УДК 004.896

Красовська Ганна Валеріївна
Кандидат технічних наук, доцент, доцент кафедри інтелектуальних та інформаційних систем, *orcid.org/0000-0003-1986-6130*
Київський національний університет ім. Тараса Шевченка, Київ

АНАЛІЗ ПІДХІДІВ ДО ПОБУДОВИ МУЛЬТИАГЕНТНОЇ ІНТЕЛЕКТУАЛЬНОЇ СИСТЕМИ КУРИРУВАННЯ КОНТЕНТОМ МВОК

Анотація. *Проаналізовано особливості організації МВОК, визначено коло питань, які необхідно вирішити для підвищення ефективності їх організації та проведення. Зазначено, що під час побудови МВОК для збирання новин в роботі спільноти студентів використовуються тільки RSS-канали та агрегатори, але цих засобів недостатньо для вирішення всього спектру задач курирування контентом. Пропонується проводити подальші дослідження в напрямку створення багатоагентної рефлекторної інтелектуальної системи, яка проводить для заданої предметної області пошук та відбирання інформації в мережі Інтернет, аналіз та категоризацію інформації із застосуванням інтроформаційних методів теорії несилової взаємодії.*

Ключові слова: масові відкриті онлайн-курси; курирування контенту; інтелектуальний інтернет-агент; багатоагентна рефлекторна інтелектуальна система; інтроформація; теорія несилової взаємодії

Вступ

Як зазначено в аналітичній записці відділу гуманітарної політики Національного інституту стратегічних досліджень [1], за останні десятиліття кардинально змінилася система генерації й передачі знань, а їх обсяг багаторазово зріс. Стрімке оновлення теоретичних та практичних знань спричиняє дефіцит висококваліфікованої робочої сили. Масштабність цього явища призвела до прийняття Європейською Комісією Європейської концепції зайнятості та Програми навчання протягом життя, як одного з основних компонентів європейської соціальної моделі, що має на меті забезпечення людей будь-якого віку, соціального статусу, національної, культурної, релігійної приналежності рівним і відкритим доступом до якісного навчання. Це потребує застосування більш гнучких форм навчання, до яких належить відкрита освіта і дистанційне навчання.

Аналіз останніх досліджень і публікацій

За останні роки в освітній процес стрімко увійшли та дещо змінили традиційне розуміння підходів до організації дистанційного навчання масові відкриті он-лайн курси (МВОК, Massive Open Online Course – MOOC).

МВОК створюється та проводиться командою викладачів (тьюторів), яка заздалегідь обирає тему курсу, планує можливу діяльність студентів різного рівня підготовки, обирає посилання на матеріали курсу. У [2] зазначається, що тьютор «підбирає джерела інформації до кожної теми з мережі. Ця інформація слабоструктурована і її може бути дуже багато (понад 100 джерел на тему). Завдання тьютора – найбільш повно представити дану тему і мінімізувати інформаційний шум». Фактично тьютор виступає в ролі куратора контенту (куратора змісту), функціями якого за визначенням [3; 4] є «виявлення, відбирання та поширення кращого і найбільш відповідного он-лайн-контенту та інших інтернет-ресурсів по конкретній темі».

У роботах [3; 5; 6] визначаються такі основні складові курирування контенту:

– **агрегація** – збирання та обмін відповідним змістом;

– **фільтрація** – відбирання найбільш актуальної і цінної інформації;

– **колаж** – об'єднання двох або більше пов'язаних елементів контенту, щоб сформувати нове повідомлення;

– **хронологічні шкали** – організація контенту в хронологічному порядку, щоб показати еволюцію ідеї;

– **соціальні комунікації** – у своїй найпростішій формі курирування змісту це обмін контентом в соціальних групах (для МВОК в спеціалізованих, фахових соціальних групах).

Слід зазначити, що особливостями МВОК, які «суттєво відрізняють їх від традиційних академічних дистанційних курсів, є те, що курс стає відправною

© Г.В. Красовська

точкою надання інформації за відповідною темою, причому ця інформація може розміщуватися в різних місцях, а сторінки курсу виступають як своєрідний агрегатор, який може змінюватися протягом навчання; матеріали курсу органічно пов'язані один з одним, але можуть мати інші зв'язки поза курсом; кожний учасник використовує інформацію відповідно до власних навчальних цілей» [7]. Тому робота куратора контенту особливо «важлива при організації масових відкритих он-лайн курсів, які повинні базуватися на новітній інформації, яка ще не пройшла етап узагальнення. Саме тут важливу роль відіграє якість інформації та куратор, який робить процес обробки її прозорим і ясним» [3].

