

Учреждение образования  
«Белорусский государственный университет  
информатики и радиоэлектроники»

Государственное учреждение  
«Администрация Парка высоких технологий»

# **Открытые семантические технологии проектирования интеллектуальных систем**

**OSTIS-2015**

## **Open Semantic Technologies for Intelligent Systems**

МАТЕРИАЛЫ  
V МЕЖДУНАРОДНОЙ  
НАУЧНО-ТЕХНИЧЕСКОЙ КОНФЕРЕНЦИИ

*(Минск, 19–21 февраля 2015 года)*

УДК 004.822+004.89-027.31  
ББК 32.813-02+73  
О-83

Редакционная коллегия :

*В. В. Голенков (отв. ред.), Л. С. Глоба, Н. А. Гулякина, И. В. Ефименко, О. П. Кузнецов, Б. М. Лобанов,  
Д. Ш. Сулейманов, А. А. Харламов, В. Ф. Хорошевский*

*Организаторы конференции:*

Учреждение образования «Белорусский  
государственный университет информатики и  
радиоэлектроники»  
Российская ассоциация искусственного интеллекта  
Государственное учреждение «Администрация Парка  
высоких технологий» (Республика Беларусь)  
Научно-технологическая ассоциация «Инфопарк»  
Объединённый институт проблем информатики  
Национальной академии наук Беларуси  
Научно-исследовательский институт «Прикладная  
семиотика» АН РТ  
Институт информатизации образования Российской  
академии образования

Международная ИТ-компания «Itransition»  
Компания «Melesta»  
Компания «Qulix Systems»  
Digital-агентство «ARTOX media»  
Компания «ВирусБлокАда»  
Компания «LogicNow»  
ЧУП «Андерсан-софт»  
ООО «Прикладные системы»  
ООО «АйтиРекс Групп»  
ЧУП «Айтими́нт»  
ЧП «Нейролаб»  
ИООО «Седон БЛР»

*Техническая и информационная поддержка:*

Международный журнал «Программные продукты и  
системы»  
Научный журнал «Информатика»

Научно-практический журнал для специалистов  
«Электроника ИНФО»  
Научный журнал «Онтология проектирования»

*Издание осуществлено по заказу государственного учреждения  
«Администрация Парка высоких технологий»*

**О-83** **Открытые семантические технологии проектирования интеллектуальных систем = Open Semantic Technologies for Intelligent Systems (OSTIS-2015) :** материалы V междунар. науч.-техн. конф. (Минск, 19–21 февраля 2015 года)/ редкол. : В. В. Голенков (отв. ред.) [и др.]. – Минск : БГУИР, 2015. – 620 с.  
ISBN 978-985-543-034-7.

Сборник включает прошедшие рецензирование статьи V международной научно-технической конференции «Открытые семантические технологии проектирования интеллектуальных систем».

Сборник предназначен для преподавателей высших учебных заведений, научных сотрудников, студентов, аспирантов, магистрантов, а также для специалистов предприятий в сфере проектирования интеллектуальных систем.

Материалы сборника одобрены Программным комитетом OSTIS-2015 и печатаются в виде, представленном авторами.

**УДК 004.822+004.89-027.31**  
**ББК 32.813-02+73**

**ISBN 978-985-543-034-7**

© УО «Белорусский государственный  
университет информатики  
и радиоэлектроники», 2015

# СОДЕРЖАНИЕ

<b>ПРЕДИСЛОВИЕ</b>	23
ПРОБЛЕМА ПОНИМАНИЯ: НАСТОЯЩЕЕ И БУДУЩЕЕ ИСКУССТВЕННОГО ИНТЕЛЛЕКТА Тарасов В.Б.	25
СЕМАНТИЧЕСКАЯ ТЕХНОЛОГИЯ КАРТИРОВАНИЯ СЕМАНТИЧЕСКИХ ТЕХНОЛОГИЙ (Наукометрический анализ конференций OSTIS) Хорошевский В.Ф., Ефименко И.В.	43
СЕМАНТИЧЕСКАЯ ТЕХНОЛОГИЯ КОМПОНЕНТНОГО ПРОЕКТИРОВАНИЯ СИСТЕМ, УПРАВЛЯЕМЫХ ЗНАНИЯМИ Голенков В.В., Гулякина Н.А.	57
СРЕДСТВА ПОДДЕРЖКИ КОМПОНЕНТНОГО ПРОЕКТИРОВАНИЯ СИСТЕМ, УПРАВЛЯЕМЫХ ЗНАНИЯМИ Шункевич Д.В., Давыденко И.Т., Корончик Д.Н., Жуков И.И., Паркалов А.В.	79
РЕАЛИЗАЦИЯ ПЛАТФОРМЫ ДЛЯ WEB-ОРИЕНТИРОВАННЫХ СИСТЕМ, УПРАВЛЯЕМЫХ ЗНАНИЯМИ Корончик Д. Н.	89
МЕТОДИКА КОМПОНЕНТНОГО ПРОЕКТИРОВАНИЯ СИСТЕМ, УПРАВЛЯЕМЫХ ЗНАНИЯМИ Шункевич Д.В., Давыденко И.Т., Корончик Д.Н., Губаревич А.В., Борискин А.С.	93
МОДЕЛИ И АЛГОРИТМЫ ИНТЕГРАЦИИ ЗНАНИЙ НА ОСНОВЕ ОДНОРОДНЫХ СЕМАНТИЧЕСКИХ СЕТЕЙ Ивашенко В.П.	111
ПРЕДСТАВЛЕНИЕ СЕМАНТИЧЕСКИХ СЕТЕЙ И АЛГОРИТМЫ ИХ ОРГАНИЗАЦИИ И СЕМАНТИЧЕСКОЙ ОБРАБОТКИ НА ВЫЧИСЛИТЕЛЬНЫХ СИСТЕМАХ С МАССОВЫМ ПАРАЛЛЕЛИЗМОМ Ивашенко В.П., Вереник Н.Л., Гирель А.И., Сейткулов Е.Н., Татур М.М.	133
КОМПОНЕНТНАЯ АРХИТЕКТУРА ИНТЕЛЛЕКТУАЛЬНОЙ СИСТЕМЫ КОНСУЛЬТАЦИОННОГО ОБСЛУЖИВАНИЯ И ОБУЧЕНИЯ РАЗРАБОТЧИКОВ ПРОГРАММ Пивоварчик О.В.	141
ПРИНЦИПЫ ПОСТРОЕНИЯ КЛИНИЧЕСКОЙ СИСТЕМЫ ПОДДЕРЖКИ ПРИНЯТИЯ РЕШЕНИЙ НА ОСНОВЕ ТЕХНОЛОГИИ OSTIS Каешко А.И., Маргунов Е.А.	149
ПРОЕКТИРОВАНИЕ БАЗЫ ЗНАНИЙ ИНТЕЛЛЕКТУАЛЬНОЙ СПРАВОЧНОЙ СИСТЕМЫ ПО АЛГЕБРЕ Шарипбай А.А., Омарбекова А.С., Нургазинова Г.Ш.	157
ПУТИ РАЗВИТИЯ ИСКУССТВЕННОГО ИНТЕЛЛЕКТА В КАЗАХСТАНЕ Шарипбай А.А.	161
СОЗДАНИЕ ОНТОЛОГИЧЕСКОЙ МОДЕЛИ ЛОКАЛЬНОЙ КОМПЬЮТЕРНОЙ СЕТИ Шарипбай А.А., Аскарова С.А., Муканова А.С.	165

ОНТОЛОГИЧЕСКАЯ МОДЕЛЬ ПРОЦЕССА ОБЕСПЕЧЕНИЯ ИНФОРМАЦИОННОЙ БЕЗОПАСНОСТИ	169
Ниязова Р.С., Буданова Н.	
КОНЦЕПЦИЯ ИНСТРУМЕНТАЛЬНОЙ ПЛАТФОРМЫ ИНФОРМАЦИОННОЙ БЕЗОПАСНОСТИ В СРЕДЕ ОБЛАЧНЫХ ВЫЧИСЛЕНИЙ С ИСПОЛЬЗОВАНИЕМ ИНТЕЛЛЕКТУАЛЬНЫХ ТЕХНОЛОГИЙ	173
Вишняков В.А., Гондаз Саз М.М., Моздуоани Шираз М.Г.	
ОНТОЛОГИЧЕСКИЙ ПОДХОД К СОЗДАНИЮ НАУЧНЫХ ИНТЕРНЕТ-РЕСУРСОВ	177
Загорулько Ю.А., Загорулько Г.Б.	
РАЗРАБОТКА И ГЕНЕРАЦИЯ ВИРТУАЛЬНЫХ СРЕД НА ОСНОВЕ СЕМАНТИЧЕСКИХ ПРЕДСТАВЛЕНИЙ	183
Грибова В.В., Федорищев Л.А.	
КЛАССИФИКАЦИЯ ДЛЯ ЗАДАЧ, ВЫЯВЛЯЕМЫХ ПРИ СИСТЕМНОМ АНАЛИЗЕ ИНТЕЛЛЕКТУАЛЬНОЙ ДЕЯТЕЛЬНОСТИ	187
Шалфеева Е.А.	
ТЕХНОЛОГИЯ РАЗРАБОТКИ РЕШАТЕЛЕЙ ЗАДАЧ ИНТЕЛЛЕКТУАЛЬНЫХ СИСТЕМ С ИСПОЛЬЗОВАНИЕМ ИНСТРУМЕНТАЛЬНЫХ СЕРВИСОВ ОБЛАЧНОЙ ПЛАТФОРМЫ IASRAAS	193
Крылов Д.А., Москаленко Ф.М., Тимченко В.А.	
МЕТОДЫ И СРЕДСТВА СИТУАЦИОННОГО УПРАВЛЕНИЯ В ЭНЕРГЕТИКЕ НА ОСНОВЕ СЕМАНТИЧЕСКОГО МОДЕЛИРОВАНИЯ	199
Массель Л.В., Массель А.Г.	
ОТ ИНТЕЛЛЕКТУАЛЬНЫХ ПОРТАЛОВ К ПОРТАЛАМ ЗНАНИЙ: О РОЛИ ИНТЕЛЛЕКТУАЛЬНЫХ СЕРВИСОВ	205
Страхович Э.В., Власов С.А., Гаврилова Т.А.	
ОБЗОР СИСТЕМ КОЛЛАБОРАТИВНОЙ ФИЛЬТРАЦИИ	209
Шереметова Е.И.	
МЕТОДЫ И ИНСТРУМЕНТЫ РАБОТЫ С БИЗНЕС-ПРАВИЛАМИ	215
Савин А.М.	
СИСТЕМЫ УПРАВЛЕНИЯ ЗНАНИЯМИ В АТОМНОЙ ПРОМЫШЛЕННОСТИ	221
Костебелова В.К.	
СЦЕНАРНЫЙ ПОДХОД ПРИ ИССЛЕДОВАНИИ ДИНАМИКИ ИНФОРМАЦИОННЫХ ПОТОКОВ В СЕТИ ИНТЕРНЕТ	225
Додонов А.Г., Ландэ Д.В., Бойченко А.В.	
ПОДХОД К ФОРМИРОВАНИЮ СЛОЖНОГО ИНЖЕНЕРНОГО РАСЧЕТА НА ПОРТАЛАХ	231
Новогрудская Р.Л., Глоба Л.С.	
МЕТАГРАФЫ КАК ОСНОВА ДЛЯ ПРЕДСТАВЛЕНИЯ И ИСПОЛЬЗОВАНИЯ БАЗ НЕЧЕТКИХ ЗНАНИЙ	237
Глоба Л. С., Терновой М. Ю., Штогрин Е. С.	

