

education computer networks. Past, present and development trends), *Obrazovatel'nye resursy i tekhnologii*, 2017, vol. 2 (19), pp. 17–25 (in Russian).

7. **Official** website of the consortium GÉANT, available at: <http://www.geant.net>.

8. GÉANT Association Compendium of national research and education networking organisations in Europe — 2015 Edition, available at: <https://compendium.geant.org>.

9. **Abramov A. G., Evseev A. V.** Set' RUNNet: navstrechu sovremennym vyzovam sfery telekommunikatsiy v nauke i obrazovanii (Network RUNNet: towards the state-of-the-art challenges in the field of telecommunications in science and education), *Informatizatsiya Obrazovaniya i Nauki*, 2017, vol. 1 (33), pp. 100–115 (in Russian).

10. **Abramov A. G., Evseev A. V.** Analiz potrebnostej organizatsiy vysshego obrazovaniya Rossii v telekommunikatsionnykh servisakh nauchno-obrazovatel'noy napravlenosti (Analysis of demands of russian organizations of higher education in telecommunication services of scientific and educational direction), *Informatizatsiya Obrazovaniya i Nauki*, 2018, vol. 2 (38), pp. 9–18 (in Russian).

11. **Official** website of the project WLCG, available at: <http://wlcg.web.cern.ch>.

12. **Martelli E., Stancu S.** LHCOPN and LHCONE: status and future evolution, *Journal of Physics: Conference Series*, 2015, vol. 664, available at: <http://iopscience.iop.org/article/10.1088/1742-6596/664/5/052025>.

13. GÉANT Annual Report 2016, available at: <https://ar2016.geant.org>.

14. **Abramov A. G., Vasilyev I. V., Porhachev V. A.** Razvitie infrastruktury autentifikatsii i avtorizatsii dlya udostoverayushchey federatsii v ramkakh proektov eduGAIN i eduroam na baze seti RUNNet (Development of the authentication and authorization infrastructure for the identity federation within the eduGAIN and eduroam projects based on the RUNNet network), *ITNOU: Informatsionnye Tekhnologii v Nauke, Obrazovanii i Upravlenii*, 2017, vol. 4, pp. 56–64 (in Russian).

15. **Abramov A. G.** Oblachnye tekhnologii na osnove otkrytykh programmnykh reshenij v universitetskoj seti RUNNet: opyt razvertyvaniya i jekspluatatsii (Cloud technologies based on open source solutions in the university network RUNNet: experience of deployment and operation), *Informatizatsiya Obrazovaniya i Nauki*, 2016, vol. 1 (29), pp. 106–117 (in Russian).

УДК 004.422

DOI: 10.17587/it.24.553-560

**А. А. Романов**, преподаватель, e-mail: [gloomsfb@gmail.com](mailto:gloomsfb@gmail.com),

**Д. Г. Волчек**, преподаватель, e-mail: [dvolchek@yandex.ru](mailto:dvolchek@yandex.ru),

**Д. И. Муромцев**, канд. техн. наук, доц., зав. каф., e-mail: [mouromtsev@corp.ifmo.ru](mailto:mouromtsev@corp.ifmo.ru),  
Университет ИТМО, Санкт-Петербург

## Семантические технологии в массовых открытых онлайн-курсах: экстракция, обогащение и связывание данных электронных учебных курсов для образовательной платформы Open edX

*Рассматривается проблема, связанная с увеличением количества учебных материалов, которые доступны в открытых онлайн-курсах. Такой рост приводит к снижению качества отдельных ресурсов, усложнению классификации и поиска, что в целом негативно отражается на конечных пользователях. Предлагается решение для платформы Open edX, основанное на принципах семантических технологий и объединяющее в себе экстракцию, обогащение и взаимосвязь учебных данных. Для исследования эффективности метода и внедрения его в учебный процесс были разработаны и протестированы специальные испытательные компоненты.*

**Ключевые слова:** семантическая сеть, образование, связанные данные в образовании, образовательная онтология, электронное образование, семантические технологии в образовании

### Введение

Идея дистанционного образования рассматривается уже достаточно давно [1, 2], и по мере развития информационных технологий эта идея воплощается в жизнь все более успешно. Спектр различных форм, используемых для реализации концепции дистанционного обучения, широк и включает в себя как достаточно примитивные способы (напри-

мер, бумажную переписку, радиотрансляции, телепрограммы), так и более успешные виды (такие как онлайн-обучение), которые обеспечивают быструю и удобную коммуникацию между субъектами образовательного процесса. Онлайн-обучение стало возможным благодаря развитию сети Интернет, росту ее популярности и, соответственно, увеличению числа пользователей. Начиная с середины 90-х годов XX века многие университеты мира продви-

гали свои образовательные программы в сети Интернет. В основной своей массе это были частные инициативы самих учебных заведений, которые университеты реализовывали собственными силами. При этом наблюдалось многообразие организации такого вида обучения, что привело к ряду позитивных и негативных последствий. К позитивным факторам можно отнести тот факт, что благодаря независимости университетов на разработчиков не накладывались какие бы то ни было ограничения. При таком подходе они могли изобретать и исследовать различные методы, средства, формы и стратегии. Однако обучающемуся было неудобно, а порой и невозможно, участвовать в образовательном процессе из-за различий вузовских систем онлайн-образования по своей структуре и форме, а также из-за направленности онлайн-курсов вуза на обучение студентов только своего вуза.

