

Цифровите библиотеки в обучението на учители

Красен Стефанов, Павел Бойчев, Елиза Стефанова, Атанас Георгиев,
Николина Николова, Александър Григоров

Резюме: Настоящата статия разглежда въпроси, свързани с ролята на съвременните европейски цифрови библиотеки в обучението на учители. В нея се дискутират изграждането в рамките на проекта **Share.TEC** на цифрово хранилище за метаданни на ресурси за обучение на учители и портал за общността, работеща с него. Описани са очакванията на учителите и обучаващите ги от системата, а също и моделът за адаптивност, който е избран да стои в основата на системата, както и неговата реализация в портала.

Въведение

Феноменът Web 2.0, актуален в момента, поставя акцент върху споделянето на цифрови ресурси и поражда дискусии относно най-добрите практики при тяхната употреба. Друг ключов компонент на този феномен е адаптивността – възможностите на системите да се адаптират към нуждите и желанията на потребителите. Важна част от възможностите за адаптивност на системите е да съветват потребителите и да им препоръчват нови ресурси и възможности. Тези нови възможности поставят сериозни предизвикателства пред подготовката на учителите – да се прилага онзи стил, който би бил най-гъвкав и приемлив за учителите, както и да се подготвят да обучават и да се учат в съвременното Интернет общество.

За да могат учителите да споделят по най-добрия възможен начин наличните им знания и ресурси, е необходимо да се справят с един основен проблем: наличните ресурси са разхвърляни между многобройни източници и недобре структурирани и това затруднява намирането им. Има голям недостиг от метайнформация, която може да подпомогне процеса на търсене на точната информация, в точното време и на точното място.

Обучението като цяло и в частност това на учители е бавен процес, за който е трудно да възприеме всеобхватно иновации като тези, предлагани от инструментите на Web 2.0. Националните образователни системи са обвързани с културните особености на съответните нации, обучението на учителите (ОУ) има локална обусловеност и за сега споделянето на цифрови ресурси на международно ниво е все още рядко и е в ембрионален стадий.

Благодарение на информационните технологии, някои от традиционните характеристики на обучението на учители претърпяват модификации. Съвременният модел на обучение включва:

- педагогически характеристики и класификация на ресурсите за ОУ;;
- обучение, ориентирано към индивидуалните потребности и специфичните компетенции;
- способност за разбиране на различни езици и различни култури;
- способност за търсене, анализиране и повторна употреба на информация в цифров вид.

Проектът **Share.TEC**: Sharing Digital Resources in the Teaching Education Community [1] е насочен към подпомагане на учителите и на преподавателите, подготвящи учители, да постигнат тези общи цели. Проектът изпълнява няколко специфични задачи: изграждане на модерна система, центрирана около потребителя; агрегиране на общоевропейски метаданни; осигуряване на персонализирано, съобразено с индивидуалните културни особености посредничество; подкрепа при развиването на перспектива сред тези, които работят в- или с обучавани учителски общности; мощна и гъвкава система за препоръчване на подходящи учебни ресурси, като се вземат предвид поведението и личните предпочитания на потребителите; удобни и ориентирани към общността характеристики на Web 2.0, като например: коментиране, оценяване и категоризиране на образователните ресурси, изграждане на общности от учители със сходни интереси.

Цифрови библиотеки

Според дефиницията една съвременна цифрова библиотека включва следните няколко задължителни компонента:

- съдържа научна и научно-популярна литература и други научни и/или образователни обекти в цифров вид, като достъпът се осъществява с помощта на информационните технологии;
- обектите са систематизирани в тематични йерархии, класификации или онтологични схеми;
- всички обекти са описани с подходящо избран формат на метаданни, който съответства на основните стандарти за цифрови библиотеки на IMS и поддържа напълно стандарта OAI-MPN.

Съдържанието може да се съхранява локално или да е достъпно по мрежата. Цифровите библиотеки могат да съдържат всякакъв вид „физически“ материали, които са били дигитализирани (цифровизирани) – книги, аудиовизуални или мултимедийни материали, фотографии, архивни документи и други, както и материали, оригинално създадени в цифров формат. Основната функция на тези библиотеки е споделянето на знания чрез:

- създаване и управление на цифрови колекции;
- осигуряване на свободен достъп до водещи световни научни постижения;
- споделяне на цифрови ресурси за учене;
- увеличаване на видимостта и разширяване на въздействието;
- демонстриране на най-добрите налични резултати и продукти;
- бързо и изчерпателно търсене и намиране на нужната информация.

Метаданните са ключът за изпълнението на всички тези функции, осигурявайки нужната информация за значението (смисъла) на оригиналните ресурси, като ги прави по-прозрачни, лесни за намиране и употреба. Те представляват допълнителни данни, които включват различни характеристики, особености, връзки и свойства на оригиналните данни. По-рано метаданните са използвани най-вече за каталогизиране на книги в традиционните библиотеки. Днес вече метаданните са ключ за търсене, намиране и използване на точните данни.

