

ЗАКЛЮЧЕНИЕ

ДИССЕРТАЦИОННОГО СОВЕТА Д 002.013.01, СОЗДАННОГО НА БАЗЕ
ФЕДЕРАЛЬНОГО ГОСУДАРСТВЕННОГО БЮДЖЕТНОГО УЧРЕЖДЕНИЯ НАУКИ
ЦЕНТРАЛЬНОГО ЭКОНОМИКО-МАТЕМАТИЧЕСКОГО ИНСТИТУТА
РОССИЙСКОЙ АКАДЕМИИ НАУК, ПО ДИССЕРТАЦИИ НА СОИСКАНИЕ
УЧЕНОЙ СТЕПЕНИ КАНДИДАТА НАУК

Аттестационное дело № _____

Решение диссертационного совета

от 20 декабря 2021 г. № 14

О присуждении Милковой Марии Александровне, гражданке Российской Федерации ученой степени кандидата экономических наук.

Диссертация «Разработка и экспериментальная апробация методики построения тематической модели при работе с научной информацией» по специальности: 08.00.13 - «Математические и инструментальные методы экономики» принята к защите 18 октября 2021 г. (протокол заседания № 12) диссертационным советом Д 002.013.01, созданным на базе Федерального государственного бюджетного учреждения науки Центрального экономико-математического института Российской академии наук, подведомственного Министерству науки и высшего образования Российской Федерации. Адрес: 117418, г. Москва, Нахимовский проспект, д. 47 (приказ № 105/нк от 11 апреля 2012 г.).

Соискатель Милкова Мария Александровна, 30 июля 1986 года рождения, в 2008 году окончила Московский государственный институт электроники и математики (технический университет), квалификация экономист-математик по специальности «Математические методы в экономике» (диплом ВСГ 2918233 от 19.06.2008), по настоящее время является научным сотрудником лаборатории 1.02 «Экспериментальной экономики» ЦЭМИ РАН.

Диссертация выполнена в лаборатории 1.02 «Экспериментальной экономики» Федерального государственного бюджетного учреждения науки Центральный экономико-математический институт Российской академии наук (117418, Москва, Нахимовский проспект, 47).

Научный руководитель — доктор экономических наук, кандидат физико-математических наук, главный научный сотрудник, научный руководитель лаборатории 1.02 «Экспериментальной экономики», руководитель научного направления «Математические и компьютерные модели, экономика знаний, инструменты и методы» Федерального государственного бюджетного учреждения науки Центральный экономико-математический институт Российской академии наук Козырев Анатолий Николаевич.

Официальные оппоненты:

1. Лугачев Михаил Иванович, доктор экономических наук, профессор, научный руководитель кафедры экономической информатики Экономического факультета Федерального государственного образовательного учреждения высшего профессионального образования Московского государственного университета имени М.В. Ломоносова.

2. Писарева Ольга Михайловна, кандидат экономических наук, доцент, заведующая кафедрой математических методов в экономике и управлении Федерального государственного бюджетного образовательного учреждения высшего образования Государственного университета управления, дали положительные отзывы на диссертацию.

Ведущая организация — Федеральный исследовательский центр «Информатика и управление» Российской академии наук (ФИЦ ИУ РАН) в своем положительном отзыве, подписанном заведующей отделом информационных технологий оценки эффективности инвестиций ФИЦ ИУ РАН, доктором экономических наук, профессором Орловой Еленой Роальдовной, утвержденном директором ФИЦ ИУ РАН академиком РАН, доктором технических наук Соколовым Игорем Анатольевичем, указывает, что представленная диссертационная работа имеет теоретическую и практическую значимость, обладает научной новизной, выполнена на актуальную тему, соответствует требованиям «Положения о присуждении ученых степеней», утвержденного Постановлением Правительства Российской Федерации от 24.09.2013 № 842.