ДЕТЕКТИРОВАНИЕ ИЗМЕНЕНИЙ В XML-ПОДОБНЫХ СТРУКТУРАХ ДОКУМЕНТОВ	241
Глоба Л.С., Молчанов Ю.Н.	
ИНТЕГРАЦИЯ КОРПОРАТИВНЫХ ДАННЫХ НА ОСНОВЕ WIKI-СИСТЕМ В УСЛОВИЯХ СЛАБОЙ СВЯЗАННОСТИ ИСТОЧНИКОВ	249
Галушка И.Н., Оксанич И.Г., Щербак С.С.	
СЕМАНТИЧЕСКИЙ ПОДХОД К ИНТЕГРАЦИИ ОБЪЕКТОВ WEB OF THINGS	255
Рогушина Ю.В., Гладун А.Я.	
РАЗРАБОТКА СРЕДСТВ ПЕРСОНИФИКАЦИИ ИНТЕЛЛЕКТУАЛЬНЫХ WEB- ПРИЛОЖЕНИЙ	265
Рогушина Ю.В.	
ОБЗОР И АНАЛИЗ МЕТОДОЛОГИЙ И МЕТОДОВ ПОСТРОЕНИЯ ОНТОЛОГИЙ	271
Хала Е.А.	
СПЕЦИФИКАЦИЯ ТЕМПОРАЛЬНЫХ ОТНОШЕНИЙ В СИСТЕМЕ «БИНАРНАЯ МОДЕЛЬ ЗНАНИЙ»	275
Плесневич Г.С., Нгуен Тхи Минь Ву	
ОТ ТЕРМИНОЛОГИЧЕСКИХ СЕТЕЙ К ТОЛКОВЫМ СЛОВАРЯМ	281
Мальковский М.Г., Соловьев С.Ю.	
ОБ ОДНОМ ПОДХОДЕ ТЕМАТИЧЕСКОГО МОДЕЛИРОВАНИЯ ТЕКСТОВ НА ЕСТЕСТВЕННОМ ЯЗЫКЕ	285
Чан Ван Ан	
ОБОСНОВАНИЕ МОДЕЛИРОВАНИЯ ПРОЦЕССОВ ПРАВДОПОДОБНЫХ РАССУЖДЕНИЙ	291
Найденова К. А.	
ЭВРИСТИЧЕСКИЙ МЕТОД УДОВЛЕТВОРЕНИЯ ОГРАНИЧЕНИЙ НА ОСНОВЕ ИХ МАТРИЧНОГО ПРЕДСТАВЛЕНИЯ	297
Зуенко А.А., Очинская А.А.	
СЕТЕВАЯ МОДЕЛЬ РАСПРОСТРАНЕНИЯ АКТИВНОСТИ В СРЕДЕ ГЕТЕРОГЕННЫХ АВТОМАТОВ	303
Жилякова Л.Ю.	
СИСТЕМА ПРИОРИТЕТНОГО ДВИЖЕНИЯ НА ПЕРЕКРЕСТКАХ «ЗЕЛЕНАЯ ВОЛНА» ДЛЯ ТРАНСПОРТНЫХ СРЕДСТВ ОПЕРАТИВНОГО НАЗНАЧЕНИЯ	309
Согоян А.Л., Шуть В.Н.	
ПРОГНОЗ РАЗВИТИЯ ИНФОРМАТИЗАЦИИ ОТЕЧЕСТВЕННОГО ОБРАЗОВАНИЯ	315
Роберт И.В.	
ИНТЕЛЛЕКТУАЛЬНЫЕ ТЕХНОЛОГИИ ОБУЧЕНИЯ В ИНФОРМАЦИОННО- ОБРАЗОВАТЕЛЬНОМ ПРОСТРАНСТВЕ	325
Тельнов Ю.Ф., Трембач В.М.	
ОПРЕДЕЛЕНИЕ УРОВНЯ УСВОЕННЫХ ЗНАНИЙ ПО ОБУЧАЮЩЕМУ КУРСУ, ПРЕДСТАВЛЕННОМУ СЕМАНТИЧЕСКОЙ СЕТЬЮ	331
Янковская А.Е., Шурыгин Ю.А., Ямшанов А.В., Кривдюк Н.М.	

WOLFRAM МАТЕМАТИКА СРЕДСТВА И ТЕХНОЛОГИИ РАЗРАБОТКИ ИНТЕЛЛЕКТУАЛЬНЫХ ОБУЧАЮЩИХ СИСТЕМ Таранчук В.Б.	339
РАЗРАБОТКА КОМПЛЕКСНОЙ СЕМАНТИЧЕСКОЙ МЕТОДИКИ К КОМПОНЕНТНОМУ ПРОЕКТИРОВАНИЮ СЛОЖНЫХ ПРОЕКТНО- ПРОМЫШЛЕННЫХ СРЕД ОБУЧЕНИЯ (ППСО) Афанасьев А.Н., Войт Н.Н.	347
ОНТОЛОГИЧЕСКИЙ ПОДХОД К АНАЛИЗУ КОМПЕТЕНЦИЙ ВЫПУСКНИКОВ ВУЗОВ Шеркунов В.В.	351
НЕЙРОСЕТЕВАЯ МОДЕЛЬ ПРОЦЕССА ПРОФЕССИОНАЛЬНОГО ОБУЧЕНИЯ МОЛОДЫХ СПЕЦИАЛИСТОВ Федяев О.И.	357
МЕТОДЫ НЕЧЕТКОГО РЕГУЛИРОВАНИЯ В ВОПРОСАХ УПРАВЛЕНИЯ ОБРАЗОВАТЕЛЬНЫМ ПРОЦЕССОМ Нишанов А.Х., Дощанова М.Ю., Мирзаев Д.А.	365
ИНТЕЛЛЕКТУАЛЬНЫЕ СРЕДСТВА АВТОМАТИЗАЦИИ УПРАВЛЕНИЕМ ОБУЧЕНИЯ Бабамухамедова М.З., Дощанова М.Ю., Джангазова К. А.	369
КОНЦЕПТУАЛЬНОЕ ЭКСПЕРИМЕНТИРОВАНИЕ В ПРОЕКТИРОВАНИИ КОНФИГУРИРУЕМЫХ ШАБЛОНОВ АВИАЦИОННЫХ ДЕТАЛЕЙ Соснин П.И., Чоракаев О.Э.	373
ОНТОЛОГИЯ ПРОЕКТИРОВАНИЯ ШАБЛОНОВ АВИАЦИОННЫХ ДЕТАЛЕЙ Гришин М.В., Ларин С.Н., Соснин П.И.	381
СЕМАНТИЧЕСКИЙ АНАЛИЗ ДИАГРАММАТИЧЕСКИХ МОДЕЛЕЙ В ПРОЕКТИРОВАНИИ СЛОЖНЫХ АВТОМАТИЗИРОВАННЫХ СИСТЕМ Афанасьев А.Н., Гайнуллин Р.Ф., Афанасьева Т.В.	385
ПРИМЕНЕНИЕ НАВИГАЦИОННОЙ СТРУКТУРЫ ЭЛЕКТРОННОГО АРХИВА ПРОЕКТНОЙ ОРГАНИЗАЦИИ В ЗАДАЧАХ ИНТЕЛЛЕКТУАЛЬНОГО АНАЛИЗА ТЕХНИЧЕСКИХ ДОКУМЕНТОВ Субхангулов Р.А., Филиппов А.А.	389
ОЦЕНКА ТЕРМИНОЛОГИЧНОСТИ ЛЕКСИЧЕСКИХ ЕДИНИЦ НА ОСНОВЕ ОНТОЛОГИИ ПРЕДМЕТНОЙ ОБЛАСТИ Андреев И.А., Башаев В.А., Клейн В.В., Мошкин В.С., Ярушкина Н.Г.	395
МЕТОДИКИ ПОСТРОЕНИЯ НЕЧЕТКИХ ОНТОЛОГИЙ СЛОЖНЫХ ПРЕДМЕТНЫХ ОБЛАСТЕЙ Мошкин В.С., Ярушкина Н.Г.	401
ОНТОЛОГИЧЕСКИЙ ПОДХОД К ФОРМИРОВАНИЮ ПРОЕКТНЫХ ЗАПРОСОВ ИНТЕЛЛЕКТУАЛЬНОГО АГЕНТА Наместников А.М., Субхангулов Р.А.	407

ОПЫТ СОЗДАНИЯ СРЕДСТВ СЕМАНТИЧЕСКОГО МОДЕЛИРОВАНИЯ И ПРОЕКТИРОВАНИЯ НА МАССОВОЙ ПРОГРАММНОЙ ПЛАТФОРМЕ Смирнов С.В.	413
РЕАЛИЗАЦИЯ УДАЛЕННОГО УПРАВЛЕНИЯ ПАРАМЕТРИЗОВАННОЙ ТРЕХМЕРНОЙ МОДЕЛЬЮ САМОЛЕТА С ПОМОЩЬЮ КЛИЕНТ-СЕРВЕРНОГО ПРИЛОЖЕНИЯ Боргест Н.М., Власов С.А., Коровин М.Д.	417
ПОДХОД К ПАРАМЕТРИЗАЦИИ ТРЕХМЕРНЫХ МОДЕЛЕЙ ДЛЯ РЕШЕНИЯ ЗАДАЧ АВТОМАТИЗАЦИИ ПРЕДВАРИТЕЛЬНОГО ПРОЕКТИРОВАНИЯ САМОЛЕТА Боргест Н.М., Коровин М.Д., Спирина М.О.	421
МЕТОДЫ ПРЕДСТАВЛЕНИЯ ЗНАНИЙ ДЛЯ РЕШЕНИЯ ЗАДАЧ МОДЕЛИРОВАНИЯ Коршиков Д.Н., Лахин О.И., Носкова А.И., Юрыгина Ю.С.	425
ОНТОЛОГИЧЕСКОЕ МОДЕЛИРОВАНИЕ ТЕХНИЧЕСКОГО ОБСЛУЖИВАНИЯ Федотова А.В., Давыденко И.Т.	429
СЕМАНТИЧЕСКИЕ АСПЕКТЫ ПРЕДСТАВЛЕНИЯ И ОБРАБОТКИ ПОИСКОВЫХ ЗАПРОСОВ В СИСТЕМЕ КОРПУС-МЕНЕДЖЕР Невзорова О.А., Мухамедшин Д.Р., Билалов Р.Р.	439
ТЕХНОЛОГИЯ АВТОМАТИЗИРОВАННОГО СИНТЕЗА ИНФОРМАЦИОННЫХ СИСТЕМ С ПОМОЩЬЮ СЕМАНТИЧЕСКИХ МОДЕЛЕЙ ПРЕДМЕТНОЙ ОБЛАСТИ Бикмуллина И.И.	445
ВЕБ-ИНТЕРФЕЙС ДЛЯ СНЯТИЯ МОРФОЛОГИЧЕСКОЙ МНОГОЗНАЧНОСТИ В КОРПУСЕ ТАТАРСКОГО ЯЗЫКА Гильмуллин Р.А. Гатауллин Р.Р.,	451
МНОГОКРИТЕРИАЛЬНЫЙ ПОДХОД К ФОРМИРОВАНИЮ ПОЛИТИКИ НАУЧНОГО ФОНДА И ЭКСПЕРТНОГО СООБЩЕСТВА Бойченко В.С., Заболеева-Зотова А.В., Петровский А.Б.	455
ОНТОЛОГИЧЕСКИЙ ПОДХОД ПРИ ПРОЕКТИРОВАНИИ ТЕХНОЛОГИЧЕСКИХ ПРОЦЕССОВ Бурдо Г.Б., Воробьева Е.В.	461
ИДЕНТИФИКАЦИЯ МОДЕЛИ ПРЕДСТАВЛЕНИЙ АГЕНТА О ПРЕДМЕТНОЙ ОБЛАСТИ В СИТУАЦИИ ВЫБОРА Виноградов Г.П., Борзов Д.А.	465
ИССЛЕДОВАНИЕ ЭФФЕКТИВНОСТИ ГЕНЕТИЧЕСКОГО АЛГОРИТМА ДЛЯ ТЕМАТИЧЕСКОГО ДОКУМЕНТАЛЬНОГО ПОИСКА Иванов В.К., Палюх Б.В.	471
МОДЕЛИРОВАНИЕ РАБОТЫ АВТОМАТИЗИРОВАННОЙ СИСТЕМЫ УПРАВЛЕНИЯ КАЧЕСТВОМ СЕТЯМИ ПЕТРИ Бурдо Г. Б., Виноградов Г. П., Сорокин А. Ю.	477
ПРИМЕНЕНИЕ НЕЙРОННЫХ СЕТЕЙ ГЛУБОКОГО ДОВЕРИЯ ДЛЯ ВЫДЕЛЕНИЯ СЕМАНТИЧЕСКИ ЗНАЧИМЫХ ПРИЗНАКОВ Головки В.А., Крощенко А.А.	481

