
КРАТКИЕ СООБЩЕНИЯ

УДК 004.021, 004.047, 004.622
DOI: 10.17586/0021-3454-2015-58-12-1022-1026

АВТОМАТИЗИРОВАННОЕ ФОРМИРОВАНИЕ ПРОФИЛЕЙ ОБЪЕКТОВ НАУЧНО-ОБРАЗОВАТЕЛЬНОЙ ДЕЯТЕЛЬНОСТИ В ИНФОРМАЦИОННОЙ СИСТЕМЕ УПРАВЛЕНИЯ УНИВЕРСИТЕТА

Д. А. ВАРЕНИКОВ, Д. И. МУРОМЦЕВ, М. Д. ШЛЕЙ,
М. Н. ЕФИМОВ, Н. А. ЖУКОВА

*Университет ИТМО, 197101, Санкт-Петербург, Россия
E-mail: varenikovda@gmail.com*

Представлены подходы к автоматизированному формированию профилей объектов научно-образовательной деятельности в информационной системе университета. Рассмотренные подходы позволяют повысить качество выборки, на основе которой строится система рекомендаций для пользователей информационной системы по подбору научного руководителя или обучающегося, научного мероприятия, грантов, публикаций и периодических изданий для опубликования научных результатов.

Ключевые слова: *научные интересы, научная деятельность, обучающиеся, наукометрические базы данных.*

Введение. Эффективность реализации процессов научной деятельности определяется качеством решения следующих задач: поиск и обеспечение доступа к научно-исследовательским работам, выполняемым по схожим тематикам, своевременное информирование о проведении научных мероприятий и планирование участия в них, обеспечение возможности публикации полученных научных результатов в высокорейтинговых изданиях.

Постановка задачи. Для анализа процесса реализации научной деятельности обучающихся (магистры и аспиранты) и преподавателей вуза была построена информационная модель верхнего уровня, представленная на рисунке. Эта модель описывает процессы взаимодействия участников и результаты их научной деятельности в информационной системе университета.

Анализ модели и предметной области позволил выделить основные проблемные области, связанные с принятием решений и поиском соответствующей информации:

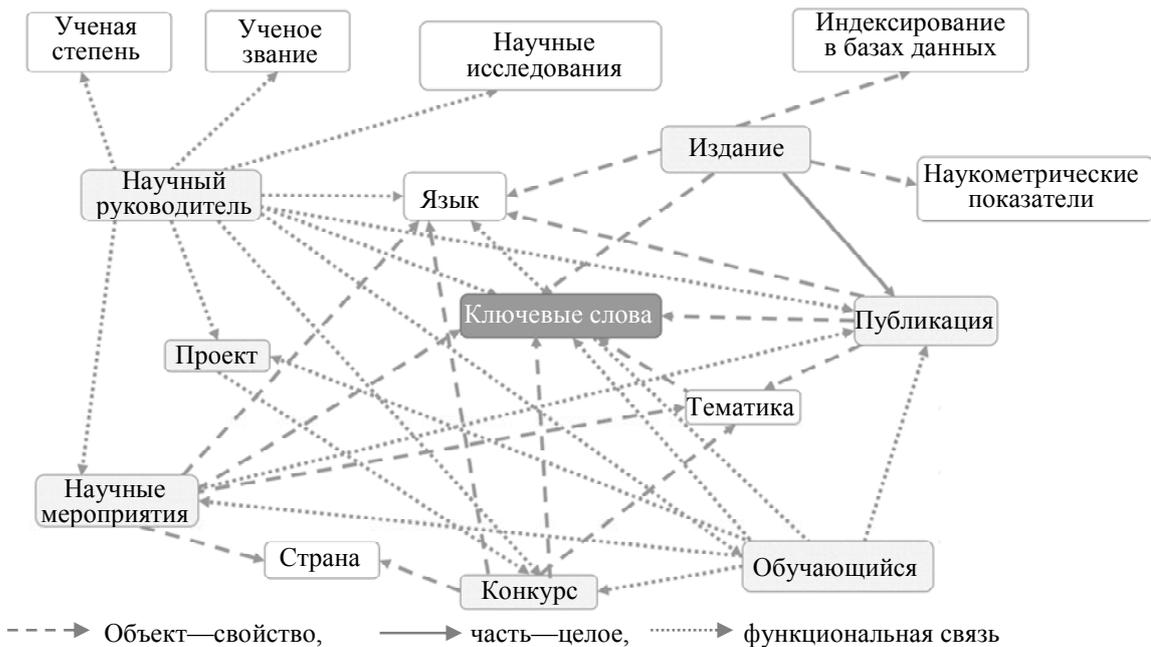
- выбор научного руководителя или обучающихся с учетом научных интересов;
- подбор периодических изданий для публикации научных результатов;
- выбор научных мероприятий для участия;
- выбор конкурсов, грантов для финансирования исследований. Проектные менеджеры университета отслеживают информацию об открываемых конкурсах и фондах и отбирают потенциально интересные для преподавателей и учащихся [1];
- поиск значимых публикаций по заданной тематике.

