

ОТЗЫВ ОФИЦИАЛЬНОГО ОППОНЕНТА

на диссертацию Милковой Марии Александровны на тему

«Разработка и экспериментальная апробация методики построения тематической модели
при работе с научной информацией»,

представленную на соискание ученой степени кандидата экономических наук по
специальности 08.00.13 – «Математические и инструментальные методы экономики»
(экономические науки)

Актуальность темы диссертационной работы

Бесспорными фактами нашего времени являются стремительный рост информации, в том числе, экономической и научно-технической, обусловленный известными феноменами Big Data, неубывающая потребность в достоверной и объективной информации со стороны ее разнообразных потребителей - с одной стороны, а с другой – очевидный прогресс в сфере методов машинного обучения, искусственного интеллекта и в целом информационно-коммуникационных технологий стимулируют развитие прикладных исследований в области ICT и разработок инструментов, упрощающих процедуры поиска и обработки тематической информации. Безусловно это относится и к удовлетворению потребностей специалистов и экспертов, связанным с проблематикой обоснования принятия управленческих решений в разнообразных сферах экономики и бизнеса. В этой связи представленная диссертация является актуальным исследованием, направленным на развитие инструментов и средств анализа текстовой информации, в том числе на создание алгоритмов, программных кодов, а также методик их настройки и использования применительно к конкретным областям человеческой деятельности.

Общая характеристика работы

Диссертация М.А. Милковой структурно оформлена в соответствии с требованиями, предъявляемыми к научным работам. Исследование изложено на 172

страницах машинописного текста. Работа состоит из введения, трех глав (включая 9 таблиц и 13 рисунков), заключения, списка литературы (228 позиций), а также семи приложений.

В целом структура диссертации обнаруживает внутреннее логическое и стилистическое единство, свидетельствующее о личном вкладе автора в выполненное исследование. Автор достаточно корректно использует известные научные методы исследования для аргументации и получения собственных результатов, выводов и рекомендаций; обоснованно приводит цитаты, делая необходимые ссылки на источники. Хотя при этом следовало бы более сбалансировано отнестись к формальному представлению элементов текста диссертации, в том числе соответствии требованиям ГОСТ Р 7.0.11-2011 (см., например, сс.61-62, 68-69, 72-73, 74-76, 102, 105-107 и др.).

Представленное автором в тексте диссертационной работы изложение методологии проведенного исследования и его результатов в основном логично и носит научный характер. Последовательность изложения и содержание представленного автореферата в основном корреспондирует со структурой диссертационной работы.

Таким образом, в целом, представленная Марией Александровной Милковой диссертационная работа может характеризоваться как целостное, последовательно изложенное, аргументированное и законченное научное произведение.

Оценка научной новизны исследования

Заявленная автором цель работы – «...адаптация тематического моделирования для выявления структуры научно-технической информации, способствующей эффективному получению знаний» (с.9) в целом достигнута (сс.67-70, 119-121, 145). Хотя нельзя не заметить, что в тексте работы отсутствует специальный структурный элемент текста, где бы концентрировано излагалась суть «адаптации». Тем не менее ее содержание можно понять из материала, изложенного на сс.67-70, а также на с.145.

Остановимся на оценке 4х пунктов научной новизны, обозначенных в тексте диссертационной работы (сс.10-11) и выносимых автором, М.А. Милковой, на защиту:

1. Автором выявлены особенности потребления информации (Глава 1), в частности, отмечается ограниченная рациональность (Г. Саймон) (с.19), когнитивные сдвиги восприятия (Д. Канеман, А. Тверски, Р. Талер) (с.20-22), экологическая рациональность (В. Смит, Г. Гигеренцер) (с.22-23), подход быстрых и экономных эвристик (Г. Гигеренцер