Чрез използването на метаданни ние се придвижваме от дигиталните библиотеки към специализирани хранилища за масиви от метаданни, комбинирани с допълнителна семантична информация (най-често под формата на онтологии и таксономии) – цифрови складове или хранилища на метаданни, които се опитват да категоризират и свържат всички възможни информационни ресурси в съответната област.

През последните няколко години в Европейския Съюз (ЕС) се реализира сериозна и амбициозна програма за изграждане на инфраструктура от дигитални библиотеки. Най-мощната инициатива е свързана с проектите Europeana [2] и DRIVER [3] в рамките на програмите eContent [4] и eContentplus [5]. В рамките на тези проекти бяха проведени широкообхватни проучвания на състоянието на цифровите библиотеки във всяка една страна от ЕС като настоящето изследване се базира на техните резултати и прави крачка напред в посока на приложение на цифровите библиотеки за обучение на учители.

Системата *Share.TEC*

Като база за осъществяване на задачите на проекта Share.TEC са използвани различни цифрови библиотеки с ресурси, подходящи за обучение на учители. В тази секция са представени архитектурата на системата Share.TEC, общ модел за описание на подобни ресурси в цифрови библиотеки, предвиждащ отчитане на метаданните, и специализирана онтология, моделираща процесите на обучение.

Използвайки съществуващите технологии, партньорите по проекта Share.TEC си поставиха за цел да се изгради портал, който да е във висока степен видим и функционален, предлагащ модерни услуги, осигуряващи персонализиран достъп до широк спектър от учебно съдържание, предназначено за подготовка на учители [6]. Централното хранилище [7] за съхраняване на метаданни на ресурси за обучение на учители (ОУ) е сърцевината на системата (фиг. 1).

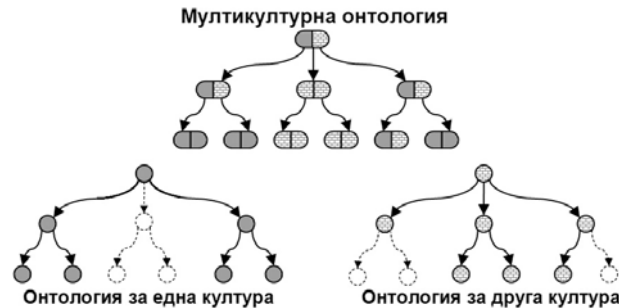


Фигура 1. Архитектура на Share.TEC системата

В рамките на проекта Share.TEC е разработена специфична онтология, наречена Онтология за обучение на учители – Teacher Education Ontology; TEO [8, 9, 10]. Целта беше да се осигури по-устойчив, гъвкав и мощен начин за класифициране на ресурси за ОУ в централното хранилище на системата.

ТЕО [11] е ориентиран към света на обучението на учители и по-специално към цифрови ресурси и добри практики от цяла Европа. Онтологията има многопластова структура (фиг. 2) с общ основен слой, който може да бъде приложен или надграден чрез слоевете на конкретни, културно специфични онтологии.

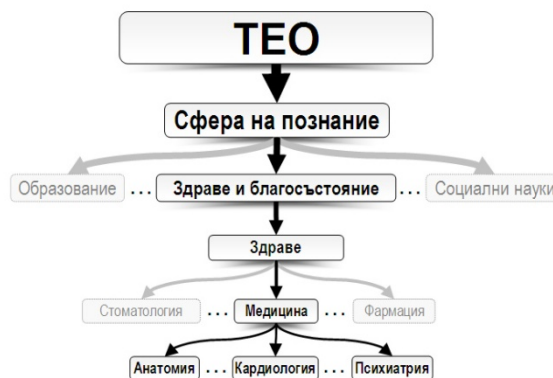
Целите на ТЕО са да поддържа педагогически обусловено цифрово съдържание; да представи потребителя и компетенциите му; да предложи многоезиково, мултикултурно и персонализирано взаимодействие с адаптивни приложения; а също и да поддържа йерархично и динамично търсене.



Фигура 2. Многопластова структура на онтологията

ТЕО се стреми да обхване онези области, които се считат съществени за описването, обмена, споделянето и разработването на ресурси, специално посветени на обучението на учители. Комплексна структура на онтологията е организирана в разклонения (фиг.3), отнасящи се до:

- цифровото съдържание;
- компетенциите (на ниво учебен предмет, мета-когнитивно и др.);
- сферата на познание;
- контекста от дейности в рамките на съответната област на обучение на учителите и
- ролята в контекста на ОУ и в системата Share.TEC.



Фигура 3. Разклонения на онтологията

Всички метаданни в хранилището следват формата на общия модел за метаданни (ОММ) [8, 12], а той е базиран на формата на стандартния модел за описание на учебни обекти (Learning Object Model, LOM) [13]. Общият модел на метаданни представлява разширение на LOM, в което са добавени секции за педагогическите характеристики на описваните ресурси. Съществената разлика се състои в педагогическите характеристики на цифровизираното съдържание.

Очаквани функционалности от *Share.TEC* според учителите

За да бъдат идентифицирани основните функционалности, желани от потребителите на разработваната система, в страните партньори в проекта Share.TEC бяха организирани семинари с учителите и преподавателите в университети, включени в обучението на учители.