По теме диссертации опубликовано 13 работ общим объемом 19,2 п.л. (вклад автора — 18 п.л.), в том числе в 3 статьях в изданиях, включенных в перечень ВАК РФ, в 1 статье в издании, индексируемом в базе данных Scopus (авторский вклад — 5,5 п.л.). В публикациях отражены основные теоретические и прикладные результаты выполненного диссертационного исследования.

Наиболее значимые публикации автора по теме диссертации:

1. Милкова М.А. Инновационный подход к поиску информации на примере патентного анализа плана импортозамещения // *Экономическая наука современной России*. 2020. № 1(88). С. 143-157. DOI: 10.33293/1609-1442-2020-1(88)-143-157 (1.7 п.л.). Соискателем разработан подход к анализу программы импортозамещения на основе патентных данных с использованием инструментария аддитивной регуляризации тематических моделей.
2. Милкова М.А., Неволин И.В., Пигорев Д.П. Извлечение ключевой информации из нормативных документов о политике продовольствия и питания в России // *Экономическая наука современной России*. 2021. № 2(93). С. 101-114. DOI: 10.33293/1609-1442-2021-2(93)-101-114 (2.2 п.л., авт. – 1.8 п.л.). Личный вклад автора – применение методов текстового анализа для реферирования информационных блоков.
3. Milkova M.A. Patent-based import substitution analysis with Additively Regularized Topic Models // *Proceedings of the 10th International Scientific and Practical Conference named after A. I. Kitov "Information Technologies and Mathematical Methods in Economics and Management (IT&MM-2020)"*. Moscow, Russia, October 15-16, 2020. CEUR Workshop Proceedings. – 2021, Vol. 2830, pp. 16-27. urn:nbn:de:0074-2830-9 (1.0 п.л.). Соискателем приведено описание методики построения тематической модели для патентного анализа документов по отраслям импортозамещения.
4. Milkova, M., Andreichikova, O., Andreichikov, A. Decision making under uncertainty: a heuristics overview and the analytic network process // *Psychology Journal of the Higher School of Economics*. 2019. Vol. 16. N 4, pp. 730–751. DOI: 10.17323/1813-8918-2019-4-730-751 (2.8 п.л., авт. – 2.0 п.л.). Личный вклад автора – аналитический обзор эвристических подходов к принятию решений в условиях неопределенности.
5. Милкова М.А. Тематическое моделирование: восприятие научной информации // *Цифровая экономика*. 2021. № 14(2). С. 31-36. DOI: 10.34706/DE-2021-02-04 (0.9 п.л.). Соискателем разработана методика построения тематической модели для выявления ключевых тем (а также ключевых терминов, основных авторов, опорных публикаций) больших коллекций научных публикаций.
6. Милкова М.А. Информация и ограниченная рациональность выбора в цифровой экономике // *Цифровая экономика*. 2021. № 13(1). С. 69-88. DOI:

10.34706/DE-2021-01-08 (3.2 п.л.). Соискателем приведен аналитический обзор поведенческих особенностей обработки информации и принятия решений, а также инструментов для исследования этих особенностей.

На автореферат поступили отзывы:

1. Савцова Анна Валерьевна, доктор экономических наук, доцент, профессор кафедры финансов и кредита ФГАОУ ВО «Северо-Кавказский федеральный университет», отмечает необходимость пояснения в автореферате, какие существуют альтернативные инструменты выявления смысла больших текстовых коллекций, в чем перед ними преимущество используемого в работе инструментария; необходимость пояснения, в какой степени указанные в списке работы по смежным темам оказали влияние на работу по защищаемой теме.

2. Перепечко Людмила Николаевна, кандидат физико-математических наук, руководитель Центра трансфера технологий ФГБУН Института теплофизики им. С.С. Кутателадзе Сибирского отделения Российской академии наук, отмечает отсутствие обоснования использования базы Semantic Scholar для выгрузки исходных данных, а также причину выбора для анализа области поведенческой и экспериментальной экономики, обращает внимание на сокращенный вид схемы формализованного алгоритма в автореферате, необходимость уточнения, по какому параметру исследовалась чувствительность результатов; обращает внимание на формулировку объекта исследования как некоторой социальной группы и предмета исследования как взаимодействие некоторых социальных групп.