и др.) (с.23-24), автоматическая конформность (В. Ключарев) (с.31), влияние эмоций (Д. Канеман, А. Санфей) (с.28-29), ограниченность внимания (Г. Саймон, Ф. Хайек, К. Камерер) (с.32-34) и возможности его манипуляции (с.37-38). Соискатель дает перспективную оценку преломления ранее сформулированных положений в контексте дальнейшего развития современных поисковых, рекомендательных систем, которые стремятся предоставить «готовые» и «быстрые» ответы на запросы пользователя (с.40; с.43-44), что, снижая затраты пользователя на поиск информации, способствует решению проблем дефицита внимания, но при этом может вести к деградации аналитического мышления. Таким образом, автор приходит к мысли, что использование современных поисковых и рекомендательных систем способствует формированию фрагментарных знаний, с эффектом искажения восприятия (с.46-47). Предлагаемый автором подход тематического моделирования, представляя структуру исследуемого направления, способствует полноте получения знаний и стимулирует дальнейший научный поиск.

2. Автором предложен формализованный подход к выявлению структуры больших коллекций научных публикаций (являющийся надстройкой над «общим» алгоритмом ARTM), который позволяет снизить трудоемкость управления многочисленными параметрами модели, обеспечивая интерпретируемость результатов ее работы (отмечено на с.63). Его суть заключается в совокупности приемов и средств, используемых для выявления ключевых тем в коллекции научных публикаций, а также характеристик каждой из тем (ключевых терминов, основных авторов, опорных публикаций). Для чего предложен набор регуляризаторов (декоррелирования, сглаживания, разреживания); порядок их включения в модель; сетка перебора значений коэффициентов регуляризации; критерии качества модели, на основе которых возможно получать интерпретируемые результаты (перплексия, средняя когерентность по 15 наиболее частотным биграммам, степень разреженности матриц); а также набор модальностей (метаинформации научных публикаций) и их весов (с.67-69; с.145).

3. Апробация авторской методики проведена на коллекции публикаций по поведенческой и экспериментальной экономике (с.72-83). Полученные в ходе её результаты – автоматически выделенные ключевые темы, представленные в виде терминов, авторов и публикаций являются интерпретируемыми и непротиворечивыми (с.83-101). Авторский подход, выделяя релевантные темы, может уступать в точности

выявленных «якорных» фраз и авторов по темам, однако позволяет быстро определять адекватную структуру научной области там, где экспертное аннотирование отсутствует.

4. На основе тематического моделирования предложен подход к осуществлению патентного поиска для оценки результатов реализации в стране программы импортозамещения (с.104-108). А именно, по словам и двухсловным терминам из отраслевых планов импортозамещения осуществлялось сглаживание тематической модели, что позволило на выходе получать подборку патентных документов по 22 отраслям промышленности, соответствующих программам импортозамещения. На основе найденных документов выделены основные категории, по которым осуществлялось импортозамещение (с.108-112), предложен рейтинг, показывающий соответствие отраслей заявленным планам импортозамещения (с. 113-114). Авторский подход патентного поиска позволяет находить релевантные документы по большому числу пунктов планов импортозамещения (в исследовании поиск ведется суммарно по 1553 пунктам планов) по различным отраслям промышленности, что невозможно осуществлять стандартными средствами поисковых и аналитических систем (например, Яндекс.Патенты, Google Patents, Patseer и др.), обладающих ограниченными возможностями многоаспектного смыслового поиска (с.107).

В целом к общим достоинствам работы можно отнести следующие моменты:

- междисциплинарный характер исследования, обусловивший попытку применить комплексный подход и разнообразный исследовательский инструментарий к решению научных задач квалификационной работы;
- Учет содержательных аспектов и применение качественных методов в рамках анализа полноты и ценности информации при принятии решений экономическими агентами;
- вполне успешную попытку эмпирической оценки научной гипотезы о повышении качества обоснования и выбора решения за счет структуризации знания о предметной области на основе тематической модели и привлечение математического аппарата для анализа нечисловой информации, что позволяет обозначить дополнительные возможности применения технологий ИИ для анализа и обоснования управленческих решений в экономике.

