

Оглавление

Е.П. Чирко, В.И. Зуев

**ОТ СОСТАВИТЕЛЕЙ**

Н. И. Баяндин

**ИНФОРМАЦИОННО-АНАЛИТИЧЕСКИЕ СИСТЕМЫ В АНТИКОРРУПЦИОННОЙ  
БОРЬБЕ**

Д. А. Богданова

**ИНТЕРНЕТ ВЕЩЕЙ – «ЦИФРОВЫМ АБОРИГЕНАМ» И ИХ РОДИТЕЛЯМ**

Ю. Н. Дрешер, Т. И. Ключенко, Е. А. Косолапова

**ИНСТРУМЕНТЫ УПРАВЛЕНИЯ ИНФОРМАЦИОННО-БИБЛИОТЕЧНОЙ  
ДЕЯТЕЛЬНОСТЬЮ**

Р. А. Румянцев, О. А. Невзорова

**НЕКОТОРЫЕ ПРОГРАММНЫЕ ИНСТРУМЕНТЫ  
ДЛЯ АВТОМАТИЗИРОВАННОГО ПОПОЛНЕНИЯ  
ТЕРМИНОЛОГИЧЕСКОГО СЛОВАРЯ ПРЕДМЕТНОЙ ОБЛАСТИ**

## ОТ СОСТАВИТЕЛЕЙ

Настоящий тематический выпуск журнала «Электронные библиотеки» включает выборочные материалы 10-й Международной научно-практической конференции «Электронная Казань 2018» (Информационные технологии в современном мире), которая состоялась 24–25 апреля 2018 года в Институте социальных и гуманитарных знаний (ИСГЗ), г. Казань. Данная конференция проводится ИСГЗ совместно с Казанским (Приволжским) федеральным университетом, Экономическим университетом в Братиславе (Словакия) и Российским экономическим университетом им. Г.В. Плеханова (г. Москва), начиная с 2009 года (портал конференции с архивами всех предыдущих: <http://e-kazan.info>).

Конференция собрала более 100 человек из разных городов России, в том числе, Москвы, Санкт-Петербурга, Ульяновска, Волгограда, Оренбурга, Рязани, Самары, Йошкар-Олы, Симферополя, Улан-Удэ, Нижнего Новгорода, Новосибирска, Махачкалы, Набережных Челнов, Чебоксар, Ярославля, Чистополя, Альметьевска и, конечно же, Казани. Приехали также специалисты из Казахстана, Словакия и США.

Основные темы, заявленные в докладах, раскрывают современные направления развития электронного обучения и информационных технологий в образовании и других сферах общественной жизни. В зоне внимания авторов: электронные ресурсы и их продвижение в информационную образовательную среду учебного заведения, инструменты поддержки учебно-методического обеспечения образовательных программ, поддержание и развитие информационных компетенций студентов. Одна из ведущих тем – использование современных технологий совмещенной реальности и интернета вещей при реализации образовательных программ, отдельных учебных курсов.

Лейтмотивом, связующим все выступления, прозвучавшие на конференции, стала тема о стремительном развитии и нарастании влияния современных информационных технологий во всех сферах жизни общества в условиях глобальных вызовов современности. В полной мере это нашло отражение и в материалах, представленных в данном выпуске.

Е. П. Чирко, В. И. Зуев

УДК 004.91

## ИНФОРМАЦИОННО-АНАЛИТИЧЕСКИЕ СИСТЕМЫ В АНТИКОРРУПЦИОННОЙ БОРЬБЕ

Н. И. Баяндин

Российский экономический университет им. Г.В. Плеханова, г. Москва

bajanick@rambler.ru

### ***Аннотация***

Рассмотрен вопрос возможности применения новых информационных технологий при организации антикоррупционных мероприятий. Показано, что уменьшение влияния человеческого фактора становится важным фактором противодействия коррупции. Обозначена программа обучения методам противодействия коррупции в рамках магистерской программы.

***Ключевые слова:*** коррупция, антикоррупционная борьба, информационные технологии, человеческий фактор, магистерская программа

В Стратегии национальной безопасности Российской Федерации [1] коррупция определена как один из главных рисков и угроз национальной безопасности в экономической и политической сферах. По оценкам специалистов в области изучения коррупции и борьбы против нее, за последние годы коррупция разрослась до глобальных масштабов, перестав быть только национальной, внутригосударственной проблемой. Она приобрела трансграничный характер, и этот факт не может не беспокоить как национальные правительства, так и международные организации, которые в последнее десятилетие активизировали деятельность по борьбе с коррупцией. В ООН принята «Декларация о борьбе с коррупцией и взяточничеством в международных коммерческих операциях» [2] и разработана Глобальная Программа Борьбы с Коррупцией. Во многих странах разрабатываются и принимаются национальные законы и государственные программы борьбы с коррупцией, создаются международные антикоррупционные комитеты. Международными неправительственными организациями проводятся широко-масштабные исследования проявления коррупции во многих странах, в которых изучаются социально-экономические предпосылки и причины ее возникновения.

В 2016 году Международной организацией стандартизации был принят международный стандарт ISO 37001-2016 «Системы менеджмента противодействия коррупции» [3]. В России борьба с коррупцией проводится в соответствии с требованиями Национального плана противодействия коррупции на 2016–2017 годы, утвержденного Указом Президента Российской Федерации от 01.04.2016 № 147 [4].

Одним из наиболее эффективных методов противодействия коррупции является устранение влияния человеческого фактора из коррупционной цепочки. Современные информационные технологии позволяют выявлять многие коррупционные признаки человека, его предрасположенности, аффилированность и др. Вопросы о личности коррупционера чрезвычайно важны, так как именно через личность можно с достоверностью установить причину коррумпированных связей в каждом конкретном случае. Во-многом причина любых действий человека кроется в нем самом и зависит от многих его личностных качеств и свойств биологического и социального плана. Многие эксперты считают, что у подавляющего большинства людей, если не у всех, честность имеет свою цену. В подтверждение этого приводится много примеров. «Так, президент США Авраам Линкольн, признанный современниками честнейшим человеком, однажды вышвырнул из кабинета человека, предложившего ему большую взятку. Когда его спросили, чем это он так выведен из себя, А. Линкольн ответил: «У каждого есть своя цена, а он подобрался слишком близко к моей!» [5].

Коррупционные преступления, основанные на личных качествах тех или иных чиновников, включают признаки, которые могут сигнализировать об их нечестности или высокой степени возможности к разложению и, как следствие, к продажности. Эту проблему понимали и в 1930-е годы [6]. Современные информационные технологии позволяют выявить эти признаки при помощи соответствующего мониторинга информационного пространства, в первую очередь, социальных сетей. Наиболее часто встречаются следующие признаки:

- наличие больших личных долгов или финансовых запросов;
- наличие дорогостоящих привычек и интересов (неумеренное употребление спиртных напитков, увлечение азартными играми и т. п.);
- неясное или уголовное прошлое;

- наличие в ближайшем окружении лиц, имеющих высокое материальное благосостояние, достигнутое за счет коррупционной или другой криминальной деятельности;
- неэтичное поведение на службе (систематическое присутствие на банкетах, проводимых организациями, контроль деятельности которых входит в компетенцию данного лица; присвоение дорогих подарков; бесплатное пользование услугами, подлежащими оплате; устройство близких на высокооплачиваемую работу в подконтрольных коммерческих организациях и т. п.);
- отсутствие реагирования на факты коррупции и др.

Распознавание этих и других признаков причастности отдельных чиновников к коррупционным преступлениям является наиболее важной частью их выявления.

В настоящее время использование информационных технологий в подготовке и принятии управленческих технологий предоставляет возможность создания системы подготовки и принятия управленческих решений с чрезвычайно высокой антикоррупционной устойчивостью. Выявление антикоррупционных угроз можно разделить на антикоррупционный мониторинг персонала, подлежащего антикоррупционному контролю, и антикоррупционный мониторинг контрагентов. Можно выделить несколько новых информационных технологий, позволяющих эффективно бороться с коррупционными проявлениями. В антикоррупционной деятельности службы безопасности любого предприятия можно выделить две группы применяемых технологий:

1. **Кадры.** Технологии, позволяющие выявить коррупционные признаки при приеме на работу (полиграф, методы психозондирования, различные виды тестирования), оценке лояльности сотрудника (мониторинг социальных сетей, DLP системы – *аббревиатура DLP расшифровывается как Data Leak Prevention, т. е. предотвращение утечек данных*), контроль переговоров и др.

2. **Партнеры\контрагенты.** Технологии, позволяющие выявить коррупционные признаки при взаимодействии с партнерами при заключении сделок (информационно-аналитические системы, базы данных, мониторинг социальных сетей).

Рассмотрим первую группу – антикоррупционный мониторинг персонала,

под которым понимаем такое направление работы, при котором изучению подвергается личность контролируемого лица (сотрудник компании, его родственники) путем изучения его бизнес-активности. В частности, проверяется нарушение сотрудником государственной компании запрета на ведение коммерческой деятельности (ст. 349.1 Трудового кодекса РФ). Проверка по данному критерию (причастность к деятельности какой-либо коммерческой компании) обеспечивается возможностями информационно-аналитической системы (ИАС) по поиску информации о контролируемом лице (запрос по фамилии, имени, отчеству и ИНН контролируемого лица). Также важным условием является проверка родственников и знакомых контролируемого лица, имеющих отношение к этой сфере бизнеса. Эта деятельность начинается в момент приема гражданина на соответствующую должность и продолжается в ходе всей его трудовой деятельности. В свете требований международного стандарта ISO 37001-2016 «Системы менеджмента противодействия коррупции» [3] следует определиться с перечнем сотрудников, подлежащих антикоррупционному контролю, а также с кругом их родственников и знакомых, которые, осуществляя свою бизнес-деятельность, могут привести к ситуации конфликта интересов (к участию в сделках с заинтересованностью). Согласно рассматриваемому стандарту и российскому законодательству, публичное должностное лицо по своему статусу относится к категории лиц, занимающих наиболее коррупционно-емкие должности, и является носителем коррупционных рисков. Согласно ч. 1 ст. 81 Федерального закона № 208-ФЗ «Об акционерных обществах» [7] и ч. 1 ст. 45 Федерального закона № 14-ФЗ «Об обществах с ограниченной ответственностью» [8] ограничения устанавливаются на сделки, в совершении которых имеется заинтересованность следующих лиц:

- членов Совета директоров (Наблюдательного совета) общества;
- лиц, осуществляющих функции единоличного исполнительного органа общества, в том числе, управляющей организации или управляющего; членов коллегиального исполнительного органа общества;
- акционеров общества, имеющего совместно с его аффилированными лицами 20 и более процентов голосующих акций общества;
- лиц, имеющих право давать обществу обязательные для него указания.

Использование возможностей ИАС является единственным способом легитимно получить требуемую информацию. При проведении антикоррупционных мероприятий могут использоваться информационно-аналитические системы «СПАРК», «СПАРК-Маркетинг», «X-Compliance», «Контур-Фокус», «Интегрум» и др. [9]. Исходя из общих понятий теории риск-менеджмента, следует сформировать два информационных модуля. В первый модуль входит превентивная информация, т. е. общедоступные сведения обо всех знакомых сотрудника компании, подлежащего антикоррупционному контролю; во второй модуль входит ситуационная информация, т. е. подлежащая проверке версия об аффилированности по критерию «наличие близких отношений» сотрудников проверяемой компании и руководства компании-контрагента, подлежащих антикоррупционному контролю.

Эти антикоррупционные задачи решаются путем проверки состава учредителей, директоров и других персоналий проверяемой компании. Выявив их круг, проверяют аффилированность между персоналом проверяемой компании и сотрудниками компании-контрагента. Кроме того, проверяется в рамках деловой репутации запятнанность компании-контрагента в коррупционных скандалах и тому подобная информация. Необходимая искомая информация должна соответствовать следующим критериям: она должна быть достаточной, чтобы выполнить антикоррупционные задачи, поставленные Стандартом и Федеральными законами; эта информация должна быть проверяема легитимными методами. Информация тогда ценна, если ее можно проверить. При приеме на работу нового и при анкетировании действующего сотрудника должны быть получены сведения обо всех их родственниках и знакомых, которые требуются упомянутыми законами. Минимум информации, достаточный для идентификации, – фамилия, имя, отчество контролируемого лица и его ИНН. Информационно-аналитическая система позволит получить достоверную информацию о бизнес-активности этого человека в рамках конкретного юридического лица или в статусе индивидуального предпринимателя. Сведения, полученные из информационных массивов ИАС, нуждаются в аналитическом осмыслении в привязке к специфике проверяемой компании. Поэтому важным фактором борьбы с коррупцией становится исполь-

зование аналитических систем типа «Семантический архив» [9]. Также рекомендуется формировать на предприятии собственные базы антикоррупционных данных, включающие информацию об аффилированности определенной компании контролируемому сотруднику проверяемой компании или его родственнику.

В принципе эта работа мало чем отличается от действий по проверке потенциальных контрагентов, и тут незаменимым подспорьем являются широко применяемые на практике ИАС «СПАРК» и «Контур. Фокус» [9]. Применяемые в «СПАРК Маркетинге» уникальные технологии обработки данных позволяют выявлять признаки нарушений и коррупционные схемы в торгах и контрактах. Эта ИАС предоставляет информацию о руководителях компаний. Возможности «СПАРК» позволяют исследовать взаимную аффилированность компаний, а также возможную аффилированность работников, имеющих отношение к организации конкурсов, и их родственникам. Эта информация вводится в свои корпоративные базы данных «Возможна коррупция – физические лица» и «Возможна коррупция – юридические лица». Весьма эффективным в этом представляется решение, предлагаемое ИАС «Контур. Фокус»: ее модуль «Фокус. Форум» предоставляет возможность пользователю системы оставлять на странице своего контрагента комментарии о различных аспектах сотрудничества с этим бизнес-партнером (своего рода «сарафанное радио», эффективность которого трудно переоценить). Например, эти отзывы могут содержать информацию о задержке платежей, недобросовестности контрагента. Заметки могут разоблачать и контрагента-коррупционера.

Таким образом, мониторинг рисков коррупции является одним из важнейших направлений деятельности компании. Это означает, что специализированное подразделение компании (в Стандарте оно названо «Службой обеспечения соблюдения антикоррупционных норм») обязано с определенной периодичностью проводить колоссальный объем работы. Обеспечить эффективное решение этих задач возможно только при широком применении информационно-аналитических систем и собственных антикоррупционных баз данных.

Для подготовки специалистов, способных профессионально работать с новыми информационными технологиями в сфере противодействия коррупции, в РЭУ им. Г.В. Плеханова с 2018 года начинается прием в магистратуру по профилю программы «Защита информационного пространства субъектов экономической деятельности».