3. Жуков Михаил Станиславович, кандидат экономических наук, инженер в области облачных технологий компании Corporate Planning AG, отмечает отсутствие в автореферате обоснования того, почему для работы выбран именно подход аддитивной регуляризации тематических моделей и чем он отличается от других известных подходов (например, Латентного размещения Дирихле, LDA).

4. Кульба Владимир Васильевич, доктор технических наук, профессор, главный научный сотрудник лаборатории №20, заслуженный деятель науки РФ и Сиротюк Владимир Олегович, доктор технических наук, доцент, ведущий научный сотрудник лаборатории №20 ФГБУН Института проблем управления им. В.А. Трапезникова Российской академии наук, отмечают перегруженность Главы 1 обзорными сведениями в области экономики информации, экономики знаний, информационных продуктов, вместо уделения большего внимания

вопросам формирования, сопровождения и использования патентной и научно-технической информации. Рекомендуют при построении тематической модели для патентных документов использовать классификаторы МПК и УДК, рассмотреть такие качества тематической модели как полнота, достоверность, доступность и своевременность данных и метаданных, неизменность структуры данных.

5. Федорец Олег Владимирович, кандидат технических наук, начальник отдела программных систем Управления информационных систем ФГБУН Всероссийский институт научной и технической информации Российской академии наук, отмечает неясное обоснование включения в раздел «Степень разработанности проблемы» нейроэкономики, поведенческой экономики, психологических и поведенческих аспектов принятия решений, высказывает пожелания раскрытия в автореферате сути экспериментального определения веса модальностей, приведения сведений о технических характеристиках компьютера и машинном времени, затраченном на построение тематических моделей; обращает внимание на отсутствие указания множества, к которому принадлежит индекс s в формуле (3) в автореферате.

6. Федоров Сергей Викторович, кандидат экономических наук, научный сотрудник центра аналитических исследований и моделирования в образовании НИИ урбанистики и глобального образования ГАОУ ВО «Московский городской педагогический университет» рекомендует уточнить степень общности формализованного алгоритма, а также возможность его применения для русскоязычных публикаций.

7. Свиридова Ольга Александровна, кандидат экономических наук, доцент кафедры математических методов в экономике ФГБОУ ВО «Российский экономический университет им. Г.В. Плеханова», рекомендует изучить удобство представления результатов тематического моделирования научных публикаций для пользователей, возможности измерения эффективности деятельности научного сообщества, возможности масштабирования результатов исследования при поступлении новых данных.

Выбор доктора экономических наук Лугачева Михаила Ивановича в качестве первого официального оппонента обосновывается тем, что он является одним из ведущих специалистов в области экономики информации, экономики информационных продуктов. М.И. Лугачев внес существенный вклад в изучение

особенностей экономических аспектов цифровой трансформации, современных информационных технологий в исследовании сложных структур.

Выбор кандидата экономических наук, доцента Писаревой Ольги Михайловны в качестве второго официального оппонента обосновывается тем, что Ольга Михайловна является одним из ведущих специалистов в области экономической кибернетики, машинного обучения. Сфера ее научных интересов касается в том числе методологии и практики организационно-экономических механизмов стимулирования научно-исследовательской деятельности. Научные работы О.М. Писаревой вносят вклад в разработку и адаптацию математического инструментария в цифровой среде, в том числе с использованием моделей обработки естественного языка.