Следующим этапом эволюции дистанционных образовательных технологий стали массовые открытые онлайн-курсы (МООК). Эта концепция была представлена в 2008 г. и через четыре года превратилась в по-настоящему влиятельный тренд [3]. В данный момент существуют несколько крупных платформ, которые на своих ресурсах размещают курсы различных учебных заведений. Одна из них — проект edX.org, в основе которого лежит свободно распространяемая система управления обучением (Learning Management System, LMS) Open edX [4]. В проекте edX.org участвуют около 100 ведущих мировых университетов, которые предоставляют свыше 1300 курсов. Обучение проходят более 10 миллионов студентов со всего мира. Сама платформа Open edX используется многими другими образовательными проектами, например, проектом "Открытое образование" (<http://openedu.ru>) или площадкой курсов Университета ИТМО (<http://courses.ifmo.ru>). С одной стороны, совокупность всех образовательных курсов представляет собой большой объем данных, которые уже невозможно обрабатывать традиционными способами, причем семантические технологии решают эту проблему, определяя общие правила для осуществления обмена данными и их многократного использования в различных приложениях, корпорациях и сообществах [5]. С другой стороны, это эффективный способ представления данных в сети Интернет, причем такой формат понятен как человеку, так и компьютеру. Семантические технологии представляют значения с помощью онтологии и обеспечивают

аргументацию, используя связи, правила, логику и условия, оговоренные в онтологии [6]. Семантические технологии помогают извлекать полезную информацию из данных, содержания документов или кодов приложений, опираясь на открытые стандарты, что позволяет говорить о том, что компьютер понимает смысл документа. Процесс построения онтологий получил название "онтологический инжиниринг" [7], и его можно считать развитием методологии инженерии знаний с помощью искусственного интеллекта.

В данной статье описан онтологический инжиниринг и основные этапы разработки по внедрению семантических технологий в МООК и их необходимость. Стоит отметить факт научной новизны, а именно — отсутствие схожих по реализации проектов в сфере образования, хотя семантические технологии широко применяются в сфере бизнеса и банковского сектора.

### Постановка проблемы и ее решение

Рост числа образовательных материалов [8], безусловно, является позитивным аспектом индустрии образования. Используя существующие ресурсы, можно удобно и быстро получить высококачественные, а иногда и уникальные знания во многих предметных областях. Однако такой объем информации приносит и определенные проблемы, такие как:

- необходимость структурирования информации;
- сложность поиска необходимых знаний;
- возникновение дублирования и наложения учебных материалов;
- необходимость построения междисциплинарных связей;
- сложность построения образовательных траекторий;
- соответствие информации учебным планам и связь с компетенциями.

Таким образом, онтологический инжиниринг и разработка ряда модулей для образовательной платформы Open edX нацелена решить указанные проблемы.

МООК состоят из нескольких базовых частей, а именно, из теоретических материалов, практических заданий, оценки знаний и финального экзамена. В различных курсах возможны вариации (взаимное оценивание, проектная работа, творческое задание и т. п.), но их число невелико. Если автоматизировать

извлечение из теоретических материалов ключевых слов или так называемых концептов предметной области, то можно рассматривать их как "единицы знаний" [9, 10]. По данному набору концептов можно определить, к какой предметной области относится теоретический материал, что он описывает, какие цели преследует. Если обогатить концептами достаточный объем курсов, станет возможным построение иерархии предметных областей и, как следствие, междисциплинарных связей. Таким образом достигается структурирование информации. Если обучающийся встречается в лекции незнакомые термины и понятия, то, используя список концептов, он может посмотреть, в каких курсах и каких лекциях они также используются и объясняются. Все это позволяет повторно использовать теоретические материалы, облегчает поиск необходимых знаний, обеспечивает построение междисциплинарных связей.