## СПИСОК ЛИТЕРАТУРЫ

1. Указ Президента Российской Федерации от 31.12.2015 г. № 683 «О Стратегии национальной безопасности Российской Федерации». URL: <http://kremlin.ru/acts/bank/40391>

2. Декларации ООН по борьбе с коррупцией и взяточничеством в международных коммерческих операциях (Резолюция Генеральной Ассамблеи 51/191 от 16 декабря 1996 г., 86-е заседание). URL: [http://www.un.org/ru/documents/decl\\_conv/declarations/bribery.shtml](http://www.un.org/ru/documents/decl_conv/declarations/bribery.shtml)

3. *Глебовский А.* Кадровые аспекты антикоррупционной работы в свете требований ISO37001-2016//Директор по безопасности. 2017. №7. С. 28–35.

4. Указ Президента Российской Федерации от 01.04.2016 № 147 «О Национальном плане противодействия коррупции на 2016–2017 годы». URL: <http://pravo.gov.ru/proxy/ips/?docbody=&nd=102393795&intelsearch>

5. *W. Steve Albrecht, Gerald W. Wernz, Timothy L. Williams.* Fraud: Bringing Light to the Dark Side of Business. Burr Ridge, Ill.: Irwin Professional Pub., 1995. 296 p. (*Альбрехт С., Вернц Г., Уильямс Т.* Мошенничество. Луч света на темные стороны бизнеса. Перев. с англ. СПб: Питер, 1995. 400 с.).

6. *Епихин А.Ю., Мозохин О.Б.* ВЧК – ОГПУ в борьбе с коррупцией в годы новой экономической политики (1921 – 1928). М.: «Кучково поле», 2007. 71 с.

7. Федеральный закон № 208-ФЗ от 26.12.1995 «Об акционерных обществах». URL: <http://kremlin.ru/acts/bank/8653>

8. Федеральный закон №14-ФЗ от 08.02.1998 «Об обществах с ограниченной ответственностью». URL: <http://kremlin.ru/acts/bank/11994>

9. *Баяндин Н.И.* Информационно-аналитическое обеспечение безопасности бизнеса. Деловая разведка. СПб.: ИЦ Интермедия, 2016. 264 с.

## INFORMATION-ANALYTIC SYSTEMS IN ANTICORRUPTION WARS

N. I. Bayandin

*Plekhanov Russian University of Economics, Moscow*

bajanick@rambler.ru

### **Abstract**

The article tells about the using of new information technologies in anticorruption wars. Human factor as key element of anticorruption wars in business are explained. The information on programs of improvement of professional skill in anticorruption wars are given.

**Keywords:** *corruption, anticorruption wars, human factor, information technologies, educational program of magisters.*

### **REFERENCES**

1. Ukaz Prezidenta Rossijskoj Federacii ot 31.12.2015 g. № 683 «O Strategii nacional'noj bezopasnosti Rossijskoj Federacii». URL: <http://kremlin.ru/acts/bank/40391>
2. Deklaracii OON po bor'be s korrupciej i vzyatochnichestvom v mezhdunarodnyh kommercheskih operacijah (Rezolyuciya General'noj Assamblei 51/191 ot 16 dekabrja 1996 p., 86-e zasedanie). URL: [http://www.un.org/ru/documents/decl\\_conv/declarations/bribery.shtml](http://www.un.org/ru/documents/decl_conv/declarations/bribery.shtml)
3. *Glebovskij A.* Kadrovyje aspekty antikorrupcionnoj raboty v svete trebovanij ISO37001-2016//Direktor po bezopasnosti. 2017. №7. S. 28–35.
4. Ukaz Prezidenta Rossijskoj Federacii ot 01.04.2016 № 147 «O Nacional'nom plane protivodejstviya korrupcii na 2016–2017 gody». URL: <http://pravo.gov.ru/proxy/ips/?docbody=&nd=102393795&intelsearch>
5. *W. Steve Albrecht, Gerald W. Wernz, Timothy L. Williams.* Fraud: Bringing Light to the Dark Side of Business. Burr Ridge, Ill.: Irwin Professional Pub., 1995. 296 p. (Al'brekht S., Vernc G., Uil'yams T. Moshennichestvo. Luch sveta na temnye sto-rony biznesa. Perv. s angl. SPb: Piter, 1995. 400 s.).
6. *Epihin A.YU., Mozohin O.B.* VCHK – OGPU v bor'be s korrupciej v gody novoj ehkonomicheskoj politiki (1921 – 1928). M.: «Kuchkovo pole», 2007. 71 s.

7. Federal'nyj zakon № 208-FZ ot 26.12.1995 «Ob akcionernyh obshche-stvah». URL: <http://kremlin.ru/acts/bank/8653>

8. Federal'nyj zakon №14-FZ ot 08.02.1998 «Ob obshchestvah s ogranichen-noj otvetstvennost'yu». URL: <http://kremlin.ru/acts/bank/11994>

9. *Bayandin N.I.* Informacionno-analiticheskoe obespechenie bezopasnosti biznesa. Delovaya razvedka. SPb.: IC Intermediya, 2016. 264 s.

### **СВЕДЕНИЯ ОБ АВТОРЕ**



**БАЯНДИН Николай Иванович** – заведующий учебно-научной лабораторией «Информационное противоборство в бизнесе», лауреат Национального форума по информационной безопасности «Инфофорум 2017».

**N. BAYANDIN**, head of the educational and research laboratory "Information wars in Business", laureate of the National Forum on Information Security "InfoForum 2017" .

email: [bajanick@rambler.ru](mailto:bajanick@rambler.ru)

*Материал поступил в редакцию 31 мая 2018 года*

УДК 004.089

## ИНТЕРНЕТ ВЕЩЕЙ – «ЦИФРОВЫМ АБОРИГЕНАМ» И ИХ РОДИТЕЛЯМ

Д. А. Богданова

Институт кибернетики и образовательной информатики  
Федерального исследовательского центра «Информатика и управление»  
Российской академии наук, г. Москва

d.a.bogdanova@mail.ru

### **Аннотация**

На основании публикаций в западной научной литературе и средствах массовой информации рассмотрено разнообразие интернет-игрушек, предлагаемых на рынке в настоящее время. Сформулированы ожидания по взаимодействию компаний-производителей с потребителями продукции. Приведены рекомендации для родителей по покупке и пользованию интернет-игрушками.

**Ключевые слова:** интернет-игрушки, безопасность интернет-игрушек, соблюдение конфиденциальности, персональные данные, защита конфиденциальности, цифровые аборигены, игрушки «в реале», «Хелло, Барби!»

### **ВВЕДЕНИЕ**

Спектр интернет-игрушек широк, он включает и «игрушки в реале», и роботизированные игрушки, и носимые устройства, и образовательные развивающие игрушки, и многие другие. Игрушки имеют разные формы, цвет, размеры – от говорящих кукол до роботов и шлемов для пейнтбола с геолокационными функциями. Эти различные типы игрушек могут использоваться детьми по-разному и собирать разные виды и количество данных [1]. Эти факторы должны учитываться при разработке рекомендаций и выборе настроек в каждом конкретном случае.

## **1. СОВРЕМЕННЫЕ ИНТЕРНЕТ-ИГРУШКИ И ПРОБЛЕМЫ КОНФИДЕНЦИАЛЬНОСТИ**

«Игрушки в реале» – это, как правило, любимые игрушки или герои любимых видеоигр, получившие физическое воплощение, имеющие определенную индивидуальность. Они оказались чрезвычайно популярными у «цифровых аборигенов».

ригенов», чувствующих себя вполне комфортно с новыми разработками, сочетающими в себе элементы физической и виртуальной игр. Зачастую «игрушки в реале» требуют использования приложения только для первоначального запуска. Так, например, «Хелло, Барби!» требует приложения для первоначального запуска, чтобы получить согласие родителей ребенка и подключиться к сети Wi-Fi. После того, как первоначальная процедура запуска завершена, кукла начинает взаимодействовать с ребенком совершенно независимо, не нуждаясь более в использовании приложения. Родители имеют доступ к личным данным ребенка через отдельный онлайн аккаунт. Роботизированные игрушки могут нередко управляться на расстоянии, иногда с помощью ручного контроллера, иногда – с помощью речевых команд. Иногда роботизированные игрушки управляются через приложения, другие – вполне интерактивны без использования приложений или контроллеров. Данные, собираемые игрушкой? могут иногда быть очень экстенсивными, т. к. игрушке для нормальной работы необходимо анализировать значительный объем информации о том, что ее окружает. Это может быть геолокационная и пространственная информация, видео-, аудио- и другие виды сенсорной информации о среде, в которой робот функционирует. Кроме того, настраиваемые предпочтения могут оказывать влияние на то, какую информацию о пользователе робот раскрывает в данный конкретный момент. Чтобы оценить возможности той или иной информации, раскрываемой роботом, родители или иные пользователи должны проанализировать, какие данные собираются и как они обрабатываются: в игрушке или вовне. Поскольку роботы становятся все более продвинутыми, им может требоваться все больше информации. Однако они могут быть способны обрабатывать информацию самостоятельно или обходиться меньшим объемом информации, используя её более эффективно.

Носимые игрушки – это, как правило, игрушки, используемые как браслеты или нарукавные повязки, или это шлемы, оборудованные сенсорами. Чрезвычайно популярные среди детей, эти игрушки могут собирать довольно много «чувствительной» информации, как, например, точные геолокационные данные, частоту пульса и другие данные о здоровье. Такая информация должна подвергаться особо строгому контролю [2].

Развивающие образовательные игрушки ориентированы на выработку у ребенка умения или формирование знания в какой-либо области. Такие игрушки обычно создаются для использования вне классной комнаты, например, зубная щетка, которая учит правильно чистить зубы, или прививают первоначальное представление о предметах, их формах и объемах. Игрушки этой категории не обязательно отличаются от других по количеству собираемой информации, хотя они, как правило, собирают более детальную информацию о конкретном умении.

В свете растущей популярности интернета игрушек очень важно, чтобы производители игрушек и разработчики приложений придерживались строгих правил в процессе разработки, продажи и поддержки интернет-игрушек и связанных с этим услуг. В настоящее время 66% родителей проводят исследования, читают обзоры, консультируются с другими родителями, спрашивая их рекомендации относительно интернет-игрушек. Из них 88% полагаются на обзоры продукции, 71% опираются на мнения, высказанные в блогах. Приведенные цифры говорят о том, что родители готовы потратить значительное время на сбор информации из различных источников прежде, чем принять решение о покупке интернет-игрушки. Иногда мнение и предпочтения детей также влияют на принятие родителем решения о покупке той или иной игрушки. В связи с этим необходимость разработки правил конфиденциальности и контроля безопасности играет ключевую роль при принятии родителями решения о покупке той или иной игрушки.

Производители интернет-игрушек собирают данные, необходимые для того, чтобы игрушка могла работать, и чтобы иметь возможность совершенствовать ее работу с течением времени. Однако перед ними стоит задача, чтобы родители, которые купят эту игрушку для своего ребенка, чувствовали себя комфортно, зная, какая информация собирается о ребенке и как она используется. Интернет-игрушки портативны, как правило, не имеют экрана и обычно сконструированы таким образом, что случайный наблюдатель не сможет отличить интернет-игрушку от неподключенной игрушки. И хотя, как правило, интернет-игрушки подразумевают онлайн сервисное сопровождение, покупают их обычно в обычных магазинах [3]. Компании должны быть открыты перед публикой относительно использования данных настолько, насколько это возможно. И хорошим первым шагом в этом направлении может быть представление краткой информации о конфиденциальности в тех пунктах, где родители пользуются услугами или

по запуску игрушки, или по использованию приложений, или в момент регистрации аккаунта. Компании могут дополнить эти краткие уведомления дополнительной четкой информацией для ситуаций, когда родители будут пытаться взаимодействовать с игрушкой, например, в процессе чтения пользовательской инструкции или посещения сайта компании-производителя игрушки. Родители должны понять уже в момент продажи, т.е. до того, как они принесут игрушку в дом, потребуется ли в дальнейшем их согласие на то, чтобы игрушка собирала персональную информацию об их ребенке. Полная информация о политике конфиденциальности, размещенная на упаковочной коробке, вряд ли окажется полезной, однако ключевые моменты смогут помочь родителям понять, прежде чем совершить покупку, не вызывает ли информация об игрушке сомнения и не имеет ли смысл провести дополнительное исследование. Например, на коробке может быть сказано, что родителям потребуется создать личный онлайн аккаунт для того, чтобы получить доступ ко всем возможностям игрушки. Или, например, может быть сказано, что потребуется согласие родителей на то, чтобы игрушка использовала персональные данные ребенка, т. к. это необходимо для запуска игрушки. Например, многие компании, продукция которых имеет политику конфиденциальности, сообщают об этом только на своем сайте, о чем родители могут не знать. Уведомление должно быть размещено во всех доступных местах, где родитель будет пользоваться сервис-услугами для инсталляции и запуска приложений. В мире интернет-игрушек, многие из которых не имеют экрана, особенно важно целостно воспринимать уведомления подобного рода. Компаниям следует инвестировать в творческие или интуитивные способы информирования родителей и детей о том, когда данные собираются или передаются, используя графические, аудио- или тактильные знаки. Например, кукла «Хелло, Барби!» имеет заметный Wi-Fi индикатор и реагирует на «пробуждающую» фразу миганием и звуковым сигналом. Извещения подобного рода будут способствовать тому, что игрушка не вызовет у родителей удивления и неприятия факта сбора информации. Когда родители купили игрушку, особенно если при покупке присутствовал ребенок, который просил ее купить, маловероятно, что они не дадут своего согласия на сбор и использование данных.

В настоящий момент значительное число интернет-игрушек использует приложение или онлайн-интерфейс, которому родители могут дать согласие на сбор данных. Таким образом, дизайн используемого мобильного приложения очень важен. Хотя мобильная операционная система может в какой-то момент не требовать его, компания-поставщик игрушки должна разместить в App Store уведомление о конфиденциальности так же, как и в самом приложении, так и на любой другой платформе, к которой родитель предположительно может иметь доступ. Механизм согласия в приложении должен быть удобным в использовании. Для того чтобы родители могли осознанно контролировать передачу данных об их ребенке, компании должны тщательным образом продумать пользовательский интерфейс и степень дробления ступеней родительского согласия. Очень важно, чтобы он был интуитивно понятен, но и не переусложнен, чтобы не ошеломить родителей. Операционные системы требуют более дробного и нюансного контроля вместо одного общего «согласия на все», даваемого однократно в процессе загрузки приложений. Провайдеры игрушек, особенно те, которые собирают персональную информацию, несут ответственность за ее сохранность.

## **2. РЕКОМЕНДАЦИИ ДЛЯ РОДИТЕЛЕЙ ПРИ ПОКУПКЕ ИНТЕРНЕТ-ИГРУШЕК**

Ниже приводятся рекомендации для родителей, на какие вопросы следует получить ответы перед покупкой игрушки и советы по пользованию для тех, кто уже купил игрушку [4].