Выбор Федерального исследовательского центра «Информатика и управление» Российской академии наук (ФИЦ ИУ РАН) в качестве ведущей организации обосновывается тем, что сфера научных исследований ФИЦ ИУ РАН, в частности, заведующего отделом информационных технологий оценки эффективности инвестиций, профессора, д.э.н. Е.Р. Орловой, сотрудников отдела к.э.н. А.В. Вершининой, к.э.н. Е.Н. Кошкиной связана с исследованиями широкого спектра в области информационной экономики, изучением применения информационных технологий и искусственного интеллекта в образовательном процесс, что позволяет определить как теоретическую, так и практическую ценность диссертационной работы М.А. Милковой.

Диссертационный совет отмечает, что на основании выполненных соискателем исследований:

предложен оригинальный подход к организации потребления информационных продуктов (научных публикаций, патентных документов) на основе тематического моделирования. Впервые предложено научное обоснование, что подход к организации информационных продуктов на основе тематического моделирования, в отличие от подхода на основе алгоритмов повсеместно распространенных поисковых систем, позволяет повысить эффективность работы научного сообщества, благодаря организации информационного взаимодействия производителей и потребителей научного знания;

разработана авторская методика построения тематической модели (в концепции аддитивной регуляризации тематической модели, АРТМ) при работе с научными публикациями, представленная в виде комплекса программ, набора инструкций. Ее 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. Предложенная методика построения модели ARTM, в отличие от общего подхода ARTM, позволяет снизить трудоемкость подбора параметров при построении тематической модели на базе научных публикаций;

предложена оригинальная тематическая кластеризация научных публикаций по Поведенческой и экспериментальной экономике: впервые выявлены машинным способом ключевые темы, в каждой из которых определены ключевые термины, основные авторы и опорные публикации. Авторский подход позволяет определять адекватную структуру научной области там, где экспертное аннотирование отсутствует;

разработан оригинальный подход к анализу программы импортозамещения на основе патентных данных с использованием инструментария аддитивной регуляризации тематических моделей. С использованием разработанного подхода получена подборка патентных документов по 22 отраслям промышленности, соответствующих программам импортозамещения. В результате анализа найденных документов впервые выделены основные категории, по которым осуществлялось импортозамещение, предложен рейтинг, показывающий соответствие отраслей заявленным планам импортозамещения (по патентным данным). Предложенный подход патентного поиска позволяет находить релевантные документы сразу по большому числу пунктов планов импортозамещения по различным отраслям промышленности, что невозможно осуществлять стандартными средствами поисковых и аналитических систем.

Теоретическая значимость исследования обоснована тем, что:

изучены особенности потребления информации (в том числе, научно-технической информации), в частности, ограниченная рациональность, когнитивные сдвиги, экологическая рациональность, влияние эмоций, склонность к выбору опции по умолчанию, ограниченность внимания. На основе выделенных особенностей разработан способ организации информационных продуктов, представленных в виде научных публикаций и патентных

документов, помогающий в выборе ценных (с точки зрения вычлененного смысла) продуктов;

раскрыто, что подход к организации информационных продуктов на основе тематического моделирования, в отличие от подхода на основе алгоритмов повсеместно распространенных поисковых систем, может быть результативно использован для повышения эффективности работы научного сообщества (с точки зрения организации информационного взаимодействия производителей и потребителей научного знания).

Значение полученных соискателем результатов исследования для практики подтверждается тем, что:

разработан и внедрен комплекс программ, снижающих трудоемкость построения тематической модели для научной информации, и тем самым позволяющий повысить эффективность деятельности экспертов при сокращении затрат на извлечение необходимой информации. Результаты работы использованы в научной и практической деятельности Национальной ассоциации трансфера технологий; прошли всестороннюю проверку специалистов компаний «Кит оценка» и используются при производстве судебных экспертиз. Подход к потреблению информационных продуктов (представленных в виде научных публикаций) внедрен в учебный процесс на кафедре Экономической информатики Экономического факультета МГУ им. Ломоносова в рамках курса «Экономика информации» для подготовки бакалавров;

создан веб-сайт с результатами моделирования <http://behavioral.site>, позволяющий осуществлять оперативную навигацию по темам в области поведенческой и экспериментальной экономики, может быть полезен специалистам, работающим в этой научно-предметной области.