Робота тьютора МВОК, як куратора контенту, повинна сприяти [3]:

– уточненню і проясненню «інформаційного безладу» та позбавлення інформаційного шуму;

– забезпеченню цінності наданої інформації;

– зміцненню довіри (надаючи своїй аудиторії якісний контент, куратор змісту формує довіру).

Крім того, курирування контету повинно забезпечувати слухачів курсу тільки ретельно відібраними частинами змісту, з додаванням відповідних коментарів чи розуміння, або виділенням окремих частин, і завжди із зазначенням творця контенту без елементів плагіату, що передбачає вказівку в роботі посилання на першоджерело, автора інформації, використання реблога, ретвіта тощо [3].

Також слід брати до уваги, що слухачами одного курсу МВОК можуть бути десятки тисяч людей, що унеможливлює реалізацію стійкого зворотного зв'язку з тьютором під час засвоєння лекційного матеріалу та виконання домашніх завдань, а також перевірки їх правильності. Ці функції покладені на спільноту самих студентів, що реалізуються через форум запитань та відповідей, взаємоперевірку та оцінювання домашніх робіт [8]. Саме співпраця, що втілює принципи нової теорії навчання конективізму, згідно якої навчання – це процес створення мережі, вузлами якої є зовнішні сутності (люди, організації, бібліотеки, сайти, книги або будь-яке інше джерело інформації) [9], є ще одною характерною рисою МВОК, яка відрізняє їх від традиційних академічних дистанційних курсів.

МВОК триває від 6 до 10 тижнів, але навіть після завершення всі матеріали курсу є доступними в Інтернет, а спільноти студентів (спільноти практики [2]) продовжують існування. Іншими словами, в МВОК під час навчання застосовується технологія «краудсорсінга», внаслідок чого утворюється «великий обсяг неструктурованої інформації. Виникає дуже щільна хмара знань, яка перевищує за своєю щільністю результат пошуку Гугл або Вікіпедії. Ці знання є сильно фрагментованими, їх складно зберігати, особливо після закінчення курсу» [7].

У роботах [2; 7; 9; 10] зазначається, що для успішного використання МВОК необхідно вирішувати такі питання:

– як обирати відкритий контент, перевіряти та оцінювати його?

– як допомогти студентам організувати діяльність з відкритим контентом?

– які технології необхідні для того щоби управляти інформаційним перенавантаженням?

Мета статті

Метою даної роботи є аналіз відомих підходів до побудови програмної системи задля реалізації ефективного пошуку, відбору, перевірки на адекватність та категоризації слабоструктурованого масиву інформації, що утворюється в Інтернет під час проведення МВОК.

Виклад основного матеріалу

Глобальна концепція організації ефективного пошуку, аналізу та розуміння величезних обсягів інформації в Інтернет реалізується в технології Semantic WEB, де обробкою, обміном інформацією та видобутком знань займаються інтелектуальні програмні агенти, які використовують онтології різних рівнів для визначення змісту термінів використаних при описі Інтернет-ресурсів, отримання логічних виведень, фактів, які не представлені в онтології буквально, але випливають з її семантики. Для створення зрозумілого програмним агентам опису ресурсу в Semantic WEB використовується формат RDF, що заснований на синтаксисі XML, та мова опису онтологій OWL [11].

На основі XML і RDF був розроблений формат RSS, призначений для легкого і швидкого обміну контентом між сайтами. Його застосування забезпечує узгоджений спосіб резюмування контенту сайтів, що зараз використовується на сайтах новин [12]. Обробку та аналіз даних в форматі RSS виконують спеціальні програми RSS-агрегатори (парсери), що дозволяють групувати публікації з різних джерел, автоматично оновлювати списки актуальних інформаційних повідомлень (стрічки новин).

Технології, що ґрунтуються на RSS, також активно використовується під час проведення МВОК [2; 7].

