

Формування комплексного тематичного опису ВНЗ

Роман Корж¹, Андрій Пелешишин²

1. Кафедра електронних засобів комп'ютерних технологій, Національний університет “Львівська політехніка”, УКРАЇНА, м. Львів, вул. С. Бандери, 12, E-mail: korzh@lp.edu.ua

2. Кафедра соціальних комунікацій та інформаційної діяльності, Національний університет “Львівська політехніка”, УКРАЇНА, м. Львів, вул. С. Бандери, 12, E-mail: apele@ridne.net

This article presents a general sequence of the formation of complex thematic image of higher education, namely: definitions for each category, forming a complete list of terms, removal of repetitions, detection synonyms are not included in the list, detection of socially significant shear terms (social tags).

Ключові слова – вищий навчальний заклад (ВНЗ), інформаційний образ, інформаційна діяльність.

Вступ

Другим етапом формування ІО ВНЗ є формування комплексного тематичного опису ВНЗ, який повинен включати не лише тематичну характеристику ВНЗ у цілому, але й тематику кожного підрозділа, що визначений на попередньому етапі.

II. Формування комплексного тематичного опису ВНЗ

Великі обсяги інформації та організаційна складність їхнього формування визначають необхідність прямої участі підрозділів. Водночас, недопустимою є наявність значного числа колізій між підрозділами та упущення окремих тематичних напрямків, які мали би бути розподілені у силу загальної тематики ВНЗ.

Загальна послідовність дій з формування комплексного тематичного опису наведена далі.

Визначення термінів для кожної категорії полягає у трансформації тематичних показників підрозділу $DTDr(Dep_i)$, $DTTh(Dep_i)$, $DTTD(Dep_i)$ у набори термінів (пошукових фраз). Показник $DTTL(Dep_i)$ до множини додається автоматично. Результатом є розширена множина термінів $DTTL^*(Dep_i)$.

$$DTTL^*(Dep_i) = \left(Term(DTDr(Dep_i)) \cup Term(DTTh(Dep_i)) \cup Term(DTTD(Dep_i)) \cup DTTL(Dep_i) \right), \quad (1)$$

де $Term$ – оператор трансформації рубрики у пошукові терміни. Трансформація включає усунення стоп-слів та загальних термінів, приведення слів до базових форм.

Формування повного переліку термінів є організаційним кроком, що полягає в зборі усієї інформації (визначених на попередньому кроці термінів) в єдину БД.

Усунення повторів – очищення БД термінів від повторів.

Виявлення синонімів, що не увійшли у перелік необхідне для повного покриття семантичного поля, у якому здійснюється пошук. На сьогодні інформаційний пошук не в повній мірі передбачає пошук за синонімами. Як наслідок, для складних предметних областей значна кількість сторінок можуть бути не виявлені лише пошуком за певним терміном, попри свою релевантність на семантичному рівні (пертинентність), адже на сторінці може згадуватися не термін, а його синоніми.

Виявлення соціально значимих дотичних термінів (social tags) є фактично розвитком попереднього кроку. БД термінів поповнюється вже не синонімами, а термінами, які відображають релевантність сторінки базовим термінам. Необхідність такого поповнення актуальна для наукоємких та технічно насичених областей, де термінами є складні і рідковживані слова, наукова термінологія. У силу близькості наповнення соціальних середовищ Інтернету текстами, що близькі до усної мови, такі спеціальні терміни зустрічаються рідко. Навзамін використовуються сленгові слова, простіші терміни, торгові марки тощо, за якими доцільно здійснювати пошук. Виявлення соціально значимих дотичних термінів є достатньо складним завданням, яке може бути частково автоматизоване через семантичні мережі, побудовані на публічно створених гіпертекстових масивах (таких як Вікіпедія).

ВИСНОВОК

Формування такого опису є неможливим лише за схемою «згори донизу», коли адміністрація ВНЗ визначає самостійно повний опис тематики підрозділів. З технічної точки зору опис зводиться до побудови множини термінів, що описують предметну область підрозділів ВНЗ і можуть безпосередньо використовуватися для інформаційного пошуку.

Література

1. Корж Р. О. Етапи формування цілісного інформаційного образу ВНЗ та їхня характеристика / Р. О. Корж, А. М. Пелешин // Інформація, комунікація, суспільство 2014 : матеріали 3-ої Міжнар. наук. конф. ІКС – 2014, 21-24 трав. 2014р., Україна, Львів, Славське / М-во освіти і науки України, Нац. ун-т «Львів. політехніка». – Л., 2014. – Бібліогр.: 2 назви.
2. Корж Р. О. Формалізація процесу формування інформаційного образу ВНЗ в соціальних середовищах Інтернету / Р. О. Корж, А. М. Пелешин // Східно-Європейський журнал передових технологій. – Харків, 2013. - №5/3 (65) – С. 4-8. Бібліогр.: 10 назв.