### **1. Очевидна ли необходимость использования интернета?**

Иногда из описания на упаковке или в интернет-магазине не совсем ясно, что игрушка была разработана для использования с беспроводным соединением с WiFi и / или Bluetooth. В некоторых случаях такое соединение может потребоваться для того, чтобы активировать игрушку. А некоторым игрушкам Интернет может быть необходим для работы.

2. Пользователям не всегда ясно, что происходит с их онлайн-данными, даже и в тех случаях, когда некоторые личные данные требуются для запуска игрушки. Поэтому полезно иметь гарантию конфиденциальности с указанием законов, которым игрушка соответствует. Родители могут также предпринять собственные шаги для того, чтобы снизить риск взлома и доступа к данным.

### **3. К кому можно обратиться, если возникнут вопросы?**



Родителям практически невозможно точно установить, каким именно стандартам и требованиям изготовитель должен соответствовать. Это происходит в результате того, что не все юридические требования должны быть отражены в упаковке. Желательно поинтересоваться, получает ли игрушка регулярные обновления с позиции требований безопасности. Полезно посетить веб-сайт производителя, чтобы узнать, есть ли служба поддержки, в которую можно будет обращаться. Имеет смысл поискать сайты или форумы, посвященные существующим проблемам.

4. Понятно ли, с чем на самом деле играет ребенок?

Желательно наблюдать или играть вместе с ребенком, когда он только начинает использовать игрушку, и продолжать следить за любыми эффектами, которые могут возникнуть в результате игры ребенка с интернет-игрушкой. Соответствует ли это собственным ценностям и идеям родителей? Например, предлагает ли игрушка определенные идеи о том, как мальчики и девочки должны вести себя? Следует спросить себя как родителя, присутствует ли готовность принять особенности игрушки, которые поощряют определенные формы поведения, и оценить, насколько сильным является это влияние.

5. Соответствует ли игрушка возрасту ребенка?

Не каждая интернет-игрушка подходит для каждой возрастной группы. В отдельных случаях они даже могут быть вредными. Указанный возрастной диапазон – это просто рекомендация производителя. Вопрос о том, какому возрастному диапазону игрушка подходит, касается не только возможных рисков, но также и развития и интересов ребенка: действительно ли подготовлен ребенок для этой игрушки?

6. Действительно ли игрушка образовательная?

Когда игрушки продаются как «обучающие», это не обязательно означает, что это правда. Игрушки предназначены в первую очередь для удовольствия, и обещания образовательной ценности со стороны некоторых производителей игрушек редко оправдываются. Так что не следует входить в заблуждение, увидев термин «образовательная». Также можно проверить, согласуется ли игрушка с методами обучения, используемыми в школе.

7. Предполагаются ли дополнительные расходы?

Некоторые интернет-игрушки будут требовать повторяющихся дополнительных расходов, если ребенок хочет продолжать в них играть. Следует проверить упаковку игрушки или веб-сайт производителя для получения большей информации об этом. Родительские форумы также могут быть хорошим источником информации. Информация о покупке или реклама в приложении должна быть понятна для родителей и храниться в не доступном для детей месте.

#### 8. Можно ли пользоваться игрушкой совместно?

Детям нравится играть вместе. Но всегда ли это возможно с интернет-игрушкой? Можно ли создать несколько аккаунтов или можно играть, используя только один аккаунт? Это может означать, что игрушкой может пользоваться только один ребенок. Например, могут ли братья и сестры играть с использованием одной и той же учетной записи? Опять же, информация может быть найдена на упаковке или на сайте производителя. Иначе можно связаться с ним напрямую, чтобы уточнить имеющуюся информацию.

#### 9. Предусматривает ли игрушка возможность свободной игры?

Предоставление ребенку возможности играть свободно предполагает, что он сможет использовать свой творческий подход. В этом случае надо быть готовым к тому, что ребенок будет делать что-то отличное от того, что предусмотрел производитель.

#### 10. Имеет ли игрушка срок годности?

Интернет-игрушки работают на основе специального программного обеспечения. Если производитель игрушек прекращает поддержку этого программного обеспечения, то пользование игрушкой может быть затруднено или стать небезопасным. Следует проверить, предоставил ли производитель любую информацию на эту тему и будет ли поддержка осуществляться и в отдаленном будущем. Также желательно проверить, сможет ли ребенок использовать свои данные для будущих версий игры или эти данные будут удалены.

Советы по пользованию уже приобретенной интернет-игрушкой:

- необходимо изменить пользовательское имя и пароль игрушки сразу после покупки;
- всегда следует использовать надежный пароль или PIN-код;
- предоставлять только те личные данные, что строго необходимы;

- полностью выключать игрушку, когда ребенок закончил играть или игрушка не используется некоторый период времени;
- когда возможно, создать учетную запись для ребенка, используя частично вымышленные персональные данные;
- желательно использовать двухступенчатую верификацию для защиты аккаунта;
- желательно использовать безопасное интернет-подключение: Wi-Fi с паролем или VPN соединение.

### **ЗАКЛЮЧЕНИЕ**

Будущее интернет-игрушек многообещающе, но необходимо сделать очень многое для создания атмосферы доверия между родителями, детьми, производителями игрушек, внедряющими самые современные методы обеспечения безопасности и сохранности данных, что будет способствовать развитию рынка интернет-игрушек.

### **СПИСОК ЛИТЕРАТУРЫ**

1. Paz K. The internet of toys: the digital lives of young children. URL: <https://www.euroscientist.com/internet-toys-digital-lives-young-children/>
2. Leamy E. The danger of giving your child 'smart toys'. URL: [https://www.washingtonpost.com/lifestyle/on-parenting/giving-your-child-internet-connected-smart-toys-could-bedumb/2017/09/29/a168218a-a241-11e7-8cfe-d5b912fab99\\_story.html?utm\\_term=.d8ef9d5d06b2](https://www.washingtonpost.com/lifestyle/on-parenting/giving-your-child-internet-connected-smart-toys-could-bedumb/2017/09/29/a168218a-a241-11e7-8cfe-d5b912fab99_story.html?utm_term=.d8ef9d5d06b2)
3. McReynolds E. et al. Toys that listen: A study of parents, children, and Internet connected toys. URL: [https://www.ftc.gov/system/files/documents/public\\_comments/2017/11/00038-141895.pdf](https://www.ftc.gov/system/files/documents/public_comments/2017/11/00038-141895.pdf)
4. Barrett B. Don't get your kid an Internet-connected toy. URL: <https://www.wired.com/story/dont-gift-internet-connected-toys/>

## THE INTERNET OF THINGS FOR "DIGITAL ABORIGINES" AND THEIR PARENTS

D. Bogdanova

Federal Research Center "Computer Science and Control"  
of the Russian Academy of Sciences, Moscow, Russia

d.a.bogdanova@mail.ru

### **Abstract**

The variety of connected toys currently offered on the market, based on publications in Western scientific literature and the media, is being considered. A list of parents' safety requirements for toys purchased for their children is given. Expectations on interaction between manufacturers and products consumers are presented. Recommendations for parents on buying and using Internet toys are presented.

**Keywords:** *connected toys, connected toys safety, confidentiality, personal data, privacy protection, digital aborigines, toys for life, "Hello, Barbie!"*.

### **REFERENCES**

1. Paz K. The internet of toys: the digital lives of young children. URL: <https://www.euroscientist.com/internet-toys-digital-lives-young-children/>
2. Leamy E. The danger of giving your child 'smart toys'. URL: [https://www.washingtonpost.com/lifestyle/on-parenting/giving-your-child-internet-connected-smart-toys-could-bedumb/2017/09/29/a168218a-a241-11e7-8cfe-d5b912fabc99\\_story.html?utm\\_term=.d8ef9d5d06b2](https://www.washingtonpost.com/lifestyle/on-parenting/giving-your-child-internet-connected-smart-toys-could-bedumb/2017/09/29/a168218a-a241-11e7-8cfe-d5b912fabc99_story.html?utm_term=.d8ef9d5d06b2)
3. McReynolds E. et al. Toys that listen: A study of parents, children, and Internet connected toys. URL: [https://www.ftc.gov/system/files/documents/public\\_comments/2017/11/00038-141895.pdf](https://www.ftc.gov/system/files/documents/public_comments/2017/11/00038-141895.pdf)
4. Barrett B. Don't get your kid an Internet-connected toy. URL: <https://www.wired.com/story/dont-gift-internet-connected-toys/>

## СВЕДЕНИЯ ОБ АВТОРЕ



**БОГДАНОВА Диана Александровна** – кандидат педагогических наук, старший научный сотрудник. Институт кибернетики и образовательной информатики Федерального исследовательского центра «Информатика и управление» Российской академии наук;

**область интересов:** дистанционное образование, интернет-безопасность, качество образовательных ресурсов;

**Diana BOGDANOVA**, PhD., Senior Science Researcher, Federal Research Center “Computer Science and Control” of the Russian Academy of Sciences;

**professional interests:** distance learning, Internet-safety, educational resources quality

email: d.a.bogdanova@mail.ru

*Материал поступил в редакцию 28 мая 2018 года*

УДК 02:005.6

## ИНСТРУМЕНТЫ УПРАВЛЕНИЯ ИНФОРМАЦИОННО-БИБЛИОТЕЧНОЙ ДЕЯТЕЛЬНОСТЬЮ

Ю. Н. Дрешер, Т. И. Ключенко, Е. А. Косолапова

ГАУ Республиканский медицинский библиотечно-информационный центр,  
Республика Татарстан, г. Казань

dresher07@yandex.ru, kluchenko@rambler.ru, rmbic@tatar.ru

### **Аннотация**

В статье говорится о том, что менеджмент качества – это эффективный метод управления организацией, а сбалансированная система показателей является ее инструментом, который позволяет повысить результативность деятельности и системно измерить эффективность работы организации на основе видения, ценностей и стратегии.

*Ключевые слова:* сбалансированная система показателей, менеджмент качества, интеграция, библиотечно-информационная деятельность, стратегические ориентиры, степень удовлетворенности потребителей, мониторинг, классификация факторов удовлетворенности, шкалы измерения, критерии, формула индекса удовлетворенности, анкетирование

### **ВВЕДЕНИЕ**

В наше время, когда вырос спрос на информацию и появилось множество технологических новаций (веб-порталы нового поколения, открытые мобильные платформы и мобильные коммуникации, вычислительные облака, семантический веб и др.), возникла новая концепция доступа к информации через библиотеки. Если раньше информация рассматривалась как общественный ресурс, который должен распределяться бесплатно, то теперь она воспринимается как товар, который может быть продан и куплен для частного потребления, и объем доступа к этому ресурсу зависит от оплаты. Черты этих изменений можно увидеть уже в новой терминологии: посетителей библиотек теперь называют потребителями, онлайн-пользователями, библиотекари строят бизнес-планы и т. п. Ввиду того, что финансирование на содержание библиотек сократилось и одновременно усилилась критика основ организации библиотечного дела, многие из этих

учреждений пришли к тому, что начали использовать двухзвенную модель: населению бесплатно, корпоративному пользователю — за плату. Разумеется, эта модель плохо сочетается с традиционным подходом к библиотечно-информационному обслуживанию как общественной службе, доступной всем независимо от достатка. Рекомендации проекта Европейской комиссии PULMAN-ХТ «Публичные библиотеки в век цифровой информации» нацеливают библиотеки на сотрудничество с другими службами цифровой информации с целью обеспечения по возможности наиболее широкого доступа к информации [9].

### **3. ПРЕДПОСЫЛКИ ВНЕДРЕНИЯ СИСТЕМЫ МЕНЕДЖМЕНТА КАЧЕСТВА В БИБЛИОТЕКАХ**

Современная ситуация, резкое обострение конкуренции на рынке из-за его глобализации, в том числе в области качества интеллектуальных ресурсов общества и образования, потребовала от библиотек осознать необходимость радикального изменения своей деятельности, организационной структуры, практического библиотечно-информационного обслуживания в условиях взаимоиспользования ресурсов, овладевать навыками использования не только информационных, но и гуманитарных технологий, осваивать инструментарий систем управления в контексте понимания того, что управление будет эффективным только тогда, когда управлять будут процессами и проектами [6, 10].

В конце XX века очень большое внимание уделяется качеству как главному фактору любых субъектов хозяйствования. Рынок производителя сменился рынком потребителя. Потребитель диктует, что, когда, в каком виде и по какой цене он хочет получить.

Применительно к библиотечной деятельности качество определяется как совокупность свойств (параметров) услуг, процессов и условий их предоставления, обеспечивающих удовлетворение соответствующих потребностей пользователей в данном виде обслуживания. А качество услуги определяется как совокупность свойств (параметров) услуги, обеспечивающих ее способность удовлетворять определенные (обусловленные и предполагаемые) потребности пользователя в соответствии с его запросами и ожиданиями [3].

Предпринятый специалистами анализ системы менеджмента качества (СМК) в контексте развития библиотек свидетельствует, что внедрение этой системы в библиотеках становится жизненной необходимостью [8]. Стандарт ИСО 9000-2015 так определяет содержание современного менеджмента качества: «Совокупность взаимосвязанных или взаимодействующих элементов организации для разработки политик (намерения и направление организации, официально сформулированные ее высшим руководством), целей и процессов для достижения этих целей» [1].

#### **4. СБАЛАНСИРОВАННАЯ СИСТЕМА ПОКАЗАТЕЛЕЙ**

В современных условиях эффективное управление качеством рассматривается как фактор системной организации учреждения, обеспечивающий его рыночный успех, конкурентоспособность, как необходимое условие его экономического развития. В то же время появляется много новых подходов к повышению результативности деятельности, каждый из которых имеет своих сторонников, пытающихся противопоставить его всем остальным подходам.

Одним из таких подходов является внедрение сбалансированной системы показателей (ССП), разработанной профессорами Гарвардского университета Д. Нортон и Р. Капланом (США). Данная система основана на причинно-следственных связях между стратегическими целями, отражающими их параметрами и факторами получения планируемых результатов. Согласно их мнению, ССП обеспечивает новый подход к стратегическому управлению компаниями любой сложности в разных отраслях деятельности [7].

Анализ интеграции ССП и методологии всеобщего управления качеством позволяет утверждать, что СМК и ССП следует применять как взаимодополняющие подходы. Результаты анализа наглядно представлены в таблице «Сравнительный анализ принципов СМК, TQM, критериев Европейской премии в области качества и сбалансированной системы показателей» [2]. Действительно, менеджмент качества, как уже отмечалось выше, является методом управления организацией, по сути, ее идеологией. Чтобы метод управления стал эффективным, работоспособным, необходим инструмент, позволяющий системно измерить эффективность деятельности организации на основе видения, ценностей и стратегии, отражающей наиболее важные аспекты ее деятельности.



