

На правах рукописи

Неволин Иван Викторович

АЛГОРИТМИЧЕСКИЙ ПОДХОД К ЦЕННОСТИ ДАННЫХ

5.2.2. «Математические, статистические и инструментальные
методы экономики» (экономические науки)

Диссертация на соискание ученой степени доктора экономических наук

Научный консультант
д.э.н., к.ф.-м.н. Козырев Анатолий Николаевич

Москва 2025

Содержание

Введение.....	5
1. Понимание ценности информации в экономической науке	22
1.1. Расширение доступа к ресурсам через знания.....	22
1.2. Управление, анализ сценариев и ситуационное моделирование.....	29
1.3. Устранение неопределённости информации при принятии решений	34
1.4. Новая парадигма работы со знаниями, информацией и данными.....	45
1.5. Три подхода к понятию количества информации и ее ценности	53
Выводы раздела	59
2. Организация рынка данных (информации)	61
2.1. Информация как публичное благо, поставляемое в частном порядке.....	63
2.2. Модели равновесия с общественным благом	67
2.3. Уравнение Сэмюэльсона и оптимальные цены общественных благ	78
2.4. Подходы к назначению цен	81
2.5. Алгоритм Гровса-Кларка	90
2.5.1. Проблемы использования линдалевских цен.....	91
2.5.2. Выявление спроса на общественное благо (алгоритм Гровса-Кларка).....	92
2.5.3. Оптимизация по Парето с распределением вычислений	97
2.6. Передача ноу-хау, протоколы и автоматизация	101
Выводы раздела	107
3. Стоимость и цены в лицензионных договорах	108
3.1. Двухкомпонентные цены (теория).....	108
3.2. Двухкомпонентные цены (практика).....	116
3.3. Сервис по уточнению двухкомпонентных цен.....	122
Выводы раздела	127
4. Распределение выигрыша с применением теории игр	129
4.1. Модель сделки по Нэшу при заключении лицензионного договора.....	130
4.1.1. Особенности лицензионных сделок.....	130

4.1.2. Формальная схема сделки по Нэшью	131
4.1.3. Интерпретация результата сделки.....	135
4.1.4. Методы дележа при передаче технологий	139
4.2. Распределение стоимости программного комплекса между его компонентами	147
4.3. Анализ портфеля программ большой размерности	155
Выводы раздела	164
5. Технологическая платформа для рынка данных.....	166
5.1. Сети изобретательства, рационализации, инноваций.....	169
5.2. Платформенное решение для рынка данных	173
5.2.1. Каталог объектов	174
5.2.2. Машинный анализ текстов и методы суммаризации	176
5.2.3. Организационный механизм загрузки данных	177
5.3. Обеспечение качества каталога.....	178
5.3.1. Модель краудсорсинга	179
5.3.2. Препятствия для внедрения модели.....	180
5.3.3. Моделирование	181
5.4. Назначение цен за использование данных.....	184
5.5. Оптимизация роялти и устаревание данных.....	185
5.6. Организационные вопросы функционирования платформы	187
5.6.1. Конкуренция с провайдерами услуг и разработчиками моделей.....	188
5.6.2. Безопасность.....	190
5.6.3. Блокчейн как возможное техническое решение для библиотеки	191
5.7. Уточнённая архитектура.....	192
Выводы раздела	194
6. Мотивированные решения на макро- и мезоуровнях по данным микроуровня.....	196
6.1. Дифференцированный подход к реализации мероприятий	197
6.2. Управление на основе данных	209
6.3. Патентная аналитика и типология подходов к реагированию на заболевания	217

6.4. Согласование данных.....	230
6.5. Цена информационного продукта.....	243
Выводы раздела	250
Заключение	252
Список сокращений и условных обозначений	260
Литература	262

Введение

Актуальность темы исследования определяется, с одной стороны, объективными потребностями экономики, где все большую роль играют информационные технологии, продукты и услуги на их основе, с другой стороны, появлением новых возможностей для алгоритмизации и оптимизации сделок, включая количественные измерения и принятие решений. Эти обстоятельства принимаются во внимание на самом высоком уровне. Так, обозначенные приоритеты национального проекта «Экономика данных» включают разработку конкретных приложений, которые позволяют говорить о развитии рынка данных: госуправление на основе цифровых платформ (ГИС, ЭДО, ЦОДы) и искусственный интеллект. Причём эти приложения имеют явный экономический аспект, поскольку затрагивают производство данных, их обмен и распространение. Вопросы доступа к данным для развития технологий просматриваются в поручениях Президента по итогам конференции «Путешествие в мир искусственного интеллекта», прошедшей в 2022г. В частности:

- рекомендовано ускорить рассмотрение проекта федерального закона, устанавливающего порядок обезличивания персональных данных. Такой закон подписан Президентом 08.08.2024, и он устанавливает порядок обезличивания с 2025 г.;

- поручено обеспечить переход системы государственной власти на федеральном и региональном уровнях к модели управления на основе автоматического сбора и анализа данных с использованием информационных платформ;

- формируются открытые базы больших данных (дата-сетов) обезличенных медицинских данных пациентов в целях их использования для развития технологий искусственного интеллекта в здравоохранении;

- ожидаются предложения по вопросам создания индустриальных центров компетенций в этой сфере с формированием справочника лучших практик;

- ожидаются предложения о предоставлении отечественным разработчикам программного обеспечения доступа к обезличенным персональным данным.

Обсуждение вопросов, связанных с рынком данных, на самом высоком уровне демонстрирует их актуальность и важность. И тесная связь этих вопросов с технологиями искусственного интеллекта объясняется важностью данных. Она проявляется не только на стадии разработки систем, но и на стадии эксплуатации. Дело в том, что лучшие системы искусственного интеллекта предусматривают процедуру дообучения. В противном случае повышается риск попадания в ловушку дрейфа данных, когда работа обученной и зафиксированной нейронной сети перестаёт соответствовать меняющейся действительности. То есть, будучи созданной, система искусственного интеллекта постоянно сопровождается вопросом об обращении с данными, в том числе, об обеспечении доступа к новым наборам и исключения из моделей устаревших.

Однако в государственном управлении используются не только технологии искусственного интеллекта. В общем случае следует говорить о системах поддержки принятия решений, которые нуждаются в данных и инструментах их обработки. И говорить следует не только о рынке данных, но и связанном с ним рынке программ – готовых средств обработки типовых данных. За десятилетия разработано множество моделей для различных областей применения. На фоне растущего количества программных реализаций, развития методологии моделирования тех или иных объектов (государственных ведомств, коммерческих предприятий, индивидов, транспортных средств и т.д.) отмечается некоторое отставание в развитии инструментов для их распространения и повторного использования. Поэтому, учитывая высокую степень связи между данными и программами, их близкие свойства в неконкурентности потребления и возможности исключения, эти объекты можно рассматривать как части общего рынка информационных продуктов. Под информационным продуктом Естественными представляются вопросы ценообразования этих продуктов, соединения спроса и предложения. На развитие методологии

построения рынка информационных продуктов с фокусом на обозначенных вопросах направлено данное диссертационное исследование.

Таким образом, актуальность темы исследования определяется развитием информационных технологий, создающих все новые возможности для сбора, хранения и обработки больших массивов данных. Большие массивы данных, которые характерны для решения задач управления в современных технических и социально-экономических системах, могут эффективно использоваться только с применением вычислительной техники и технологий, требующих умения с ними обращаться. Следовательно, положение о совместимости данных с программами не только в техническом смысле, но и в смысле соответствия ожиданиям от повышения информированности принимаемых решений является существенным для исследования рынка данных.

Степень разработанности темы – сложный и многоплановый вопрос, так как необходимо различать три аспекта рассматриваемой темы. Первый из них связан с математическими подходами к измерению количества информации и её ценности. Второй аспект связан с использованием информации в экономике, третий – с развитием информатики, цифровых информационных технологий. При этом развитие математики обычно опережает её эффективное использование на несколько десятилетий. Развитие информационных технологий опережает их эффективное использование в экономике не так сильно, обычно в пределах десяти лет. В основном это связано с тем, что для эффективного использования новых информационных технологий они должны стать достаточно дешевыми, ими должно овладеть большое число людей, не являющихся специалистами в области информатики, но хорошо понимающих те сферы деятельности, где их применение дает большой экономический эффект. А потому говорить о степени разработанности темы имеет смысл применительно к каждому из трех перечисленных аспектов.

Говоря об измерении количества информации и её ценности, следует указать на работы К. Шеннона, Р. Хартли, А.Н. Колмогорова, А.Х. Шеня, А.А. Харкевича, Д.С. Чернавского, О.Д. Чернавской, Р. Ховарда, Р.Л.

Стратановича, Г. Кэстлера. В диссертационном исследовании особенный акцент сделан на алгоритмическом подходе к понятию «количества информации», введённому А.Н. Колмогоровым в 1965 году как альтернативы комбинаторному подходу Р. Хартли и вероятностному подходу Шеннона-Винера. Примечательно, что во всех трех подходах речь идет не о том, что такое информация, а о том, как измерять её количество, причем в конкретном контексте. Клод Шеннон решал задачи, связанные с передачей информации по каналам связи, позже он писал, что неправильно назвал свою теорию «теорией информации», надо было назвать её «теорией сигналов». Создатель науки кибернетики – Норберт Винер – акцентировал внимание на обратных связях и уменьшении энтропии. Между его подходом и подходом Шеннона имелись определенные различия. Но это не мешало А.Н. Колмогорову считать эти подходы эквивалентными. То же самое относится к последующим попыткам уточнить и усовершенствовать подход К. Шеннона – ничего существенного здесь не произошло.

Алгоритмический подход А.Н. Колмогорова к определению количества информации радикально отличается от комбинаторного и вероятностного подходов, как минимум, по двум существенным признакам. Во-первых, он явно связывает количество информации с её полезностью при решении конкретных задач. Так, карта местности содержит больше информации, когда она более полно и точно отражает особенности местности и расположенные на ней объекты, а не тогда, когда в ней больше пикселей. Во-вторых, описание и применение алгоритмического подхода многократно сложнее. Он связан с математической логикой, теорией алгоритмов, сложностью вычислений и не только. Его развитию посвящено много работ отечественных и зарубежных математиков, среди которых особо стоит упомянуть А.Х. Шеня и Н.К. Верещагина, которые много сделали для развития фрагментарно высказанных идей А.Н. Колмогорова по наполнению теории.

Об использовании информации в экономике можно говорить с нескольких позиций: принятия решений в условиях различной информированности,

цен на информацию и протоколов её передачи. Классическими работами по асимметрии информации являются работы Д. Акерлофа, К. Эрроу. В развитии методов принятия решений большую роль сыграли, Т. Саати, А.В. Андрейчиков, О.Н. Андрейчикова, Е.Г. Гольштейн, Д.Б. Юдин, Дж. Фон Нейман. Последний в этом ряду известен фундаментальными работами в области теории игр, которая предложила формальный аппарат для исследования ситуаций противоборства и кооперативного поведения. В этой же части следует отметить работы Л. Шепли и Дж. Нэша как важный математический аппарат для моделирования сделок и разделения общей стоимости. Работы Д.Б. Юдина и Е.Г. Гольштейна относятся, скорее, к ситуациям принятия решений в условиях определённости, хотя используемый ими аппарат допускает введение вероятностей в целевые функции. О работах по передаче информации (знаний) и протоколах таких сделок также следует говорить последовательно. В данном месте «знания» и «информация» используются как синонимы в силу того, что сами авторы, говоря о накоплении и использовании дополнительных сведений употребляли тот или иной термин. В контексте человеческой деятельности такое разделение является уместным. Однако возрастающая роль роботов – промышленных и программных – позволяет взглянуть на терминологию иначе. Х. Вэриэн отмечает, что «знания» и «информация» являются промежуточными степенями на пути от «данных» к «действию». Человек не может миновать ступень «знаний» на пути к «правильному» действию, а машина может. Учитывая сказанное, следует отметить работы Ф. Махлупа, В.Л. Макарова и Г.Б. Клейнера, Д. Детеринга, Г. Демсетца, Э. Томпсона, Э. Кларка, М. Бабайова, А. Смолина, А.Н. Козырева, Р. Коуза, Э. Линдаля, У. Ои, Р. Шмалензе, Х. Вэриэна. Работы перечисленных авторов, как элементы мозаики, вписываются в общую канву диссертационного исследования. С именем Махлупа связывают первенство в распознавании знаний как товара. В.Л. Макаров предложил формальную модель производства знаний и вместе с Г.Б. Клейнером систематизировал роль знаний в экономике. Работы Э. Линдаля, Г. Демсетца, Э. Томпсона, Э. Кларка и Д. Детеринга развивают модели рынков с общественным благом,

поставляемым в частном порядке, и исследуют способы назначения цен. М. Бабайов и А. Смолин внесли вклад в развитие протоколов передачи информации. А.Н. Козырев развил методы стоимостной оценки нематериального. У. Ои, Р. Шмалензе и Р. Коуз упоминаются в контексте двухкомпонентных цен. Х. Вэриэн может рассматриваться в этом ряду как персона, объединившая теоретические работы по экономике знаний, ценообразованию и информационным технологиям в практической деятельности. И эта же практическая деятельность имеет непосредственное отношение к третьему аспекту степени проработанности темы.

О степени разработанности темы исследования в аспекте информатики позволяют судить работы не только конкретных исследователей, но и целых коллективов – общественных организаций и фирм. Поддержку решений о социально-экономических системах осуществляют ситуационные центры. Имитационное моделирование является широко распространённым инструментом для реализации аналитических функций таких центров, и здесь следует отметить работы В.Л. Макарова, А.Р. Бахтизина и Е.Д. Сушко как авторов имитационных моделей для государственного управления, работы Н.И. Ильина по системе распределённых ситуационных центров. Программную реализацию платформы для обмена знаний долгие годы исследуют С.И. Паринов и М.Р. Когаловский. Значительный функционал современных аналитических систем связан с машинным анализом текстов, и в этой области значимый вклад внесли Т. Хоффман, Д. Блей, К.В. Воронцов. Наконец, в систематизации технологий, лучших практик их использования в экономике важная роль принадлежит организациям по стандартизации – как международной, так и отечественной. В документы попадают лишь общие формулировки, и значительный пласт практической деятельности закрыт для независимых исследователей ввиду коммерческого интереса компаний, которые широко внедряют в свою работу цифровые технологии. Среди них можно выделить гигантов ИТ и финансового сектора, и их главный результат состоит в соединении спроса и предложения посредством технологий. Понимание этого дают работы уже упомянутого Х.

Вэриэна, который известен не только как автор ряда экономических работ, но и как главный экономист одного из лидеров ИТ-сектора в мире.

Цель настоящего исследования – разработка методологии и набора конкретных методов, содействующих более широкому использованию существующих информационных продуктов (данных и программ) и созданию новых сервисов для поддержки процессов распространения таких продуктов.

Объектом исследования является рынок информационных продуктов (данных и программ) как совокупность отношений, связанных с оборотом данных и программного обеспечения.

Предметом исследования являются экономические механизмы для поддержки сделок на рынке информационных продуктов, в том числе предназначенные для соединения спроса и предложения, для назначения цен.

Соответствие темы диссертации требованиям Паспорта специальностей ВАК (по экономическим наукам). Тема и результаты диссертационной работы соответствуют следующим пунктам паспорта научной специальности 5.2.2 «Математические, статистические и инструментальные методы в экономике»: 1. Теоретические и методологические вопросы применения математических, статистических, эконометрических и инструментальных методов в экономических исследованиях; 2. Типы и виды экономико-математических и эконометрических моделей, методология их использования для анализа экономических процессов, объектов и систем; 4. Разработка и развитие математических и компьютерных моделей и инструментов анализа и оптимизации процессов принятия решений в экономических системах; 8. Оптимизационные модели в экономике; 9. Теоретико-игровые модели в экономических исследованиях; 12. Имитационное моделирование. Разработка и оценка имитационных моделей экономических процессов; 18. Развитие и применение инструментария проектирования, разработки и сопровождения информационных систем в интересах субъектов экономической деятельности.

Для достижения поставленной цели определены следующие **задачи исследования**:

1. Исследование роли информационных продуктов в экономике и связанных с ней предпосылок для формирования ценности в конкретных приложениях.
2. Исследование моделей рынка информационных продуктов.
3. Исследование механизмов ценообразования на рынке информационных продуктов.
4. Критический анализ доступных моделей и механизмов.
5. Исследование и критический анализ существующих подходов к реализации платформенных решений для поддержки функционирования рынка информационных продуктов.
6. Формирование общей методологии для проектирования платформенных решений для рынка информационных продуктов.
7. Апробация отдельных элементов методологии на примере создания информационного продукта и его стоимостной оценки.

Теоретическая ценность работы заключается в совершенствовании научно-методологических основ управления нематериальными ценностями на основе алгоритмического подхода, применения принципа двойственности, привнесенного Л. В. Канторовичем из функционального анализа в экономику, к оптимизации условий договоров о передаче ноу-хау и других договоров лицензионного типа.

Практическая значимость полученных результатов заключается в возможности применения созданной методологии для разработки новых информационных сервисов, прогнозирования спроса на наукоемкую продукцию, поддержки сделок по передаче информационных продуктов, в том числе, в части согласования цены сделки. Полученные методические результаты ложатся в основу инструментов разделения стоимости информационных продуктов между их компонентами – данными и компьютерными программами. Это позволяет, в частности, уточнять размер компенсации при нарушении исключительных прав, а также выполнять стоимостную оценку неторгуемых на рынке активов.

Практическая значимость работы подтверждается успешным применением предлагаемых методов для решения задач ценообразования в сделках реального сектора экономики. Полученные результаты также имеют потенциал для широкого применения ввиду увеличения общественных излишков благодаря дифференцированному ценообразованию.

Методика проектирования платформы для поддержки рынка данных может быть использована для реализации конкретных инициатив при выполнении национального проекта «Экономика данных» Российской Федерации. Методика, однако, фокусируется на экономических механизмах. В том числе, в части расчёта цены сделки. За рамками диссертационной работы остаются задачи, связанные с техникой и технологиями. К их числу относится, например, трудоёмкая задача разработки онтологии, которая поспособствовала бы соединению спроса и предложения на рынке данных.

Методология и методы исследования

В работе используются математические методы, включая элементы теории дележей (решение по Шепли), методы оптимизации, статистические и эконометрические методы. Для анализа массивов документов использовались электронные библиотеки со встроенными аналитическими возможностями (библиометрические и патентные информационные системы).

Использование математического аппарата и инструментальных средств для поддержки принятия решений развиты в работах Андрейчикова, Андрейчиковой, Макаровым, Бахтизиным, Машковой.

Для демонстрации роли дополнительной информации в снижении ошибок принятия решений и связанным с ней развитием отрасли знаний построена модель отраслевой динамики. В её основе лежит методика, развивающая идеи самосогласованного поля, изложенная в работах И.Д. Грачева.

Значительный по своей важности для диссертационного исследования блок результатов посвящён ценообразованию на информационные продукты. Отправной точкой для него послужили модели стоимостной оценки, в том числе, с использованием методов оптимизации развиты в работах Кларка,

Детеринга, Козырева, Бабайова. Теоретическим обоснованием для них служат работы Нэша, Шепли и Коуза.

В качестве информационной базы исследования использованы:

- библиометрические базы и электронные научные библиотеки для понимания состояния и тенденций развития научных направлений, поиска и анализа массивов научной литературы;
- патентные библиотеки как источник данных для формирования представления о тенденциях в прикладных разработках по теме исследования;
- нормативно-правовая база Российской Федерации для формирования представления о приоритетах научно-технического развития, отраслевых стандартов разработки инструментальных средств;
- результаты социально-экономических наблюдений как источник данных для демонстрации возможностей применения инструментальных средств для выявления функциональных зависимостей и их использования в задачах принятия решений.

Научная новизна диссертационного исследования состоит в создании методологической основы для создания платформенных решений для поддержки рынка информационных продуктов путём сочетания программных сервисов и экономических механизмов, которые содействуют соединению спроса и предложения. При выполнении диссертационного исследования получены следующие новые научные результаты.

1. Построена и исследована экономико-математическая модель для анализа совместной динамики отраслей экономики с выделенной отраслью производства знаний. В классических публикациях анализируется выгода от вложений в науку и технологии через изменение коэффициентов производственной функции. В предложенной модели отрасли могут детализироваться до отдельной фирмы, и их решения состоят в выборе размера инвестиций в производство знаний. Последние, в свою очередь, влияют на доступ к ресурсам (как отдельной отрасли или фирмы, так и экономики в целом), а также на снижение ошибок принятия решений.

2. В экономический дискурс введена колмогоровская сложность как объективная мера для различения потребителей и поставщиков на рынке информационных продуктов. Ценовая дискриминация на рынке информационных продуктов является известным теоретически обоснованным результатом. Для различения потребителей – на практике и в теоретических моделях – опираются на типы. При этом возникает проблема вычисления данного типа. Колмогоровская сложность даёт объективную меру для 1) дифференциации программ обработки данных и 2) дифференциации потребителей на рынке данных по тому, доступом к какой программе они обладают. В случае отсутствия ограничения на ресурсы, такая типологизация является однозначной в том смысле, что существует «наилучшая» программа для обработки данных. В случае колмогоровской сложности с ограничением на ресурсы оказывается, что пользователи разных программ могут быть отнесены к одному типу, если они обрабатывают строку одинаковым образом. Типологизация по колмогоровской сложности имеет применение к борьбе с арбитражем на рынке программных продуктов, поскольку учитывает состояние строки после обработки – именно то, для чего приобретаются данные на рынке информационных продуктов.

3. Предложена модель рынка данных, поддерживаемая цифровой платформой с онтологическим представлением цифровых объектов, обеспечивающая соединение спроса и предложения с использованием ценовых алгоритмов. Полных аналогов платформ, которые позволили бы соединить спрос и предложение на рынке данных, не обнаружено. Некоторым приближением могут служить электронные библиотеки – публикаций, программного обеспечения, поскольку помогают распространять информационные продукты, поддерживают их повторное использование. Однако при соединении спроса и предложения используются рекомендательные сервисы, основанные на профайлинге пользователей (анализе профиля пользователя по индивидуальным характеристикам и истории активности на платформе). Также электронные библиотеки не предлагают сервисы для оптимизации цены сделки. При анализе предложенной платформы исследованы подходы к построению сервисов для

хранения, обработки цифровых объектов и их распространения, учитывающие правовые ограничения и технические возможности.

4. Путём обобщения известной процедуры назначения цен за информацию на случай повторяющихся взаимодействий развит математический аппарат — основа сервиса ценообразования на цифровой платформе. Существует протокол продажи информации с использованием информирующих сигналов для потребителей разных типов. Теоретически обосновано, что при однократном взаимодействии этот протокол, который полагается на ценовую дискриминацию, позволяет полностью изъять потребительский излишек при соблюдении условий индивидуальной рациональности и совместимости со стимулами, исключая, таким образом, возможность арбитража со стороны потребителя. На рынке данных возможны ситуации повторяющихся покупок с некоторой периодичностью, что позволяет потребителю актуализировать доступный ему аналитический сервис. В этой связи предложена модификация условий индивидуальной рациональности и совместимости со стимулами в протоколе передачи информации для расширения на случай повторяющихся взаимодействий.

5. Формализована математическая процедура оптимизации двухкомпонентных цен на рынке данных с использованием алгоритма достижения Парето-оптимальных состояний в задаче разделения ресурсов. Цена, назначаемая протоколом передачи информации из предыдущего пункта, может выступать параметром функции, описывающей двухкомпонентные цены (тарифы). Такие цены, с одной стороны, теоретически оправданы в случае возрастающей отдачи, к которому относятся высокие технологии, и, с другой стороны, позволяют вычислить равновесное для рынка состояние. Известный алгоритм распределения ресурсов для достижения Парето-оптимального состояния, который автором уже применялся в контексте лицензионной торговли, адаптирован к двухкомпонентным ценам.

6. Предложена процедура стоимостной оценки неторгуемого на рынке нематериального актива, основанная на аппарате теории игр. Потребитель на

рынке данных обладает средствами их обработки – той или иной программой, которая характеризуется колмогоровской сложностью. Цены на данные можно наблюдать на рынке. Программа является уникальным активом, стоимость которого заранее не известна, но может быть полезна для ряда задач: использования для целей амортизации, вклада в уставный капитал вновь создаваемого общества, использования в качестве залога. Стоимость такой программы устанавливается по известной цене данных с поправкой на синергетические эффекты по известной процедуре расчёта компонентов Вектора Шепли.

7. Теоретические положения о построении рынка данных приложены к примеру разработки многокомпонентной системы принятия решений, которая предполагает несколько уровней обработки данных. Модельная ситуация о выработке политики в области здравоохранения предполагает использование разнообразного программного инструментария и неоднородных данных. И инструменты, и данные характеризуются различными условиями доступа: одни являются открытыми и бесплатными, другие оцениваются по затратам, третьи – по рыночным ценам. Все эти компоненты важны для корректности работы верхнеуровневой агентной модели, которая представляется неторгуемым на рынке активом. Стоимость компьютерной программы, реализующей модель, оценивается по описанной выше процедуре.

Положения, выносимые на защиту

1. Модель самосогласованного поля позволяет интерпретировать вложения в развитие имеющихся и получение новых знаний в терминах обмена доступных благ на сведения и технологии, которые способствуют расширению доступных отраслям ресурсов. (Подраздел 1.1.)
2. Концепция колмогоровской сложности в экономических моделях с дифференцированным ценообразованием на информацию может быть использована для параметризации товаров и потребителей на рынке информационных продуктов. (Подразделы 2.6.)

3. Цифровая платформа с ценовыми сервисами, онтологией представления данных и механизмами контроля качества данных позволяют автоматизировать сделки на рынке данных. (Раздел 5)
4. Обобщение известной процедуры построения ценового меню на рынке данных позволяет построить вычислить цены на отдельные срезы данных при их покупке частями при повторяющихся взаимодействиях (например, по подписке). (Подраздел 2.6.)
5. Математическая процедура оптимизации использования ресурсов может быть адаптирована к случаю двухкомпонентных цен на рынке данных. (Подраздел 3.3.)
6. С привлечением концепции колмогоровской сложности к процедуре определения выигрыша участников кооперативной игры возможно построение метода стоимостной оценки неторгуемого на рынке нематериального актива. (Подразделы 4.2. и 4.3.)
7. Совместное применение процедур дифференцированного ценообразования на рынке данных и стоимостной оценки неторгуемого на рынке нематериального актива позволяет оценить стоимость разработки систем поддержки принятия решений. (Подраздел 6.5.)

Достоверность и обоснованность полученных результатов обеспечивается опорой на формализованный математический аппарат, логическую последовательность выводов. Результаты – как сторонние, взятые в качестве отправной точки и привлекаемые к рассуждениям по мере необходимости строгого изложения, так и авторские – опубликованы в рецензируемых изданиях. Эмпирические данные и наблюдения получены в процессе научно-исследовательской работы и взаимодействия с представителями реального сектора экономики в ходе решения конкретных прикладных задач. Полученные результаты получили подтверждение в ходе численной проверки на компьютерных моделях, а также в условиях реальной экономики при сопровождении сделок на микроэкономическом уровне.

Апробация и внедрение результатов исследования

Результаты исследования прошли апробацию в виде научного обсуждения на научных семинарах ЦЭМИ РАН, следующих международных и Всероссийских научных конференциях:

- 48-е заседание международной научной школы-семинара «Системное моделирование социально-экономических процессов» имени академика С.С. Шаталина, Воронеж, Россия, 8-10 октября 2025
- IV международная научно-практическая конференция «Развитие теории и механизмов повышения устойчивости, инновационности и конкурентоспособности пространственного развития экономики регионов» (SICSEDR-25), г. Санкт-Петербург, Россия, 5-6 июня 2025
- VI Международный научный форум «Шаг в будущее: искусственный интеллект и цифровая экономика», г. Москва, Россия, 21 марта 2024
- 43-е заседание международной научной школы-семинара «Системное моделирование социально-экономических процессов» имени академика С.С. Шаталина, Воронеж, Россия, 13-18 октября 2020
- 42-е заседание международной научной школы-семинара «Системное моделирование социально-экономических процессов» имени академика С.С. Шаталина, Ростов-на-Дону, Россия, 1-6 октября 2019
- V САНКТ-ПЕТЕРБУРГСКИЙ МЕЖДУНАРОДНЫЙ ЭКОНОМИЧЕСКИЙ КОНГРЕСС (СПЭК-2019) «ФОРСАЙТ «РОССИЯ»: БУДУЩЕЕ ТЕХНОЛОГИЙ, ЭКОНОМИКИ И ЧЕЛОВЕКА», Санкт-Петербург, Россия, 3 апреля 2019
- IV САНКТ-ПЕТЕРБУРГСКИЙ МЕЖДУНАРОДНЫЙ ЭКОНОМИЧЕСКИЙ КОНГРЕСС (СПЭК-2018) «ФОРСАЙТ «РОССИЯ»:

НОВОЕ ИНДУСТРИАЛЬНОЕ ОБЩЕСТВО. БУДУЩЕЕ», Санкт-Петербург, Россия, 2-3 апреля 2018

- 13th International Conference on Current Research Information Systems (CRIS-2016), Сен-Эндрюс, Великобритания, 9-11 июня 2016
- IX Международный Форум «Интеллектуальная собственность - XXI век», Москва, Россия, 25-29 апреля 2016
- Второй Санкт-Петербургский международный экономический конгресс (СПЭК-2016), Санкт-Петербург, Россия, 22-23 марта 2016
- euroCRIS Membership Meeting, Париж, Франция, 10-13 мая 2015
- 12th International Conference of Current Research Information Systems (CRIS-2014), Рим, Италия, 13-15 мая 2014

Результаты исследования используются в материалах курса «Экономика интеллектуальной собственности» для студентов Экономического факультета ГАУГН. Результаты исследования также использованы в деятельности организаций («Вертолеты России», «Рособоронэкспорт»).

Публикации.

Результаты исследования ранее представлены в печатных работах в рецензируемых научных изданиях. Всего по теме исследования опубликовано 31 работа общим объёмом 60,4 печатных листов (авторский вклад составляет 22 печатных листа), в том числе 6 работ (общим объёмом 7,7 п.л., авторский вклад – 2,7 п.л.) в изданиях, индексируемых в системе Web of Science или Scopus, 17 работ из Перечня ВАК (общим объёмом 20,7 п.л., авторский вклад – 13,8), 10 (общим объёмом 14,9 п.л., авторский вклад – 12,6) из которых опубликованы в журналах категории К-2.

Структура диссертации. Диссертация состоит из Введения, шести разделов, Заключения, Списка сокращений и условных обозначений, списка литературы.

В первом разделе обосновывается привлечение алгоритмического подхода к определению количества информации для описания процессов,

связанных с приобретением дополнительных сведений в контексте принятия решений. Второй раздел излагает теоретические основы ценообразования на фрагменты сведений – наборы данных – и протоколы их передачи. В третьем разделе раскрываются механизмы двухкомпонентного ценообразования, в том числе, инструмент улучшения цены сделки. Четвёртый раздел обращается к проблеме разделения стоимости сложного, многокомпонентного информационного продукта между его составляющими. В пятом разделе обосновывается методический подход к построению цифровой платформы для поддержки рынка данных. Наконец, шестой раздел на конкретном примере иллюстрирует разработку многокомпонентного информационного продукта – неторгуемого на рынке актива – и разделение стоимости между его составляющими.

Общий объём работы составляет 288 страниц. Список литературы включает 260 источников.

1. Понимание ценности информации в экономической науке

1.1. Расширение доступа к ресурсам через знания

Тематика важности знаний для экономики столь обширна, что необходимо по возможности подробнейшим образом задать контекст знаний в данном диссертационном исследовании. На многообразии функций знаний в экономике указали В.Л. Макаров и Г.Б. Клейнер, и принятую ими классификацию следует выбрать в качестве отправной точки к описанию контекста. Этими функциями являются [195]:

- «непосредственный продукт деятельности;
- предмет непосредственного конечного потребления;
- производственный ресурс, используемый в процессе производства продукции;
- предмет и средство распределения и/ или рыночных транзакций;
- средство транзакций;
- орудие или инструмент управления;
- средство консолидации общества и воспроизводства общественных институтов.»

В меньшей степени данное исследование касается последней функции, хотя отдельные её аспекты в части рекламы, искажения информации и влияния на общественные нормы затрагиваются некоторыми разделами. Главные вопросы исследования связаны с ролью знаний в расширении возможностей управления и снижении ошибок принятия решений. Поэтому именно в таком контексте следует начать обзор работ о важности знаний в экономике.

Вероятно, Ф. Хайек в числе первых выразил мнение о том, что рынок в целом является инструментом, который «обрабатывает» распределённые в обществе знания наилучшим образом, обеспечивая быструю адаптацию общества к меняющимся условиям [44]. Причём, речь идёт не только об информации, которая поддаётся статистическому учёту. Автор говорит именно о всём доступном обществу и распределённом между отдельными лицами знании. И ключевым механизмом в этом инструменте являются цены. Не углубляясь в

спор о преимуществах рынка перед центральным планированием, который красной нитью проходит через цитируемую работу Хайека, следует отметить тезис о том, что знание является распределённым в обществе ресурсом, и для его обработки необходим некоторый институт (для Хайека таким институтом является рынок).

Понимание знания как ресурса и одновременно продукта в экономике расширил Ф. Махлуп. Его результат связан с измерением и классификацией секторов экономики, производящих и распространяющих знание [70]. Значимость этого шага видна по тому, как знания стали учитываться в моделях экономического роста. Если Солоу использовал лишь некоторый параметр роста технологической переменной [107], то последовавшие за статьёй Махлупа работы Ромера [91], Лукаса [68], Агийона [3] явным образом идентифицируют инвестиции именно в знания и человеческий капитал как в форму воплощения знаний. На отличительные характеристики знаний как товаров ясно указали в своих работах К. Эрроу, И. Нонака и Х. Такеучи [6, 81], которые обратили внимание на экстерналии, проблемы распространения знаний как по причинам недофинансирования, так и из-за неформализуемости личностных (неявных) знаний. Наконец, Э. Брюнйольфссон и Л. Хитт [15] продемонстрировали важность сочетания инвестиций в ИТ и в человеческий капитал с развитием организационных инноваций. Важность ИТ в контексте диссертационного исследования объясняется возрастающими данными и повышением сложности¹ инструментов их обработки.

Всё сказанное позволяет модельным образом продемонстрировать многогранную роль знаний и инноваций (то есть новинок, которые позволяют получить коммерческую ценность) в экономическом росте. Во-первых, они создают новые рынки. Во-вторых, открывают доступ к новым ресурсам. В-третьих, повышают эффективность сложившихся отраслей и предприятий.

¹ Поскольку далее в работе «сложность» приобретает ещё один оттенок в разговоре об измерении информации, здесь уместно сделать оговорку о сложности программного обеспечения. В данном конкретном предположении подразумевается знание математического аппарата обработки данных и многообразия используемых программных библиотек, сочетание которых обеспечивают корректные вычисления.

Взаимодействие этих эффектов позволяет исследовать модель перетока ресурсов [154, 218, 155].

Пусть в модели с дискретным временем существует рынок товаров и услуг. Его динамика описывается вектором-столбцом \bar{A}_i , компоненты которого отражают выпуск отдельных фирм в момент времени i . Вместе с этим рынком существует сектор исследований и разработок, выход которого описывается вектором-столбцом \bar{B}_i . Причём каждая компонента отражает результат каждого отдельного института или конструкторского бюро в момент времени i . Источником ресурсов для экономики является «природа», чьи запасы в момент времени i представлены скаляром Π_i . Каждая фирма, каждый институт или конструкторское бюро предъявляют свои требования к ресурсам, и оценки этих ресурсов описываются векторами $\bar{\xi}$ и $\bar{\eta}$, соответственно. Если ввести предельную себестоимость ресурсов «природы» μ , то динамику системы можно описать системой уравнений:

$$\bar{A}_{i+1} = \bar{A}_i - \text{diag}(\bar{\xi}) \times \bar{A}_i + \left\langle \frac{\bar{A}_i^T \times \bar{\xi} + \Pi_i \times \mu}{\bar{A}_i^T \times \bar{I} + \Pi_i} \right\rangle \times \bar{A}_i \quad (1.1)$$

$$\Pi_{i+1} = \Pi_i - \mu \Pi_i + \left\langle \frac{\bar{A}_i^T \times \bar{\xi} + \Pi_i \times \mu}{\bar{A}_i^T \times \bar{I} + \Pi_i} \right\rangle \times \Pi_i \quad (1.2)$$

$$\bar{B}_{i+1} = \bar{B}_i + \text{diag}(\bar{\eta}) \times \bar{B}_i - \left\langle \frac{\bar{B}_i^T \times \bar{\eta}}{\bar{B}_i^T \times \bar{I}} \right\rangle \times \bar{B}_i \quad (1.3),$$

Второе слагаемое в формулах (1.1)-(1.3) характеризует индивидуальную эффективность каждого из экономических агентов в использовании ресурсов. Третье слагаемое характеризует общую эффективность рынка в использовании ресурсов относительно идеального случая. Таким образом, из текущего состояния рынка – товаров и услуг, интеллектуальных разработок – определяют общие возможности участников, и далее, каждый, распоряжаясь доступными ресурсами, обеспечивает приращение (или убывание) своей части. Вместе последние два слагаемых для каждого отдельного участника описывают следующую закономерность: действуешь в среднем лучше рынка – выигрываешь, хуже – проигрываешь.

Для моделирования необходимо обеспечить связь секторов экономики с результатами НИОКР и ресурсами «природы». Будем считать, что после каждого периода из экономики изымается часть продукта для вложения в исследования и разработки – $\gamma \times \Delta(\bar{A}^T \times \bar{I}) = \gamma \times (\bar{A}_i^T \times \bar{I} - \bar{A}_{i-1}^T \times \bar{I})$. В свою очередь, исследования и разработки увеличивают объём доступных ресурсов. При этом, однако, в экономику могут возвращаться не все результаты, а лишь некоторая их часть – $\theta \times \Delta(\bar{B}^T \times \bar{I}) = \theta \times (\bar{B}_i^T \times \bar{I} - \bar{B}_{i-1}^T \times \bar{I})$. Заметим, что объём знаний не сокращается при передаче результатов НИОКР, в отличие от капитала секторов экономики. Тогда формулы (1.1)-(1.3) примут вид:

$$\bar{A}_{i+1} = \bar{A}_i - \text{diag}(\bar{\xi}) \times \bar{A}_i + \left\langle \frac{\bar{A}_i^T \times \text{diag}(\bar{\xi}) + \Pi_i \times \mu}{\bar{A}_i^T \times \bar{I} + \Pi_i} \right\rangle \times \bar{A}_i - \frac{\gamma \times \Delta(\bar{A}^T \times \bar{I})}{\bar{A}_i^T \times \bar{I}} \times \bar{A}_i \quad (1.4)$$

$$\Pi_{i+1} = \Pi_i - \mu \Pi_i + \left\langle \frac{\bar{A}_i^T \times \text{diag}(\bar{\xi}) + \Pi_i \times \mu}{\bar{A}_i^T \times \bar{I} + \Pi_i} \right\rangle \times \Pi_i + \theta \times \Delta(\bar{B}^T \times \bar{I}) \quad (1.5)$$

$$\bar{B}_{i+1} = \bar{B}_i + \text{diag}(\bar{\eta}) \times \bar{B}_i - \left\langle \frac{\bar{B}_i^T \times \bar{\eta}}{\bar{B}_i^T \times \bar{I}} \right\rangle \times \bar{B}_i + \frac{\gamma \times \Delta(\bar{A}^T \times \bar{I})}{\bar{B}_i^T \times \bar{I}} \times \bar{B}_i \quad (1.6).$$

Динамику системы можно проиллюстрировать числовым примером. Пусть на каждом рынке – реальной экономики с вектором \bar{A}_i и сектором НИОКР с вектором \bar{B}_i – отслеживается динамика всего 200 экономических агентов (по 100 в каждом секторе). В таком упрощённом случае векторы $\bar{\xi}$ и $\bar{\eta}$ могут интерпретироваться как ошибки оценивания, которые возникают, в том числе, по причине искажения сигнальной функции патентов. Пусть в начальный момент времени компоненты векторов $\bar{\xi}$ и $\bar{\eta}$ равномерно распределены от 0,01 до 0,1 и от 0,001 до 0,01, соответственно. Компоненты векторов \bar{A}_0 и \bar{B}_0 распределены на отрезках от 0,1 до 1 и от 0,01 до 0,1, соответственно, $\theta = 0,9$, $\gamma = 0,1$, Общие запасы «природы» $\Pi_0 = 100$. Компоненты векторов $\bar{\xi}$ и $\bar{\eta}$ не являются постоянными в ходе численного эксперимента, а распределены по логистическому закону.

Типичные результаты моделирования представлены рисунками 1.1-1.4. Графики демонстрируют базовый сценарий (рисунки 1.1-1.2) и вариант с повышенной отдачей от сектора НИОКР (рисунки 1.3-1.4). Последний отличается от базового значением одного единственного параметра – коэффициент θ

для него в 5 раз больше. В таком случае сценарии иллюстрируют динамику двух систем с одинаковым выходом результатов, но различным процентом их использования в экономике.