Кроме теоретических материалов концептами можно обогащать и контент, так как он служит источником знаний, которые должен усвоить студент. Анализируя, какие предметные области изучены хорошо, а с какими обучающемуся еще нужно поработать, можно составить рекомендации для конкретного студента. Таким образом достигается построение индивидуальных образовательных траекторий. Оценивая степень того, насколько концепты теоретических материалов покрывают концепты тестов, можно оценить их соответствие друг другу, а значит, в случае несоответствия, сообщить создателям курса о необходимости корректировки. Таким образом, извлечение концептов поможет решить следующий ряд проблем.

1. *Построение областей знаний.* Путем анализа состава и порядка появления концептов, используемых в курсе, возможно семантическое построение предметных областей [11].

2. *Навигация по курсам.* Анализируя связи между концептами и пересечения различных предметных областей, можно осуществлять нелинейную навигацию, в том числе с учетом других курсов, доступных на платформе Open edX.

3. *Связывание курсов.* Если обучающемуся встретился незнакомый концепт, можно узнать, к какой предметной области относится концепт, где еще он используется или описывается, какие другие концепты с ним связаны, какие концепты ответственные за его появление или, наоборот, к появлению каких концептов он приводит.

4. *Индивидуализация.* Концепты, представленные в семантическом виде и формирующие предметные области, позволяют учащемуся

иметь в своем распоряжении не один конкретный курс, а все курсы, связанные между собой. Обучающийся может удобно и просто "путешествовать" по различным разделам многочисленных курсов, получая информацию в той последовательности и том объеме, которые удобны конкретно ему.

5. *Рекомендации.* Концепты можно получать не только из образовательного контента, но и из тестов. Таким образом, можно предоставлять рекомендации учащимся о том, что предметная область изучена недостаточно хорошо и требуется повторить материал. Рекомендации могут быть даны и создателям курсов, если, например, тест и образовательные материалы недостаточно коррелируют между собой с точки зрения использования концептов.

## Модель онтологии MOOK

Представление данных MOOK в формальном, семантическом виде в целях их адекватной трактовки человеком и надлежащей машинной обработки — это достаточно сложная и распространенная задача [12]. Данная проблема может быть решена, в том числе благодаря онтологическому моделированию и семантическим технологиям [13]. Разработанная авторами онтология онлайн-курсов состоит из 25 классов, 13 свойств-объектов и 19 свойств-значений. Разработанная онтология (рис. 1) основана на онтологиях высшего уровня:

- AIIISO (<http://purl.org/vocab/aiiso/schema>), которая предоставляет классы и свойства-объекты для описания внешней организационной структуры учреждения, занимающегося образовательной деятельностью;
- BIBO (<http://purl.org/ontology/bibo/>), которая предоставляет основные концепты и свойства для описания цитат и библиографических ссылок;
- FOAF (<http://xmlns.com/foaf/spec/>), которая описывает участников какого-либо процесса, виды их деятельности и виды отношений между собой или другими объектами;
- TEACH (<http://linkedscience.org/teach/ns/teach.rdf>), которая является словарем, содержащим термины, относящиеся к образовательному процессу.

Онтология состоит из классов, описывающих структуру онлайн-курса, и классов, которые содержат описание концептов и предметных областей. Определения этих классов приведены ниже.