РАЗРАБОТКА ОТКРЫТЫХ ТЕХНОЛОГИЙ ПРОЕКТИРОВАНИЯ ИНТЕЛЛЕКТУАЛЬНЫХ РОБОТОТЕХНИЧЕСКИХ СИСТЕМ	487
Прокопович Г.А., Сычёв В.А., Герасюто С.Л.	
ПОИСК В СТАТИЧЕСКИХ РОЯХ	489
Воробьев В.В.	
ПОСТРОЕНИЕ ИНТЕЛЛЕКТУАЛЬНЫХ МУЛЬТИАГЕНТНЫХ КОБОРГ-СИСТЕМ	495
Соловьев В.И.	
АДКРЫТЫЯ КАМПАНАНТЫ WWW.CORPUS.BY ДЛЯ НАТУРАЛЬНА- МАЎЛЕНЧАГА ІНТЭРФЕЙСУ	499
Гецэвіч Ю.С., Лабанаў Б.М., Лысы С.І., Гюнтар А.В., Дзенісюк Д.А., Захар’еў В.А.	
РАСПРАЦОЎКА КАМПАНАНТА РАСПАЗНАВАННЯ МАЎЛЕННЯ ДЛЯ НАТУРАЛЬНА МАЎЛЕНЧАГА ІНТЭРФЕЙСУ	507
Нікалаенка К.А., Кайгародава Л.І., Гецэвіч Ю.С.	
РАЗДЕЛЬНОЕ МОДЕЛИРОВАНИЕ РЕЧЕВОГО СООБЩЕНИЯ В ВИДЕ ГОЛОСОВЫХ, ФОНЕТИЧЕСКИХ И ПРОСОДИЧЕСКИХ ПАРАМЕТРОВ	513
Азаров И.С., Петровский А.А.	
ПРИМЕНЕНИЕ МЕТОДОВ ИНТЕЛЛЕКТУАЛЬНОГО АНАЛИЗА ВРЕМЕННЫХ РЯДОВ ДЛЯ ЗАДАЧИ КЛАСТЕРИЗАЦИИ ПОЛЬЗОВАТЕЛЕЙ ПО ГОЛОСУ	519
Вагин В.Н., Ганишев В.А.	
ГОЛОСОВАЯ ИДЕНТИФИКАЦИЯ ПОЛЬЗОВАТЕЛЯ В СИСТЕМАХ КОНТРОЛЯ ДОСТУПА	525
Меньшаков П.А., Мурашко И.А.	
ИНФОРМАЦИОННАЯ МОДЕЛЬ ЛИЧНОСТИ	529
Харламов А.А.	
ОБ ОДНОМ МЕТОДЕ ПРИМЕНЕНИЯ ОБОБЩЕННОЙ ОНТОЛОГИИ ДЛЯ АНАЛИЗА ЕСТЕСТВЕННО-ЯЗЫКОВЫХ СООБЩЕНИЙ ПОЛЬЗОВАТЕЛЕЙ СЕТИ ИНТЕРНЕТ	533
Деева Н.В., Вишневский С.Я.	
ЧАСТОТНЫЙ МЕТОД КЛАССИФИКАЦИИ ТЕКСТОВ С ЛЕКСИЧЕСКИМ РАЗБОРОМ СЛОВА	537
Третьяков Ф.И., Серебряная Л.В.	
ПРОВЕРКА ИНФОРМАТИВНОСТИ КЛАССИФИКАЦИОННЫХ ПРИЗНАКОВ В ЗАДАЧЕ АВТОМАТИЧЕСКОЙ КЛАССИФИКАЦИИ ТЕКСТОВ НА ЕСТЕСТВЕННОМ ЯЗЫКЕ	541
Глазкова А.В.	
ВЕРИФИКАЦИЯ МОДЕЛЕЙ ПРОЦЕССОВ В ДИНАМИЧЕСКИХ СИСТЕМАХ ПО МЕТОДУ MODEL CHECKING	545
Королев Ю.И.	
ИССЛЕДОВАНИЕ И РАЗРАБОТКА СИСТЕМЫ АБСТРАКТНОЙ АРГУМЕНТАЦИИ С ВЕРОЯТНОСТНЫМИ СТЕПЕНЯМИ ОБОСНОВАНИЯ	549
Дервянко А. В., Моросин О. Л.	



МЕТОДЫ ТЕМАТИЧЕСКОЙ КЛАСТЕРИЗАЦИИ ПРИМЕНИТЕЛЬНО К АНАЛИЗУ НОВОСТНЫХ СТАТЕЙ	555
Солошенко А.Н., Орлова Ю.А., Заболевая-Зотова А.В.	
РАЗРАБОТКА ИНТЕЛЛЕКТУАЛЬНОГО ОБУЧАЮЩЕГО ИНТЕРФЕЙСА ДЛЯ ДЕТЕЙ С ИСПОЛЬЗОВАНИЕМ СЕНСОРА LEAP MOTION	561
Розалиев В.Л., Вяхирев А.А., Заболевая-Зотова А.В.	
АВТОМАТИЗИРОВАННОЕ ВЫДЕЛЕНИЕ КИСТЕЙ РУК ЧЕЛОВЕКА ДЛЯ РАСПОЗНАВАНИЯ ЖЕСТОВОЙ РЕЧИ	565
Розалиев В.Л., Агафонов Г.В., Кириченко М.И.	
АВТОМАТИЗАЦИЯ РАСПОЗНАВАНИЯ ВНЕШНЕГО ВИДА ЧЕЛОВЕКА ПО ТЕКСТУ НА ЕСТЕСТВЕННОМ ЯЗЫКЕ И СИНТЕЗ ПОРТРЕТНОГО ИЗОБРАЖЕНИЯ	571
Орлова Ю.А., Долбин А.В., Кипаева Е.В.	
КОНТЕКСТНО-ЗАВИСИМЫЙ АНАЛИЗ ПОРТРЕТНЫХ ФОТОГРАФИЙ	577
Алексеев А.В., Орлова Ю.А.	
МЕТОД ХРАНЕНИЯ ИЕРАРХИЧЕСКИХ СЕМАНТИЧЕСКИХ СЕТЕЙ В РЕЛЯЦИОННОЙ БАЗЕ ДАННЫХ	581
Ванясин Н.В., Сидоркина И.Г.	
НР-ЗАДАЧА БАЛАНСИРОВКИ НАГРУЗОЧНОЙ СПОСОБНОСТИ СЕТИ В САПР	585
Сидоркина И.Г. Сорокин О.Л.	
АППАРАТНЫЙ ИСКУССТВЕННЫЙ ИНТЕЛЛЕКТ ДЛЯ СЕТЕВОЙ ИНФРАСТРУКТУРЫ САПР И СИСТЕМ УПРАВЛЕНИЯ	589
Шелеметьев А.М., Шелеметьева Я.В., Сидоркина И.Г.	
РЕШЕНИЕ ЗАДАЧИ ПРОГНОЗИРОВАНИЯ ВОЗРАСТАНИЯ НАГРУЗКИ В СИСТЕМЕ ДИСТАНЦИОННОГО ОБУЧЕНИЯ «ISPRING LEARN»	593
Сокольников А.М., Сидоркина И.Г.	
ОПЫТ РАЗРАБОТКИ СИСТЕМЫ ИМИТАЦИИ, ОСНОВАННОЙ НА АГЕНТАХ	597
Замятина Е.Б., Каримов Д.Ф. , Митраков А.А.	
РАЗРАБОТКА УЧЕБНОЙ ПРОГРАММНОЙ ПЛАТФОРМЫ ДЛЯ РОБОТА iROBOT CREATE	605
Бармина Е.И., Ланин В.В., Плетнёв А.О.	
<b>АВТОРСКИЙ УКАЗАТЕЛЬ</b>	<b>609</b>

# TABLE OF CONTENTS

<b>FOREWORD</b>	23
PROBLEM OF UNDERSTANDING: PRESENT AND FUTURE OF ARTIFICIAL INTELLIGENCE Tarassov V.B.	25
SCIENTOMETRICS OF THE DOMAIN: OSTIS-CONFERENCES CASE STUDIES Khoroshevsky V.F., Efimenko I.V.	43
EMANTIC TECHNOLOGY OF COMPONENT DESIGN OF SYSTEMS, MANAGED BY KNOWLEDGES Golenkov V.V., Guliakina N.A.	57
SUPPORT TOOLS KNOWLEDGE-BASED SYSTEMS COMPONENT DESIGN Shunkevich D.V., Davydenko I.T., Koronchik D.N., Zukov I.I., Parkalov A.V.	79
IMPLEMENTATION OF WEB-PLATFORM FOR SYSTEMS BASED ON KNOWLEDGES Koronchik D. N	89
THE METHODOLOGY OF KNOWLEDGE BASED SYSTEM COMPONENT DESIGN Shunkevich D.V., Davydenko I.T., Koronchik D.N., Hubarevich N.U., Boriskin A.S.	93
MODELS AND ALGORITHMS OF INTEGRATION OF KNOWLEDGE BASED ON HOMOGENEOUS SEMANTIC NETWORKS Ivashenko V.P.	111
SEMANTIC NETWORKS REPRESENTATION AND ALGORITHMS FOR THEIR ORGANIZATION AND SEMANTIC PROCESSING ON MASSIVELY PARALLEL COMPUTERS Ivashenko V.P., Verenik N.L., Girel A.I., Seitkulov Y.N., Tatur M.M.	133
A COMPONENT-BASED ARCHITECTURE OF INTELLIGENT SYSTEMS OF CONSULTATION PROVIDING AND TRAINING OF SOFTWARE DEVELOPERS Pivovarchyk O.	141
PRINCIPLES OF CONSTRUCTION CLINICAL DECISION SUPPORT SYSTEM BASED ON OSTIS TECHNOLOGY Kayeshko A.I., Marhunou Y.A.	149
KNOWLEDGE BASE DESIGN INTELLECTUAL REFERENCE SYSTEMS IN ALGEBRA Sharipbay A.A., Omarbekova A.S., Nurgazinova G. Sh.	157
THE WAYS OF ARTIFICIAL INTELLIGENCE EVOLUTION IN KAZAKHSTAN Sharipbay A.A.	161
CREATING AN ONTOLOGICAL MODEL FOR THE LOCAL COMPUTER NETWORK Sharipbay A.A., Askarova S.A., Mukanova A.S.	165
ONTOLOGIC MODEL OF PROCESS OF ENSURING INFORMATION SECURITY Niyazova R.S., Budanova N.	169

