

СОВРЕМЕННЫЙ КЛАССИФИКАТОР ДЛЯ ПРЕДМЕТНОЙ ОБЛАСТИ

О.Б. ПОПОВА

*Кубанский государственный технологический университет,
350002, Российская Федерация, Краснодар, ул. Московская, 2,
электронная почта: popova_ob@mail.ru*

Сегодня актуальной задачей является развитие системы научных коммуникаций через создание эффективных информационных систем, которые осуществляют наукометрические расчёты разного уровня, помогают осуществлять поиск научной информации из разных источников, организуют удобное и эффективное хранение научно-технической информации и системы ссылок на актуальные публикации. Среди них можно встретить современные базы научного цитирования РИНЦ, Scopus, Web of Science. Для удобного хранения информации, осуществления поиска и отражения дисциплинарных и междисциплинарных особенностей предметной области в таких системах используются классификаторы разного уровня. Новые тенденции в этой области - объединение в классификаторе нескольких функций. Сейчас в электронных журналах с высоким рейтингом используют классификаторы, которые объединяют сортировку материала по тематике, удобный поиск статей, распределение по тематической компетенции рецензентов. Полной картины перспективы развития определённой предметной области такие классификаторы не определяют, только определённые векторы развития. Для выявления всей картины необходимо прочитать множество статей по предметной области. Автор предлагает разработать методику получения современного классификатора предметной области, по внешнему виду которого можно получить представление о самой предметной области и о наукометрических показателях.

Ключевые слова: современный классификатор, предметная область, множество альтернатив, метод классификации, бинарное дерево, ключевые слова, поиск, научное исследование.

Сегодня приходится работать с большими объёмами информации. Решается огромное число прикладных задач и конкретную предметную область становится всё труднее описать. Для сортировки полученных знаний внутри предметной области используются различные классификаторы, которые могут указать лишь конкретные перспективные направления или некоторый симбиоз из них. У них есть ряд общих недостатков:

- не учитывают все направления;
- имеют два три уровня;
- слишком общие;

- не могут быть использованы как элемент управления для поиска данных;
- не подходят для точного наукометрического расчёта.

В некоторых классификаторах предполагается выбор очень общих кодов классификатора, то есть можно выбрать самый верхний уровень, что усложняет способ соотнесения знания к определенной группе. Такой классификатор используется в eLIBRARY при выборе «Раздела тематического рубрикатора». Обычно такие классификаторы используются совместно с полем «Ключевые слова». Конечно, это улучшает поиск, так как грубый поиск завершается более точным, но вместе эти две системы поиска не дают информацию о предметной области. Они представляют собой множество векторов (научных направлений) и огромное множество точек (разные варианты и комбинации слов). Всё это трудно визуализировать самостоятельно и представить в виде графиков. Если возникнет потребность самостоятельного изучения предметной области, то придётся прочитать огромное число статей, понять суть, структурировать и представить предметную область, что при современных объёмах информации очень трудно сделать. Так же трудно осуществить и наукометрические расчёты, составить карты науки и определить где каким научным направлением в основном занимаются. Вот почему многие международные наукометрические базы предоставляют платные услуги для проведения таких расчётов для журналов (Scopus, Web of Science), которые должны так же удовлетворять определённым условиям. Такие условия позволяют получить более менее вменяемые результаты при расчёте, то есть нужно журналу иметь хорошие аннотации, интерес для международного научного сообщества, списки литературы на уровне, широкий охват рецензентов журнала и его редакторов и так далее. Чем «хуже» журнал, тем меньше «точность» расчёта. Честнее было бы с точки зрения поиска и расчётов сделать тематические наукометрические базы с более эффективными классификаторами, которые совмещены с ключевыми словами для поиска, где использованы «плохие» и «хорошие» журналы.

Важность предлагаемого исследования по указанной проблеме заключена в том, что необходимо сформировать новые подходы в получении более точных и полных классификаторов предметных областей, получить новые структуры в организации поисковых запросов по ключевым словам, получить новые способы хранения данных с участием современного классификатора. Всё это позволит развить существующие направления по расчёту наукометрических показателей журналов, конкретных учёных, коллективов, научных школ, научных направлений и так далее. Такое исследование позволит расширить возможности практического применения научных результатов по теме гранта 16-03-00382 от 18.02.2016 "Мониторинг исследовательской деятельности образовательных учреждений в условиях информационного общества".