Анализ различных подходов к реализации ССП, их достоинств и недостатков в различных организациях в контексте библиотечно-информационной деятельности позволил выделить подход Х. Фридага и В. Шмидта (известные немецкие бизнес-консультанты по внедрению ССП), акцентирующих свое внимание на интеллектуальном капитале персонала организации, и подход Л. Мейселя, поместившего показатели эффективности управления персоналом в отдельный блок оценки [4, 7, 11]. Основываясь на результатах анализа достоинств и недостатков ряда ССП, специалисты Республиканского медицинского библиотечно-информационного центра Министерства здравоохранения Республики Татарстан (РМБИЦ) провели подготовительную работу с целью внедрения собственной системы. Начали с исследования ключевого элемента успеха центра — удовлетворенности потребителей его продукцией и услугами, с проведения мониторинга и измерения этой удовлетворенности [4, 5]. Основным аргументом для принятия такого решения: информация, полученная в результате мониторинга и измерения удовлетворенности потребителей послужит основанием для определения стратегических ориентиров (миссия, ценности, видение) библиотечно-информационного центра.

Исследование осуществлялось в несколько этапов:

- 1) качественная классификация факторов удовлетворенности;
- 2) определение устойчивости измерения, выраженного в однозначности информации.

Исследование проводилось в соответствии с концептуальной моделью удовлетворенности потребителей (ISO/TS 10004:2010). Принимались во внимание косвенные показатели удовлетворенности, использовался метод мозгового штурма для систематизации требований суждений о качестве предоставляемой продукции и услуг (библиотечная, библиографическая, информационная услуги и информационно-библиотечная продукция специалистам здравоохранения (ИРИ и ДОР), информационно-правовая услуга). Набор систематизированных требований к суждениям оценивался группой экспертов в составе 23 библиотечных специалистов.

В ходе исследования:

- предложена методика разработки анкет, предназначенных для изучения удовлетворенности потребителей информационно-библиотечной и издательскими услугами;
- на основе трех критериев (обоснованность, устойчивость, точность) определялась надежность шкал измерения удовлетворенности потребителей;
- экспериментально установлено, что наиболее простым методом оценки обоснованности шкалы измерения удовлетворенности является метод повторного измерения экспертами;
- во избежание большого разброса результатов при анализе критерия устойчивости удовлетворенности потребителей потребовалась перефразировка вопросов, первоначально включенных в таблицу удовлетворенности;
- точность шкалы измерения удовлетворенности потребителей достигалась посредством сравнительного анализа величины относительной устойчивости измерения и абсолютной ее погрешности;
- выведена формула для определения индекса удовлетворенности потребителей в условиях библиотечно-информационного центра.

Полученные результаты исследования будут использованы для совершенствования системы менеджмента качества РМБИЦ и деятельности центра. Они могут быть применены в других организациях, внедривших СМК.

## **ЗАКЛЮЧЕНИЕ**

Следует отметить, что система сбалансированных показателей повышает способность организации развивать оригинальные и полезные цели и решения. Это организационный контекст, который определяет эффективность управления отдельно взятой организации.

## **СПИСОК ЛИТЕРАТУРЫ**

1. ГОСТ ISO 9000–2015. Система менеджмента качества. Основные принципы и словарь. URL: <http://docs.cntd.ru/document/1200124393>
2. Андросенко Н.В. Интеграция сбалансированной системы показателей и методологии всеобщего управления качеством // Экономика качества. 2012. № 1 (1). 4 с. URL: <http://www.eq-journal.ru>
3. Ахмадова Ю.А., Галимова Е.Г. Менеджмент качества и библиотека: учеб.-практ. пособие. М.: Либерея; Бибинформ, 2007. 88 с.

4. *Дрешер Ю.Н., Квашнина Е.А.* Модели сбалансированной системы показателей: достоинства и недостатки // Внедрение системы менеджмента качества в деятельность библиотек – главное условие достижения устойчивых конкурентных преимуществ: сб. науч. трудов / РМБИЦ МЗ РТ. Казань, 2016. С. 159–174.

5. *Дрешер Ю.Н., Квашнина Е.А., Медведев В.М.* Обоснованность измерения индекса удовлетворенности потребителей информационно-библиотечной услугой // Внедрение системы менеджмента качества в деятельность библиотек – главное условие достижения устойчивых конкурентных преимуществ: сб. науч. трудов / РМБИЦ МЗ РТ. Казань, 2016. С. 110–147.

6. *Дубровина Л.А.* Минимум управления, максимум управляемости. Руководителям библиотек о всеобщем управлении на основе качества. М.: Фаир-Гранд, 2004. 398 с.

7. *Каплан Р., Нортон Д.* Сбалансированная система показателей. От стратегии к действию; пер. с англ. 2-е изд., испр. и доп.; М.: ЗАО «Олимп-Бизнес», 2006. 320 с.

8. *Квашнина Е.А.* Система менеджмента качества — движущая сила развития библиотек // Внедрение системы менеджмента качества в деятельность библиотек – главное условие достижения устойчивых конкурентных преимуществ: сб. науч. трудов / РМБИЦ МЗ РТ. Казань, 2016. С. 30–58.

9. Публичные библиотеки в век цифровой информации: Проект рекомендаций Европейской комиссии PULMAN-XT Extending the European research Network for Public Libraries Museums and Archives (машинопись); пер. с англ. М.: Минкультуры России; Архив отдела библиотек, 2004.

10. *Суббетто А.И.* Политика качества, в том числе политика образования как база решения проблем выхода из кризиса и устойчивого развития России в XXI веке // Политика качества образования и проблема квалиметрического мониторинга в сфере образования: науч. докл. М.: Исследовательский центр проблем качества подготовки специалистов, 1998. С. 6–24.

11. *Фридаг Х.Р., Шмидт В.* Сбалансированная система показателей; пер. с нем. М. В. Лапшинова. М.: Изд-во «Омега-Л», 2006. 144 с.

## INFORMATION AND LIBRARY MANAGEMENT TOOL

Ju. N. Dresher, T. I. Klyuchenko, E. A. Kosolapova

*State Autonomic Institution «Republican Medical Library-Information Center»,  
Republic of Tatarstan, Kazan*

dresher07@yandex.ru, kluchenko@rambler.ru, rmbic@tatar.ru

### **Abstract**

The article says that quality management is an effective method of managing an organization, and a balanced system of indicators is its tool that allows to increase the effectiveness of activities and systematically measure the effectiveness of the organization based on vision, values and strategy.

**Keywords:** *the balanced scorecard, quality management, integration, library-information activity, strategic guideline, the degree of customer satisfaction, monitoring, classification of satisfaction factors, measurements scales, criteria, satisfaction index formula, questionnaire.*

### **REFERENCES**

1. GOST ISO 9000–2015. Sistema menedzhmenta kachestva. Osnovnye principy i slovar'. URL: <http://docs.cntd.ru/document/1200124393>
2. *Androsenko N.V.* Integraciya sbalansirovannoj sistemy pokazatelej i metodologii vseobshchego upravleniya kachestvom // *EHkonomika kachestva*. 2012. № 1 (1). 4 p. URL: <http://www.eq-journal.ru>
3. *Ahmadova Yu.A., Galimova E.G.* Menedzhment kachestva i biblioteka: ucheb.-prakt. posobie. M.: Liberiya; Bibinform, 2007. 88 s.
4. *Dresher Yu.N., Kvashnina E.A.* Modeli sbalansirovannoj sistemy pokazatelej: dostoinstva i nedostatki // *Vnedrenie sistemy menedzhmenta kachestva v deyatel'nost' bibliotek – glavnoe uslovie dostizheniya ustojchivyh konkurentnyh preimushchestv: sb. nauch. trudov / RMBIC MZ RT. Kazan', 2016. S. 159–174.*
5. *Dresher Yu.N., Kvashnina E.A., Medvedev V.M.* Obosnovannost' izmereniya indeksa udovletvorennosti potrebitelej informacionno-bibliotechnoj uslugoj // *Vnedrenie sistemy menedzhmenta kachestva v deyatel'nost' bibliotek – glavnoe uslovie dostizheniya ustojchivyh konkurentnyh preimushchestv: sb. nauch. trudov / RMBIC MZ RT. Kazan', 2016. S. 110–147.*

6. *Dubrovina L.A.* Minimum upravleniya, maksimum upravlyaemosti. Rukovoditel'nyam bibliotek o vseobshchem upravlenii na osnove kachestva. M.: Fair-Grand, 2004. 398 s.

7. *Kaplan R., Norton D.* Sbalansirovannaya sistema pokazatelej. Ot strategii k dejstviyu; per. s angl. 2-e izd., ispr. i dop.; M.: ZAO «Olimp-Biznes», 2006. 320 s.

8. *Kvashnina E.A.* Sistema menedzhmenta kachestva — dvizhushchaya sila razvitiya bibliotek // Vnedrenie sistemy menedzhmenta kachestva v deyatel'nost' bibliotek — glavnoe uslovie dostizheniya ustojchivykh konkurentnykh preimushchestv: sb. nauch. trudov / RMBIC MZ RT. Kazan', 2016. S. 30–58.

9. Publichnye biblioteki v vek cifrovoj informacii: Proekt rekomendacij Evropejskoj komissii PULMAN-XT Extending the European research Network for Public Libraries Museums and Archives (mashinopis'); per. s angl. M.: Minkul'tury Rossii; Arhiv otдела bibliotek, 2004.

10. *Subbetto A.I.* Politika kachestva, v tom chisle politika obrazovaniya kak baza resheniya problem vyhoda iz krizisa i ustojchivogo razvitiya Rossii v XXI veke // Politika kachestva obrazovaniya i problema kvalimetriceskogo monitoringa v sfere obrazovaniya: nauch. dokl. M.: Issledovatel'skij centr problem kachestva podgotovki specialistov, 1998. S. 6–24.

11. *Fridag H.R., Shmidt V.* Sbalansirovannaya sistema pokazatelej; per. s nem. M. V. Lapshinova. M.: Izd-vo «Omega-L», 2006. 144 s.

## СВЕДЕНИЯ ОБ АВТОРАХ



**ДРЕШЕР Юлия Николаевна** – доктор педагогических наук, профессор, директор Государственного автономного учреждения «Республиканский медицинский библиотечно-информационный центр», Республика Татарстан, г. Казань. **Julia N. DRESHER**, Doctor of Education, Professor, Director of SAI “Republican Medical Library Information Centre”, Republic of Tatarstan, Kazan.

email: dresher07@yandex.ru



**КЛЮЧЕНКО Тамара Ивановна** – доктор педагогических наук, профессор, главный библиограф Государственного автономного учреждения «Республиканский медицинский библиотечно-информационный центр», Республика Татарстан, г. Казань.

**Tamara Ivanovna KLUCHENKO**, Doctor of Education, Professor, chief bibliographer of SAI “Republican Medical Library Information Centre”, Republic of Tatarstan, Kazan.

email: kluchenko@rambler.ru



**КОСОЛАПОВА Евгения Александровна** – Заведующая сектором отдела качества Государственного автономного учреждения «Республиканский медицинский библиотечно-информационный центр», Республика Татарстан, г. Казань

**Evgenia A. KOSOLAPOVA**, Head of the quality department of SAI “Republican Medical Library Information Centre”, Republic of Tatarstan, Kazan..

email: rmbic@tatar.ru, evgkvashnina@yandex.ru

*Материал поступил в редакцию 11 мая 2018 года*

УДК 004

## НЕКОТОРЫЕ ПРОГРАММНЫЕ ИНСТРУМЕНТЫ ДЛЯ АВТОМАТИЗИРОВАННОГО ПОПОЛНЕНИЯ ТЕРМИНОЛОГИЧЕСКОГО СЛОВАРЯ ПРЕДМЕТНОЙ ОБЛАСТИ

<sup>1</sup>Р. А. Румянцев, <sup>1,2</sup>О. А. Невзорова

<sup>1</sup>Институт вычислительной математики и информационных технологий,  
Казанский (Приволжский) федеральный университет, г. Казань, ул. Кремлевская,  
д. 35, 420008

<sup>2</sup>Институт прикладной семиотики Академии наук Республики Татарстан,  
г. Казань, ул. Лево-Булачная, 36А, 420111

romanrar93@gmail.com, onevzoro@gmail.com

### **Аннотация**

Описано приложение OntoDictionary, которое предназначено для работы с научными математическими статьями и онтологиями, созданными в редакторе Protege. Приложение способно создавать словарь онтологии, разбивать его элементы на концепты и обрабатывать их в булевом поиске. Имеется функционал для выделения определённых именных групп из математических статей. Новизна заключается в создании и методе обработки именных групп, содержащих формулы. Формулы обрабатываются независимо от их типа. Построен отбор кандидатов в термины. По всему функционалу произведён ряд экспериментов с онтологией математического знания OntoMathPRO, которая также была разработана в Казанском федеральном университете.

**Ключевые слова:** математическое знание, онтология, концепт, поисковый индекс, именная группа, кандидаты в термины.

### **ВВЕДЕНИЕ**

Как известно (см., например, [1]), в области искусственного интеллекта под онтологией понимается формальная система понятий и взаимосвязей между ними, описывающая определённую предметную область.

Онтологии можно классифицировать по уровню детализации, «природе» предметных действий, степени разработки и сопровождения. Развитие онтологий

актуально сейчас, так как они используются в системах автоматизированного сбора информации, управлении корпоративными информационными ресурсами, порталах и обучающих системах [2]. Особенно быстро онтологии совершенствуются в способах доступа к данным и методах автоматического построения вывода.

Отметим ряд языков проектирования онтологий. Например, LOOM, OKBC, OCML, Flogic, LBase – примеры традиционных языков. Ontolingua, CycL, SHOE – специальные языки спецификаций онтологий. XML, RDF, RDFS, OWL – это языки, основанные на веб-стандартах, их использование является наиболее современным и актуальным [3–5]. В настоящее время большинство мировых онтологических проектов использует RDF и OWL.

В настоящей статье рассмотрена задача создания удобного, современного программного инструмента для автоматизированного пополнения терминологического словаря предметной области и максимально полного анализа концептов онтологии.

## **5. ПОСТАНОВКА ЗАДАЧИ**

Нашей целью является разработка информационной системы для автоматизированного пополнения терминологического словаря предметной области. Такая система должна иметь следующий функционал:

- конвертация содержимого онтологии в реляционное представление, сохраняющая изначальную древовидную структуру данных;
- реализация булевого поиска по концептам онтологии;
- разработка метода выделения из текста заданных NP-моделей;
- построение технологии отбора кандидатов в термины.

## **6. ОБЗОР ВЕДУЩИХ СОВРЕМЕННЫХ ИНСТРУМЕНТОВ ДЛЯ РАБОТЫ С ОНТОЛОГИЯМИ**

*Редактор Protege* [6] разработан на языке Java в Стэнфордском университете и впервые представлен в 1987 году. С тех пор Protege не переставал развиваться: постоянно публиковались новые версии приложения, предназначенные для создания, просмотра и обновления онтологий. Строение онтологии является таким же, как и иерархия каталога. Имеется графический интерфейс. Основное



преимущество данного редактора – возможность адаптации приложения к моделям разных форматов (UML, OWL, XML, RDF, текстовый и ряд других). Такой функционал получил технологическую реализацию, благодаря фреймовой структуре представления знания [7] и ряду плагинов [8].