CONCEPTION INSTRUMENTAL PLATFORM INFORMATION SECURITY IN CLOUD COMPUTING WITH INTELLIGENCE TECHNOLOGIES	173
Vishniakou U.A., Gongas Sas M.M., Mosdurani Shiras M.G.	
ONTOLOGICAL APPROACH TO DEVELOPMENT OF SCIENTIFIC INTERNET RESOURCE	177
Zagorulko Yu.A., Zagorulko G.B.	
DEVELOPMENT AND GENERATION OF VIRTUAL ENVIRONMENTS ON BASE OF SEMANTYC FORMS	183
Gribova V.V., Fedorischev L.A.	
A CLASSIFICATION FOR THE TASKS REVEALED DURING INTELIGENT ACTIVITY SYSTEM ANALYSIS	187
Shalfeeva E.	
A TECHNOLOGY FOR DEVELOPMENT OF PROBLEM SOLVERS OF INTELLIGENT SYSTEMS WITH THE USE OF IACPAAS CLOUD PLATFORM TOOLS	193
Krylov D.A., Moskalenko Ph.M., Timchenko V.A.	
METHODS AND TOOLS OF CONTINGENCY MANAGEMENT IN ENERGY SECTOR BASED ON SEMANTIC MODELING	199
Massel L.V., Massel A.G.	
FROM ENTERPRISE INFORMATION PORTALS TO KNOWLEDGE MANAGEMENT PORTALS: THE ROLE OF INTELLIGENT SERVICES	205
Strakhovich E.V., Vlasov S.A., Gavrilova T.A.	
COLLABORATIVE RECOMMENDER SYSTEMS	209
SHEREMETOVA E.I.	
TOOLS AND METHODS IN BUSINESS RULES APPROACH	215
SAVIN A.M.	
NUCLEAR KNOWLEDGE MANAGEMENT	221
Kostebelova V.K.	
A SCENARIO APPROACH IN THE RESEARCH OF DYNAMICS OF INFORMATION STREAMS IN THE INTERNET	225
Dodonov A.G., Lande D.V., Boychenko A.V.	
APPROACHES OF THE FORMATION OF COMPLEX ENGINEERING CALCULATIONS ON THE PORTAL	231
Novogradskaya R.L., Globa L.S.	
METAGRAPH BASED REPRESENTATION AND PROCESSING OF FUZZY KNOWLEDGBASES	237
Globa L.S., Ternovoy M.Y., Shtogrina O.S.	
CHANGES DETECTION OF XML DOCUMENTS	241
Globa L.S., Molchanov Y.N.	
ENTERPRISE DATA INTEGRATION METHODS UNDER CONDITIONS OF LOW SOURCES RELATEDNESS	249
Galushka I.M, Oksanich I.G., Shcherbak S.S.	

SEMANTIC APPROACH TO THE WEB OF THINGS OBJECTS INTEGRATION Rogushina J., Gladun A.	255
KNOWLEDGE-ORIENTED MEANS OF SEMANTIC SEARCH INTO THE WEB Rogushina J.	265
REVIEW AND ANALYSIS METHODOLOGIES AND TECHNIQUES OF ONTOLOGY BUILDING Khala C.A.	271
SPECIFICATION OF TEMPORAL RELATIONS IN THE SYSTEM "BINARY MODEL OF KNOWLEDGE" Plesniewicz G.S., Nguyen Thi Minh Vu	275
FROM TERMINOLOGICAL NETWORKS TO THE EXPLANATORY DICTIONARIES Malkovsky M.G., Soloviev S.Y.	281
ABOUT AN APPROACH OF TOPIC MODELING FOR TEXT IN NATURAL LANGUAGE Tran Van An	285
MOTIVATION OF MODELING COMMONSENSE REASONING PROCESS Naidenova X. A.	291
HEURISTIC METHOD OF CONSTRAINT SATISFACTION BASED ON MATRIX REPRESENTATION OF CONSTRAINTS Zuenko A.A., Ochinskaya A.A.	297
NETWORK MODEL OF ACTIVITY PROPAGATION AMONG HETEROGENEOUS AUTOMATA Zhilyakova L.Yu.	303
THE SYSTEM PRIORITY TRAFFIC AT INTERSECTIONS "GREEN WAVE" FOR EMERGENCY VEHICLES Sogoyan A.L., Shuts V.N.	309
NATIONAL EDUCATION INFORMATIZATION DEVELOPMENT PROGNOSIS Robert I.V.	315
INTELLIGENT TECHNOLOGIES OF TUTORING IN INFORMATION-EDUCATIONAL SPACE Telnov Yu.F., Trembach V.M.,	325
APPLICATION OF COGNITIVE GRAPHICS TOOLS BASED ON THE 3-SIMPLEX IN INTELLIGENT TRAINING-TESTING SYSTEMS Yankovskaya A.E., Shurygin Y.A., Yamshanov A.V., Krivdyuk N.M.	331
WOLFRAM MATHEMATICA TOOLS AND TECHNOLOGIES OF DEVELOPMENT OF INTELLIGENT TUTORING SYSTEMS Taranchuk VB	339
INTELLECTUAL TECHNOLOGY DEVELOPMENT OF COMPONENT DESIGN OF DESIGN AND INDUSTRIAL LEARNING ENVIRONMENTS Afnas'ev A.N., Voit N.N.	347

ONTOLOGICAL APPROACH TO THE ANALYSIS OF COMPETENCIES OF UNIVERSITY GRADUATES Sherkunov V.V.	351
NEURAL NETWORK MODEL OF THE PROCESS OF PROFESSIONAL EDUCATION OF YOUNG SPECIALISTS Fedyayev O.I.	357
METHODS OF INDISTINCT REGULATION IN MANAGEMENT PROBLEMS EDUCATIONAL PROCESS Nishanov A.H., Doshchanova M. Yu., Mirzaev D.A.	365
INTELLECTUAL MEANS OF AUTOMATION OF MANAGEMENT OF TRAINING Babamukhamedova M.Z., Doshchanova M. Yu., Djangazova ., K.A.	369
CONCEPTUAL EXPERIMENTING IN DESIGNING THE CONFIGURED PATTERNS OF AVIATION PARTS Sosnin P.I., Chorakaev O.E.	373
ONTOLOGY OF DESIGNING FOR TEMPLATES OF AIRCRAFT PARTS Grishin M.V., Larin S.N., Sosnin P.I.	381
SEMANTIC ANALYSIS DIAGRAMMATICAL MODELS IN THE DESIGN OF COMPLEX AUTOMATED SYSTEMS Afanasjev A.N., Gainullin R.F., Afanasjeva T.V.	385
APPLICATION OF NAVIGATION STRUCTURE OF DIGITAL ARCHIVE OF PROJECT ORGANIZATION IN TASKS OF THE INTELLECTUAL ANALYSIS OF CAD DOCUMENTS Subkhangulov R.A., Filippov A.A.	389
ALGORITHMS FOR EVALUATION OF WORD COMBINATIONS OR WORDS MEMBERSHIP DEGREE TO TERM LIST BASED ON SUBJECT AREA ONTOLOGY Andreev I.A., Bashaev V.A., Klein V.V., Moshkin V.S., Yarushkina N.G.	395
METHODS OF CONSTRUCTION OF FUZZY ONTOLOGIES OF COMPLEX SUBJECT AREAS Moshkin V.S., Yarushkina N.G.	401
ONTOLOGICAL APPROACH TO THE FORMATION OF PROJECT REQUESTS OF INTELLIGENT AGENT Namestnikov A.M., Subkhangulov R.A.	407
THE EXPERIENCE OF SEMANTIC MODELING AND DESIGNING TOOLS ON WIDELY USED PLATFORM Smirnov S.V.	413
IMPLEMENTATION OF REMOTE CONTROL OVER A PARAMETERIZED THREE-DIMENSIONAL MODEL OF THE AIRPLANE BY A CLIENT-SERVER APPLICATION Borgest N.M., Vlasov S.A., Korovin M.D.	417
APPROACH TO THE PARAMETRIZATION OF THREE-DIMENTIONAL MODELS FOR SLOVING THE TASK OF THE AIRCRAFT PRELIMINARY DESIGN AUTOMATION Borgest N.M., Korovin M.D., Spirina M.O.	421

METHODS OF KNOWLEDGE REPRESENTATION TECHNIQUES FOR USE IN MODELING Korshikov D, Lakhin O, Noskova A, Yurygina Yu	425
ONTOLOGICAL MODELING OF MAINTENANCE Fedotova A.V., Davydenko I.T.	429
SEMANTIC ASPECTS OF SEARCH REQUEST REPRESENTATION AND PROCESSING IN CORPUS-MANAGER SYSTEM Nevzorova O.A., Mukhamedshin D.R., Bilalov R.R.	439
TECHNOLOGY OF AUTOMATED SYNTHESIS THE INFORMATION SYSTEMS USING SEMANTIC MODELS OF SUBJECT AREA Bikmullina I.I.	445
WEB-SITE FOR HANDY MORPHOLOGICAL DISAMBIGUATION IN TATAR LANGUAGE CORPUS Gilmullin R.R., Gataullin R.R.	451
MULTICRITERIA APPROACH TO FORMING POLICY OF SCIENTIFIC FOUNDATION AND EXPERT COMMUNITY Boychenko V.S., Zaboлева-Zotova A.V., Petrovsky A.B.	455
ONTOLOGICAL APPROACH THE DESIGN PROCESS Burdo G.B., Vorobyeva E.V.	461
IDENTIFICATION MODEL REPRESENTATIONS OF THE AGENT ABOUT THE SUBJECT IN A SITUATION OF CHOICE Vinogradov, P., Borzov D. A.	465
STUDY THE EFFECTIVENESS OF GENETIC ALGORITHM FOR DOCUMENTARY SUBJECT SEARCH Ivanov V.K., Palyukh B.V.	471
MODELING OF AUTOMATED QUALITY CONTROL BY PETRI NETS Burdo G.B., Vinogradov G. P., Sorokin A. Y.	477
APPLYING DEEP BELIEF NEURAL NETWORKS TO EXTRACTION VALUEBLE SEMANTIC FEATURES Golovko V.A., Kroshchanka A.A.	481
DEVELOPMENT OF OPEN TECHNOLOGY DESIGNING OF INTELLIGENT ROBOTIC SYSTEMS Prakapovich R.A., Sychyou U.A., Gerasuto S.L.	487
THE SEARCHING TASK IN THE STATIC SWARM Vorobiev V.V.	489
CREATION OF INTELLEAGENT MULTIAGENT COBOR–SYSTEMS Soloviev V. I.	495
WWW.CORPUS.BY: OPEN-SOURCE COMPONENTS FOR NATURAL LANGUAGE INTERFACES Hetsevich Y.S., Lobanov B.M., Lysy S.I., Hiuntar E.V., Denisjuk D.A., Zakharyeu V.A.	499