Основные направления исследований в отечественной и мировой науке – это создание готовых классификаторов разного уровня – международные, межгосударственные, национальные или межотраслевые, отраслевые, региональные, системные, которые позволяют автоматизировать процесс управления и обработки информации при решении задач различного уровня при создании эффективного документооборота. В настоящее время в «России Общероссийский классификатор — классификатор, принятый Росстандартом и обязательный для применения при межотраслевом обмене информацией и в общероссийских унифицированных формах документов (УФД). Общероссийский классификатор не должен противоречить соответствующим международным классификаторам. Общероссийские классификаторы должны соответствовать требованиям Единой системы классификации и кодирования технико-экономической и социальной информации РФ (ЕСКК) ТЭСИ. Национальные и отраслевые классификаторы являются нормативными документами, внесение изменений в которые должны согласовываться с Госкомстатом и Комитетом РФ по стандартизации, метрологии и сертификации» [1].

«Перечень общероссийских классификаторов содержится в Общероссийском классификаторе информации об общероссийских

классификаторах (ОКОК)" [2]. При составлении классификаторов используют один из трёх методов:

- иерархический метод классификации;
- фасетный метод классификации;
- дескрипторный метод классификации.

Иерархический метод, по мнению коллектива, является наиболее перспективным методом классификации, который может наиболее ёмко и точно описать предметную область. К сожалению, такой метод в основном используют в системных классификаторах, так как на момент разработки основных международных и национальных классификаторов способы представления иерархических структур не позволяли сделать это на нужном уровне. Теперь же это можно сделать более эффективно, устранив недостатки метода. Но требуется также внести изменение в ряд концепций этого метода. Например, при «построении классификации выбор последовательности признаков зависит от вероятности обращения к тому или иному признаку», при этом «наиболее вероятным обращениям должны соответствовать высшие уровни классификации» [1]. Мы же предлагаем убрать данное условие, так как оно не позволяет создавать стационарных деревьев поиска с разным уровнем вложений, с удобным механизмом добавления элементов. Так как это условие позволяет быстро найти актуальную информацию на момент поиска (что широко используется сейчас в современных поисковых системах, например, Google), но не позволяет описать очень подробно предметную область целиком (с течением времени она разрастается, но структура не переформируется).

Фасетный метод классификации сейчас широко применяется для создания классификаторов, которые позволяют выбрать одно из научных направлений в определённой области, например, классификаторы современных журналов, классификаторы перспективных направлений для различных фондов и так далее. «Недостатками фасетного метода классификации являются неполное использование ёмкости, нетрадиционность и иногда сложность применения» [1]. Поэтому такие классификаторы позволяют лишь разделить

документы на большие тематические группы, но не производить эффективный поиск среди всех документов.

Дескрипторный метод классификации сейчас широко применяется для организации поиска информации в информационных системах (в электронных журналах разного уровня, в наукометрических базах данных разного уровня), который основан на пересекающемся последовательном поиске по всей базе данных информационной системы. Этот поиск самый простой в реализации и самый неэффективный. Его эффективность ещё больше снижается, когда параллельно требуется организовать добавление элементов с их последующей сортировкой и использование наукометрических данных после поиска в расчётах.

Сейчас существуют и появляются новые методы поиска, сортировки и анализа данных [3-14], которые более эффективны и которые теперь можно реализовать в таких крупных информационных системах. Тем более, что задача поиска теперь не самая важная в них, но теперь важны и статистическая обработка информации, и автоматизированный анализ предметных областей для планирования и прогнозирования развития того или иного направления науки.