Поставяйки пред участниците въпроса: *Какви пречки срещате, когато търсите цифрови материали?* голяма част от тях отговориха:

- *Повечето цифрови ресурси нямат описание и трябва да отворя и разгледам файла, за да реша дали е полезен за мен.*
- *Информацията не е добре класифицирана.*
- *Понякога трябва да видиш целия материал, за да разбереш дали е използваем в твой случай.*

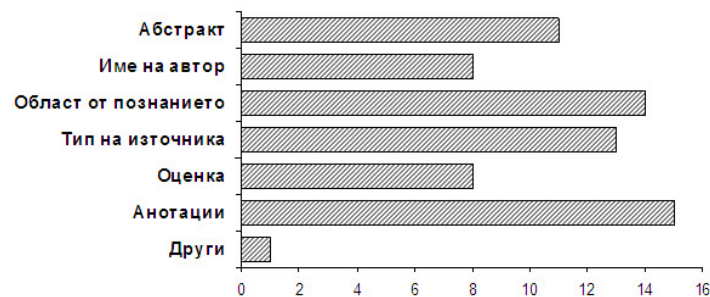
Изказаните от участниците проблеми, затвърдиха нашето първоначално убеждение, че има нужда от хранилища на цифрови ресурси, които да бъдат класифицирани и описани с допълнителни метаданни.

Но достатъчно ли е такова описание на нашите потенциални потребители? От натрупаните впечатления от работа с учители и обучаващите ги имахме основание да изкажем хипотеза, че не е достатъчно и, че съществена роля играе професионалната общност. Хипотезата се потвърди от твърденията, изказани от участниците, относно професионалните общности, където доминираха изказвания като:

- *За мен, като учител, е много важно да прочета коментари за дадени ресурси, за да придобия представа за тяхното качество и възможности за използване*
- *Отнема много време и усилия да разработиш/създадеш нещо сам; понякога е дори невъзможно*
- *Мнението на колегите, с които работя, е важно за мен*

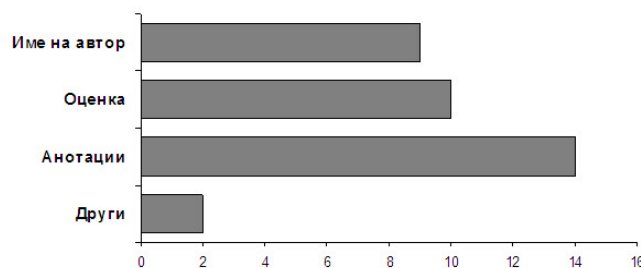
Т.е. учителите и обучаващите ги очакват от система като Share.TEC като основна функционалност не само намирането на класифициран ресурс, а и възможност за споделяне на опит посредством мнения, коментари, оценки на колеги и други.

За да може да бъде подпомогнат чрез системата избора на даден ресурс от учителите, искаме да разберем мнението им по въпроса *Каква информация ще ги убеди да изберат ресурс?* Участниците в семинара показаха (фиг. 4), че за тях най-голямо значение има анотацията на ресурса и дали той е в интересувашата ги познавателна област. Голямо значение имат също типа на ресурса, резюмето (абстракт), автора му и рейтинга (оценката), който е получил от другите потребители.



Фигура 4. Степен на значимост на характеристики при избор на ресурс

Аналогични резултати (фиг. 5) се получиха и в отговор на въпроса *Коя информация допринася за разбиране качеството на ресурс?* И в този случай потребителите отдават най-голяма роля на анотацията на ресурса, и почти еднаква значимост на автора му и оценката, направена от другите от общността.



Фигура 5. Степен на значимост на характеристики при оценяване качеството на ресурс

Много от анкетиранияте се интересуват от избора на другите потребители в общността. Над 75% се интересуват какво са изтегляли другите, а почти 30% – какво са гледали.

Анализирайки резултатите [14] от семинарите, можем да заключим, че онтологията за обучение на учители и общият модел за метаданни, разработени в рамките на проекта Share.TEC, осигуряват подходяща среда за класификация на цифрови и печатни ресурси за обучение и самообучение на учители. „Удовлетворени сме, че най-сетне ще имаме възможности да събираме и лесно да намираме необходими материали”, сподели един от участниците в семинара. Обратната информация, получена от участниците, показва, че ТЕО и ОММ могат да удовлетворят изискванията на потребителите, и че те са подходяща база за разработване на портала Share.TEC. В резултат от проучванията и семинарите с учители бяха идентифицирани основните функционални характеристики на портала Share.TEC и в съответствие с тях, започна разработването на потребителски интерфейс и прототип на системата.

Модел за адаптивност в Share.TEC

Share.TEC представлява архив включващ анотации на ресурси от областта на обучение на учители в Европа, разположени в множество цифрови библиотеки. Този архив предоставя достъп до всички тези ресурси с отчитане на културните особености. Общността от потребители на системата включва хора с много различни квалификации, идеи и очаквания, различни изисквания и очаква тя да отговори на многообразието от езици и култури. Затова съществено важно за мисията на Share.TEC е разработваната система да има възможности за поддръжка на разнообразието. Ключовият аспект, който лежи в основата на тази поддръжка, е адаптивността.