COMPONENT DESIGN FOR SPEECH RECOGNITION OF NATURAL LANGUAGE INTERFACE	507
Nikolaenko K.A., Kaigorodova L.I. , Hetsevich Y.S.	
SEPARATE MODELING OF SPEECH USING VOICE, PHONETICAL AND PROSODIC CHARACTERISTICS	513
Azarov E., Petrovsky A.	
APPLICATION OF TIME SERIES ANALYSIS FOR SPEAKER CLUSTERING	519
Vagin V.N., Ganishev V.A.	
VOICE USER IDENTIFICATION IN ACCESS CONTROL SYSTEMS	525
Menshakov P.A., Murashko I.A.	
PERSON PHYSIOGNOMY INFORMATIONAL MODEL	529
Kharlamov A.A.	
ABOUT ONE METHOD OF USE OF GENERALIZED ONTOLOGY FOR THE ANALYSIS OF THE NATURAL LANGUAGE MESSAGES OF INTERNET USERS	533
Deeva N.V. , Vishneuski S.Y.	
TEXT CLASSIFICATION FREQUENCY METHOD WITH WORD LEXICAL ANALYSIS	537
Tretyakov F.I., Serebryanaya L.V.	
CLASSIFICATION FEATURES INFORMATIONAL CONTENT TESTING FOR AUTOMATIC NATURAL TEXTS CLASSIFICATION TASK	541
Glazkova A.V.	
VERIFICATION OF MODELS OF PROCESSES IN DYNAMIC SYSTEMS USING MODEL CHECKING METHOD	545
Eremeev A.P., Korolev Y.I.	
RESEARCH AND DEVELOPMENT OF THE ABSTRACT ARGUMENTATION SYSTEM WITH PROBABILISTIC DEGREE OF JUSTIFICATION	549
Derevyanko A., Morosin O.	
THEMATIC CLUSTERING METHODS APPLIED TO NEWS ARTICLES ANALYSIS	555
Soloshenko A.N., Orlova Yu.A., Zaboлева-Zotova A.V.	
THE DEVELOPMENT OF INTELLIGENT LEARNING INTERFACE FOR CHILDREN USING SENSOR LEAP MOTION	561
Rozaliev V.L., Vyakhirev A.A., Zaboлева-Zotova A.V.	
AUTOMATED ALLOCATION OF THE HANDS OF THE PERSON TO RECOGNIZE SIGN LANGUAGE	565
Rozaliev V.L., Agafonov G.V., Kirichenko M.I.	
AUTOMATED RECOGNITION OF THE APPEARANCE OF MAN ON THE NATURAL LANGUAGE AND SYNTHESIS OF PORTRAIT IMAGES	571
Orlova Yu.A., Doldin A.V., Kipaeva E.V.	
CONTEXT-SENSITIVE ANALYSIS OF PORTRAIT IMAGES	577
Alekseev A.V., Orlova Yu.A.	
HIERARCHICAL SEMANTIC NETWORK STORAGE METHOD FOR RELATIONAL DATABASES	581
Vanyasin N.V., Sidorkina I.G.	

---

NP-BALANCING TASK RATED ABILITY OF NETWORK IN CAD Sidorkina I.G. Sorokin O.L	585
HARDWARE ARTIFICIAL INTELLIGENCE FOR CADS NETWORK INFRASTRUCTURE AND FOR CONTROL SYSTEMS Shelemeteva Y.V., Shelemetev.A.M., Sidororkina I.G.	589
SOLUTION TO THE PROBLEM OF PREDICTING LOAD INCREASES IN LEARNING MANAGEMENT SYSTEM "ISPRING LEARN" Sokolnikov A.M., Sidorkina I.G.	593
THE EXPIERANCE OF AGENT-BASED SIMULATION SYSTEM IMPLEMENTATION Zamyatina E.B., Karimov D.F. , Mitrakov A.A.	597
EDUCATIONAL SOFTWARE PLATFORM DEVELOPMENT FOR iROBOT CREATE Barmina E., Lanin V., Pletnev A.	605
<b>AUTHOR INDEX</b>	609

---





# OSTIS-2015

(Open Semantic Technologies for Intelligent Systems)

УДК 004.934.5

## АДКРЫТЫЯ КАМПАНАНТЫ WWW.CORPUS.BY ДЛЯ НАТУРАЛЬНА-МАЎЛЕНЧАГА ІНТЭРФЕЙСУ

Гецэвіч Ю.С.\*, Лабанаў Б.М.\*, Лысы С.І.\*, Гюнтар А.В.\*, Дзенісюк Д.А.\*, Захар'еў В.А.\*\*

*\* Аб'яднаны інстытут праблем інфарматыкі Нацыянальнай акадэміі навук Беларусі,  
г. Мінск, Беларусь*

**yury.hetsevich@gmail.com**

**lobanov@newman.bas-net.by**

**stanislau.lysy@gmail.com**

**lena205593@gmail.com**

**d.denissyuk@gmail.com**

*\*\* Беларускі дзяржаўны ўніверсітэт інфарматыкі і радыёэлектронікі, г. Мінск, Беларусь  
**delfvad@gmail.com***

У дадзеным артыкуле апісаны агульныя падыходы да праектавання, рэалізацыі і выкарыстання адкрытых кампанентаў інтэрнэт-рэсурсу [www.corpus.by](http://www.corpus.by) для натуральна-маўленчага інтэрфейсу з мэтай паляпшэння і ўдасканалення працы сістэмы сінтэзу маўлення па тэксце з адначасовым прадстаўленнем магчымасці інтэграцыі з праектам OSTIS. Прыводзяцца апісанні мэтаў і задачай, спосабаў працы і даступнай функцыянальнасці некаторых рэалізаваных сэрвісаў.

**Ключавыя словы:** сінтэз маўлення па тэксце, натуральна-маўленчы інтэрфейс, кампанент, інтэрнэт-сэрвіс.

### Уводзіны

Дзеля вырашэння задачы ўсталявання ўзаемадзеяння паміж чалавекам і машынай пры дапамозе маўлення навукоўцы ва ўсім свеце працуюць над стварэннем сістэм сінтэзу і распазнавання маўлення. У апошні час сістэмы сінтэзу маўлення па тэксце (ССМТ) ужо дасягнулі пэўнай дасканаласці і шырока выкарыстоўваюцца. Аднак нельга сказаць, што задача сінтэзу маўлення з'яўляецца вырашанай, асабліва ў дачыненні да беларускай мовы. У сувязі з гэтым аўтарамі было прынята рашэнне стварыць інтэрнэт-рэсурс, які б аб'ядноўваў шэраг інтэрнэт-сэрвісаў, кожны з якіх у сваю чаргу вырашаў бы пэўныя свае ўласныя задачы (напрыклад, прыкладныя лінгвістычныя задачы), але ў той жа час дапамагаў распрацоўшчыкам ва ўдасканаленні працы сінтэзатара маўлення па тэксце. Такі падыход дае магчымасць праз ужыванне сэрвісаў вялікай колькасцю карыстальнікаў правесці тэставанне і збор неабходнай інфармацыі, пры дапамозе якіх можа ўдасканальвацца якасць працы сінтэзатара маўлення па тэксце [Лобанов, 2008], [Гецэвіч, 2012].

У дакладзе прадстаўлены агульныя падыходы да праектавання, рэалізацыі і выкарыстання інтэрнэт-сэрвісаў у мэтах паляпшэння працы ССМТ з прадстаўленнем магчымасці інтэграцыі з іншымі праектамі (у прыватнасці з праектам OSTIS) праз інтэрнэт-сайт [www.corpus.by](http://www.corpus.by) [[www.corpus.by](http://www.corpus.by), 2012]. Прыводзяцца апісанні мэтаў і задачай, спосабаў працы і даступнай функцыянальнасці некаторых рэалізаваных сэрвісаў.

### 1. Агульныя падыходы да праектавання інтэрнэт-сэрвісаў

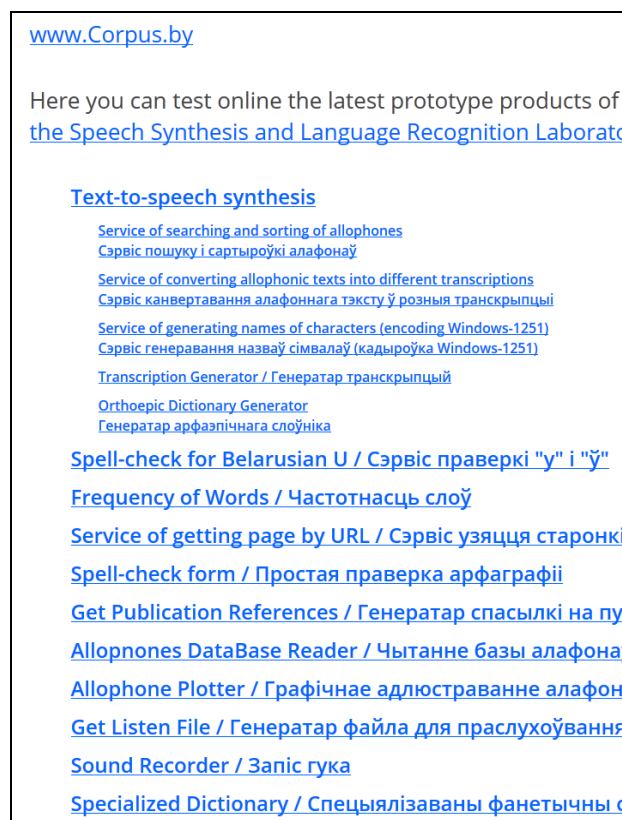
Распрацоўка ССМТ з'яўляецца складанай задачай, якая ахоплівае вялікую колькасць падзадач, а таксама сутыкаецца з шэрагам пабочных прыкладных лінгвістычных задач, якія патрабуюць вырашэння. Задачы, якія паўстаюць пры стварэнні ССМТ, можна ўмоўна падзяліць на тры групы:

- перадапрацоўку ўваходных тэкстаў;
- непасрэдна сінтэз маўлення;
- апрацоўку і выкарыстанне выніковых даных.

Інтэрнэт-рэсурс [www.corpus.by](http://www.corpus.by) мае сваёй мэтай рашэнне як непасрэднай задачы сінтэзавання

маўлення, так і шэрагу дадатковых задач перадапрацоўкі тэксту і апрацоўкі вынікаў працы сінтэзатара маўлення. Дзеля дасягнення дадзенай мэты было прынята рашэнне распрацаваць для гэтага рэсурсу сістэму інтэрнэт-сэрвісаў, кожны з якіх вырашаў бы пэўную ўласную задачу, а ў спалучэнні з іншымі сэрвісамі – неабходныя падзадачы ССМТ. На сённяшні дзень распрацаваны альбо знаходзяцца ў праежжавай стадыі распрацоўкі і даступны бясплатна ў Інтэрнэце наступныя сэрвісы:

- сэрвіс пошуку і сартыроўкі алафонаў;
- сэрвіс канвертавання алафоннага тэксту ў розныя транскрыпцыі;
- сэрвіс генерацыі назваў сімвалаў;
- сэрвіс генерацыі транскрыпцый;
- сэрвіс генерацыі арфаэпічнага слоўніка;
- сэрвіс праверкі “y” і “ў”;
- сэрвіс падліку частотнасці слоў;
- сэрвіс генерацыі спасылкі на публікацыю;
- сэрвіс графічнага адлюстравання алафонаў і алафонных фраз;
- сэрвіс генерацыі файла для праслухоўвання;
- сэрвіс запісу гуку;
- сэрвіс спецыялізаванага фанетычнага слоўніка; і іншыя.