Степень обоснованности и достоверности научных положений и выводов, сформулированных в диссертации, определяется использованием в ней современных и

апробированных в научной практике методов анализа текстовой информации (тематического моделирования). Полученные в ходе моделирования результаты являются адекватными, хорошо интерпретируемыми и непротиворечивыми.

На основании анализа существенного количества научных источников (228 позиций) в области поведенческой экономики, экономики знаний, экономики информации, нейроэкономики дается научное обоснование применимости развиваемого в работе инструментария для оказания помощи в процессе разработки принятия решений в заданной предметной области.

Основные положения и результаты диссертационного исследования были ранее изложены автором в 13 научных публикациях и доложены на 5 научных конференциях и семинарах. Три работы опубликованы автором в журналах из перечня ВАК РФ, одна работа – в издании, индексируемом в базе данных Scopus.

Учитывая вышесказанное, можно утверждать, что научные положения, выводы и результаты соискателя Милковой М.А. являются вполне обоснованными и достоверными, предварительно прошедшими апробацию.

Теоретическая и практическая значимость исследования заключается в развитии подходов к обработке текстовой информации для принятия решений о выборе информационных продуктов. В частности, речь идет о принятии решений, связанных с контентным поиском в области экономики информации, экономики знаний. Авторский алгоритм тематического моделирования, как элемент общей методики построения тематической модели предметной области позволяет организовать информационное взаимодействие производителей и потребителей научного знания и тем самым повышать эффективность работы научного сообщества.

Результаты авторской разработки могут быть полезны разработчикам поисковых систем, т.к. результаты экспериментов по подбору параметров тематической модели выявления тематики научных публикаций, а также программный код выложены в репозитории Github (с. 121): <https://github.com/behavioral-econ-codes/Publications>. Созданный автором веб-сайт с результатами моделирования <http://behavioral.site>, позволяющий осуществлять оперативную навигацию по темам в области поведенческой и экспериментальной экономики, может быть интересен специалистам, работающим в этой научно-предметной области.

Замечания по диссертационной работе

При экспертизе диссертационной работы нельзя обойти и ряд проблемных моментов и дискуссионных вопросов в изложении представленного исследования и его результатов. К наиболее существенным следует отнести:

- 1) Нечеткость в идентификации и обозначении научной проблемы, увязанной с сформулированным названием и поставленной целью исследования (сс. 1, 9), а также несоответствии представленного в автореферате диссертационной работы состава и содержания научной новизны проведенного исследования (сс.8-9, автореферат) его раскрытию во 2-м разделе автореферата (сс.12-25, автореферат).
- 2) В тексте диссертационной работы не приводится акцентированной и детальной характеристики оригинальности предложенной соискателем общей методики построения тематической модели, поскольку ей посвящены лишь ссылка в тексте (с. 70) и иллюстрация в Приложении 1, с.145, поэтому полноценно и содержательно методику можно оценить лишь на основе изложения авторского подхода «рассыпанного» в материалах Главы 2 и в контексте примеров её использования в материалах Главы 3. Это сказывается на степени аргументированности доводов соискателя об элементах научной новизны полученных результатов исследования (сс.10-11).
- 3) При описании прикладных аспектов реализации авторского подхода следовало явно указать и подробнее раскрыть, с какими экономическими задачами, решаемыми экономическим агентом при принятии решения, связаны примеры построения тематической модели для совокупности обрабатываемой метаинформации о научных публикациях по направлениям поведенческой и экспериментальной экономики (параграф 3.1) и для набора патентной информации о зарегистрированных объектах интеллектуальной собственности в сфере импортозамещения (параграф 3.2).
- 4) При изложении материала диссертационного исследования соискатель свободно оперирует набором разнообразных терминов и категорий из различных областей научного знания. Однако для адекватного понимания сущности основных авторских положений, доводов и выводов целесообразно было бы дать определения ряда понятий: «внимание» (уточнив его сущность и характеристику как ресурса и товара, сс. 4-5 и далее в параграфе 1.4), «информационный продукт» (раскрыв контекст утверждения, что цель настоящей диссертации в использовании методов тематического моделирования, чтобы способствовать принятию решений о потреблении информационных продуктов, с. 62).