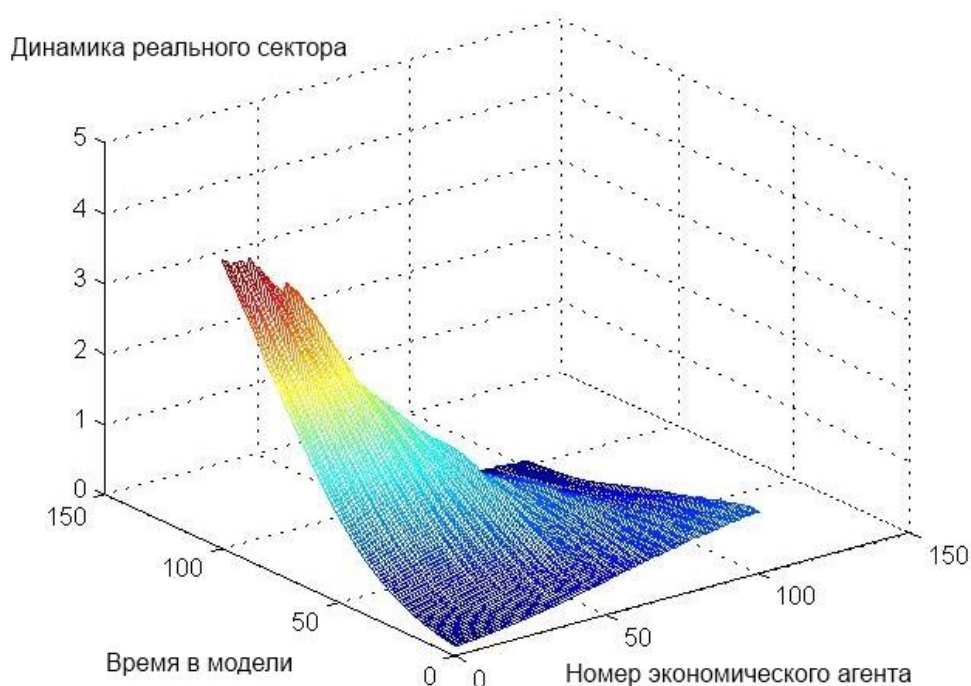


Рисунок 1.1. Динамика капитала агентов реального сектора в базовом сценарии. Источник: [154]

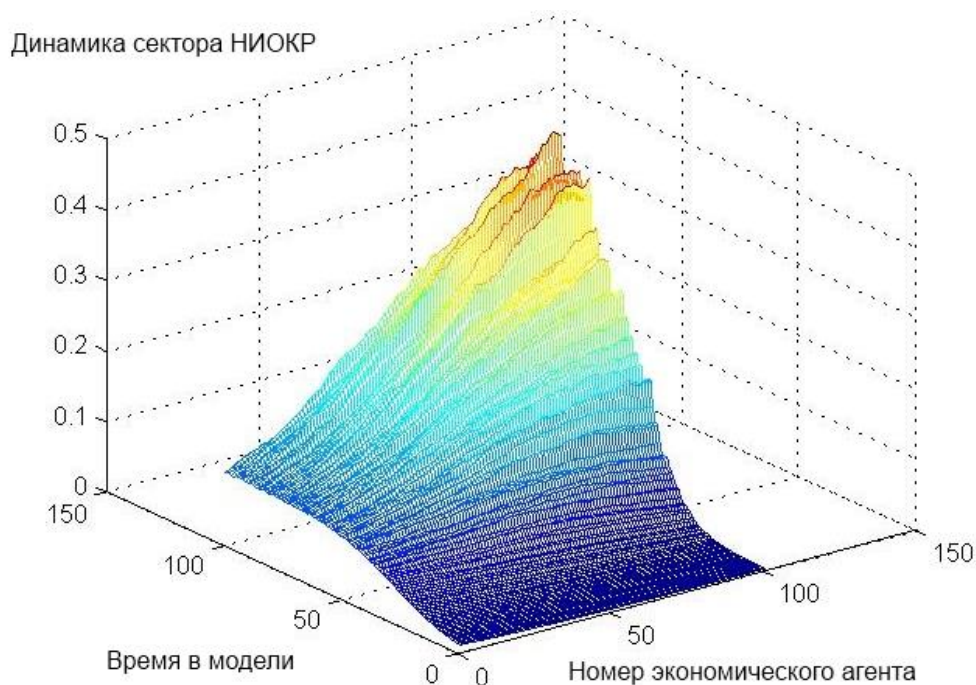


Рисунок 1.2. Динамика капитала агентов сектора НИОКР в базовом сценарии. Источник: [154]

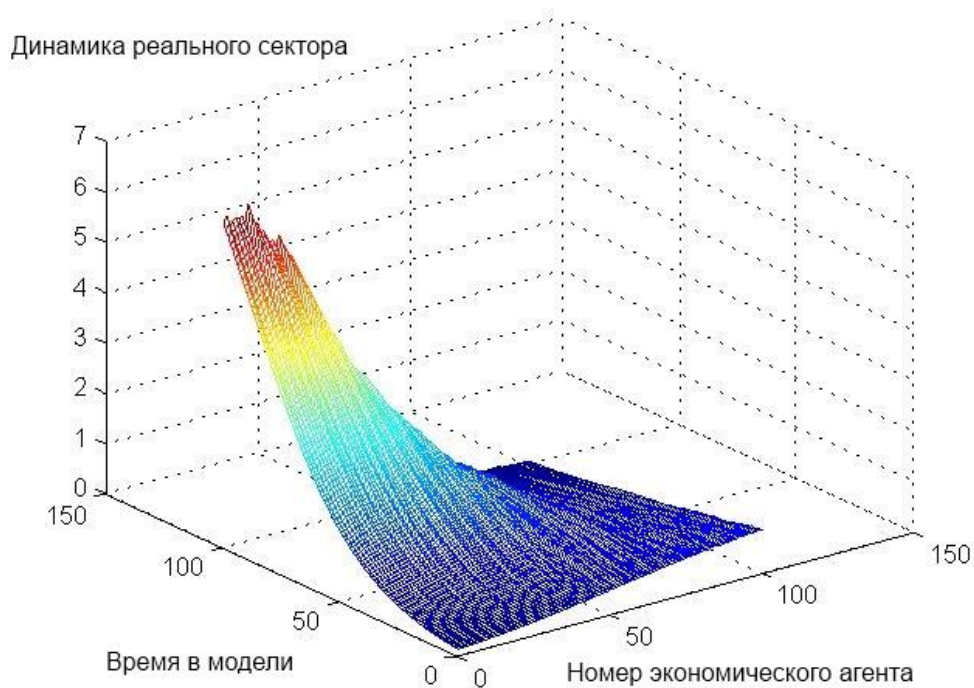


Рисунок 1.3. Динамика капитала агентов реального сектора в сценарии повышенной отдачи сектора НИОКР. Источник: [154]

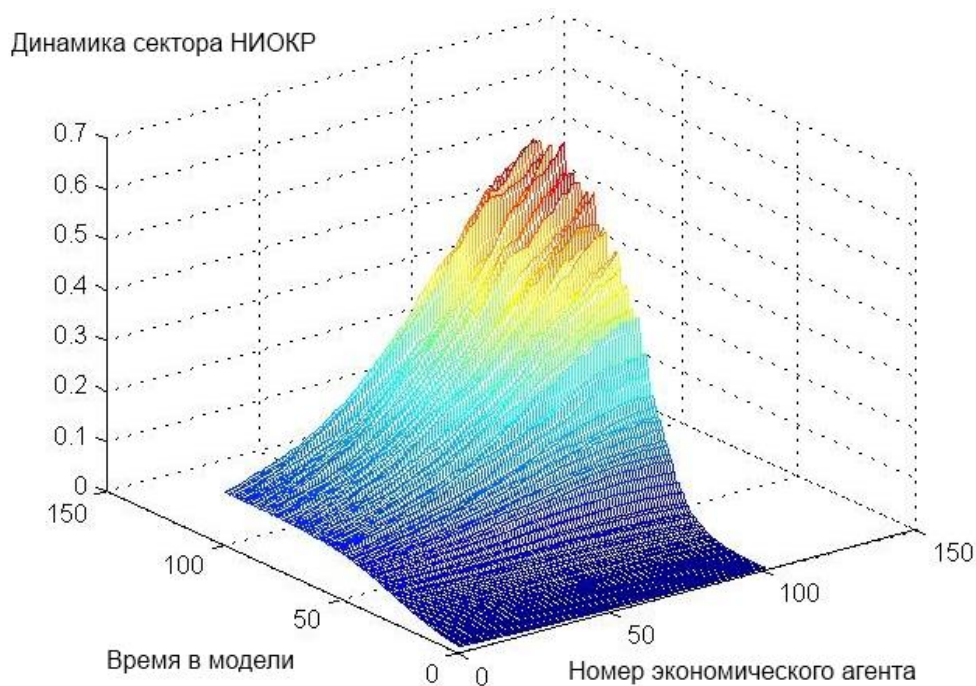


Рисунок 1.4. Динамика капитала агентов сектора НИОКР в сценарии повышенной отдачи сектора НИОКР. Источник: [154]

Те фирмы – экономические агенты, которые лучше приспособлены к внедрению инноваций и активно их используют, добиваются лучших результатов. Этот результат соответствует уже опубликованным работам. Например, выводам Д. Аджемоглу и Дж. Робинсона о том, что неравенство в доступе к

знаниям и инновациям создаёт различия в росте [129]. И это различие является одной из предпосылок в теории технологической дистанции и моделях диффузии инноваций.

Следует отметить, что внедрение знаний в экономику – дело не одного лишь бизнеса. В этом процессе высока и роль научно-исследовательских организаций [163, 173]. Этой стороне вопроса уделяется пристальное внимание, и, вероятно, дополнительным стимулом к исследованиям в области коммерциализации технологичных разработок стали показатели реформы США в части обращения университетов со своими результатами, связанной с именами Б. Бэя и Б. Доула [67]. В России, однако, сохраняется проблема передачи технологий из научного сектора в промышленный, о чём свидетельствует результативность малых инновационных предприятий, созданных с участием университетов и научных институтов [250].

Хотя уравнения (1.4)-(1.6) интерпретируются в терминах предприятий, они могут быть адаптированы к анализу эффектов от вложений в экономику знаний на макроскопическом уровне. Представление отраслей в агрегированном виде – через скаляр её выпуска, а не вектор выпуска отдельных предприятий – позволяет использовать метод межотраслевого анализа для анализа экономических процессов. Так, коэффициенты технологической матрицы могут использоваться для вычисления масштабов перетока ресурсов между отраслями [148]. А таблицы «затраты-выпуск» уже адаптированы к исследованиям цен на промежуточные продукты, в том числе, в контексте шоковых воздействий на экономическое развитие [207, 144]. Современные научные исследования критически зависят от высокотехнологичного оборудования, и торговые ограничения, например, существенно сказываются на доступности приборной базы [247]. Сказанное предметно иллюстрирует связь научных исследований с продукцией других отраслей экономики и важность вопроса ресурсного обеспечения. Таким образом, при определении масштаба перетока ресурсов в сектор «исследования и разработки» необходимость учёта цен в отраслях – поставщиках промежуточной продукции не вызывает сомнений. И развитие

предложенного в данном разделе инструментария в направлении анализа макроэкономических эффектов может заключаться в соединении с моделями реакции отраслевых цен на шоковые факторы, подобные уже указанной [144].

Продемонстрированная математической моделью зависимость результативности фирм от доступных им знаний является первым слоем. Да, принимаемые решения зависят от доступных знаний. Но следующий слой заключается в способах обращения со знанием. Он замечен в работе Брюнйольфссон и Хитт, которые указали на важность сочетания факторов [15]. В действительности экономические агенты различаются не только в своих представлениях об объектах, субъектах, системе или среде, с которыми сталкиваются по роду своей деятельности, но также способами принятия решений, и это тоже часть знания. Причём эти способы тесно связаны с накопленными сведениями, и это демонстрируется следующим подразделом.

1.2. Управление, анализ сценариев и ситуационное моделирование

Совершенствование государственного управления на основе цифровых технологий – одно из актуальных направлений, в котором разрабатываются и принимаются к реализации специальные программы развития. В период 2019-2024 гг. они укладывались в национальный проект «Цифровая экономика», в период 2025-2030 гг. – в национальный проект «Экономика данных и цифровая трансформация государства». Цели программ не ограничиваются расширением списка государственных услуг, предоставляемых в электронном виде. В подробностях вопросы совершенствования государственного управления обсуждаются на специализированных конференциях и совещаниях, материалы которых ёмко и содержательно формулируют суть решаемых проблем, методы решения и возможные ограничения для реализации решений. Изложение этого раздела ведётся по ранее материалам обзора одного из таких мероприятий, опубликованного автором данной диссертационной работы [209]. По мере необходимости это изложение дополняется комментариями и ссылками на дополнительную литературу.

Комментируя актуальную практику реализации программ цифровизации, С.Д. Бодрунов отмечает следующие тенденции в цифровизации государственного управления: прозрачность расходования бюджетных средств и рациональное использование ресурсов. А.В. Подойницын предлагает следующее направление в развитии этих тенденций: перейти от прогнозирования показателей к их планированию с опорой на экономико-математические модели и суперкомпьютерные вычисления. При этом, однако, подразумевается не тотальное планирование в экономике – только индикативное, внедрение в практику государственного управления проектного планирования. При таком подходе планирование в государственном управлении, по мнению А.В. Подойницына, выглядит следующим образом.

На первом этапе устанавливаются значения целевых показателей. На втором этапе формируются планы и задачи для достижения целевых показателей. На третьем этапе при известных межотраслевых связях оцениваются ресурсы и выпуски отраслей экономики для достижения целей. Собственно, на втором и третьем этапе конкретные проекты обретают свои контуры. Наконец, под проекты выделяются необходимые ресурсы, в том числе, осуществляется эмиссия денежных средств в объеме, необходимом для реализации проектов. В такой схеме естественным образом возникает задача для суперкомпьютеров – задача обратного счета, перехода от целей к экономическим задачам через межотраслевой баланс.

При этом отмечено, что работоспособность такого подхода зависит от организационных решений. В том числе, необходим государственный контроль над эмиссией, низкие процентные ставки для кредитования проектов, федеральный орган исполнительной власти для макроэкономического планирования, стабильный валютный курс.

Такая постановка вопроса об оптимальном использовании ресурсов и широком применении вычислительной техники в госуправлении концептуально близка идеям ОГАС (Общегосударственной автоматизированной системы учёта и обработки информации), которые неразрывно связаны с

именами А.И. Китова и В.М. Глушкова [192]. Параллель между звучащими сегодня предложениями и упомянутой системой находит своё продолжение в распределённом характере систем планирования и вычислительной инфраструктуры. Звучащие сегодня голоса аргументируют использование искусственных обществ и компьютерного моделирования при анализе траекторий перехода между состояниями реальных социально-экономических систем. Такие инструменты позволяют ответить на следующие вопросы: «Какие меры и в какой последовательности наиболее эффективны для перехода к новому образованию, новой науке, новому производству? Как корректировать отклонения перехода? В каких метриках его отслеживать?». Суперкомпьютерные технологии возникают в моделировании социально-экономических процессов естественным образом, однако, в списке высокопроизводительных систем top500 преобладает отечественный коммерческий сектор, из научных – только суперкомпьютер МГУ им М. В. Ломоносова. На фоне высокой стоимости систем и вопроса их загрузки также выдвигается тезис о приоритете задачи аудита и эффективного использования созданных в стране суперкомпьютеров перед задачей приобретения новых вычислительных мощностей. И в этом месте вновь возникает пересечение с ОГАС и ВЦКП (Вычислительного центра коллективного пользования). Именно их опыт распределения общих ресурсов позволяет утверждать, что эта задача – совместное использование вычислительной мощности имеющихся в стране суперкомпьютеров – может быть решена и заслуживает самого пристального внимания на уровне государственной власти, особенно при существующих ограничениях в поставках вычислительной техники. Накопленный опыт могут дополнить актуальные разработки о координации в использовании общих ресурсов [136].

Концепция государственного управления в части контроля и принятия решений на различных уровнях власти с широким применением информационно-аналитических систем также предполагает активное использование ситуационных центров. В этой сфере инициатива исторически распространялась сверху вниз – по уровню и масштабу управленческих задач. Исторически

первым стал ситуационный центр Президента России, затем – Правительства Российской Федерации. Потом такие центры стали появляться в федеральных министерствах и у руководителей субъектов Федерации. На сегодняшний день ситуационные центры созданы и в корпоративном секторе, и в образовательных учреждениях. Распространение центров и взаимодействие организаций уже позволяют констатировать существование системы ситуационных центров. И программный функционал для таких центров развивается в направлении охвата следующих сфер моделирования: стратегические направления развития страны, отраслей, развитие территорий, общественно-политической жизни. Последнее влечет за собой необходимость моделирования социальных взаимодействий, социальных сетей. Современный инструментарий позволяет преодолеть рамки социологических опросов и перейти к исследованию общества в социальных сетях. Отмечается, что задачи ситуационных центров для государственного и корпоративного управления выходят за рамки мониторинга. Всё более остро ставятся «задачи многофакторного анализа, ситуационного прогнозирования, стратегического планирования, управления проектами и программами [166]. Для прогнозирования в технических решениях может заключаться в широком внедрении предиктивной аналитики, которая также находится в числе приоритетов стратегии научно-технического развития России [168].

На фоне развития концепции государственного управления на основе системы распределённых ситуационных центров с использованием широкого инструментария для анализа социально-экономических процессов и планирования исследователи разрабатывают разнообразные компьютерные модели для поддержки принятия решений. В частности, МЧС России демонстрирует инструмент для оптимизации имеющихся ресурсов для защиты населения и территорий от чрезвычайных ситуаций. Исследователи региональной экономики инструментальными средствами решают задачу экономического районирования страны для снижения неоднородности территориального развития. Эта задача остается актуальной до сих пор: после СССР она потребовала

решения в результате появления независимых государств, но и с 2010-х годов, когда появилась новая сетка экономических районов, сохраняется проблема экономической дезинтеграции и конкуренции регионов.

Известен инструмент для сценарного анализа последствий внешнего воздействия в области сельского хозяйства. На региональном уровне выделяются поставщики и потребители сельскохозяйственной продукции в разрезе девяти товарных групп, моделируется внешний рынок. Далее в модели частичного равновесия с явным представлением технологий оцениваются выпуск, цены, объемы внешней торговли. Сценарии анализа включают регулирование экспорта и импорта сельхозпродукции, санкционное воздействие, климатические изменения.

Связанной с вопросами питания является модель распространённости анемии в зависимости от доступности сбалансированной продуктовой корзины. Модель связывает данные макро-, мезо- и микроуровней для исследования вопроса о рисках развития анемии при различных сценариях развития продовольственного сектора экономики.

Большую группу инструментов анализа составляют демографические модели – в том числе, со специализацией на конкретных субъектах Федерации. Зачастую такие модели являются агентными, где каждый индивид описывается набором характеристик, а численность населения оценивается в динамике по результатам естественного и миграционного прироста населения. При этом агенты следуют различным образовательным, трудовым, семейным и миграционным траекториям в соответствии с predetermined функциональными зависимостями.

Сталкиваясь с задачами принятия решений, исследователи ищут возможности для повышения информированности рассматриваемых сценариев. Так, оригинальная методика к решению задач импортозамещения предложена на основе сочетания информации из патентных и финансовых информационных ресурсов [133]. И примеры решения управленческих задач, перспективы развития ситуационного моделирования требуют обращения к

методологическим основам принятия решений для лучшего понимания того места, которое знания занимают в этой процедуре.

1.3. Устранение неопределённости информации при принятии решений

Разрабатываемые для государственного управления информационно-аналитические системы с опорой на модельные комплексы преследуют общую цель – снижения ошибок при принятии политических решений. Но эта же цель имеет место и в корпоративном секторе: дополнительный анализ призван снизить ошибки фирм или индивидуальных действий. «Ошибки», однако, не всегда понимаются как отклонение от некоторого идеального конечного результата. Такое возможно в случае детерминированных ситуаций. В остальном можно говорить об отклонении от общей динамики. В случае государства общим фоном является международный уровень, в случае региона – страна, в случае предприятия – отрасль. Иллюстрацией этого служит модель из Подраздела 1.1, где индивидуальная ошибка принятия решений может интерпретироваться в терминах отклонения показателей конкретной фирмы от показателей рынка в целом. То, как экономические агенты обрабатывают доступное им знание, непосредственно связано с принятыми ими процедурами принятия решений. Для последних характерна дифференциация по степени участия человека. Одни процедуры полностью полагаются на человека с трудно формализуемым правилом решения, другие – на сочетание человека и формальных методов, третьи – исключительно на алгоритмизуемые методы. Важно сказать об обработке знаний формальными процедурами, чтобы яснее представить связь между информационным продуктом и ценностью, которую он приносит потребителю. В контексте диссертационного исследования ценность реализуется в результате принятия информированного решения.

Задача принятия решений (ЗПР) может быть сформулирована в терминах выбора одного или нескольких лучших вариантов действий (альтернатив) из имеющегося набора [132]. Ключевыми элементами выбора являются чёткая цель и критерии сравнения вариантов решения между собой. Причём

процедура их количественного измерения открывает широкие возможности для использования компьютерных программ в ЗПР. На применимость того или иного метода принятия решений существенно влияют состав, объём и достоверность доступных сведений, которые можно разделить на четыре основные категории:

1. сведения об альтернативных вариантах;
2. сведения о критериях выбора;
3. сведения о предпочтениях лиц, принимающих решение (ЛПР);
4. сведения об окружении задачи.

Важность информации и вычислительных процедур в ЗПР позволяет проиллюстрировать их формальное описание. В самом общем случае ЗПР представима в виде [130]:

$$\langle S, G, A, C, M, F, P, R, T, V \rangle,$$

где S – формулировка проблемы принятия решений;

G – главная цель выбора;

A – множество доступных альтернатив;

C – множество критериев выбора;

M – множество шкал для измерения качества альтернатив по критериям;

F – отображение множества допустимых альтернатив в множество их оценок по критериям;

P – система предпочтений ЛПР;

R – решающее правило, отражающее систему предпочтений;

T – время, отводимое на анализ решения;

V – устойчивость информации категорий A, C, P во времени.

Каждая из компонент раскрывается ниже.

Формулировка проблемы (S). Возможны различные формулировки проблемы принятия решений, однако, в общем случае принимаемая формулировка S зависит от вида множеств A, C и системы предпочтений P .

Главная цель выбора (G). Формулировка цели – важная, но непростая задача. Проблема принятия решения часто характеризуется неопределённостью цели и даже множественностью целей. Конфликт целей зачастую не устраним, и именно поэтому принято говорить о выборе главной цели – той, относительно которой рассматриваются доступные альтернативы. Хотя от неопределённости в сложных ситуациях принятия решений полностью уйти не удаётся, возможно её снижение привлечением дополнительной информации.

Множество альтернатив (A). Множество исследуемых альтернатив A может быть дискретным или непрерывным, конечным или бесконечным. Рассмотрение пустых множеств A не имеет смысла, так как в таких случаях задачи выбора не существует. Существуют задачи с составными альтернативами, в которых варианты решений составляются из более мелких элементов, а выбрать нужно лучшую комбинацию. ЗПР со сложными (составными) альтернативами включают в себя процедуру синтеза альтернативных вариантов решений из неделимых элементов, при этом возникают проблемы измерения качества как простых элементов, так и составных, которые появляются в процессе синтеза и часто не существуют априори.

Способ описания множества альтернатив оказывает существенное влияние на постановку ЗПР и методику ее решения. Наиболее распространенными способами описания множества альтернатив являются следующие [159]:

1. Система ограничений (математических равенств и/или неравенств) на переменные целевых функций, отражающих качество альтернативных вариантов.
2. Перечисление альтернативных вариантов или их генерация случайным образом.
3. Специальный формальный алгоритм генерации альтернативных вариантов, обладающих определенными свойствами.

Алгоритмические процедуры и математические конструкции, которые участвуют в способах описания множества альтернатив, указывают на

естественность использования программного обеспечения в случае высокой трудоёмкости построения множества A .

Множество критериев выбора (C). По количеству критериев выбора (мощности множества C) ЗПР можно разделить на задачи с единственным критерием и многокритериальные. Задачи с единственным критерием принципиально отличаются от многокритериальных тем, что в них не возникает конфликтов между критериями. Чем больше критериев в конкретной задаче выбора, тем выше вероятность возникновения конфликтов между оценками альтернатив по критериям. Большая часть реальных проблем относится к многокритериальным ЗПР.

Множество шкал для измерения качества альтернатив по критериям (M). Понятие качество альтернативы отражает степень удовлетворения рассматриваемого объекта предъявляемым к нему требованиям со стороны ЛПР. Причём «качество» включает в себя представления об объективных свойствах альтернативы и субъективных предпочтениях ЛПР. Соотношение объективной и субъективной составляющих может быть различным в зависимости от типа критерия и компетентности ЛПР. По типу критерии можно разделить на измеряемые и неизменяемые.

Отображение множества допустимых альтернатив в множество их оценок по критериям (F). Отображение представляет собой формальное правило «оцифровки» альтернатив в том смысле, что каждый из доступных вариантов получает некоторую измеримую характеристику качества. По виду F различают задачи в условиях определенности (детерминированный вид F), задачи в условиях риска (вероятностный вид F) и задачи в условиях неопределенности (теоретико-множественный вид F). *Задачи в условиях определенности* также называются задачами *оптимизации*. Если отображение F имеет вероятностный вид, то ЗПР называют *задачей в условиях риска*. Если информация о проблеме является неполной, неточной, неколичественной либо формальные модели слишком сложны или отсутствуют, то ЗПР называют *задачами в условиях неопределенности*.

Система предпочтений ЛПР (P). Выбор из заданного множества альтернатив возможен, если ЛПР имеет в распоряжении информацию, необходимую для сравнения этих альтернатив в терминах главной цели. Для задач в условиях неопределенности и риска характерна неопределенность целей, справиться с которой можно, привлекая дополнительную информацию о субъективных предпочтениях ЛПР. При этом предполагается, что ЛПР имеет некоторую систему предпочтений на множестве рассматриваемых альтернатив и ее необходимо выявить в процессе анализа проблемы.

Предпочтения ЛПР могут быть выражены количественными или порядковыми оценками. В первом случае говорят о количественных (кардинальных) предпочтениях, во втором – о порядковых (ординальных). Количественные оценки дают больше возможностей для выражения тонких различий альтернатив по критериям качества.

Решающее правило, отражающее систему предпочтений (R). Под термином «решающее правило» обычно подразумевается способ обработки информации из категорий A, C, M, F, P, V , позволяющий достигнуть поставленной цели G (выбрать лучшую альтернативу из множества A). В задачах со многими критериями решающее правило описывает метод обобщения оценок альтернатив по разным критериям с учетом важности последних. Конкретный вид решающего правила зависит от формулировки проблемы, способов представления информации о критериях и альтернативах, а также от используемых шкал измерения предпочтений. Решающее правило представляет собой основу любого метода принятия решений.

Время, отводимое на принятие решения (T). Ресурс времени, которое отводится на принятие решения, определяющим образом влияет на выбор метода решения поставленной проблемы. Когда решение не является срочным (ресурс времени достаточно большой), существует возможность собрать необходимую информацию, получить консультации экспертов, построить математические модели, проанализировать разнообразные варианты и последствия решения. Иногда решения нужно принимать быстро, т.е. ресурс времени

сильно ограничен (переговоры, аварийные ситуации, военные действия). В таких случаях успех зависит от уровня подготовленности ЛПР к решению поставленной проблемы. Для многих ЗПР с ограниченным ресурсом времени информацию можно собрать заранее.

Устойчивость информации категорий A, C, P во времени (V). Длительность реализации решения, временной промежуток между решением и его последствиями, вообще говоря, могут быть связаны с изменением доступных альтернатив, критериями выбора и системой предпочтений.

В представленном формальном описании ЗПР явным образом выделяются элементы (A, C, P), существенно зависящие от доступной информации, и сферы применения программных средств для поддержки принятия решений (F, R , а также A в части генерации альтернатив). Особенно выделяется в этом ряду элемент F – отображение множества допустимых альтернатив в множество критериальных оценок задачи принятия решений. Как сказано выше, по этому элементу задачи делятся на задачи в условиях определенности, задачи в условиях риска и задачи в условиях неопределенности. Для решения задач в условиях определенности применяются методы оптимизации, в том числе методы математического программирования [222, 141, 249]. Задачи в условиях риска (или стохастической неопределённости) разрешаются с использованием статистических методов, в том числе Метода Монте-Карло, Марковских цепей [141], привлечением имитационного моделирования (134, 135).

Неопределенность в задачах принятия управленческих решений обусловлена целым рядом причин, начиная от объективной невозможности точно предсказать будущее и заканчивая конфликтами между субъективными предпочтениями ЛПР, который далеко не всегда точно знает, что на самом деле лучше или важнее. Для условий неопределенности характерно наличие многих критериев, среди которых могут присутствовать объективно неизмеряемые. Кое-что, однако, можно формализовать и в этом случае. Например, известны теоретико-игровые модели, которые представляют внутренний конфликт ЛПР о критериях в виде борьбы виртуальных игроков, каждый из

которых характеризуется своей собственной полезностью [42]. Введение виртуальных игроков позволяет разделить роли, которые сочетаются в одном индивидуе, в контексте конкретной ситуации принятия решения: например, руководитель фирмы, собственник фирмы, частный заёмщик и т.п. В общем случае, однако, решение задач в условиях неопределенности основано на привлечении информации о субъективных предпочтениях ЛПР. И большая часть управленческих решений на уровне предприятия и отрасли соответствует этому классу задач. В связи с высокой практической значимостью многокритериальных ЗПР следует также остановиться на методах их решения.

Учет нескольких критериев в процессах выбора приводит к возникновению потенциальных конфликтов, являющихся источниками неопределенности. В данном случае, однако, неопределенность возникает не от недостатка информации, а от ее избытка, так как можно получить различные упорядочения альтернатив по разным критериям. При этом вопрос о лучшем решении остается открытым. Для получения некоего единого порядка в таких ситуациях можно применить один из двух возможных приёмов: 1 – оставить один самый важный критерий, исключив из рассмотрения остальные (или переведя их в ограничения); 2 – получить некоторый единый критерий для оценки качества альтернатив путем агрегирования частных критериев.

Методы индивидуального принятия решений при многих критериях можно разделить на следующие группы [132]:

- лексикографические методы;
- аксиоматические методы многокритериальной теории полезности;
- методы сравнения многомерных альтернатив (методы доминирования, компенсации, порогов несравнимости);
- методы построения обобщенного критерия;
- метод идеальной точки;
- вербальные методы;
- методы теории нечетких множеств;

- интеллектуальные методы.

Представленные методы отличаются между собой по степени формализации используемых процедур, в том числе, в части сравнения альтернатив. К более строгим и формальным методам можно отнести лексикографические, методы многокритериальной теории полезности, метод идеальной точки, методы построения обобщённого критерия. Для своей работы они требуют количественных данных и упорядоченных предпочтений ЛПР. Конечно, предпочтения носят субъективный характер, и, в частности, лексикографические методы страдают именно от сильной зависимости исхода от субъективных предпочтений относительно ранжирования критериев. Тем не менее, эти методы опираются на количественные показатели и структуру взаимосвязей между элементами ЗПР. В основе лексикографических методов лежит предположение об упорядочивании критериев по их значимости [230]. Методы многокритериальной теории полезности основываются на аксиомах рационального выбора [233, 259]. Метод идеальной точки основан на использовании метрики, позволяющей оценить близость реальных альтернатив к идеальному варианту. Если на множество альтернатив x существует альтернатива $x^* \in X$, векторная оценка которой в пространстве критериальных оценок R^n удовлетворяет требованиям $\forall i = \overline{1, n} | \xi_i = \min f_i(x)$, то x^* является наилучшей альтернативой. Во многих случаях такая альтернатива реально не существует, поэтому в качестве лучшей предлагается выбрать альтернативу, векторная оценка которой наиболее близка к идеальной точке $\xi = \{\xi_1, \xi_2, \dots, \xi_n\}$. Методы построения обобщённого критерия предполагают сведение многокритериальной задачи к однокритериальной на основе тех или иных принципов [229, 221, 206, 251], наиболее известными из которых являются [132]:

1. Принцип гарантированного результата.
2. Принцип дифференциальной оптимальности.
3. Принцип компромисса.
4. Принцип равномерной оптимальности.

5. Принцип интегральной оптимальности.
6. Принцип справедливого компромисса (мультипликативный).
7. Критерий Байеса – Лапласа.

Менее строгие, хотя также алгоритмизируемые методы характеризуются своей возможностью работы со слабо структурированными данными, не количественными оценками и качественной информацией. Так, методы принятия решений на базе теории нечетких множеств способны обрабатывать различные аспекты ЗПР, а том числе, оценивать качество альтернатив, выявлять и отбирать предпочтения [164, 162, 165, 187, 225, 139, 201, 260, 140, 128]. Однако их слабым метом остаётся выбор вида функций принадлежности и интерпретация результатов. Вербальные методы изначально разрабатывались для слабоструктурированных задач, отличительными особенностями которых являются: качественная информация (в противовес количественной); неизвестные зависимости между переменными. Методы данной группы ориентированы на вербальный язык описания проблем, и основоположником их развития принято считать О.И. Ларичева [188, 189]. Интеллектуальный подход к принятию решений реализуется экспертной системы с заблаговременно заложенными в неё знаниями. Поиск решений моделирует рассуждения на основе заданных фактов и базы знаний. Наибольшее практическое распространение получили экспертные системы логического и продукционного типа с дедуктивными выводами (системы первого поколения) [131, 254, 220, 191, 232]. Методы сравнения многомерных альтернатив следуют парадигме парнодоминантного выбора, и наиболее известными представителями данной группы являются методы ELECTRE (Elimination et Choix Traduisant la Realite – исключение и выбор, отражающие реальность), предложенные Бернардом Руа [95, 242, 241], в которых на основе парного сравнения вычисляются индексы согласия и несогласия с гипотезой о превосходстве одной альтернативы над другой. Здесь осуществляется последовательное ранжирование множества заданных альтернатив на основе проверки гипотез о превосходстве, а не решается задача построения обобщённого критерия.

В отношении обработки альтернатив одни методы больше полагаются на работу с критериями, другие – на попарное сравнение альтернатив. Хотя альтернативы соотносятся друг с другом через призму принятых критериев во всех случаях, часть методов в большей степени сфокусирована на отборе критериев: их приоритезации и сведения многокритериальной задачи к однокритериальной. Принимая во внимание трудоёмкость процедур многокритериальной оптимизации с одной стороны и возможность их алгоритмизации – с другой, исследователи разработали программные средства для поддержки принятия решений. В частности, на практике используется программа, реализующая шаги метода анализа иерархий [203]. Метод, разработанный Томасом Саати [245, 97], основан на представлении задачи принятия решений в виде иерархической модели с последующим определением приоритетов всех ее элементов на основе предпочтений ЛПР. Метод хорошо зарекомендовал себя для анализа управленческих решений. Он предполагает декомпозицию проблемы на иерархию составляющих, оценку элементов нижних уровней относительно элементов высших уровней на основе суждений ЛПР, вычисление приоритетов элементов и их иерархический синтез. Вершиной иерархии является цель выбора, а нижний уровень содержит альтернативы. Математический аппарат метода базируется на матричной алгебре, что позволяет эффективно компьютеризировать процесс. При наличии взаимосвязей между элементами задачи, нарушающих принцип иерархической композиции (например, взаимозависимости критериев), используется развитие метода анализа иерархий – метод аналитических сетей [96, 244]. Он позволяет оценить приоритеты элементов сетевой структуры с произвольными связями относительно заданной цели. Элементы задачи группируются в кластеры, связи между которыми устанавливаются на основе специфики решаемой задачи.

Большинство из представленных методов требуют множества вычислительных операций: обращение и взвешивание матриц парных сравнений, расчета индексов согласия/несогласия в одних случаях и решение задач оптимизации, операции с нечёткими множествами – в других. Это приводит к тому,

что ручной расчёт в задачах с хотя бы десятью альтернативами и критериями становится крайне неэффективным. Этим объясняется разработка специализированного программного обеспечения для поддержки процессов принятия решений. Такие программы могут не только включать специализированные инструменты (например, простейшие версии статистических пакетов там, где требуются вероятностные оценки), но также возможности для интеграции с базами данных и средства работы с большими данными. Значительные усилия разработчиков также направлены на разработку визуального интерфейса, что позволяет сделать системы более доступными для менее подготовленных в техническом отношении экспертов, что особенно важно в ситуациях группового принятия решений.

Представленный обзор методов принятия решений демонстрирует «расщепление» знаний на «данные» – доступные ЛПР сведения – и «программы» – методы обработки данных. Первое вряд ли вызывает недоумение, поскольку мысль о том, что знания вытекают из данных в результате специальной обработки, уже высказывалась в некоторых работах [205]. Способ такой обработки тоже может интерпретироваться в терминах знания. Возражение может вызвать ассоциация компьютерной программы со знанием – сущностью, которая в философском плане неотделима от индивида. На него можно ответить следующим образом. Способ обработки данных может иметь несколько форм представления: алгоритм/ последовательность операций, математическое представление вычислительных процедур, программная реализация вычислительных процедур. То есть программа является способом представления знания. Однако в дальнейшем изложении диссертационного исследования регулярно используется термин «информационный продукт». Им могут являться данные, компьютерные программы и сочетания первых со вторыми. Таким образом, для последовательного приближения к рынку информационных продуктов и исследованию процессов, происходящих на этом рынке, требуется упорядочивание терминологии с разграничением мест для знаний, информации, данных и компьютерной программы.

1.4. Новая парадигма работы со знаниями, информацией и данными

Развитие информационных технологий и алгоритмов работы с данными приводят к существенным изменениям не только в деятельности людей, работающих в информационной сфере, но и в используемом в ней профессиональном жаргоне – языке, на котором эта деятельность описывается ее непосредственными участниками. В частности, это касается таких общеупотребительных слов, как «данные», «информация» и «знания».

Следует отметить, что изменения затрагивают не столько определения данных понятий, сколько их взаимосвязь и смысловое наполнение. В практической деятельности формальные определения часто оказываются избыточными, поскольку для понимания удобнее оперировать образными представлениями. Для визуализации различий в информатике используется концепция «пирамиды информации». На рисунке 1.5 представлена интерпретация данной пирамиды, предложенная специалистами Google [113].

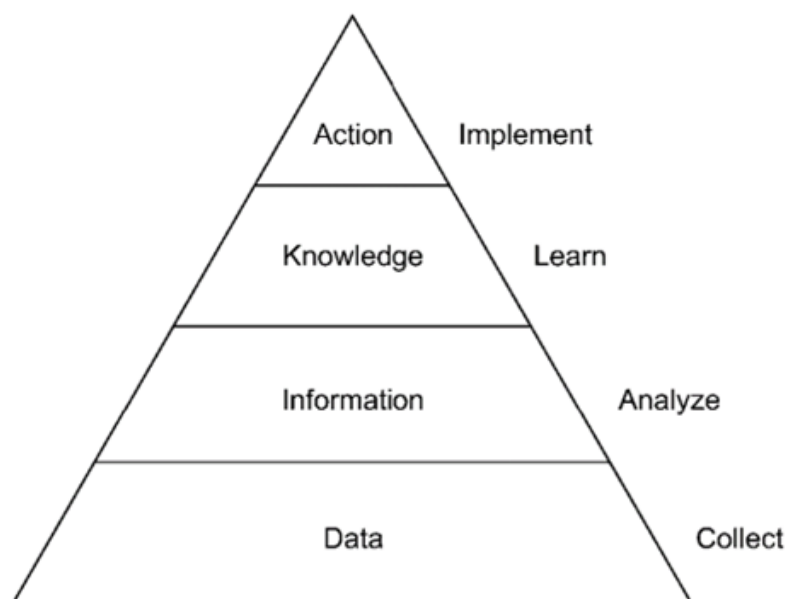


Рисунок 1.5. Пирамида информации. Источник: [113]

Основание пирамиды образуют собственно данные – первичная информация в машиночитаемом формате, собираемая техническими устройствами (видеорежистраторами, сканерами и т.п.). Даже в случае работы исключительно с изображениями вариантов реализации существует множество.

Объединяющим признаком является предназначенность для автоматизированной обработки. Упрощённо, данные — это отдельные, неинтерпретированные сигналы или символы. Следующий уровень – информация, адаптированная для человеческого восприятия, которая формируется путём преобразования исходных данных с использованием специализированных алгоритмов. Важно, что на этом уровне в данных появляется некоторая структура, доступная для восприятия человеком. Верхний уровень пирамиды, обозначенный как «знания» (Knowledge), составляет усвоенная и интерпретированная информация, служащая основой для принятия решений. Часто она соотносится с опытом (который включает и другие «знания») и ценностями индивида. Предложенная трактовка может рассматриваться как узкая с философской точки зрения, однако она отражает отраслевую специфику информатики. Семантика рассматриваемых понятий варьируется в зависимости от контекста, включая предметную область, временной период и решаемые задачи. Это особенно характерно для их использования в качестве терминов в различных научных и прикладных дисциплинах – экономике, юриспруденции, биологии, химии, физике и других. Подробный анализ вариантов употребления представлен в учебном пособии ИМИТ ПетрГУ².

Попытки дать общее определение термину «информация» предпринимались неоднократно и всегда неудачно. Не случайно такие великие ученые, как Винер, Шеннон и Колмогоров предпочитали давать определение понятию «количество информации», но не давали определения понятию «информация», хотя о своем понимании сути дела высказывались и приводили примеры. Так, в статье [184] определению количества информации предшествует рассуждение, которое раскрывает смысл алгоритмического:

«Реальные объекты, подлежащие нашему изучению, очень (неограниченно?) сложны, но связи между двумя реально существующими объектами исчерпываются при более простом схематизированном их описании. Если

² https://math-it.petrstu.ru/users/semenova/Informatika/DOC/MUM_65/Lectons/Information_PP.pdf

географическая карта дает нам значительную информацию об участке земной поверхности, то все же микроструктура бумаги и краски, нанесенной на бумагу, никакого отношения не имеет к микроструктуре изображенного участка земной поверхности» [184, с. 8].

Тут напрашивается аналогия с тем, как Дитмар Деттеринг в своей диссертации [25] описывает техническое определение медиаконтента со ссылкой на [114], хотя сам Вэриан не использует этот термин, он дает определение термина «information goods», то есть «информационные продукты», а термин «контент» использует при описании их ценности и правовой охраны. Согласно этому описанию, медиаконтент включает в себя все, что можно оцифровать. Например, если речь идет о книге, то это касается не книги как таковой, а только содержащегося в ней текста, и видеокассета сама по себе также не может быть оцифрована, но, возможно, изображения и звуки, которые она хранит. «Таким образом, тексты, изображения, звуки, базы данных и, конечно же, компьютерные программы представляют собой просто *состояния* носителя – среды». [25]

В цитате из статьи Колмогорова речь идет о том, что носитель информации, содержащейся в карте местности, не имеет отношения к микроструктуре изображенного на ней участка местности. Отсюда следует, что она не имеет отношения и к информации о микроструктуре этой местности, о чем можно и не напоминать, пока речь о физических объектах, а не о больших и очень больших объемах неструктурируемых данных. Но далее он говорит и об этом, хотя эпоха больших данных и отделения их от информации тогда еще не наступила.

«Практически нас интересует чаще всего количество информации в индивидуальном объекте x относительно индивидуального объекта y . Правда, уже заранее ясно, что такая индивидуальная оценка количества информации может иметь разумное содержание лишь в случаях достаточно больших количеств информации» [184].

Потом он переходит от общего рассуждения к разбору двух небольших примеров.

«Не имеет, например, смысла спрашивать о количестве информации в последовательности цифр

0110

относительно последовательности

1100.

Но если мы возьмем вполне конкретную таблицу случайных чисел обычного в статистической практике объема и выпишем для каждой ее цифры цифру единиц ее квадрата по схеме

0123456789

0149656941

то новая таблица будет содержать примерно

$$\left(\log_2 10 - \frac{8}{10}\right)n$$

информации о первоначальной (n — число цифр в таблицах)» (Колмогоров, 1965).

Здесь $\left(\log_2 10 - \frac{8}{10}\right)$ — количество информации, содержащейся в одной ячейке новой таблицы относительно содержания такой же ячейки старой таблицы. Сложение количества информации из отдельных ячеек здесь правомерно, поскольку их содержание никак не связано. Величина $\log_2 10$ — количество информации о содержании ячейки первой таблицы, если во второй это 0 или 5. Если во второй таблице другая цифра, то не хватает одного бита. А в среднем получается то, что в формуле, но здесь это уже приблизительно.