Оценка достоверности результатов исследования выявила:

идея базируется на анализе существенного количества научных источников (228 позиций) в области поведенческой экономики, экономики знаний, экономики информации, нейроэкономики (речь об идее использования инструментария тематического моделирования для оказания помощи в принятии решений о потреблении информационных продуктов в заданной предметной области);

использованы в диссертационной работе:

а). (в качестве информационной базы исследования) база научных публикаций Semantic Scholar (всего 37352 научных публикаций до 2020 года включительно),

официальная база Федерального института промышленной собственности Роспатента (всего 152718 документов, за период январь 2016 - июнь 2019 гг.), официальные планы импортозамещения Минпромторг для 22 отраслей промышленности;

б). современные и апробированные в научной практике методы анализа текстовой информации (тематическое моделирование, концепция аддитивной регуляризации тематических моделей). Полученные в ходе моделирования результаты являются адекватными, хорошо интерпретируемыми и непротиворечивыми.

Личный вклад соискателя состоит в:

- включенном участии на всех этапах процесса, непосредственном участии соискателя в получении исходных данных и научных экспериментах, личное участие в апробации результатов исследования;

- подготовке основных публикаций по выполненной работе: всего опубликовано 13 работ общим объемом 19,2 п.л. (вклад автора — 18 п.л.);

- исследовании особенностей потребления информации в контексте цифровизации, принятия решений экономическими агентами о потреблении информационных продуктов;

- разработке формализованного подхода к выявлению структуры больших коллекций научных публикаций на основе тематического моделирования: предложен комплекс программ, проведены эксперименты, представлен набор инструкций;

- апробации предложенного подхода для выявления ключевых тем в коллекции научных публикаций по поведенческой и экспериментальной экономике, а также характеристик каждой из тем: ключевых терминов, основных авторов, ключевых публикаций;

- разработке подхода к анализу программы импортозамещения на основе патентных данных с использованием инструментария аддитивной регуляризации тематических моделей, что позволило оценить соответствие отраслей промышленности заявленным планам импортозамещения (по патентным данным).

В ходе защиты диссертации академиком РАН, д.э.н., проф. В.М. Полтеровичем В.М. было высказано замечание о том, что в исследование было бы желательно включить эксперимент, где была бы поставлена задача по поиску

информации с помощью стандартных алгоритмов (например, Google Scholar) и алгоритмов, предложенных в диссертации. Д.т.н., проф. Акопов А.С. высказал замечание о том, что необходимо было включить описание технической реализации предлагаемых алгоритмов в текст автореферата.

Соискатель Милкова М.А. согласилась с замечаниями, ответила, что данные задачи не входили в цели исследования, но будут учтены в дальнейших работах.

На заседании 20 декабря 2021 г. диссертационный совет Д 002.013.01 принял решение: за результаты, представленные в диссертационной работе методикой построения тематической модели при работе с научной информацией, которая имеет важное значение в области экономики информации, экономики информационных продуктов, присудить Милковой Марии Александровне ученую степень кандидата экономических наук.

При проведении тайного электронного голосования из 21 члена диссертационного совета, участвовавших в заседании и указанных в явочном листе, в голосовании приняли участие 20 человек, из них 10 докторов наук (отдельно по каждой научной специальности рассматриваемой диссертации), из 27 человек, входящих в состав совета, проголосовали следующим образом:

«ЗА» - 20, «ПРОТИВ» - нет.

« 23 » декабря 2021 г.

Председатель

диссертационного совета Д 002.013.01,
академик РАН, д.ф.-м.н., профессор

Макаров Валерий Леонидович

Ученый секретарь

диссертационного совета Д 002.013.01,
к.э.н.

Ставчиков Александр Иванович