**Система Ontolingua** [9] так же, как редактор Protege, разработана в Стэнфордском университете и является самым ранним инструментом конструирования онтологий. Этот инструмент включает сервер, состоящий из архива онтологий многих предметных областей, и язык представления знаний (Ontolingua). Изменение онтологий реализовано через веб-приложение, где присутствует возможность визуализации концептов. Предусмотрено использование множественного наследования. Работа с форматами производится посредством сервера: предусмотрена конвертация онтологий из одного языка в другой.

**Редактор OntoEdit** [10] написан на языке Java и разработан в университете Karlsruhe. Это приложение функциональностью и структурой во многом похоже на Protege, но его коды закрыты. Программа имеет удобный интерфейс и подробную документацию. Кроме стандартных методов в интерфейс входят визуализация запросов и создание аксиом. Имеется возможность создания онтологий двумя способами: на основе логического вывода и с использованием методологии.

**Редактор OilEd** [11] создан в университете Манчестера и написан на языке OIL. В данный момент также поддерживается язык DAML+OIL. Редактор имеет функционал для вычисления логического противоречия классов и скрытых отношений концептов. Однако данное приложение не способно полноценно заменить аналоги, описанные выше, по следующим причинам:

- отсутствуют средства для слияния и изменения онтологий;
- нет возможности работы с большими данными;
- отсутствует контроль версий.

Следовательно, OilEd – своеобразный блокнот, в котором можно создавать, просматривать и исследовать онтологии, но не на профессиональном уровне.

**OntoSaurus.** Это приложение разработано в виде интернет-браузера, совмещает в себе как серверную, так и клиентскую части. На сервере хранятся базы

знаний LOOM [12]. Клиентская часть реализует красивый интерфейс с возможностью просмотра и изменения онтологий. Важно отметить, что основное предназначение программы – просмотр онтологий.

**ODE (Ontological Design Environment)** [13] – это инструмент для построения онтологий, отличительной чертой которого является связь с пользователем на уровне концептов, то есть данная программа предоставляет администратору таблицы (отношений, атрибутов, концептов), затем на основании введённой информации создаёт код в Loom и Ontolingua. Утверждается, что некоторым пользователям проще проектировать онтологии именно таким способом.

## **7. ОБЗОР ОСНОВНЫХ ИСПОЛЬЗУЕМЫХ МОДЕЛЕЙ**

### **а. Модель реляционного представления онтологии**

Реляционное представление информации имеет вид совокупности двумерных таблиц, состоящих из фиксированного количества столбцов и изменяемого количества строк. Для быстрого поиска по таблице применяются индексы. Кроме того, каждая таблица должна содержать первичный ключ, то есть поле, которое имеется во всех записях и при этом однозначно определяет их. Благодаря такой структуре, процедура поиска не зависит от иерархии информации, содержащейся в таблице.

В целях выполнения научно-исследовательской работы для хранения информации использована реляционная форма данных, так как описанная выше модель обладает подходящими свойствами, а именно:

- таблица легко представима на плоскости, то есть имеет двухмерную систему координат;
- обеспечена атомарность, то есть каждое поле является более неделимым;
- столбцы должны иметь разные названия;
- все ячейки одного столбца должны иметь одинаковые параметры, а именно, тип и длину;
- все строки в таблице различны;
- последовательность записей и полей может быть любой.

Для повсеместного эксплуатирования реляционных таблиц определены нормальные формы, то есть критерии, которым должна отвечать любая из таких

таблиц. Нормализация – это процесс приведения к этим обозначенным стандартам. Иначе говоря, он сводится к упрощению схемы базы данных. В нашем проекте все три таблицы были преобразованы к 1-ой и 2-ой нормальным формам, которые соответственно гласят:

- любое поле неделимо и не имеет повторяющихся групп;
- таблица уже приведена к 1-ой нормальной форме, и каждое её поле зависит от первичного ключа.

В приложении OntoDictionary применена реляционная база данных mysql, которая была создана с помощью инструмента MySQL Workbench версии 6.3. Архитектура базы состоит из трёх таблиц (Concepts, Dictionary, ModelsNP) и изображена на рис. 1.

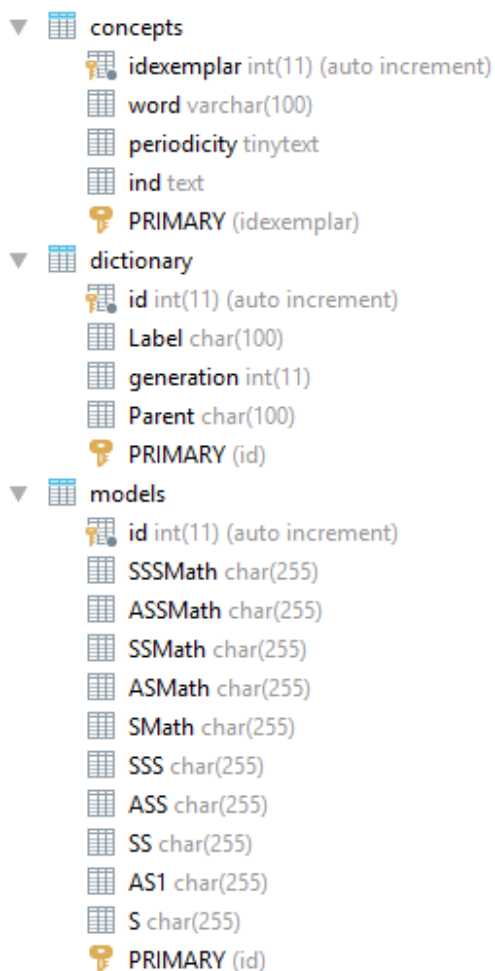


Рисунок 1. Архитектура базы данных

Таблица *concepts* содержит поля:

*Idexemplar* – тип целочисленный, авто инкремент для таблицы, первичный ключ;

*Word* – символьная строка, размером не более 100 символов;

*Periodicity* – строковая переменная, размером не более 255 байт;

*Ind* – текстовая переменная.

Таблица *dictionary* содержит поля:

*Id* – тип целочисленный, авто инкремент для таблицы, первичный ключ;

*Label* – символьная строка, размером не более 100 символов;

*Generation* – целочисленная переменная;

*Parent* – символьная строка, размером не более 100 символов.

Таблица *Models* содержит поля:

*Id* – тип целочисленный, авто инкремент для таблицы, первичный ключ;

*SSSMath* – символьная строка, размером не более 255 символов, которая предназначена для хранения моделей вида *SS<sub>2</sub>S<sub>2</sub>Math*;

*ASSMath* – символьная строка, размером не более 255 символов, которая предназначена для хранения моделей вида *ASS<sub>2</sub>Math*;

*SSMath* – символьная строка, размером не более 255 символов, которая предназначена для хранения моделей вида *SS<sub>2</sub>Math*;

*ASMath* – символьная строка, размером не более 255 символов, которая предназначена для хранения моделей вида *ASMath*;

*SMath* – символьная строка, размером не более 255 символов, которая предназначена для хранения моделей вида *SMath*;

*SSS* – символьная строка, размером не более 255 символов, которая предназначена для хранения моделей вида *SS<sub>2</sub>S<sub>2</sub>*;

*ASS* – символьная строка, размером не более 255 символов, которая предназначена для хранения моделей вида *ASS<sub>2</sub>*;

*SS* – символьная строка, размером не более 255 символов, которая

предназначена для хранения моделей вида  $SS_2$ ;

AS – символьная строка, размером не более 255 символов, которая предназначена для хранения моделей вида AS;

S – символьная строка, размером не более 255 символов, которая предназначена для хранения моделей вида S.

В разработанной базе данных можно совершать стандартные для реляционной модели операции, а именно: выделить (SELECT), вставить (INSERT), обновить (UPDATE), удалить (DELETE). Первая из них отбирает данные, с которыми будут производиться дальнейшие действия, а остальные три изменяют состояние базы данных. Пример:

- команда "SELECT \* FROM dictionary" покажет всю информацию из таблицы dictionary;
- команда "INSERT INTO dictionary (Label, generation, Parent) VALUES ('" + label + "', '" + generation + "', '" + parent + "')" вставит в таблицу dictionary в поля Label, Generation, Parent соответственно значения переменных label, generation, parent;
- команда "UPDATE models SET SSSMath = ' " + NP + "' WHERE id = '" + j + "'" обновит значение поля SSSMath в таблице models, где поле id будет равно j. В результате, ячейка таблицы примет значение переменной NP, указанной в запросе.

#### **b. Модель информационного поиска**

Под информационным поиском мы понимаем процесс отбора информации из неструктурированного хранилища. Отбор осуществляется с помощью параметров, определяемых пользователем в зависимости от ожидаемого результата. Свободная структура хранилища означает отсутствие чёткой иерархии, фильтрации переменных. Стоит отметить, что сами параметры из хранилища всё же должны обладать необходимым строением, например, относиться к известному типу данных.

Существует несколько моделей поиска (булева, векторная, с использованием вероятностей и другие), однако на сегодняшний день среди них нет доминирующей, поэтому модель подбирается индивидуально под конкретную задачу.

В приложении OntoDictionary применена булева модель поиска, так как в данном случае она даёт точный результат. Модель считается наиболее простой,

---

при этом для описываемого инструмента она является достаточно эффективной. Она была изобретена в середине двадцатого века и до сих пор используется в крупных информационных службах.

Определение этой модели подразумевает возможность осуществления запросов, состоящих из совокупности слов и простейших логических операций, таких, как “И”, “ИЛИ”, “НЕ”. Эти операции необходимы для более детального отбора информации. Реализована модель посредством наличия поискового индекса по концептам. Для удобства применён интерфейс с возможностью выбора одного из четырёх вариантов поиска:

- первый вариант – поиск элементов, содержащих концепт из запроса пользователя;
- второй вариант – поиск элементов по сочетаниям из двух концептов, заданных в запросе;
- третий вариант – поиск элементов, имеющих хотя бы один из двух концептов, заданных пользователем;
- четвертый вариант – поиск всех элементов, которые не содержат концепт, заданный в запросе.

В нашем исследовании описанная выше система поиска построена на данных, взятых из словаря. Любой термин дробится на более мелкие части, концепты. В результате операции поиска пользователю предоставляется список, где каждый термин отвечает набору концептов, заданных в запросе.

Отметим, что понятия, хранящиеся в словаре, являются уникальными, поэтому вычислять частоту использования терминов не имеет смысла. А частота употребления концептов вычисляется программой и сохраняется в специально отведённом поле базы данных.

### с. Семантические модели

Именная группа (Noun Phrase) является либо словосочетанием, в котором определено главное слово, либо одним именем существительным без зависимых слов. Основное слово называют ядром. В случае наличия зависимых слов оно определяет их характеристику. В текстах такие группы используются глаголами в роли объектов.

Чтобы вычислить NP (Noun Phrase), необходимо попытаться разрешить референциальные ссылки (местоимения). Пример:

*-Диссертация содержит новый результат /Она содержит его.*

В данном примере местоименные ссылки во второй фразе разрешаются на основе данных из первой фразы, т. к. «диссертация» и «новый результат» – это именные группы, удовлетворяющие грамматическим ограничениям. Видов NP очень много, так как размер словосочетаний в русском языке не ограничен. Как следствие, в нашей работе типы NP были отобраны самостоятельно, исходя из встречаемости в текстах выбранной тематики. Отметим, что OntoDictionary сконцентрирован на обработке математических научных статей, поэтому рассматриваются варианты, где в составе именной группы могут встречаться арифметические символы и формулы. Как следствие, нами было принято решение о разработке ранее не существующих NP-моделей, включающих формулы. Их проектирование, а также внедрение в информационную систему характеризуют научную новизну настоящей работы.

Использовались следующие обозначения:

S – имя существительное;

S<sub>2</sub> – имя существительное в родительном падеже;

A – прилагательное;

Math – формула.

С помощью этих обозначений нами были выделены именные группы:

AS – словосочетание из прилагательного и существительного, расположенных в соответствующем порядке;

$SS_2$  – словосочетание из двух существительных, второе из которых находится в родительном падеже;

$ASS_2$  – словосочетание из прилагательного и двух существительных; необходимо, чтобы второе существительное было в родительном падеже;

$SS_2S_2$  – словосочетание из трёх существительных, расположенных в соответствующем порядке; два последних слова должны находиться в родительном падеже;

$S$  – любое существительное без зависимых слов;

$ASMath$  – словосочетание из согласованных слов: прилагательного, существительного и формулы, расположенных в соответствующем порядке;

$SS_2Math$  – словосочетание из двух существительных и формулы, расположенных в указанном порядке;

$ASS_2Math$  – словосочетание из прилагательного, двух существительных и формулы, представленных в указанном порядке;

$SS_2S_2Math$  – словосочетание из четырёх слов: трёх последовательно расположенных существительных и формулы; два последних существительных должны быть в родительном падеже;

$SMath$  – словосочетание из существительного и формулы, представленных в соответствующем порядке.

Для определения именной группы и её характеристик был применен свободно распространяемый морфологический анализатор *mystem*, разработанный компанией Яндекс. Данный анализатор был выбран, так как он прост в использовании и обладает необходимым функционалом. Для работы с ним было применено несколько команд:

“-*n*” – применяется для вычисления нормальной формы слов, то есть вывода входного слова в единственном числе, именительном падеже, настоящем времени и так далее (в зависимости от части); если нормальную форму нельзя однозначно определить, анализатор предлагает несколько её вариантов;

“-*nig*” – используется для определения части речи и его семантических характеристик; в случае неоднозначности также выводится несколько вариантов.



## 8. ПРОГРАММНЫЕ ИНСТРУМЕНТЫ

Нами было создано веб-приложение, в котором объединены три разработанных инструмента. Для каждого из них имеется удобный интерфейс для взаимодействия с пользователем.

### а. Инструмент конвертирования онтологии в реляционное представление

Задача этого инструмента состоит в том, чтобы обработать выбранную пользователем онтологию и сконвертировать её в реляционное представление. Интерфейс этого инструмента представлен на рис. 2. Программная реализация инструмента состоит из нескольких функций. Первая из них – TransformationFromOwlToTxt типа void, которая принимает на вход онтологию в качестве параметра. Работа функции осуществляется посредством библиотек apache.log4j, semanticweb.owlapi и semanticweb.elk.owlapi.ElkReasonerFactory.