5) Можно отметить некоторую несбалансированность текста диссертационной работы, например, предлагаемая и защищаемая автором методика сжата до схемы на 1й странице в приложении (с. 145), а результат тематического моделирования на примере публикаций в области поведенческой и экспериментальной экономики представлен в основном тексте диссертационной работы на 19 страницах (сс. 83-101). Следовало бы уделить больше внимания характеристике степени общности выводов соискателя при освещении предложения по включению модальностей (параметров информационного запроса) для традиционного подхода обработки совокупности исходных текстов в конкретном примере с поиском в коллекции научных публикаций (с. 10-11 и далее в параграфе 2.2.7), а также при указании в п.2-3 заключения результатов апробации подхода к управлению ограниченной рациональностью при формализации знаний в условиях перенасыщения текстовой информации на основе применения аддитивной регуляризации тематических моделей (с. 119-121).

Имеются ряд редакционных замечаний к оформлению текста диссертационного исследования. Например: не указано, кем из представителей предметной области верифицированы полученные результаты научных публикаций в области поведенческой экономики (с. 13); не ясно, на что «использование достижений поведенческих наук в сочетании с цифровыми технологиями способно оказывать существенное воздействие» (с. 47); ошибочное обозначение формул уравнений 7-9 (с. 56); ошибочно указана ссылка для примера запросов к базе MongoDB в Приложении 1 (фактически - Приложении 2) (с. 74), некорректность в указании ссылок на источники (с.37) и др.

Вместе с тем, указанные недостатки и отмеченные неточности не снижают значимости полученных результатов и не влияют на общую положительную оценку диссертационного исследования Милковой М.А.

Заключение о соответствии диссертации критериям, установленным «Положением о присуждения ученых степеней», утвержденным постановлением Правительства Российской Федерации №842 от 24.09.2013

Диссертационная работа Марии Александровны Милковой на тему «Разработка и экспериментальная апробация методики построения тематической модели при работе с научной информацией» представляет собой логически завершенное самостоятельное

научно-квалификационное исследование, обладающее научной новизной, имеющее теоретическую и практическую значимость.

Диссертация соответствует паспорту специальности 08.00.13 «Математические и инструментальные методы экономики», в части пунктов 2.6. «Развитие теоретических основ методологии и инструментария проектирования, разработки и сопровождения информационных систем субъектов экономической деятельности: методы формализованного представления предметной области, программные средства, базы данных, корпоративные хранилища данных, базы знаний, коммуникационные технологии»; 2.8. «Развитие методов и средств аккумуляции знаний о развитии экономической системы и использование искусственного интеллекта при выработке управленческих решений».

Работа соответствует требованиям пп. 9-14 «Положения о присуждении ученых степеней», утвержденного Постановлением Правительства Российской Федерации от 24 сентября 2013 г. № 842, предъявляемым к кандидатским диссертациям.

Считаю, что автор диссертационного исследования, Милкова Мария Александровна, заслуживает присуждения ей ученой степени кандидата экономических наук по специальности 08.00.13 – «Математические и инструментальные методы экономики» (экономические науки).

Официальный оппонент:

Кандидат экономических наук,
доцент, заведующая кафедрой
математических методов в экономике и
управлении ФГБОУ ВО
«Государственный университет
управления»



Писарева Ольга Михайловна

Подпись Писаревой О.М. заверяю:
Ученый секретарь Ученого совета
д.э.н., профессор М.А. Жукова



«03» декабря 2021 г.

Адрес: 109542, г. Москва, Рязанский проспект, 99,
ГУУ, Лабораторный корпус, лк-219
Телефон: +7 (495) 371-05-28
E-mail: om_pisareva@guu.ru