Дальнейшее изложение материала, включая определение понятия «количество информации», касается формальной схемы, где объекты пронумерованы, причем в двоичной системе. Впрочем, переход к двоичной форме представления — вопрос удобства. Так проще связать излагаемый материал с альтернативными походами к определению количества информации. Количество информации существенно зависит от принятого способа программирования. Это связано с тем, что количество информации с длиной символического представления строк — как и той, которая содержит обрабатываемые сведения, так

и той, которая представляет набор команд (программу). Различия в методах программирования влияют, в том числе, на трудоёмкость представления каждого символа в отдельности. При использовании одинаковых методов программирования в разных информационных системах этим можно пренебречь. Однако при передаче данных между информационными системами это обстоятельство может проявить себя в виде технической поправки в количестве информации. Именно технической – без искажения смысла сообщения (собственно информации). Формально это иллюстрирует следующая обширная цитата из [184].

«В соответствии с только что сказанным предлагаемое далее определение величины

$$I_A(x : y)$$

будет сохранять некоторую неопределенность. Разные равноценные варианты этого определения будут приводить к значениям, эквивалентным лишь в смысле $I_{A_1} \approx I_{A_2}$ т. е.

$$|I_{A_1} - I_{A_2}| \leq C_{I_{A_1}I_{A_2}},$$

где константа $C_{I_{A_1}I_{A_2}}$ зависит от положенных в основу двух вариантов определения универсальных методов программирования A_1 и A_2 .

Будем рассматривать «нумерованную область объектов», т.е. счетное множество

$$X = \{x\},$$

каждому элементу которого поставлена в соответствие в качестве «номера» $n(x)$ конечная последовательность нулей и единиц, начинающаяся с единицы. Обозначим через $l(x)$ длину последовательности $n(x)$. Будем предполагать, что:

- соответствие между X и множеством D двоичных последовательностей описанного вида взаимно однозначно;

- $D \subset X$, функция $n(x)$ на D общерекурсивна [1] (*примечание И.Н. – это цитирование не относится к списку литературы диссертации, а к списку оригинальной диссертации А.Н. Колмогорова*), причем для $x \in D$

$$l(n(x)) \leq l(x) + C,$$

где C – некоторая константа;

3) вместе с x и y в X входит упорядоченная пара (x, y) , номер этой пары есть общерекурсивная функция номеров x и y и

$$l(x, y) \leq C_x + l(y),$$

где C_x зависит только от x .

Не все эти требования существенны, но они облегчают изложение. Конечный результат построения инвариантен по отношению к переходу к новой нумерации $n'(x)$, обладающей теми же свойствами и выражающейся общерекурсивно через старую, и по отношению к включению системы X в более обширную систему X' (в предположении, что номера n' в расширенной системе для элементов первоначальной системы общерекурсивно выражаются через первоначальные номера n). При всех этих преобразованиях новые «сложности» K и количества информации остаются эквивалентными первоначальным в смысле \approx .

«Относительной сложностью» объекта y при заданном x будем считать минимальную длину $l(p)$ «программы» p получения y из x . Сформулированное так определение зависит от «метода программирования». Метод программирования есть не что иное, как функция

$$\Phi(p, x) = y,$$

ставящая в соответствие программе p и объекту x объект y .»

Тут уместно напомнить, что цитируемый текст написан в 1965 году, когда уже были вычислительные машины, но не было того, что сегодня принято называть большими данными. Соответственно, не было необходимости делать различие между данными и информацией. Впрочем, разделение информации,

хранящейся в памяти вычислительной машины, на команды и данные было сделано много раньше. Такое разделение было зафиксировано как один из двух признаков в формуле изобретения при подаче Конрадом Цузе заявки [126] на патент в 1936 году [175]. Но тогда это касалось только вычислительной техники, которую еще предстояло создать.

В настоящее время разделение доступной для восприятия человеком информации и собираемых различными устройствами данных, как минимум, целесообразно. Вопрос в том, где проходит граница, или, точнее, где заканчивается та область информации, воспринимаемой человеком не только в принципе, но и реально, учитывая ограниченность времени и внимания. Сложно утверждать, что человек может без использования технических средств просмотреть несколько десятков тысяч статей или патентных заявок. Тут на помощь приходят аналитические системы, позволяющие установить связи между патентами, представить их в виде схем и диаграмм. Несколько хуже обстоит дело с анализом массива научных статей, поскольку система представления научных публикаций гораздо слабее формализована. Но и тут за последнее время много сделано. А потому здесь имеет смысл говорить об алгоритмическом подходе. При этом роль «доступной для восприятия» играет та информация, которую пользователю предоставляет аналитический сервис.

Переходя к вопросу о разделении информации и знания, стоит отметить, что в экономической литературе можно найти немало попыток дать «четкое» разграничение между этими понятиями в духе «знания – это усвоенная информация», как и сознательных отказов от таких попыток. Первопроходец экономики знаний Фриц Махлуп [70] не только понимал термин «знания» (Knowledge) существенно шире, чем это представлено в «пирамиде информации», но и смотрел на него под иным углом зрения. Он был первым экономистом, кто посмотрел на знания как на производимый и потребляемый продукт. Во время выхода его книги [70] данных в том понимании, как они представлены в «пирамиде информации», просто не было, то есть не было таких объемов и специальных алгоритмов для обработки, хотя приборы и

вычислительные машины, разумеется, уже были. Данные хранились преимущественно на аналоговых носителях, а не в двоичном коде. Завершая этот абзац, можно сказать: различия между знаниями, информацией и данными Махлуп не делал по той причине, что его интересовало производство и распространение информации, а не принятие решений на ее основе.

Управление на основе знаний – область исследований, обозначаемая в англоязычной литературе как Knowledge management. В русскоязычной экономической литературе это направление обычно называют «управление знаниями», так как буквальный перевод «знаниевое управление» плохо согласуется с привычными нормами языка. Впрочем, перевод англоязычного термина «digital economy» как «цифровая экономика» воспринимается нормально. А то, что речь идет об управлении на основе знаний, видно не только из содержания многих научных работ этого направления, но и из того факта, что наиболее ярких его представителей можно найти в сообществе «Value based management». Тут без промежуточного based получается двусмысленность. Судя по всему, старый мем поизносился и был сменен без ущерба для смысла.

В той предметной области, где появилась пирамида информации, различие между знаниями и информацией принципиально. Его проще всего показать на таких примерах, как знание правил дорожного движения и принятие решения по управлению автомобилем при виде дорожного знака или сигнала светофора. Знак и сигнал светофора – информация, а знание правил дорожного движения водителем – знание. Но робот, управляющий автомобилем, реагирует на знак или сигнал не так, как человек, его путь от данных к действию короче, так как переход от данных к информации в воспринимаемой человеком форме – лишнее звено. Реально для управления автомобилем нужна непосредственная связь между «картинками», то есть данными, фиксируемыми приборами, и конкретными устройствами или системами (двигателем, тормозами и т.д.). Тут четко можно увидеть «алгоритм», сопоставляющий «картинки», используемые при обучении нейросети, и наборы сигналов, отправляемыми на устройства управления автомобилем. Кавычки здесь использованы

по той причине, что «картинки» здесь не для восприятия человеком и «алгоритм» не написан человеком на основе его понимания.

Естественно возникает вопрос о том, как менялось представление об информации и ее ценности в экономической науке, когда и как в экономической науке сформировалось современное представление об информации, позволяющее понимать процессы, происходящие в реальной экономике и не путать термины, используемые в профессиональном жаргоне, с их бытовыми двойниками.

1.5. Три подхода к понятию количества информации и ее ценности

«Информация» является общеупотребимым термином и в широком обращении апеллирует к интуитивному пониманию того, что составляет содержание термина. Однако существует подход, который развивает теорию информации, отталкиваясь от определения Кастлера: «Информация – это запомненный выбор одного варианта из множества возможных и подобных ему» [90]. Ключевым предположением направления, получившего название «динамической теории информации», является самоорганизация в системе (вообще говоря, неважно – нейронов или группы людей) [256]. При этом взаимодействия проходят в двух подсистемах – «левой» и «правой», если проводить аналогию с полушариями головного мозга. Их также можно ассоциировать с «Системой 2» и «Системой 1», известными в поведенческой психологии [170]. Левая подсистема является детерминированной и отвечает за хранение усвоенных образов. Правая подсистема содержит элемент случайности, что позволяет, с одной стороны, генерировать новую информацию, а с другой стороны – обеспечивает недетерминированность связи между образом и символом. Теория динамической информации (на уровне взаимодействия нейронов) находит многочисленные практические подтверждения. В работе 2017 года на основе формализма теории проводятся параллели с системами искусственного интеллекта, которые ассоциируются с левой подсистемой. На тот момент времени генеративные модели не были так широко распространены и большой вес формальных классификаторов (например, метод опорных векторов)

свидетельствовал в пользу такой связи. Однако бурное распространение больших языковых моделей после 2022 года, графических генеративных сетей позволяет провести аналогии и с правой подсистемой: она также с элементом случайности генерирует новые образы по доступным системе символам, а дискриминатор, чью аналогию можно уподобить левой подсистеме, выдаёт оценку реалистичности результата.

Положения и формализм динамической теории информации представляют интерес в контексте данной работы по нескольким причинам. Во-первых, предполагается существование связи между сигналом и образом. Фактически, это эквивалентно переходу между данными и собственно информацией (в виде текстов, изображений – «контента»). Этот переход имеет важное значение для рынка информационных продуктов (данных, программ, информации), поскольку затрагивает важные вопросы об оценке целого по доступным фрагментам и принятия решений в условиях неполной или асимметричной информации – сугубо экономические проблемы. Во-вторых, сигналы, несущие сведения об образе, имеют ненулевую ценность. Причём эта ценность имеет субъективный характер, поскольку в образовании связи «образ – сигнал» вмешивается некоторая стохастическая величина – «шум». И, далее, результат обработки сигналов (данных, распространяемых на рынке) существенно зависит от доступных программных средств (вообще говоря, участники рынка по-разному обрабатывают данные). В-третьих, идея самосогласованности системы в установлении связей перекликается с идеей самосогласованного рынка. В динамической теории информации связи между символами имеют тенденцию к укреплению – формированию «лучших» связей. На самосогласованном рынке в состязательной процедуре также выигрывает тот, кто наиболее адекватен сложившейся ситуации. И можно показать, что этот выигрыш связан с меньшими ошибками участника рынка, и накопление информации способствует снижению ошибок.

1.5.1. Подходы к измерению количества информации.

Переход к рынку информационных продуктов естественным образом влечёт за собой вопрос о передаче информации: как измерять её количество и как определять её цену. В части измерения количества информации установился консенсус, чего нельзя сказать об определении информации. Широко распространённым является вероятностный подход, предложенный К. Шенноном [102]. Свою роль в его распространении сыграло развитие вычислительной техники и средств связи, для которых вероятностные характеристики сигналов оказались существенными для кодирования и анализа полноты полученных сведений.

Вероятностный подход стал развитием комбинаторного подхода, поэтому краткий обзор способов для измерения информации следует начать именно с него. Предложенный Хартли [43] комбинаторный подход отталкивается от вопроса о том, сколько информации передаёт получателю конкретное сообщение из наперёд заданного количества знаков. Естественно предположить, что сообщение позволяет получателю выделить конкретную строку из $N = m^n$ возможных, где n – количество знаков, m – количество возможных значений для каждого знака. И количество информации следует измерять некоторой аддитивной функцией f так, чтобы приращение информации было пропорционально приращению числа знаков в сообщении. Эти предпосылки приводят к выражению $df = k \cdot dN/N$ и, далее, $f = \log_a N$, откуда для двоичного представления совсем близко до $I = \log_2 N$

Шеннон ввёл понятие энтропии для определения количества информации, исходя из предположения, что строки сообщений не являются равновероятными [102]. И количество информации представляется разностью в изменении энтропии о строке при получении дополнительного сообщения:

$$I(x: y) = H(y) - H(y|x),$$

где $I(x: y)$ – информация в сообщении x относительно строки y ,

$H(y) = -\sum_y p(y) \log_2 p(y)$ – энтропия строки y (при двоичном кодировании),

$H(y|x) = -\sum_y p(y|x) \log_2 p(y|x)$ – условная энтропия строки y относительно x .

А.Н. Колмогоров обратил внимание на то, что связи между строками могут быть не только случайными [184]. Это наблюдение привело к формулировке количества информации в терминах программного представления данных. Пусть множество объектов $X = \{x\}$ пронумеровано, и $D = \{n(x)\}$ – множество двоичных последовательностей – номеров элементов x , представленных двоичной записью. При этом требуется взаимно однозначное соответствие между X и D . Функция $l(x)$ обозначает длину последовательности $n(x)$. Предполагая $l(n(x)) \leq l(x) + \text{const}$ на $D \subset X$ и $l(x, y) \leq C_x + l(y)$ при $x \in X, y \in X$, вводят понятие программы p и метода программирования $\varphi(p, x) = y$, ставящий в соответствие объект y программе p и объекту x .

Информацией при этом называется величина $I_A(x:y) = K_A(y) - K_A(y|x)$, где

$$K_A(y|x) = \begin{cases} \min_{A(p,x)=y} l(p) \\ \infty, \neg \exists p | A(p, x) = y \end{cases}.$$

Метод программирования A является «наилучшим» в том смысле, что любой другой метод связан с программой большей длины, т.е. $K_A(y|x) \leq K_\varphi(y|x) + C_\varphi$.

Следует отметить, что «случайность» остаётся в теории Колмогорова. Это ясно видно из комментария, который даётся к определению (α, β) -стохастичности [258]:

«Пусть α, β – натуральные числа. Число x будем называть (α, β) -стохастическим, если существует такое конечное множество $A \subset N$, что $x \in A, K(A) \leq \alpha, K(x) \geq \log_2 |A| - \beta$;

здесь через $|A|$ обозначена число элементов множества A . Первое неравенство (если α невелико) означает, что множество A достаточно просто. Второе (если β невелико) означает, что элемент x является "элементом общего положения" в множестве A .»

Сильно упрощая, строка с маленькой колмогоровской сложностью случайна для множества, элементы которого сложны с вычислительной точки зрения.

Однако гораздо важнее в теории Колмогорова то, что данное определение вовлекает понятие программирования. Далее, при моделировании рынка информационных продуктов, в формализме используются случайные величины: товар (программа, массив данных или их сочетание) и сигнал о качестве товара имеют характер случайности и связаны друг с другом. В этом отношении, алгоритмический подход позволяет обойти вопрос о характере связи между строками (товаром – информационным продуктом и сигналом о нём), обоснования случайности. Принципиальным является то, что программа – способ обработки сигнала о товаре – может быть своей у каждого участника рынка, и это принципиально. Сам А.Н. Колмогоров глубоко прорабатывал вопрос о вычислении колмогоровской сложности (длины строки) [142]. В приведённом выше определении отсутствие программы приводит к бесконечной сложности строки. Действительно, в общем случае вычисления сталкиваются с проблемой останова: не существует алгоритмической процедуры, которая позволяет проверить, закончилось выполнение программного кода или нет до его запуска. Поэтому для приближения к реальным вычислениям вводится понятие колмогоровской сложности с ограничением на ресурсы $K^{B(k)}(y)$, т.е. значения, получаемого в пределах заданного времени с некоторым отклонением от $K(y)$. Фактически, различия в доступе к ресурсам позволяет дифференцировать покупателей на рынке информационных продуктов, и использование этой возможности при формальном анализе ценообразования показано ниже.

1.5.2. Несколько слов о ценности информации

Наряду с количеством информации измеряются другие её характеристики – например, незашумлённость, ценность информации. О последней следует дать комментарий. В управленческом смысле информация, будучи объектом, который передаётся и сообщает какие-то сведения, обладает некоторой

полезностью. Например, позволяет уменьшить ошибку при совершении действия, принятия какого-либо решения. В литературе, однако, чаще встречается понятие о ценности информации, а не её полезности.

В ряде таких определений можно указать на концепцию Р. Ховарда [49], согласно которой ценность информации пропорциональна разности в ценности решений, принимаемых при наличии информации и в её отсутствие:

$$V(I) = U(d') - U(d),$$

где $V(I)$ — ценность информации I ; d — решение, принимаемое до получения информации I ; d' — решение, принимаемое после получения информации I ; $U(d)$ и $U(d')$ — функции ценности соответствующих решений.

А.А. Харкевич [253] определяет ценность информации через разницу вероятностей в достижении цели. Субъект в этом случае присутствует в явном виде. Если в случае определения, данного Ховардом, функции ценности решений могут принимать денежное выражение (например, чистая приведённая стоимость проекта по разработке месторождения до проведения разведки и после), то концепция Харкевича предполагает целеполагание некоторым субъектом. Ценность информации не объективна, а имеет место для кого-то конкретного. Итак, согласно Харкевичу,

$$V(I) = \log_2 p_1/p_0,$$

где p_0 — вероятность достижения цели до получения информации, p_1 — вероятность достижения после.

Также можно указать на формальный подход Стратановича, который характеризует ценность диагностической информации [185]. Модельная ситуация, которой соответствует определение, предполагает наличие объекта с набором возможных состояний, измеримых диагностических признаков, процедуры проверки (измерения значений признаков). Последняя выявляет информационные состояния. Результаты проверок — наборы случайных исходов (диагностических признаков). Считается, что с каждой проверкой связаны две конкурирующие гипотезы о результате R_{kj}^v и R_{kj}^u . Им соответствуют априорные и апостериорные вероятности исходов. Если априорные средние потери

до проверки обозначить через $L(R_{kj}^u)$, а апостериорные – через $L(R_{kj}^v/R_{kj}^u)$, то ценность получаемой информации (при стратегии минимизации потерь) вычисляется во формуле:

$$V(I) = L(R_{kj}^u) - L(R_{kj}^v/R_{kj}^u).$$

Перечень концепций можно было бы продолжить, однако, гораздо важнее не обозреть их, а подчеркнуть сходство. Во-первых, речь идёт о сравнении двух состояний – до и после измерения. Во-вторых, ценность информации может быть отрицательной, и в таком случае к ней применяется термин «дезинформация». Эти наблюдения используются далее при построении модели рынка информационных продуктов.

Выводы раздела

Историю исследований об информационных продуктах как товарах можно проследить по нескольким линиям: экономика знаний, экономика интеллектуальной собственности. Последняя подчёркивает отличия: для придания знаниям свойств товаров им необходимо обеспечить редкость. В случае информационных продуктов это чаще обеспечивается техническими средствами: форматами представления, протоколами обработки данных. При этом ценность таких продуктов заключается в той роли, которую они играют в принятии решений: это устранение неопределённости и реализация формальных процедур выбора альтернативы. Экономически значимым следствием информированных решений в экономике являются расширение доступа к природным ресурсам и снижение индивидуальных ошибок экономических агентов – конкретные механизмы экономического роста.

Продукты можно разделить по их ценности для потребителя или по тому, в какой степени они устраняют неопределённость о конкретной ситуации (или, эквивалентно, насколько больше информации извлекается из данных). Естественным в этой связи выглядит алгоритмический подход к определению количества информации, поскольку он опирается на тесную связь между данными и программным обеспечением для их обработки, в том числе, с учётом ограничения на вычислительные ресурсы. Опора на него предлагает

различать сферы применения продуктов по колмогоровской сложности, и значение такого взгляда на предмет раскрывается в следующем разделе.

2. Организация рынка данных (информации)

Выбранный формализм в определении количества информации удобен своей операциональностью – возможностью соединения с результатами, полученными в разделах теории игр: совместимыми со стимулами механизмами и теорией дележей. К рынку информационных продуктов применимы результаты, полученные для товаров, неуничтожимых в потреблении. Прежде всего, это необходимость ценовой дискриминации для поставки блага на уровне, необходимым для наиболее полного удовлетворения потребностей потребителей. Этот теоретический результат соответствует фактическим способам оплаты на рынке программных продуктов. Коробочная лицензия по единой цене, несомненно, тоже имеет место. И такой вариант продажи характерен для продуктов, чей функционал не поддается изменениям. И отсутствие возможностей для ценовой дискриминации именно из-за невозможности скорректировать предложения для групп пользователей подталкивает продавца к единой для всего рынка цене, в том числе, и по необходимости соблюдения антимонопольного законодательства. Согласно теории монополистической конкуренции, товар в таком случае должен распространяться по средним издержкам.

Однако в случае программных продуктов существуют широкие возможности для дифференциации групп пользователей и предложения им индивидуальных цен. Причём при распространении по подписке (данная концепция известна под названием *software as a service (SaaS)* и противопоставляется коробочной лицензии) позволяет уточнять группировку потребителей за счёт повторяющихся взаимодействий. Распространение по подписке позволяет продавцам использовать разнообразные механизмы оплаты в противовес плоской цене. Во-первых, может предлагаться так называемое «ценовое меню». В этом случае продукт дифференцируется по функциональности: набор товаров получается из комбинации программных сервисов, и каждый из них предлагается соответствующей группе потребителей по цене, равной (в идеальном случае) их готовности к оплате. Однако помимо ценового меню следует указать

на возможности для применения двухкомпонентных тарифов. Поэтому в качестве второго примера механизмов можно привести встроенные покупки. Большинству он знаком по компьютерным играм. Потребитель приобретает лицензию на продукт, однако, улучшение пользовательского опыта предлагается за дополнительную плату (например, игрокам предлагаются улучшенные игровые предметы – артефакты, которые позволяют получить дополнительные очки или улучшить внешний вид игровых персонажей). Наконец, в-третьих, после приобретения подписки потребителю могут предлагаться дополнительные программные сервисы, что также может рассматриваться как вариант двухкомпонентного тарифа. Например, за дополнительную плату пользователь имеет возможность расширить ограничения на выгрузку данных из информационной системы.

С известной особенностью товаров на информационном рынке связан вопрос о том, как сформировать предложение потенциальному покупателю. Эта проблема известна также, как парадокс Эрроу: если покупателю неизвестны все сведения о товаре (знании, информации, данных), возникает вопрос о том, за что платить; если покупателю раскрывается полный набор сведений – возникает вопрос о том, зачем платить, ведь сам товар уже может считаться переданным при ознакомлении. На практике подобную проблему обходят пробным периодом использования программного продукта – ограничением сроком лицензии/ подписки на полнофункциональную версию. При этом, однако, остаётся вопрос о том, как установить цены для групп потребителей и обеспечить сделки по этим ценам. Необходимо исключить ситуации, при которых потребители оплачивают по более низким ценам, чем это ожидается, «притворяясь» потребителями иного типа. С этой целью разрабатывают механизмы, совместимые со стимулами. Среди них, в частности, выделяются работы, основанные на использовании налога Кларка (варианты схемы Викри-Гровса-Кларка), и процедуры передачи информирующих сигналов. Причём в случае ценового меню такими сигналами могут выступать сами цены, сообщая информацию о качестве продукта (набора данных).

2.1. Информация как публичное благо, поставляемое в частном порядке

Приближение к рынку информационных продуктов следует начать с определения товара. Цифровые технологии позволяют копировать и тиражировать ряд продуктов без ущерба для их качества практически по нулевым издержкам: будь то тексты, музыкальные, аудиовизуальные произведения, изображения или компьютерные программы. Причём потребление продукта одним потребителем не ухудшает его качества, как и не делает продукт недоступным для потребления другими. В этом отношении товары имеют признаки общественных благ, на что указывается в литературе. Но при обсуждении общественных благ часто обращается внимание на «исключительность». Неконкурентность в потреблении – свойство, при котором потребление одним потребителем не делает благо недоступным для других потребителей, – вообще говоря, неравносильно доступности блага всем потребителям без ограничений. Результаты интеллектуальной деятельности, к которым относятся компьютерные программы и базы данных, получают свойства товаров именно в результате ограничения на тиражирование посредством института интеллектуальной собственности. Правообладатели получают монополию на распространение³ в обмен на раскрытие своих результатов. Поэтому конкуренция в потреблении и возможности ограничения доступа к продукту могут стать основой классификации продуктов – одной из многих. Такую классификацию предложили Адамс и МакКормик [2] – хотя первенство в идее о том, что понятия «исключительности» и «неконкурентности» следует различать при обсуждении общественных благ, они приписывают Нитзану и Остром [80]. Классификация благ по указанным признакам представлена в Таблице 2.1.

³ Правообладатель законодательно наделяется целым пучком прав, но в контексте рынка сфокусируемся на распространении.

Таблица 2.1. Классификация благ в пространстве «исключимость – конкурентность в потреблении». Источник: перевод по [2]

	Возможно исключение	Исключение невозможно
Конкуренция в потреблении	Частные блага	Ресурсы общей собственности
Подверженность перегрузкам	Клубные блага	Нерыночные нечистые общественные блага
Отсутствие конкуренции в потреблении	Рыночные общественные блага	Общественные блага, не подлежащие продаже

В левом верхнем углу (конкурентные в потреблении с возможностью исключения) находятся частные блага. Конкурентные в потреблении продукты с невозможностью исключить каждого потребителя в отдельности названы ресурсами общей собственности. В качестве примера приводится рыба в океане. Это благо остаётся частным (каждая конкретная выловленная рыба достанется одному потребителю), но исключить кого бы то ни было из потребления нельзя, поскольку «каждый может выловить столько рыбы, сколько захочет ввиду отсутствия имущественных прав за пределами национальных вод».

Блага, подверженные перегрузке, имеют общее свойство ухудшения потребительских свойств при достижении некоторого порогового количества потребителей. Авторы классификации указывают на то, что группа потребителей делит благо между собой, как делит и расходы. Причём вопрос об исключении является скорее политическим, нежели экономическим. Примером клубного блага может являться бассейн, чья заполненность сверх меры затрудняет плавание. В качестве нерыночного не строго общественного блага можно указать на автомобильные дороги, перегрузка которых приводит к заторам.

Именно с последней строкой – неконкурентных в потреблении благ – Адамс и МакКормик связывают распространение технологий. Их пример относится к кабельному телевидению. Прежде только телевизионные компании могли принимать спутниковые сигналы. И такое (кабельное) телевидение относилось к рыночному общественному благу. Когда потребители, покупая

спутниковые антенны – частные блага, получили возможность принимать сигналы самостоятельно, это телевидение перешло в разряд общественных благ, не подлежащих продаже (чистых общественных благ). Компании ответили на это шифровкой сигнала, вернув свой товар в категорию рыночных общественных благ. Следуя этой классификации, Детеринг, относит весь медиаконтент к категории рыночных общественных благ [25]. По аналогии к данной категории можно было бы отнести и данные с программным обеспечением – информационные продукты. Но такая позиция игнорировала бы современные технологические реалии.

Разнообразие информационных продуктов позволяет указать примеры для каждой таблицы, предложенной Адамсом и МакКормиком. Удивлять тут может даже пример частного блага, хотя выше и говорилось о возможности копирования цифровых продуктов без потери качества и тиражирования с нулевыми издержками. Однако с 1993 года технологии и потребители так изменились, что на рынке информационных продуктов нашлось место и частному благу.

Частными благами представляются криптографические продукты на архитектуре блокчейн. К ним относятся криптовалюты и NFT. Последние следует рассматривать не в контексте предметов искусства, в каком они привлекли большой интерес СМИ, а в контексте виртуальных вселенных и компьютерной индустрии – именно эта область составляет большой по масштабам рынок⁴. Игровая индустрия предложила пользователям уникальные предметы, которые дают игровым персонажам специальные возможности (или уникальный внешний вид), и рынок принял это предложение. Пользователи приобретают такие виртуальные предметы (по своей сути – строчки кода – компьютерные программы) за реальные (фиатные) деньги и также – за реальные деньги продают их.

⁴ Количество пользователей в виртуальной вселенной Fortnite уже превышает 380 миллионов.

В качестве ресурса общей собственности можно указать на отдельные варианты генеративного искусственного интеллекта. Такие модели получили широкое распространение в конце 2023 г., прежде всего, благодаря удобному пользовательскому интерфейсу ChatGPT. Модели обучения первых открытых для широкой круга пользователей⁵ версий опирались только на тексты, курируемые разработчиками. Версия ChatGPT-4.1 стала попыткой выйти за сформированный разработчиками набор документов и расширить обучение на тексты пользователей системы. И здесь произошло то, что можно назвать ухудшением продукта: система стала выдавать неудовлетворительные ответы. Таким образом, оставаясь открытым (неисключимость в потреблении), потребляемый пользователями продукт стал сильно отличаться от выпущенной разработчиком ChatGPT-4.1. В этом смысле можно говорить о конкурентности в потреблении – пользователи, которые раньше опробовали модель, могли получить корректно работающий продукт в рамках конкретных тем для своих запросов.

Примеры благ с перегрузками понятны большинству пользователей интернет-сервисов или программ, распространяемых по модели SaaS. Перегрузки, по аналогии с физическим благом, возникают при избыточном для системы количестве одновременных обращений пользователей. При этом наращивание вычислительной и коммуникационной инфраструктуры переводит эти блага в разряд неконкурентных в потреблении.

При всём разнообразии групп, на которые можно разбить информационные продукты, модели получения дохода от их распространения на практике сводятся к двум: продажа лицензий и рекламные поступления при условии бесплатного доступа пользователей к продукту. Во втором случае также говорят об оплате пользователей вниманием, но денежные платежи с их стороны в пользу продавца отсутствуют. Оплата вниманием – отдельный вопрос,

⁵ Для регистрации, однако требовалась сим-карта, и абоненты не из каждой страны могли регистрироваться свободно. Вскоре появились сервисы, которые предлагали обойти наложенное техническое ограничение.

который переводит дискуссию в тему «Экономики внимания», что выходит за рамки данной работы. Для целей данного исследования интересен вопрос денежной оплаты поставщику со стороны потребителя, и поэтому следует иметь в виду только две модели доступа – с оплатой и бесплатно. В части развития методов ценообразования на рынке информационных продуктов представляют интерес модели первой группы, и далее приводится обзор основных теоретических моделей ценообразования.

2.2. Модели равновесия с общественным благом

Свойства «неисключительности» и «неконкурентности в потреблении» общественных благ можно записать формально – для использования в математических моделях и строгого анализа отношений, связанных с их распространением. Идею того, что «каждый потребляет одно то же самое количество», можно выразить равенством $x_{ik} = x_k$, где x_{ik} – потребление агентом i блага k , а x_k – доступный уровень блага k . Однако переход от равенства к неравенству позволяет расширить сферу применения получаемого формализма на те случаи, когда неравенство в уровне потребления реализуется по инициативе самих потребителей. Для примера можно обратиться к использованию веб-браузера. Распространение этого продукта не исключает ни одного из потребителей. Напротив, разработчик намеренно распространяет браузер бесплатно и без ограничений, стремясь тем самым получить наибольший охват. Ценность для разработчика реализуется не в установке браузера, а в возможности сбора данных об активности пользователя в сети Интернет и использования этих данных в разработке и продвижении других своих продуктов. Наряду с неисключительностью следует отметить и неконкурентность в потреблении: установка браузера на какое-либо устройство не ограничивает установки для других пользователей. Но уровень потребления блага разными пользователями, несомненно, различается. Это касается, в первую очередь, использование встроенных функций браузера: просмотр документов, отличных от HTML-формата (например, PDF); ведение каталога закладок; ограничение на сбор и передачу данных о поведении в сети; сохранение данных учётных записей и

т.п. Следующий слой относится к установке дополнительных надстроек (также бесплатных для пользователей и доступных без ограничений), но зачастую они поставляются сторонними разработчиками, и в таком случае более уместно говорить об эффектах дополненности. Но суть от этого не меняется: при наличии полнофункциональной версии продукта («и технической возможности в смысле совместимости программного обеспечения, свободной физической памяти и т.п.» – добавим для строгости) пользователи могут ограничиться базовыми функциями и стандартными предустановками. Соответственно, изолированное от «неисключительности» свойство «неконкурентности» следует описывать соотношением $x_{ik} \leq x_k$

2.2.1. Равновесие Нэша

Равновесие Нэша в игре n лиц – это такое состояние, когда ни один игрок не может улучшить свое положение, если остальные игроки не откажутся от выбранных ими стратегий. В этом смысле существующая стратегия каждого игрока в равновесии Нэша оптимальна. Равновесие по Нэшу, или решение по Нэшу, используется во многих ситуациях с игровым элементом, однако в ситуации с коллективными или общественными благами оно абсолютно неудовлетворительно, поскольку дает заведомо неоптимальное решение. Иначе говоря, действуя сообща, игроки могли бы получить результат, существенно лучший для каждого из них.

Для понимания ситуации сначала удобно предельно упростить задачу, рассматривая экономику всего лишь с двумя агентами, одним общественным благом и одним частным благом, причем производство общественного блага отождествить с затратами частного блага. Такое предположение относительно производства вполне корректно, поскольку его выполнения можно добиться путем замены переменных. Вместе с тем, при сделанных предположениях можно использовать наглядные геометрические конструкции, известные как диаграмма Кольма и треугольник Кольма [63]. В том и другом случае распределение благ – точка некоего симплекса. Но диаграмма Кольма, вообще говоря, трехмерна, а треугольник Кольма двумерен, рисовать его проще.

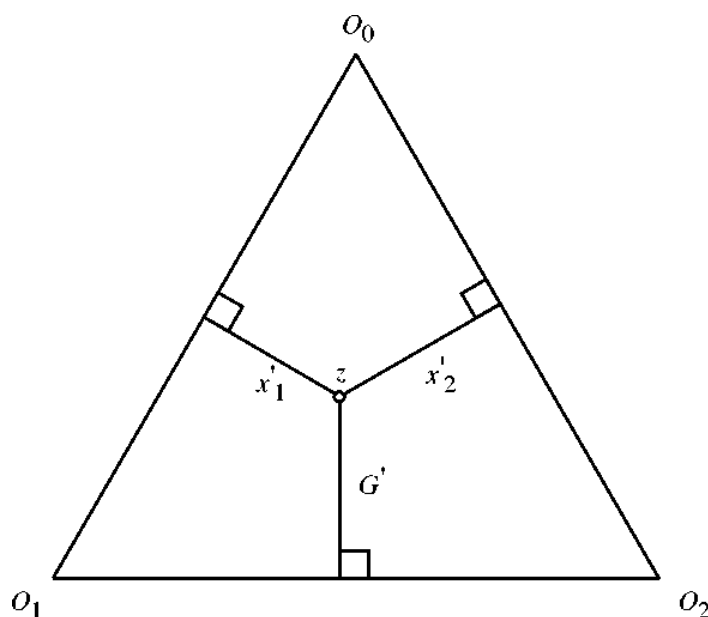


Рисунок 2.1. Допустимое распределение (x'_1, x'_2, G') в треугольнике Кольма.
Источник: [63]

В треугольнике Кольма эксплуатируется тот факт, что сумма высот равностороннего треугольника постоянна. Поэтому можно понимать одну из высот как потребление общественного блага, а две другие – как потребление частного блага первым и вторым агентом, соответственно. Если x_1 – потребление частного блага первым агентом, x_2 – вторым агентом, а G – потребление общественного блага, то должно выполняться общее ограничение:

$$x_1 + x_2 + G = \text{const.}$$

Если функции полезности агента i обозначить $U_i(x_i, G)$, то условие оптимальности по Парето записывается в виде уравнения Сэмюэльсона

$$\sum_i \frac{\frac{\partial U_i(x_i^*, G^*)}{\partial G}}{\frac{\partial U_i(x_i^*, G^*)}{\partial x}} = \sum_i MRS_i(x_i^*, G^*) = 1 = MRT,$$

где MRS_i – предельная норма замещения частного блага общественным благом для агента i , а MRT – норма преобразования частного блага в общественное благо, равная в данном случае 1 для обоих агентов.

Данное условие можно переписать в виде:

$$\alpha_1^* \frac{\partial U_1(x_1^*, G^*)}{\partial x_1} = 1; \alpha_2^* \frac{\partial U_2(x_2^*, G^*)}{\partial x_2} = 1; \alpha_1^* \frac{\partial U_1(x_1^*, G^*)}{\partial G} + \alpha_2^* \frac{\partial U_2(x_2^*, G^*)}{\partial G} = 1,$$

где

$$\alpha_i^* = \frac{1}{\partial U_1(x_1^*, G^*)}.$$

Векторы $(q_1, g_1, q_2, g_2) \in R_+^4$, удовлетворяющие условию:

$$\alpha_1 q_2 = 1; \quad \alpha_2 q_2 = 1; \quad \alpha_1 g_1 + \alpha_2 g_2 = 1;$$

при некоторых неотрицательных α_i , образуют выпуклый конус Q в пространстве R_+^4 . Уравнение Сэмюэльсона означает, что

$$\text{grad}U = \left(\frac{\partial U_1(x_1^*, G^*)}{\partial x_1}, \frac{\partial U_1(x_1^*, G^*)}{\partial G}, \frac{\partial U_2(x_2^*, G^*)}{\partial x_2}, \frac{\partial U_2(x_2^*, G^*)}{\partial G} \right) \in Q.$$

Если же это условие не выполняется, то состояние не является оптимальным по Парето. Его можно улучшить сразу для обоих агентов. Для этого надо спроектировать $\text{grad}U$ на конус Q , а затем вычесть проекцию из $\text{grad}U$. В результате получим направление, в одинаковой мере улучшающее положение обоих экономических агентов. В самом деле, разность между градиентом и его проекцией $q^* = (q_1^*, g_1^*, q_2^*, g_2^*)$ на конус – это вектор $\text{grad}U - q^*$, составляющий с градиентом острый угол, т.е. их скалярное произведение $(\text{grad}U - q^*, \text{grad}U)$ положительно. При этом величина $(\text{grad}U - q, \text{grad}U - q)$ достигает при $q = q^*$ минимума по всем q из Q .

Этот минимум достигается при выполнении условий:

$$\frac{\partial U_1(x_1^*, G^*)}{\partial x_1} - q_1^* = \frac{\partial U_2(x_2^*, G^*)}{\partial x_2} - q_2^*; \quad \frac{\partial U_1(x_1^*, G^*)}{\partial G} - g_1^* = \frac{\partial U_2(x_2^*, G^*)}{\partial G} - g_2^*,$$

откуда следует, что направление движения – это либо одновременное увеличение вложений в общественное благо, либо одновременное их уменьшение.

С вычислительной точки зрения процесс обмена мало отличается от того, что используется при обмене ресурсами. На Рисунке 2.2 это направление идет в точности в центр «линзы», образуемой линиями уровня функций полезности экономических агентов. В точке оптимума по Парето имеет место касание линий уровня. Если при этом выполняются еще и бюджетные

ограничения, то имеет место равновесие Линдаля. Такое состояние изображено на Рисунке 2.3.

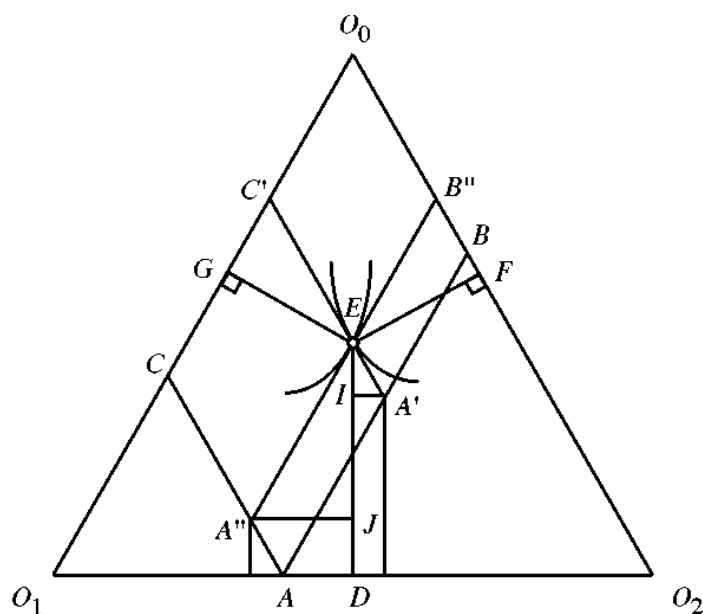


Рисунок 2.2. Индивидуально оптимальное, но не оптимальное по Парето состояние (равновесие Нэша). Источник: Автор по [63]

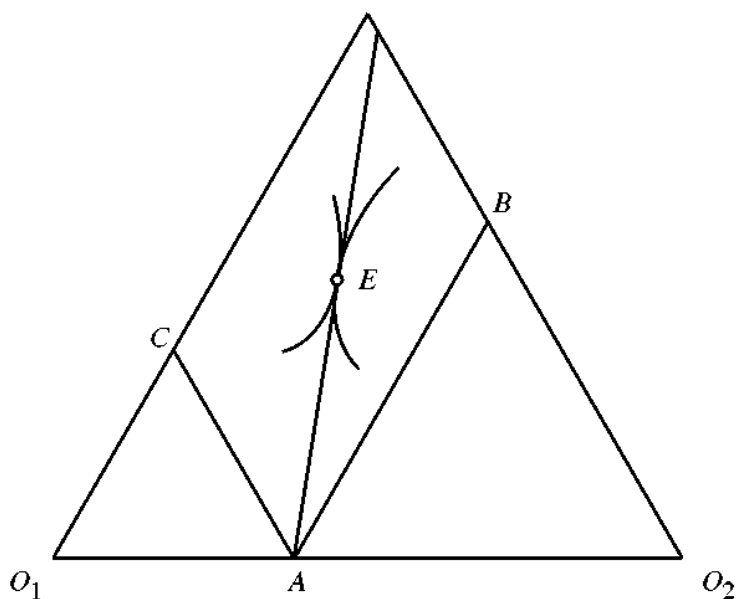


Рисунок 2.3. Равновесие Линдаля. Источник: [63]

Чтобы распространить предложенную выше вычислительную схему на случай произвольного числа продуктов и произвольного числа агентов, следует несколько изменить обозначения. Частные блага, число которых теперь

произвольно, составляют множество J . Общественные блага получают номера k из K . Номера агентов по-прежнему обозначаем через i , где i принадлежит I . Потребление агентом i частного блага j обозначим x_{ij} , потребление общественного блага k обозначим G_k . Далее, q_i – вектор размерности пространства частных благ, g_i – вектор размерности пространства общественных благ. Конус Q теперь будет определяться условиями

$$\alpha_i q_i = q; i \in I; \sum_{i \in I} \alpha_i g_i = g, q \in S(J), g \in S(K).$$

Здесь $S(J)$, $S(K)$ – симплексы в пространстве частных и общественных благ. Далее все аналогично случаю, рассмотренному выше, а именно, проектируем градиент функции $U = (U_1, U_2, \dots, U_{|I|})$ на конус Q , вычитаем из него проекцию и получаем направление, движение по которому локально выгодно всем экономическим агентам. В случае, если функции полезности кусочно-линейны, движение осуществляется до тех пор, пока градиент вычисляется однозначно. Далее движение возможно, если перейти к субдифференциалам и решать на каждом шаге чуть более сложные экстремальные задачи.

2.2.2. Равновесие Штакельберга

В концепции равновесия по Штакельбергу предполагается, что на рынке действуют два продавца, причем один из них – лидер или ведущий, а второй – ведомый. В принципе, однако, ведомых поставщиков может быть и больше. Важно, что лидер один, а увеличение числа ведомых не имеет принципиального значения, их можно считать за одного игрока, поведение которого вычислимо. Лидер исходит из того, что расширением своих поставок он провоцирует снижение цен, а потому делает отрасль менее прибыльной и вынуждает конкурента снижать свои поставки.