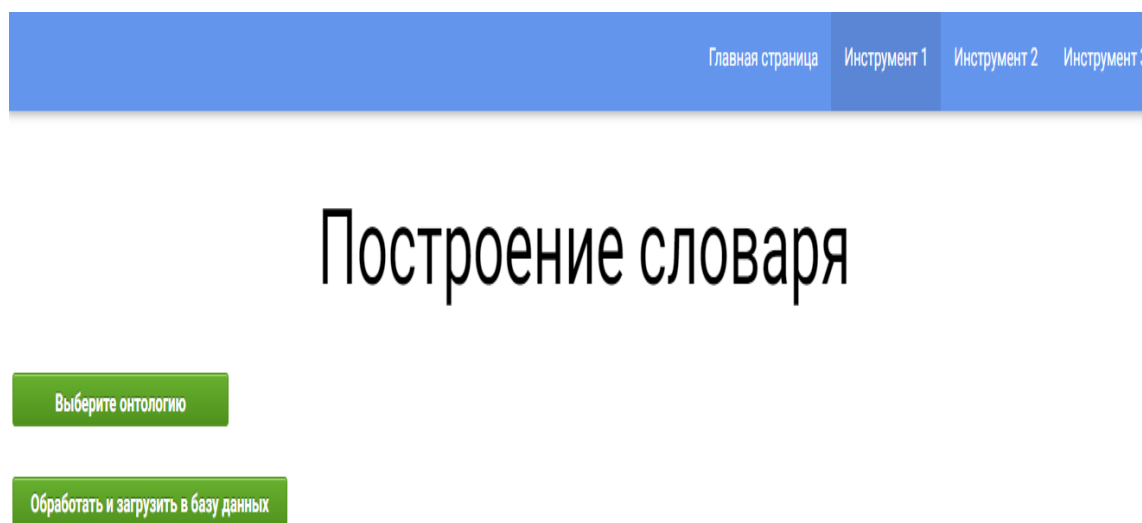


Рисунок 2. Интерфейс инструмента для построения словаря

Owlapi используется для парсера файлов такого формата. Парсер работает с использованием аргумента, reasoner. Поскольку в экспериментах на вход принималась усеченная онтология OntoMathPro, имеющая достаточно большой объём, то был использован специальный reasoner для работы с большими данными – ElkReasonerFactory.

Функция TransformationFromOwlToTxt реализует следующий алгоритм:

1. на основе входного файла создаётся новый объект класса SimpleHierarchy;
2. с использованием System.setOut указывается, что будущий выходной поток будет перенаправляется с консоли в указанный текстовый файл;
3. создаётся экземпляр ElkReasonerFactory;
4. создаётся экземпляр класса OWLClass, которому поочередно присваиваются значения классов онтологии;
5. у экземпляра класса SimpleHierarchy вызывается метод printHierarchy, принимающий в качестве аргументов экземпляр ElkReasonerFactory и экземпляр OWLClass;
6. в результате вызова printHierarchy печатается иерархия классов онтологии.

Результатом работы данной функции является файл input.txt с иерархией классов. Затем вызывается функция getStringSpaceCount для определения поколения каждого элемента иерархии. Алгоритм функции getStringSpaceCount:

1. создаётся пустой целочисленный список;
2. файл input.txt читается построчно;
3. в цикле, в каждой строке подсчитывается количество пробелов перед первым словом;
4. число, полученное на третьем шаге, добавляется в список.

Итогом работы данной функции является целочисленный список, числа в котором служат номерами поколений. Затем вызывается функция parents, которая определяет имя родительского элемента. Для самого верхнего элемента иерархии она укажет, что нет родителя. Функция принимает на вход два списка: лейблов (названий элементов) и поколений, а на выходе выдаёт список родительских элементов.

Последняя функция инструмента, writeDictionaryToDb, записывает все три списка (Labels, Generations, Parents) в таблицу Dictionary базы данных.

Алгоритм функции writeDictionaryToDb:

1. создаётся соединение с базой данных;

2. создаётся цикл for;
3. в цикле вычисляются переменные из трёх списков и вставляются в базу данных;
4. соединение с базой данных закрывается.

Как итог, имеем заполненную таблицу в базе данных. Затем данные через контроллер поступают из базы данных в соответствующую модель, Dictionary. Модель отправляет данные в представление. В результате пользователю выдается обновлённая таблица.

#### **в. Инструмент поискового индекса по концептам онтологии**

Данный инструмент предназначен для упрощения операции поиска элементов словаря. Интерфейс инструмента представлен на рис. 3. Используется механизм разбиения элементов на более мелкие части, концепты. Элементом словаря могут быть словосочетание, устойчивое выражение, условное обозначение или просто одиночное слово (например, термин). Концепт – это слово, семантическая единица языка. Таким образом, любой элемент построенного словаря предстаёт как совокупность логически связанных концептов.

За логическую связность в большинстве случаев отвечает форма слова. В данной задаче было принято решение абстрагироваться от всех возможных форм слов, то есть привести все концепты словаря к канонической форме. Такое решение позволило реализовать поиск, независимый от окончаний слов.

Приложение принимает на входе обработанный морфологическим анализатором файл, который содержит все концепты словаря. Таким образом, как промежуточный результат работы инструмента обработки OWL-файла, был создан файл с иерархией классов. Этот файл необходимо обработать морфологическим анализатором MyStem командой “-n”.

## Булев поиск по словарю

Выберите размеченный файл

Загрузить и построить индекс

Концепт1:     Концепт2:     Поиск 'или'

Концепт1:     Концепт2:     Поиск 'и'

Концепт1:     Поиск 'не'

Концепт1:     Стандартный поиск

---

| ID | Элемент | Поколение | Родитель |
|----|---------|-----------|----------|
|----|---------|-----------|----------|

Рисунок 3. Интерфейс инструмента для булевого поиска по словарю

После того, как программа получает на вход нужный файл, она начинает обрабатывать его следующим образом:

1. Поскольку MyStem возвращает файл в формате, изображённом на рис. 4, необходимо очистить его от «мусора». Необходимо извлечь из файла только слова, то есть каждую вторую строку. Для этого создаётся список `clearList`, и в функции `afterSteamList` он заполняется;
2. С помощью `HashMap` строится `frequencyList`, необходимый для вычисления частоты каждого концепта;
3. Создаётся и заполняется `ponduplic` – список без дубликатов;
4. Создаётся список `index`, где будут храниться индексы для каждого слова из списка `ponduplic`;

```
\n_____  
элемент{элемент}  
  
—  
математического{математический}  
  
—  
знания{знание}  
\n_____  
геометрический{геометрический}  
  
—  
объект{объект}  
\n_____  
координатная{координатный}  
  
—  
система{система}  
\n_____  
барицентрические{барицентрический}  
  
—  
координаты{координата}  
\n_____
```

Рисунок 4. Структура файла после морфологического анализатора

5. Заполняется список `index`, в функции `indexList`, принимающей на вход в качестве аргументов два списка: `index` и `labels` – список названий элементов из словаря. Заполняется `index` в цикле, где он сравнивается с элементами списка `nonduplic`:

- 5.1. создаётся строка индексов `ind=""` для каждого слова;
- 5.2. находится индекс первого вхождения символа "{" в подстроке;
- 5.3. с помощью `substring` выбирается имя атрибута из подстроки, то есть строка считывается с нулевого элемента по номер элемента, найденный на шаге 5.2;
- 5.4. производится обход списка `Labels`; в цикле проверяется, содержит ли строка `labels` имя атрибута, полученного на шаге 5.3; если строка включает в себя искомое слово, то включается номер этой строки в строку индексов;
6. Добавляется строка индексов `ind` для определенного слова в общий список `index`;

7. От трёх переменных (списков `frequencyList`, `nonduplic`, `index`) вызывается функция записи концептов в таблицу `concepts` базе данных; в результате на данном этапе имеется полностью заполненная таблица для поиска элементов по индексу.

Как видно из интерфейса на рис. 3, имеется 4 вида поиска элементов по вводу концептов: стандартный, поиск «и», поиск «или», поиск «не».

Алгоритм стандартного поиска:

1. принимается концепт, введённый пользователем;
2. производится поиск этого концепта по таблице `concepts`;
3. в случае нахождения искомого концепта считывается значение его поля `ind`, то есть список `id` элементов таблицы `Dictionary`, которые содержат данный концепт;
4. из таблицы `Dictionary` выводятся строки с выбранными на шаге 3 значениями поля `id`;
5. если искомый концепт не находится, возвращается пустая таблица.

Алгоритм поиска «и»:

1. принимаются два концепта, введённые пользователем;
2. для каждого из этих концептов реализуется поиск соответствующего значения по таблице `concepts`;
3. в случае нахождения обоих концептов считываются их значения с полей `ind`; эти значения представляют набор целых чисел, разделённых пробелами; строятся два множества чисел, соответствующие значениям полей `ind` для каждого концепта; затем находится пересечение этих множеств, по которому строится запрос в таблицу `Dictionary`;
4. из таблицы `Dictionary` выводятся строки, значения `id` которых находятся в пересечении множеств, найденным на шаге 3;
5. если один из концептов не найден, то возвращается пустая таблица.

Алгоритм поиска «или»:

1. принимаются два концепта, введённые пользователем;

2. для каждого из этих концептов реализуется поиск соответствующего значения по таблице `concepts`;

3. в случае нахождения обоих концептов считываются их значения с полей `ind`; эти значения представляют набор целых чисел, разделённых пробелами; строятся два множества чисел, соответствующие значениям полей `ind` для каждого концепта; затем находится объединение этих множеств, по которому строится запрос в таблицу `Dictionary`;

4. из таблицы `Dictionary` выводятся строки, значения `id` которых находятся в объединении множеств, найденных на шаге 3;

5. если один из концептов не найден, то программа будет действовать, как при стандартном поиске, где результат ищется лишь по одному концепту.

Алгоритм поиска «не»:

1. принимается концепт, введённый пользователем;

2. производится поиск этого концепта по таблице `concepts`;

3. в случае нахождения искомого концепта считывается значение его поля `ind`, то есть список `id` элементов таблицы `Dictionary`, которые содержат данный концепт;

4. из таблицы `Dictionary` выводятся все строки, кроме тех, которые были выбраны на шаге 3;

5. если искомый концепт не находится, возвращается всё содержимое таблицы `Dictionary`, без исключений.

### **с. Инструмент извлечения терминологии из математических статей**

Данный инструмент предназначен для работы с кандидатами в термины. Функционал представляет корректную обработку данных с трёх входных файлов и поиск по полученным результатам этой обработки. Интерфейс инструмента представлен на рис. 5.

## Просмотр кандидатов в термины

Загрузите научную статью

Список id

Выберите файл после анализатора

Концепт модели:

---

SSSMath    ASSMath    ASMath    SSMath    SMath    ASS    AS    SS

---

Рисунок 5. Интерфейс инструмента для просмотра и фильтрации кандидатов в термины

Для начала работы необходимо выбрать файл в формате xml и нажать кнопку «Обработать файл». После проверки на корректность файла запустится следующий алгоритм:

1. создаётся файл text.txt, куда будет записан результат работы функции;
2. с помощью библиотеки org.w3c.dom создаются экземпляры её классов Document, NodeList и Element; структура данных Document необходима для представления входного файла в нужном для библиотеки виде;
3. в экземпляр класса NodeList, названный nodeList, записывается всё содержимое файла, находящееся в теге “p”; отметим, что входному файлу не требуется определённая XML-разметка: программа универсальна, так как поддерживает любую структуру документа формата XML;
4. создаётся цикл в зависимости от размера nodeList;
5. в цикле из nodeList создаются экземпляры класса Element, которые построчно записываются в выходной файл;
6. закрывается соединение с записывающим файлом;



7. обновляется страница приложения, и появляется надпись, что файл успешно загружен.

Как итог работы функции, имеем текстовый файл с наличием всего текста из документа формата XML. Данный файл требуется обработать морфологическим анализатором MyStem, командой “-nig”. Затем пользователю следует выбрать заранее подготовленный список формул и их id и нажать кнопку «Загрузить список id». Этот файл будет загружен в корень и будет использован в функции, выделяющей NP-модели.

Следующая функция getNPResult принимает в качестве входного параметра файл «text.txt», обновлённый морфологическим анализатором. Для отсутствия путаницы теперь файл будет называться «Newtext.txt». Чтобы запустить функцию, необходимо выбрать файл и нажать кнопку «Выделить NP-модели».

Алгоритм функции getNPResult следующий:

1. обрабатывается загруженный файл со списком id для формул, то есть запускается цикл while:

1.1 пока не наступит конец файла, считывается строка;

1.2 если строка содержит подстроку «text=», то она добавляется в созданный список строк;

1.3 на выходе из цикла получается список lines, состоящий исключительно из нужных строк.

2. для каждой колонки таблицы models создаётся свой список, который будет заполняться в ходе реализации программы;

3. обрабатывается файл «Newtext.txt»:

3.1 вычисляется количество строк в файле;

3.2 строится массив, заполняющийся строками из файла;

4 создаётся цикл, где происходит работа с элементами массива (строками):

4.1 если строка содержит подстроку «S,», обозначающую, что в данной строке находится существительное, то делаются следующие шаги,

иначе программа переходит к следующей строке массива;

4.2 если следующая строка содержит подстроку «MathID», обозначающую, что в данной строке находится формула, то выполняется шаг 4.3, иначе сразу реализуется шаг 4.4;

4.3 из списка `lines` выбирается строка, из которой копируется формула и вставляется в элемент массива вместо её `id`;

4.4 на шаге 4.1 было выявлено, что текущий элемент является существительным; тогда проверяется, является ли предыдущий элемент существительным и находится ли текущий элемент в родительном падеже; если проверка показала, что это так, проверяется, стоит ли согласованное прилагательное перед ними, то есть существует ли такой объект массива и содержит ли он подстроку «А,» (обозначение прилагательного), а также совпадают ли его семантические характеристики (род, число, падеж) с семантическими характеристиками существительного, которое идёт следом (относительно текущего элемента, это предыдущий элемент); если всё совпало, то элементы массива добавляются, в зависимости от наличия формулы, соответственно в список `ASSMath` либо `ASS`;

4.5 аналогичным образом с помощью сравнений выделяются остальные NP-группы; выделим два семантических правила, которые использовались при программировании алгоритма:

- если рядом находятся два существительных, то второе из них обязательно должно быть в родительном падеже; прочие семантические характеристики у них не сравниваются;
- если рядом находятся прилагательное и существительное, то всегда проверяется, чтобы они были согласованными; для этого необходимо, чтобы у них совпали род, число и падеж;

5. полученные списки NP вставляются в таблицу `models` базы данных; для удобства просмотра было принято решение построения такой архитектуры, где строка таблицы представляла бы 10 разных элементов, а не 1 элемент с десятью свойствами, как принято в классическом представлении; для записи в такую таблицу используется следующий алгоритм:

- 5.1. из 10-ти списков NP вычисляется размер самого большого;
  - 5.2. в зависимости от этого размера в таблице базы данных создаётся число пустых строк;
  - 5.3. при записи списков NP значения пустых строк обновляются.
6. полученные значения из таблицы показываются пользователю.

Также имеется функционал для отбора из построенных NP-моделей кандидатов в термины. Для использования метода отбора необходимо ввести концепт в поле для ввода и нажать кнопку «Поиск кандидатов».