Малюнак 1 – Знешні інтэрфейс інтэрнэт-рэсурсу [www.Corpus.by](http://www.Corpus.by)

Знешні інтэрфейс пачатковай старонкі інтэрнэт-рэсурсу [www.Corpus.by](http://www.Corpus.by), на якой знаходзяцца

спасылкі на ўсе пералічаныя вышэй сэрвісы, прадстаўлены на малюнку 1.

## 2. Фарміраванне патрабаванняў да праектуемых інтэрнэт-сэрвісаў

Мэтай распрацоўкі інтэрнэт-сэрвісаў з’яўляецца павышэнне якасці працы сінтэзатара маўлення па тэксце для беларускай мовы. Аўтарамі артыкула было заўважана, што гэтага можна дасягнуць некалькімі спосабамі: па-першае, праз вынясенне пэўных элементаў існуючай ССМТ длявольнага карыстання ў Інтэрнэт у выглядзе інтэрнэт-сэрвісаў, якія б мелі пэўную мэтавую аўдыторыю; па-другое, праз распрацоўку пэўных інтэрнэт-сэрвісаў, якія б маглі быць у далейшым інтэграваны ў ССМТ, а пакуль удасканалення і адтэставання ў Інтэрнэце. Такім чынам можна атрымаць наступнае:

- непасрэдня вынікі працы кожнага сэрвісу на адвольна зададзеных карыстальнікам уваходных даных. Мноства адпаведнасцяў «уваходныя даныя – выніковыя даныя» можа быць выкарыстана як для аналізу правільнасці працы таго ці іншага элемента ССМТ, так і для стварэння аўтаматызаванага тэставання працы сэрвісу і ССМТ;
- зваротную сувязь ад мэтавай аўдыторыі і экспертаў, у супрацоўніцтве з якімі вядзецца распрацоўка;
- зваротную сувязь ад звычайных зацікаўленых у тэматыцы карыстальнікаў.

Прымаючы да ўвагі пералічаныя вышэй мэты, можна сфармуляваць наступныя асноўныя патрабаванні да праектуемых інтэрнэт-сэрвісаў:

- простасць, зручнасць і інтуітыўная зразумеласць карыстальніцкага інтэрфейсу;
- захаванне ўсіх даных, пададзеных карыстальнікам на ўваход;
- захаванне ўсіх выніковых даных;
- аператыўнае аўтаматызаванае апаўшчэнне распрацоўшчыкаў аб памылках у працы сэрвісаў і сістэмы сінтэзу маўлення.

Праектаванне інтэрнэт-сэрвісаў уключае ў сябе як вызначэнне патрабаванняў, так і апісанне статычнай структуры і ўзаемадзеяння элементаў сістэмы. У якасці графічнай мовы дакументавання выкарыстоўваецца натацыя UML 2.0 (Unified Modeling Language).

На малюнку 2 прадстаўлена схема кампанентаў узаемадзеяння карыстальніка з шэрагам службовых сэрвісаў сістэмы адкрытых кампанентаў для пабудавання маўленчага інтэрфейсу – [www.Corpus.by](http://www.Corpus.by). Пяройдзем да разгляду функцыянальных магчымасцей праектуемых інтэрнэт-сэрвісаў.

## 3. Вызначэнне варыянтаў выкарыстання інтэрнэт-сэрвісаў

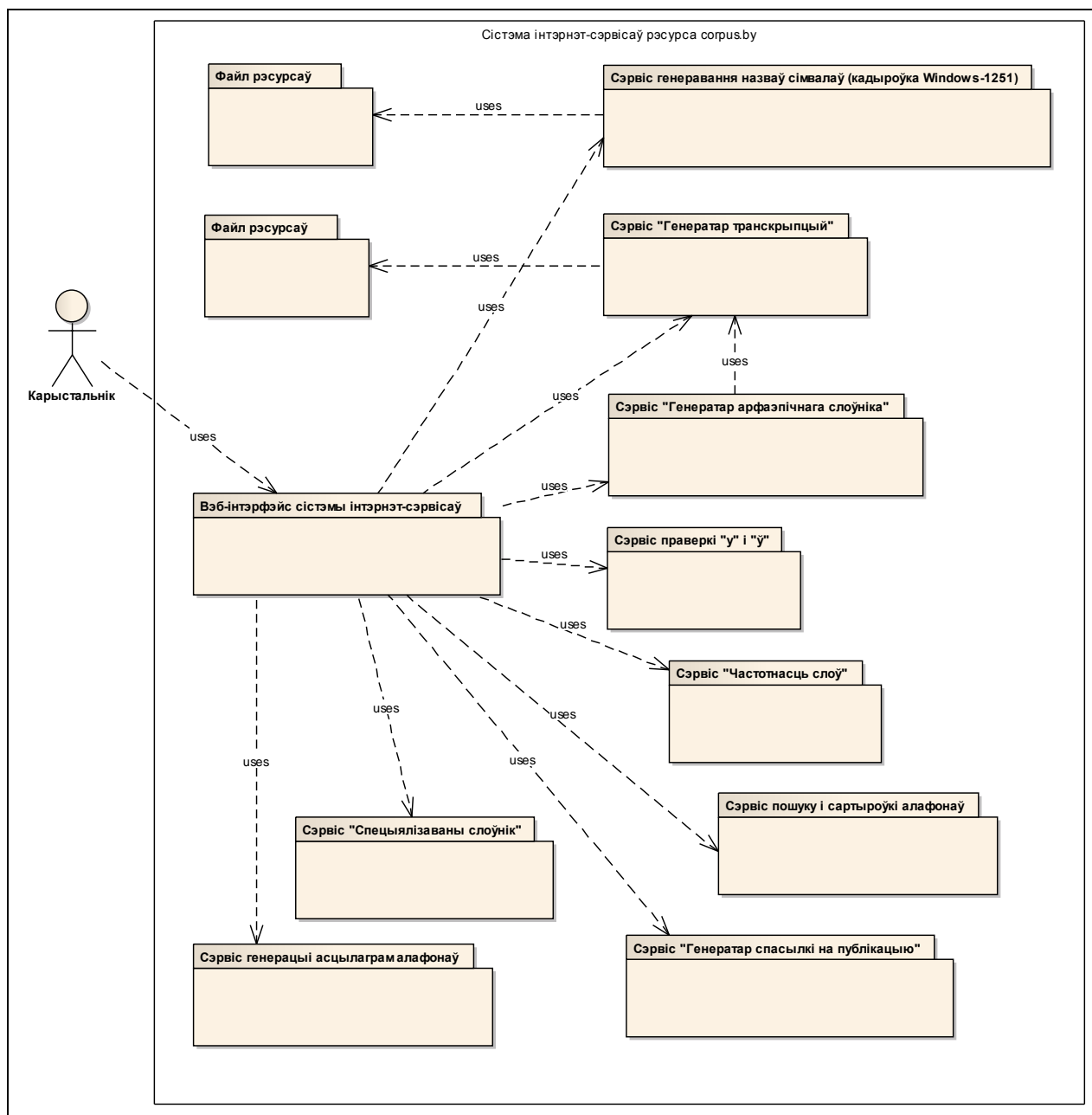
Праектуемая сістэма інтэрнэт-сэрвісаў прадугледжвае выкарыстанне яе як знешнімі

карыстальнікамі, так і распрацоўшчыкамі. Выкарыстанне інтэрнэт-сэрвісаў аднымі і другімі можа як супадаць, так і адрознівацца, бо шэраг распрацоўшчыкаў з’яўляюцца адначасова і карыстальнікамі ў сваіх працоўных мэтах.

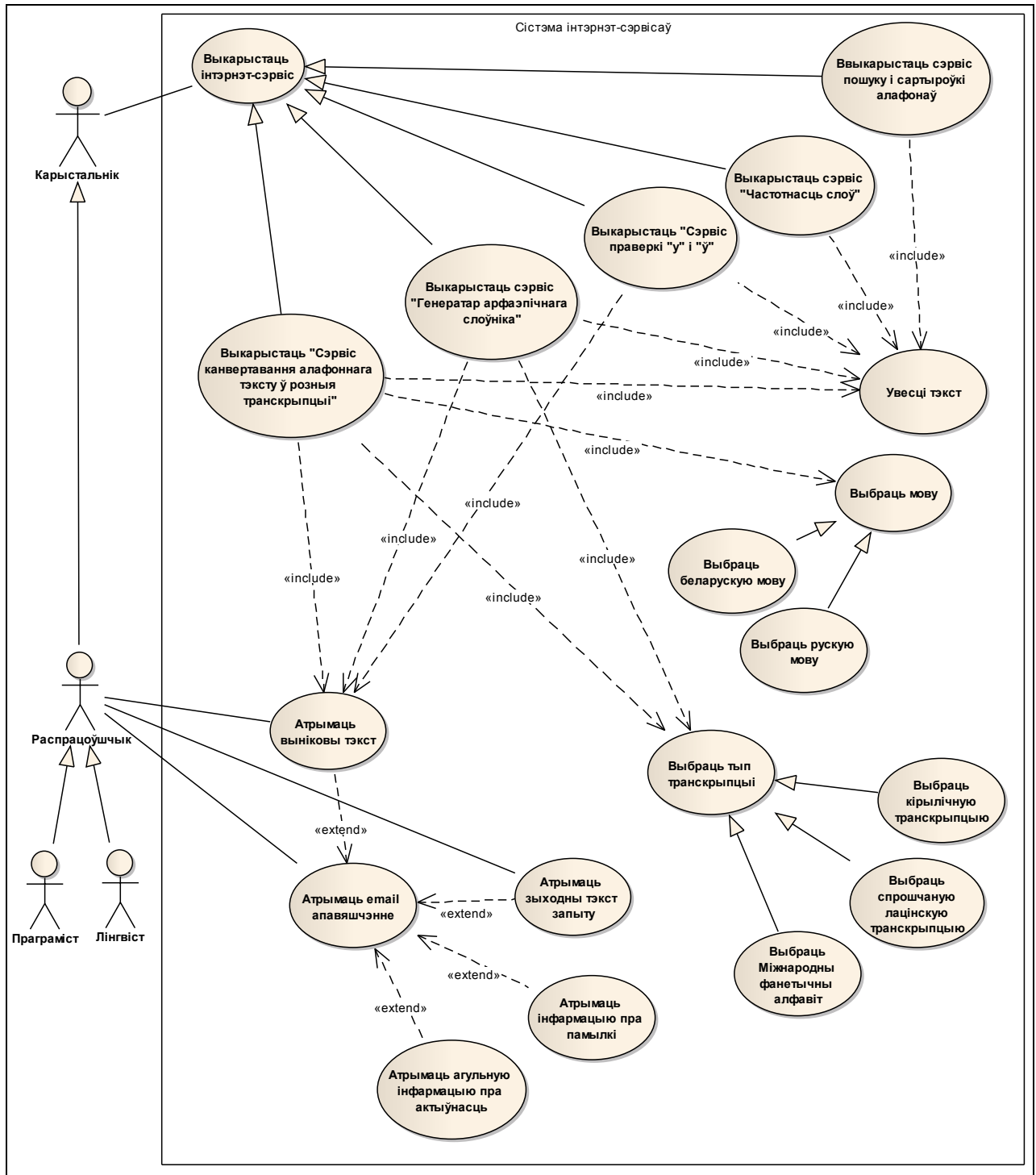
На малюнку 3 прадстаўлена схема варыянтаў выкарыстання сістэмы інтэрнэт-сэрвісаў карыстальнікамі і распрацоўшчыкамі.

