества:

- Она обеспечивает непротиворечивость, уменьшает уровень сложности и повышает качество путем создания и поставки сервисов на одной общей платформе.
- ASAP является масштабируемой методологией с поддержкой всех видов поставки сервисов от развертывания RDS до крупномасштабных проектов внедрения.
- Методология ASAP реализует директивный и комплексный подход. ASAP предлагает управляемые процедуры для проектных групп, результаты для руководителей проектов и акселераторы для всех пользователей.
- ASAP предлагает инновации в области поставки с ориентацией на Agile-развертывание, визуализацию приложений и на использование процессов развертывания на базе облаков. В то же время эта методология позволяет применять практики поставки рентабельных проектов в формате Q-Gate по проекту для традиционных проектов и RDS. В основе ASAP лежит сбор и повторное использование знаний. ASAP в простом формате предоставляет доступ к знаниям, полученным в ходе реализации тысяч проектов внедрения решений SAP.

Таблица 1. Дорожная карта методологии ASAP

Этапы	Simplified Rapid Deployment Solution Experience/ Assemble to Order Project	Agile ASAP 8 Methodology	Standard ASAP 8 Methodology
1	2	3	4
Подготовка проекта	Этап предусматривает первоначальное планирование и подготовку к проекту. Каждый проект имеет свои уникальные цели, масштаб и приоритеты. Результаты, описанные на этом этапе, помогают в эффективном завершении начальных этапов, также на этом этапе готовятся установки гибкого управления проектом, плана проекта и графика проекта.		
Концептуальный проект	Проверка области видимости - целью этапа является достижение общего понимания того, как компания намеревается запустить SAP для поддержки своего бизнеса. Основное внимание в нем уделяется быстрой настройке среды, доступной для проведения валидационного семинара с бизнес-пользователями, для подтверждения объема и определения требований к дельтам, которые будут реализованы на следующем этапе с целью улучшения базового уровня, обеспечиваемого предварительно собранной RDS.	Lean blueprint - Цель этого этапа - достичь общего понимания того, как компания намеревается использовать SAP для поддержки своего бизнеса. Основное внимание в нем уделяется быстрой настройке среды проверки решений на семинарах по валидации с бизнес-пользователями для подтверждения области и определения требований к дельтам, которые будут реализованы на следующем этапе, чтобы улучшить базовую сборку системы.	Бизнес-план - цель этапа является достижение общего понимания того, как компания намерена использовать SAP для поддержки своего бизнеса. В стандартной методологии ASAP 8 результатом является «Бизнес-план», подробная документация результатов, собранных во время семинаров по требованиям.
Реализация	Цель - реализовать все требования к дельта-процессам бизнес-процессов, определенные на предыдущем этапе. Команда конфигурирует, разрабатывает, тестирует и документирует решение в виде серии последовательных итераций - наиболее ценных функций в первую очередь. До того, как решение будет выпущено на следующий этап, он будет полностью интегрирован и проверен конечными пользователями.		
Заключительная подготовка	Цель - завершить работу по переключению (включая техническое и нагрузочное тестирование, обучение конечных пользователей, управление системой и репетиционные операции), чтобы завершить вашу готовность к работе. Заключительная фаза подготовки также служит для решения всех остальных важных проблем. После успешного завершения этого этапа вы готовы запустить свой бизнес в своей действующей SAP-системе.		
Продуктивный старт и поддержка	Целью - переход от ориентированной на проект среды предварительного производства к оперативной производственной деятельности и постоянная поддержка бизнес-пользователей в целях содействия их переходу в новую среду.		
Эксплуатация и оптимизация	Цель этого этапа заключается в точной настройке стандартов жизненного цикла приложений, процессов и процедур, установленных во время проекта, и их согласования с операционными потребностями. Центральной оперативной платформой является SAP Solution Manager, которая использует документированное решение для системных операций.		

*Подписано к печати 29.05.2017 г. Бумага офсетная №1.
Формат 60X84 1/8 Объем 32,3 усл.печ.л.
Тираж 50 экз. Заказ0168
Отпечатано в типографии ТОО «Мастер ПО»
010008, г. Астана, ул. Пушкина, 15-76
(ул. Кенесары, 93-76) тел.: 8 (7172) 223-418
e-mail.: masterpo08@mail.ru*