Алгоритм отбора кандидатов в термины:

1. принимается введённый пользователем концепт модели;
2. производится поиск этого концепта по всем столбца таблицы models;
3. если какое-то поле содержит искомый концепт, то строка с этим полем считывается и добавляется в список result;
4. в списке result проверяются все поля каждой строки; если поле не содержит нужного концепта, то для текущего запроса его значение становится пробелом;
5. в результате запроса изображается отфильтрованная таблица models;
6. результат запроса записывается в файл «Результат поиска.txt»;
7. если искомый концепт не находится, возвращается пустая таблица.

## 9. ПРОВЕДЁННЫЕ ЭКСПЕРИМЕНТЫ С OntoMathPro

**Эксперимент с применением инструмента конвертирования онтологии в реляционное представление.** Этот эксперимент проводился с усечённой версией специализированной онтологии математической области OntoMathPro. Результат обработки файла представлен на рисунках 6–9, на которых, соответственно, изображены фрагмент онтологии, часть выходного текстового файла, фрагмент таблицы, заполненной в результате работы инструмента, и фрагмент таблицы, отображенной в приложении.

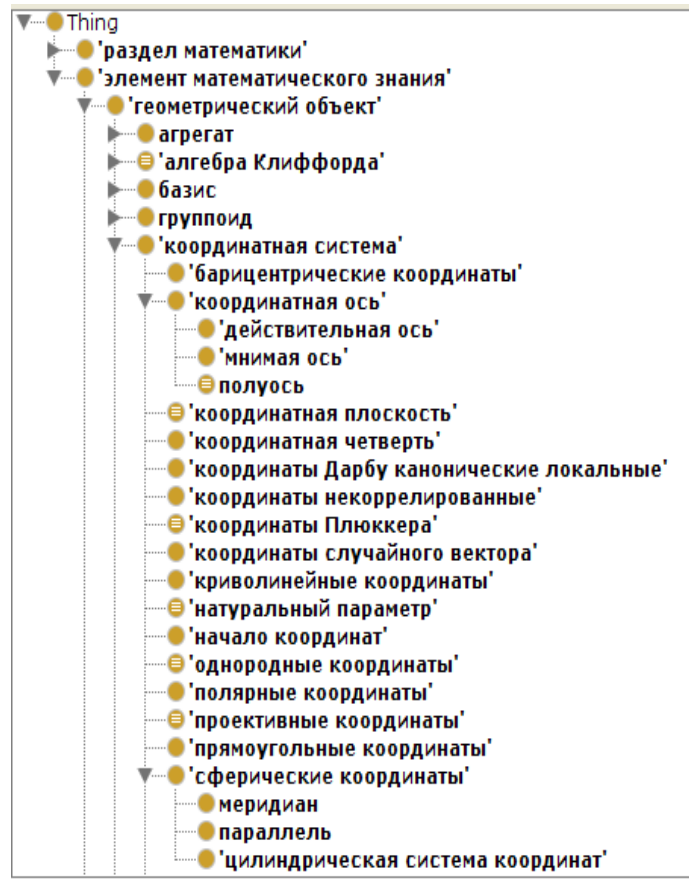


Рисунок 6. Фрагмент, использованной в ходе экспериментов, онтологии

```
элемент математического знания
  геометрический объект
    координатная система
      барицентрические координаты
      проективные координаты
      полярные координаты
      координатная плоскость
      координатная четверть
      координаты Плюккера
      координаты Дарбу канонические локальные
      сферические координаты
        меридиан
        цилиндрическая система координат
        параллель
      координаты случайного вектора
      координаты некоррелированные
      координатная ось
        полуось
        действительная ось
        мнимая ось
      натуральный параметр
      цилиндрические координаты
      криволинейные координаты
      однородные координаты
      начало координат
      декартова система координат
```

Рисунок 7. Фрагмент выходного файла

| id | Label                               | generation | Parent                         |
|----|-------------------------------------|------------|--------------------------------|
| 1  | элемент математического знания      | 1          | по                             |
| 2  | геометрический объект               | 2          | элемент математического знания |
| 3  | координатная система                | 3          | геометрический объект          |
| 4  | барицентрические координаты         | 4          | координатная система           |
| 5  | проективные координаты              | 4          | координатная система           |
| 6  | полярные координаты                 | 4          | координатная система           |
| 7  | координатная плоскость              | 4          | координатная система           |
| 8  | координатная четверть               | 4          | координатная система           |
| 9  | координаты Плуккера                 | 4          | координатная система           |
| 10 | координаты Дарбу канонические ло... | 4          | координатная система           |
| 11 | сферические координаты              | 4          | координатная система           |
| 12 | меридиан                            | 5          | сферические координаты         |
| 13 | цилиндрическая система координат    | 5          | сферические координаты         |
| 14 | параллель                           | 5          | координатная система           |
| 15 | координаты случайного вектора       | 4          | сферические координаты         |
| 16 | координаты некоррелированные        | 4          | сферические координаты         |
| 17 | координатная ось                    | 4          | координатная система           |
| 18 | полуось                             | 5          | координатная ось               |
| 19 | действительная ось                  | 5          | координатная ось               |
| 20 | мнимая ось                          | 5          | координатная система           |
| 21 | натуральный параметр                | 4          | координатная ось               |
| 22 | цилиндрические координаты           | 4          | координатная ось               |
| 23 | криволинейные координаты            | 4          | координатная ось               |
| 24 | однородные координаты               | 4          | координатная ось               |
| 25 | начало координат                    | 4          | координатная ось               |

Рисунок 8. Заполненная в ходе экспериментов таблица Dictionary

| ID | Элемент                                 | Поколение | Родитель                       |
|----|---|-----------|--------------------------------|
| 1  | элемент математического знания          | 1         | по                             |
| 2  | геометрический объект                   | 2         | элемент математического знания |
| 3  | координатная система                    | 3         | геометрический объект          |
| 4  | барицентрические координаты             | 4         | координатная система           |
| 5  | проективные координаты                  | 4         | координатная система           |
| 6  | полярные координаты                     | 4         | координатная система           |
| 7  | координатная плоскость                  | 4         | координатная система           |
| 8  | координатная четверть                   | 4         | координатная система           |
| 9  | координаты Плуккера                     | 4         | координатная система           |
| 10 | координаты Дарбу канонические локальные | 4         | координатная система           |

Рисунок 9. Отображение заполненной таблицы Dictionary

На всех четырёх рисунках показаны одни и те же данные, и по ним сделан вывод, что программа работает корректно и успешно конвертирует данные в реляционное представление.

**Эксперимент с применением инструмента для поискового индекса по концептам онтологии.** Этот эксперимент проводился со словарём, построенным в предыдущем эксперименте. Часть результата обработки этого файла представлена на рис. 10, где видно, что в базу данных была записана верная информация. Чтобы удостовериться в правильности построения индекса, разберём конкретный пример – слово «координатная», находящееся на шестой строчке рис. 10. Видно, что оно имеет частоту, равную 4, и используется в элементах словаря с id, равным 3, 7, 8, 17. Если посмотреть на рис. 8, то можно удостовериться, что это верная информация.

| idexemplar | word                                | periodicity | ind   |
|------------|-------------------------------------|-------------|---|
| 1          | элемент {элемент}                   | 7           | 1 30 579 762 1922 1973 2349                           |
| 2          | математического {математический}    | 12          | 1 762 992 1168 1189 1233 1245 1413 1574 1734 2392 ... |
| 3          | знания {знание}                     | 1           | 1   |
| 4          | геометрический {геометрический}     | 2           | 2 1379  |
| 5          | объект {объект}                     | 2           | 2 1189  |
| 6          | координатная {координатный}         | 4           | 3 7 8 17  |
| 7          | система {система}                   | 13          | 3 13 26 220 323 325 336 338 340 805 806 807 808       |
| 8          | барицентрические {барицентрический} | 1           | 4   |
| 9          | координаты {координата}             | 11          | 4 5 6 9 10 11 15 16 22 23 24                          |
| 10         | проективные {проективный}           | 1           | 5   |
| 11         | координаты {координата}             | 1           | 6   |
| 12         | полярные {полярный}                 | 16          | 7 1055 1070 1078 1118 2042 2043 2044 2097 2219 223... |
| 13         | координаты {координата}             | 1           | 8   |
| 14         | координатная {координатный}         | 1           | 9   |
| 15         | плоскость {плоскость}               | 5           | 10 924 1205 1664 2014                                 |
| 16         | координатная {координатный}         | 1           | 10  |
| 17         | четверть {четверть}                 | 1           | 10  |
| 18         | координаты {координата}             | 1           | 11  |
| 19         | Плюккера {плюккер?}                 | 1           | 12  |
| 20         | координаты {координата}             | 1           | 13  |
| 21         | Дарбу {дарба?   дарб?}              | 18          | 3 4 5 6 7 8 9 10 11 13 15 16 17 22 23 24 25 26        |
| 22         | канонические {канонический}         | 4           | 14 200 222 640  |
| 23         | локальные {локальный}               | 1           | 15  |
| 24         | сферические {сферический}           | 1           | 15  |
| 25         | координаты {координата}             | 1           | 16  |

Рисунок 10. Заполненная в ходе экспериментов таблица Concepts

Также были проведены успешные эксперименты с четырьмя типами поиска. По результатам всех экспериментов сделан вывод, что инструмент работает корректно и реализует алгоритм булевого поиска на основе концептов.

**Эксперимент с применением инструмента для извлечения терминологии из математических статей.** Этот эксперимент проводился с входными файлами, фрагменты которых представлены на рисунках 11–13. На рис. 11 показано, как встречаются id формул среди текста. На рис. 12 видно, что в начале каждой строки пишется id, а само значение формулы находится в поле text тега Math. На рис. 13 приведён пример синтаксиса файла после обработки MyStem. Результат обработки файлов представлен на рис. 14, где изображён фрагмент заполненной таблицы базы данных. На рис. 15 показано отображение заполненной таблицы на страницу инструмента.

```
<?xml version="1.0" encoding="UTF-8"?><?latexml searchpaths="/opt/mocassin/patched-tex,/home/linglab/projects/thirdparty/a
d/><ERROR> <para xml:id="p2"> <p>20052 (513)<ERROR class="undefined">\nomer<ERROR><!-- %Remove in Eng. %\nomer{Vo.
есны свойства периферического спектра линейных операторов.Периферический спектр оператора MathIDA, некоторая степень котор
нтированный путь MathIDBI.Число звеньев пути называется длиной этого пути.Однозвенный путь из MathIDBJ в MathIDCA (петля)
*** -->общему делителю разностейMathIDDD.Известна и другая характеристика индекса импримитивности ([2], с. 131):MathIDDE е
OR class="undefined"><XMTok role="UNKNOWN" font="italic">\ov</XMTok></ERROR>.Тогда MathIDFA будет единственным с точностью
ает уравнению MathIDFJ<ERROR class="undefined"><XMTok role="UNKNOWN" font="italic">\ov</XMTok></ERROR>.Утверждение следует
твMathIDHF,то MathIDHG.</p> </para> <para xml:id="p12"> <p>Оценим сверху норму однорангового возмущения MathIDHH. Пус
го возмущения MathIDJE имеем оценку</p> <equation xml:id="S0.Ex3"> MathIDJF </equation> <p>При MathIDJG, пол
./../mocassin/patched-tex/ivm12.tex Line 200 **** --> <para xml:id="p15"> <p>Аналогом неприводимой матрицы с неотрицат
естно ([1], с. 337), что свойство s) эквивалентно следующему свойству:существует такой скаляр MathIDBDI, что для всякого M
thIDBFH иMathIDBFI будут единственными с точностью до постоянногомножителя собственными векторами, отвечающими соответстве
belref="LABEL:4"/>).</p> </para> <para xml:id="p20"> <p>При фиксированном MathIDBHJ имеем</p> <equation xml:id="S0
ara xml:id="p22"> <p>В общем случае для потенциально компактного положительного оператора MathIDBJH,неприводимого над з
заданных насвязном хаусдорфовом компакте MathIDCAI.</p> </para> </item> </itemize> </para> <para xml:id="|
```

Рисунок 11. Фрагмент статьи формата xml, использованной в эксперименте

Также был проведен эксперимент по извлечению кандидатов в термины на основе концептов. Был использован концепт «Элемент». В итоге были извлечены все кандидаты в термины, содержащие этот концепты. Фрагмент результирующей таблицы представлен на рис. 16, где также видно, какому классу NP-моделей принадлежат отобранные кандидаты.

A F2 <Math mode="inline" tex="A" xml:id="p6.m1" text="A"><XMath><XMTok role="UNKNOWN" font="italic">A</XMTok></XMath></Math>

B F1 <Math mode="inline" tex="\\lambda:\\lambda|=r(A)\\)" xml:id="p6.m2" text="conditional-set@(lambda, absolute-value@(lambda) = r \* A)">

C F2 <Math mode="inline" tex="r(A)" xml:id="p6.m3" text="r \* A"><XMath><XMAApp><XMTok meaning="times" role="MULOP">\*</XMTok><XMTok role="U

D F2 <Math mode="inline" tex="A" xml:id="p6.m4" text="A"><XMath><XMTok role="UNKNOWN" font="italic">A</XMTok></XMath></Math>

E F2 <Math mode="inline" tex="A" xml:id="p7.m1" text="A"><XMath><XMTok role="UNKNOWN" font="italic">A</XMTok></XMath></Math>

F F2 <Math mode="inline" tex="n" xml:id="p7.m2" text="n"><XMath><XMTok role="UNKNOWN" font="italic">n</XMTok></XMath></Math>

G F2 <Math mode="inline" tex="a\_{i j}" xml:id="p7.m3" text="a \_ (i \* j)"><XMath><XMAApp><XMTok role="SUBSCRIPTOP" scriptpos="post2"/><XMTok

H F2 <Math mode="inline" tex="n" xml:id="p7.m4" text="n"><XMath><XMTok role="UNKNOWN" font="italic">n</XMTok></XMath></Math>

I F2 <Math mode="inline" tex="p\_{1}, \\dots, p\_{n}" xml:id="p7.m5" text="list@(p \_ 1, \\dots, p \_ n)"><XMath><XMAApp><XMTok meaning="list" rol

J F2 <Math mode="inline" tex="a\_{i j}" xml:id="p7.m6" text="a \_ (i \* j)"><XMath><XMAApp><XMTok role="SUBSCRIPTOP" scriptpos="post2"/><XMTok

BA F2 <Math mode="inline" tex="A" xml:id="p7.m7" text="A"><XMath><XMTok role="UNKNOWN" font="italic">A</XMTok></XMath></Math>

BB F2 <Math mode="inline" tex="p\_{i}" xml:id="p7.m8" text="p \_ i"><XMath><XMAApp><XMTok role="SUBSCRIPTOP" scriptpos="post2"/><XMTok role="U