При этом лидер «знает», что конкурент (ведомый) действует по Курно, т.е. максимизирует свою прибыль, исходя из предположения, что поставки продукта на рынок лидером заданы.

В случае с общественным благом ситуация в целом похожа, также есть ведущий и ведомый. Однако, когда ведущий вкладывает средства в

общественное благо, пользу от этого получает и он сам, и ведомый. Далее ведомый действует так, как будто уровень общественного блага задан. Он максимизирует свою прибыль, исходя из этого условия, т.е. на вербальном уровне его задача воспринимается так же, как и в случае с обычными продуктами.

Рациональное поведение лидера состоит в том, чтобы максимизировать прибыль с учетом поведения конкурента (ведомого). Формально это можно записать в виде задачи на максимум прибыли по одной переменной q_1 :

$$\pi_1(q_1, q_2(q_1)) \rightarrow \max_{q_1}, q_2(q_1) = \operatorname{argmax}_{q_2} \pi_2(q_1, q_2).$$

Здесь предполагается, что $q_2(q_1)$ определяется однозначно, хотя для этого надо наложить дополнительные условия на зависимость между предложением и ценами. Равновесие Штакельберга – такая пара (q_1^*, q_2^*) , что

$$\pi_1(q_1^*, q_2(q_1^*)) = \max_{q_1} \pi_1(q_1, q_2(q_1)), q_2(q_1) = \operatorname{argmax}_{q_2} \pi_2(q_1, q_2).$$

Разумеется, равновесие Штакельберга также приводит к не самому лучшему решению, поскольку, зная уровень общественного блага, обеспечиваемый ведущим, ведомый будет вести себя оппортунистически. Иначе говоря, он будет стараться использовать свой частный продукт только на свое благо, а не на общественное. Зная это, ведущий не будет вкладывать свой продукт в общественное благо. А в результате оба проиграют.

Решение можно радикально улучшить, если слегка изменить правила игры. А именно: ведущий делает ход, объявляя субсидию в определенном объеме на каждый рубль, вложенный ведомым в общественное благо, т.е. берет на себя некое обязательство. В результате ведомый больше заинтересован во вложениях в общественное благо. А дальше все по той же схеме, ведущий должен выбрать оптимальный размер субсидии, исходя только из своих интересов. Решение, принятое в результате, будет много лучше, чем в равновесии по Нэшу или прямолинейно по Штакельбергу.

2.2.3. Равновесие Линдаля

Идеальным выходом из противоречия могла бы служить такая система цен на цифровые продукты и иные общественные блага, поставляемые в

частном порядке, при которой цены копий индивидуализированы для разных потребителей или, как минимум, для разных групп потребителей. В принципе, назначение разных цен для разных покупателей – давно известная практика, получившая название ценовой дискриминации и запрещенная в большинстве стран с рыночной экономикой. В частности, в США эта практика запрещена актом Клейтона (дополнение к антитрестовскому закону). Считается, что ценовая дискриминация – инструмент, позволяющий монополисту увеличить свою прибыль за счет потребителей. Отчасти это верно. Однако в случае цифровых продуктов, когда издержки изготовления дополнительных копий продукта близки к нулю, увеличение прибыли производителя (монополиста) происходит в основном за счет расширения числа покупателей. Это легко показать графически.

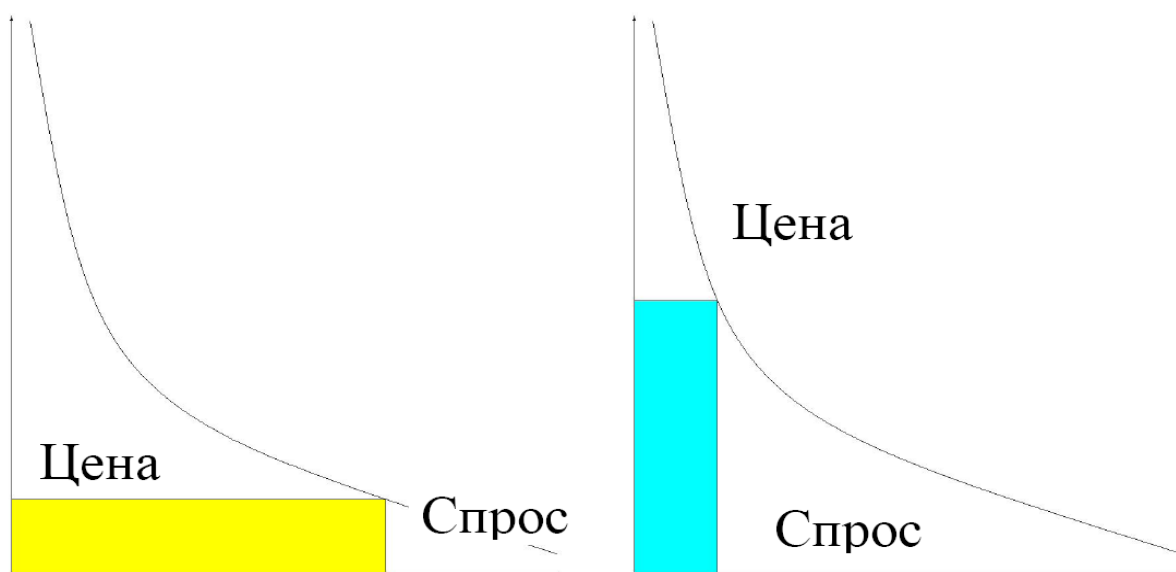


Рисунок 2.4. Независимость выручки от цены при обратной пропорциональности спроса и цены блага. Источник: А.Н. Козырев.

Выручка от продаж (произведение цены на спрос) представлена прямоугольником, вписанным в пространство под гиперболой. Поскольку спрос обратно пропорционален цене, их произведение постоянно, т.е. выручка не зависит от цены. Примечательно, что в рамках предположения о такой зависимости между спросом и ценой, потенциальные возможности получения выгоды от ценовой дискриминации почти безграничны. А именно: при назначении

разных цен для разных категорий покупателей выручка может быть представлена как площадь фигуры, получаемой путем объединения нескольких вписанных прямоугольников. Если принять дополнительное предположение о том, что покупатели очень мелкие, число их стремится к бесконечности, а цена для каждого из них индивидуальна, то выручка стремится к бесконечности.

Разумеется, предположение об обратной зависимости спроса от цены – не вполне соответствует действительности. В реальной экономике спрос повышается при снижении цены, но зависимость может быть не столь простой.

Теоретически можно построить совершенный экономический механизм, обеспечивающий доступ к представленным в цифровой форме продуктам. Его идейную основу могла бы составить модель равновесия Линдаля с системой индивидуализированных цен на такие продукты. В такой системе все потребляют продукт на одном уровне – самом высшем, но каждый платит за продукт по цене, которая соответствует полезности продукта именно для него. За счет этого полностью используется потенциал неконкурентности в потреблении, а суммы, собираемые с потребителей и направляемые на разработку таких продуктов, максимальны. Формальное описание модели равновесия Линдаля предполагает формальное описание экономики с общественными благами.

В рассматриваемой экономике имеется конечное число потребителей. Их совокупность обозначается символом I . Произвольный потребитель i – элемент этого множества. Число благ тоже конечно, K_1 – множество всех общественных благ, K_2 – множество всех частных благ (ресурсов). Вектор x_i состоит из двух частей $\{x_k\}$ и $\{x_{ik}\}$. Первая его часть, соответствующая потреблению общественных благ, для всех i одинакова, вторая – индивидуальна для каждого i . В целом потребительский набор i -го потребителя имеет вид

$$x_i = (\{x_k\}_{k \in K_1}, \{x_{ik}\}_{k \in K_2}) = (x^{(1)}, x_i^{(2)}).$$

Множество допустимых потребительских наборов i -го потребителя X_i удобно представить в виде:

$$X_i = X^{(1)} \times X_i^{(2)},$$

так что

$$\mathbf{x}_i = (\mathbf{x}_i^{(1)}, \mathbf{x}_i^{(2)}) \in X_i$$

тогда и только тогда, когда

$$\mathbf{x}^{(1)} \in X^{(1)} \text{ и } \mathbf{x}_i^{(2)} \in X_i^{(2)}.$$

Состояние (\mathbf{x}, \mathbf{y}) экономики с общественными благами является допустимым, если выполнены следующие соотношения (напомним, что начальные запасы общественных благ мы считаем равными нулю):

$$\begin{aligned} \mathbf{x}_i &\in X_i, \forall i \in I, \\ x_k &= \sum_{j \in J} y_{jk}, \forall k \in K_1, \forall i \in I, \\ \sum_{i \in I} x_{ik} &= \sum_{j \in J} y_{jk} + \sum_{i \in I} w_{ik}, \forall k \in K_2, \\ g_j(y_j) &\geq 0, \forall j \in J. \end{aligned}$$

Здесь J – совокупность всех производителей, y_{jk} – производство продукта k производителем j . При этом отдельные компоненты могут быть отрицательными, что означает затраты соответствующего продукта в производстве.

Первое из условий означает, что вектор потребления каждого агента находится в пределах его физических возможностей потребления. Второе условие – баланс по общественным благам, их производство всеми производителями в точности равно их потреблению каждым потребителем. Третье условие – баланс по частным благам. Их суммарное потребление всеми потребителями равно сумме начальных запасов и производства. При этом учитываются затраты частных благ в производстве других благ (частных или общественных). Наконец, последнее условие – технологические ограничения на производство.

В том, что касается цен, все обстоит ровным счетом наоборот: цены $\{q_{ik}\}$ на каждое общественное благо индивидуальны, а цены $\{p_k\}$ на частные блага для всех одинаковы.

Задача потребителя i при бюджете β_i имеет следующий вид:

$$u_i(x_i) \rightarrow \max_{x_i}$$

$$\sum_{k \in K_1} q_{ik} x_k + \sum_{k \in K_2} p_k x_{ik} \leq \beta_i,$$

$$x_i \in X_i.$$

Задача производителя имеет обычный вид

$$p y_j \rightarrow \max_{y_j}$$

$$q_j(y_j) \geq 0, \forall j.$$

Четверка $(\bar{p}, \bar{q}, \bar{x}, \bar{y})$ называется равновесием Линдаля, если:

1. (\bar{x}, \bar{y}) — допустимое состояние экономики с общественными благами;
2. каждый набор \bar{x}_i является решением соответствующей задачи потребителя при ценах \bar{p} и индивидуализированных ценах общественных благ $\{\bar{q}_{ik}\}_{k \in K_1}$ и доходах

$$\beta_i = \bar{p} w_i + \sum_{j \in J} \gamma_{ij} \bar{p} \bar{y}_j + S_i;$$

3. технология \bar{y}_j является решением задачи производителя при ценах \bar{p} .
4. сумма индивидуализированных цен общественного блага равна цене производителя:

$$\sum_{i \in I} \bar{q}_{ik} = \bar{p}_k \quad \forall k \in K_1.$$

Иначе говоря, сумма индивидуальных цен потребления общественного блага должна быть равна цене его производства. Проблема, однако, в том, что такие цены очень трудно рассчитать. Еще труднее добиться их реального применения, в том числе по причине возможной перепродажи продукта, манипулирования информацией о потребностях и другими формами оппортунистического поведения. Поэтому необходимо дополнение модели игровыми конструкциями, делающими оппортунистическое поведение невыгодным.

2.3. Уравнение Сэмюэльсона и оптимальные цены общественных благ

Строгое математическое доказательство необходимости дифференциации цен на общественное благо требует полной формализации рассматриваемой ситуации. Дифференцированные (линдалевские) цены следуют из уравнения Сэмюэльсона, а само уравнение получается как необходимое условие достижения оптимальности по Парето, т.е. из решения некоторой оптимизационной задачи или набора таких задач.

Поскольку цены в моделях экономического равновесия – это предельные цены, они соответствуют производным функций полезности по соответствующим переменным. А переменные в функциях полезности – это частные и общественные блага. Найдя эти производные, можно представить, какими должны быть цены на обычные (частные) и на общественные блага. Остается только проделать то же рассуждение формально, используя математику.

Число потребителей в рассматриваемой экономике произвольно, но конечно. Их совокупность далее обозначается символом I . Произвольный потребитель i – элемент этого множества. Число благ тоже конечно, в данном случае K_1 – множество всех общественных благ, K_2 – множество всех частных благ или ресурсов, используемых в производстве. Вектор потребления \mathbf{x}_i каждого потребителя состоит из двух частей $\{x_k\}$ и $\{x_{ik}\}$. Первая его часть соответствует потреблению общественных благ, она для всех i одинакова, вторая часть – индивидуальна для каждого i . В целом потребительский набор i -го потребителя имеет вид

$$\mathbf{x}_i = (\{x_k\}_{k \in K_1}, \{x_{ik}\}_{k \in K_2}) = (\mathbf{x}^{(1)}, \mathbf{x}_i^{(2)}).$$

Множество допустимых потребительских наборов i -го потребителя X_i удобно представить в виде:

$$X_i = X^{(1)} \times X_i^{(2)},$$

так что

$$\mathbf{x}_i = (\mathbf{x}^{(1)}, \mathbf{x}_i^{(2)}) \in X_i$$

тогда и только тогда, когда

$$\mathbf{x}^{(1)} \in X^{(1)} \text{ и } \mathbf{x}_i^{(2)} \in X_i^{(2)}.$$

Состояние (\mathbf{x}, \mathbf{y}) экономики с общественными благами допустимо, если выполнены следующие соотношения (предполагается, что начальные запасы общественных благ равны нулю):

$$\begin{aligned} \mathbf{x}_i &\in X_i, \forall i \in I, \\ x_k &= \sum_{j \in J} y_{jk}, \forall k \in K_1, \forall i \in I, \\ \sum_{i \in I} x_{ik} &= \sum_{j \in J} y_{jk} + \sum_{i \in I} w_{ik}, \forall k \in K_2, \\ g_j(y_j) &\geq 0, \forall j \in J. \end{aligned}$$

Здесь J – совокупность всех производителей y_{jk} – производство продукта k производителем j . При этом отдельные компоненты могут быть отрицательными, что означает затраты соответствующего продукта в производстве.

Первое из условий означает, что вектор потребления каждого агента находится в пределах его физических возможностей потребления. Второе условие – баланс по общественным благам, их производство всеми производителями в точности равно их потреблению каждым потребителем. Третье условие – баланс по частным благам. Их суммарное потребление всеми потребителями равно сумме начальных запасов и производства. При этом учитываются затраты частных благ в производстве других благ (частных или общественных). Наконец, последнее условие – технологические ограничения на производство.

Допустимое состояние экономики $(\hat{\mathbf{x}}, \hat{\mathbf{y}})$ оптимально по Парето тогда и только тогда, когда оно доставляет решение следующему набору оптимизационных задач ($i_0 = 1, \dots, m$):

$$\begin{aligned} u_{i_0}(x_{i_0}) &\rightarrow \max_{\mathbf{x}, \mathbf{y}}, \\ \mathbf{x}_i &\in X_i, \forall i \in I, \end{aligned}$$

$$\begin{aligned}
u_i(\mathbf{x}_i) &\geq u_i(\hat{\mathbf{x}}_i), \quad \forall i \neq i_0, \\
g_j(\mathbf{y}_j) &\geq 0, \quad \forall j \in J, \\
x_k &= \sum_{j \in J} y_{jk}, \quad \forall k \in K_1, \\
\sum_{i \in I} x_{ik} &= \sum_{j \in J} y_{jk} + \sum_{i \in L} w_{ik}, \quad \forall k \in K_2.
\end{aligned}$$

Последнее равенство – материальный баланс для общественных благ. В нем и заключается отличие этой задачи от соответствующей задачи для классической экономики с частными благами.

Лагранжиан, соответствующий приведенному выше набору задач, с точностью до констант $u_i(\hat{\mathbf{x}}_i)$ имеет вид:

$$\begin{aligned}
L &= \sum_{i \in I} \lambda_i u_i(\mathbf{x}_i) + \sum_{j \in J} \mu_j g_j(\mathbf{y}_j) + \\
&+ \sum_{k \in K_1} \sigma_k \left(\sum_{j \in J} y_{jk} - x_k \right) + \sum_{k \in K_2} \sigma_k \left(\sum_{j \in J} y_{jk} + \sum_{i \in L} w_{ik} + \sum_{i \in I} x_{ik} \right).
\end{aligned}$$

Если $\hat{\mathbf{x}}_i \in \text{int}X_i$, а функции полезности $u_i(\cdot)$ и производственные функции $g_j(\cdot)$ дифференцируемы для всех i, j , то можно получить условия оптимальности по Парето, приравнявая нулю частные производные лагранжиана.

Для любой из задач ($i_0 = 1, \dots, m$) существуют (не все равные нулю) множители Лагранжа $(\lambda_i, \mu_j, \sigma_k)$ такие, что $\lambda_i \geq 0, \forall i, \mu_j \geq 0, \forall j$, и

$$\frac{\partial L}{\partial x_k} = 0, \quad \forall k \in K_1, \quad \frac{\partial L}{\partial x_{ik}} = 0, \quad \forall i \in I, \quad \forall k \in K_2, \quad \frac{\partial L}{\partial y_k} = 0, \quad \forall j \in J, \quad \forall k \in K.$$

Формально это следует из теоремы Д. Фрица. При естественном дополнительном условии регулярности (линейной независимости градиентов ограничений в каждой задаче i_0) можно выбрать такие множители Лагранжа, что $\lambda_{i_0} = 1$. Содержательно регулярность означает, что для каждого потребителя существует частное благо, увеличение потребления которого желательно, а для каждого производителя существует частное благо, без затрат которого невозможно увеличить потребление общественных благ. Формально это

выглядит так: для каждого i существует $k \in K_2$, такой что $\partial u_i(\mathbf{x}_i)/\partial x_{ik} > 0$, а для каждого j существует $k \in K_1$, такой что $\partial g_j(\mathbf{y}_j)/\partial y_{jk} < 0$.

$$\sum \frac{\partial u_i(\hat{\mathbf{x}}_i)/\partial x_k}{\partial u_i(\hat{\mathbf{x}}_i)/\partial x_{ik_0}} = \frac{\partial g_j(\hat{\mathbf{y}}_j)/\partial y_{jk}}{\partial g_j(\hat{\mathbf{y}}_j)/\partial y_{jk_0}}, \quad \forall i \in I, \quad \forall j \in J, \quad \forall k \in K_1,$$

$$\frac{\partial u_i(\hat{\mathbf{x}}_i)/\partial x}{\partial u_i(\hat{\mathbf{x}}_i)/\partial x_{ik_0}} = \frac{\partial g_j(\hat{\mathbf{y}}_j)/\partial y_{jk}}{\partial g_j(\hat{\mathbf{y}}_j)/\partial y_{jk_0}}, \quad \forall i \in I, \quad \forall j \in J, \quad \forall k \in K_2,$$

Первое из полученных равенств называют уравнением Сэмюэльсона. Оно говорит, что сумма предельных норм замещения общественного блага на частное в потреблении равна предельной норме замещения общественного блага на частное в производстве. Иначе говоря, сумма индивидуальных цен для потребителей общественного блага равна цене его производства. Именно это соотношение называется уравнением Сэмюэльсона. Для частных благ цены потребления одинаковы для всех потребителей и равны ценам производства. Ничего подобного для общественных благ требовать нельзя, так как это несовместимо с оптимумом.

Следует также отметить, что Сэмюэльсон и его последователи предполагали, что производство и уровень потребления общественных благ для всех потребителей одинаковы. В более поздних моделях равенство было заменено операцией взятия максимума.

2.4. Подходы к назначению цен

Модели ценообразования на частное благо – хорошо разработанный и давно изученный в экономической литературе вопрос. Основой для рассуждений здесь является модель экономики Вальраса, в которой производители определяют количество выпускаемого ими продукта, максимизируя прибыль, а потребители выбирают количество приобретаемого продукта, максимизируя свою полезность. Единая для всех цена, индивидуальные значения производства и предложения при этом считаются равновесием, если ни один из участников рынка не заинтересован отклоняться от принятого решения.

Детеринг, исследуя экономику медиаконтента, обращает внимание на то, что следование теории – как рыночной конкуренции, так и монополии,

олигополии или рыночной конкуренции – приводит к потерям общественного благосостояния в случае торгуемых на рынке общественных благ [25]. Основная причина состоит в тесной связи между ценой и предельными издержками, которая установлена на теоретическом уровне. Но должна ли цена быть нулевой, если предельные издержки каждой отдельной единицы товара стремятся к нулю? А если продукт общий для всех (компьютерная программа, кинофильм или музыка), что означает увеличение его поставки: дополнительные строчки кода, дополнительные минуты производства?

Отправной точкой в разрешении этих вопросов служит допущение о том, что потребителям с различной готовностью к оплате следует назначить различные цены за использование одного и того же продукта. Эти цены называются в экономике ценами Линдаля, и они определяют равновесие в модели, отличной от Вальраса. Как и у последнего, в модели Линдаля цена частного блага общая для всех потребителей, и каждый выбирает объём потребления, максимизируя индивидуальную полезность, но производитель общественного блага даёт один для всех потребителей объём потребления по цене, которая, с одной стороны, равна предельным издержкам производства и, с другой стороны, соответствует сумме индивидуальных цен (платежей). Одно из следствий модели Линдаля состоит в том, что с ростом числа потребителей поставщик общественного блага может собрать больше денег с рынка, так что сумма будет соответствовать предельным издержкам поставки продукта в большем объёме. Говоря в терминах информационных продуктов, можно сказать о доступности более функционального программного обеспечения. Аргументация Детеринга в ответ на несправедливость такой схемы оплаты, при которой, казалось бы, одинаковые индивиды платят разные суммы денег, может быть усилена в случае информационных продуктов. Общий аргумент состоит в том, что, обеспечивая поставку общественных благ, государство проводит дифференцированную налоговую политику – по размеру предприятий, отраслям, товарам, доходам граждан. Дополнение в случае информационных продуктов связано с тем, что их потребители имеют различные возможности по

реализации преимуществ, полученных в ходе потребления блага. Например, потребители общей базы данных имеют различные инструменты для её обработки. Соответственно, сочетания данных и инструментов приводят к различным реализациям значений колмогоровской сложности, получению информации и её полезности в принятиях решений. Также свою роль может играть время: для одних потребителей важна возможность наиболее раннего доступа к продукту и, соответственно, их готовность к оплате при этом возрастает в надежде на ускорение разработок.

Индивидуальные цены имеют параллели с ценовой дискриминацией по Пигу. Классическое понимание ценовой дискриминации включает в себя три степени. При ценовой дискриминации первой степени каждый потребитель платит максимально возможную для него цену. Дискриминация второй степени предполагает сегментацию рынка: цены зависят от объёма потребления или функциональных характеристик продукта. При ценовой дискриминации третьей степени производитель разделяет группы потребителей – по возрасту, статусу, региону проживания и т.п. – и назначает свои цены для каждой группы потребителей. Однако цены Линдаля не укладываются в эту схему. Действительно, в равновесии Линдаля потребители сохраняют некоторый излишек (этого не происходит при ценовой дискриминации первой степени). Потребители получают доступ к одному и тому же благу по индивидуальным ценам Линдаля (при ценовой дискриминации второй степени качество или объём предложения меняются от сегмента к сегменту). Также в отличие от ценовой дискриминации третьей степени, цены Линдаля никого не исключают из потребления и позволяют достичь более высокого уровня благосостояния.

Хотя цены Линдаля имеют ряд преимуществ и, как показано в теории, являются единственным способом достижения баланса Сэмюэльсона, при котором вклад каждого отдельного потребителя точно равен его предельной полезности, на практике возникает вопрос о способах назначения таких цен. Эти обстоятельства заставляют обратиться к моделям ценообразования на продукты, имеющие черты общественных, коллективных или клубных благ. Эти

модели обсуждаются уже в следующем подразделе. Однако обращение к моделям стоит предварить комментарием о ценах на информационные продукты, которые можно отнести к частным благам.

Как говорилось выше, индивидуальные продукты, конкурирующие в потреблении, находят своё место в игровой индустрии или виртуальных вселенных. Можно говорить о приобретении такого продукта по предельным издержкам разработки. Однако следует учитывать, что виртуальные миры подконтрольны одной фирме – монополисту, предлагающему потребителям соответствующие товары. Поэтому следует также указать на равенство цены предельной выручке. Но и этим ситуация не исчерпывается. Уникальные игровые предметы, введённые в оборот разработчиком, становятся самостоятельным товаром на P2P-рынке (peer-to-peer) среди игроков или пользователей виртуальной вселенной. И эти предметы продаются на аукционах – совсем как предметы искусства. Аналогия с художественными произведениями становится особенно уместной в контексте стимминга⁶: пользователи готовы приобретать визуально привлекательные предметы (одежда игровых персонажей, модифицированные изображения стандартных игровых предметов и т.п.) для использования в своих трансляциях, выводя, таким образом, эстетическую составляющую на более высокую позицию по сравнению с функциональной. Поэтому аукционы первой цены – реально используемый на практике механизм ценообразования таких благ.

Модели ценообразования

Поскольку основным фокусом работы являются условия для совершения сделок с использованием информационных продуктов, их стоимостная оценка и активное вовлечение в гражданский оборот, обзор методов ценообразования следует ограничить рыночными общественными благами (по

⁶ Деятельность, при которой пользователь или игрок транслирует игровой процесс в сеть. Популярные игроки могут собрать большую аудиторию на свои трансляции в реальном времени или их записи. Доход от трансляций имеет несколько составляющих. Во-первых, зрители вносят добровольные пожертвования во время трансляции или покупают подписку на платных площадках. Во-вторых, во время трансляций пользователи имеют возможность заказать товары, недоступные для продажи в другое время. В-третьих, возможен рекламный доход. От английского stream – «поток», «трансляция».

классификации Адамса-МакКормака) или, что эквивалентно, общественными благами, поставляемыми в частном порядке. Классические модели относятся к рынку однородного блага, что является очень сильным, хотя и мало реалистичным допущением, поскольку потребитель в действительности сталкивается с набором продуктов заменителей: одни отличаются составом данных, другие – функциональностью программного обеспечения, третьи – их сочетанием. Модели пространственной конкуренции, которые берут своё начало с теории монополистической конкуренции Чемберлина [19], хотя исследуют поставку неоднородной продукции, в большей степени затрагивают вопрос оптимального для благосостояния разнообразия продукции. Они обозначаются лишь кратко в конце подраздела.

Исторически первой моделью поставки общественных благ в частном порядке в условиях конкуренции принято считать ту, которая изложена в работе Томпсона [110]. Предполагается рынок однородного товара, неконкурентного в потреблении. Потребитель сталкивается с уменьшающейся предельной полезностью каждой последующей единицы блага (например, минуты эфирного времени). С другой стороны, множество производителей обеспечивают выпуск блага с возрастающими предельными издержками. На рынке отсутствуют барьеры для доступа новых производителей. Наконец, производители дифференцируют цены между индивидами. Каждый индивид, однако, платит одинаковую фиксированную цену за каждую единицу приобретаемого им блага. В равновесии требуется 1) равенство предельных издержек предельным доходам производителей и 2) равенство цены предельной полезности потребителя.

В равновесии, однако, конкуренция производителей приводит к уровню поставки, который оказывается ниже оптимального по Парето при монопольных поставках. В условиях конкуренции производители снижают цены, что приводит к снижению общего уровня и, соответственно, недопоставки блага. Условие о том, что потребители приобретают каждую единицу блага по фиксированных для них (индивидуальных) ценах приводят к тому, что в условиях

убывающей предельной полезности даже интрамаргинальные единицы блага приобретаются по цене последней единицы. Это приводит к увеличению его потребительского излишка. На это снижение цены со стороны потребителей Томпсон предлагает реагировать введением специального налога. Многоуровневые связи между ценой и спросом (на уровне производителей и потребителей) приводят к мало реалистичным предположениям о том, что потребитель при каждом решении о покупке учитывает не только соотношение цены со своей предельной полезностью, но и воздействие этого решения на снижение цены для всех потребителей. Но наиболее значимое для текущего диссертационного исследования положение состоит в том, что цены в равновесии индивидуальны и соответствуют предельной готовности к оплате (по предельной полезности).

Модель Демсетца [24] опирается не те же предположения, что и модель Томпсона, предполагая исключение неплательщиков без издержек. Например, при помощи технических средств. Ключевым отличием является то, что потребители не участвуют в формировании цен – они принимают цены производителей. Равновесие модели соответствует равновесию Линдаля: индивидуальные цены равны индивидуальным предельным полезностям, и сумма всех предельных полезностей равна предельным издержкам производства. Равновесие имеет общие черты с ценовой дискриминацией третьей степени, поскольку потребители различаются по группам – каждая из них платит свою цену. При этом группа с низким спросом может получать благо бесплатно, если её включение не увеличивает издержки.

Окленд [83] исключил из предпосылок своей модели возможность разделения потребителей на группы, что делает невозможной последующую ценовую дискриминацию. В остальном его модель соответствует ситуации Демсетца: однородное общественное благо поставляется на рынке с возможностью исключения потребителей. Если все потребители имеют одинаковые предпочтения, то и цена продукта распределяется между ними пропорционально их числу. Этот вырожденный случай (неразличимость потребителей)

соответствует равновесию Линдаля. Но, если предпочтения различны, поставщик может предложить набор цен – каждая соответствует предпочитаемому объёму потребления. Эта модель также позволяет оценивать потери благосостояния в специальных состояниях рынка и эффективность финансирования поставки общественного блага, например, через взимаемые налоги.

Основополагающие модели поставки общественных благ в частном порядке органично сочетаются с концепцией индивидуальных цен по Линдалю. Общим ограничением моделей является случай однородной продукции. Реальные рынки, однако, предлагают потребителям выбор и возможность замещения продукта не полной копией, но достаточно близким аналогом. Так, на рынке информационных продуктов потребители ищут нужный им функционал, реализацию которого обеспечивают разные продукты. Если от продукта ожидаются относительно простые операции, потребитель вполне может ограничиться использованием бесплатной версии или свободно распространяемым программным обеспечением. Теория монополистической конкуренции, предложенная Чемберлином, позволяет исследовать рынки неоднородных продуктов, но обзор соответствующих моделей представляется избыточным для целей данного диссертационного исследования по нескольким причинам. Во-первых, равновесные модели рынков представляются некоторой рамкой в контексте общей работы о сделках на рынке информационных продуктов. Базовая проблема видится в том, что на этом рынке сходятся не просто потребитель и поставщик, а зависимые друг от друга агенты. Например, один предлагает данные, другой – программу для их обработки. Их взаимодействие оказывается полезным для получения новой информации и, соответственно, снижения ошибок принятия решений. При этом ни первый, ни второй могут не иметь действительно массового рынка, и каждый из участников ограничивается несколькими потенциальными контрагентами. В такой ситуации всё ещё можно говорить о моделях для поиска равновесия, например, на олигополистическом рынке, но найденные значения не следует интерпретировать как императив при совершении сделки. Скорее, модели равновесия показывают ориентир, от

которого можно отталкиваться в конкретных сделках. Далее, имея некоторую отправную точку, можно двигаться в направлении улучшения цены. Как этого добиться, сказано ниже в Разделе 3. Агенты на рынке информационных продуктов не только не сталкиваются с однородностью благ, но и сами не являются компаниями чистой игры. Современные гиганты ИТ-отрасли предлагают целый спектр товаров и услуг (технологии искусственного интеллекта, финансы, служба доставки, беспилотные технологии и т.п.), поэтому остаётся место для моделей дополняющих товаров. Например, моделей, которые развивают работы Маршала о поставке блага-сырья для двух различных рынков однородных продуктов. Моделей может быть много, и обзреть их все не представляется возможным. Задача подраздела уже выполнена: информационные продукты описаны в терминах теории общественных благ, и показаны основные следствия такого описания.

Одно из таких следствий – индивидуальные цены – являются второй причиной, по которой следует ограничить обзор. Индивидуальные цены позволяют снизить потери благосостояния, как показывает теория. В методе ценообразования, который предлагает диссертационная работа, индивидуальные цены становятся результатом взаимодействия на рынке информационных продуктов, при котором стороны приходят в оптимальное по Парето состояние. И ценовая дифференциация также обнаруживает себя в моделях неоднородных продуктов, о которых следует сказать в заключении данного подраздела.

Анализ монополистической конкуренции, выполненный Вэриэном [145], показывает, что она сталкивается с той же проблемой, что и монополия или рынок общественных благ – с достижением оптимального уровня производства по удельным ценам. Равновесие на рынке монополистической конкуренции всё ещё связано с потерями общественного благосостояния, и дифференцирование помогло бы в их сокращении. Но простой перенос результатов моделей Томпсона, Демсетца и Окленда невозможен, поскольку на многопродуктовом рынке неизбежно возникает вопрос о балансе между разнообразием и качеством товаров. Для исследования этих вопросов привлекают, например,

модели Штайнера [109], Ротенберга [94] и Биба [11]. В модели поставки бесконечного количества возможных вариантов продукта, финансируемого за счёт рекламы, рынок неизбежно приходит к дублированию продуктов – самых популярных. Так что монополистическое решение оказывается даже лучше для общества в целом. В модели Спенса и Оуэна [108] также обеспечивается недостаточное разнообразие предложения – исключаются продукты с высокими производственными затратами. В моделях пространственной конкуренции – Сэлопа [98] и Хотеллинга [47] – можно добиться оптимального сочетания разнообразия и качества при дифференцированном ценообразовании.

Моделям ценообразования на информационные продукты и на программное обеспечение в частности посвящены и отечественные работы, которые следует отметить в данном подразделе диссертации. О.В. Капустина исследовала равновесные модели с линейными функциями полезности в условиях неоднородности потребительских предпочтений [171]. Базовая модель, где каждая группа потребителей сталкивается со своей ценой, которая полностью изымает потребительский излишек, исследуется в трёх модификациях: детерминированные предпочтения, частные стохастические предпочтения и общие стохастические предпочтения. Последние модификации различаются заданием уровня качества продукта: в частном случае неизвестны предпочтения потребителей о качестве, а уровень качества считается известным, в общем случае оба параметра определены лишь с помощью функции распределения. Максимизируя прибыль, производитель в каждой постановке выбирает оптимальное количество версий информационного продукта и цены. При определённых сочетаниях условий, рынку предлагается только одна версия. В модели со стохастическими предпочтениями количество версий зависит от величины предельных издержек тиражирования продукта. При этом, однако, существует класс функций полезности, для которых гарантируется наличие нескольких версий продукта во всех рассматриваемых случаях: функции с достижимым по качеству верхним пределом.

А.С. Румянцев исследовал вопрос ценообразования на программное обеспечение со стороны разработчика [243]. Предложены варианты ценообразования при известной функциональной зависимости между требованиями заказчика и стоимостью продукта (фактически, речь об известной смете работ), а также при известных вероятностных распределениях о превышении стоимости разработки и сопровождения продукта. Для борьбы с незаконным копированием продукта («пиратством») предложен вариант двухкомпонентной цены⁷: оплата базовой версии и оплата разработки дополнительных функций, набор которых индивидуален для разных групп потребителей. Результаты носят достаточно абстрактный характер, поскольку 1) функциональные зависимости (можно даже использовать термин «оператор» - зависимости только обозначаются) не раскрываются и не исследуются, а также 2) упрощается производственный процесс (например, через игнорирование повторного использования доступных программных модулей). Индивидуальные цены в работе получаются в результате дифференцированных затрат на разработку продуктов разного функционала и различных ситуаций использования продуктов.

Обзор работ по ценообразованию на рынке информационных продуктов выявляет общий консенсус об использовании дифференцированных цен, и его следует развивать в направлении механизмов взимания оплаты. Этому вопросу посвящена оставшаяся часть данного раздела.

2.5. Алгоритм Гровса-Кларка

Проблема достижения равновесия Линдаля в экономике с общественными или коллективными благами до последнего времени считалась практически неразрешимой по целому ряду причин, о которых частично уже сказано выше. Однако в последнее время стали появляться идеи игровых решений, часть из которых связана с новыми информационными технологиями, позволяющие взглянуть на эту проблему по-новому, особенно в тех случаях, когда речь идет о цифровых продуктах, которым исключительность придается

⁷ Двухкомпонентные цены (или двухкомпонентные тарифы) рассматриваются в Разделе 3 данной диссертационной работы.

искусственно. В космической отрасли свойство редкости искусственно придается сигналам спутникового телевидения и радио. Делается это с помощью технических средств. Тем самым свойство редкости придается тому контенту, который несут сигналы спутника.

2.5.1. Проблемы использования линдалевских цен

Для начала имеет смысл несколько расширить представление о том, что можно называть ценами Линдаля, или линдалевскими ценами, когда речь идет не об экологии и безопасности, а о цифровых продуктах типа фильмов, программного обеспечения и т.д. В широком смысле слова – это индивидуализированные цены на потребление таких благ, причем степень индивидуализации может быть различной. Важно, чтобы в сумме они давали цену, по которой можно приобрести требуемый цифровой продукт и обеспечить им всех участников. В идеале это индивидуальные цены для каждого потребителя, при которых он выбирает уровень потребления на самом высоком уровне (среди всех возможных), а в сумме цены стимулируют производство до того же уровня. Разумеется, на практике это – слишком идеалистическая картина. Даже при широком понимании термина остается много проблем с использованием таких цен в реальной экономике и даже в теории.

Во-первых, есть так называемая проблема арбитража, когда экономические агенты покупают продукт (или право его потребления) по низкой индивидуализированной цене, а потом перепродают его по более высокой другому лицу, у которого индивидуализированная цена высока. В силу этой причины ценовая дискриминация возможна только в специфических случаях, когда между потребителями с разными индивидуализированными ценами есть какой-то барьер, например, языковой.

Во-вторых, если даже удастся решить проблему арбитража, индивидуализированные цены очень трудно рассчитать. В частности, не вполне понятно, как заставить экономических агентов сообщать правдивую информацию о предельной цене, при которой они готовы потреблять общественное благо на данном уровне.

В-третьих, необходимо каким-то образом заставить каждого экономического агента приобретать данное благо по индивидуализированной цене, предназначенной именно для него, не отказываясь от условий, которые были сформулированы, когда определялись цены.

Три названные проблемы далеко не исчерпывают возможный перечень, но и они столь серьезны, что проблема долго считалась неразрешимой. Однако в последнее время стали появляться идеи решения, если не всех, то хотя бы некоторых из перечисленных проблем. Некоторые из этих идей опираются на новые информационные технологии, другие от них не зависят. С них мы и начнем, а именно с некоторых модификаций алгоритма Гровса-Кларка выявления спроса на цифровые продукты.

2.5.2. Выявление спроса на общественное благо (алгоритм Гровса-Кларка)

Игровые механизмы, обеспечивающие автоматическое наказание за оппортунистическое поведение, уже появляются в теоретических работах. Наиболее известный и перспективный из них – алгоритм выявления спроса на общественное благо Гровса-Кларка. Центральным звеном его является так называемый налог (или осевой сбор) Кларка [20]. Этот алгоритм непрерывно совершенствуется и дополняется новыми элементами [40]. Например, в неопубликованной работе⁸ Дитмара Детеринга покупателями таких продуктов являются не отдельные потребители, а клубы или (в более старой терминологии) потребительские кооперативы, приобретающие право на использование такого продукта для своих членов по определенным правилам и автоматически наказывающие за оппортунистическое поведение путем взимания осевого сбора (pivot tax) согласно этим правилам.

Процесс выявления спроса по Кларку (сокращенно именуется DRP – Demand-revealing process) может определить совокупную готовность группы заплатить за общественные блага. В клубной аранжировке право каждого

⁸ How to Charge Lindahl Prices for IP goods by Dietmar Detering, Ph.D. Режим доступа: <https://www.medieninhalte.de/article.htm>

участника использовать цифровое благо по затратам воспроизводства может быть понято как общественное благо этого клуба. Однако предшествующие применения DRP были ограничены дискретными или непрерывными решениями относительно воспроизводства общественных благ или других действий, которые не зависят от или затрагивают других лиц вне голосующей группы. Например, применения DRP позволяли определить, оправдано ли производство определенного общественного блага или нет. Кроме того, они позволяли определить надлежащее количество инвестиций в общественное благо для максимизации благосостояния.

На простом примере можно показать, как работал бы DRP при непосредственном применении для урегулирования покупки дисконтным клубом права использования программного обеспечения. Компьютерная программа *megasoft* доступна в розничной продаже за 50 US\$. Таблица 2.1 показывает готовность к плате потребителей 1-10, представляя всех членов дисконтного клуба. Поставщик оценивает в \$100 полную готовность клуба заплатить за право каждого участника использовать *megasoft* бесплатно и делает клубу соответствующее предложение. Если оно принято, все *могут* использовать *megasoft*, и все *должны* заплатить по \$10.

Затем клуб голосует по предложению, используя DRP. Здесь каждый участник не только указывает, какое решение он одобряет, но также и насколько он одобряет этот конкретный результат. Например, участник 1 готов к плате \$60 и приобрел бы *megasoft* за \$50 в розничной продаже. При коллективной покупке он сэкономил бы \$40 по сравнению с розничной ценой. Таким образом, он голосует этой стоимостью в пользу коллективной покупки. Участники 2-9 ведут себя так же.

Конечно, никто из голосующих членов клуба не захочет преуменьшить свою готовность заплатить в пользу нужного результата. Чем больше готовности заплатить раскрывает одобрение одной стороны, тем более вероятно эта сторона должна победить. Однако в Таблице 2.1 также показано, как DRP гарантирует, что выборщики не собираются преувеличивать свою готовность

заплатить. Это делается путем взимания осевого сбора – ключевой компонент DRP Кларка – с тех выборщиков, которые раскачивают голосование своим утверждением о готовности заплатить в одиночку. В приведенном примере решение против приобретения *megasoft* без учета голосования участника 1 даст \$60 к \$55. Это показывает, что участие члена 1 сдвигает голосование к своему выигрышу, принося всем остальным потерю \$5. Плата этих \$5 осевого сбора оставляет участника 1 с чистым выигрышем \$35 (готовность платить за исход \$40 за вычетом осевого сбора \$5).

Выборщик, который преувеличивает свою готовность заплатить за определенный результат, тем не менее, пытается качать голосование, даже если в результате его качания голосования чистый убыток всем другим выборщикам выше его личной выгоды. При взимании с него осевого сбора, равного чистому убытку всех других выборщиков, он проиграл бы материально, так как осевой сбор в таких случаях превышает его выигрыш от покачивания голосования.

Таблица 2.1. Подсчёт результатов голосования. Источник: переведено по материалам Д.Детеринга.