В інформаційному центрі компанії "ЕЛВІСТІ" (м. Київ, http://visti.net/) розроблено систему контент-моніторингу InfoStream, яка вирішує завдання автоматизованого збирання новин з web-сайтів, їх опрацювання та систематизації. Системою

InfoStream використовується єдине програмне забезпечення – сканер, що проводить моніторинг Інтернет-ресурсів і корпоративних Інтернет-мереж. Збирання даних виконується шляхом обходу мережевих ресурсів відповідно до інструкцій на метамові системи, які створюються експертами [13].

У роботі [14] описується алгоритмічна модель асоціативно-семантичного аналізу текстів природною мовою, реалізовано асоціативний аналіз контексту в семантичному розборі речень тексту, запропонована технологія асоціативно-семантичної обробки текстів з використанням лінгвістичної бази знань та алгоритмічних блоків, які реалізують різні фази лінгвістичного аналізу та різні елементи смислової обробки структур тексту. Результати даного дослідження втілені в системах "Рефератор", "VitaminE", системі семантичної фільтрації текстів, системі фільтрації Internet-повідомлень з використанням лінгвістичних методів аналізу текстів, системі тематичної рубрикації та кластеризації текстів природною мовою.

У [15] наведено результати теоретичних досліджень та експериментальна перевірка методів моделювання інтелектуальної діяльності людини з використанням інтрофізичного підходу. Інтелектуальні системи, що розробляються в рамках інтрофізичного підходу, отримали назву рефлекторних.

Було запропоновано рефлекторний алгоритм обробки природно-мовних текстів і створено компілятор природно-мовних текстів, який забезпечує формування результуючого формалізованого подання семантичної складової вхідного тексту. Експериментальним шляхом доведено ефективність запропонованого підходу при вирішенні природно-мовного доступу до баз даних [15], довільних текстів будівельної тематики для створення ресурсної нормативної бази [16].

Висновки

В результаті проведеного дослідження можна зробити висновок, що під час побудови МВОК використовуються тільки RSS-канали та агрегатори для збираня «новин» в роботі спільноти студентів, але цих засобів недостатньо для вирішення всього спектру задач курирування контентом. Автор вважає перспективним проводити подальші дослідження в напрямку створення рефлекторної інтелектуальної програмної системи, яка проводить для заданої предметній області:

– пошук та відбір в мережі Інтернет інформації з використанням інтелектуальних Інтернет-агентів;

– аналіз та категоризацію інформації, яка додається до наявної бази із застосуванням інтроформаційних методів теорії несилової взаємодії;

– вдосконалення методів пошуку та аналізу інформації в процесі самонавчання.

Стартом для розвитку бази знань такої системи є організація МВОК, під час створення якого тьютор (експерт) формує концептуальну модель знань (онтологію) даної предметної області. Розроблена модель вдосконалюється та насичується під час проведення МВОК спільнотою студентів (краудсорсінг) під контролем тьютора.

Побудована за такими принципами рефлекторна інтелектуальна програмна система («процесор знань») дозволить надалі проводити оперативний відбір, автоматичний аналіз інформації про наукові та технічні новації в даній предметній області, що публікується на Інтернет-ресурсах, та збагачення результатами, що отримані під час аналізу наявної бази знань. Це може бути корисним під час проведення наукових досліджень, для розвитку програм навчальних курсів, які мають враховувати найсучасніші тенденції в галузі, розробки вимог до знань та вмінь галузевих стандартів освіти тощо.

Список літератури

1. *Освіта протягом життя: світовий досвід і українська практика. Аналітична записка // [Електронний ресурс]. Режим доступу: http://www.niss.gov.ua/articles/252/*
2. *Кухаренко В.М. Навчальний процес у масовому відкритому дистанційному курсі // Теорія і практика управління соціальними системами. – 2012. — № 1. – с.40-50.*
3. *Кухаренко В. М. Методи роботи куратора змісту / В. М. Кухаренко // Інформаційні технології в освіті. – 2013. – Вип. 16. – С. 100-107. – Режим доступу: http://nbuv.gov.ua/UJRN/itvo_2013_16_14*
4. *Chaney Paul. Content Curation Fundamentals, Part 1 of a Series «Blogger Network Електронний ресурс – Режим доступу: http://blog.utalkmarketing.com/content-strategy/content-curationfundamentals-part-1-of-a-series/*
5. *Kelly David. Is Content Curation in Your Skill Set ? It Should Be.: Learning Solutions Magazine Електронний ресурс – Режим доступу: http://bit.ly/V2ImLl*
6. *Farber Eugene. Content Curation : The Ultimate Guide Електронний ресурс – Режим доступу: http://www.contentstrategyhub.com/content-curation-guide*
7. *Бугайчук К. Л. Масові відкриті дистанційні курс. / Бугайчук К. Л.. // Інформаційні технології і засоби навчання. – 2011. – №6(26). // [Електронний ресурс]. Режим доступу: http://www.journal.iitta.gov.ua*
8. *Koller Daphne: What we're learning from online education. / Daphne Koller // TEDGlobal 2012 20:40 Filmed Jun 2012 [Електронний ресурс]. – Режим доступу: https://www.ted.com/ talks/daphne_koller_what_we_re_lear-ning_from_ online_education*