Рисунок 12. Фрагмент списка id, использованного в эксперименте

Спектральные {спектральный=A=(вин, мн, полн, неод|им, мн, полн) }

свойства {свойство=S, сред, неод=(вин, мн|род, ед|им, мн) }

матриц {матрица=S, жен, неод=род, мн}

с {с=FR=|с=S, сокр=(пр, мн|пр, ед|вин, мн|вин, ед|дат, мн|дат, ед|род, мн|род, ед|твор, мн|твор, ед|им, мн|им, ед) }

вещественными {вещественный=A=твор, мн, полн}

неотрицательными {неотрицательный=A=твор, мн, полн}

элементами {элемент=S, муж, неод=твор, мн}

установленные {устанавливать=V, пе=(прош, вин, мн, прич, полн, сов, страд, неод|прош, им, мн, прич, полн, сов, страд) |установленный=A=(вин, мн, полн, неод|им, мн, полн) }

Перроном {перрон=S, муж, неод=твор, ед}

и {и=CONJ=|и=INTJ=|и=S, сокр=(пр, мн|пр, ед|вин, мн|вин, ед|дат, мн|дат, ед|род, мн|род, ед|твор, мн|твор, ед|им, мн|им, ед) |и=PART=}

фробениусом {фробениус?=S, фам, муж, од=твор, ед|фробениус?=S, имя, муж, од=твор, ед}

в {в=FR=|в=S, сокр=(пр, мн|пр, ед|вин, мн|вин, ед|дат, мн|дат, ед|род, мн|род, ед|твор, мн|твор, ед|им, мн|им, ед) }

начале {начало=S, сред, неод=пр, ед}

прошлого {прошлое=S, сред, неод=род, ед|прошлый=A=(вин, ед, полн, муж, од|род, ед, полн, муж|род, ед, полн, сред) }

столетия {столетие=S, сред, неод=(вин, мн|род, ед|им, мн) }

были {быть=V, нп=прош, мн, изъяв, несов|быль=S, жен, неод=(пр, ед|вин, мн|дат, ед|род, ед|им, мн) }

обнаружены {обнаруживать=V, пе=прош, мн, прич, кр, сов, страд}

сначала {сначала=ADV=}

у {у=FR=|у=INTJ=|у=S, сокр=(пр, мн|пр, ед|вин, мн|вин, ед|дат, мн|дат, ед|род, мн|род, ед|твор, мн|твор, ед|им, мн|им, ед) }

интегральных {интегральный=A=(пр, мн, полн|вин, мн, полн, од|род, мн, полн) }

операторов {оператор=S, муж, од=(вин, мн|род, мн) |оператор=S, муж, неод=род, мн}

а {а=CONJ=|а=PART=|а=INTJ=|а=S, сокр=(пр, мн|пр, ед|вин, мн|вин, ед|дат, мн|дат, ед|род, мн|род, ед|твор, мн|твор, ед|им, мн|им, ед) }

затем {затем=ADV=}

и {и=CONJ=|и=INTJ=|и=S, сокр=(пр, мн|пр, ед|вин, мн|вин, ед|дат, мн|дат, ед|род, мн|род, ед|твор, мн|твор, ед|им, мн|им, ед) |и=PART=}

у {у=FR=|у=INTJ=|у=S, сокр=(пр, мн|пр, ед|вин, мн|вин, ед|дат, мн|дат, ед|род, мн|род, ед|твор, мн|твор, ед|им, мн|им, ед) }

более {более=ADV=|много=ADV=срав}

общих {общий=A=(пр, мн, полн|вин, мн, полн, од|род, мн, полн) }

операторов {оператор=S, муж, од=(вин, мн|род, мн) |оператор=S, муж, неод=род, мн}

Рисунок 13. Фрагмент файла после обработки морфологическим анализатором



| ASSMath   |
|---|
| спектральный радиус оператор "conditional-set@(lambda, absolute-value@(lambda) = r * A)"                    |
| квадратный матрица порядок "conditional-set@(lambda, absolute-value@(lambda) = r * A)"                      |
| направлять граф матрица "list@(p_1, ldots, p_n)"  |
| однозвенный? путь из "p_i"  |
| периферический спектр матрица "A"   |
| характеристический многочлен матрица "A"  |
| общий делитель разность "Delta * lambda = lambda ^ n + a_1 * lambda ^ (n_1) + cdots + a_t * lambda ^ (n_t)" |
| направлять граф матрица "A"   |
| положительный решение уравнение "A ^ * * v = r * A * v"   |
| собственный вектор матрица "A ^ *"  |
| направлять граф матрица "A"   |
| квазивнутренний? точка решетка "B"  |
| квазивнутренний? точка решетка "B"  |
| это неравенство при "r * A = 1"   |
| спектральный радиус оператор "A"  |
| квазивнутренний? точка решетка "B"  |
| единственный решение уравнение "x = A * x + f"  |

Рисунок 14. Столбец таблицы models, заполненной в эксперименте

| SSSMath  | ASSMath  | ASMath  | SSMath                      | SMath             | ASS                          | AS                    | SS             |
|--|--|---|-----------------------------|-------------------|------------------------------|-----------------------|----------------|
| произведение вектор в "C ^ n"  | спектральный радиус оператор "conditional-set@(lambda, absolute-value@(lambda) = r * A)" | ненулевой элемент "a_(i * j)"   | вектор матрица "A"          | плоскость "n"     | собственный значение этот    | периферический спектр | перрон и       |
| и из равенство "A ^ * * v = r * A * v"                                       | квадратный матрица порядок "conditional-set@(lambda, absolute-value@(lambda) = r * A)"   | ориентированный путь "list@(overline@(p_i * p_i_1), ldots, overline@(p_i_1 * p_j))" | индекс импримитивность? "m" | точка "a_(i * j)" | этот матрица и               | некоторый степень     | фробениус? в   |
| кронекер? из равенство "absolute-value@(D ^ (k-1) * x) = absolute-value@(x)" | направлять граф матрица "list@(p_1, ldots, p_n)"   | спектральный радиус "overline@(p_i * p_j)"  | идеал решетка "B"           | матрица "A"       | Другой характеристика индекс | спектральный радиус   | начало прошлое |

Рисунок 15. Отображение фрагмента заполненной таблицы models

|   |         |        |        |       |     |    |    |     |   |
|---|---------|--------|--------|-------|-----|----|----|-----|---|
| SSSMath                                   | ASSMath | ASMath | SSMath | SMath | ASS | AS | SS | SSS | S |
| элемент $\text{varphi}_k = D^{(k-1)} * u$ |         |        |        |       |     |    |    |     |   |
| элемент $v$                               |         |        |        |       |     |    |    |     |   |
| ненулевой элемент $a_{(i * j)}$           |         |        |        |       |     |    |    |     |   |

---

Рисунок 16. Извлечённые кандидаты для концепта «элемент»

По итогам экспериментов сделан вывод, что инструмент работает корректно: правильно строит NP-модели и извлекает кандидатов в термины.

### ЗАКЛЮЧЕНИЕ

В статье представлен разработанный программный инструмент для работы с онтологиями, который может быть использован для автоматизированного пополнения терминологического словаря предметной области. Проведенные эксперименты показали корректность полученных результатов. В дальнейшем планируется усовершенствовать данный инструмент с помощью анализа семантических отношений между классами.

### БЛАГОДАРНОСТИ

Работа выполнена за счет средств субсидии, выделенной Казанскому федеральному университету для выполнения государственного задания в сфере научной деятельности, проект 1.2368.2017/ПЧ.

### СПИСОК ЛИТЕРАТУРЫ

1. Портал искусственного интеллекта. 2017. URL: <http://www.aiportal.ru/articles/other/ontology.html>
2. Ле Хоай, Тузовский А. Ф. Использование онтологии в электронных библиотеках // Изв. Томского политехнического университета. – 2012. – Т. 320, № 5. – С. 36–42.
3. Обзор языка онтологии OWL. 2017 URL: <https://www.w3.org/TR/owl-features/> (Дата обращения: 20.12.2017)
4. Словарь RDF. Язык описания 1.1 URL схемы RDF. URL: <https://www.w3.org/TR/rdf-schema/> (Дата обращения: 20.12.2017)

5. Структура описания ресурсов (RDF). 2017. URL: <https://www.w3.org/RDF/> (Дата обращения: 20.12.2017)

6. *Musen M.* Domain Ontologies in Software Engineering: Use of Protege with the EON Architecture // *Methods of Information in Medicine*. – 1998. – P. 540–550.

7. *Chaudhri V., Farquhar A., Fikes R., Karp P., Rice J.* OKBC: A Programmatic Foundation for Knowledge Base Interoperability // *Fifteenth National Conf. on Artificial Intelligence. AAAIPres/ The MIT Press, Madison, 1998*. – P. 600–607.

8. *Noy N., Sintek M., Decker S., Crubezy M., Ferguson R., Musen M.* Creating Semantic Web Contents with Protege-2000 // *IEEE Intelligent Systems, March/April, 2001* – P. 60–71.

9. *Farquhar A., Fikes R., Rice J.* The Ontolingua Server: A Tool for Collaborative Ontology Construction // *Int. J. of Human-Computer Studies*. 1997. – Vol. 46, No 6. – P. 707–728.

10. *Sure Y., Erdmann M., Angele J., Staab S., Studer R., Wenke D.* OntoEdit: Collaborative ontology development for the Semantic Web // *In Proc. of the Int. Semantic Web Conference (ISWC 2002), Sardinia, Italia, June 2002*.

11. *Bechhofer S., Horrocks I., Goble C., Stevens R.* OilEd: A Reasonable Ontology Editor for the Semantic Web // *Joint German/Austrian Conf. on Artificial Intelligence (KI'01). Lecture Notes in Artificial Intelligence LNAI 2174, Springer-Verlag, Berlin, 2001*. – P. 396–408.

12. *MacGregor R.* Inside the LOOM classifier // *SIGART Bulletin*. – 1991. – Vol. 3, No 2. – P. 70–76.

13. *Fernandez M, Gomez-Perez A., Pazos J.* A Building a Chemical Ontology Using Methodology and the Ontology Design Environment // *IEEE Intelligent Systems, Jan./Feb., 1999*. – P. 37–46.

## SOME SOFTWARE INSTRUMENTS FOR THE AUTOMATED REPLENISHMENT OF THE TERMINOLOGICAL DICTIONARY OF THE SUBJECT OBLAST

<sup>1</sup>R. A. Rumyantsev, <sup>1,2</sup>O. A. Nevzorova

<sup>1</sup>*Institute of Computational Mathematics and Information Technologies, Kazan (Volga Region) Federal University, Kazan, 35 Kremlyovskaya Street, 420008*

<sup>2</sup>*Institute of Applied Semiotics, Academy of Sciences of the Republic of Tatarstan, Kazan, Levo-Bulachnaya Street, 36A, 420111*

### **Abstract**

The article describes the application OntoDictionary, which is designed to work with scientific mathematical articles and ontologies created in Protege. The application is able to create an ontology dictionary, split its elements into concepts, and process them in Boolean search. There is a functional for the selection of certain nominal groups from mathematical articles. The novelty lies in the creation and method of processing nominal groups containing formulas. Formulas are processed regardless of their type. The selection of candidates for terms has been constructed. Throughout the functional, a number of experiments have been performed with the ontology of mathematical knowledge of OntoMathPRO, which was also developed at the Kazan Federal University.

**Keywords:** *mathematical knowledge, ontology, concept, search index, Noun Phrase, candidates in terms.*

### **REFERENCES**

1. Portal of artificial intelligence. 2017 URL: <http://www.aiportal.ru/articles/other/ontology.html>
2. *Le Howay, Tuzovsky A.F.* Using ontology in electronic libraries // Proc. of the Tomsk Polytechnic University. – 2012. – № 5. – S. 36–42.
3. OWL Web Ontology Language Overview. 2017 URL: <https://www.w3.org/TR/owl-features/>
4. RDF Vocabulary Description Language 1.1 RDF Schema URL: <https://www.w3.org/TR/rdf-schema/>
5. Resource Description Framework (RDF). 2017. URL: <https://www.w3.org/RDF6>.

6. *Musen M.* Domain Ontologies in Software Engineering: Use of Protege with the EON Architecture // *Methods of Information in Medicine*. – 1998. – P. 540–550.

7. *Chaudhri V., Farquhar A., Fikes R., Karp P., Rice J.* OKBC: A Programmatic Foundation for Knowledge Base Interoperability // *Fifteenth National Conf. on Artificial Intelligence*. AAAIPres/ The MIT Press, Madison, 1998. – P. 600–607.

8. *Noy N., Sintek M., Decker S., Crubezy M., Ferguson R., Musen M.* Creating Semantic Web Contents with Protege-2000 // *IEEE Intelligent Systems*, March/April, 2001 – P. 60–71.

9. *Farquhar A., Fikes R., Rice J.* The Ontolingua Server: A Tool for Collaborative Ontology Construction // *Int. J. of Human-Computer Studies*. 1997. – Vol. 46, No 6. – P. 707–728.

10. *Sure Y., Erdmann M., Angele J., Staab S., Studer R., Wenke D.* OntoEdit: Collaborative ontology development for the Semantic Web // *In Proc. of the Int. Semantic Web Conference (ISWC 2002)*, Sardinia, Italia, June 2002.

11. *Bechhofer S., Horrocks I., Goble C., Stevens R.* OilEd: A Reasonable Ontology Editor for the Semantic Web // *Joint German/Austrian Conf. on Artificial Intelligence (KI'01)*. *Lecture Notes in Artificial Intelligence LNAI 2174*, Springer-Verlag, Berlin, 2001. – P. 396–408.

12. *MacGregor R.* Inside the LOOM classifier // *SIGART Bulletin*. – 1991. – Vol. 3, No 2. – P. 70–76.

13. *Fernandez M, Gomez-Perez A., Pazos J.* A Building a Chemical Ontology Using Methodology and the Ontology Design Environment // *IEEE Intelligent Systems*, Jan./Feb., 1999. – P. 37–46.

## СВЕДЕНИЯ ОБ АВТОРАХ



**РУМЯНЦЕВ Роман Анатольевич** – Магистрант Института вычислительной математики и информационных технологий Казанского (Приволжского) федерального университета, Республика Татарстан, г. Казань

**Roman A. RUMYANTSEV** – graduate student in the Institute of Computational Mathematics and Information Technologies of the Kazan (Volga Region) Federal University, Republic of Tatarstan, Kazan

email: romanrar93@gmail.com



**НЕВЗОРОВА Ольга Авенировна** – кандидат технических наук, доцент кафедры информационных систем Института вычислительной математики и информационных технологий Казанского (Приволжского) федерального университета, Республика Татарстан, г. Казань

**Olga A. NEVZOROVA** – Candidate of Technical Sciences, Associate Professor of the Information Systems Department of the Institute of Computational Mathematics and Information Technologies of the Kazan (Volga Region) Federal University, Republic of Tatarstan, Kazan

email: onevzoro@gmail.com

*Материал поступил в редакцию 15 мая 2018 года*