Член клуба	Готовность платить	Голос «за»	Голос «против»	Осевого сбора
1	\$60	\$40		\$5
2	\$40	\$30		
3	\$35	\$25		
4	\$10	\$0		
5			\$10	
6			\$10	
7			\$10	
8			\$10	
9			\$10	
10			\$10	
Сумма	\$145	\$95	\$60	\$5

Понятно, что затраты принятия решения в DRP, когда все вовлечены в голосование по каждому отдельному предложению, очень высоки. И есть еще две проблемы, которые линдалевское ценообразование должно преодолеть:

(1) Очень высока вероятность того, что предложение будет отклонено. Поставщики попытаются получить максимально высокую цену, но могут только предположить, как высока готовность клуба платить за право использования их цифрового блага. В вышеупомянутом примере поставщик, возможно, попросил только \$50 лишь затем, чтобы быть уверенным, что его предложение принимается, но за \$50 он сделал бы меньше блага, чем он сделал фактически. Также возможно, он попытался бы попросить больше, может быть, \$200, но выборщики отклонят такое предложение, оставляя поставщика вообще без дохода. Такой результат означает трату ресурсов, потому что рассматриваемое цифровое благо уже существует и должно использоваться каждым согласным оплачивать затраты на его воспроизводство.

(2) Чтобы предложение было принято, цена должна быть не выше, чем коллективная готовность заплатить за конкретное право использования. Однако это означает, что доходы поставщиков или нулевые, поскольку их предложение было отклонено, или они положительны, но все же ниже, чем справедливая компенсация за обеспеченную ими ценность.

Из-за этих проблем доход поставщиков был бы ниже того, чего они должны достигнуть при истинном линдалевском ценообразовании, а потребители остались бы отлученными от многих цифровых благ. Решение приходит в виде усовершенствованной процедуры, которая была первоначально предложена для решения других проблем DRP: по любому конкретному предложению голосует только представительная (случайная) выборка участников клуба. Тогда показанная выборкой готовность платить за предложение определяет цену, которую все не участвующие в голосовании потребители должны заплатить за предлагаемое использование права независимо от их фактической личной готовности заплатить. Однако голосующие типовые участники все еще или отлучаются от права использования, или, если они приняли предложение, платят цену, определенную до того, как они голосовали. (См. рисунок 2.5),



Рисунок 2.5: Блок-схема линдалевского ценообразования цифровых благ в «Клубе» (Источник: переведено по схеме Д.Детеринга)

Таким образом, каждый член клуба, не являющийся частью голосующей выборки, получает право использовать предлагаемые цифровые блага по затратам воспроизводства. Участники платят за права цену, которая усредняется по всем купленным предложениям, и в среднем для всех членов клуба цены каждого предложения линдалевские – это способ дать компенсацию владельцам интеллектуальной собственности справедливым и максимизирующим благосостояние способом. Участники исключаются из этого только для тех цифровых благ, по которым они голосовали как часть голосующей выборки.

Этот алгоритм вполне применим, например, при продаже программ спутникового телевидения. В пользу этой идеи говорят три важных обстоятельства. Во-первых, потребителей спутникового вещания можно исключать из потребления, так как они имеют персональные коды, без которых прием

сигнала невозможен. Во-вторых, их можно разбить на страты, в каждой из которых собраны любители примерно одинаковой продукции, готовые платить примерно одинаковую цену. Наконец, в-третьих, их достаточно много, исключение некоторых выборщиков не скажется на выручке. Выборщики набираются по одному или по несколько из каждой страты, а потом играют в игру, описанную выше. Если они, голосуя деньгами, не обеспечивают окупаемость передачи, то исключаются из клуба получателей сигнала, т.е. отключаются. Если они голосуют так, что передача окупается, то производится перерасчет, а для зрителей из каждой страты назначается линдалевская цена.

2.5.3. Оптимизация по Парето с распределением вычислений

Разработка инструментальных средств на основе сочетания grid и математических методов оптимизации – это дальнейшее переосмысление и переоценка применительно к новым условиям имеющегося задела в области алгоритмов для решения задач математического программирования с блочной структурой. К числу этих новых условий, прежде всего, следует отнести сам факт появления grid-технологий и связанных с ними возможностей. Помимо этого, к числу новых условий относится появление разного рода социальных сетей и большого количества цифровых продуктов, что позволяет говорить о новой цифровой экономике. Для такой экономики новый смысл приобретает понятие обмена. В частности, это касается обмена файлами. Современный пользователь интернета вполне мог бы сказать: «Если у нас по яблоку, и мы обменялись, то у нас снова по одному яблоку, а если мы обменялись файлами, их у нас стало по два».

Приведенная выше формулировка совпадает с известной американской поговоркой про обмен яблоками или идеями с точностью до одного слова. А именно, слово «идеями» заменено на слово «файлами». Однако файл может быть художественным фильмом или дорогим программным продуктом, распространяемым на коммерческой основе. То и другое – объекты исключительных прав. Фильм охраняется в целом как сложное произведение, кроме того, охраняются музыка к фильму (как музыкальное произведение), сценарий (как

литературное произведение), а также костюмы и т.д. Программа для ЭВМ также охраняется в рамках авторского права как произведение литературы. Поэтому обмен типа – я тебе файл с фильмом, а ты мне файл с программным продуктом (или другим фильмом) – уже не выглядит так безобидно, как обмен идеями. Нарушаются исключительные права правообладателей и на фильм, и на программу (если в законодательстве страны нет соответствующего изъятия из авторского права). Если же речь идет об обмене через торрент-трекер, где число получателей файла может исчисляться миллионами, то дело принимает качественно иной оборот. Оно становится предметом внимания международных антипиратских организаций, а в отдельных случаях может обсуждаться на уровне президентов США и России. Иначе говоря, ситуация качественно отличается от той, что возникала при обмене идеями вместо яблок.

Эти качественные изменения в представлении об обмене нематериальными продуктами – прямое следствие появления цифровых технологий, торрентов и сетей различного типа. Именно они привели к тому, что свойство идемпотентности сложения информации проявилось в столь очевидной и отнюдь не безобидной форме. Реакцией на это может быть беспрецедентное ужесточение законодательства об авторском праве и об информации, хотя возможен альтернативный подход, а именно, признание бесполезности борьбы с обменом файлами. К этому склоняются многие общественные деятели в разных странах мира [190], особенно когда речь идет об обмене научными результатами⁹.

Так или иначе, необходимо исследовать возможности оптимизации потребления цифровых продуктов и в случае, если они понимаются как общественные блага, и в случае, когда они получают правовую охрану и становятся коллективными благами. В том и другом случае возможность построения

⁹ Keeping science open: The effects of intellectual property policy on the conduct of science. Royal Society report, London, 2003. Текст доступен по: <http://royalsociety.org/Keeping-science-open-the-effects-of-intellectual-property-policy-on-the-conduct-of-science/>

алгоритмов оптимизации с применением параллельных вычислений предполагает формализацию постановки задачи и анализ формальной модели.

В случае, когда речь идет об экономике с общественными благами, задачи, решаемые в каждом блоке (задачи потребителей), связаны не только по ресурсам, но и по ценам потребления общественных благ. Эти цены, как уже говорилось выше, индивидуализированы, при этом сумма индивидуализированных цен потребления должна быть равна цене производства общественного блага – продукта (в смысле его разработки, а не тиражирования). Формально задача потребителя i имеет следующий вид:

$$\begin{aligned}
 u_i(x_i) &\rightarrow \max_{x_i} \\
 \sum_{k \in K_1} q_{ik} x_k + \sum_{k \in K_2} p_k x_{ik} &\leq \beta_i, \\
 x_i &\in X_i.
 \end{aligned}$$

Здесь K_1 – множество всех общественных благ, K_2 – множество всех частных благ (ресурсов). Вектор x_i , как и прежде, состоит из двух частей $\{x_k\}$ и $\{x_{ik}\}$. Первая его часть, соответствующая потреблению общественных благ, для всех i одинакова, вторая – индивидуальна для каждого i . В том, что касается цен, все обстоит ровным счетом наоборот: цены $\{q_{ik}\}$ на каждое общественное благо индивидуальны, а цены $\{p_k\}$ на частные блага для всех одинаковы. Если же нас интересует только та часть экономики, где циркулируют общественные блага, то частные блага и цены на них можно убрать из рассмотрения и останется набор оптимизационных задач, связанных только по ценам

$$\sum_{i \in I} \bar{q}_{ik} = \bar{p}_k \quad \forall k \in K_1.$$

Иначе говоря, сумма индивидуальных цен потребления общественного блага должна быть равна цене его производства. Формально это очень похоже на задачу, в которой отдельные блоки связаны по ресурсам. Но есть два принципиальных отличия. Одно из них содержательное, оно касается смысла

обмена, а другое формальное – наличие общей части $\{x_k\}$ среди переменных, по которым осуществляется оптимизация потребления каждым агентом.

Чтобы избавиться от формального несоответствия задаче обмена ресурсами, достаточно уйти от понятия общественных благ к понятию коллективных благ с соответствующим изменением переменных. Тогда переменные $\{x_k\}$ в целевых функциях всех агентов заменяются на $\{x_{ik}\}$, где k из множества K . И появляются дополнительные ограничения $\{x_{ik}\} \leq \{x_k\}$ для каждого i . Обратный переход еще проще, достаточно все эти неравенства заменить равенствами, чтобы получить набор задач, эквивалентный исходному набору (для экономики с общественными благами). К этому можно добавить, что именно выполнение всех этих неравенств в виде равенств гарантирует оптимальность по Парето соответствующего состояния экономики, т.е. производства и потребления общественных и частных благ всеми экономическими агентами. Иначе говоря, выход неравенств на равенства желателен, к этому надо стремиться.

Гораздо сложнее обстоит дело с противоречиями содержательного характера. При обмене ресурсами агенты обмениваются своими наборами ресурсов, которые им в каком-то смысле принадлежат. В данном же случае место агентов занимают продукты. Обмен ценами на потребление между продуктами – абстракция, не имеющая экономического смысла. Продуктам ничего не может принадлежать, у них нет интересов и прав. Однако выход из этого институционального тупика все же просматривается. Он связан с идеей кооперации.

В качестве экономических агентов можно рассматривать не продукты как таковые, что бессмысленно, а клубы потребителей определенных продуктов или потребительские кооперативы. То и другое имеет прототипы в обычной (не цифровой) экономике. Среди кооперативов наиболее близким аналогом того, о чем тут идет речь, могут служить построенные на идеях А.В. Чаянова [255] кооперативы, в которые объединяются фермеры Голландии. Эти кооперативы специализированы по видам услуг. Например, один кооператив

нужен для совместной организации вспашки земли, другой – для обработки растений ядохимикатами, третий – для мелиорации и т.д. Один фермер может состоять в нескольких таких кооперативах. Точно так же одни и те же потребители цифровых продуктов могут состоять в разных клубах, или потребительских кооперативах нового типа. Эти «клубы», или потребительские кооперативы цифровой экономики, создаются на базе совпадения вкусов и пристрастий, а по организационным принципам и ответственности близки к фермерским кооперативам, но при этом опираются на сетевые технологии и параллельные вычисления, в том числе, возможно, на grid-технологии.

С вычислительной точки зрения процесс обмена индивидуализированными ценами на коллективное благо мало отличается от того, что используется при обмене ресурсами. Исключение может составлять форма субдифференциалов. Это связано с типом ограничений. Более интересно – понять, что и как может оптимизироваться в процессе обмена ценами, а также – за счет чего. Но как раз последнее почти очевидно. Оптимизация возможна, прежде всего, за счет предоставления всем потенциальным потребителям коллективного блага права его потребления на максимальном уровне, т.е. на уровне его производства. Второй возможный источник – оптимизация уровня производства коллективного блага при условии, что оно производится для всех.

2.6. Передача ноу-хау, протоколы и автоматизация

Не углубляясь в обзор экономико-математических моделей распространения информации и протоколов передачи информации, отметим работу, изначально ориентированную на практическое применение. Как отмечает А.Н. Козырев, работа [7] «представляет собой несомненный прорыв в области математического моделирования торговли информацией, поскольку в ней были впервые представлены не модели – математические метафоры, повторяющие на языке математики давно известные практикам истины, а работающие алгоритмы, совместимые со стимулами и позволяющие реально оптимизировать продажу информации» [176]. Представим математическую постановку из процитированной работы и укажем направление её развития.

Предполагается, что продавец обладает частной информацией ω о состоянии мира – сведениями об объекте (в оригинальной работе – характеристиками пользовательской аудитории, которые позволяют настроить контекстную рекламу). В свою очередь покупатель обладает частной информацией θ о своём типе – о том, к какой группе потребителей он относится. При этом покупатель заинтересован в приобретении информации, поскольку она позволяет ему выбрать наилучшее действие $a \in A$, чтобы максимизировать полезность $u(\theta, \omega, a)$. Общим знанием считаются совместное вероятностное распределение $\mu(\theta, \omega)$ случайных величин θ и ω , а также функция полезности $u(\theta, \omega, a)$. В оригинальной работе показано, что наилучшим механизмом, который позволяет продавцу извлечь весь потребительский излишек, является ценовое меню. Строгая формулировка звучит таким образом: если для любого набора (θ, ω) возможно изъять потребительский излишек у ненадёжного покупателя, то этим механизмом является ценовое меню, в котором покупатель каждого типа получает предназначенные ему сведения об объекте по цене, установленной для всех покупателей его типа. И оптимальные цены устанавливаются в результате решения задачи относительно $x_\theta(q)$ и t_θ :

$$\begin{aligned} & \max_{t_\theta} \sum_{\theta} \mu(\theta) t_\theta \\ & \sum_q v_\theta(q) x_\theta(q) - t_\theta \geq v_\theta(p), \forall \theta \\ & \sum_q v_\theta(q) x_\theta(q) - t_\theta \geq \sum_q v_{\theta'}(q) x_{\theta'}(q) - t_{\theta'}, \forall \theta' \neq \theta \\ & \sum_q x_\theta(q) q(\omega) = \mu(\omega), \forall \theta, \omega \\ & x_\theta(q), t_\theta \geq 0, \forall \theta, q. \end{aligned}$$

В представленных выше соотношениях используются следующие обозначения:

- ω – состояние мира (частный сигнал);
- θ – тип покупателя (частный сигнал);
- μ – совместное распределение пары частных сигналов;

- t_θ – плата покупателя типа θ за предназначенный ему контракт;
- q – апостериорная оценка вероятности сигнала ω ;
- p – априорная оценка вероятности сигнала ω , $p(\omega) = \mu(\omega)$;
- $x_\theta(q)$ – вероятность оценки q ;
- $v_\theta(q) = \max_{a \in A} \sum_{\omega} q(\omega) u(\theta, \omega, a)$ - ценность q для покупателя θ .

Согласно представленной схеме, продавец может вычислить значения t_θ и предложить покупателю набор контрактов, каждый из которых состоит в следующем. Покупателю предлагается заплатить одну из цен t_θ . После выбора контракта и оплаты покупатель может наблюдать значение случайной величины, коррелированной с ω . Важно, что вероятностные распределения величин в каждом из контрактов имеют свои распределения – близкие к априорным оценкам покупателей p . То есть покупатель каждого типа наблюдает реализацию конкретного значения из специально подобранного для него (для его типа) распределения в обмен на персональную (для его типа) цену. Причём цена подбирается таким образом, чтобы, во-первых, соответствовать индивидуальной рациональности покупателя (ожидаемая ценность от получения сигнала выше, чем в его отсутствие), во-вторых, обеспечить совместимость со стимулами (ожидаемая ценность от выбора контракта для своего типа выше, чем от контракта чужого типа) и, в-третьих, апостериорные оценки покрывают всё вероятностное пространство.

Сами авторы не стали развивать свои идеи на случай повторяющихся взаимодействий, а также поставили вопрос о дополнительности единиц информации. Обобщение на случай повторяющихся взаимодействий, как показано ниже, требует расширения условия индивидуальной рациональности и учёта изменений в представлениях продавца о типе покупателя. Дополнительность единиц информации остаётся без рассмотрения, однако, предлагается учёт соответствия между сигналом (товара – строки данных) и программой для его обработки.

Развитие механизма

Пусть модельная ситуация имеет следующий вид, описанный в [219]. Продавец обладает набором данных. Доступная ему база может распространяться разными частями: срезами по полям и строкам базы, их комбинациями. Соответственно, срезы различаются информативностью об объекте. Потенциальный покупатель, тип которого θ , ($\theta \in \Theta$), не наблюдаем для продавца, готов востребовать лишь тот срез базы данных, который может технически обработать. Это означает, что доступные ему вычислительные мощности 1) достаточны для хранения среза данных и 2) позволяют получить результат требуемого уровня точности в течение заданного промежутка времени. Для демонстрации отдельных срезов данных из общего набора ω , ($\omega \in \Omega$), предъясняется случайная величина, коррелированная с имеющейся у продавца базой данных. Наблюдения случайной величины проявляются для покупателя в виде сигналов – строк данных. Сигналы о содержимом имеющейся у продавца базы данных предоставляются покупателю в виде меню, в котором за получение сигнала о случайной величине предлагается заплатить определенную цену t_θ^j . Покупатель может n раз обратиться к меню и выбирать за раз по одной случайной величине, информацию о которой он хотел бы получить. Наблюдая сигналы, покупатель может 1) уточнять уже полученные данные и 2) обоснованно составить представление об имеющемся у продавца наборе. При составлении меню продавец назначает цену, максимизируя ожидаемую выручку, при условии, что приобретение сигнала выгодно покупателю по сравнению с его отсутствием (условия индивидуальной рациональности), и покупатель каждого типа θ , ($\theta \in \Theta$), платит именно за интересующие его сигналы, не искажая свой тип (условие совместимости со стимулами). Тогда задачу назначения цен за различные элементы меню на шагах j , $j = \overline{1, n}$, относительно $x_\theta^j(q)$ и t_θ^j можно сформулировать в следующем виде:

$$\max_{t_\theta} \sum_{\theta \in \Theta} \sum_{j=1}^n \mu(\theta) t_\theta^{(j)}$$

$$\sum_q v_\theta^{(1)}(q) x_\theta^{(1)}(q) - t_\theta^{(1)} \geq v_\theta(p), \forall \theta \in \Theta$$

$$\sum_q x_\theta^{(j)}(q)q^{(j)}(\omega) = \mu(\omega), j = 2, \dots, n, \forall \theta \in \Theta, \omega \in \Omega$$

$$\sum_q v_\theta^{(j)}(q)x_\theta^{(j)}(q) - t_\theta^{(j)} \geq \sum_q v_\theta^{(j-1)}(q)x_\theta^{(j-1)}(q) - t_\theta^{(j-1)}, j = 2, \dots, n, \forall \theta \in \Theta$$

$$\sum_q v_\theta^{(j)}(q)x_\theta^{(j)}(q) - t_\theta^{(j)} \geq \sum_q v_{\theta'}^{(j)}(q)x_{\theta'}^{(j)}(q) - t_{\theta'}^{(j)}, j = \overline{1, n}, \forall \theta, \theta' \in \Theta, \theta \neq \theta'$$

$$x_\theta^{(j)}(q), t_\theta^{(j)} \geq 0, \forall \theta, q, j,$$

где $\mu(\theta)$ — распределение типа покупателя θ , $t_\theta^{(j)}$ — цена меню, выбранного на шаге j , для покупателя типа θ , $v_\theta^{(j)}(q)$ — ценность покупателя типа θ от апостериорного распределения $q = q^{(j)}(\omega)$ сведений ω на шаге j в зависимости от дискретно получаемых сигналов, $x_\theta^{(j)}(q)$ — вероятность апостериорного распределения q при наблюдаемом сигнале, $\mu(\omega)$ — распределение имеющейся у продавца информации, $v_\theta(p)$ — ценность покупателя от изначального распределения $p = \mu(\omega)$ информации ω . Технически на практике оптимизация может проводиться путем выбора подаваемых сигналов.

В представленной выше модификации сохраняется возможность для обновления вероятностных распределений типа покупателя на каждом шаге взаимодействия j по результатам совершенных ранее актов приобретения информации. В том числе, может быть использовано правило Байеса [16]. Трудность, однако, при этом состоит в техническом усложнении модели и пересмотре предпосылок исходной работы о механизме с использованием ценовых сигналов. Дело в том, что использование $\mu^{(j)}(\theta)$ в целевой функции предполагает, что продавец, решая оптимизационную задачу, должен заранее иметь представления об изменении вероятностных $\mu(\omega, \theta)$ в зависимости от шага j . Во-первых, это требует введения вероятностей для соответствующих распределений $\mu^{(j)}(\omega, \theta)$, а также априорных и апостериорных (для каждого шага) значений этих вероятностей. Во-вторых, возникает вопрос о том, действительно ли продавцы имеют траекторию изменения своих ожиданий в сделках о продаже информации. Тем более таких, которые поддерживаются (или полностью осуществляются) компьютерными программами.

В действительности протокол продажи данных с возможностью многократного выбора из меню сигналов позволяет покупателю определиться с составом приобретаемых срезов и в последующем при оформлении подписки платить лишь за те сведения, которые подходят именно для его деятельности. Продавец также получает выгоду: во-первых, он не раскрывает сразу всю базу данных, а во-вторых, каждому типу покупателя передаёт подходящие именно для его деятельности сведения, а не общую базу данных, строки которой могут найти применение в различных ситуациях.

Наконец, следует сказать о параметризации оптимизационной задачи через доступные покупателям программы, переходя к колмогоровской сложности строк (данных). Как целевая функция, так и условия (совместимости со стимулами и индивидуальной рациональности) имеют денежную размерность. Поэтому, считая множество Θ списком возможных программных средств, возможности для введения длины строк (колмогоровской сложности и количества информации) – лишь в функции полезности. Слово «лишь» отмечает ограниченность доступных для модификации зависимостей. Однако, по существу, детализация полезности позволяет уточнить её содержание. Денежная полезность $u(\theta, \omega, a)$ – результат решения (действия $a \in A$) при наличии информации, полученной в результате обработки сигнала ω программой θ (с ограничением на ресурсы). При этом, однако, необходимо соотноситься с реальностью в том отношении, что каждой программе $\theta \in \Theta$ соответствует некоторая стоимость её владения (разработки и обслуживания) c_θ . И она также должна включаться в функцию полезности, чтобы усилить этот аспект условия совместимости со стимулами. А именно: приобретение набора данных менее сложных (в смысле определения А.Н. Колмогорова) по соответствующей цене (скорее, более высокой, которая оправдывает их подготовку продавцом и, таким образом, их «упрощение») должно входить в противоречие с более высокой стоимостью владения c_θ той программой, которая настроена на обработку более сложных наборов (скорее, более дорогой в разработке и/или требовательной в ресурсах). Соответственно, при дальнейшем развитии данного

протокола передачи информации следует переходить от $u(\theta, \omega, a)$ к $u(I_\theta(\omega), a, c_o(\theta))$.

Выводы раздела

Информационные продукты вписываются в каждую из предложенных Адамсом групп благ по характеристикам конкурентности в потреблении и возможности исключения потребителей. Для случаев неконкурентности потребления подходят модели ценообразования, которые имеют отношение к общественным или клубным благам. Важный теоретический результат, уже известный экономической науке, состоит в использовании ценовой дискриминации для компенсации издержек поставщика, с одной стороны, и для изъятия потребительского излишка с целью исключения арбитража, с другой. Также доказано, что меню ценовых сигналов позволяет полностью изъять излишек потребителей. Этот результат развит в диссертационном исследовании в двух аспектах.

Во-первых, для различения типологии потребителей предложено использование колмогоровской сложности. Такая параметризация позволяет различать потребителей на рынке данных по функционалу их программного обеспечения и, таким образом, ассоциировать разницу в ценах, уплачиваемую потребителями разных типов, с разницей в стоимостной оценке их программного обеспечения и, соответственно, с ценностью извлекаемой из данных информации.

Во-вторых, расширение модели на случай повторяющихся взаимодействий позволяет охватить ситуации регулярных покупок данных – доступ по подписке.

Цены, о которых шла речь в данном разделе, сформулированы в терминах разовых платежей. Но уже случай повторяющихся покупок показывает, что так бывает не всегда. И, если говорить о лицензионной торговле, цены обычно включают несколько платежей. Вопрос двухкомпонентных цен обсуждается в следующем разделе, где, в частности, показано, что они допускают параметризацию одним числом – например, платой «за доступ» («за подключение»).

3. Стоимость и цены в лицензионных договорах

3.1. Двухкомпонентные цены (теория)

Информационный продукт в современном понимании – это, как правило, комбинация данных в нечитаемом человеком формате и программных сервисах, позволяющих получать из этих данных воспринимаемые человеком образы в виде звуков, картинок, схем, текстов, фильмов и многого другого, включая вкусовые и тактильные ощущения. При этом и данные, и программы – цифровые продукты. Они не обладают сами по себе свойством «редкости», присущим обычным рыночным продуктам. Как показано выше, в терминах экономической теории они обладают свойством неконкурентности в потреблении [100], то есть потребление такого блага одним экономическим агентом не мешает потреблению другими. Таким продуктам свойство «редкости», превращающее их в товары, придается с помощью правовых, организационных и технических средств. Кроме того, продаются обычно не сами цифровые продукты, а лицензии на их использование или услуги с их использованием. Это касается, в частности, электронных книг, на примере которых удобно проиллюстрировать эффекты рынка.

Приобретая электронную книгу, покупатель соглашается с условием лицензионного договора, который он обычно не читает, но обязуется соблюдать. Теоретически о таком рынке можно говорить как о рынке лицензий, хотя для потребителя этот рынок максимально приближен к условиям рынка обычных товаров. С точки зрения экономической теории здесь мы имеем типичный рынок общественных благ, поставляемых в частном порядке [24]. Оптимальные цены для такого рынка – цены Линдаля, то есть индивидуальные для каждого потребителя. Но реально такие цены установить слишком сложно или даже невозможно, хотя алгоритмы, позволяющие получать почти оптимальные решения, существуют (алгоритм Гровса-Кларка и др. см. [25]). На практике книги – типичный продукт с возрастающей отдачей на масштаб производства, причем это связано не с обучением в процессе производства, а с большими начальными затратами, которые распределяются на весь тираж и,

соответственно, входят в цену для всех покупателей. Однако для покупателя бумажной книги, которую он приобретает как вещь, покупка книги – покупка вещи, а для покупателя электронной – это покупка лицензии или услуги (если предоставляется только возможность читать со своего компьютера без скачивания).

Для более сложных продуктов типа информационных систем рынок устроен заметно сложнее, тут часто используется схема двухкомпонентного ценообразования, когда покупатель сначала платит за базовый функционал, то есть фактически за вход, а потом за дополнительные возможности. Внешне это похоже на классический пример двухкомпонентного ценообразования – «парк Микки-Мауса» – Диснейленд в Калифорнии. Там посетители платят одинаковую плату за вход, а потом отдельно за каждый аттракцион, который захотят посетить. При этом вполне очевидно, что, заплатив за вход, посетитель получает набор возможностей, то есть набор реальных опционов, которые может «исполнить» при наличии денег и времени. Тот факт, что при входе посетители платят одинаковую цену, следует подчеркнуть по той причине, что набор приобретаемых возможностей у них далеко не одинаков. Далее они платят за каждый аттракцион по одинаковым ценам.

Пример с парком аттракционов известен по, как считается, первой формализованной модели монополии с двухкомпонентной ценой (тарифом) [84]. В модели присутствует монополист, который поставяет каждую единицу блага с нулевыми предельными издержками. Также на рынке присутствуют однородные потребители с идентичными функциями спроса. Оптимальной по Парето оказывается цена, состоящая из фиксированной платы и переменной цены за использование каждым отдельным продуктом (в примере – аттракционом) в размере предельных издержек. Это решение критикуется за возможное исключение потребителей, для которых фиксированная плата оказывается слишком высокой, и за однородность потребителей [101].

Однако истоки теоретических работ о двухкомпонентном ценообразовании следует искать в научной литературе на несколько десятилетий раньше

работ Ои и Шмалензе. Исследуя вопрос о монопольных поставках продукта с возрастающей отдачей на масштаб, Хотеллинг обосновал продажу по предельным ценам производителя [48]. Однако в этом случае предельные издержки оказываются ниже средних, и сумма платежей от всех потребителей оказывается ниже общих издержек. Следовательно, аргументирует Хотеллинг, выпадающие доходы монополиста должно компенсировать государство через собираемый налог. Принципиальным выводом этих рассуждений является неизбежность вмешательства государства в рынки в возрастающей отдачей (или, что эквивалентно, в рынки с убывающими предельными издержками). Почти десятилетие спустя Коуз обратил внимание на то, что в основе рассуждений лежит предположение о совершенно информированном регуляторе, который может предложить схему налогообложения, не приводящую к перераспределению благосостояния [21]. В качестве реакции на работу Хотеллинга Коуз исследовал возможности рынка для ценообразования без государственного вмешательства. При этом обращалось внимание на несколько принципиальных положений, которые не нашли достаточного отражения в работах Хотеллинга и его последователей. Два из них особенно выделяются на общем фоне. Во-первых, некоторые издержки производителя являются общими для всех потребителей. Во-вторых, фиксированные расходы связаны с ранее понесёнными издержками на факторы, которые в момент продажи товара на рынке не приносят доход. Руководствуясь этими положениями, Коуз приводит аргументы о том, что двухкомпонентное ценообразование с одинаковой для всех потребителей платой в качестве первой компоненты и цены за объём потребляемых товаров или услуг в качестве второй компоненты претендуют на роль рыночного механизма, не требующего государственного вмешательства.

Но дискуссия о предельных издержках остаётся в стороне, когда дело доходит до экономико-математического исследования двухкомпонентных тарифов. Ни Ои, ни Шмалензе не упоминают работу Коуза, но предлагают формальный подход к оценке выгоды и издержек участников рынка в случае монополии, полагающейся на двухкомпонентный тариф. Ои, скорее,

отталкивается от работ Пигу по ценовой дискриминации, проводя аналогию между двухкомпонентным тарифом и ценовой дискриминацией третьей степени. Шмалензе отталкивается от работы Ои, отказываясь от однородности потребителей. Для различных случаев (модельной постановки задачи и соотношения параметров) им показана возможность неотрицательной цены за доступ к благу. Позже к неоднородности спроса также добавлен параметр ограничения на мощности производителя/ поставщика услуг, что особенно важно для случая информационных продуктов, распространяемых, например, по подписке. Это сделано в работе [29].

Анализ показывает, что двухкомпонентные тарифы доминируют в отраслях с высокими постоянными издержками (телекоммуникации, ЖКХ), позволяя фирмам максимизировать прибыль через ценовую дискриминацию, но могут приводить к компромиссам между эффективностью и справедливостью в распределении излишков. В этом контексте – ценовая дискриминация и двухкомпонентное ценообразование – примечательна работа [177], где прослеживается аналогия с составлением меню тарифов для достижения оптимального на рынке состояния. Меню тарифов уже упоминались ранее в Разделе 2.6. о протоколах передачи информации. Исходным положением в работе Козырева является вопрос о поиске оптимума в экономике с невыпуклым производственным множеством или в экономике с возрастающей отдачей на масштаб. Производственная функция и функция полезности в таком случае перестают отделяться гиперплоскостью. Однако возможно их отделение криволинейной поверхностью. Причём такая поверхность может быть задана, как огибающая кусочно-линейных функций (или, эквивалентно, как выпуклая оболочка множества, ограниченного системой кусочно-линейных функций).

Рассматривается семейство двухкомпонентных тарифов, параметризованное двумя функциями одной переменной – $f(t)$ и $g(t)$, где t – индекс тарифного плана в упорядоченном семействе. Функция $f(t)$ задаёт переменную часть тарифа – плату за единицу продукта, а $g(t)$ – фиксированную часть – плату за доступ к благу. Предполагается, что с ростом t функция $f(t)$ убывает,

а $g(t)$ – возрастает. Таким образом, совокупные затраты потребителя на объем x при выбранном тарифе t описываются функцией:

$$h(x, t) = f(t)x + g(t).$$

Пусть бюджетное ограничение потребителя определяется одной величиной: запасом невозпроизводимого ресурса («денег») в размере γ . После приобретения продукта в количестве x у потребителя остаётся количество невозпроизводимый ресурс в количестве:

$$y = \gamma - [f(t)x + g(t)].$$

Таким образом, потребительский набор принимает вид:

$$(x, y) = (x, \gamma - [f(t)x + g(t)]).$$

Рациональный потребитель выбирает оптимальный тариф \bar{t} , минимизирующий затраты на желаемый объём блага x :

$$\bar{t} = \operatorname{argmin}_{t \geq 0} [f(t)x + g(t)].$$

Если функции f и g дифференцируемы, а уравнение $\dot{f}(t)x + \dot{g}(t) = 0$ имеет единственное решение, переменную t можно исключить, получив функцию минимальных затрат от объёма потребления x .

Для дальнейшего теоретического анализа и получения содержательных выводов привлекается функциональная зависимость с нужными свойствами гладкости и минимальным числом параметров. А именно, рассуждения развиваются для двухкомпонентного тарифа следующего вида:

$$h(x, t) = \alpha t^{-\beta} x + t,$$

где t – фиксированный платеж (интерпретируемый как начальные вложения или абонентская плата), а $\alpha t^{-\beta} x$ – переменные затраты. Параметр α отражает базовую стоимость продукта, а β – эластичность переменных затрат от масштаба вложений.

Потребитель выбирает t из условия:

$$t = \operatorname{argmin}_{t \geq 0} [\alpha t^{-\beta} x + t],$$

что приводит к решению:

$$t = (\alpha \beta x)^{\frac{1}{1+\beta}}.$$

Подстановка этого выражения в $h(x, t)$ исключает номер тарифа в семействе t и даёт функцию затрат с понижающимися затратами на масштаб вида:

$$\hat{h}(x) = h\left(x, (x\alpha\beta)^{\frac{1}{1+\beta}}\right) = (\alpha\beta x)^{\frac{1}{1+\beta}} \left(\frac{\beta + 1}{\beta}\right),$$

которая зависит только от объёма x и параметров α, β . Параметр α выполняет, в первую очередь, нормализующую роль, согласуя единицы измерения объёма продукта и денежных затрат. Параметр β имеет содержательную экономическую интерпретацию: он определяет баланс между фиксированной и переменной компонентами затрат. Чем выше β , тем сильнее эффект от масштаба и тем больше доля фиксированных затрат в двухкомпонентном тарифе.

Эти рассуждения можно распространить на случай множества потребителей. Пусть на рынке присутствуют n потребителей, где i -й потребитель приобретает объём x_i , где $i = \overline{1, n}$. Его оптимальный фиксированный платёж составит:

$$t_i = (x_i \alpha \beta)^{1/(1+\beta)},$$

откуда

$$x_i = \alpha^{-1} \beta^{-1} t_i^{(1+\beta)}.$$

Суммарный платёж (тариф) потребителя $i, i = \overline{1, n}$, равен $t_i + t_i/\beta$, а общий доход продавца составит $t + \frac{t}{\beta} = \sum_{i=1}^n \left(t_i + \frac{t_i}{\beta}\right)$.

Предложенная А.Н. Козыревым модель предлагает инструмент предметного анализа двухкомпонентных тарифов. Явный вид функциональных зависимостей позволяет проследить логику рассуждений и поддерживает анализ чувствительности получаемых выводов к значениям параметров. Так, например, варьируя параметр тарифов β , можно охватить случаи вплоть до «безлимитного тарифа» (при $\beta \rightarrow \infty$), когда потребитель оплачивает только доступ к благу (подключение). Соответственно, одна и та же модель применима к рынкам с издержками разной структуры. Также с помощью модели удаётся

проиллюстрировать рынок с возрастающей отдачей. Бюджетное множество становится невыпуклым при $\alpha = 1$, $\beta = 1$ и $\gamma = 1$. И в этом случае семейство двухкомпонентных тарифов обеспечивает нужную отделимость множеств и, как следствие, оптимальность выпуска на рынке с возрастающей отдачей.

Особенной ценностью модели в контексте диссертационного исследования является возможность расширения на случай множества потребителей. Множественность тарифов в контексте рынка с возрастающей отдачей переключается с ценовым меню на информирующие сигналы, о которых шла речь в Подразделе 2.6. И параметризация тарифов единственным числом открывает возможности для сочетания двух моделей (двухкомпонентного тарифа и ценового меню) в общей методологии ценообразования на информационные продукты: за формированием ценового меню в соответствии с процедурой из Подраздела 2.6. фактически может стоять не набор чисел-цен, а элементы из семейства двухкомпонентных тарифов. При этом, однако, следует отметить ограничение, которое касается периодических платежей за доступ к данным по подписке. В модели Подраздела 2.6. с каждой выгрузкой может быть связан свой платёж, и его величина – результат решения оптимизационной задачи. В двухкомпонентном тарифе эти платежи для потребителя i словно «сжимаются» в один: t_i/β . Таким образом, прояснение этого вопроса остаётся открытым для дальнейших исследований. Параметризация может оказаться более сложной и в случае экстерналий. Например, сетевой эффект в ценообразовании на программные продукты играет свою роль, и исследования проблемы назначения цен по мере присоединения новых пользователей активно развиваются [157].

Дальнейшее усложнение модели двухкомпонентного тарифа может касаться дифференцирования параметров затрат для разных типов потребителей: переход от единых для всех потребителей значений параметров α и β к индивидуальным α_i и β_i . Также возможно заложить некоторую степень риска в ожидания (и предпочтения) потребителей относительно использования продукта. А именно, в какой степени реализуется объём потребления,

запланированный на момент приобретения блага по выбранному тарифу. Возможны случаи, когда пользователи переоценивают свои возможности, о чём сказано ниже в Подразделе 3.2.

Резюмируя оценку теоретического результата об описании двухкомпонентного тарифа, следует заключить, что представленная параметризация является полезным инструментом анализа. Её ключевая ценность заключается в демонстрации фундаментальных экономических механизмов, лежащих в основе двухкомпонентного тарифа: экономии от масштаба, возможности дифференцированного ценообразования и возникновения невыпуклости. Эта модель укладывается в цепочку процедуры назначения индивидуальных цен на рынке данных, что открывает возможности для автоматизации сделок – совершения их без длительных переговоров и процедур согласования при посредничестве одного лишь программного робота. Цепочка начинается с выбора двухкомпонентного тарифа – каждый параметризуется одним числом. Конкретные элементы из семейства двухкомпонентных тарифов выбираются для включения в ценовое меню согласно процедуре, описанной в Разделе 2. Однако возможна ситуация, при которой обратившийся к продавцу покупатель в действительности не укладывается в априорную типологию потребителей. Он близок одному из типов, но в точности не соответствует ему. Теоретически в Разделе 2 продавец может выбрать какое угодно конечное количество типов, но на практике это связано с большим объёмом вычислений относительно их предпочтений и вероятностных оценок. Поэтому реальный продавец может ограничиться разумным числом типичных представителей типов и сформировать ценовое меню именно для этого случая. Когда приходит потенциальный покупатель, он может выбрать за основу один из элементов ценового меню и сформировать своё предложение на основе своих действительных предпочтений. Процедура построения такого предпочтения описана ниже в Подразделе 3.3. Но перед этим картину двухкомпонентных тарифов следует дополнить научными исследованиями и практическими выводами, получаемыми на их основе. Хорошо проработанной, в том числе в условиях асимметрии информации,

является практика лицензионной торговли, поэтому она становится отправной точкой дальнейших рассуждений.

3.2. Двухкомпонентные цены (практика)

Являясь объектами интеллектуальной собственности, информационные продукты позволяют обратиться к практике лицензионной торговли – типичной сфере применения двухкомпонентных тарифов. Устоявшаяся схема лицензионного вознаграждения включает паушальный платёж (плату «за вход») и текущие платежи/роялти (плату за каждый случай использования прав интеллектуальной собственности). Такое сочетание выплат связано скорее не с необходимостью оплаты рынком постоянных издержек продавца, на что указывал Коуз [21], и дифференцирования покупателей по их предпочтениям. Рынка для прав¹⁰ на конкретную технологию может и не существовать. Но даже в том случае, когда покупатель (лицензиат) – единственный, структура вознаграждения включает оба компонента. Для этого есть несколько причин.

- единовременный (паушальный) платёж и роялти являются инструментами для балансирования специфических рисков сторон [179],
- роялти позволяют лицензиару манипулировать точкой безразличия лицензиата (участвовать в соглашении или нет) в своих интересах [28],
- паушальный платёж функционирует как реальный опцион, предоставляющий право на использование технологии [111].

Если весь рынок прав на конкретную технологию ограничивается одним покупателем, разговор о меню тарифов теряет смысл. Если покупателей больше, можно различать их по вариантам использования прав и по секторам экономики. Практика лицензионной торговли обращает внимание на таблицы

¹⁰ В этом месте следует различать «права на технологию» и «технологию». Дело в том, что технология может найти применение в производстве массового продукта. В этом смысле «технология» получает рынок – полученный с её помощью продукт получается массовым и распространяется, как обыкновенное частное благо. «Права на технологию» может позволить себе гораздо меньше покупателей. Возможно, только один. Если говорить о микроэлектронике, то в мире не так много компаний, способных производить оборудование для производства микросхем. Чуть больше производителей современных микросхем. И, наконец, микросхемы используются в действительно массовых устройствах. Но технология в области литографии может быть востребована только в начале этой производственной цепочки.

стандартных отраслевых ставок роялти, что существенно облегчает решение вопроса о переменной составляющей двухкомпонентного тарифа, хотя обращение к таблицам и не лишено критики [179].

Поставка информационного продукта не всегда сопровождается двухкомпонентным тарифом. Ранее в Подразделах 2.1 и 2.3 представлены теоретические работы, которые обосновывают равновесные цены без обращения к двухкомпонентным тарифам. В предыдущем подразделе показана сводимость двухкомпонентного тарифа к одной цене при манипулировании параметром β , который влияет на соотношение между разовым платежом и оплатой потребления каждой единицы блага. Вопрос о моментальной оплате и возможном будущем потреблении остро стоит перед потребителями на рынке информационных продуктов. Исследования показывают, что потребители не всегда способны соотнести выгоды и издержки, связанные с двухкомпонентным тарифом, ошибаются в оценке своего будущего потребления. Так, в [61] указывается на то, что потребители более склонны к выбору фиксированных тарифов, даже если это и связано с некоторой переплатой по сравнению с вариантом за объём использования. Для объяснения причин привлекаются психологические эффекты: желание подстраховаться, взять на себя дополнительные обязательства по использованию продукта и т.п. В то же время, фокус потребителей на первоначальном платеже (плате за подключение, за доступ к продукту) приводит к недооценке ими переменных издержек на потребление единиц блага [41]. Выявленные эффекты относятся скорее к области нерационального поведения и психологии. И продавцы на ИКТ-рынке пользуются этим – особенно, когда речь заходит о продуктах, связанных с впечатлениями.

Для примера можно обратиться к информационным объектам с криптографической защитой. Вероятно, наиболее типичный образ в этой сфере – криптовалюта. Однако, по существу, за уникальным идентификатором, записанным в неизменяемый реестр, может стоять что угодно. Поскольку в такие реестры (чаще – блокчейн) записаны транзакции, разблокировать которые могут только владельцы конкретных ключей шифрования, с ними стали

ассоциировать информационные продукты, претендующие на уникальность и эксклюзивность. Это и предметы искусства, и игровые предметы, способные наделять игровых персонажей уникальными способностями. В случае биткоина – классической открытой и не администрируемой цепочки блоков – потребитель делает плату «за вход», приобретая специальное оборудование для майнинга криптовалюты, а потребление каждой единицы блага оплачивается текущими расходами на работу оборудования. Бурный рост этой деятельности связан с возможностью торгов криптовалютой за фиатные деньги. При том, что эта деятельность имеет известные издержки для общества, связанные с расходом природных ресурсов, она во многом полагается на нерациональные ожидания потребителей относительно окупаемости своих вложений.