На схеме варыянтаў выкарыстання прыведзены наступныя сэрвісы: “Сэрвіс канвертавання алафоннага тэксту ў розныя транскрыпцыі”, “Генератар арфаэпічнага слоўніка”, “Сэрвіс праверкі “у” і “ў”, сэрвіс “Частотнасць слоў”, “Сэрвіс пошуку і сартыроўкі алафонаў”. У кожным

з іх карыстальнік можа падаць на ўваход пэўны тэкст альбо паслядоўнасць сімвалаў і атрымаць на выхадзе неабходны выніковы тэкст у адпаведнасці з задачай, якую вырашае дадзены сэрвіс. У «Сэрвісе канвертавання алафоннага тэксту ў розныя транскрыпцыі» і сэрвісе «Генератар арфаэпічнага слоўніка» карыстальнік можа абраць адзін з трох тыпаў транскрыпцыі: кірылічную, спрошчаную лацінскую ці транскрыпцыю ў фармаце Міжнароднага фанетычнага алфавіту. У «Сэрвісе канвертавання алафоннага тэксту ў розныя транскрыпцыі», як і ў шэрагу іншых, не пазначаных на схеме сэрвісаў, магчымы выбар мовы ўваходнага тэксту: беларускай альбо рускай.



Малюнак 2 – Схема кампанентаў узаемадзеяння карыстальніка з сістэмай інтэрнэт-сэрвісаў www.Corpus.by



Малюнак 3 – Схэма варыянтаў выкарыстання сістэмы інтэрнэт-сэрвісаў www.Corpus.by

Распрацоўшчыкі могуць выкарыстоўваць інтэрнэт-сэрвісы для аналізу працы сэрвісаў і прыняцця рашэнняў па дапрацоўцы і выпраўленні памылак іх працы, а таксама для вырашэння шляхоў развіцця, пошуку сродкаў і спосабаў павышэння эфектыўнасці працы праз назіранне за пэўнымі тэндэнцыямі і заканамернасцямі ў карыстальніцкіх запытах да сэрвісаў.

Дадзеныя сэрвісы праектуюцца як для звычайных карыстальнікаў, так і для адмысловай мэтавай аўдыторыі: лінгвістаў-экспертаў і ўласна

распрацоўшчыкаў праграмных сродкаў: дадзеных сэрвісаў і ССМТ. Такім чынам, распрацоўшчыкі (лінгвісты і праграмісты) з'яўляюцца і прыватным выпадкам карыстальніка, але між іншага валодаюць і сваімі ўласнымі спецыфічнымі магчымасцямі выкарыстання.

Звычайныя карыстальнікі маюць магчымасць бясплатна выкарыстоўваць інтэрнэт-сэрвісы рэсурсу www.Corpus.by у адвольных мэтах. Прывядзём апісанне некалькіх распрацаваных і даступных у інтэрнэце сэрвісаў.

## 4. Апісанне сэрвісаў: функцыянальныя магчымасці, спосабы выкарыстання

### 4.1. Генератар файлаў для праслухоўвання

Сэрвіс “Generator of Listen File” ставіць сваёй мэтай прывесці хаця б па адным прыкладзе для адлюстравання кожнага канкрэтнага алафона ці дыфона (для дыфонаў пакуль рэалізаваны толькі выпадак “санорны зычны – галосны”) [Listen file, 2014]. У аснове дадзенага сэрвісу ляжаць спісы слоў, якія былі спецыяльна складзены экспертам. Для зручнасці карыстання і ўспрымання, падабраныя словы згрупіраваны па тэматыках, паколькі адвольны набор слоў, зусім не звязаных між сабой, часта выклікае пэўныя складанасці – карыстальнік падсвядома пачынае суадносіць слова з яго прадметнай вобласцю, што адцягвае ўвагу ад пастаўленай перад ім задачы і запавольвае працу.

Для таго, каб атрымаць спіс слоў для алафонаў, размеркаваных па тэматыках, неабходна націснуць на кнопку “Read Data from Xslx File (Allophones)” (малюнак 4а), адпаведна для дыфонаў – “Read Data from Xslx File (Diphones)” (малюнак 4б). Акрамя прыкладаў карыстальнік таксама атрымае інфармацыю пра колькасць тэматык і слоў, а таксама сярэдняю колькасць слоў у адной тэматыцы.

Функцыянальнасць сэрвісу “Generator of Listen File” дае магчымасць паляпшаць працу сінтэзатара маўлення шляхам начытвання адабраных і разбітых па тэматыках слоў і выразання алафонаў, якіх на дадзены момант няма ў базе сінтэзатара і якія маюць ня вельмі добрую якасць. Акрамя таго дадзены сэрвіс можа выкарыстоўвацца для больш хуткага і зручнага стварэння новых “галасоў” для сінтэзатара.

**Generator of Listen File**

**Generator of Listen File**

Read Data from Xslx File (Allophones)

Read Data from Xslx File (Diphones)

Read Import Txt File

Read Import Xslx File

**Read Data from Xslx File (Allophones)**

Topics = 42, Words = 719, Words per Topic = 17.1

**Тэматычны дамен: будоўля..##**

аббудава+ць..#  
аббі+ць..#  
аббіва+нне..#  
абпіло+ўванне..#  
адту+ліна..#  
бу=йнамашта+бных..#  
бэ=лечна-кансо+льны..#  
гу=каізаля+цыя..#  
да+х хі+сткі..#  
жыллѐ+..#  
не+калькі а+рак..#  
пабудава+ць э+лінг..#

**Тэматычны дамен: ваенная тэматыка..##**

правядзі+ а=нтывае+нную а+кцыю..#  
радыёгра+ма..#  
раззбрае+нне..#  
ружжо+..#

а)

**Generator of Listen File**

**Generator of Listen File**

Read Data from Xslx File (Allophones)

Read Data from Xslx File (Diphones)

Read Import Txt File

Read Import Xslx File

**Read Data from Xslx File (Diphones)**

Topics = 37, Words = 805, Words per Topic = 21.8

**Тэматычны дамен: абстрактыныя назо+ўнікі..##**

абплята+нне..#  
адда+насць..#  
аддале+нне..#  
адту+ліна..#  
аслупяне+нне..#  
бясці+лле..#  
выбе+льванне..#  
ка+шлянне..#  
любо+ў..#  
рва+нь..#  
імкне+нне..#

**Тэматычны дамен: будоўля..##**

аб'е+ктавы..#  
аб'е+мны..#  
аб'е+місты..#  
аббудава+ць..#  
абпіло+ўванне..#

б)

Малюнак 4 – Прыклад адлюстравання слоў для рэпрэзентацыі а) алафонаў; б) дыфонаў

## 4.2. Генератар спасылкі на публікацыю

Стварэнне бібліяграфічных спасылак – неад’емная частка навуковай працы. Каб спрасціць гэтую задачу для карыстальнікаў, быў створаны сэрвіс “Get a publication reference!” [Publication references, 2014]. Дадзены сэрвіс распрацаваны для аўтаматычнай генерацыі спасылак.

Зайшоўшы на старонку сэрвісу, карыстальніку неабходна пазначыць фармат спасылкі, тып крыніцы (кніга, артыкул часопісу, канферэнцыя, вэб-сайт) і мову афармлення публікацыі (беларуская, руская ці англійская) (малюнак 6).

Наступным крокам з’яўляецца ўвод даных карыстальнікам. Абавязковыя для запаўнення палі пазначаны чырвонай зорчак (\*). Таксама прадугледжана магчымасць прагляду прыкладаў запаўнення, для чаго неабходна націснуць на кнопку “Ачысціць усё і паказаць прыклады”. Прыклад запаўнення формы данымі паказаны на малюнку 7.

Малюнак 6 – Пазначэнне фармату, тыпу і мовы публікацыі

Малюнак 7 – Запаўненне формы данымі

Па заканчэнні запаўнення формы неабходна націснуць на кнопку “Атрымаць спасылку на публікацыю”. Вынік з’явіцца ў асобным полі ніжэй (малюнак 8).

На дадзены момант для карыстальнікаў даступны толькі ВАК-фармат публікацый. Плануецца пашырыць магчымасці сэрвісу праз павелічэнне колькасці тыпаў спасылак і далучэнне наступных фарматаў: Chicago, MLA, APA.

### Publication references

Гецэвіч, Ю.С. Стварэнне сэрвіса арфаэпічнага генератара слоўнікаў / Ю.С. Гецэвіч, А.В. Гюнтар, С.І. Лысы [і інш.] // Тезі доповідей міжнародной конференції «Діалекты в синхроніі та дяхроніі : загальнаслов’янський контекст» (Київ, 2–4 квітня 2014 року) / За ред. П.Ю. Гриценка. Ін-т укр. мови НАН України. Київ : КММ, 2014. — С. 101-106.

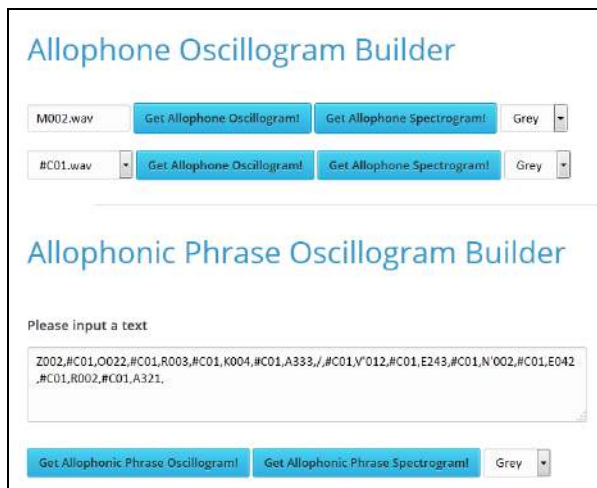
Малюнак 8 – Згенераваная спасылка на публікацыю

Праца з генератарам бібліяграфічных спасылак дазволіла даведацца больш дакладна пра іх змест і структуру, што з’яўляецца добрай асновай для распрацоўкі сістэмы распазнавання спасылак у тэкстах беларускай і рускай моў. Інфармацыя якую можна атрымаць з бібліяграфічных спасылак пры распазнаванні, можа быць карыснай для бібліятэк, выдавецтваў ці навуковых устаноў, бо яна змяшчае звесткі пра аўтараў і рэдактараў, назвы публікацый, назву і месца выдавецтва, год выдання і інш.

## 4.3. Сэрвіс графічнага адлюстравання алафонаў і алафонных фраз

Сэрвіс графічнага адлюстравання алафонаў і алафонных фраз “Allophone Plotter” – гэта яшчэ адзін са службовых сэрвісаў сістэмы адкрытых кампанентаў для пабудавання маўленчага інтэрфейсу – www.Corpus.by [Allophone Plotter, 2014]. Яго галоўнай задачай з’яўляецца графічнае адлюстраванне фізічнага сігналу ў часавым альбо частотным выглядзе адпаведнай фанетычнай адзінкі – алафона, абранага карыстальнікам сэрвісу, альбо платаванне цэлай фразы па фанетычным тэксце, які мае форму радка з паслядоўна запісаных алафонаў. Для адлюстравання сігналу ў часовай вобласці выкарыстоўваецца графік асцылаграмы – залежнасці амплітуды сігналу ад часу. Для частотнага прадстаўлення выкарыстоўваецца графік спектраграмы – залежнасці амплітуды ці энергіі сігналу ад часу і ад частаты адначасова. У гэтым выпадку графік выглядае як двухмерны каляровы малюнак сігналу з часам, які адкладаецца па восі абсцыс, частатой – па восі ардынат, а ўзровень энергіі сігналу характарызуецца інтэнсіўнасцю колеру на малюнку.