9. Siemens George. Researching open online courses. / Siemens G. // [Електронний ресурс]. – Режим доступу: http://www.elearnspace.org/blog/2011/07/04/researching-open-online-courses/

10. Vicki D. Questioning the Future of the Open Student / Davis Vicki. // EDUCAUSE Review, vol. 45, no. 4 – 2010. – pp. 22–28.

11. Андон Ф.И., Гришанова И.Ю., Резниченко В.А. Semantic web как новая модель информационного пространства интернет // Пробл. програмув. — 2008. — N 2-3. — с. 417-430.

12. Ландэ Д.В. Сканер системы контент-мониторинга InfoStream / Д.В. Ландэ // Открытые информационные и компьютерные интегрированные технологии. – Харьков: НАКУ "ХАИ", 2005. – Вып. 28. – С. 53 – 58

13. Ландэ Д.В. Редкостный Синтез Сайтов / Ландэ Д.В., Морозов А.Ю.. // "Мой Компьютер". – 2003. – №25 // [Електронний ресурс]. Режим доступу: http://dwl.visti.net/art/rssart/

14. Анісімов А.В., Марченко О.О., Никоненко А.О. Алгоритмічна модель асоціативно-семантичного контекстного аналізу текстів природною мовою // Пробл. програмув. — 2008. — N 2-3. — С. 379-384.

15. Тесля Ю.Н. Введение в информатику природы /Ю.Н. Тесля: монография. – К.: Кондор, 2010. – 256 с.

16. Тесля Ю. М., Шабала Є. Є. Рефлекторна система формалізації змісту довільних текстів будівельної тематики / Ю. М. Тесля, Є. Є. Шабала // Управління розвитком складних систем. – 2013. – Вип. 16. – С. 119-123.

Стаття надійшла до редколегії 23.10.2017

Рецензент: д-р техн. наук, проф. В.І. Кудін, Київський національний університет імені Тараса Шевченка, Київ.

Красовская Анна Валерьевна
Кандидат технических наук, доцент, доцент кафедры интеллектуальных информационных систем,
orcid.org 0000-0003-1986-6130
Киевский национальный университет им. Тараса Шевченко, Киев

АНАЛИЗ ПОДХОДОВ К ПОСТРОЕНИЮ МУЛЬТИАГЕНТНОЙ ИНТЕЛЛЕКТУАЛЬНОЙ СИСТЕМЫ КУРИРОВАНИЯ КОНТЕНТОМ МООК

Аннотация. *Проанализированы особенности организации МООК, определен круг вопросов, которые необходимо решить для повышения эффективности их организации и проведения. Отмечено, что при построении МООК для сбора новостей в работе сообщества студентов используются только RSS-каналы и агрегаторы, но этих средств недостаточно для решения всего спектра задач курирования контентом. Предлагается проводить дальнейшие исследования в направлении создания многоагентной рефлекторной интеллектуальной программной системы, которая проводит для заданной предметной области поиск и отбор информации в сети Интернет, анализ и категоризацию информации с применением интроформационных методов теории несилового взаимодействия.*

Ключевые слова: массовые открытые онлайн-курсы; курирование контентом; интеллектуальный интернет-агент; многоагентная рефлекторная интеллектуальная система; интроформация; теория несилового взаимодействия

Krasovska Hanna
Candidate of technical sciences, associate professor, assistant professor of the department of intellectual and informational systems,
orcid.org/0000-0003- 1986-6130
Taras Shevchenko National University of Kyiv, Kyiv