В построенной модели можно выделить следующие ключевые объекты:

- научные руководители и учащиеся – пользователи информационной системы;
- научные мероприятия – конференции, конгрессы, семинары, круглые столы и пр.;

- конкурсы, ориентированные на материальную поддержку преподавателей и учащихся;
- публикации в периодических изданиях;
- издания (периодические).

Как видно из представленной модели, основные информационные объекты, рассматриваемые в настоящей статье, связаны через ключевые слова, которые формируют профиль объекта или так называемую „область научных интересов“. Методы частотного анализа и информационные технологии позволяют выполнять анализ связей между научными интересами участников, проводимыми исследованиями, актуальными конкурсами и мероприятиями. По результатам анализа могут быть сформированы рекомендации для решения поставленных задач.



Использование предложенного подхода позволяет сформировать научный профиль пользователя информационной системы, от полноты информации о профиле зависит точность формируемых рекомендаций, а следовательно — эффективность развития научной деятельности вуза. Решение этой задачи имеет свои особенности, так как при формировании профиля используется множество источников информации, и при анализе информации необходимо правильно определять приоритеты тех или иных научных интересов.

Формирование научного профиля пользователя в информационной системе университета происходит за счет самостоятельного ввода информации при заполнении личного профиля в информационной системе, а также автоматического сбора сведений по ключевым словам. Автоматический сбор сведений основан на формализации и последующей интеграции информации из наукометрических баз данных [2], а также анализе поведения пользователей в информационной системе, его научно-практических результатов с использованием методов частотного анализа и алгоритмов нечеткого поиска. Под наукометрическими базами данных понимают библиографические и реферативные базы данных, а также инструмент для отслеживания цитируемости научных статей [3].

Автоматическое наполнение профилей ключевыми словами позволяет значительно расширить выборку для дальнейшего предоставления рекомендаций по поставленным задачам. Научные интересы пользователя представлены множеством ключевых слов K :

$$K = K_p \cup K_a,$$

где K_p — множество ключевых слов, указанных пользователем, а K_a — множество ключевых слов, автоматически выбранных с учетом частоты их появления;

$$K_a = \{q \in Q : f(q) > c\},$$

Q — множество автоматически полученных ключевых слов, $f(q)$ — частота появления ключевого слова q , c — пороговое значение частоты появления ключевого слова;

$$Q = \bigcup_{i=1}^h Q_i,$$

h — число источников, на основании которых формируется Q за счет:

1) посещения пользователем информационной системы. Q_1 — множество ключевых слов, полученных по результатам посещения страниц, содержащих ключевые слова и тематики:

$$Q_1 = \bigcup_{j=1}^m L_j,$$

L_j — множество ключевых слов j -й страницы, m — число страниц (в статье рассматриваются страницы информационной системы, которые посещает пользователь);