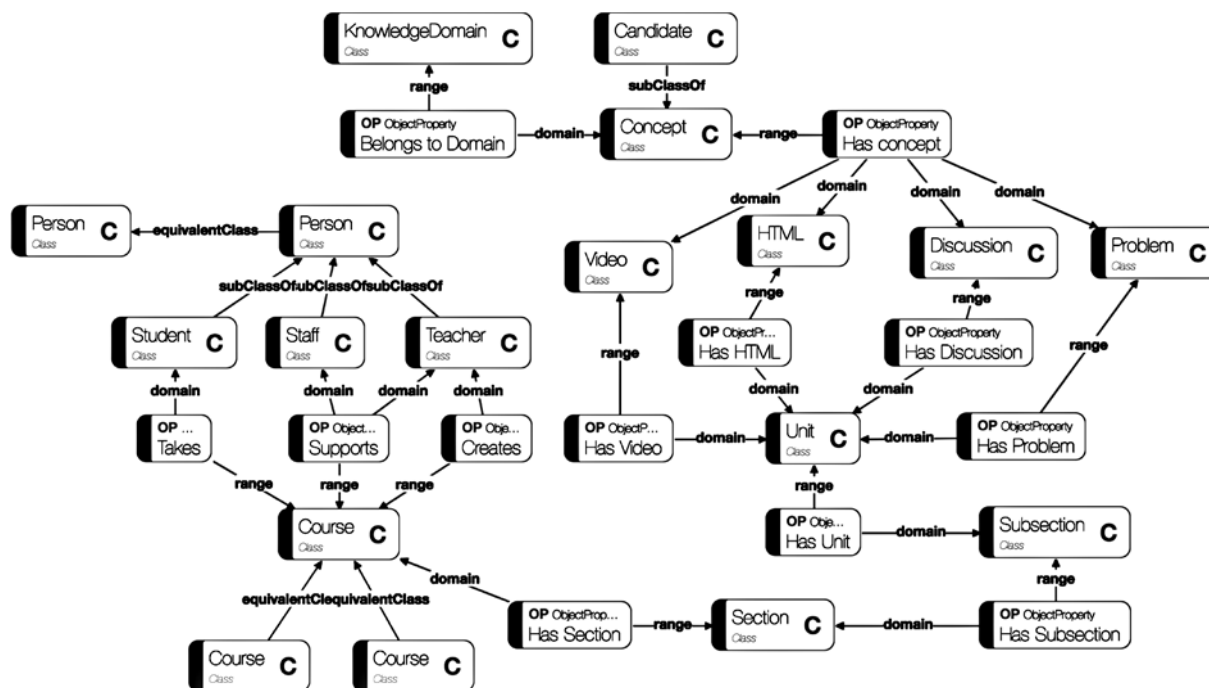


Рис. 1. Схема онтологии курсов платформы Open edX

- **Course (Курс)** — главный класс рассматриваемой онтологии. Он соответствует онтологиям AIIISO:Course и TEACH:Course. Атрибуты указанного класса включают в себя: даты начала и окончания курса; число часов в неделю, необходимое для успешного прохождения курса; название, описание и изображение для обозначения курса.
- **Section (Секция)** — класс, описывающий верхний уровень структуры курса. Он содержит следующие атрибуты: заголовок, дату начала и видимость пользователям (секции — это учебные разделы, отображающиеся поэтапно, обычно еженедельно).
- **Subsection (Подсекция)** — класс, описывающий элементы, входящие в секции. Данный класс включает следующие атрибуты: заголовок, дату начала, видимость и сроки выполнения (например, дата выполнения тестового задания, однако сроки могут и отсутствовать).
- **Unit (Юнит)** — класс, описывающий элементы, входящие в подсекции. В него входят следующие атрибуты: заголовок и видимость.
- **HTML** — класс, содержащий основной образовательный материал (текстовые лекции) в формате HTML.
- **Video (Видео)** — класс, необходимый для описания учебных материалов в формате видео и содержащий следующие атрибуты: ссылку на видео, хостинг (ресурс), субтитры.
- **Problem (Тест)** — класс для различных форм тестовых заданий. Атрибуты: число баллов, число попыток и тип заданий.
- **Person (Пользователь)** — класс для описания информации обо всех пользователях, вовлеченных в образовательный процесс MOOK платформы. Он соответствует онтологии FOAF:Person и содержит следующие подклассы:
  - **Teacher (Преподаватель)** — класс, описывающий создателей и разработчиков курса;
  - **Staff (Персонал)** — класс, описывающий участников, вовлеченных в поддержку курса (отвечающих на вопросы и предложения и решающих проблемы пользователей);
  - **Student (Студент)** — класс, описывающий участников, изучающих курс.
- **Concept (Концепт)** — класс, описывающий термины, входящие в учебные материалы:
  - **Concept Candidate** (Кандидат в концепты) — класс для хранения информации об автоматически полученных в результате обработки естественного языка концептах. Класс включает следующие атрибуты: вес концепта (его значимость), число вхождений в различные разделы и рейтинг;
  - **Concept** (Концепт) — класс, содержащий описание концептов и их связи с внешними источниками данных (такие как DBpedia, Wikidata).