При анализе готовых классификаторов [2, 15] стало понятно, что там присутствуют лишь те, которые помогают автоматизировать документооборот в государственных структурах или организациях, в которых проводятся стандартные процедуры и экономические расчёты, которые можно стандартизировать. Для научных организаций такие классификаторы не сформированы. Каждая научная организация или научный журнал составляет их самостоятельно. Поэтому сейчас от точности, ёмкости и эффективности такого классификатора зависит эффективность работы информационной системы, в которой он будет использоваться. Тем более это очень важно для наукометрических баз и для электронных журналов.

Сформулируем задачу в рамках проблемы, на решение которой направлено исследование. Необходимо провести анализ существующей

современной иерархической системы (дерева поддержки принятия решений – бинарного дерева системы вопросов и ответов) и составить современный классификатор предметной области для методов оптимизации, который может быть использован в научных организациях, образовательных организациях, государственных организациях и программных продуктах. Составление методологии получения современного классификатора предметной области на примере полученного для методов оптимизации.

Научная новизна исследования заключается в том, что современный классификатор может быть использован как заранее сформированный ключ для поиска информации и для выбора информации для соответствующих наукометрических расчётов; может быть использован для одновременного добавления элементов в информационную структуру с сортировкой добавляемых в неё элементов. Предлагаемый подход к получению классификатора нов, так как он в себе совмещает сразу же несколько выполняемых функций - сортировка по заданным признакам, поиск по ключам классификатора, анализ данных на разных уровнях (статистические расчёты, визуализация предметной области целиком, прогнозирование и планирование дальнейших направлений развития в предметной области). Предлагаемая постановка проблемы оригинальна, так как сейчас в основном указанные функции (сортировка, поиск, анализ наукометрических данных) исследуются отдельно друг от друга. Достигается эффективность каждой функции без учёта их совместного влияния. Они рассматриваются как отдельные системы, тогда как они взаимозависимы, и их нужно рассматривать и исследовать как одну систему. При таком подходе эффективность системы будет значительно выше, чем у известных сегодня систем (их эффективность складывается из суммы эффективностей каждой выполняемой функции).

Комплексность исследования состоит в том, что оно практически полностью охватит все основные аспекты проблемы, так как позволит разработать новый более эффективный метод классификации предметной области, который учитывает современные разработки и современные способы

их реализации. Этот метод можно опробовать на примере актуальной предметной области – методы оптимизации. Методология получения современного классификатора, пример такого классификатора и полученные эмпирические данные можно будет использовать в научно-исследовательском и учебном процессе. Таким путем устранится проблема отсутствия более эффективного метода классификации, который удовлетворяет современному состоянию науки и техническим достижениям. Будет также устранён пробел в наличии эффективных современных классификаторов в научно-исследовательской области.

В исследовании использованы следующие методологические принципы:

- единства теории и практики (будет получена методология и сам современный классификатор);
- логического сравнения различных признаков и образов;
- визуализации информации естественным интеллектом.

Предполагается использовать комбинацию разработанных автором методов [3-14], которые будут трансформированы под решение поставленной научной задачи. Это новый метод представления дерева принятия решений [3], через структуру бинарного дерева системы вопросов и ответов [3, 4], и новый метод усиления интеллекта [3] для осуществления выбора альтернативы из большого числа известных альтернатив. Все эти методы доказали свою эффективность в решении задач (автоматизация процесса выбора [4-14]), которые до них решались вручную. Такие методы позволяют хорошо визуализировать предметную область - методы оптимизации, и соединить формирование структуры с последующим эффективным поиском в ней. Для составления методологии получения современного классификатора будут использоваться принципы и способы представления материала, которые были использованы автором для получения правил формирования бинарного дерева системы вопросов и ответов [3, 4]. Такой подход позволит проработать основные аспекты задачи исследования – получение современного

классификатора, методологии его получения и обнародование результатов исследования с дальнейшим их использованием на практике.

Ожидаемые результаты исследования позволят представить отдельно рассматриваемые процессы сортировки, поиска, анализа наукометрических данных, как единую функционирующую систему. Совмещая эти процессы вместе, можно достичь более эффективных результатов и внедрить современные разработки в предметных областях науки – работа с научными данными и анализ наукометрических данных.