2) анализа схожих интересов пользователей, посетивших одинаковые страницы. Q_2 — множество ключевых слов пользователей со схожими интересами [4, 5]

$$Q_2 = \left\{ K_v : v \in \bigcup_{j=1}^m V_j, \text{Jaccard}(K_v, K_p) > s \right\},$$

K_v — множество ключевых слов пользователя v , V_j — множество пользователей, посетивших j -ю страницу, за исключением рассматриваемого, $\text{Jaccard}(K_v, K_p) = \frac{|K_v \cap K_p|}{|K_v \cup K_p|}$ — мера

Жаккара [6] ($0 \leq \text{Jaccard}(K_v, K_p) \leq 1$), s — пороговое значение схожести;

3) анализа тематики публикаций пользователя. Q_3 — множество ключевых слов, полученных на основе анализа публикационной активности пользователя

$$Q_3 = \bigcup_{j=1}^g K_{l_j},$$

K_{l_j} — множество ключевых слов j -й публикации, g — число публикаций пользователя;

4) анализа схожести интересов соавторов публикаций пользователя. Q_4 — множество ключевых слов соавторов публикаций пользователя

$$Q_4 = \left\{ K_{Ca} : Ca \in \bigcup_{j=1}^g Ca_j, \text{Jaccard}(K_{Ca}, K_p) > s \right\},$$

K_{Ca} — множество ключевых слов заданного пользователя; Ca_j — множество соавторов публикации j , за исключением рассматриваемого пользователя;

5) получения сведений о подписке пользователя на рассылку в информационной системе. Q_5 — множество ключевых слов, указанных пользователям для получения рассылки;

6) анализа профиля пользователя в наукометрических базах данных. Q_6 — множество ключевых слов пользователя, полученных из баз

$$Q_6 = \left\{ K_{Apr} : Apr \in \bigcup_{j=1}^d Apr_j \right\},$$

K_{Apr} — множество ключевых слов пользовательского профиля Apr , Apr_j — множество авторских профилей наукометрической базы данных [7] j для рассматриваемого пользователя, d — количество наукометрических баз данных.

Заключение. В настоящей работе предложен подход к автоматизации процесса формирования научных профилей, который позволил значительно расширить выборку, на основе которой строятся рекомендации пользователям информационной системы по выбору научного руководителя или обучающегося, научного мероприятия, грантов, публикаций и периодических изданий для публикации научных результатов. Полнота полученных данных позволила оптимизировать учет публикаций специалистами и как следствие — повысить качество отчетных данных. Предложенный подход реализован в информационной системе управления университета.

СПИСОК ЛИТЕРАТУРЫ

1. Попова И. А., Громов Г. Ю. Подходы к созданию эффективной информационной системы управления университетом // Сб. матер. XX Всерос. науч.-метод. конф. „Телематика—2013“. СПб, 2013.
2. Наукометрические базы данных [Электронный ресурс]: <<http://pspu.ru/university/biblioteka/prepodavatelam/indeksy-nauchnogo-citirovaniya>>.
3. Коляда А. С., Гогунский В. Д. Автоматизация извлечения информации из наукометрических баз данных // Управління розвитком складних систем. 2013. № 16. С. 96—99.
4. Ефимов М. Н., Шлей М. Д., Вареников Д. А. Метод определения рекомендаций для пользователей информационной системы на основе их научных интересов и активности // Матер. VIII Междунар. науч.-практ. конф. „Научно-образовательная информационная среда XXI века“. Петрозаводск, 2014. С. 74—77.
5. Ефимов М.Н., Шлей М. Д., Вареников Д. А. Система определения научных интересов пользователей // Тр. XXI Всерос. науч.-метод. конф. „Телематика'2014“. 2014. С. 87—88.
6. Kim J.-O., Mueller W. Factor Analysis: Statistical Methods and Practical Issues. SAGE Publications, 1978. 88 p.
7. Вареников Д. А., Муромцев Д. И., Шлей М. Д. Подходы автоматизации обработки данных наукометрических баз данных // Компьютерные инструменты в образовании. 2015. № 2. С. 3—13.