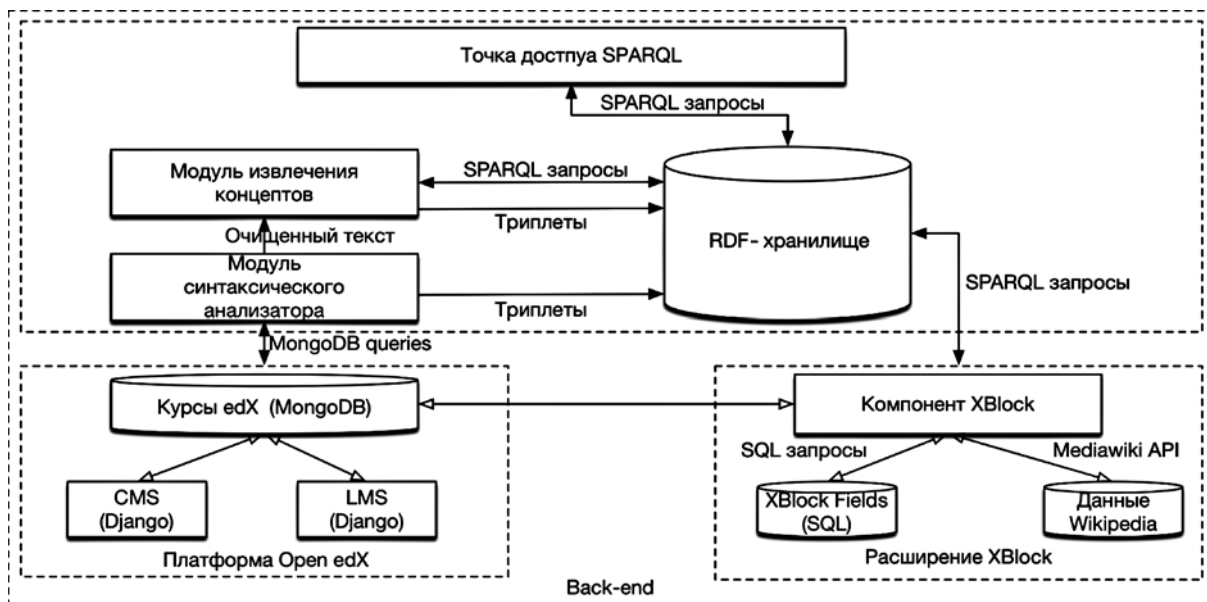


Рис. 2. Общая архитектура разработки

## Программная реализация

Как показано на общей схеме архитектуры (рис. 2), компонент XBlock взаимодействует с хранилищем платформы Open edX посредством MongoDB-запросов. Таким образом, с помощью разработанного расширения происходит обработка данных, связанных со структурой курсов, лекциями, учебными материалами в формате HTML, концептами, полученными за счет алгоритма экстракции ключевых слов, и концептами, связанными с лекциями вручную. Расширение XBlock, разработанное с использованием языка Python, преобразует данные в триплеты (формат N-Triples <http://www.w3.org/TR/n-triples/>) и экспортирует их в семантическое ядро системы на основе RDF4J (<http://rdf4j.org/>).

**Компонент XBlock.** XBlock является стандартной точкой расширения платформы Open edX со своим API. Для хранения данных, по-

лученных от пользователя, используются формы, которые по умолчанию сохраняют информацию в предустановленной базе данных SQL. Указанная БД была расширена в связи с необходимостью динамического поиска данных по уже существующим концептам, "привязанным" к другим лекциям. Кроме того, динамический поиск связан с терминами из Википедии посредством MediaWiki API. Компонент XBlock (рис. 3) решает задачу связывания лекций с концептами, чтобы затем установить соответствие данных с онтологией.

**Модуль синтаксического анализатора.** На рис. 4 представлена структура модуля синтаксического анализатора базы данных MongoDB Open edX. После обработки анализатором данные преобразуются в зависимости от одного из нижеуказанных типов.

- *Структура.* В отсутствие документации о структуре хранилища и с учетом того факта, что используется документно-ориентированная СУБД MongoDB, основной задачей было выявить все связи между учебными материалами курсов, распределенные по трем следующим коллекциям.

1. В коллекции `Active_versions` хранится краткая информация и текущая опубликованная версия (`published-branch`) курса, причем версия имеет идентификатор `ObjectId`, используемый для связи со второй коллекцией.

