

«Утверждаю»

Директор ФИЦ ИУ РАН,  
академик  
И.А. Соколов



« 6 » 11 2021 г.

## ОТЗЫВ ВЕДУЩЕЙ ОРГАНИЗАЦИИ

на диссертацию Милковой Марии Александровны «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»

### 1. Актуальность темы диссертационного исследования

Диссертационная работа посвящена актуальной в настоящее время проблеме управления вниманием к информации в условиях ее перенасыщения. В эпоху цифровой трансформации темпы накопления информации так велики, что это порождает значительные сложности ее восприятия и анализа. С экономической точки зрения принципиальное значение здесь имеет вопрос о распределении наиболее дефицитного в настоящее время ресурса – внимания, дефицит которого растет вместе с ростом объема информации. Данный факт касается информации разного рода, в том числе и научной. Задача извлечения ключевой информации из больших массивов с целью принятия на ее основе решений в настоящее время решается с использованием информационно-коммуникационных технологий. Однако цели, которые преследуют поисковые и рекомендательные системы при извлечении ключевой информации отличны от целей, которые актуальны для эффективной работы научного сообщества. В

первом случае – это предоставление быстрых и фрагментарных ответов, во втором случае – стремление к полноте, возможности получения структуры исследуемой области. Тематическое моделирование в данном контексте позволяет справиться с дефицитом внимания, представить дорожную карту предметной области. С учетом сказанного, рассматриваемую в диссертации тему разработки и экспериментальной апробации методики построения тематических моделей при работе с научной информацией можно признать актуальной научной задачей, представляющей теоретический и практический интерес.

## **2. Научная новизна исследования и полученных автором результатов**

В диссертационной работе были получены следующие основные результаты, обладающие научной новизной.

- Приведено научное обоснование оправданности использования тематического моделирования для конфигурирования среды с целью восприятия информации так, чтобы снизить затраты на поиск ключевой информации; обеспечить полноту и ценность (с точки зрения смысла) извлекаемой информации; минимизировать сдвиги в восприятии (сместить фокус внимания с первой попавшейся, эмоциональной информации) (с.42-47).

- Предложен и апробирован в экспериментах формализованный алгоритм аддитивной регуляризации тематических моделей на базе научных публикаций по поведенческой и экспериментальной экономике. Представлен комплекс программ, набор инструкций, что позволяет использовать авторские наработки в других областях научных знаний (с.68-69; с.119-120; с.145).

- На основе формализованного алгоритма выявлены основные темы поведенческой и экспериментальной экономики (на базе англоязычных научных публикаций из базы Semantic Scholar). Каждая из тем представлена автоматически выявленными ключевыми фразами, основными авторами, наиболее цитируемыми публикациями (с.83-101). Проведен анализ чувствительности результатов моделирования. Показано, что результаты устойчивы к изменениям начального приближения матриц (с.102-103).

- Создан онлайн ресурс – сайт <http://behavioral.site> для осуществления пользовательской навигации по основным темам, авторам и публикациям научной области Поведенческая и экспериментальная экономика (с.117; с.170-172).

- Проведен анализ программы импортозамещения на основе патентных данных с использованием инструментария аддитивной регуляризации тематических моделей (с.104-108). Предлагаемый 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. Подход также позволяет выделять основные категории, по которым осуществлялось импортозамещение по отраслям (с.109-112), составить рейтинг отраслей (на основе патентных данных) (с. 113-114).

### **3. Обоснованность положений и выводов, сформулированных в диссертационной работе**

Диссертационная работа представляет собой целостное и самостоятельное исследование, имеющее логичную структуру. Использование автором тематического моделирования к решению задач извлечения ключевой информации из больших массивов текстовых данных (научных публикаций, патентных данных) соответствует современным подходам в области анализа естественного языка. Предложенная методика построения тематической модели является в достаточной мере обоснованной, что подтверждается хорошо интерпретируемыми, релевантными результатами моделирования. Кроме того, стоит отметить, что решение описанных в диссертации задач имеет прямое отношение к экономической науке, так как представляет собой способ организации информационных продуктов с точки зрения их потребления, оказывает помощь в принятии решения о потреблении информационных продуктов на основе их ценности.