Сэрвіс можа выкарыстоўвацца экспертамі-фанетыстамі, лінгвістамі, студэнтамі філалагічных і педагогічных вузаў ці проста зацікаўленымі асобамі для знаёмства з “выглядам” і фізічнымі характарыстыкамі алафонаў (рэалізацыямі фанем), у працэсе вывучэння беларускай мовы. Напрыклад, спектраграма можа вельмі моцна дапамагчы ў вывучэнні як фанетыкі мовы ў агульным, так і асобных гукаў мовы ў прыватнасці. Кожны алафон мае свае асаблівасці з пункту гледжання яго фізічных параметраў, якія з’яўляюцца вынікамі дзеяння інтра- і экстралінгвістычных фактараў мовы і якія складаюць “жывую непаўторную карціну” гэтага гука, добра бачную на частотна-часавым плане, які сабой уяўляе спектраграма.

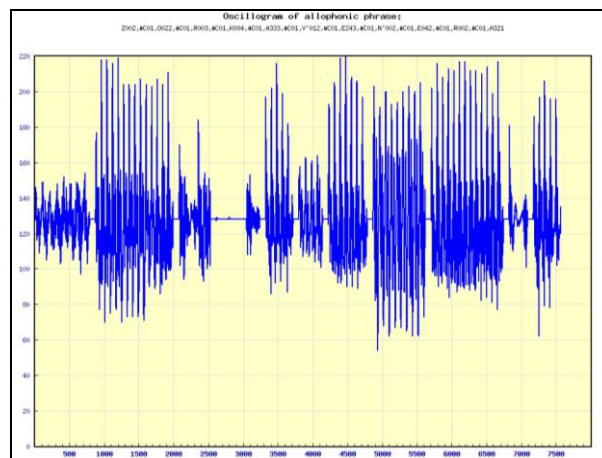


Малюнак 9 – Знешні інтэрфейс “Сэрвісу графічнага адлюстравання алафонаў і алафонных фраз”

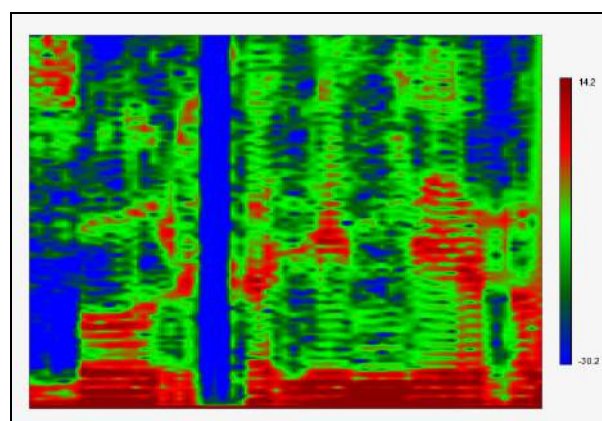
Асцылаграма фанетычнай фразы можа эфектыўна выкарыстоўвацца для вывучэння прасодыкі (інтанацыйнага склада) мовы, ад якой вельмі моцна залежыць успрыманне маўлення чалавекам. Як вядома, прасодыка складаецца з трох кампанентаў: энергетыкі – цяжучай змены сілы гуку, рытмікі – змены працягласці гукаў і паўз і мелодыкі – руху частаты асноўнага тону. Усе гэтыя моўныя з’явы яркава назіраюцца на асцылаграме. Напрыклад, праз адлюстраванне рытмікі асцылаграма можа падказаць карыстальніку, дзе адносна агульнай працягласці фразы, знаходзіцца слоўны ці фразавы націск, таму што ў гэтым месцы на графіку павінна назірацца большая амплітуда сігналу.

На сённяшні момант агульны сцэнар працы з сэрвісам выглядае наступным чынам. Карыстальнік заходзіць на старонку “Сэрвісу графічнага адлюстравання алафонаў” (знешні інтэрфейс старонкі прадстаўлены на малюнку 9). На гэтай старонцы ён можа наплатаваць асобны алафон ці алафонную фразу. Асобны алафон можна ўвесці ўручную ў тэкставым полі ці выбраць з фанетычнай базы праз выпадаючае меню. Алафонную фразу карыстальнік павінен уводзіць у спецыяльна адведзенае тэкставае поле, у якім дазволены шматрадковы ўвод. Пасля ўвода асобнага алафона ці алафоннага тэксту карыстальнік павінен націснуць адпаведную кнопку для пабудавання асцылаграмы ці спектраграмы. Пры выбары платавання спектраграмы, у карыстальніка ёсць магчымасць выбару карты колеру: чорна-белай ці каляровай. Далей трэба пачакаць некаторы час пакуль сэрвіс апрацуе запыт і згенеруе адказ у выглядзе html-старонкі, у якую ўбудаваны малюнак сігналу ў фармаце png, і якую браўзер адлюструе карыстальніку. Вынікі працы сэрвісу прадстаўлены на малюнку 10.

Асаблівацю сэрвісу з’яўляецца магчымасць апрацоўкі сігналу і будавання графічных “партрэтаў гукаў жывой мовы on-line”, а таксама яго арыентацыя на беларускую мову.



а)



б)

Малюнак 10 – Вынікі працы “Сэрвісу графічнага адлюстравання алафонаў і алафонных фраз”: а) асцылаграма алафоннай фразы; б) спектраграма алафоннай фразы

На дадзены момант аўтары не маюць інфармацыі аб аналагічных бясплатных on-line сэрвісах для лінгвістаў ці людзей, зацікаўленых у вывучэнні беларускай мовы, якія валодалі б такой жа самай функцыянальнасцю. У далейшым плануецца выкарыстанне гэтага сэрвісу ў якасці дадатковай функцыянальнасці для больш высокаўзроўневых сэрвісаў рэсурсу [www.Corpus.by](http://www.Corpus.by).

## Заклучэнне

У выніку праведзенай працы былі распрацаваны агульныя падыходы да стварэння сістэмы інтэрнэт-сэрвісаў [www.Corpus.by](http://www.Corpus.by) у мэтах паляпшэння і ўдасканалення працы сістэмы сінтэзу маўлення па тэксце з адначасовай магчымасцю як вырашэння розных прыкладных задач, так і інтэграцыі з іншымі праектамі, у прыватнасці з праектам OSTIS. У дадзеным артыкуле апісаны асаблівасці праектавання, рэалізацыі і выкарыстання сістэмы, прыведзены апісанні некаторых распрацаваных сэрвісаў, а менавіта: сэрвісу “Генератар файлаў для праслухоўвання”, сэрвісу “Генератар спасылкі на публікацыю”, “Сэрвісу графічнага адлюстравання алафонаў і алафонных фраз”.

## Бібліяграфічны спіс

[Allophone Plotter, 2014] Allophone Plotter [Electronic resource]. – 2014. – Mode of access : <http://corpus.by/readVoiceDBAllophones/>. – Date of access : 12.12.2014.

[Listen file, 2014] Generator of Listen File [Electronic resource]. – 2014. – Mode of access : <http://corpus.by/genListenFile/>. – Date of access : 12.12.2014.

[Publication references, 2014] Get a publication reference! [Electronic resource]. – 2014. – Mode of access : <http://corpus.by/publicationReference/>. – Date of access : 12.12.2014.

[www.Corpus.by, 2012] Text-to-Speech PHP-Based Synthesizer [Electronic resource]. – 2012. – Mode of access : <http://corpus.by/>. – Date of access : 12.12.2014.

[Гецэвіч, 2012] Гецэвіч, Ю.С. Алгарытмы лінгвістычнай апрацоўкі тэкстаў для сінтэзу маўлен-ня на беларускай і рускай мовах: дыс. ... канд. тэхн. навук / Ю.С. Гецэвіч. – Мінск, 2012. – 193 с.

[Лобанов, 2008] Лобанов, Б.М. Компьютерный синтез и клонирование речи / Б.М. Лобанов, Л.И. Цирульник. – Минск : Белорусская наука, 2008. – 344 с.

### WWW.CORPUS.BY: OPEN-SOURCE COMPONENTS FOR NATURAL LANGUAGE INTERFACES

Hetsevich Y.S.\*, Lobanov B.M.\*, Lysy S.I.\*,  
Hiuntar E.V.\*, Denisyuk D.A.\*,  
Zakharyeu V.A.\*\*

\* *United Institute of Informatics Problems,  
National Academy of Sciences, Minsk, Belarus*

**yury.hetsevich@gmail.com,  
lobanov@newman.bas-net.by,  
stanislau.lysy@gmail.com, lena205593@gmail.com,  
d.denisyuk@gmail.com**

\*\* *Belarusian State University of Informatics and  
Radioelectronics, Minsk, Belarus*

**delfvad@gmail.com**

This paper describes general approaches to designing, implementation and usage of open-source components of [www.Corpus.by](http://www.Corpus.by) internet-resource for Natural Language Interfaces. These components serve to improve the performance of a text-to-speech system (TTS), while at the same time providing the opportunity of integration with the OSTIS project. The paper also define goals and objectives, ways of operation and available functionality of some services that have been implemented.

## INTRODUCTION

In order to solve the problem of speech interaction between people and machines, scientists all over the world are working on creating speech synthesis and recognition systems. Recently, TTS systems have reached certain degree of accuracy and become widely used, but it is impossible to say that the task of text-to-speech synthesis is completely resolved, especially it concerns the Belarusian language. In this regard, the authors decided to create an Internet resource that would bring together a number of Internet-services,

each aimed at solving certain tasks and at the same time helping developers to make the TTS system performance more optimal and precise.

## MAIN PART

The tasks facing the developers of a text-to-speech system may be conventionally divided into three groups: preprocessing of input texts, speech synthesis itself and processing of resulting data.

The internet-resource [www.Corpus.by](http://www.Corpus.by) is aimed at addressing the challenge of text-to-speech synthesis itself, as well as at solving a range of additional tasks of text-preprocessing and processing of the results of the text-to-speech synthesizer's operation. In order to achieve this aim, the decision was made to develop a system of Internet-services, all of each would solve its own certain task and, when combined with all the rest services, would also solve the essential subtasks of the text-to-speech system. The following services are currently developed or are on the intermediate stage of development and are easily available on the Internet:

- “Service of searching and sorting of allophones”;
- “Service of converting allophonic texts into different transcriptions”;
- “Service of generating names of characters (encoding Windows-1251)”;
- “Transcription Generator”;
- “Orthoepic Dictionary Generator”;
- “Spell-check for Belarusian U”;
- “Frequency of Words”;
- “Get Publication References”;
- “Allophone Oscillogram Builder”;
- “Allophone Plotter”;
- “Get Listen File”;
- “Sound Recorder”;
- “Specialized Dictionary”;
- etc.

## CONCLUSION

As a result of the work carried out by the authors, in order to improve and make more precise the performance of the text-to-speech system, and at the same time to provide free electronic services to the population in order to resolve a wide range of applied tasks, general approaches to creation of the system of internet-services [www.Corpus.by](http://www.Corpus.by) were developed.