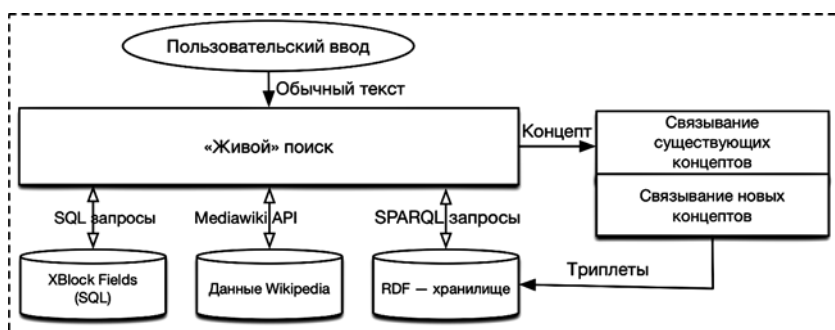


Рис. 3. Структура компонента XBlock

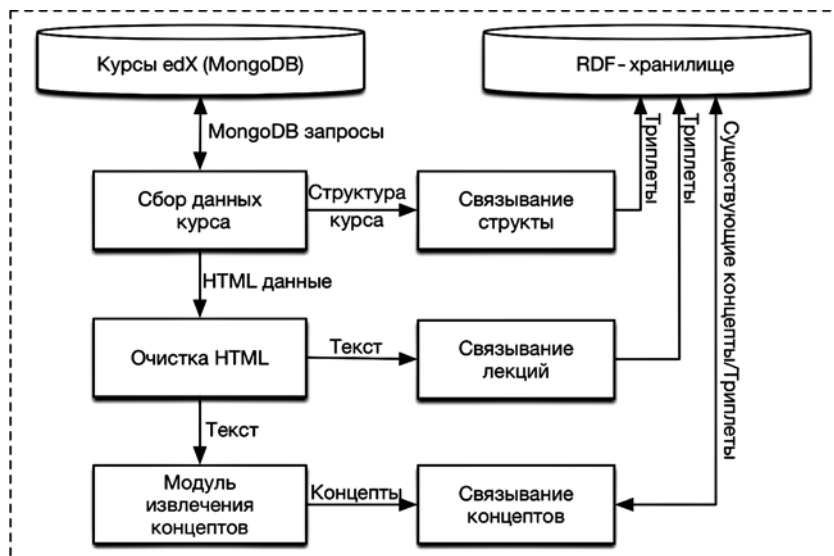


Рис. 4. Структура синтаксического анализатора

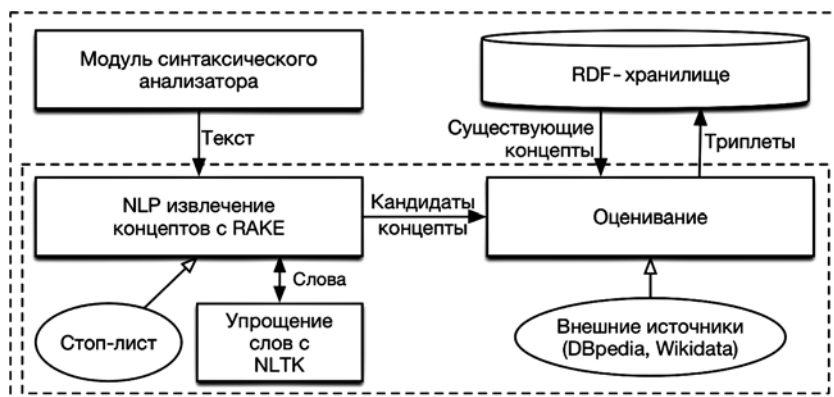


Рис. 5. Структура модуля извлечения концептов

2. В коллекции Structures находится вся история изменений и структура всех курсов, при этом один набор данных содержит сразу все данные о курсе. Следовательно, определить, к чему относится информация (лекция, видео или тест), возможно лишь по параметру типа данных (*block\_type*). Кроме того, необходимо учитывать иерархию курса в системе Open edX, т. е. вложенность всех учебных материалов (называемых юнитами) и полный путь от родителей до потомков. Идентификатор потомков позволяет выявить связь с третьей коллекцией.

3. В Definitions хранятся все данные курса, такие как текст лекций, учебные тесты, обсуждения, видео и их субтитры и т. д.

- *HTML-данные.* Все данные лекций хранятся в формате HTML. Для создания читаемого текста модуль определяет все HTML-теги, прочие нечитаемые последовательности и запрещенные символы, чтобы установить

корректные связи с онтологией и затем выполнить импорт в RDF-хранилище.

- *Концепты.* Концепты разделяются на два типа. К первому типу относятся концепты, добавленные "вручную" и связанные с лекциями с помощью XBlock компонента, а ко второму — концепты, автоматически выявленные из лекционного материала посредством модуля извлечения концептов (рис. 5).

**Модуль извлечения концептов.** Связывание лекций и концептов "вручную" создателем курса — длительный процесс, в особенности для уже полностью готового курса. В связи с этим был применен метод автоматического извлечения концептов на основе нейролингвистического программирования. В основу соответствующего модуля были положены готовые способы извлечения ключевых слов (RAKE — Rapid Automatic Keyword Extraction, инструментарий Быстрого Автоматического Извлечения Ключевых Слов) и NLTK (Natural Language Toolkit, Инструментарий Естественного Языка), а также дополнительные преобразования, связанные с более точным выделением ключевых элементов текста и его обработкой, в частности, с определением падежей, рода, числа и т. д. В результате были получены кандидаты в концепты, оцененные по открытым источникам данных (DBpedia, Wikidata), по уже связанным ранее концептам, и по добавленным "вручную" авторами курсов.