Нечто похожее можно увидеть и в игровой индустрии, о чём сказано выше. Потребители приобретают компьютерную игру (плата «за вход» может принимать не только форму цены за конкретную игру, но также приобретения специальной игровой приставки и подключения к сетевой инфраструктуре для распространения видеоигр), а затем получают доступ к магазину с игровыми предметами. Рациональное объяснение покупки виртуальных предметов за реальные (фиатные) деньги трудно придумать. Его признаки демонстрируют некоторые потребители, которые приобретают такие предметы в надежде продать их другим игрокам в аукционах.

Встроенные покупки не являются полностью корректным примером двухкомпонентной цены. В представленном выше примере игрок получает возможность пользоваться продуктом после его покупки без дополнительной оплаты за игровое время. Однако часто встроенные покупки предлагаются в многопользовательских онлайн-играх и виртуальных вселенных. В этих случаях пользователь после покупки получает возможность создать игрового персонажа и «бродить» по игре. Но является ли это полноценным продуктом с точки зрения самого игрока? «Слабый» персонаж не позволяет игроку выполнять интересные ему игровые задания. Развитие игрового персонажа часто связано с выполнением однотипных и не всегда захватывающих задач, что и

является одним из факторов, подталкивающих пользователя к покупке игровых предметов, «усиливающих» персонаж. Бывают варианты, когда дополнительные уровни игры доступны за отдельную плату. Поэтому на том уровне обсуждения, который включает психологические эффекты потребителя, опытное благо со встроенными покупками может интерпретироваться в терминах двухкомпонентных тарифов.

Более строгие примеры информационных продуктов с двухкомпонентным тарифом можно найти в секторе B2B. Облачные сервисы для бизнеса полностью укладываются в классическую схему. Очевидным является пример сервиса по распознаванию изображений. Потребитель приобретает лицензию за использование облачного продукта, которая включает дисковое пространство на сервере продавца и применяемую модель для извлечения данных с изображения. Текущие платежи – цена за единицу блага – связаны с каждым отдельным изображением. Доступ к такому сервису может приобретаться на время (месяц, квартал, год), и схема оплаты может заранее фиксировать количество обращений за период, а договор на предоставление доступа – предусматривать единовременный платёж за доступ к сервису и за обработанные изображения. Но этот платёж детализируется по составляющим в приложении к договору. Возможен, однако, и такой вариант, когда время доступа не ограничено. Тогда договор может предусматривать единовременный платёж по факту подписания и периодическую оплату за оказание услуг по распознаванию изображений, привязанную к месяцам или кварталам года. По истечении периодов потребитель получает счёт на оплату в размере, который соответствует количеству обращений к системе. Нюанс бессрочного, однако, состоит в том, что счета на оплату могут включать платежи, направленные на покрытие издержек продавца на поддержание инфраструктуры. Облачные сервисы – достаточно гибкие к масштабированию вычислительных ресурсов под конкретных потребителей. Если часть хранилища и процессоров зарезервировано для конкретного потребителя, это означает невозможность для продавца предложить их на открытом рынке. При этом работа оборудования связана с

текущими издержками. В противном случае, невостребованная вычислительная инфраструктура может быть остановлена и отключена от сети. Соответственно, переменные издержки продавца и упущенная выгода (альтернативная полезность) могут покрываться текущими платежами покупателя.

Если информационный продукт представляет собой исключительно данные, то и в этом случае остаётся место для применения двухкомпонентного ценообразования. Современные программы для анализа данных зачастую предусматривают загрузку через программные интерфейсы (API). Таким образом, пользователи программных инструментов могут с заданной периодичностью обновлять обрабатываемые массивы, анализируя меняющуюся систему (техническую или социально-экономическую). В этом случае потребитель приобретает подписку на использование программного интерфейса, а затем оплачивает каждое обращение к хранилищу продавца.

Существующим моделям двухкомпонентных цен, однако, всё же присущи ограничения, на которые указывал Вэриэн [145].

1. Асимметрия информации между продавцом и потребителем относительно стоимости продукта и его функциональных возможностей.
2. Ненаблюдаемость предпочтений и технологических возможностей участников сделки.
3. Процедурная сложность лицензионных сделок, требующая участия мультидисциплинарных команд (юристов, инженеров, экономистов).

На асимметрии информации и сигналах в экономике следует дополнительно остановиться. Причиной тому служит то, что сигналы позволяют устранить или сгладить асимметрию информации, а в их роли могут выступать не только цены, но и сведения, распространение которых изначально не связывалось с информационным воздействием на рынок.

В научной литературе принято говорить о неполноте информации, асимметрии информации и связанных с ними эффектах. Фундаментальной работой

в этой области является статья Акерлофа о «рынке лимонов»¹¹ [5]. Продавцы предлагают товар разного качества и имеют полную информацию о каждой единице товара. Продавцы наблюдают на рынке лишь ограниченное количество характеристик – например, внешний вид и цену товара. Продавцы, действуя рационально, выдают товар низкого качества за товар более высокого качества. Это становится возможным, поскольку покупатель не может оценить состояние товара до покупки. При этом цена не несёт в себе достаточно информации о качестве товара. Если цена дифференцирована по качеству товара, продавцы товара низкого качества назначают столь же высокую цену, что и аналоги высокого качества. Потребители теряют доверие к рынку и уходят с него. Если на рынке устанавливается средняя цена (сумма цен, взвешенная на количество товара каждого возможного качества) – продавцы высококачественного товара уходят с рынка. В результате, потребителям доступны для выбора только товары низкого качества. Таким образом, реализуется ухудшающий (неблагоприятный) отбор.

Описанная ситуация характерна для опытных благ, информации как товара и объектов интеллектуальной собственности. В последнем случае ситуация усугубляется длительностью процесса передачи технологий. Лицензионный договор включает в себя не только перечень передаваемых прав, но также описывает условия передачи: перечень сотрудников, которые проходят обучение, площадка обучения, критерии успешности процедуры передачи технологии и т.п. От продавца (лицензиара) требуются дополнительные усилия для того, чтобы покупатель (лицензиат) освоил технологию и успешно применял её в своих производственных процессах. Но, если продавец уже получил достаточное, по его мнению, вознаграждение, он теряет заинтересованность в содействии передаче технологии. Этот пример иллюстрирует ещё одно из возможных последствий асимметрии информации – моральный риск, более

¹¹ «Лимонами» в США принято называть подержанные автомобили. При выборе покупатель видит красивые и аккуратные автомобили, не наблюдая их техническое состояние в полной мере. После покупки может выявиться дефект автомобиля. Это сочетание красивого внешнего вида товара и, возможно, «кислого» выражения лица покупателя наводит именно на ассоциацию с лимоном.

рискованное или недобросовестное по отношению к условиям сделки поведение после заключения договора. В случае лицензионной торговли эта проблема смягчается или полностью нейтрализуется оформлением вознаграждения через текущие платежи (роялти). Такой механизм привязывает размер вознаграждения к размеру продаж и, следовательно, к достаточным усилиям лицензиара к передаче технологии. Огромный пласт литературы по передаче технологиям посвящён именно этой теме.

Несколько выровнять положение продавца и покупателя могли бы цены, выступая в роли сигналов о качестве товара. Они, однако, не всегда могут выполнить эту роль, как показывает пример «рынка лимонов». Дополнительные меры в виде гарантийного обслуживания автомобилей, репутации продавца, обязательной проверки технического состояния (за счёт третьего лица) могли бы исправить ситуацию. Но в таком случае именно эти меры берут на себя сигнализирующую функцию.

3.3. Сервис по уточнению двухкомпонентных цен

Рассмотрим рынок информационных продуктов в предельно упрощённом случае. Товар на таком рынке представляет собой комбинацию данных и программных сервисов для их обработки. Для производителя характерна возрастающая отдача: вложения в основные факторы производства (труд или капитал) приводят к улучшениям продукта, которые кратно увеличивают потенциальный рынок и, соответственно, выпуск (количество проданных лицензий). Как показано [177], хотя на рынке с возрастающей отдачей затрудняется поиск равновесия в виде решения задач производителя и потребителя, становится возможным подобрать двухкомпонентный тариф для покупателя конкретного типа. Тариф включает некоторый фиксированный платёж (интерпретируется как плата за подключение) и цену за единицу продукции. Такие тарифы уже применяются в цифровой экономике. Достаточно указать на видеоигры, где пользователь платит за игру, а затем – за отдельные игровые элементы (артефакты), на рынок информационных продуктах в сегментах B2C или B2B. В последнем случае, например, помимо платы за подключение к

сервису пользователю предлагается за дополнительную плату увеличить объём выгружаемых данных или расширить набор доступных аналитических сервисов. Известные модели ценообразования не допускают возможности адаптации к потребителю путём формирования индивидуального предложения, выходящего за рамки существующего ценового меню. Существующие возможности предполагают анализ профиля пользователя и формирование предложения на основе отнесения этого самого пользователя к той или иной группе из принятой классификации. Однако есть возможности коррекции предложения для типичного представителя группы, учитывая особенности конкретного пользователя. Причём в соответствии со строгой формальной процедурой.

Пусть потребитель и продавец оказались в условиях переговоров о цене информационного продукта. На массовом рынке такое сложно представить, но в случае информационных систем с возможностью индивидуального сопровождения сделок и послепродажного обслуживания такое имеет место. Существует начально предложение об условиях оплаты, и стороны ищут возможность улучшить (как минимум, не ухудшить) своё положение. Пусть каждый участник решает задачу линейного программирования – максимизации целевой функции в условиях существующих ограничений. Такой подход восходит к публикации [227] и является широко распространённым в экономико-математическом моделировании [228, 198, 196, 197]. Этот же подход широко распространён в среде профессиональных оценщиков именно для случая неторгуемых на рынке активов – например, работающих предприятий. В профессиональной оценке этот подход получил название функционального [73], но его ограничение состоит в учёте только финансовых ограничений, которые определяют возможности организации по привлечению кредитов возможности по использованию временно свободных денег. Ресурсы, используемые непосредственно в производстве, при наличии денег можно купить, поэтому они не могут быть жесткими.

Далее, пусть потребитель сталкивается с двухкомпонентным тарифом. Этот тариф может быть сформирован продавцом по уже описанным процедурам. Во-первых, двухкомпонентный тариф может идентифицироваться лишь одним параметром t , как показано выше в Подразделе 3.1. Во-вторых, такие тарифы могут быть дифференцированы продавцом, как на то указывают работы [101, 177], и процедура выработки таких тарифов – ценового меню – описана в Разделе 2. Предполагается, что ценовое меню дискретно и не может покрыть континуум возможных типов потребителей. Отчасти отклонение предпочтений потребителя от конкретного тарифа из меню продавца может быть связано с индивидуальными особенностями в оценке значимости первоначального платежа и переменных издержек на потребление блага, о которых говорилось в предыдущем подразделе. Тогда в конкретной ситуации потребитель выбирает «ближайший» к нему тариф и предлагает продавцу свой – более предпочтительный – вариант тарифа. Вопрос состоит в том, как покупателю сформулировать предложение, приемлемое для продавца.

В [181] описан алгоритм решения задачи распределения ресурсов, который можно применить к задаче оптимизации роялти. Для этого задачи агентов переписываются в каноническом виде

$$\begin{cases} (c^s, x^s) \rightarrow \max \\ M^s x^s + E\xi = b(t^s) \end{cases}, s=\{le,lr\}, \text{ где } E - \text{единичная матрица } n \times n, \xi - \text{набор}$$

дополнительных переменных. Двойственная к ней задача имеет вид

$$\begin{cases} (y^s, b(t^s)) \rightarrow \min \\ \left[M^s \quad E \right] y^s \geq \begin{bmatrix} c^s \\ \dot{c} \end{bmatrix} \end{cases}, \text{ где } y^s - \text{двойственная переменная, } \dot{c} - \text{нулевой } n\text{-}$$

мерный вектор, а штрих обозначает операцию транспонирования. Согласно [181], для поиска оптимума по Парето, необходимо:

- (1) для каждого агента вычислить такое решение двойственной задачи \bar{y}^s , что оно является наименее удалённым от луча с направляющим вектором $l = \bar{y}^{le} + \bar{y}^{lr}$ в смысле евклидовой метрики;

- (2) проверить базисное множество – исключаются строки матрицы $[M^s \ E]$, соответствующие неактивным ограничениям двойственной задачи, и включаются строки, соответствующие активным ограничениям двойственной задачи. Если a_j^s - j-тый столбец матрицы $[M^s \ E]$, а c_j^s - j-тый элемент вектора $\begin{bmatrix} c^s \\ \dot{c} \end{bmatrix}$, то a_j^s входит в базис, когда $(a_j^s, y^s) = c_j^s$;
- (3) для агента $s=\{le,lr\}$ вычислить матрицу B^s , обратную к базисной матрице, составленной из базисных столбцов a_j^s ;
- (4) вычислить вектор $e^s = \bar{y}^s - \frac{(\bar{y}^s, l)}{(l, l)} l$ для каждого $s=\{le,lr\}$;
- (5) вычислить все числа $\lambda^s \mid B^s b^s (t^s + \lambda^s e^s) \geq 0, s=\{le,lr\}$;
- (6) вычислить новые доли $\bar{t}^s = t^s + \min(\lambda^s) e^s, s=\{le,lr\}$;
- (7) при новых долях повторить шаги (1)-(6).

Для того, чтобы показать сходимость данного алгоритма, необходимо доказать, что набор $(\bar{y}^s / \sum_{i=1}^n \bar{y}_i^s, \bar{t}^s)$ является равновесием в смысле Розенмюллера [235]: точка (p, t) , где $\sum_{i=1}^n p_i = 1$, называется равновесием, если $(p, t) \leq (p, a)$ и $z(t) = \max_{(p, t) \leq (p, a)} z(t)$, где a – начальное распределение долей.

Существование равновесия следует из монотонности предпочтений, описываемых функцией $z(t^s) = (c^s, x_*^s)$, где x_*^s - решение задачи при данном значении параметра t^s .

Для иллюстрации допускается существование в экономике только двух продуктов: информационного и «денег», которые характеризуют возможности потребления: со стороны производителя – потребления ресурсов, со стороны потребителя – потребления других благ. Информационный продукт обозначен x , деньги – t . Можно оперировать функциями общего вида – производственной $V(t, p) = x$ и функцией полезности $U(x, t) = f(t) + \beta g(p, x)$, где p – цена

за единицу продукта, β – коэффициент, который отражает индивидуальную полезность от потребления продукта $g(p, x)$ по сравнению с полезностью от подключения $f(t)$. Однако для наглядности используем функции постоянной эластичности:

$$\begin{aligned} V(t, p) &= x = A(t + p)^\gamma, \\ U(x, t) &= (T - t) + \beta px, \end{aligned}$$

где A, γ, β – коэффициенты, T – допустимый для потребления объём блага «деньги» (можно также интерпретировать как «начальные запасы денег»). Пусть потребителю предлагается информационный продукт в количестве x_0 по двухкомпонентному тарифу $\{t_0, p_0\}$. Будем считать это предложение исходным. Для поиска лучшего предложения построим градиенты функций в пространстве (t, p) . Для этого исключим x из $U(x, t)$, используя выражение $x = A(t + p)^\gamma$, считая, что предложение $\{x_0, t_0, p_0\}$ принадлежит производственной функции и функции полезности одновременно. Тогда

$$\begin{aligned} \nabla V(t_0, p_0) &= (A\gamma(t_0 + p_0)^{\gamma-1}; A\gamma(t_0 + p_0)^{\gamma-1}), \\ \nabla U(t_0, p_0) &= (-1 + \gamma\beta A p_0 (t_0 + p_0)^{\gamma-1}; \beta A (t_0 + p_0)^\gamma [1 + \gamma p_0 / (t_0 + p_0)]). \end{aligned}$$

Будем иметь в виду, что предложение $\{x_0, t_0, p_0\}$ соответствует некоторому набору (меню) тарифов, который представляет собой ограниченное по предложению t и по p множество. И тогда существуют предельные значения, то есть

$$\begin{aligned} \exists t_{min}, t_{max} | t_{min} \leq t \leq t_{max}, \forall (t, p) \in S, \\ \exists p_{min}, p_{max} | p_{min} \leq p \leq p_{max}, \forall (t, p) \in S, \end{aligned}$$

где S – множество тарифов. Ограниченность значений позволяет воспользоваться процедурой повышения уровней выпуска (для производителя) и полезности (для потребителя), следуя формальному алгоритму, который известен из задачи оптимизации ресурсов [181] и позднее применён к сделкам именно путём перехода к безразмерным величинам [208]. Вводя замену, перейдём к безразмерным переменным $\tau \in [0, 1]$ и $\rho \in [0, 1]$:

$$\tau = \frac{t - t_{min}}{t_{max} - t_{min}},$$

$$\rho = \frac{p - p_{min}}{p_{max} - p_{min}}.$$

Далее, переход к новому тарифу описывается в пространстве (τ, ρ) . На первом этапе вычисляется единичный вектор равноудалённого от $\nabla V(\tau_0, \rho_0)$ и $\nabla U(\tau_0, \rho_0)$ луча:

$$e = \frac{1}{|\nabla V(\tau_0, \rho_0) + \nabla U(\tau_0, \rho_0)|} (\nabla V(\tau_0, \rho_0) + \nabla U(\tau_0, \rho_0)).$$

На втором этапе выбираются проекции векторов градиента $\nabla V(\tau_0, \rho_0)$ и $\nabla U(\tau_0, \rho_0)$, ортогональные лучу, заданному вектором e :

$$\nabla V_+(\tau_0, \rho_0) = \nabla V(\tau_0, \rho_0) - (\nabla V(\tau_0, \rho_0) \circ e)e,$$

$$\nabla U_+(\tau_0, \rho_0) = \nabla U(\tau_0, \rho_0) - (\nabla U(\tau_0, \rho_0) \circ e)e,$$

где запись вида $(a \circ b)$ означает скалярное произведение векторов a и b . Наконец, новый тариф вычисляется по формуле $(\tau_0, \rho_0) + \nabla U_+(\tau_0, \rho_0)$ с точностью до масштаба. Строго говоря, вектор $\nabla U_+(\tau_0, \rho_0)$ задаёт направление сдвига к новому тарифу, и новое значение следует искать, учитывая ограничения сторон. Таковыми могут являться, например: бюджетное множество потребителя, предельное по входной плате (или удельной цене) значение тарифа из меню, предложенного производителем.

Выводы раздела

Использование двухкомпонентных цен на рынке информационных продуктов расширяет дискуссию об индивидуальных ценах, начатую во втором разделе, в направлении адаптации условий сделки к конкретному потребителю. Поставщик вправе разбить взимаемую цену на два платежа: фиксированный и за единицу блага. Имея ожидания об индивидуальных ценах и представления о предпочтениях групп потребителей, поставщик может предложить наилучшее с его точки зрения меню двухкомпонентных тарифов. Но эти ожидания могут не совпасть – и скорее всего не совпадут – с предпочтениями реальных потребителей. В этом случае для конкретной сделки потребитель может выбрать наиболее близкий ему вариант из ценового меню и, следуя строгой формальной процедуре, сформулировать более предпочтительный

вариант тарифа, который гарантированно окажется для поставщика не хуже других вариантов изначального ценового меню.

Что особенно важно, так это возможность реализации такой процедуры в виде программы – ценового алгоритма. Использование таких алгоритмов не запрещено антимонопольным законодательством и, таким образом, может стать рабочим инструментом для сопровождения и поддержки сделок на рынке данных. Но алгоритмизация возможна не только в случае двух участников или двух дополняющих друг друга товаров – данных и программы. Возможности существуют и в стоимостной оценке многокомпонентных товаров, коими являются современные информационные продукты – наборы программных библиотек и неоднородных, а зачастую и распределённых, данных. И этот аспект подробно рассматривается в следующем разделе.

4. Распределение выигрыша с применением теории игр

Главная задача раздела – изложение метода стоимостной оценки неторгуемого на рынке актива, программы в использовании. Существенным предположением метода является возможность оценить совокупный эффект, который получается от совмещения набора данных и программы (способа обработки набора данных). Свою роль играет также известная цена данных – приобретаемого на рынке товара. Метод разделения стоимости портфеля прав ИС и НМА по его составляющим изложен в [179] и позднее применён к портфелю большой размерности с группировкой элементов портфеля в кластеры, прежде всего, по принципам правовой защиты и участия в общих проектах [178]. Этот метод опирается на результаты раздела теории игр о кооперативном поведении. Поэтому последовательное приближение к авторскому вкладу начинается изложением некоторых положений из теории игр, важных в контексте стоимостной оценки прав (Подраздел 4.1.). Затем представлен метод разделения стоимости портфеля между его компонентами, предложенный А.Н. Козыревым и В.Л. Макаровым (Подраздел 4.2.). В том же Подразделе 4.2. исходный метод интерпретирован в терминах колмогоровской сложности и извлечения дополнительной информации для повышения ценности набора данных. В этом состоит авторский вклад: колмогоровская сложность естественным образом позволяет интерпретировать стоимость программы в терминах усреднения её вклада в извлечение информации из набора данных. Наконец, в Подразделе 4.3. предложен авторский пример стоимостной оценки программы для случая вычисления компенсации за неправомерное использование. Основой для этого примера послужила совместная работа с А.В. Костиным об оценке товарных знаков [186]. Авторский вклад состоит в формализации моделей и переноса конфликта из области товарных знаков в сферу программных продуктов. Подраздел 4.3. также опирается на методику, предложенную А.Н. Козыревым [178], однако, автор предлагает её развитие в части способа объединения элементов портфеля в кластеры. Для снижения размерности задачи предлагается выделить группу «уникальных», но чрезвычайно ценных программ.

Критерием для этого служит то, в какой степени программы позволяют уменьшить колмогоровскую сложность набора данных (из портфеля).

4.1. Модель сделки по Нэшу при заключении лицензионного договора

Модель сделки по Нэшу – один из вариантов двусторонних игр, в которых допускается кооперирование между игроками. Это значит, что могут заключаться соглашения, допускается совместный выбор смешанных стратегий, а полезность может передаваться от одного игрока к другому (хотя и не всегда линейно). Тем самым изначально подразумевается, что предпочтения игроков описываются с использованием функций полезности. При этом полезность обязательно измеряется в деньгах, скорее, наоборот. Типичный пример применения сделки по Нэшу как раз касается случая, когда стороны, находящиеся в заведомо неравном положении, делят некоторую сумму денег. В этом случае большую долю получает та сторона, для которой предельная полезность денег ниже. В конкретном примере, приведенном в [226], выигрыши сторон сделки обратно пропорциональны предельной полезности денег для них. Этот или похожий эффект можно наблюдать в лицензионных сделках.

4.1.1. Особенности лицензионных сделок

Применительно к лицензионным договорам сказанное выше относительно общей схемы может означать, что все платежи, получаемые лицензиатом в виде единовременного (паушального) платежа и в виде роялти, приводятся к одному моменту времени. Кроме того, полезность сделки для него зависит не только от получаемых платежей. Предоставляя лицензиату лицензию, он фактически отдает ему часть рынка. Лицензиат, напротив, выплачивает паушальный платеж и роялти, но приобретает часть рынка, ранее занятого лицензиатом, и не только. Он может выйти на какие-то другие рынки, что обычно при предоставлении лицензий зарубежным компаниям. Наконец, деньги могут иметь разную предельную полезность для сторон по договору.

Чтобы применить модель сделки по Нэшу к лицензионному договору, надо рассмотреть для начала несколько упрощенную ситуацию, когда есть

минимальная цена, на которую согласен лицензиар, есть максимальная цена, которую готов заплатить лицензиат. Разность между минимальной ценой лицензиара и максимальной ценой лицензиата, если она положительна, обозначается в англоязычной литературе термином *equity of the deal*. Однозначного перевода этого термина на русский язык пока нет, так как перевод «доля участия в сделке» (вариант онлайн-переводчика) или вариации этого перевода с прилагательным «справедливое» – это просто не о том. Реально речь идет о том, что могло бы переводиться как «чистая выгода сделки» по аналогии с переводом *equity* как «чистые активы», когда речь идет об активах компании. Именно эту величину делят стороны сделки при условии, что сама сделка оптимальна по Парето, то есть нельзя улучшить положение одного участника, не ухудшая положение другого, меняя те или иные параметры.

4.1.2. Формальная схема сделки по Нэшу

Рассмотрим взаимодействие двух игроков и будем считать его нетривиальным в том смысле, что множество возможных исходов игры не является пустым. При описании предпочтений игроков функциями полезностей каждому исходу можно поставить в соответствие пару чисел – значения полезностей игроков. Таким образом, функции полезности позволяют отобразить множество исходов в подмножество евклидовой плоскости R^2 – замкнутое, ограниченное сверху и выпуклое. Последнее следует из возможности существования смешанных стратегий. Такое подмножество S плоскости R^2 называют *допустимым множеством*. Имеется ввиду, что в ходе взаимодействия (по мере развёртывания игры) участники могут прийти к таким исходам, которые принесут им полезности u и v , что $(u, v) \in S, \forall u, v$. В антагонистических играх выигрыш одного участника связан с выигрышем другого, поэтому в таких случаях рост u связан со снижением v и наоборот. При решении задачи о сделке требуется найти такую точку подмножества S , которая одновременно удовлетворяла бы обоим игрокам.

В общем случае стратегии поведения игроков нельзя предугадать заранее, хотя возможно установить принципы формирования таких стратегий.

Например, каждый из игроков стремится выбрать для себя наилучший результат, выбирает вариант поведения из доступных ему, играет до тех пор, пока не может улучшить своё положение и т.п. Эти принципы перекликаются с аксиомами Нэша для построения функции, разрешающей задачу о сделке. Но ещё до перехода к аксиоматике можно указать на существование пороговых значений полезностей. Это значения, которые могут получить игроки в результате односторонних действий независимо от поведения другого участника – максиминные величины u^* и v^* . Для биматричной игры с матрицами (A, B) , где действия x одного игрока определяют его полезность u , а действия y другого игрока определяют полезность v выражения принимают вид:

$$u^* = \max_x \min_y xAy^T \quad (4.1)$$

$$v^* = \max_y \min_x xBy^T, \quad (4.2)$$

(где x и y выбираются из множеств всех смешанных стратегий).

Если считать известными подмножество S и максиминные значения (u^*, v^*) , возникает вопрос о поиске правила, которое приписывает такой тройке (S, u^*, v^*) «решение задачи о сделках»:

$$\varphi(S, u^*, v^*) = (\bar{u}, \bar{v}).$$

Условия для такой функциональной зависимости задают формулировки, известные как аксиомы Дж. Нэша. Они являются частью формальной стороны задачи о сделках и не учитывают такие особенности, как рыночная власть, умение вести переговоры и другие индивидуальные особенности игроков в каждом конкретном случае. Таких аксиом шесть [226]:

N1 (индивидуальная рациональность), $(\bar{u}, \bar{v}) \geq (u^*, v^*)$.

N2 (допустимость), $(\bar{u}, \bar{v}) \in S$.

N3 (оптимальность по Парето). Если $(u, v) \in S$ и $(u, v) \geq (\bar{u}, \bar{v})$, то $(u, v) = (\bar{u}, \bar{v})$.

N4 (независимость от посторонних альтернатив). Если $(\bar{u}, \bar{v}) \in T \subset S$, и $(\bar{u}, \bar{v}) = \varphi(S, u^*, v^*)$, то $(\bar{u}, \bar{v}) = \varphi(T, u^*, v^*)$.

N5 (независимость от линейного преобразования). Пусть T получается из S с помощью линейного преобразования

$$u' = \alpha_1 u + \beta_1, \quad v' = \alpha_2 v + \beta_2.$$

Тогда, если $\varphi(S, u^*, v^*) = (\bar{u}, \bar{v})$, то

$$\varphi(T, \alpha_1 u^* + \beta_1, \alpha_2 v^* + \beta_2) = (\alpha_1 \bar{u} + \beta_1, \alpha_2 \bar{v} + \beta_2).$$

N6 (симметрия). Предположим, что S таково, что

$$(u, v) \in S \leftrightarrow (v, u) \in S.$$

Предположим также, что $u^* = v^*$ и $\varphi(S, u^*, v^*) = (\bar{u}, \bar{v})$. Тогда $\bar{u} = \bar{v}$.

Аксиомы N1, N2 и N3 очевидны и интуитивно понятны. Аксиома N4 утверждает, что если точка (\bar{u}, \bar{v}) есть решение задачи о сделках, то при расширении допустимого множества ни одна из прежних точек не может стать решением новой задачи – либо останется (\bar{u}, \bar{v}) , либо решением станет одна из новых точек. Аксиома N5 отражает естественное безразличие к выбору вида функции полезности. Наконец, аксиома N6 отражает изменения отношения к сделке (приемлемость или нет) при изменении состава участников: «она может оказаться неприемлемой в случае сделок между неодинаковыми игроками, например, между лицом и коллективом» [226].

Считается, что такая аксиоматика не является единственной. Но введение новых аксиом требует пересмотра всего списка. Так, возможна формулировка следующей аксиомы:

N7 (монотонность). Если $T \subset S$, то

$$\varphi(T, x^*, x^*) \leq \varphi(S, x^*, y^*).$$

Однако эта аксиома несовместна с N2 и N3 [226]. Но и после согласованием аксиом потребуются поделаться дальнейшие рассуждения для построения цельной теории, вытекающей из принятых положений. Для представленного выше списка аксиом N1—N6, например, существует следующая теорема, которую, несомненно, потребуются переформулировать в другой аксиоматике.

Теорема. *Существует единственная функция φ , определенная для всех задач о сделках (S, x^*, y^*) и удовлетворяющая аксиомам N1—N6.*

Поэтому наиболее разумным было бы положиться на аксиоматику Нэша и двигаться дальше в построении и исследовании сделок.

В случае, когда граница подмножества S является гладкой в точке (\bar{u}, \bar{v}) (то есть имеет касательную), прямая, на которой линейная функция

$$h(u, v) = (\bar{v} - v^*)u + (\bar{u} - u^*)v$$

принимает постоянное значение, совпадает с этой касательной.

Наклон касательной к S в любой точке отражает соотношение, в котором полезность может быть перераспределена между игроками. Согласно схеме Нэша, дополнительная полезность делится между участниками именно в этой пропорции. Если полезность не является линейно трансферабельной, может существовать лишь одна точка, где перераспределение происходит в заданном соотношении, что проиллюстрировано на Рисунке 4.1.

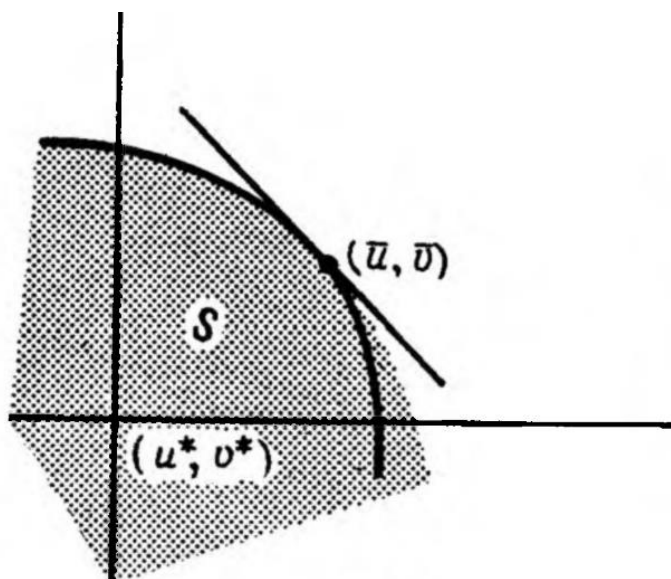


Рисунок. 4.1. Пропорции перераспределения полезности меняются от точки к точке. Источник: [226, с 155]

Ситуация существенно упрощается, когда полезность линейно трансферабельна. В этом случае (возможно, после соответствующего изменения шкал полезности) можно считать, что полезность передается в соотношении один к одному. Это означает, что игрок может передать одну единицу полезности другому игроку, теряя при этом ровно одну единицу. Таким образом,

подмножество S включает все точки, лежащие на прямой $u + v = k$ или ниже неё (где k – максимальная совокупная полезность, которую игроки могут получить вместе), и одновременно выше и правее максиминной точки (u^*, v^*) (см. Рисунок 4.2).

Решение Нэша для этого случая — точка (\bar{u}, \bar{v}) , где

$$\bar{u} = (u^* - v^* + k)/2 \text{ и } \bar{v} = (v^* - u^* + k)/2.$$

При этом $\bar{u} + \bar{v} = k$, а $\bar{u} - \bar{v} = u^* - v^*$. Следовательно, относительные позиции игроков (разность их полезностей) остаются неизменными, а их абсолютные полезности возрастают максимально возможным образом, поскольку избыток полезности делится между ними поровну.

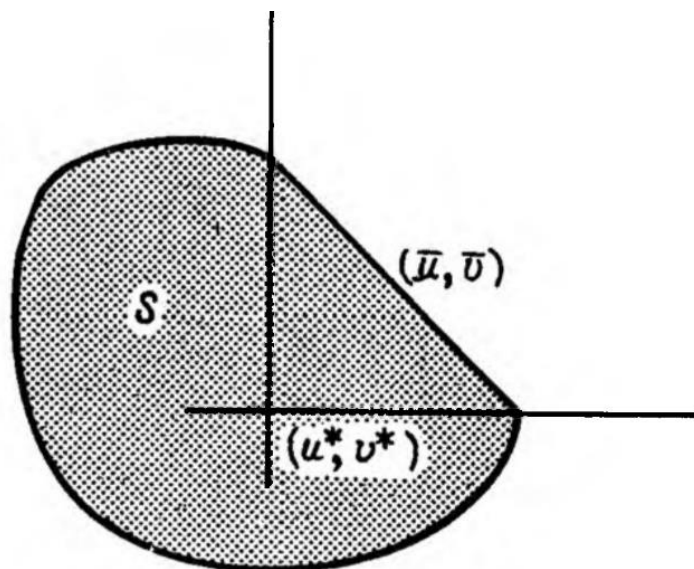


Рисунок. 4.2. Пропорции перераспределения полезности сохраняются на некотором промежутке. Источник: [226, с 155]

4.1.3. Интерпретация результата сделки

В примере, используемом в учебниках [226], двум игрокам предлагают 100 долларов, если они смогут решить, как поделить эти деньги между собой. Первый игрок очень богат, а капитал второго ограничен 100 долларами. Полезность суммы денег, которые надо поделить, пропорциональна её логарифму. Каким является решение задачи о дележе в данном случае? Дальнейшие рассуждения имеют следующий вид:

Так как первый игрок очень богат, можно предположить, что полезность x долларов, где $x \leq 100$, пропорциональна x . Так как второй игрок имеет только 100 долларов, полезность, которую он получает от x долларов, равна

$$\log(100 + x) - \log 100 = \log \frac{100 + x}{100}$$

(Мы, конечно, полагаем $u^* = v^* = 0$.) Таким образом, множество S будет выпуклой оболочкой точки $(0,0)$ и дуги кривой, задаваемой уравнением

$$v = \log \frac{200 - u}{100}.$$

Так как эта функция выпукла, S просто является областью, ограниченной этой кривой (см. Рисунок 4.3); конечно, интерес представляет только пересечение этой области с положительным квадрантом.

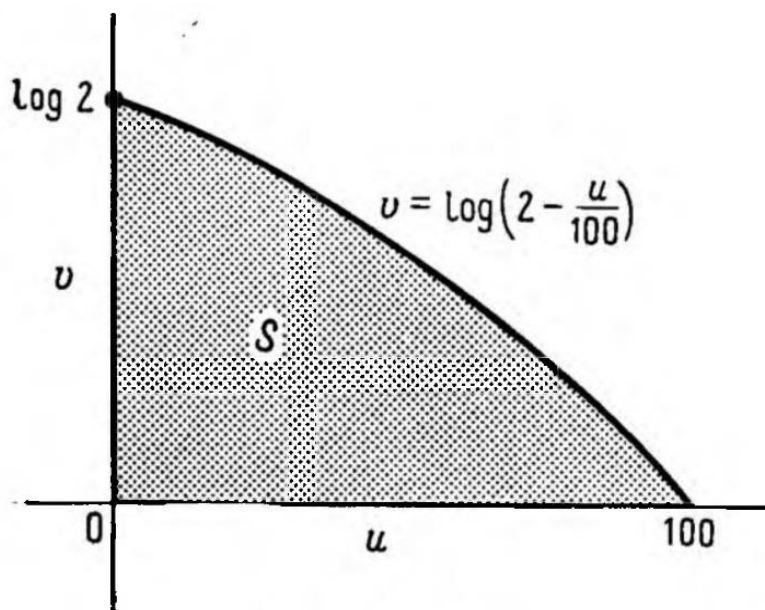


Рисунок. 4.3. Допустимое множество для численного примера. Источник: [226, с 156]

Теперь ищем точку в S , максимизирующую uv , т. е. значение u , при котором достигается максимум функции

$$g = u \log \frac{200 - u}{100}.$$

После дифференцирования и приравнивания нулю производной получаем уравнение

$$\frac{u}{200 - u} = \log \frac{200 - u}{100}.$$

Решая его, получаем приближенно $u = 54,4$; иначе говоря, игрок I должен получить 54,40 доллара, а игрок II должен получить 45,60 доллара, то есть несколько меньше половины.

В учебнике [226] этот результат преподносится как кажущийся странным: так как из него следует, что «богатый игрок должен получить больше, чем бедный, о котором можно утверждать, что он больше нуждается в деньгах». Это утверждение неявно предполагает возможность сравнения полезностей принципиально разных игроков, что противоречит принятой схеме. В действительности, в исходной постановке полезность денег у игрока II (капитал в 100 долларов) убывает быстрее, чем у игрока I (очень богатого). Следовательно, в общей игре существенно менее богатый игрок стремится получить хоть что-то и готов пойти на уступки, в то время как первый игрок менее чувствителен к колебаниям в несколько десятков долларов.

При всей своей схематичности, описываемый пример достаточно хорошо описывает ситуацию, когда одна сторона сделки получает преимущество по той причине, что она менее чувствительна к изменению условий в ту или иную сторону. Если говорить конкретно о лицензионных договорах, то напрашивается пример, когда одна сторона – крупная фирма, другая – изобретатель-одиночка, программист, придумавший интересную всем компьютерную игру (например, Тетрис), и так далее. Эффект сделки – *equity of the deal* – может измеряться в миллиардах долларов, но одиночке будет предложено несколько десятков тысяч долларов за всё, то есть за уступку исключительных прав и титула (если речь идет о патенте). И дело здесь не только в том, что предельная полезность денег для него больше, чем для крупной фирмы.

Дополнительных оснований для того, чтобы не «переплачивать» одиночке, может быть несколько. Справедливость дележа здесь тоже играет свою роль, хотя дело совсем не в предельной полезности. Во-первых, крупная

фирма на момент сделки, как правило, уже имеет сильные позиции на рынке конечного продукта, она вкладывает в сделку не деньги или не только деньги, но и доступ к большому списку или сетевому сообществу покупателей своей продукции или услуг. Иногда такой актив называется клиентелой. Но дело не в названии, а в том, что такие активы играют существенную роль в сделках, связанных с передачей исключительных прав, включая позицию при распределении общего выигрыша от совершаемой сделки. Во-вторых, крупная фирма должна принимать во внимание тот факт, что такая сделка – одна из ряда аналогичных сделок, поэтому распределение выигрыша не просто создает прецедент, а формирует норму для последующих сделок. Вот тут вступает в действие та самая предельная полезность, хотя к справедливости дележа это имеет очень слабое отношение. Для одиночки, получившего результат (изобретение, программу и т.д.) не в рамках служебных обязанностей или служебного задания, деньги, как правило, очень нужны сейчас (на момент совершения сделки). А потому у него нет возможности торговаться жестко и настаивать на справедливом дележе *equity of the deal*. О служебных изобретениях, программах для ЭВМ и т.д. речь вообще не идет.

Существенно иначе обстоит дело, когда сделка о передаче исключительных прав на результаты интеллектуальной деятельности заключается между большой и сравнительно маленькой фирмой, имеющей, тем не менее, хорошие перспективы на завоевание существенной части рынка.

Подводя итог сказанному выше, стоит отметить, что модель сделки по Нэшу – это распределение, обратно пропорциональное предельной полезности, причем не только в приведенном выше примере с дележом 100 долларов между богачом и бедняком. Сфера применения решений такого типа много шире, как и возможных обоснований такого решения. Основная трудность для применения такой схемы – численное определение соотношения предельной полезности для разных сторон сделки, так как для каждого из них полезность измеряется в своих единицах, если вообще как-то измеряется.

Но и в случае, если соотношение предельных полезностей удастся посчитать по какой-то схеме, принятой обеими сторонами, полученная величина может служить лишь некоторым ориентиром. Реальные параметры сделки определяются в ходе переговоров, где не последнюю роль играет умение такие переговоры вести и вовремя приводить подходящие аргументы.

4.1.4. Методы дележа при передаче технологий

В ходе переговоров стороны приходят к некоторому распределению общей выгоды, опираясь на следующие наиболее распространённые процедуры [211]:

- метод освобождения от роялти,
- правило 25%,
- экспертные системы,
- по аналогии,
- нормативные требования.

Метод освобождения от роялти предполагает известной ставку роялти и, строго говоря, больше подходит для оценки интеллектуального вклада, а не разделения выгоды. Последнее отражается ставкой роялти. «Правило 25%» реализует принцип разделения выгоды от сотрудничества в пропорции 25 % – лицензиару и 75 % – лицензиату. Оно вытекает из эмпирического наблюдения, в соответствии с которым, с одной стороны, такое распределение возникало в реальных сделках и, с другой стороны, примерно соответствует точке безразличия для лицензиата между вариантами покупки технологии и проведением исследований и разработок своими силами.

В основе экспертных систем лежат согласованные суждения о степени важности конкретной технологии для успеха конкретного проекта с учётом доступных альтернатив. Такие системы могут быть реализованы, например, с привлечением метода анализа иерархий, или квалиметрии. Также на экспертные суждения полагается распределение по аналогии, при котором эталонная сделка с известными характеристиками служит базой для расчёта вознаграждения.

Наконец, нормативные требования – это установленные какими-то документами (законодательными или внутрифирменными) правила расчёта вознаграждения. Правила подобного рода характерны для бухгалтерских целей и для расчёта с авторами служебных изобретений. Например, утратившее ныне силу Постановление Правительства РФ от 04.06.2014 № 512 увязывало размер вознаграждения со средней заработной платой работника.