Современный классификатор предметной области – методы оптимизации и методология его получения могут быть использованы:

- в программных продуктах автоматизирующих поисковые исследования;
- в электронных журналах по различной тематике, например, "Методы оптимизации";
- в тематических наукометрических базах;
- в различных фондах, поддерживающих работу учёных.

Работа выполнена при финансовой поддержке Российского гуманитарного научного фонда № 16-03-00382 от 18.02.2016 в рамках темы “Мониторинг исследовательской деятельности образовательных учреждений в условиях информационного общества”.

ЛИТЕРАТУРА

1. Классификатор. Материал из Википедии — свободной энциклопедии. 2016. <https://ru.wikipedia.org/wiki/Классификатор>.
2. Категория: Ведомственные коды и классификаторы РФ. Материал из Википедии — свободной энциклопедии. 2016. https://ru.wikipedia.org/wiki/Категория:Ведомственные_коды_и_классификаторы_РФ.
3. Popova O., Popov B., Karandey V., Evseeva M. Intelligence amplification via language of choice description as a mathematical object (binary tree of question-answer system) // Procedia – Social and Behavioral Sciences. – 2015. – V. 214. – С. 897–905; URL: <http://www.sciencedirect.com/science/article/pii/S1877042815061030>.

4. Системный анализ процесса выбора метода оптимизации информационной системы: монография / О.Б. Попова, Б.К. Попов, В.И. Ключко; ФБГОУ ВПО «Кубан. гос. технол. ун-т». – Краснодар: Издательский Дом – Юг, 2012 – 135 с.

5. Попова О.Б., Попов Б.К., Ключко В.И. Проблема сокращения времени выбора методов управления большими системами (БС) // Современные проблемы науки и образования. – 2013. – №1. – С. 163; URL: <http://www.science-education.ru/ru/article/view?id=8371>.

6. Попова О.Б. Системный подход к исследованию процесса оптимизации. Деп. в ВИНТИ №83-В2010 от 17.02.2010.

7. Попова О.Б., Попов Б.К., Ключко В.И. Анализ связей в реальной и технической системах процесса оптимизации // Международный журнал экспериментального образования. – 2013. – №10-2. – С. 405-408; URL: <http://cyberleninka.ru/article/n/analiz-svyazey-v-realnoy-i-tehnicheskoy-sistemah-protssessa-optimizatsii>

8. Попова О.Б., Попов Б.К., Ключко В.И. Анализ связей в реальной и технической системах процесса оптимизации // Международный журнал экспериментального образования. – 2013. – №10-2. – С. 405-408; URL: <http://elibrary.ru/download/85589730.pdf>

9. Попова О.Б., Попов Б.К. Интеллектуальная информационная система выбора «Оптимэль». Патент на изобретение RUS № 2564641 от 27.05.2014.

10. Попова О.Б. Структура технической системы процесса выбора метода оптимизации. Деп. в ВИНТИ №243-В2012 от 25.05.2012.

11. Попова О.Б. Системный анализ и управление процессом оптимизации. Деп. в ВИНТИ №256-В2010 от 07.05.2010.

12. Попова О.Б. Участие процесса оптимизации в развитии сложных технических систем. Деп. в ВИНТИ №257-В2010 от 07.05.2010.

13. Попова О.Б., Попов Б.К. Связи в исследуемой системе процесса оптимизации. Деп. в ВИНТИ №112-В2012 от 22.03.2012.

14. Попова О.Б., Попов Б.К. Анализ процесса оптимизации. Определение понятий. Деп. в ВИНТИ №111-В2012 от 22.03.2012.

15. Категория: Ведомственные коды и классификаторы РФ. Материал из Википедии — свободной энциклопедии. 2016. https://ru.wikipedia.org/wiki/Категория:Ведомственные_коды_и_классификаторы_РФ.