### Результаты внедрения компонента

Разработанные модули и компонент XBlock были успешно интегрированы в платформу Open edX и протестированы на уже размещенных онлайн-курсах. В результате работы ошибок не выявлено, а экстракция всех данных и выявление концептов прошли успешно, что позволило получить следующие количественные метрики: курсов — 6, секций — 64, подсекций — 315, юнитов — 941, страниц лекций — 1005; время работы (выполнение MongoDB-запросов и создание локального RDF-хранилища) составило в среднем 0,72 с после 1000 запусков; пол-

### Статистические данные по структуре онлайн-курсов

Курс	Секции	Подсекции	Юниты	Концепты	Уникальные концепты
Applications of Linear Algebra Part 1	8	49	50	1072	559
Applications of Linear Algebra Part 2	9	58	86	1610	811
Introduction to Differential Equations	18	45	106	1890	639
Introduction to Differential Equations	5	28	181	3346	1078
LAFF On Programming for Correctness	8	39	181	1871	654
Pre-University Calculus	9	71	238	1950	770

ное время работы (MongoDB-запросы, SPARQL-запросы к DBpedia, выявление концептов и их анализ) заняло около 2 ч; общее число триплетов — 32 590. Подробная информация по полученным данным представлена в таблице.

### Заключение

Задачи по разработке и документированию онтологии выполнены в полной мере, экстракция данных из онлайн-курсов платформы Open edX осуществлена. Автоматическое выявление концептов также реализовано, при этом в дальнейшем следует улучшить качество выявления терминов с точки зрения лексики языка. Однако уже полученный набор данных, представленный в семантическом виде, позволяет использовать их для формирования:

- междисциплинарных связей;
- рекомендаций при прохождении курсов;
- нелинейной траектории изучения материала по одному курсу с учетом всех курсов, представленных на учебной платформе;
- рекомендаций для создателей курса на основе анализа корреляции между концептами учебных материалов и тестов.

Прделанная работа является основой для создания интерфейсов визуализации динамических переходов между учебными материалами, рекомендательных сервисов для студентов и разработчиков курсов. Все материалы представлены в открытом репозитории лаборатории "Интеллектуальные методы обработки информации и семантические технологии" на сайте <https://github.com/ailabitmo/edX-ontology>.

### Список литературы

1. Андреев А. А. Введение в дистанционное обучение. 1997. С. 9—13.

2. Полат Е. С. Дистанционное обучение: организационные и педагогические аспекты // Информатика и образование. 1996. № 3. С. 87—91.

3. Бадарч Д., Токарева Н., Цветкова М. МООК: реконструкция высшего образования // Высшее образование в России. 2014. № 10. С. 135—145.

4. Wilson L., Gruz A. MOOCs — international information and education phenomenon? // Bulletin of the Association for Information Science and Technology. 2014. Т. 40. № 5. С. 35—40.

5. Харабет Я. К. Автоматическое выделение количественных конструкций в русскоязычных научно-популярных текстах // XVIII Объединенная научная конференция "Интернет и современное общество" (IMS—2015). Санкт-Петербург. 2015. С. 23—25.

6. Гаврилова Т. А. Онтологический подход к управлению знаниями при разработке корпоративных информационных систем // Новости искусственного интеллекта. 2003. № 2. С. 24—30.

7. Гаврилова Т. А. Онтологический инжиниринг // Докл. Восьмой науч.-практ. конф. "Реинжиниринг бизнес-процессов на основе современных информационных технологий. Системы управления знаниями" (РБП-СУЗ-2005). 2005. С. 79—82.

8. Хорошевский В. Ф. Семантические технологии: ожидания и тренды // Открытые семантические технологии проектирования интеллектуальных систем: Тр. II Междунар. науч.-техн. конф. Беларусь, Минск: БГУИР. 2012. С. 143—158.

9. Ховлова М. В. Исследование лексико-синтаксической сочетаемости в русском языке с помощью статистических методов (на базе корпусов текстов) // АКД. СПб. 2010. С. 26—32.

10. Большакова Е. И., Васильева Н. Э. Формализация лексико-синтаксической информации для распознавания регулярных конструкций естественного языка // Программные продукты и системы. 2008. № 4. С. 103—106.

11. Палагин А. В., Петренко Н. Г. К вопросу системно-онтологической интеграции знаний предметной области // Математические машины и системы. 2007. Т. 1. № 3—4.

12. Vas R. Educational ontology and knowledge testing // The Electronic Journal of Knowledge Management. 2007. Т. 5. № 1. С. 123—130.

13. Berners-Lee T., Hendler J., Lassila O. The semantic web. Scientific American. 2001. № 5. P. 28—37.

A. A. Romanov, Teacher, e-mail: gloomspb@gmail.com,  
D. G. Volchek, Teacher, e-mail: dvolchek@yandex.ru,  
D. I. Mouromtsev, Professor, e-mail: mouromtsev@corp.ifmo.ru,  
ITMO University, Laboratory of Information Science and Semantic Technologies,  
St. Petersburg, Russian Federation

## Towards Semantic MOOC: Extracting, Enriching and Interlinking E-learning Data in Open edX Platform

*The number of educational materials available through open online courses gradually increases. Their growing number negatively affects quality of certain materials and makes it harder for end users to classify information and search for data. The offered solution is based on the principles of semantic technologies for Open edX platform. It includes extraction, enrichment and interlinking of educational data. For the purposes of method effectiveness study and method implementation, special components were developed and tested.*