Достоверность и обоснованность научных выводов, полученных в диссертационном исследовании, подтверждены на различных международных, всероссийских научно-практических конференциях, симпозиумах и научных семинарах; публикациями по теме диссертации; использованием теоретических и прикладных разработок как отечественных, так и зарубежных ученых; объемным перечнем использованных информационных источников.

### **4. Теоретическая и практическая значимость результатов, полученных автором**

Теоретическая значимость результатов исследования состоит в систематизации и развитии подходов к восприятию и работе с информацией человеком. В частности, подход к организации текстовой информации с использованием тематического моделирования и принятие на его основе решений оказывает существенный вклад в экономику информации, экономику знаний.

Практическая значимость результатов исследования заключается в разработке инструментария для упорядочения работы с научной информацией в условиях информационного взрыва. Предложенная методика работы с научной информацией представляет практическую ценность при проведении исследований – получении представления о структуре направления, выявлении ключевых авторов и опорных публикаций. Подход позволяет повысить эффективность работы научного сообщества путем организации информационного взаимодействия производителей и потребителей научного знания.

## **5. Рекомендации по использованию результатов и выводов диссертации**

Основные положения и результаты диссертационного исследования могут быть использованы преподавателями ВУЗов при составлении курса «Экономика информации», а именно, методы машинного обучения могут использоваться для целей организации потребления информационных продуктов; курса «Поведенческая экономика», а именно, выделенные ключевые термины, авторы и публикации направления Поведенческая и экспериментальная экономика могут послужить удобной отправной точкой для дальнейшего изучения области студентами, аспирантами и молодыми учеными.

Предложенный формализованный алгоритм аддитивной регуляризации тематических моделей позволяет выделять ключевые темы и их характеристики в научных публикациях в любой научной области, что может быть использовано при проведении библиографических исследований, оказывать помощь молодым исследователям при изучении предметной области.

Подход по проведению патентного анализа на основе тематического моделирования позволяет применять его, например, при заключении международных контрактов, когда необходимо проводить семантический патентный поиск (поиск по смыслу) сразу по множеству категорий.

## **6. Замечания по диссертационной работе**

Вместе с тем представленная работа не лишена некоторых недостатков.

Несмотря на то, что результаты автоматического выделения ключевых тем (и их характеристик - ключевых терминов, авторов, ссылок) в целом адекватны, их нельзя назвать точными. Например, отсутствуют такие рубрики как

поведенческая теория игр, социальные нормы, структурные эконометрические модели.

В тексте диссертации отсутствует информация о том, кто может быть потенциальным пользователем созданного ресурса (веб сайта). Существует ли в настоящее время потребность в таком ресурсе? Ученый, как правило, начинает свой поиск с ресурса Google Scholar (или аналогичных поисковых ресурсов), находит интересующие работы по поисковым словам, оценивает количество ссылок на них, в каких журналах они опубликованы – как правило этого для целей исследователя достаточно.

Приведенный анализ программы импортозамещения на основе патентных данных никак не соотносится с результатами импортозамещения на основе других показателей. Неясно, каким образом полученные результаты соотносятся с действительностью.

Отмеченные замечания не отменяют общую положительную оценку диссертационного исследования.

## **7. Заключение**

Диссертационная работа Милковой Марии Александровны представляет собой завершённую научно-квалификационную работу на актуальную тему, выполненную на высоком уровне. Полученные автором результаты обладают научной новизной и оригинальностью, имеют существенную теоретическую и практическую значимость. Цель диссертационного исследования достигнута, а поставленные в работе задачи успешно решены соискателем. Опубликованные в рецензируемых изданиях работы и автореферат в полной мере отражают содержание и основные выводы, представленные в диссертационном исследовании.

Можно утверждать, что диссертация удовлетворяет требованиям «Положения о присуждении ученых степеней» (утв. Постановлением Правительства Российской Федерации от 24 сентября 2013 г. № 842), а ее автор, Милкова Мария Александровна, заслуживает присуждения ей ученой степени кандидата экономических наук по специальности 08.00.13 – «Математические и инструментальные методы экономики».

Отзыв подготовлен заведующей отделом «Информационные технологии оценки эффективности инвестиций», доктором экономических наук, профессором, Орловой Еленой Роальдовной.