Сведения об авторах

- Денис Анатольевич Вареников** — Университет ИТМО; департамент информационных технологий, отдел информационной поддержки проектной деятельности; зам. начальника отдела; E-mail: varenikovda@gmail.com
- Дмитрий Ильич Муромцев** — канд. техн. наук, доцент; Университет ИТМО, международная лаборатория „Интеллектуальные методы обработки информации и семантические технологии“; руководитель международной лаборатории; E-mail: mouromtsev@mail.ifmo.ru
- Михаил Дмитриевич Шлей** — канд. техн. наук; Университет ИТМО; департамент информационных технологий, отдел информационной поддержки проектной деятельности; начальник отдела; E-mail: mikhail.shlei@gmail.com
- Максим Николаевич Ефимов** — Университет ИТМО; департамент информационных технологий, отдел информационной поддержки проектной деятельности; программист
- Наталья Александровна Жукова** — канд. техн. наук; Университет ИТМО, международная лаборатория „Интеллектуальные методы обработки информации и семантические технологии“; научный сотрудник

Рекомендована кафедрой информатики и прикладной математики-1

Поступила в редакцию 01.11.15 г.

Ссылка для цитирования: Вареников Д. А., Муромцев Д. И., Шлей М. Д., Ефимов М. Н., Жукова Н. А. Автоматизированное формирование профилей объектов научно-образовательной деятельности в информационной системе управления университета // Изв. вузов. Приборостроение. 2015. Т. 58, № 12. С. 1022—1026.

**AUTOMATED PROFILE FORMATION FOR OBJECT OF RESEARCH AND EDUCATIONAL ACTIVITY
IN THE INFORMATION SYSTEM OF THE UNIVERSITY MANAGEMENT**

**D. A. Varenikov, D. I. Mouromtsev, M. D. Shley,
M. N. Efimov, N. A. Zhukova**

ITMO University, 197101, St. Petersburg, Russia

E-mail: varenikovda@gmail.com

Approaches to automated profile formation for object of research and educational activity in the information system of the University are presented. The approaches under consideration allows to increase the quality of the sampling used as a base for recommendations system for users of the information system for selection of research manager or a student, research activity, grants, papers, and periodicals for publication of scientific results.

Keywords: scientific interest, research activity, students, scientometric databases.

Data on authors

- Denis A. Varenikov** — ITMO University; Department of Information Technologies, Division of Information Support of Project Management; Deputy Head of the Division; E-mail: varenikovda@gmail.com
- Dmitriy I. Mouromtsev** — PhD, Associate Professor; ITMO University, International Laboratory of Intellectual Methods of Information Processing and Semantic Technologies; Head of the Laboratory; E-mail: mouromtsev@mail.ifmo.ru
- Mikhail D. Shley** — PhD; ITMO University; Department of Information Technologies, Division of Information Support of Project Management; Head of the Division; E-mail: mikhail.shlei@gmail.com
- Maksim N. Efimov** — ITMO University; Department of Information Technologies, Division of Information Support of Project Management; Programmer
- Nataliya A. Zhukova** — PhD; ITMO University, International Laboratory of Intellectual Methods of Information Processing and Semantic Technologies; Scientist

For citation: *Varenikov D. A., Mouromtsev D. I., Shley M. D., Efimov M. N., Zhukova N. A.* Automated profile formation for object of research and educational activity in the information system of the University management // *Izvestiya Vysshikh Uchebnykh Zavedeniy. Priborostroenie.* 2015. Vol. 58, N 12. P. 1022—1026 (in Russian).

DOI: 10.17586/0021-3454-2015-58-12-1022-1026