**Keywords:** semantic web, education, linked data in education, educational ontology, e-learning, educational semantic technologies

DOI: 10.17587/it.24.553-560

### References

1. **Andreev A. A.** *Vvedenie v distancionnoe obuchenie* (Introduction to distance learning), 1997 (in Russian).
2. **Polat E. S.** Distancionnoe obuchenie: organizacionnye i pedagogicheskie aspekty (Distance learning: organizational and pedagogical aspects), *Informatika i Obrazovanie*, 1996, no. 3, pp. 87–91 (in Russian).
3. **Badarch D., Tokareva N., Cvetkova M.** MOOC: rekonstrukcija vysshego obrazovanija (MOOC: reconstruction of higher education), *Vysshee Obrazovanie v Rossii*, 2014, no. 10 (in Russian).
4. **Wilson L., Gruzd A.** MOOCs — international information and education phenomenon?, *Bulletin of the Association for Information Science and Technology*, 2014, vol. 40, no. 5, pp. 35–40.
5. **Harabet Ja. K.** Avtomaticheskoe vydelenie kolichestvennykh konstrukcij v russkojazychnykh nauchno-populjarnykh tekstah (Automatic selection of quantitative structures in the Russian popular science texts), *XVIII Ob#edinennaja nauchnaja konferencija "Internet i sovremennoe obshhestvo" (IMS—2015)*, Sankt-Peterburg, 2015, pp. 23–25 (in Russian).
6. **Gavrilova T. A.** Ontologicheskij podhod k upravleniju znanijami pri razrabotke korporativnykh informacionnykh sistem (Ontological approach to knowledge management in the development of corporate information systems), *Novosti Iskusstvennogo Intellekta*, 2003, no. 2 (in Russian).
7. **Gavrilova T. A.** Ontologicheskij inzhiniring (Ontology engineering), *Dokl. Vos'moj nauchno-prakticheskoy konferencii "Re-*
8. **Horoshevskij V. F.** Semanticheskie tehnologii: ozhidaniya i trendy (Semantic technologies: expectations and trends), *Otkrytye semanticheskie tehnologii proektirovanija intellektual'nykh sistem: Tr. II Mezhdunar. nauch.-tehn. konf. Belarus'*, Minsk, BGUIR, 2012, pp. 143–158 (in Russian).
9. **Hohlova M. V.** *Issledovanie leksiko-sintaksicheskoy sochetamosti v russkom jazyke s pomoshh'ju statisticheskikh metodov (na baze korpusov tekstov)* (Lexico-syntactic compatibility in the Russian language with the help of statistical methods (based on text corpora)), AKD, SPb., 2010 (in Russian).
10. **Bol'shakova E. I., Vasil'eva N. Je.** Formalizacija leksiko-sintaksicheskoy informacii dlja raspoznavanija reguljarnykh konstrukcij estestvennogo jazyka (The formalization of lexico-syntactic information to recognize regular structures in natural language), *Programmnye Produkty i Sistemy*, 2008, no. 4, pp. 103–106 (in Russian).
11. **Palagin A. V., Petrenko N. G.** K voprosu sistemno-ontologicheskoy integracii znanij predmetnoj oblasti (To the issue of system-ontological integration of knowledge of subject area), *Matematicheskie Mashiny i Sistemy*, 2007, vol. 1, no. 3–4 (in Russian).
12. **Vas R.** Educational ontology and knowledge testing, *The Electronic Journal of Knowledge Management of*, 2007, vol. 5, no. 1, pp. 123–130.
13. **Berners-Lee T., Hendler J., Lassila O.** The semantic web, *Scientific American*, 2001, vol. 284, no. 5, pp. 28–37.

---

---

### Адрес редакции:

107076, Москва, Стромьинский пер., 4

Телефон редакции журнала (499) 269-5510

E-mail: it@novtex.ru

Технический редактор *Е. В. Конова*.

Корректор *Е. В. Комиссарова*.

Сдано в набор 06.06.2018. Подписано в печать 23.07.2018. Формат 60×88 1/8. Бумага офсетная.

Усл. печ. л. 8,86. Заказ ИТ818. Цена договорная.

Журнал зарегистрирован в Министерстве Российской Федерации по делам печати, телерадиовещания и средств массовых коммуникаций.

Свидетельство о регистрации ПИ № 77-15565 от 02 июня 2003 г.

Оригинал-макет ООО "Авансед солюшнз". Отпечатано в ООО "Авансед солюшнз".

119071, г. Москва, Ленинский пр-т, д. 19, стр. 1.