Теперь, после описания эмпирики в разделении платежей, можно перейти непосредственно к исследованию применения теоретико-игрового подхода к расчёту величины и распределению вознаграждения. Начать следует с публикации [8], которая, хотя и вышла в свет более десятилетия назад, обозревает принципиальные сферы использования теоретико-игрового инструментария в анализе инновационной активности. Авторы выделяют три уровня для формализации игровых ситуаций: внутриорганизационный, межфирменный и метаорганизационный. В рамках поставленной цели исследования интерес представляют первые два, где возникает задача дележа между фирмами и между работниками фирмы. В частности, из выделенных авторами задач следует обратить внимание на стимулирование инновационной активности, распространение знаний внутри фирмы, выбор оптимальных параметров лицензионного договора, взаимодействие правообладателя и инновационных фирм, кооперацию партнёров в научно-технологическом проекте. Упомянутая статья выделяет больше ситуаций, которые могут быть проанализированы с помощью теории игр, но не во всех из них возникает задача расчёта долей участников. Например, этого не происходит при выработке национальной инновационной политики.

Более того, даже релевантные для контекста ситуации ставят гораздо больше задач, чем расчёт вознаграждения. Например, кооперация партнёров в научно-техническом проекте также поднимает вопрос об устойчивости партнёрства, и тогда к анализу привлекается модель с выходом одного из участников (или присоединения игроков с разной переговорной силой), где вопрос отдачи от партнёрства до известной степени является экзогенным: стратегия

каждого участника уже предполагает наличие некоторой функции выигрыша с известными значениями для каждой стадии и каждой конфигурации партнёрства.

«Стимулирование инновационной активности» – игра по типу «принципал-агент», в которой формируется размер и структура вознаграждения работника (агента) – доля от успеха или фиксированное вознаграждение – для успешной реализации проекта. Инструментом анализа является модель Штакельберга, и получены результаты, которые в целом соотносятся с практикой: схема с фиксированным вознаграждением оказывается предпочтительной. С практической точки зрения, такая схема предпочтительна, поскольку позволяет избежать ситуаций, при которых инженеру-исполнителю требуется выплатить крупную сумму – гораздо более высокую, чем получает менеджмент компании. Высокие выплаты сотрудникам создают риск внутрифирменных конфликтов и выхода проекта из зоны рентабельности.

«Распространение знаний внутри фирмы» исследуется в соответствии с задачей, известной как Дилемма заключённого. Основной вопрос состоит в том, что сотрудник, обладающий уникальным знанием, извлекает выгоду из асимметрии информации и не заинтересован в распространении знаний внутри фирмы.

«Выбор оптимальных параметров лицензионного договора» – раздел о выборе структуры вознаграждения: будет ли это паушальный платёж, роялти или их сочетание. К этой литературе следует добавить работы об оптимальных ставках роялти. Зачастую в них используются модели Курно или Штакельберга, а равновесия в разделении общей выгоды максимизируют личные полезности игроков.

«Взаимодействие правообладателя и инновационных фирм» – общее название для ситуаций, при которых один изобретатель рассматривает возможность лицензирования на рынке нескольких фирм.

Инструмент анализа в таком случае – модели Курно или Штакельберга. Существует, однако, вариация постановки, которая также предполагает кооперацию на первом шаге игры и некооперативное поведение – на втором.

«Кооперация партнёров в научно-технологическом проекте» – задачи, в которых возникают постановки как из некооперативных игр, так и кооперативных. В частности, Дилеммой заключённого иллюстрируются ситуации, при которых один из участников консорциума передаёт меньше знаний, чтобы сохранить за собой конкурентное преимущество. Эта постановка примечательна тем, что подобные ситуации возникают при взаимодействии неравных с технологической точки зрения партнёров: у передающей стороны велик соблазн связать провал передачи технологий с недостаточной подготовкой принимающей стороны, не считаясь с усилиями передающей стороны. Другой пласт задач исследуется с помощью кооперативных игр – собственно, размер долей участников (вознаграждения) исследуется именно ими. И в решении таких задач чаще всего используются две концепции: ядро и Вектор Шепли.

В работе [51] рассматривается передача технологий в рамках международного партнёрства и исследуются две постановки задачи: с полной информацией и с асимметрией информации. Доли участников получают точную оценку в первом случае и предельную (верхнюю для лицензиата и нижнюю для лицензиара) во втором. В основе анализа лежит идея равновесия, и статья примечательна не только как характерный пример исследования на основе равновесия. Среди прочего, в этой статье работа [69] упоминается как первая попытка исследовать кооперацию в передаче технологий методами теории игр. Как показывают [8], это не совсем соответствует действительности. Далее, одним из результатов заявлено распределение выгоды между лицензиатом и лицензиаром в пропорции 70% на 30%. Этот результат не являлся новым в начале 2000-х, поскольку ранее в среде профессиональной оценки из статистических данных к тому моменту уже было «Правило 25%» [179].

Обзор [8] даёт представление о распространённости моделей некооперативной теории игр для решения задачи об оценке вознаграждения. Причём в

модельных ситуациях стороны являются антагонистами, что лишь отчасти отражает отношения участников. Неявным образом предполагается, что технология позволит добиться заявленного результата, и общие контуры проекта известны. Пусть и с некоторой вариацией в масштабах – известны модификации с различными уровнями технологии, степенью вовлечённости участников в сделку. Иными словами, модели некооперативного поведения оставляют за рамками обоснование решения о сотрудничестве. Напротив, модели кооперативного поведения исследуют возможности для сотрудничества: условия участия в общем проекте, условия прерывания сотрудничества и распределение результата не всегда является целью анализа. Между тем, абстрактные параметры, трудно поддающиеся измерению на практике, возникают и в моделях кооперативной теории игр. В частности, функция выигрыша – именно она является предметом количественного анализа в кооперативных играх о технологическом сотрудничестве – может включать коэффициенты взноса и затрат участников, коэффициент синергетического эффекта [56]. Коэффициенты представляют собой параметры общих ресурсов и затрат каждого участника при попытке реализовать проект самостоятельно. В модели они служат для демонстрации некоторых эффектов сотрудничества, но с практической точки зрения не имеют решающего значения, поскольку усложняют анализ сотрудничества. Вместо оценки выигрыша от различных сочетаний сотрудничества, принимается положение о функциональном виде неаддитивной функции выигрыша, включающей коэффициенты, которые требуют измерений без явного указания на процедуру их количественной оценки.

Если рассматривать постановку задачи научно-технического сотрудничества с оценкой вклада участников, более продуктивным оказывается подход, изложенный в [178]. Участники сталкиваются с выбором из числа потенциальных партнёров. Причём различные сочетания приводят к разной стоимости проектов. Далее, Вектор Шепли, оперируя стоимостями, усредняет вклад каждого из участников. При этом, однако, следует обратить внимание на интерпретацию результатов. Например, конкуренция двух технологически сильных

игроков за финансовые ресурсы третьего игрока приводит к следующим оценкам выигрышей в кооперативной игре: игрок с финансовыми ресурсами получает оценку $2/3$, каждый из игроков с технологией получает оценку $1/6$. Это необязательно означает кооперацию трёх игроков в одном проекте. Также это может интерпретироваться как сотрудничество инвестора с долей в проекте $2/3$ и одного из изобретателей с долей $1/3$, которую потенциальный партнёр получает с вероятностью $1/2$.

Таким образом, отвечая на первый из поставленных вопросов, следует указать на то, что кооперативная теория игр позволяет ответить на вопрос об общем размере вознаграждения за технологию, в то время как некооперативная теория используется для анализа структуры платежей – размеров паушального платежа и роялти. Однако расчёт компонентов Вектора Шепли может оказаться полезным при разделении лицензионного платежа между фирмами, которые действуют в некотором консорциуме на стороне лицензиара: текущие платежи (роялти) следует разделять между фирмами пропорционально компонентам вектора. Это видно, в частности, на примере работы [178], в которой один из разделов посвящён разбиению стоимости портфеля прав интеллектуальной собственности по отдельным компонентам. Об эмпирических правилах стоит сказать, что они являются скорее реакцией на потребность в достаточно простом, разумном и нетрудоёмком методе, который позволяет относительно быстро прийти к соглашению участникам сделки. Действительно, относительно простые математические методы работают в контексте полной информации и с ограниченным набором параметров, удобных для той или иной модели, а более сложные конструкции с учётом асимметрии информации, совместимости со стимулами и детальной проработкой важности факторов сделки выходят за рамки категории «простой и нетрудоёмкий» для их применения участниками переговоров.

Второй вопрос – о ранжировании методов – также поставлен в работе [39], где формулировка проблемы приобретает следующий вид: каким было бы справедливое (fair) распределение долей технологического проекта между

его участниками. Модель предполагает трёх игроков, каждый из которых владеет некоторым патентом. Патент игрока А является ключевым, и он может войти в партнёрство с игроками В или С, чьи патенты дополняют разработку А. В качестве способов вычисления долей участников называются 1) переговоры, где сильный игрок А выбирает партнёра и назначает ему любую долю по типу игры «Ультиматум»; 2) структурная пропорциональность, при которой игроки на каждом уровне поровну делят выгоду – 50% на долю игрока А и 50% на долю консорциума из В и С (каждый из них также получает одинаковую долю – по 25% общей выгоды); 3) выборочная пропорциональность, при которой игрок А делит выгоду поровну лишь с одним участником – В или С; 4) чистая пропорциональность, при которой доля участника есть среднее арифметическое от их количества – 33%; 5) Вектор Шепли, когда игрок А получает $\frac{2}{3}$, а каждый из игроков В и С – по $\frac{1}{3}$ общей выгоды. Следует заметить, что авторам известно об эмпирическом «Правиле 25%», поскольку они сообщают именно о таком распределении по итогам консультаций с практиками. Далее, авторы показывают размеры вознаграждения – как валового, так и чистого (с учётом вложений в сделанные разработки) – численными примерами и приводят результаты лабораторного эксперимента. Использовалась вариация игры «Ультиматум»: на первом этапе сильный игрок А распределяет доли с игроком В, на втором – А с игроком С, на третьем А делает окончательное предложение либо игрокам В и С одновременно, и тогда они узнают о существовании конкурента, либо одному из игроков. Результат эксперимента оказался отличным от того, который наблюдается в классической игре «Ультиматум» – самая большая доля распределений приходилась на равномерное. Именно оно признавалось участниками эксперимента в дополнительном вопросе после игры. Распределений по Шепли, что примечательно, среди справедливых вариантов не было.

Результаты, полученные в эксперименте [39] удивляют на фоне того, что равномерное распределение не является характерным для социальных систем. В этом контексте можно вспомнить, например, о Законе Ципфа, Кривых

Лоренца, которые продемонстрировали свои возможности в целом ряде количественных исследований. В целом, как показывает исследование первого вопроса о методах разделения выгоды, каждый способ имеет свою сферу применения – одни оценивают общий объём вознаграждения, другие – способ выплаты. Причём в разных модельных постановках. Соответственно, строгое исследование в части ранжирования методов требует: во-первых, специальной постановки, которая может быть рассмотрена как в кооперативном, так и в некооперативном разделах теории игр; во-вторых, выбора критерия ранжирования методов; в-третьих, исследования решений с учётом соответствия критерию. Последние два пункта следует пояснить. В статье [39] ранжирование происходит по критерию «справедливости». Формальные методы такой критерий просто не улавливает. В таком случае можно говорить о «совместимости со стимулами» (обеспечивает/ не обеспечивает), вероятности успеха проекта (степени достижения запланированных результатов при той вовлечённости участников, которая способна обеспечить то или иное распределение общей выгоды) и т.д. То есть строгое сравнение результатов требует подготовки к формальным процедурам, выполнения соответствующего анализа, что выходит за рамки данной работы. А с опорой на одни лишь экспериментальные исследования (опросы, наблюдения за дележами и т.д.) ответить на данный вопрос не представляется возможным. Итак, продемонстрирован эксперимент с неожиданным результатом, который, однако, нельзя игнорировать. В то же время существуют исследования, которые показывают силу распределения согласно Вектору Шепли, работоспособность эмпирики (в том числе «Правила 25 %») – эффективность каждого метода в своих условиях без попытки их сравнения в серии экспериментов.

Теоретические работы о дележе общей выгода при передаче технологий, а также эмпирические правила непосредственно связаны с проблемой кооперации для разработки высокотехнологичной продукции, к которой относится и программное обеспечение. Несомненно, правила дележа не являются единственным фактором успешной разработки. Кооперации научно-

образовательных организаций, бизнеса и государства для осуществления инновационной деятельности уделяется много внимания, прежде всего, со стороны государственного управления. Так, определённый эффект оказало Постановление Правительства Российской Федерации от 9 апреля 2010 года №218 «Об утверждении Правил предоставления субсидий на развитие кооперации российских образовательных организаций высшего образования, государственных научных учреждений и организаций реального сектора экономики в целях реализации комплексных проектов по созданию высокотехнологичных производств» [224]. Обновлена Стратегия научно-технического развития Российской Федерации¹², которая также охватывает вопросы взаимодействия науки и бизнес-сообщества [168]. Исследования о ценообразовании, доле выручки при передаче технологий и распределения стоимости многокомпонентного продукта между участниками производственной кооперации вносят свой вклад в инновационное развитие в виде конкретных правил разделения ресурсов и анализа рациональных аргументов для сотрудничества.

4.2. Распределение стоимости программного комплекса между его компонентами

Теперь, после обстоятельного обзора ситуаций, которые затрагивают кооперацию при создании высокотехнологичной продукции, настало время обратиться к вопросу стоимостной оценки продукта и отдельных его компонентов. Цены на данные подробно обсуждались в Разделе 2. Рассуждения Раздела 3 о двухкомпонентном ценообразовании применимы как к составному информационному продукту (например, информационно-аналитическая система с доступом по подписке), так и к однокомпонентному (например, доступ к базе данных по подписке). Но как быть в тех случаях, когда требуется выделить стоимость отдельных составляющих многокомпонентного информационного продукта? Такое возможно, например, при переносе интерфейса информационно-аналитической системы на другой продукт. Сколько могут стоить права

¹² Утверждена Указом Президента России № 145 от 28 февраля 2024 г

на его использование отдельно от системы? Можно было бы пойти по пути ценового меню. Но также представляет интерес стоимостной оценки актива, не торгуемого на рынке. Иными словами, диссертационное исследование следует явным образом распространить и на те случаи, когда возможности для выявления спроса ограничены.

При стоимостной оценке программ и баз данных необходимо учитывать синергетические эффекты, которые реализуются благодаря их сочетанию. Интуитивно понятно, что программный комплекс с набором данных стоит больше, чем его компоненты. Хорошим примером здесь могут служить системы искусственного интеллекта. Одна нейронная сеть мало пригодна для принятия решений. Также и размеченный набор данных – массив в миллионы записей мало полезен в качестве составного произведения (к ним в законодательстве приравниваются базы данных). Однако при их соединении получается полезный инструмент классификации. И колмогоровская сложность иллюстрирует тезис о том, что ценность сочетания данных и программы тем выше, чем больше информации возможно извлечь из строки, подающейся на вход системы – нейронной сети, за которой стоят процедуры обучения и обучающая выборка. Актуальными сегодня также являются примеры CRM-система и платформы электронной коммерции. Программа для управления взаимоотношениями с клиентами (CRM, Customer Relationship Management) сама по себе является не более чем общей конструкцией/заготовкой/шаблоном для оцифровки общей логики взаимодействия с клиентами (также говорят «бизнес-процессов»). Её ценность (и, далее, стоимость) раскрывается только при объединении с полноценной базой данных конкретных клиентов, их истории покупок и взаимодействий, а также конкретного контекста пользователя этой программы. То же наблюдение справедливо для платформы интернет-магазина, но только если не касаться особенностей ведения бизнеса: правил наполнения каталогов, обслуживания логистики, системы лояльности и т.п. Изолированная механика интернет-магазина, воплощённая в программе, без базы данных о товарах может оцениваться практически по затратам разработчика.

Возможны случаи отклонения стоимости и в обратную сторону, которые также называют «уничтожением стоимости» или «эффектом каннибализма» [179]. Однако, как отмечается А.Н. Козыревым и В.Л. Макаровым, «от оценщика обычно требуют, чтобы стоимость целого была равна суммарной стоимости частей. В частности, это касается стоимости портфелей ИС и НМА, включающих программы и базы данных». Это требование ставит перед необходимостью решения проблемы методологического характера: необходимо разработать чёткое правило, которое позволяет корректно распределять стоимость портфеля интеллектуальных прав между его компонентами. Решение этой проблемы предложено в уже упомянутой работе [179], и оно состоит в использовании аппарата теории игр, а именно, Вектора Шепли. Следует кратко описать это решение для того, чтобы затем применить его к стоимостной оценке компьютерной программы – неторгуемого на рынке актива при известных ценах на данные. И далее, в подразделе 4.3. это решение развивается на случай «большого» портфеля прав с иллюстрацией конкретной техники снижения размерности задачи именно в контексте информационных продуктов, в чём состоит вклад автора диссертации. При этом контекст отсутствия рыночной сделки усиливается примером нарушения исключительных прав и расчётом компенсации для использования в судебном процессе.

Обращение к решению по Шепли в ситуации разделения стоимости портфеля интеллектуальных прав между его компонентами начинается с абстрагирования от конкретных свойств объектов в портфеле. Важными считаются только денежные потоки и вклады объектов в их генерацию. Все объекты нумеруются натуральными числами i от 1 до n , образуя множество элементов $N = \{1, \dots, n\}$. Предполагается, что подмножества элементов $S \subseteq N$ (например, «только программа», «только база данных», «программа и база данных вместе») возможно рассчитать его стоимость $v(S)$. Это стоимость, которую генерирует использование выделенного набора объектов, при условии, что все остальные элементы из N исключаются из портфеля.

Функция $v(S)$, $S \subseteq N$, где $v(\emptyset) = 0$, а $v(N)$ – стоимость всего портфеля, называется *характеристической функцией игры*. Полный расчёт $v(S)$ для всех возможных комбинаций объектов на практике чрезвычайно трудоёмок, но, как будет показано далее, это не всегда требуется, чему способствуют техники снижения размерности.

Решение по Шепли — это вектор $\varphi[v] = (\varphi_1[v], \dots, \varphi_n[v])$, который задаёт аддитивное распределение общей стоимости портфеля $v(N)$ между n по известной неаддитивной функции $v(S)$, $S \subseteq N$. Этот вектор удовлетворяя трём интуитивно понятным аксиомам:

S1 («аддитивность»): аддитивные оценки объектов, который вносят ощутимый вклад в ведение бизнеса, в совокупности равны интегральной оценке портфеля, или $\sum_{i \in S} \varphi_i[v] = v(S)$. Бесполезные в коммерческом отношении объекты получают нулевую оценку.

S2 («независимость от перестановок»): Оценка объекта не должна зависеть от его порядкового номера в портфеле.

S3 («устойчивость к корректировке»): Если неаддитивная функция стоимости представима в виде суммы двух независимых вкладов $(v + u)$, то и компоненты Вектора Шепли должны характеризоваться соответствующим разложением: $\varphi[v + u] = \varphi[v] + \varphi[u]$. Это означает, что правило распределения устойчиво к корректировкам в правило стоимостной оценки объектов.

Этих аксиом достаточно для однозначного определения значений вектора $\varphi[v]$. Решение по Шепли записывается следующей формулой:

$$\varphi_i[v] = \sum_{T \subset S: i \in T} \frac{(n-t)!(t-1)!}{n!} [v(T) - v(T \setminus i)], \quad i = \overline{1, n},$$

где сумма берётся по всем подмножествам S , содержащим объект i , $t = |T|$ – количество объектов (размер) подмножества T , $n!$ – факториал числа n (произведение чисел от 1 до n), $T \setminus i$ – обозначение множества всех элементов T за исключением i .

Содержательный смысл формулы состоит в вычислении среднего предельного вклада объекта i во все возможные коалиции (комбинации объектов).

Каждая компонента Вектора Шепли показывает, насколько в среднем возрастает стоимость того или иного набора объектов при добавлении к нему выбранного (оцениваемого по формуле) объекта (например, базы данных). Однако, как показал обзор методов дележа при передаче технологий (Подраздел 4.1.4), применение Вектора Шепли скорее выбивается из распространённой практики. Несмотря на своё преимущество в обработке эффектов синергии, популярность решения по Шепли в среде профессиональных оценщиков практически не сдвинулась с места за прошедшие двадцать с небольшим лет. Так, в своей книге, изданной в 2003 году, А.Н. Козырев и В.Л. Макаров отмечали: «На практике решение по Шепли, как и другие арбитражные схемы либо вообще не применяется, либо применяется в упрощенном виде, поскольку практические оценщики этой техникой не владеют» [179]. Далее, после приведённой цитаты в книге следует описание оценочной практике, согласно которой, активы разделяют на амортизируемые и не амортизируемые, приписывая им различные стоимости на основе различий в генерируемых денежных потоках. Эта техника может рассматриваться как сильно упрощённое, робкое, но всё же движение в сторону решения по Шепли.

В контексте данной диссертационной работы, однако, вопросы амортизации активов, как и постановки информационных продуктов, на баланс выносятся за скобки. Принципиальным является именно синергетический эффект, наблюдаемый в многокомпонентных информационных продуктах. Для начала, применение решения по Шепли демонстрируется на примере с низкой размерностью, который позволяет проиллюстрировать ход рассуждений, в том числе, с привлечением концепции колмогоровской сложности, которая ёмко и формально иллюстрирует зависимость извлекаемой из данных информации от используемой программы.

Выше, в Подразделе 2.6., показано: продавец на информационном рынке определяет цены на информирующие сигналы, оптимизируя свой ожидаемый доход. При этом, однако, изымается потребительский излишек. Можно считать, что, следуя этой процедуре, устанавливаются цены на данные. В

процедуре получения информации по Колмогорову участвуют сигнал об объекте (строка), сам объект (также строка) и программа. Как они соотносятся с товарами на рынке информационных продуктов? Модельная ситуация предполагает, что продавец обладает некоторым набором данных (частной информацией) и различным типам потребителей посылает сигналы об этом наборе. Причём продавец может дифференцировать предпочтения относительно данного набора: разные срезы набора данных по-разному востребованы разными типами потребителей. Эти срезы и являются товарами, а сведения о составе этих срезов передаются потребителям в виде цен – информирующих сигналов. А именно: наблюдая цену t_θ , потребитель типа θ формирует апостериорные ожидания о получении ценности $\sum_{q \in Q} v_\theta(q) x_\theta(q)$. Следующий слой рынка информационных продуктов – получение и обработка среза данных. Последний рассматривается как сигнал (строка) из определения Колмогорова. Потребитель каждого типа обрабатывает полученный срез данных доступным ему программным средством. Если потребитель типа θ получает строку s и имеет в своём распоряжении программу π_θ , реализующую метод программирования $A(\pi_\theta, s)$, то принятая им строка s характеризуется им как объект колмогоровской сложности $K^\theta(s)$. Программы π_θ могут рассматриваться в качестве активов, в то же время они являются товарами на рынке информационных продуктов. Схему ценообразования на программы следует рассматривать с помощью модели, отличной от той, что представлена для случая торговли информацией. Дело в том, что аналогия с информирующими о функционале программы не вполне уместна. Её функционал можно продемонстрировать на имеющемся у покупателя наборе данных (именно для его обработки и приобретается программа), не передавая исполняемый код на его вычислительные мощности: операции можно выполнить на стороне продавца.

Если программа создана для обработки срезов данных конкретной структуры, её применимость ограничена. Заметим, что и данные имеют ценность только при наличии средств их обработки. Эти наблюдения естественным образом подводят рассуждения к формализации с использованием

элементов кооперативной теории игр и распределения стоимости согласно решению по Шепли. Пусть для простоты существуют три программы $\{\pi_1, \pi_2, \pi_3\}$ и один набор данных s , который сообщает некоторые сведения о состоянии мира ω (как в модели из Подраздела 2.6). Различные программы, как говорилось, позволяют извлечь разное количество информации. Зададимся вопросом об оценке стоимости программы в использовании. Ответ на него даёт применение решения по Шепли – правило распределения стоимости между владельцами активов (программ и данных) при их совместной работе.

Если потребитель с программой типа θ (программой π_θ) обрабатывает строку s , извлекаемая им полезность положительна зависит от получаемого количества информации $I_\theta(s; \omega) = K(\omega) - K^\theta(s)$, где $K(\omega)$ – колмогоровская сложность объекта (состояния мира), одинаковая для всех потребителей. Это допущение соответствует теории принятия решений: в Подразделе 1.3. показана влияние данных и программ на компоненты задачи принятия решений. Считается, что данные приобретаются именно для того, чтобы повысить качество решения. Таким образом, «шум» или сведения с негативным эффектом исключаются из рассмотрения: оценка их полезности не является положительной с точки зрения покупателя. Положительная связь полезности и информации формально записывается выражением

$$u(\theta) = u(\theta, \omega, a) = u(I_\theta(s; \omega), a) = u(K(\omega) - K^\theta(s), a) = u(\theta, s).$$

Такая запись позволяет определить оценки $V(Y) = V(s, \pi_\theta | \theta \in Y)$ наборов Y с данными и программами естественным образом. Поскольку колмогоровская сложность по определению включает строку и программу для её обработки, отсутствие любого из этих компонентов оставляет колмогоровскую сложность объекта ω без изменений. Это учитывается при построении оценок $V(Y)$ наборов Y .

Если в наборе Y отсутствуют программы, колмогоровская сложность объекта не меняется, новая информация не поступает, и оценка оказывается нулевой:

$$V(s) = u(\emptyset, s) = u(K(\omega) - K(\omega), a) = u(0, a) = 0,$$

где символ пустого множества иллюстрирует отсутствие программы для обработки данных. Без программы сложность объекта ω остаётся постоянной независимо от переданных данных s , что соответствует отсутствию новой информации.

Ситуация нулевой оценки набора Y реализуется и в случае отсутствия данных:

$$V(\pi_\theta | \theta \in Y) = u(K(\omega) - K(\omega), a) = u(0, a) = 0.$$

Если в наборе Y потребителю доступны программы нескольких типов, то оценка набора определяется полезностью «наилучшей» из программ (той, которая позволяет получить наименьшую колмогоровскую сложность):

$$V(\omega, \pi_\theta | \theta \in Y) = \max_{\theta \in Y} u(\theta) = \max_{\theta \in Y} u(K(\omega) - K^\theta(s), a) = \max_{\theta \in Y} u(I_\theta(s; \omega), a).$$

Следует отметить, что в таком случае компоненты Вектора Шепли естественным образом интерпретируются через усреднение выигрыша потребителя в результате использования той или иной программы. Если принять нулевую оценку для пустого набора $V(\emptyset) = 0$, и записать компоненты Вектора Шепли:

$$\varphi_i = \sum_{Y: i \in Y} \frac{(n-y)!(y-1)!}{n!} [V(Y) - V(Y \setminus i)], \quad i = \overline{1,4},$$

то разности в квадратных скобках для наборов с программами π_1 и π_2 , например, можно привести к выражению:

$$V(s, \pi_1, \pi_2) - V(s, \pi_2) = u(I_1(s; \omega), a) - u(I_2(s; \omega), a),$$

что фактически отражает дополнительную ценность потребителя от извлечения дополнительной информации в результате использования более совершенного (в смысле $K^1(s) < K^2(s)$) программного обеспечения.

Если в численном примере задать $u^*(1) = 1$, $u^*(2) = 0,9$, $u^*(3) = 0,8$, то компоненты Вектора Шепли принимают значения $\varphi_1 = 0,38$, $\varphi_2 = 0,31$, $\varphi_3 = 0,27$, $\varphi_4 = 0,72$. Мы можем вспомнить, что $\varphi_4 = t_1 + t_2 + t_3$, и, исходя из денежной оценки, пропорциональным образом оценить стоимости программ, как $(t_1 + t_2 + t_3) \cdot \varphi_\theta / \varphi_4$, $\theta = \overline{1,4}$.

4.3. Анализ портфеля программ большой размерности

Обратимся к случаю разделения стоимости портфеля прав на программы в случае большой размерности – большого количества объектов в портфеле. Как отмечено выше в Подразделе 4.2., вычисление компонентов Вектора Шепли требует рассмотрения всех возможных комбинаций из элементов портфеля. Но с ростом размерности растёт и количество таких комбинаций и, соответственно, количество вычислений. А.Н. Козыревым предложен подход анализа подобных ситуаций [178]. В его основе лежит процедура сокращения размерности задачи, которая включает такие операции, как отделение независимых элементов портфеля и объединение отдельных компонентов портфеля в кластеры. Так, объекты следует разделять по проектам, в которых они участвуют. С каждым из проектов связан свой денежный поток, и он разделяется только между входящими в проект объектами. На практике сложность представляют именно такие объекты, которые в разных сочетаниях используются в нескольких проектах. Тогда можно прибегнуть к другим способам кластеризации. При этом, однако, следует руководствоваться следующими принципами распределения стоимости между кластерами или компонентами портфеля [178].

1. Рыночная стоимость целого принимается равной стоимости частей.
2. Необходимо выявить эффекты «каннибализма» – частичного или полного обесценения нематериального актива в силу наличия иного актива, частично или полностью принимающего на себя те же функции, связанные в том числе с правовой охраной.
3. Необходимо выявить положительный синергетический эффект от использования оцениваемых объектов в различных сочетаниях в других проектах.
4. Синергетические эффекты, если они есть, должны быть либо выделены в отдельный «актив», либо, наоборот, их суммарное влияние на стоимость разнесено по отдельным объектам оценки.

5. Объекты оценки группируются по: (а) инвестиционным проектам, в которых эти объекты используются; (б) способу и надежности правовой охраны; (в) территории, на которой обеспечена правовая охрана и есть потенциальная возможность использования.

В соответствии с указанными принципами строится формализованная процедура разбиения стоимости, которая включает следующие этапы [178]:

1. Формирование кластеров.
2. Определение веса выделенных кластеров в рамках одного проекта.
3. Построение коалиционной структуры и присваивание весов разным «коалициям» из кластеров.
4. Вычисление эффектов синергии и каннибализма.
5. Присвоение кластерам скорректированных весов.
6. Вычисление компонентов Вектора Шепли по скорректированным весам.

Приведённая процедура может быть адаптирована к случаю оценки размера компенсации при нарушении исключительных прав. Изначально такая процедура предложена для анализа нарушения прав на товарные знаки [186, 210]. Её, однако, можно перенести на случай программных продуктов. Основанием для этого служат следующие соображения. Во-первых, как и в случае с товарными знаками, программные компоненты в различных комбинациях могут использоваться в разных программных продуктах. При этом, однако, следует отметить трудности подтверждения контрафакта. Товарные знаки наблюдаемы на упаковке. Программный код скрыт от пользователя, и для подтверждения нарушений может потребоваться специальная экспертиза.

Во-вторых, современные компьютерные программы представляют собой результат работы коллектива авторов, так что их названия становятся чем-то вроде наименования, близкого по своим функциям к товарному знаку, хотя и без соответствующей правовой охраны. Хорошим примером здесь могут служить программные библиотеки, которые распространяются под уникальными наименованиями и предназначены для встраивания в масштабные программные продукты.

В терминах программ адаптированная процедура имеет следующее представление. Пусть программы T_1, T_2, \dots, T_n Фирмы 1 используются в продуктах R_1, R_2, \dots, R_m . И пусть Y_k – множество программ, используемых при написании кода продукта R_k . Пусть у Фирмы 1 нет других объектов правовой охраны. Фирма 2 нарушает права на программу T_i из Y_k . Соответственно, оценка исключительных прав на программу T_i в рамках доходного подхода по денежному потоку от продукции R_k привела бы к завышению суммы компенсации. Отчасти это справедливо для оценки по затратам или рыночным сделкам, поскольку отдельные элементы программы – фрагмент кода, депонированный соответствующим образом – также могут охраняться в качестве программ для ЭВМ. Однако самое главное то, что использование нескольких программ из портфеля правообладателя для написания кода одного продукта может создавать эффект «каннибализма», так как коммерческий эффект продукта связывается с большим количеством объектов с правовой охраной. Здесь следует сделать специальную оговорку, которая отмечает специфику программ. Заявитель может оформить документы на одну программу для ЭВМ, указывая фрагмент её листинга (программного кода на языке программирования), или документы на несколько составных частей этой программы, указывая фрагменты листинга общей программы как программные реализации двух различных алгоритмов, обеспечивающих различный функционал (хотя и в рамках более длинного программного кода).

Теория дележей позволяет взглянуть на проблему выделения вклада отдельного объекта в общую стоимость портфеля прав под другим углом – не в терминах показателей (денежных потоков, затрат и т.п.), связанных с отдельными элементами. Теоретически следует рассматривать все сочетания программ для ЭВМ, которые используются Фирмой 1. Практически это реально, когда в портфеле на так уж много объектов исключительных прав. Если их много, то количество возможных сочетаний растет очень быстро и рассмотреть их все становится физически невозможно. И именно в этом месте можно обратиться к процедуре разбиения стоимости из [178] и Вектору Шепли.

В рассматриваемом случае в роли компонентов портфеля выступают зарегистрированные программы для ЭВМ, исключительные права на которые принадлежат Фирме 1. Все они пронумерованы числами от 1 до n , множество из номеров естественно обозначить через N . Роль коалиций играют различные наборы программ, необязательно связанные с одним продуктом. По этой причине для произвольного набора программ целесообразно ввести отдельное обозначение S . Игра в форме характеристической функции – это функция, определенная на совокупности всех подмножеств множества N и сопоставляющая каждой коалиции $S \subset N$ максимальный возможный выигрыш для этой коалиции $v(S)$. В нашем случае под $v(S)$ логично понимать стоимость гипотетического портфеля программ с номерами из множества S , при условии, что остальных программ в портфеле Фирмы 1 нет. Если бы такие величины для каждого возможного портфеля программ были известны, то стоимость программы для ЭВМ T_i определялась бы по формуле

$$\varphi_i[v] = \sum_{S \subset N, S \ni i} \frac{(n-s)!(s-1)!}{n!} [v(S) - v(S \setminus i)], i = \overline{1, n}, \quad (4.3)$$

где s – число элементов в множестве S , а суммирование идет по всем S , включающим i .

Сразу стоит отметить, что воспользоваться формулой (4.3) непосредственно крайне сложно, поскольку число возможных подмножеств множества N растет очень быстро с ростом n . А вычислять заранее значения $v(S)$ для всех $S \subset N$ бессмысленно. У нас с каждым продуктом R_k связан определенный набор программ Y_k . Логично предположить, что именно с этими программами связан денежный поток от продукта R_k , остальные программы не влияют на этот денежный поток от данного продукта, а он в свою очередь не влияет на их стоимость. А потому для каждого k логично ввести функцию $v_k(S)$ с носителем Y_k , определенную на всех подмножествах $S \subset N$, но удовлетворяющую равенству $v_k(Y_k \cap S) = v_k(S)$ для любого S . Если мы сможем построить функции v_k для каждого k , то можно решить задачу дележа для каждого продукта отдельно, используя формулу (1) для v_k . Разумеется, при этом ненулевые

значения получают только компоненты вектора Шепли с номерами знаков из Y_k .

Окончательный результат получим в виде суммы

$$\varphi_i[v] = \sum_{k=1}^m \sum_{S \subset N, S \ni i} \frac{(n-s)!(s-1)!}{n!} [v_k(S) - v_k(S \setminus i)], \quad i = \overline{1, n}. \quad (4.4)$$

Процесс оценки программы T_i , таким образом, включает 3 стадии:

1) Разделить денежные потоки Фирмы 1 между всеми продуктами R_1, R_2, \dots, R_m .

2) Для каждого k от 1 до m построить функцию выигрыша v_k , сопоставляющую каждому набору программ $Y_j \subset Y_k$ их стоимость $v_k(Y_j)$, связанную с денежным потоком от продукта R_k .

3) Рассчитать стоимость отдельных объектов по формуле (4.4).

Наиболее спорным в этой конструкции выглядит предположение, что можно реально построить функции v_k . При большом количестве программ следует провести укрупнение объектов – объединить неразрывно связанные друг с другом программы в группы. А независимые элементы считать без учёта эффектов синергии и «каннибализма». Также надо учитывать тот факт, что

$$v_k(S) - v_k(S \setminus i) = 0, \quad i \notin Y_k,$$

а потому

$$\begin{aligned} & \sum_{S \subset N, S \ni i} \frac{(n-s)!(s-1)!}{n!} [v_k(S) - v_k(S \setminus i)] \\ &= \sum_{S \subset Y_k, S \ni i} \frac{([Y_k] - s)!(s-1)!}{[Y_k]!} [v_k(S) - v_k(S \setminus i)], \end{aligned}$$

где $[Y_k]$ – число элементов множества $[Y_k]$. В нашем случае это количество программ для ЭВМ, используемых для реализации одного конкретного продукта R_k . И дополнительный денежный поток любая из программ даёт в том случае, если отвечает потребностям какой-то дополнительной категории покупателей. Поэтому можно оценить влияние разных программных блоков, используя традиционные инструменты маркетинга.

Предлагаемая схема расчетов применима отнюдь не только к случаю фирмы, чей портфель интеллектуальной собственности состоит из одних лишь программ для ЭВМ. В данном случае это упрощающее предположение принято, чтобы лучше показать суть перехода от вербальных рассуждений к математическим конструкциям и обратно. В общем случае следовало бы предпринять дополнительные действия по разграничению портфеля программ для ЭВМ и других исключительных прав.

Обратимся к вопросу группировки программ ещё раз - как в контексте снижения размерности задачи (количества объектов и, соответственно, количества их комбинаций), так и в контексте конкретных конфликтных ситуаций на рынке. В судебных процессах возможны ситуации, когда правообладатель заявляет максимально возможный размер компенсации за нарушение интеллектуальных прав. Такое поведение выглядит вполне рациональным, поскольку 1) соответствует предельным ожиданиям заявителя и 2) ставит ответчика в позицию, при которой ему нужно предлагать убедительные аргументы для снижения размера компенсации. Один из возможных механизмов завышения состоит в том, что правообладатель, заявляя о компенсации за нарушение прав в двукратном размере стоимости прав, исходит из оценки всего денежных потоков от продажи продукта R_k . Если для обеспечения функциональности продукта R_k используется несколько программ Y_k , а нарушитель использовал лишь некоторую часть $W \subset Y_k$, то сумма компенсации может быть завышенной и несправедливой с точки зрения ущемления интересов нарушителя. Возникает задача корректного выделения стоимости прав на нарушенные программы для ЭВМ W из общего набора Y_k .

Теория дележей подсказывает, что права на более трудные в разработке и более функциональные программы должны стоить дороже по сравнению с правами на программы, которые достаточно просты в разработке. То есть при участии в похожих проектах дороже те программы, чья уникальность выше. Это можно продемонстрировать на следующем примере [214].

Пусть имеется всего шесть активов – наборов данных и программных средств в различных пропорциях. Если наборы данных неразличимы между собой и программы выполняют одни и те же операции, то, согласно формуле Шепли, более ценным набор данных оказывается, если он доступен лишь в одном экземпляре. И это интуитивно понятно – более редкий актив имеет более высокую ценность. В Таблице 4.1 представлены оценки компонентов Вектора Шепли для ситуаций, когда множество активов состоит из 1 набора данных и 5 программ, 2 наборов данных и 4 программ, 3 наборов данных и 3 программ. При этом наборы данных неотличимы между собой, как и программы. Можно считать, что на рынке доступны идентичные по функционалу продукты и каждая программа в сочетании с данными имеет единичную ценность.

Таблица 4.1. Оценки активов в случае неразличимых наборов данных и программ. Источник: Автор.

	Оценка набора данных	Оценка программы
1 набор, 5 программ	0,833	0,033
2 набора, 4 программы	0,333	0,083
3 набора, 3 программы	0,167	0,167

Можно усложнить ситуацию. Пусть существует выделенный набор данных, а также выделенная программа. Они являются «наилучшими» в том смысле, что в сочетании друг с другом дают единичную ценность, и доступны лишь в единичном экземпляре. Также будем варьировать соотношение программ и доступных наборов данных, имея в виду, что сочетание «наилучшего» актива с обычным имеют ценность 0,9 – на 10% хуже «наилучшего сочетания». Тогда компоненты Вектора Шепли на наборы данных и программ выглядят так, как представлено в таблице 4.2.

Таблица 4.2. Оценки активов в случае «наилучших» наборов данных и программ. Источник: Автор.

	Оценка «наилучшего» набора данных	Оценка обычного набора данных	Оценка «наилучшей» программы	Оценка обычной программы
1 наилучший набор, 4 обычные программы, 1 наилучшая программа	0,800	НЕТ	0,080	0,030
1 наилучший набор, 1 обычный набор, 3 обычные программы, 1 наилучшая программа	0,350	0,300	0,125	0,075
1 наилучший набор, 2 обычных набора, 2 обычные программы, 1 наилучшая программа	0,200	0,150	0,200	0,150

Таблица 4.2 отражает дополнительную премию за доступ к «наилучшим» активам. Сравнение таблиц 4.1 и 4.2 позволяет оценить величину этой премии по сравнению со случаем неотличимых по качеству активов.

Проведённый анализ математически иллюстрирует более высокую стоимость «редкого». Отчасти это можно проиллюстрировать и в терминах примера из Подраздела 4.1.3: более редкое имеет более высокую стоимость замещения и, соответственно, более высокий вес по сравнению с более простым в реализации решением. На самом деле эта ситуация открывает очень большое поле для исследований, но вернёмся к примеру с оценкой размера компенсации. Учитывая оговорку о трудности разработки, все программы T_1, T_2, \dots, T_n

делятся на две группы: «уникальные» (множество U) и «неуникальные» (множество NU). В первом приближении каждая программа претендует на пропорциональную часть стоимости портфеля. То есть $v(U) = v(NU) = 1/2v(T_1, T_2, \dots, T_n)$, где $v(T_1, T_2, \dots, T_n)$ – стоимость всего портфеля. Стоимость «неуникальных» программ корректируются на понижающий коэффициент, подчёркивая разницу в стоимости «уникальных» (наиболее ценных) и неуникальных (менее ценных) программ для ЭВМ. Например, при коэффициенте 0,5 компоненты вектора Шепли оказываются равными $\varphi(U) = 0,625$ и $\varphi(NU) = 0,375$. Внутри множеств можно разнести стоимость между входящими в них программами или провести очередной раунд ранжирования. В случае программ сильную связь с поправочными коэффициентами имеют различия в результатах. Так, если программа позволяет извлечь «больше» информации из того же набора данных в той же ситуации использования, она оценивается выше, поскольку выше её вклад в принятие информированного решения. Естественным количественным выражением этой оценки представляется запись в терминах колмогоровской сложности. Результаты обработки одной и той же строки данных разными программами также являются строками, степень близости которых и вклад которых в устранении неопределённости можно оценить. И эта оценка может использоваться для разделения программ на «уникальные» и «неуникальные», для определения соответствующего коэффициента.

Если же в судебном порядке заявлена сумма компенсации с привязкой к стоимости денежного потока, генерируемого продажами конкретного продукта R_k , то разделять программы на «уникальные» и «неуникальные» следует внутри множества Y_k и эксперту следует внимательно рассмотреть доводы в пользу значения для поправочного коэффициента.