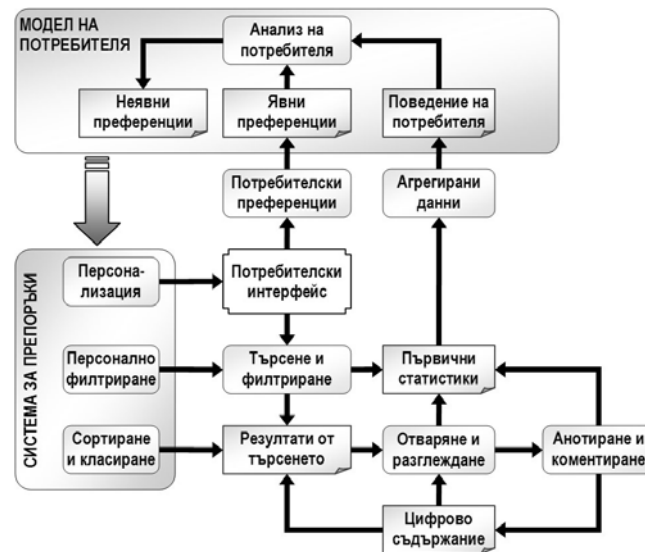
В този контекст три основни взаимосвързани елемента играят ключова роля: адаптивност, моделиране на потребителя и потребителски интерфейс.

На Фигура 6 е показан общ модел за адаптивност, включващ тези три ключови елемента и структуриран около тях, но отчитащ и представящ по-подробно специфичните процеси и данни, свързани с тях. Адаптивността е постоянен и итеративен процес, който продължава през целия жизнен цикъл на системата.

Две основни потребителски дейности, свързани с адаптивността в цифровите библиотеки, са настройките на предпочитанията и търсенето на цифрово съдържание.

Когато един потребител установява настройките си (като например език или област на интереси), те се съхраняват в модела на този потребител. Всеки потребител е асоцииран с негов собствен модел, който съдържа три клъстера от данни:

- *Явно зададени предпочитания* от потребителя, които се използват за настройка на системата;
- *Косвени предпочитания на потребителя*, които се дефинират автоматично от системата чрез анализиране на данните за поведението на потребителя и неговите настройки.
- *Обобщени данни за поведението на всички потребители*, които се генерират от натрупаните статистически данни за използването на системата или тяхна агрегирана версия.



Фигура 6. Архитектурата, реализираща адаптивността

При операциите за търсене и филтриране на намерените резултати в хранилището се създават два клъстера от данни. Единият клъстер е списъкът от резултати при търсене: потребителят може да разглежда неговите елементи, да отваря тези, които представляват интерес за него и освен това да коментира или аотира ресурси. Всички тези дейности водят до натрупване на данни във втория клъстер, който съдържа необработени данни като използвани за търсене ключови думи, намерени резултати, отваряни записи, написани коментари и анотации, рейтинги и други.

За да могат да бъдат използвани тези данни, те трябва да бъдат агрегирани и обработени във форма, която описва потребителското поведение. Както беше посочено по-горе, тези данни ще бъдат анализирани за генериране на косвените потребителски предпочитания.

Последното множество от елементи в архитектурата за адаптивност е свързано със системата за препоръчване. Тези елементи използват данни от потребителския модел, а именно явните и косвените потребителски предпочитания, и решават как да се адаптира функционалността и поведението на системата, така че да съответстват най-добре на тези предпочитания.

Системата за препоръчване включва няколко модула, които управляват различни аспекти на адаптивността. **Настройващият модул** е отговорен за адаптиране на потребителския интерфейс. Това включва предпочитания за език, цвятова схема и разположение на компонентите на интерфейса. **Модулът за интелигентно филтриране** предоставя допълнителни филтри, в зависимост от потребителските интереси (явно зададени или косвено извлечени). Този модул също инициализира стойностите на критериите за филтриране. **Модулът за класиране** (рейтинг, rating) влияе върху реда на резултатите при сортиране, предлагащ на по-предни места тези записи, които отговарят в по-голяма степен на предпочитанията на потребителя.

Системата за препоръчване в Share.TEC предлага множество от функции, подпомагащи потребителят да намери най-подходящите ресурси, съобразно профила му и историята на неговите действия в системата.

Функцията за класиране предлага персонализирано подреждане на намерените резултати от търсенето. Тя може да използва няколко различни схеми за класиране, например цифрови стойности в някакъв диапазон (например 1-100), използване на звездички, статистически графики или колони и т.н. Потребителят може да избере какъв тип за класиране да използва. Желаната схема може да бъде избрана в потребителския профил или да бъде посочена изрично при всяко търсене.

Към всеки резултат е прикачена функция, изчисляваща колко близко е до предпочитанията на потребителя. Всеки потребител може допълнително да укаже за всеки резултат доколко точна е оценката за релевантността му.

Функцията за отбелязване маркира всеки резултат доколко той е бил препоръчван от други потребители. Възможни са няколко опции както и комбинация от тях:

- препоръчван чрез разглеждане от други потребители;
- препоръчван чрез разглеждане от други потребители, които разглеждат подобни ресурси;
- препоръчван чрез разглеждане от други потребители със сходни профили.

