

ИНСТРУМЕНТЫ СЕМАНТИЧЕСКОЙ НАВИГАЦИИ ДЛЯ ФОРМИРОВАНИЯ РЕСУРСОВ ДИСТАНЦИОННОГО ОБРАЗОВАНИЯ

Чеушев К.В.¹, Глухова Л.А.²

¹ Centre for Studies on Inclusive Education, г. Минск, Беларусь,
cheushevcyrill@gmail.com

² Белорусский государственный университет информатики и радиоэлектроники, г. Минск, Беларусь,
glukhova@bsuir.by

Abstract. The problem of forming information sources (IS) of system of distance education is analyzed. An approach is proposed to improve and accelerate the process of creating information systems, based on the use of the instruments of the «Semantic navigation» complex. The usage of these tools will help to make free the subject area expert from routine work at reprocessing and systematization of initial IS components and to make better the quality of selection of determinants and structures on automatic reprocessing of initial IS base.

Одной из важных задач в сфере образования Беларуси является обеспечение электронными средствами обучения (ЭСО) и электронными информационными ресурсами (ИР) учебных курсов и программ, создание возможностей участникам образовательного процесса оперативно получать необходимую для обучения информацию, средства самообразования и тренинга с удалённых компьютеров [2]. В целом к настоящему времени уже осуществлены первичное накопление электронных образовательных документов и их сборка в хранилища данных, обеспеченные механизмами поиска и структурирования. В данный момент актуальной является задача выстраивания из множества накопленных первичных компонентов целостной системы образовательных ИР с развитыми возможностями поиска и навигации, полным и точным удовлетворением информационной потребности пользователя (ИПП), реализацией методических концепций, соответствующих учебным целям.

Реализация данной задачи предполагает внесение в образовательный ресурс высокой степени декомпозиции на элементы знаний, методически обусловленное связывание их между собой и пересборку в единое целое, соответствующее конкретным целям обучения [2].

Для её решения в Центре информационных ресурсов и коммуникаций Белгосуниверситета была создана технология проектирования баз знаний (БЗ) и информационных систем на их основе. Данная технология реализует математический аппарат семантических сетей, опирается на СУБД и XML-технологии с целью семантической разметки ИР, пользовательские веб-интерфейсы. В семантической сети ИР представлены элементами знаний с высокой степенью декомпозиции и многообразными семантическими связями. Это, вместе с механизмом иерархического наследования и правилами вывода, позволило обеспечить высокие функциональные возможности и гибкость БЗ как образовательного ИР [1].

Одной из главных практических трудностей создания образовательных ИР с высокой степенью декомпозиции, семантической разметкой и сборкой в БЗ является высокая трудоёмкость работ экспертов предметной области. Это особенно актуально для высшего образования, поскольку для узких предметных областей нелегко найти и специалистов, которых

можно было бы привлечь к формированию высококачественных ИР дистанционного обучения. При этом ИР желательно актуализировать как можно чаще.

Преодолеть это затруднение и вывести на качественно новый уровень технологию формирования ИР дистанционного образования могли бы инструменты автоматизации семантической переработки имеющихся первичных ИР. Эти инструменты, с одной стороны, помогли бы освободить эксперта от рутинного труда по пересмотру больших массивов текстовых данных ради нахождения ключевых слов семантической разметки ИР, которые обеспечили бы полноту и точность удовлетворения ИПП, а с другой – освободили бы интеллектуальный ресурс эксперта для воплощения его методического замысла.

Развитием описанной выше технологии формирования образовательных ИР в виде баз знаний являются предлагаемые в докладе методики и инструменты, называемые «инструменты семантической навигации» [3]. Идея их применения состоит в том, чтобы выполнить лингвистический разбор и первичное индексирование текстов ИР (а работа эксперта ведёт и к вторичному семантическому индексированию) и применить в форме информационно-аналитической системы комплекс интеллектуальных функций. Данные функции могут использовать:

- конечные пользователи для доступа к ИР;
- эксперты для интеллектуальной переработки первичных компонентов ИР в более высокоразвитый вторичный ресурс.

Под семантической навигацией понимается процесс движения пользователя от какой-то позиции, возможно семантически даже отдалённой от искомого результата (т.е. допускается нечёткое начальное представление о цели), в пространстве ИР, при котором он имеет возможность итерационно корректировать свои цели и замысел и переходить в соответствии с ними к другим подобластям ИР. Отличие от навигации в классическом понимании (по заранее установленным ссылкам) состоит в том, что направление движения в пространстве ИР не определено заранее. Оно обеспечивается функциональными аналитическими механизмами непосредственно в моменты взаимодействия пользователя с информационно-аналитической системой.



Для осуществления набора главных поисковых и аналитических функций, которые помогли бы эксперту качественно и быстро выполнить его работу, а конечному пользователю обеспечили бы высококачественную интеллектуальную поддержку целенаправленного движения в ИР, необходимо и достаточно три функциональных механизма:

- структурирование дисперсной среды ИР согласно текущей ИПП;
- экстрагирование семантических детерминантов;
- собственно семантическая навигация как переход от одних элементов ИР к другим.