ANALYSIS OF APPROACHES TO BUILDING THE MULTI-AGENT INTELLECTUAL CONTENT CURATION SYSTEM FOR MOOC

Abstract. *The purpose of this paper is to analyze the existing approaches to building a multi-agent intellectual system for implementing effective search, selection, validation and categorization of weakly structured information sets that are generated on the Internet during the mass open online courses (MOOC). The features in organization of MOOC were analyzed, the matters that needed to be resolved for improvement of efficiency of the organization and conduction of the courses was determined. It is noted that when constructing an MOOC to collect news in the work of the student community, only RSS-channels and aggregators are used, but these tools are not sufficient for the solution of the whole range of problems with content curation. The author considers it promising to conduct further research in the direction of creating a multi-agent reflex intellectual software system that searches for and selects for a given subject area information from the Internet, analysis and categorization of information using introformation methods of the theory of non-coercive interaction.*

Keywords: massive open online courses; content curation; intelligent Internet agents; multi-agent reflex intelligent system, intrormation, the theory non-coercive interaction.

References

1. Osvita protiahom zhyttia: svitovyj dosvid i ukrains'ka praktyka. Analitychna zapyska. Electronic source – Retrieved from http://www.niss.gov.ua/articles/252/

2. Kukharenko, V.M. (2012). Initial process in mass open distance cource. Theory and practice of social systems management, 1, 40-50.

3. Kukharenko, V.M. (2013). Methods of content manager work. Informatsijni tekhnolohii v osviti, 6, 100-107. – Retrieved from http://nbuv.gov.ua/UJRN/itvo_2013_16_14

4. Chaney, Paul. Content Curation Fundamentals, Part 1 of a Series "Blogger Network". Electronic source – Retrieved from http://blog.utalkmarketing.com/content-strategy/content-curationfundamentals-part-1-of-a-series/

5. Kelly, David. Is Content Curation in Your Skill Set ? It Should Be.: Learning Solutions Magazine. Electronic source – Retrieved from http://bit.ly/V2ImLl

6. Farber, Eugene. Content Curation : The Ultimate Guide. Electronic source – Retrieved from http://www.contentstrategyhub.com/content-curation-guide

7. Buhajchuk, K.L. (2011). Mass open distance cource. Information systems and education tools, 6(26). Electronic source – Retrieved from http://www.journal.iitta.gov.ua

8. Koller, Daphne: What we're learning from online education. / Daphne Koller // TEDGlobal 2012 20:40 Filmed Jun 2012 Electronic source – Retrieved from https://www.ted.com/ talks/daphne_koller_what_we_re_lear-ning_from_ online_education

9. Siemens, George. (2011) Researching open online course. Electronic source – Retrieved from http://www.elearnspace.org/blog/2011/07/04/researching-open-online-courses/

10. Vicki, D.(2010). Questioning the Future of the Open Student. EDUCAUSE Review, vol. 45, 4, 22–28.

11. Andon, F.Y. & Hryshanova, Y.Yu. & Reznychenko V.A. (2008). Semantic web as new model of information space internet. Problems of programming, 2-3, 417-430.

12. Lande, D.V. (2005). Scaner of the system of content-monitoring InfoStream. Open information amd computer integrated technologies, Khar'kov, Ukrayna: NAKU "KhAY", 28, 53–58

13. Lande, D.V. & Morozov, A.Yu. (2003). Rare synthesis of sites. "Moj Komp'iuter",25 [Electronic source]. – Retrieved from http://dwl.visti.net/art/rssart/

14. Anisimov, A.V. & Marchenko, O.O. & Nykonenko, A.O. (2008). Algometric model of associated semantic contecst analisys of the textes in natural language. Problems of programation, 2-3, 379-384.

15. Teslya, Y.M. (2010). Introduction to informatics of nature. Kyiv, Ukraine : Maklaut.

16. Teslia, Yu.M. & Shabala, Ye.Ye. (2013). Reflector system of content formalization of ree textes in building topic. Management of development of complex systems, 16, 119-123.

Посилання на публікацію

APA *Krasovska, Hanna (2017). Analysis of approaches to building the multi-agent intellectual content curation system for MOOC. Managment of Development of Complex Systems, 32, 81 – 85. [in Ukrainian]*

ДСТУ *Красовська Г.В. Аналіз підходів до побудови мультиагентної інтелектуальної системи курирування контентом МВОК [Текст] / Г.В. Красовська // Управління розвитком складних систем. – 2017. – № 32. С. 81 – 85.*