Функцията за препоръчване е свързана с възможностите на системата за препоръчване да предлага различни типове от ресурси на главната страница на потребителя, например:

- ресурси оценени най-високо от потребители със сходни предпочитания и интереси;
- най-често разглеждани ресурси;
- потребители които са най-близко до текущият потребител (със сходни предпочитания и интереси, с най-близка история на търсене и разглеждане на ресурси, или по друг критерий).

Системата за препоръчване поддържа и някои социални функции. Намирането на сходни потребители е процес на намиране на потребители с подобни интереси и поведение. Той използва класифициране на потребителите по стереотипи (или кръгове на подобие, които групират потребителите по техните степени на близост). Интерфейсът скрива сложността на този процес и предлага опростен и лесен начин за търсене. Системата предлага също възможности за търсене на други потребители по множество параметри, в зависимост от това кой търси и каква начална информация е налична. Например, потребителят може да търси определен потребител по име или електронен адрес, или група от потребители в същата страна или град, или тип училище, или потребители със сходни интереси.

Търсенето на ресурси чрез използване на близост до други потребители позволява да се откриват такива ресурси, които са оценени високо от потребители с близки интереси и поведение до текущия.

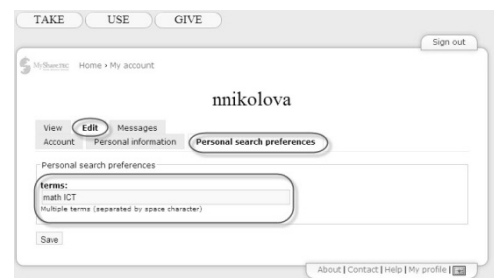
Функцията за намиране на потребители с близки интереси или поведение е основа и на социалните функции за създаване на групи по интереси. Например, потребителят може да поиска от системата “Уведоми ме, ако някой подобен на мен се присъедини към Share.TEC”.

Реализация на модела

Моделът за адаптивност и свойствата, описани в предходната секция, намират широко приложение и са реализирани в системата Share.TEC.

За да подпомогне възможно най-добре потребителите, порталът Share.TEC [15] е разработен така, че да позволява автоматичната персонализация на интерфейса на потребителя в съответствие с неговия език, област на знания и интереси (фиг.7), предпочитания, роля в професионалната общност и история на предходните му действия. Всички заявки могат да бъдат извършвани на всеки от езиците на страните-партньори в проекта Share.TEC (български, холандски, английски, италиански, испански и шведски) и в съответствие с контекста на потребителя чрез:

- опростен или комплексен набор от параметри на заявки за филтриране;
- автоматично установяване на ключови стойности на параметри в съответствие с профила на потребителя (с възможност за промяна от потребителя в последствие);



Фигура 7. Интереси в профила на потребителя

- възможност за установяване на няколко стойности за някои параметри.

На базата на персоналния профил, анализ на историята в системата и потребителското поведение, Share.TEC порталът предлага богат набор от адаптивни свойства, включително система за препоръчване. В тези свойства се включват: (1) автоматично класиране на резултатите от търсене в съответствие с профила на потребителя; (2) изчерпателни препоръки към потребителя, отчитащи оценките, отметките и анотациите на другите потребители; (3) индивидуална форма, контролирана от потребителския профил за задаване на оценки, отметки и анотации.

Потребителски профил (Модел на потребителя)

Моделът на потребителя се състои от два основни компонента:

- *явно зададен профил*, който включва персонална информация за потребителя (къде работи, предпочитан език, ниво на експертност и др.), която е малко вероятно да бъде променяна често;
- *косвена информация за потребителя*, базирана на начина на използване на системата, отчитаща оценки на ресурси, история на търсене на ресурси, анотиране и разглеждане на ресурси и др.

Допълнително системата използва статистическа информация относно начините за използването ѝ от потребителите, което позволява прилагането на различни алгоритми за изчисляване на близост между ресурси, потребители и други обекти.

След регистриране всеки нов потребител получава статут *начинаещ* (специален тип потребител с предварително установени предпочитания и начини за използване на системата), тъй като в този момент на системата не е известно нищо и няма как да бъдат прилагани алгоритмите за адаптивност при липса на данни. Този статут се променя веднага след като в системата се натрупат достатъчно данни за него.

Тъй като повечето данни, формиращи модела на потребителя, се менят динамично, с цел повишаване на ефективността на системата, самият потребителски профил не се съхранява в цифровата библиотека, а във вътрешна за портала база от данни. Потребителският профил обаче е тясно интегриран с онтологията (ТЕО) [11], която се съхранява в цифровата библиотека. Специфични данни за потребителя като професионална област, област на експертност, контекст на работа като учител, които са свързани със съответните клонове на онтологията, могат да приемат стойности само както е указано в съответните възли от онтологията. Това позволява многоезичност в системата Share.TEC, тъй като стойностите във възлите на ТЕО са представени чрез множество стойности (по една за всеки поддържан език). Когато потребителят редактира профила, той избира от списък със стойности, извлечени от онтологията на неговия (или предпочитания) език.