REFERENCES

1. Klassifikator. Material iz Vikipedii — svobodnoj jenciklopedii. 2016. <https://ru.wikipedia.org/wiki/Klassifikator>.

2. Категория:Vedomstvennye kody i klassifikatory RF. Material iz Vikipedii — svobodnoj jenciklopedii. 2016. https://ru.wikipedia.org/wiki/Kategorija:Vedomstvennye_kody_i_klassifikatory_RF.
3. Popova O., Popov B., Karandey V., Evseeva M. Intelligence amplification via language of choice description as a mathematical object (binary tree of question-answer system). *Procedia – Social and Behavioral Sciences*, 2015, V. 214, C. 897–905; URL: <http://www.sciencedirect.com/science/article/pii/S1877042815061030>.
4. Popova O.B., Popov B.K., Kljuchko V.I. The system analysis of the process to choose the method of optimization of the information system: monograph (The system analysis of the process to choose the method of optimization of the information system: monograph). – Krasnodar: ООО «Izdatel'skij Dom-Jug», FGBOU VPO «KubGTU», 2012. – 135 s.
5. Popova O.B., Popov B.K., Kljuchko V.I. *Sovremennye problemy nauki i obrazovaniya* (Modern problems of science and education), 2013, №1, S. 163; URL: <http://www.science-education.ru/ru/article/view?id=8371>.
6. Popova O.B. Sistemnyj podhod k issledovaniju processa optimizacii. Dep. v VINITI №83-V2010 ot 17.02.2010.
7. Popova O.B., Popov B.K., Kljuchko V.I. *Mezhdunarodnyi zhurnal eksperimental'nogo obrazovaniia* (International Journal of Experimental Education), 2013, №10-2, S. 405-408; URL: <http://cyberleninka.ru/article/n/analiz-svyazey-v-realnoy-i-tehnicheskoy-sistemah-protsessah-optimizatsii>.
8. Popova O.B., Popov B.K., Kljuchko V.I. Analiz svyazey v real'noj i tehničeskoj sistemah processa optimizacii // *Mezhdunarodnyj zhurnal jeksperimental'nogo obrazovaniia*. – 2013. – №10-2. – S. 405-408; URL: <http://elibrary.ru/download/85589730.pdf>.
9. Popova O.B., Popov B.K. Patent na izobretenie (The patent for invention) RUS № 2564641 ot 27.05.2014.
10. Popova O.B. Dep. v VINITI (Submitted to VINITI) №243-V2012 ot 25.05.2012.
11. Popova O.B. Dep. v VINITI (Submitted to VINITI) №256-V2010 ot 07.05.2010.
12. Popova O.B. Dep. v VINITI (Submitted to VINITI) №257-V2010 ot 07.05.2010.
13. Popova O.B., Popov B.K. Dep. v VINITI (Submitted to VINITI) №112-B2012 ot 22.03.2012.
14. Popova O.B., Popov B.K. Dep. v VINITI (Submitted to VINITI) №111-V2012 ot 22.03.2012.
15. Категория:Vedomstvennye kody i klassifikatory RF. Material iz Vikipedii — svobodnoj jenciklopedii. 2016. https://ru.wikipedia.org/wiki/Kategorija:Vedomstvennye_kody_i_klassifikatory_RF.

*MODERN CLASSIFIER FOR THE SUBJECT AREA***O.B. POPOVA**

*Kuban State Technological University,
2, Moskovskaya st., Krasnodar, Russian Federation, 350002;
e-mail: popova_ob@mail.ru*

Today, an actual task is to develop the scientific communication system through the establishment of effective information systems which scientometric calculations of different levels, helping to search for scientific information from various sources, organize convenient and efficient storage of scientific and technical information systems and links to relevant publications. Among them you can find modern base Science Citation RISC, Scopus, Web of Science. For convenient storage of information, the search and the reflection disciplinary and interdisciplinary characteristics of the subject area in such systems are used classifiers of various levels. New trends in this area - the association qualifier in several functions. Now electronic journals with the high rating usually use classifications that integrate the sorting of material on the subject, easy search articles, distribution of reviewers by thematic competences. The full picture of the development perspectives a particular the subject area, such classifiers are not determine, only certain vectors of development. To reveal the whole picture should be read many articles on the subject area. Author proposes to develop a method of preparation of the modern classifier the subject area, by appearance of which you can get an idea about the most the subject area and scientometric indicators.

Key words: modern classifier, subject area, a set of alternatives, classification method, a binary tree, keywords, search, scientific research.