Перечисленные функциональные механизмы, будучи собранные в последовательности и в иерархию, обеспечивают все нужные смысловые операции разных уровней. При этом в каждую из них на более высоком уровне рассмотрения обычно входят две другие с предыдущего иерархического уровня. Например, операцию кластеризации как структурирование довольно высокого уровня можно выразить цепью операций экстрагирования и семантической навигации более низкого уровня [3].

Под структурированием понимается выделение из дисперсной среды ресурса групп (кластеров) элементов, «откликающихся» на одну или несколько информационных потребностей. Это может быть поиск, и тогда результатом станет кластер найденных первичных компонентов ИР (а потребность сформулирована запросом), или параллельная группа поисков по разным основаниям, и тогда мы получим группы кластеров. Если эти основания заданы заранее, то это задача классификации, а если заранее смысловые основания структурирования не заданы и их нужно определять в процессе, то это кластеризация. Эти операции позволяют выделить из ИР только те группы элементов, которые потенциально удовлетворяют текущей ИПП и только ей. Кластеризация в приведенной модели не только самостоятельная аналитическая задача, но и нередко решается как служебная в другом процессе. Например, построение кластеризованного дайджеста содержит в себе кластеризацию как служебную задачу, хотя по сути является экстрагированием.

Экстрагирование семантических детерминантов – самый интеллектуально насыщенный функциональный механизм группы. Он во многом определяет успех всех аналитических цепочек. Прежде всего это ранжирование элементов ресурса и выделение верхней части ранжированного списка, что есть по сути селекция тех элементов – семантических детерминантов, – которые наиболее полезны для получения планируемого результата. В работе эксперта за выделением значимых элементов следует ещё и придание им структуры (формирование элементов вторичного семантического индекса компонентов ИР). Роль операций экстрагирования, как и других элементов рассматриваемой группы, двойственна: они могут иметь самостоятельное значение (выделение ключевых термов текстов, построение аннотаций, дайджестов кластеров), а могут входить в состав других операций. Например, кластеризация основана на экстрагировании групп ключевых термов, которые могут составлять «смысловое ядро» кластера.

Семантическая навигация как элемент рассматриваемой группы функциональных механизмов прежде всего предполагает переход от одних элементов ИР к другим согласно его методической потребности или конечного пользователя по учебному материалу. И если структурирование и экстрагирование сужают рассматриваемое пространство ИР, то навигация расширяет его. Ключевой элемент этой операции – поиск (или отбор) подобного. Подобие может измеряться по разным основаниям (скажем, поиск стилистически похожих текстов), но чаще всего имеется в виду смысловое подобие. Операция семантической навигации на уровне поиска похожих компонентов ИР включает в себя пару других функциональных механизмов: чтобы найти подобные тексты, из «эталона» экстрагируются ключевые термы и по ним выполняется операция структурирования, поиск и при необходимости классификация или кластеризация найденного.

Семантическая навигация может стать действенным механизмом поддержки формирования высококачественных образовательных ИР. Если создать необходимый набор интеллектуальных инструментов – функций для всех элементов рассмотренной группы, собрать их в более высокоуровневые цепочки, связать их в единое целое с интерфейсными решениями, то образуется мощный комплекс, работающий по принципу взаимодействия эксперта и аналитической системы достижению полезного результата, с положительной обратной связью, реализующий цепь «синергических транзакций» взаимодействия человек-машина. Начав с посылки даже самого малого смыслового сигнала, хоть как-то отражающего замысел, эксперт быстро получит отклик от системы (например, кластер компонентов ИР со сниппетами, подсветкой автоматически выявленных ключевых термов, дайджестом). С полученным результатом он может провести действия, уточняющие потребность (отобрать наиболее релевантные и представительные тексты или предложения, поднять веса некоторым термам и т. п.) и запустить следующую аналитическую цепь, которая даст уже более полный и точный результат – и так далее до достижения нужного качества и достаточной полноты решения поставленной задачи.

Литература

1. Зеков, М. Г. Информатизация школьного образования / М. Г. Зеков. – Минск: Зорны верасень, 2006 – 288 с.
2. Курбацкий, В. Н. От базы данных выпускных работ слушателей к базам знаний по школьным предметам / В. Н. Курбацкий, В. Н. Чеушев // XVI конференция-выставка «Информационные технологии в образовании»: Сборник трудов участников конференции. Ч. II. – М.: МИФИ, 2006. – С. 29-30.
3. Чеушев, К. В. Семантическая навигация в больших информационных массивах / К. В. Чеушев // Компьютерные системы и сети: материалы 52-й научной конф. аспирантов, магистрантов и студентов. (Минск, 25 – 30 апреля 2016 года). – Минск: БГУИР, 2016. – С. 89-91.