Динамичната част от потребителския профил се използва главно от препоръчващата система за да предложи подходящи ресурси и други съвети на потребителя. За целта динамично се преизчисляват разстоянията между ресурсите и потребителите, като се използват подходящо подбрани метрики. Повече подробности относно реализацията на препоръчващата система могат да бъдат разгледани в [16].

Настройка на процеса на търсене

Процеса на търсене в системата Share.TEC използва интензивно модела на потребителя и онтологията ТЕО. Той е базиран на машината за търсене Solr [17]. Компонентът за търсене осъществява семантично разширяване на заявката за търсене, преминавайки през последователност от стъпките по-долу:

- *Инициализация* - анализ на заявката и достъп до данните от модела на потребителя;
- *Разширяване на базата на предпочитанията на потребителя* - в заявката се добавят изрази явно зададени в профила на потребителя с повишено тегло, така че записи, в които те се срещат ще бъдат посочени на по-предно място в резултата;
- *Разширяване на базата на онтологията* - към заявката се добавят и изрази, които са родители в онтологията на зададените в заявката изрази;
- *Многоезично разширяване* - ако в заявката се срещат ключови думи или фрази от онтологията на даден език, към нея се добавят и техните синоними на другите езици, така че заявката да открие ресурси описани на различни езици;
- *Разширяване за препоръчване* - заявката отчита списъка на най-често разглежданите ресурси в съответствие с профила на потребителя, за да ги класира на по-предно място в списъка с намерените резултати.

Статистически данни за потребител

За да може да се реализира адаптивността, системата трябва да натрупва и съхранява за всеки потребител данни, свързани с неговите явни предпочитания, как той използва системата, културния и езиков контекст, както и други статистически данни свързани с анализа и съпоставянето на неговите действия и наличните данни в онтологията и модела за метаданни. Когато данните от тези източници са в конфликт помежду си, се дава приоритет на явно зададената информация от потребителя.

За съхраняване на данните, свързани с начините за използване на системата, се използва масив от числови индикатори, свързани със съответни елементи от потребителския интерфейс. Стойността на всеки индикатор се увеличава при всяка операция на потребителя, свързана със съответния елемент

(заявка, анотация, оценка и др.). Множеството от тези индикатори се използва за пресмятане на различните типове данни, формиращи съхраняваната статистика за всеки потребител.

Многоезичност

Тази функция е реализирана на основата на предоставените възможности от системата Drupal [21], с която е разработен портала на ShareTEC. Тя поддържа неограничен брой езици, за които при наличие на подходящ превод, всички основни системни компоненти (имена, съобщения, менюта и други) от потребителския интерфейс се визуализират на избран език.

Системата ShareTEC осигурява многоезичност и поради многоезичността заложена в модела на метаданните и в онтологията.

Система за препоръчване

Системата за препоръчване в ShareTEC включва всички модули и функции, описани в предходната секция, където бе представен модела за адаптивност. Разработката на системата е базирана на системата с отворен код Apache Mahout съвместно с използването на машината за търсене Solr.

Системата Apache Mahout [18] включва в себе си няколко основни алгоритъма за препоръчване, както и за клъстеризация и класификация на големи обеми от данни. Тя използва системата с отворен код Apache Hadoop за максимално ефективно изпълнение на тези алгоритми върху голям обем данни в разпределена среда [18, 19]. Системата Mahout включва няколко алгоритъма за клъстеризация базирани на модела Map-Reduce (k-Means, fuzzy k-Means, Canopy, Dirichlet, and Mean-Shift), алгоритми за класификация (Distributed Naive Bayes и Complementary Naive Bayes), алгоритми за еволюционно програмиране, както и библиотеки за матрични и векторни пресмятания.

Съхраняване, индексирание и обработка на данните

В Share.TEC системата за препоръчване и машината за търсене използват различни източници на данните, като записите в цифровата библиотека следващи общия модел на метаданни (ОММ) [8], онтологията за обучение на учители (ТЕО) [8, 11], както и явно и неявно зададената информация в модела на потребителя. Докато ОММ записите и онтологията ТЕО се съхраняват в цифровата библиотека, останалите данни свързани с модела на потребителите и статистиките от използването на системата се съхраняват в отделна база от данни в Share.TEC портала.

Машината за търсене Solr [17] е конфигурирана да използва като индекси отделните полета от ОММ и ТЕО. Тези индекси се изграждат постепенно и се обновяват при добавянето на всеки нов ОММ запис в цифровата библиотека.

В базата от данни на Share.TEC портала се съхраняват данни описващи всяко действие в системата – разглеждане на ресурс, оценяване, отмятане или коментиране на ресурс, заявка за търсене и др. Това се реализира чрез създаване на специални задачи в подсистемата Hadoop, които работят в асинхронен режим с базата от данни в портала. Използват се други допълнителни задачи в Hadoop за обработка на данните свързани с модела на потребителя, изчисляване на близостта между потребителите въз основа на предефинираните метрики [20] и за изчисляване на стереотипите (пръстените на близост).

Функции за препоръчване

Тук са представени всички реализирани в системата основни функции за препоръчване, заедно с конкретни примери от системата.

Препоръчване на всички потребители

Някои ресурси се препоръчват на всички потребители, включително и на анонимните. Такива са ресурсите с най-често разглеждани описания на метаданните им. Системата може също да показва и най-високо оценените ресурси.

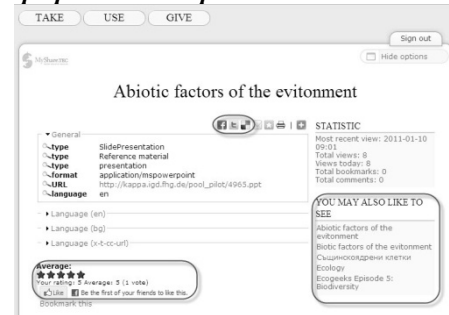
Достъпно за всички потребители е реализираното препоръчване, базирано на съдържанието. Когато потребител разглежда описанието на метаданните на ресурс, той може да щракне върху връзката “Виж повече като този” и системата да покаже списък от подобни ресурси получен чрез текстово сходство на документи. За тази цел настоящата реализация използва полетата с метаданни заглавие и описание, но може да се конфигурира да използва и също други полета на ОММ.

Препоръчване, базирано на потребителски профил

Системата препоръчва ресурси, които съответстват на явните предпочитания в потребителския профил (работен език, област на знания, професионална област, професионален опит, обучение, учителска практика и други). Реализацията използва Solr машината за търсене, за да намери съответните ресурси. Автоматично се генерира заявка, която използва механизъм за повишаване на значимостта на изрази, които съответстват на профила на потребителя. Факторите за това повишаване са избрани по подходящ начин, така че резултатите в началото на списъка най-добре отговарят на явните предпочитания в потребителския профил.

Препоръчване на най-често разглежданите ресурси съобразно профила на потребителя

В този случай системата препоръчва ресурсите, които са разглеждани най-често (фиг. 8) от други потребители, близки до текущия потребител, съгласно неговия профил. Реализацията възприема препоръки базирани на потребителите. Първо на базата на предварително дефинирани метрики се изчислява близостта на текущия потребител до други [20] и тези с най-близък профил се разглеждат като *съсед*. Разстоянията между потребителите вече асинхронно е изчислено от подзадача в Nadoor. Тогава препоръчващата система избира ресурси, които са най-често разглеждани от потребителите *съсед* и неразглеждани от текущия потребител.



Фигура 8. Препоръчване на ресурси

Препоръчване на потребители съобразно профила на потребителя

Системата препоръчва потребители, които са най-близко до текущия потребител съобразно неговия профил. Реализацията използва предварително изчислено разстояние между потребителите и намира *съседите* на текущия потребител по същия начин както ресурсите.

Препоръчване на потребители съобразно разглежданите ресурси

В този случай системата препоръчва потребители, които са близко до текущия потребител съобразно разглежданите ресурси и метаданните им.

Търсене и рейтинг на резултатите от търсенето

Търсенето е една от централните функционалности на Share.TEC портала. То играе важна роля в адаптивността на системата. Реализацията използва разширените възможности на Solr машината за търсене като пълнотекстово търсене, търсене с филтриране, оценяване на намерените резултати и други.

Търсенето с филтриране дава на потребителя възможността да определи по-точно намерените резултати като използва различни класификации за филтриране. Потребителят след това може допълнително да филтрира (фиг.9) списъка с резултати от заявката, например по език, формат и т.н.

Друго важно свойство на реализацията на търсене е рейтинга на намерените резултати (фиг.9). По подразбиране (при базово ниво на адаптивност) резултатите при търсене са сортирани по уместност съобразно заявката за търсене. При максималното ниво на адаптивност при търсене се вземат под внимание предпочитанията на потребителя и интересите му, явно изразени в потребителския му профил. Заявката се разширява и използва механизъм за повишаване на значимостта на изрази, които съответстват на профила. След това, използвайки онтологията ТЕО, заявката може да бъде допълнена да връща повече смислени резултати, но с по-малка релевантност. Не на последно място се използва препоръчващата система за ресурси за получаване списъка на най-гледаните ресурси според потребителския профил. Така заявката се разширява с цел повишаване значимостта на тези ресурси.



Фигура 9. Рейтинг от други потребители и филтриране

В настоящият момент системата е напълно работоспособна и бе успешно демонстрирана на няколко специализирани конференции и конгреси.

Заклучение

В тази статия представихме един съвременен модел за адаптивност в цифрови библиотеки, насочени към обучението на учители. Моделът е разширен с възможност за използване на различни препоръчващи системи. Уверени сме, че разработените модели ще бъдат приложими и полезни не само в рамките на проекта *Share.TEC*, но и за разработването на други продукти, свързани както с обучението на учители, така и с разработването на цифрови библиотеки в сферата на образованието.

Системата ShareTEC бе тествана успешно по време на няколко специални сесии с участието на водещи специалисти от цяла Европа. Също така бяха проведени продължителни пилотни експерименти в седем Европейски страни. Системата бе успешно демонстрирана по време на редица конгреси и международни конференции, както и на специализирани демонстрационни сесии за представители на най-значимите организации свързани с обучението на учители в Европа. Убедени сме, че системата ще намери широко приложение и в България, за което говори сериозният успех на всички досегашни експерименти с учители и че цифровите библиотеки са реално място за обмен на ресурси и опит на виртуални професионални общности от обучаващи.

Библиография

1. Сайт на проекта Share.TEC , <http://www.sharetecproject.eu> (последно посетен на 8.01.2011г.)
2. Сайт на проекта Europeana, <http://version1.europeana.eu/web/europeana-project/> (последно посетен на 8.01.2011г.)
3. DRIVER project site, <http://www.driver-repository.eu/> (последно посетен на 8.01.2011г.)
4. eContent program, <http://cordis.europa.eu/econtent/> (последно посетен на 8.01.2011г.)
5. eContentplus program, http://ec.europa.eu/information_society/activities/econtentplus/index_en.htm
6. Stefanov, K., Boytchev, P., Grigorov, A., Georgiev, A., Petrov, M., Gachev, G., and Peltekov, M. (2009), "Share.TEC System Architecture", In Proceedings of 1st International conference S3T'2009 Software, Services and Semantic Technologies, Ed. D. Dicheva, R. Nikolov, E. Stefanova, pp. 92-99, ISBN 978-954-9526-62-2
7. Stefanov, K., Boytchev, P., Grigorov, A., Georgiev, A., Petrov, M., Gachev, G., and Peltekov, M. (2009), "Share.TEC Repository System", In Proceedings of 1st International conference S3T'2009 Software, Services and Semantic Technologies, Ed. D. Dicheva, R. Nikolov, E. Stefanova, pp. 84-91, ISBN 978-954-9526-62-2
8. Share.TEC Deliverable D2.3 Ontology and Metadata Models: release versions, http://www.share-tec.eu/content/1/c6/04/41/02/DEL2_3Finalreleaseversion.pdf (последно посетен на 8.01.2011г.)
9. Alvino, S., S. Bocconi, J. Earp, L. Sarti (2008) A Teacher Education Ontology for Sharing Digital Resources across Europe, in Proceedings of the 5th International TENCompetence Open Workshop "Stimulating Personal Development and Knowledge Sharing", ed. R. Koper, K. Stefanov and D. Dicheva, pp. 26-29, ISBN: 978-954-92146-5-9
10. Teacher Education Ontology (TEO): version 1, http://www.share-tec.eu/content/1/c6/04/41/02/D2_1_TEO_v1.pdf (последно посетен на 8.01.2011г.)
11. Alvino S., Bocconi S., Boytchev P., Earp J., Sarti L. (2009), "Sharing Digital Resources in Teacher Education: an Ontology-based Approach", In Proceedings of 1st International conference S3T'2009 Software, Services and Semantic Technologies, Ed. D. Dicheva, R. Nikolov, E. Stefanova, pp.116-123, ISBN 978-954-9526-62-2
12. Share.TEC Project Deliverable D2.2 Common Metadata Model (CMM): version 1, http://www.share-tec.eu/content/1/c6/04/41/02/D2_2_Common_Metadata_Model.pdf (последно посетен на 8.01.2011г.)
13. IEEE Standard for Learning Object Metadata 1484.12.1-2002. IEEE LTSC, http://ltsc.ieee.org/wg12/files/LOM_1484_12_1_v1_Final_Draft.pdf (последно посетен на 8.01.2011г.)
14. Е. Стефанова, Н. Николова, Е. Ковачева, К. Стефанов, П. Бойчев (2010) Дигитални библиотеки с ресурси за обучение, съобразени с нуждите на потребителите, списание Littera et Lingua, бр. Лято 2010 , ISSN 1312-6172, <http://www.slav.uni-sofia.bg/liljournal/index.php/bg/component/content/article/383/596-estefanovaetal-summer2010> (последно посетен на 8.01.2011г.)
15. Share.TEC портал, <http://portal.share-tec.eu> (последно посетен на 8.01.2011г.)
16. Tosato , P. "Metrics for Resource-User Matching", Published as deliverable D4.2 of Share.TEC project (2009), <http://www.share-tec.eu/content/1/c6/04/41/02/17D42Metricsforresourceusermatching.pdf> (последно посетен на 8.01.2011г.)
17. Apache Solr home page, <http://lucene.apache.org/solr/> (последно посетен на 8.01.2011г.)
18. Mahout project home page, available at: <http://lucene.apache.org/mahout/> (последно посетен на 8.01.2011г.)
19. S. Owen, R. Anil (2009) Mahout in Action, MEAP Edition, Manning Publications
20. Tosato, P. "Metrics for Resource-User Matching", Share.TEC project deliverable D4.2 (2009), <http://www.share-tec.eu/content/1/c6/04/41/02/17D42Metricsforresourceusermatching.pdf> (последно посетен на 8.01.2011г.)
21. Drupal.org – Leading Portal and Content Management System, <http://drupal.org/> (последно посетен на 12.01.2011 г.)