Итак, экспертная процедура разделения стоимости портфеля прав на программы для ЭВМ между отдельными составляющими выглядит следующим образом.

1) Установить соответствие между продуктами и используемыми программами, права на которые были нарушены. Иными словами, для каждого

R_k , упомянутого в претензии правообладателя (исковом заявлении), определить множество Y_k .

2) Установить множество программ, права на которые были нарушены $W_k \subset Y_k$.

3) Классифицировать программы внутри каждого набора Y_k на «уникальные» и «неуникальные».

4) Разнести стоимость прав на набор Y_k , заявленную правообладателем, между входящими в него программами равномерным образом.

5) Выбрать поправочный коэффициент для отражения меньшей стоимости группы «неуникальных» программ.

6) С учётом поправочного коэффициента рассчитать стоимость групп «уникальных» и «неуникальных» программ по формуле (4.4).

7) Разнести полученную по формуле (4.4) стоимость прав внутри групп между входящими в них программами равномерным образом.

8) Сумму стоимостей для программ из W_k , полученную по данной процедуре, считать базой для определения размера компенсации.

Выводы раздела

Приобретая информационный продукт, пользователи фактически получают некоторый набор программных библиотек и данных, соединение которых и обеспечивает предсказуемый результат. О цене на продукт говорилось выше – в Разделах 2 и 3, но что можно сказать о его компонентах? Для их стоимостной оценки требуется выйти за рамки одного продукта и исследовать все варианты их использования. Оказывается, теоретико-игровой подход позволяет получить наиболее согласованные с теорией результаты. Параметризация через колмогоровскую сложность при этом не является необходимой, но оказывается полезной для иллюстрации разной степени востребованности программных функций одного и того же продукта в разных вариантах использования. Так, настроенная на обработку некоторого массива программа сможет обрабатывать данные и в «усечённом» виде. Обратное тоже верно: один и тот же массив данных может обрабатываться разными программами, приводя

к разным результатам. Однако какими бы ни были варианты использования оцениваемого информационного продукта, все они вносят свой вклад в стоимость, если они связаны с извлечением дополнительной информации.

5. Технологическая платформа для рынка данных

Вопрос о поддержке рынка информационных продуктов во многом связан с формированием среды, которая обеспечивала бы коммуникации между участниками рынка. Программная платформа представляется естественной средой таких коммуникаций по нескольким причинам. Во-первых, распространение посредством программной платформы допускает использование различных вариантов дистрибуции (коробочная лицензия, доступ по подписке, приобретение сопутствующих распределённых вычислительных ресурсов и т.п.) и ценообразования (разовый платёж, двухкомпонентные цены, периодические платежи и т.п.). Во-вторых, технические средства позволяют подключить к платформе различные сервисы, которые в реальном времени отслеживают объём использования продукта. В-третьих, платформы подобного рода уже используются для распространения программных библиотек и наборов данных. Наконец, в-четвёртых, теоретические аспекты взаимодействий в такой среде продолжают развиваться. Можно отметить не только работы о научных коммуникациях и поддерживающих их программных сервисах [86, 85, 174], но также работы в области стандартов для представления и описания программных моделей [194]. Методика построения базовой архитектуры такой платформы следует работе [215], и отраслевые стандарты позволяют сформировать её понимание. Следует привести примеры тех из них, которые относятся к автоматизированным системам, искусственному интеллекту и информационной безопасности:

- ГОСТ Р 59795-2021 Информационные технологии. Комплекс стандартов на автоматизированные системы. Автоматизированные системы. Требования к содержанию документов [239]
- ГОСТ Р ИСО/МЭК 24668-2022 Информационные технологии. Искусственный интеллект. Структура управления процессами аналитики больших данных [240]
- ГОСТ Р 59276-2020 Системы искусственного интеллекта. Способы обеспечения доверия. Общие положения [237]

- ГОСТ Р 56939-2016 Защита информации. Разработка безопасного программного обеспечения. Общие требования [236]
- ГОСТ Р 59925-2021 Информационные технологии. Большие данные. Техническое задание. Требования к содержанию и оформлению [238]

Перечисленные стандарты демонстрируют выжимку рекомендаций по результатам анализа лучших практик экспертами технических комитетов – специалистами в своей предметной области. Список стандартов в общих чертах характеризует общие принципы построения информационных систем, обеспечения их безопасности, а также аспекты, связанные с использованием технологий искусственного интеллекта. Далее даётся обзор основных пунктов перечисленных стандартов.

Стандарт [239] фиксирует требования к содержанию основных документов, разрабатываемых при создании автоматизированных систем, причём перечень документов оказывается достаточно широким. Он предполагает документирование разработки с самых ранних этапов – от пояснительных записок к эскизным проектам. Предусматривает проработку систем по широкому перечню направлений, в том числе, аппаратной реализации, информационному обеспечению, программному обеспечению, математическому обеспечению.

Далее, стандарты в области больших данных определяют процессы по обороту и аналитике больших данных. Они важны, поскольку именно данные являются основой любой системы с технологиями искусственного интеллекта. Эти стандарты предлагают эталонную модель процессов аналитики больших данных, вопросы обеспечения доверия к системам искусственного интеллекта со стороны потребителей результатов работы этих систем и регулирующих организаций, требования к техническому заданию в области оперирования большими данными. Процессы аналитики разделены на пять категорий: процессы заинтересованных сторон внутри организации, процессы развития компетенций, процессы управления данными, процессы развития аналитики, процессы интеграции технологий. Для каждого из 19 процессов из указанных 5

категорий стандартом зафиксированы название, краткое описание, цели, результаты. Также описана модель оценки качества процесса.

Вопрос доверия к системам искусственного интеллекта сводится к выбору существенных характеристик для оценки той или иной заинтересованной стороной. В общем случае теория гарантирует надёжность работы систем на основе статистических методов и нейронных сетей по типу перцептрона в заданной окрестности данных, на которых выполнялось обучение. Однако и в этих случаях необходимо помнить о доработках и дообучении моделей для подстройки модели к меняющейся действительности. При выборе существенных характеристик систем искусственного интеллекта рекомендуется руководствоваться следующими принципами:

- достаточность набора характеристик для принятия решения о возможности использования системы ИИ при решении конкретной прикладной задачи;
- простота и возможность измерения значений характеристик;
- отсутствие перекрытия между используемыми характеристиками;
- соответствие установившимся понятиям и терминологии;
- возможность последующего уточнения и детализации характеристик.

Существенные характеристики систем искусственного интеллекта в соответствии с ГОСТ Р ИСО/МЭК 9126-93 объединены в следующие группы: функциональные возможности, надёжность, эффективность, практичность, сопровождаемость (имеются в виду возможности проверять, изменять систему), мобильность [237]. Проверку доверия к системе по выбранным признакам осуществляют разработчики систем, потребители, регулирующие организации. И проверка может выполняться на трёх уровнях архитектуры: физическом, инфраструктурном, прикладном.

Сразу две из важнейших характеристик доверия к системам искусственного интеллекта – функциональные возможности и надёжность – непосредственно связаны с вопросами информационной безопасности, и в этой связи целесообразно обратиться к специальному стандарту [236]. Этот стандарт

устанавливает общие требования к содержанию и порядку выполнения работ, связанных с созданием безопасного (защищенного) программного обеспечения и формированием (поддержанием) среды обеспечения оперативного устранения выявленных пользователями ошибок программного обеспечения и уязвимостей программы. Среди прочего, стандарт обязует разработчика обеспечить безопасность среды разработки программного обеспечения, создать/выбрать руководство по разработке, реализовать и проводить проверки мер по разработке безопасного программного обеспечения. Упомянутые меры относятся к различным этапам разработки, в том числе, проектирования архитектуры, конструирования программного обеспечения, тестирования, инсталляции и приёмке, эксплуатации.

Однако помимо стандартов следует указать на конкретные решения для соединения спроса и предложения на рынке знаний – принципиальной задачи, в том числе, и на рынке программных продуктов. Платформенные решения для поддержки открытых инноваций предлагают одновременно опыт технического решения задачи о соединении поставщика с потребителя, а также базу для теоретических исследований.

5.1. Сети изобретательства, рационализации, инноваций

Широко известными и реализованными на практике являются платформы открытых инноваций по типу состязаний решений на краудсорсинговых платформах [169], корпоративных платформ Witology или «Биржа идей». Однако следует сослаться на КСИРИН (корпоративная сеть изобретательства, рационализации, инноваций) как на прототип системы с набором программных сервисов, которая направлена на соединение спроса и предложения на рынке знаний и инновационных решений [137, 182]. Её отличие состоит в систематическом использовании реальных опционов различного типа, в том числе внутренних «патентов» КСИРИН, действующих только внутри сети, а также потенциала для встраивания сервисов выявления предпочтений на основе процесса Гровса-Кларка, оптимизации на основе распределенных градиентных методов и т.д.

Сеть КСИРИН допускает несколько вариантов реализации: корпоративный (в рамках одной крупной компании), холдинговый (для группы компаний) и наиболее масштабный – открытая социальная сеть (ССИРИН), участником которой может стать любой пользователь, согласный соблюдать установленные администратором площадки правила.

В сети КСИРИН предусмотрен оборот трёх групп электронных объектов:

1. *Знания в широком смысле.* К ним относятся не только изобретения, ноу-хау и рационализаторские предложения, но и такие специфические объекты, как описания проблем и связи между проблемами и их решениями. В то время как другие платформы (открытых инноваций, рационализаторских предложений, патентные библиотеки) оперируют исключительно «запросами», «задачами» или «предложениями», КСИРИН приравнивает формулировку проблемы и предполагаемые связи между проблемами и решениями к ценным активам, за которые полагается вознаграждение. Это призвано стимулировать не только изобретательские и рационализаторские решения, но также постановку качественных задач, соединение спроса на знания с предложением.

2. *Внутренние электронные «деньги»* (в терминах цифровой экономики их также можно назвать «токенами»). Это инструмент внутренних расчётов. Использование токенов удобно для транзакций, не связанных с прямым переходом прав собственности, так как их обмен внутри сети не влечёт налоговых последствий. Ценность токенов обеспечивается возможностью их конвертации в реальные деньги по установленным правилам, и именно в момент этой конвертации формируется налогооблагаемая база. Принципиальная опора на токены открывает перспективы для выстраивания экономических процессов, поддерживающих оборот реальных ценностей, а не виртуальных баллов и бонусов, которые могут встречаться в системах поощрения.

3. *Электронные документы, подтверждающие права.* Подразумеваются сложные правовые конструкции, такие как опционы, пакеты опционов и

внутренние патенты (которые, по сути, также являются пакетами опционов). Эти документы устанавливают реализацию прав при наступлении определённых событий. Примером может служить опцион типа put на продажу ноу-хау по результатам успешной демонстрации технической разработки, в отношении которой установлен режим коммерческой тайны. Подобные документы позволяют провести параллель с государственной патентной системой, поскольку её также присущи опционы: право на подачу заявки, на её отзыв, на переход к экспертизе по существу и т.п. Но частный характер такой системы позволяет создать механизм для решения задач, стоящих именно перед КСИРИН.

Доступный пользователям сети (игрокам/ участникам) функционал определяется их ролями, что позволяет управлять информационными потоками, достигая целей создания КСИРИН. Основными типами пользователей являются:

- Изобретатели и рационализаторы. Во всех системах подобного рода этот тип пользователей является самой многочисленной группой, куда входит большинство участников, предлагающих свои идеи. Отличие КСИРИН, однако, состоит в способе вознаграждения и расширения функционала по работе с электронными документами.

- Эксперты. Их роль состоит в оценке поступающих предложений и формулировке сложных задач. Строго говоря, задачи оценки предложений и формулировки задач предполагают разную квалификацию и опираются на различные внутренние бизнес-процессы, но в концепции КСИРИН эти функции выполняют участники типа «эксперты» без дальнейшего деления.

- Менеджеры и администраторы. Отвечают за функционирование и администрирование системы. Задачи пользователей этого типа практически не отличаются от тех, что возникают в системах подобного рода. Особенности могут касаться форм поддержки пользователей и разрешения споров.

Пользователь каждого типа взаимодействует с системой посредством специализированного программного интерфейса. Предметом сделок в системе

являются не сами знания, а связанные с ними электронные документы – опционы. Когда пользователь предлагает идею, с ней связывается пакет опционов – набор возможностей, которые можно трансформировать в другие внутренние активы системы и в конечном счёте в реальные деньги.

Ключевых принципов в этом обмене два. Во-первых, вознаграждение выплачивается не деньгами, а набором прав, для реализации которых требуются дополнительные усилия. Это предотвращает неоправданные и чрезмерные траты, так как деньги возникают только после успешной реализации опциона. Во-вторых, пакет опционов выдаётся не автоматически, а по результатам экспертизы. Процедуры оценки могут быть разными: от классического рассмотрения специалистом до краудсорсинговых методов (голосование, рейтинги), что обеспечивает как глубину анализа, так и демократизм и защиту от недобросовестных экспертов.

Далее опционов могут участвовать в дальнейшем обороте и даже объединяться. Например, «внутренний патент на ноу-хау объединяется с внутренним патентом на формулировку задачи и внутренним патентом на знание, что это ноу-хау дает решение этой задачи. В результате такого объединения возникает новая стоимость, которой не было до объединения. Обладатели всех трех патентов получают опционы на доли в этом новообразовании. В рамках системы (КСИРИН или ССИРИН) это все еще пакет опционов» [137].

Изложение концепции КСИРИН в самых общих чертах опускает вопрос ценообразования на внутренние объекты сети. Необходимость специального ценообразования признаётся, и постулируются параллели между внутренними товарами площадки с объектами общественных благ, из чего делается вывод о необходимости назначения персональных цен. Как может выглядеть методология назначения персональных цен на внутренние опционы, в цитируемой работе не раскрывается. Напротив, в отличие от работы по КСИРИН, данное диссертационное исследование, напротив, обращается именно к методологии, что видно по трём предыдущим разделам, и предлагает конкретный способ реализации таких цен на практике – построение платформы для поддержки

оборота информационных продуктов, включающей набор сервисов для соединения спроса и предложения, в том числе, ценовые сервисы

5.2. Платформенное решение для рынка данных

Решение проблемы доступности данных, их повторного использования, распространения программных модулей для анализа социально-экономических процессов, поднятой, в частности, в работах [183, 212], может состоять в создании информационной системы, работоспособность которой поддерживают экономические механизмы и программные сервисы. Именно последние являются концептуальной основой системы. Это они позволяют получить больше информации о данных. Именно они влияют на ценность данных. Тем не менее, следует предложить общую схему информационной системы, которая помогает сформировать базовое представление о взаимодействии отдельных сервисов. Такая схема является ранее опубликованным авторским результатом [216], и здесь даётся её уточнение с учётом раскрытых выше ценовых сервисов и развития информационной архитектуры.

Работу информационной системы обеспечивают три крупных блока (Рисунок 5.1): пользовательский интерфейс, программные сервисы и каталог данных. Стоит отметить, что в системе не предусматривается хранение пользовательских массивов – каталог содержит лишь основные сведения о них, метаданные. Интерфейс системы позволяет пользователю формировать запросы – на загрузку, редактирование и чтение метаданных; на поиск массивов; запрос и улучшение цен за доступ к данным. Эти запросы обрабатываются сервисами, которые объединены в три базовые группы по функциональному признаку: а) для взаимодействия с каталогом данных, б) для запуска информационных посредников (роботов); в) для обеспечения взаимодействия между пользователями. Сервисы группы (а) обеспечивают запись, редактирование, чтение из каталога, в том числе для целей поиска и навигации. Сервисы группы (б) решают задачи назначения и улучшения цен, облегчающие пользователям выбор условий для распространения своих массивов. Наконец, сервисы группы (в) поддерживают обмен информационными сообщениями между пользователями и

системой: совершение юридически значимых действий (например, присоединение к пользовательскому соглашению), контакт пользователей для целей обмена данными, поддержка краудсорсинговых механизмов для обеспечения качества метаданных в каталоге. Ниже частично раскрывается функционал этих сервисов.

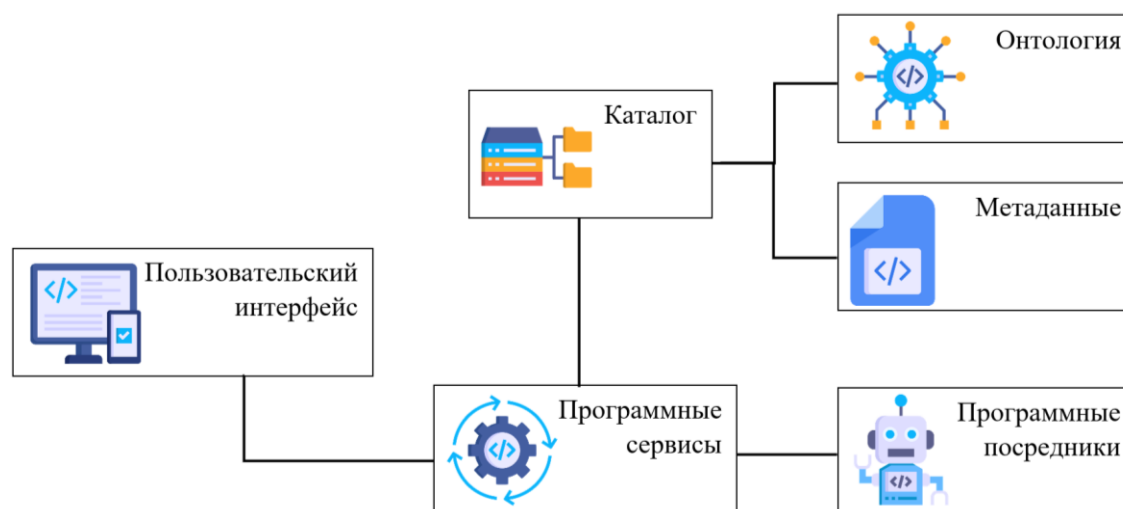


Рисунок 5.1. Ключевые компоненты информационной системы для поддержки рынка данных. Составлено автором с использованием изображений kliwir art на Flaticon.

С каталогом метаданных, который представляет собой базу данных, связана онтология. База содержит расширенное описание доступных массивов данных – не только введенный пользователем текст для характеристики своего массива. Онтология помогает расширить описание в части области применения данных – решаемых с их помощью задач, построения социально-экономических моделей.

Далее описываются основные характеристики блоков в общих чертах.

5.2.1. Каталог объектов

Экономические и социальные данные группируются по известным признакам. Основой для группировки могут служить, например, указания службы государственной статистики: демография, труд, промышленность, связь, транспорт и т.п. Причём по формам официальной статистики известно, какой показатель, собираемый на уровне наблюдений, участвует в расчёте того или иного агрегированного показателя. Уже существующий подход к группировке

данных по характеризующим ими объектам подсказывает возможность построения онтологии социально-экономических данных для облегчения навигации между ними. О том, почему речь идёт именно об «онтологии», а не о «таксономии» данных, говорится ниже при описании сервиса, облегчающего навигацию пользователей по каталогам данных. Эта (перспективная) онтология предполагает описание детальных данных – на уровне индивидов, домохозяйств, работников, предприятий. Агрегированные данные также могут включаться в базы, чтобы исследователи имели возможность анализировать макроэкономические процессы. Но именно детальная и актуальная информация даёт известную точность анализа. Детальные данные в некоторых разрезах уже доступны исследователям, в том числе, через специализированные информационные системы или курируемые научными организациями массивы данных. С опорой на уже доступные данные базовая классификация, построенная на принятой методической основе, может быть расширена до онтологии через идентификацию сферы применения массивов.

Привязка данных к онтологии – один из способов облегчить пользователям поиск необходимых им для работы массивов. Подход навигации через систему классификаторов давно известен и хорошо зарекомендовал себя. Но современные поисковые машины ориентируются на соответствие контента запросам пользователя. Причём вербальные запросы – уже сложившаяся и хорошо зарекомендовавшая себя практика. Однако для вербальных запросов нужен набор ключевых слов, ассоциированных с данными. Стандартный подход – описание массива данных пользователем на этапе их регистрации в системе – не исключает обогащение этого описания метаданными, извлечёнными из публикаций. На этапе загрузки сведений о доступном пользователю массиве данных существует возможность загрузки полных текстов тех публикаций, где описано использование этих массивов. Далее, известен сервис, который извлекает ключевые слова из публикаций и из контекста цитирований [85]. Причём контекст цитирования может обозначить сам пользователь. Таким образом, метаданные массивов можно дополнить ключевыми словами публикаций и –

более точно – ключевыми словами, которые обозначают контекст цитирования. И здесь уместно вернуться к тому, почему в отношении массивов данных вводится именно термин «онтология», а не «таксономия». Связь между объектами даёт для этого все основания. С одной стороны, имеются социально-экономические процессы или явления, полные тексты с описанием использования массивов данных. С другой стороны – сами массивы. И онтология устанавливает случаи использования данных – отношения между объектами.

5.2.2. Машинный анализ текстов и методы суммаризации

Обогащение онтологии – типами данных, компьютерных моделей, контекстами использования и связей между перечисленными – на основе анализа цитирований требует внедрение в платформенное решение технологий машинного анализа текстов. Методы извлечения ключевых терминов текстов имеют уровень развития, достаточный для реализации соответствующих сервисов для платформы, поддерживающей рынок данных (см., например, обзор [31]) Эти методы можно разделить на статистические [99], основанные на машинном обучении [120], структурные (лингвистические и графовые) (Beliga, Martincic-Ipsic, Meštrović, 2015) и гибридные – синтезирующие различные подходы [87].

Методы резюмирования (суммаризации) текстов – методы выделения ключевых предложений, которые наиболее ярко и ёмко (в смысле понимания человеком) характеризуют заданный объём информации. Подходы к резюмированию текста могут быть классифицированы по типу используемой техники суммаризации (*extractive summarization* – выделение из текста наиболее значимых, возможно, не связанных между собой единиц; *abstractive summarization* – генерация связного краткого содержания), по числу обрабатываемых документов (отдельный документ или многодокументный анализ), а также по зависимости или независимости от языка. Говорят даже о тематическом моделировании – структурировании массива текстов с выделением их общей темы, подтем (возможно, неявных) и семантической характеристики анализируемого массива. Базовыми подходами к тематическому моделированию являются

вероятностный латентно-семантический анализ [46] и модель скрытого размещения Дирихле [13]. Методы тематического моделирования и суммаризации текстов показывают надёжные результаты и уже применяются, в том числе, для навигации по научным базам [156], и нормативно-правовым актам [204].

5.2.3. Организационный механизм загрузки данных

Говоря о наполнении каталога, следует отметить несколько принципиальных пунктов, которые позволяют избежать некоторых организационных трудностей. Во-первых, загружаются не сами массивы данных, а только их метаданные. Сами по себе массивы могут исчисляться большим количеством требуемой для хранения физической памяти, что уже создаёт трудности. Но загрузка своего массива в библиотеку также вызывает у пользователя справедливый вопрос, как контролировать доступ к данным. Владельца каталога можно заподозрить в бесконтрольном распространении массива данных и даже в его модификации. Конечно, специальные технические средства могут предоставить некоторые гарантии по недопущению таких действий, но целесообразно ограничиться хранением в каталоге только метаданных. Это не только позволит сэкономить ресурсы хранилища, но также освобождает администраторов каталога от подозрений во вторичном интересе со стороны пользователя: целью системы является распространение сведений о массиве, а не массива данных как такового. Пользователь остаётся в своём праве распоряжаться имеющимся у него массивом по своему усмотрению.

Во-вторых, пользователь загружает сведения о своём массиве данных самостоятельно. Такой шаг позволяет избежать возможной юридической претензии за незаконное распространение объектов интеллектуальной собственности – базы данных, о чём говорилось выше, и метаданных. Последнее может показаться абсурдным, но прецедент создан: научные электронные библиотеки признаются судом виновными в распространении аннотаций к научным статьям как самостоятельных объектов авторского права, при отсутствии

специального соглашения с автором¹³. Таким образом, механизм загрузки должен предусматривать самостоятельную загрузку метаданных пользователем с обязательным присоединением к пользовательскому соглашению.

В-третьих, каталог данных предлагает пользователю рекомендательную услугу по назначению цены за доступ к имеющемуся у него массиву. О таком механизме сказано ниже, но важно отметить, что пользовательское соглашение также должно включать информирование о рекомендательном характере цены, предложений по её изменению, и окончательное решение остаётся за пользователем. Напротив, при взаимодействии двух пользователей – именно пользователей, а не человека с машиной – достигнутые ими предварительные соглашения освобождаются от такой оговорки и могут быть использованы сторонами для обязывания к сделке на достигнутых договорённостях в судебном порядке, если такое потребуется.

5.3. Обеспечение качества каталога

Загрузка пользователями сведений об имеющихся у них массивах данных – это первый шаг на пути качества сервиса, который исследователи получают в своё распоряжение. Открытые массивы данных имеют известные недостатки. Во-первых, наблюдения могут публиковаться «как есть», то есть с выбросами или некорректными данными. Во-вторых, данные могут распространяться в неудобном для обработки формате. В-третьих, могут иметь место ошибки в описании и уточнении области применения. Проверка пользователем и использование в анализе, таким образом, свидетельствуют о том, что часть работы по повышению качества данных уже выполнена. В случае, когда сохраняется потребность в сырых – необработанных результатах наблюдений

¹³ Можно указать, например, на постановление Девятого арбитражного апелляционного суда № 09АП-26164/2023 № 09АП-24271/2023 от 18.05.2023г. по делу № А40-64528/22, где приводится аргументация такого решения. Также можно указать на постановление Суда по интеллектуальным правам от 28.09.2023г. по делу № А40-269679/2021, в котором размещение названий и аннотаций не приравнивается к нарушению исключительных прав. При этом, что характерно, информационные цели и цитирование не сыграли своей роли в доказательстве, в другом – сыграли. То есть опасения в судебном преследовании за размещения информационных материалов о массивах данных не являются безосновательными.

– следует обращаться к первоисточнику. Таким образом, ссылка на первоисточник также составляет метаданные массива, учтённого в системе.

Второй шаг к более качественному сервису – устранение ошибок в написании и обогащении метаданных со стороны потребителей массивов. Если пользователь заметил ошибку в описании данных, ему предлагается возможность подать соответствующую заявку на исправление. Причём заявка становится заметна всем пользователям системы, и её либо одобряет автор загруженного массива, либо участники специальной отметкой о полезности данной заявки повышают её валидность. Это один из вариантов работы краудсорсингового механизма, но могут быть и другие реализации. Также пользователи могут загрузить свои статьи об использовании массива данных, что далее, через сервис для определения контекста цитирования, приводит к расширению и изменению веса ключевых слов, ассоциированных с конкретным массивом данных. Перспективными для системы (и полезными для других пользователей) также могут оказаться отметки потребителей массивов о проблемах дрейфа данных, дублирования массивов, индикация рисков открытого массива данных.

Следует заметить, что в повышении качества каталога могут участвовать не только участники платформы, но и некоторый сервис на основе машинного анализа текстов, основанный на технологиях, кратко обозначенных в Подразделе 5.2.2. Действительно, технологии могут выявлять сферы применения данных или компьютерных моделей, выявлять опечатки и неточности и т.п. Результаты работы программных сервисов, в свою очередь, становятся предметом для проверки пользователями.

5.3.1. Модель краудсорсинга

При использовании краудсорсинга существуют, по крайней мере, два принципиальных вопроса: мотивация участников и контроль качества. Современные информационные системы библиотечного типа автоматически устанавливают связи между объектом и вариантом его использования – цитированием, если обращаться к примеру научной литературы. Цитирования важны в

контексте онтологии объектов в смысле Подраздела 5.2.1. Они полезны для выявления контекстов использования данных и компьютерных моделей. Пусть потенциальные связи между объектами, сформированные человеком или машинным образом, помещаются в специальный список, называемый Списком задач. Эта пара объектов рассматривается как потенциальная связь. Но прежде, чем пара от пользователя будет помещена в список задач, он должен проверить некоторое количество пар из Списка задач. Для ясности предположим, что это число равно трем. При изучении пары пользователь должен сделать вывод, является ли корректной парой набор «данные – модель» или нет. Ответ по каждой паре также хранится в Списке задач в качестве атрибута пары. В результате одна новая пара увеличит Список задач, а решения по трем парам принесут им дополнительные баллы. Три пользователя должны дать свои оценки каждой паре из Списка задач, после чего решение об установлении связи по каждой паре принимается по трём голосам по правилу простого большинства. Однако количество голосов, необходимых для принятия решения, определяется оператором платформы, и оно может превышать три. Когда пара из Списка задач набирает два голоса «за», устанавливается связь между объектами, и пара удаляется из Списка задач. Когда два голоса «против», пара удаляется. Технически вновь созданные пары могут храниться в отдельной базе данных или в общей, но с присвоением специальной метки.

Анализ возможностей применения методов краудсорсинга для повышения качества данных в библиотеках показывает, что сами по себе они (методы) оказываются недостаточными [77]. Подробности такого анализа изложены ниже в Подразделе 5.3.3. В этой связи технологии искусственного интеллекта, связанные с анализом текстов и частично изложенные в Подразделе 5.2.2, способны дополнить краудсорсинг. В том числе, результат работы «толпы» может быть использован для формирования обучающего набора данных.

5.3.2. Препятствия для внедрения модели

При разработке методов краудсорсинга следует учитывать два аспекта. Первый – встраивание краудсорсинга в бизнес-модель оператора платформы.

Второй – недостаточная активность пользователей в качестве работников краудсорсинга.

Организация краудсорсинга связана с вознаграждением участников. Это одна сторона первого аспекта. Другая его сторона может быть связана с переводом платного функционала в бесплатный. Это хорошо иллюстрирует пример электронной научной библиотеки eLibrary. В этой системе полное управление персональным профилем доступно по платной подписке. Таким образом, функции редактирования цитирований станут бесплатными. Возникает естественный вопрос - скажется ли это изменение на количестве платных подписок. В настоящий момент расширенные возможности платной подписки доступны не авторам, а исследовательским организациям. Оплатив подписку, организация получает доступ к полному профилю каждого своего сотрудника.

Еще одно препятствие – вовлечение пользователей – также существенно влияет на результативность механизма. Пользователи системы могут столкнуться с большим объёмом работы, и в этой связи возникает закономерный вопрос о том, как быстро все предполагаемые пары из Списка задач будут проверены методами краудсорсинга. Чтобы разобраться в этом вопросе, следует сравнить два показателя. Первый – это скорость модерации пользователями, стимулируемая исключительно внутренней мотивацией к улучшению их личных профилей (описаний предлагаемых рынку наборов данных или компьютерных моделей). Второй – скорость роста Списка задач. Если бы внутренняя мотивация могла подтолкнуть пользователей к модерированию большего количества задач, чем поступает в Список задач, администратор платформы мог бы сэкономить деньги на вознаграждении. Внутренняя мотивация в модерации предполагает анализ того, как авторы распределяют свое внимание и усилия.

5.3.3. Моделирование

Если в информационной системе (платформе для рынка данных) существуют активные пользователи, готовые потратить некоторое время на то, чтобы улучшить качество метаданных, следует исследовать вопрос об их

вовлечении в работу, организованную на принципах краудсорсинга. Продемонстрируем пример такого анализа на примере численного моделирования методом Монте-Карло, чтобы ответить на вопрос об объеме работы, которую могли бы выполнить заинтересованные пользователи. При этом следует помнить, что проверка пар между объектами для обогащения метаданных не является аналогией работы пользователей-экспертов КСИРИН, которые чётко устанавливают связь между спросом и предложением. Работа экспертов похожа на деятельность технологического брокера и, таким образом, оценивается по другим принципам.

Рассмотрим численный эксперимент с оценкой достоверности пар через механизм краудсорсинга. Предположим, что автоматический алгоритм помещает в Список задач N пар. Пользователи могут увеличить этот список ещё на n пар. В этом случае один пользователь j отправляет n_j пар, проверяя взамен k пар из Списка задач. Предположим, что для подтверждения (или отклонения) связи требуется m голосов, большинство из которых поддерживает соответствующее решение. Адекватный алгоритм ранжирования заданий и решение большинства до получения m голосов способствуют уменьшению количества сравнений, которые необходимо выполнить. Однако в худшем случае каждая из $(N + n)$ пар должна быть проверена m раз. При этом пользователи могут обеспечить только $k \times n$ проверок. Таким образом, все работы из списка задач будут выполнены только в том случае, если $m \times (N + n) \leq k \times n$, или $N/n \leq (k - m)/m$.

Имитационная модель вычисляет вероятностное распределение отношения (N/n) , а в качестве количественных значений привлекает наблюдения соответствующих показателей в электронной библиотеке eLibrary, которая позволяет редактировать связи цитирования. Таким образом, при моделировании используются абсолютные значения цитирований в предположении о постоянной доле ошибок как для отдельного автора, так и для электронной библиотеки в целом. В ходе одного моделирования компьютерная программа вычисляет n_j для 4 000 активных пользователей ($j = 1..4000$) в соответствии с

эмпирическим распределением. Четыре тысячи активных пользователей – достоверное предположение. Для сравнения: в 2016-ом году eLibrary сообщало на своём интернет-ресурсе о 240 000 зарегистрированных пользователях, 93 000 из которых управляли своими профилями. После инициализации данных значение $N/(n_1 + \dots + n_{4000})$ рассчитывается для конкретной симуляции. Каждый запуск модели создает уникальный набор n_j и приводит к новому значению для $N/(n_1 + \dots + n_{4000})$. В общей сложности было выполнено пять тысяч симуляций, и это число дает распределение вероятностей для (N/n) , достаточное для оценки усилий, необходимых для исправления ссылки на источник. Результат моделирования распределения вероятностей (N/n) показан на рисунке 5.2.

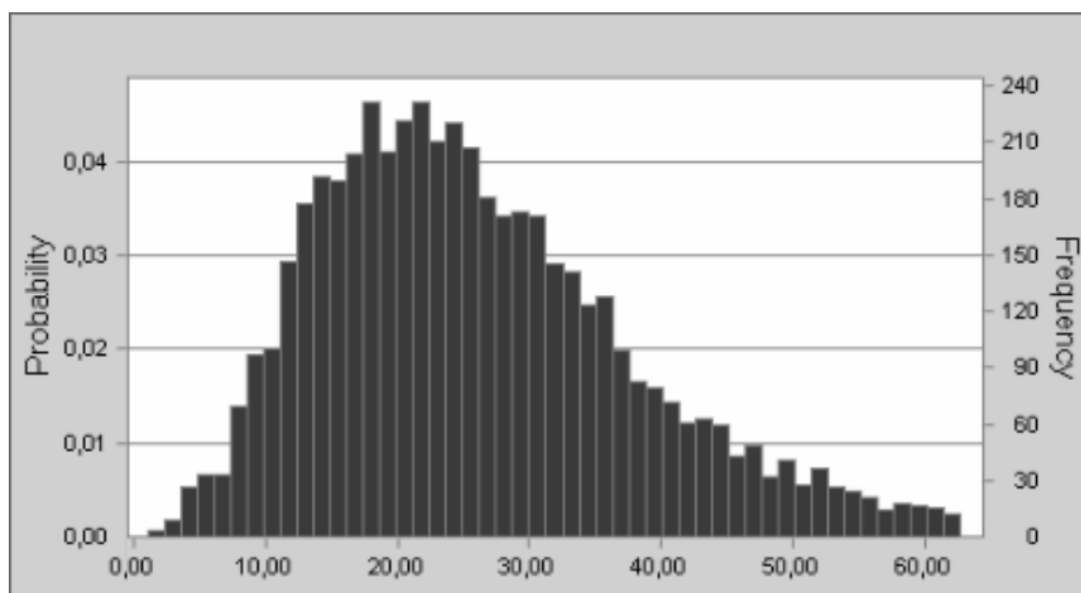


Рисунок 5.2. Вероятностное распределение для соотношения (N/n) . Источник: [77]

Анализируя процентиля распределения, можно констатировать, что (N/n) превышает значение 35,75 в 20% случаев. То есть с вероятностью 80% значение $(k - m)/m = 35,75$ является достаточным для того, чтобы Список задач был исчерпан 4 000 работниками краудсорсинга, руководствующимися исключительно внутренней мотивацией. Однако одна заявка на включение в Список задач от пользователя потребовала бы от него проведения $k = (36 \times 3 + 3) = 111$ экспертиз, если правило принятия решения требует $m = 3$ голосов

для того, чтобы подтвердить или отклонить пару. Издержки проведения 111 экспертиз одним пользователем чрезвычайно высоки и, вероятно, подорвут внутреннюю мотивацию. Таким образом, следует отвергнуть гипотезу о способности внутренней мотивации быть единственным механизмом краудсорсинга в системах подобного типа. Без внешней мотивации пользователи не смогут выполнить все доступные задачи в разумные сроки. Но оператор системы не должен отказывать пользователям с внутренней мотивацией. Вместо этого необходимо должным образом управлять двумя типами работников краудсорсинга – внутренне заинтересованными и оплачиваемыми работниками. Те, кто хочет улучшить метаданные своих товаров, изучают пары из Списка заданий, прежде чем их заявки будут добавлены в список. Остальные пользователи проверяют пары за вознаграждение. Внешняя мотивация включает тех, кто не видит потенциал для обогащения метаданных своих информационных продуктов, но кто хочет получить какое-либо вознаграждение. Такая дифференциация пользователей разумна, и в качестве дополнительной выгоды для активных пользователей можно было бы подумать о том, чтобы переместить их заявки в начало Списка задач.

5.4. Назначение цен за использование данных

Массивы данных могут распространяться за вознаграждение. Загружая информацию об имеющихся данных, пользователь сам назначает цену за доступ к ним. Она может быть нулевой. Может иметь некоторую положительную оценку. По мере накопления в системе информации о сделках/ контактах между пользователями становится возможной ценовая дискриминация – назначение разных цен разным группам потребителей данных (Козырев, 2011). В случае заинтересованности пользователь может воспользоваться функцией составления ценового меню для пользователей разного типа, сигнализируя о качестве массива данных, скрываемого за тем или иным сигналом (с соответствующей ценой). Термин «сигнал» в отношении массива данных остаётся неясным без развёрнутого комментария. В следующем абзаце понятие «информации» не привязано к какому-либо из упомянутых ранее

определений. В общем, подразумевается уточнение сведений об объекте, устранение неопределённости, и даже обывательское представление о термине приемлемо для понимания сути сигналов в процессе передачи информации.

Известен протокол передачи информации, который предполагает, что цена является для покупателя сигналом о качестве того, что за ней скрывается. В случае с информацией это более чем уместно по сравнению с «частными благами» или материальными товарами. Последние можно как-то измерить и оценить перед покупкой. С информацией дело обстоит сложнее: раскрытие информации часто подразумевает её передачу. Поэтому, чтобы «направить» покупателя именно к той информации, которую он ищет, в теории вводятся сигналы. Причём эти сигналы могут отождествляться с ценой (цена при этом и является сигналом о содержании информации), но могут передаваться за отдельную плату. Соответственно, приобретая сигнал, потребитель устраняет неопределённость относительно состава приобретаемой информации.

Простейший сценарий появления сигналов о качестве массивов данных – это распространение ограниченных фрагментов. Безусловно, описание массива в системе включает часть сведений и соответствует принятой онтологии. Но предположим, пользователю нужен не весь массив наблюдений не по всем показателям, а только их части. Например, из 50 имеющихся показателей пользователю нужно только 15, и он согласен приобрести их только в том случае, если вариация значений показателя под номером 6 по всем наблюдениям окажется достаточной для его целей. Соответственно, речь может идти о доступе к наблюдениям по одному-двум показателям, а затем (в случае положительного тестирования) – к наблюдениям по 15 показателям. Соответственно, возникает вопрос о том, как оценивать доступ к фрагментам массивам данных при повторяющихся взаимодействиях. И в этом вопросе также есть достижения, на которые можно опереться, как показывает Подраздел 2.6.

5.5. Оптимизация роялти и устаревание данных

Говоря о платном доступе к данным, нельзя не упомянуть инструменты оптимизации цен для пользователей – покупателей данных. Контент имеет тенденцию к сокращению срока коммерческой жизни. Так, спрос на популярные произведения в Интернете «затухает» в течение довольно непродолжительного времени, и только отдельные произведения пользуются популярностью в течение многих лет. По литературным произведениям, например, разумный срок правовой охраны оценивается в 8-15 лет – гораздо короче установленных законом жизни автора и 70 лет после его смерти [89]. Можно предположить, что такая же зависимость существует и для данных: более актуальные сведения ценятся сильнее тех, что описывают период десятилетней давности. Соответственно, цена за доступ к данным также может дифференцироваться не только по типу получателя, но и по времени. Следует ещё раз отметить – автор массива данных может указать безвозмездный вариант доступа. Но, если некоторые массивы доступны на платной основе, следует обсудить и то, как сделать этот – ценовой – барьер доступа разумным.

Предположим, что изначально доступ к данным планируется открывать за некоторую плату, которая остаётся неизменной в течение всего срока размещения. С учётом дисконтирования стоимость денежного потока ограничена, и можно говорить о том, как распределить эту величину между потенциальными пользователями в течение времени. Эту ситуацию можно моделировать через игру двух лиц – автора массива и обобщённого потребителя, которые при разных предпочтениях делят единичные ресурсы (ежегодные взносы за использование) в течение времени. Аппарат решения такой игровой ситуации предполагает решение с более высокими (относительно базового уровня) ценами в первые периоды и более низкими – в последние (вплоть до бесплатного доступа). Причём техника реализации алгоритма решения игры предполагает участие программы, имитирующей действия покупателя (при априорно заданных ожиданиях о его предпочтениях) [208]. Соответственно, сервис оптимизации цен за доступ к данным предполагает наличие вероятностного распределения о темпах «устаревания» данных – изменения спроса с течением

времени. И это распределение может уточняться по мере спроса на данные. К примеру, если спрос в начальные периоды оказывается ниже, автор массива может получить рекомендацию о снижении цены. Идея подключения программных сервисов для «нащупывания» цен на нематериальные продукты изложена в [78], и варианты их развития, в том числе, по встраиванию в КСИ-РИН и разделения платежей между участниками производственной кооперации (цепочки добавленной стоимости) описаны в [213].

Существуют и другие механизмы организации платного доступа, основанные на понимании данных как клубного блага [25]. В таком случае следует подключать механизмы выявления истинных предпочтений через систему голосования с тем, чтобы участники «клуба» оплачивали приоритетный доступ к данным в течение некоторого предустановленного периода, по окончании которого данные переходят в разряд общественного блага (а база данных – в категорию общественного достояния, говоря юридическими терминами) с открытым доступом всем желающим без взимания какой-либо платы.

5.6. Организационные вопросы функционирования платформы

Наряду с основными элементами цифровой платформы следует отметить некоторые вопросы, связанные с её функционированием. В частности, препятствия к обороту данных, вопросы безопасности, технология реализации цифровой платформы. Третий из перечисленных вопросов имеет скорее инженерное значение, однако, в данном исследовании – экономическом по своей сути – следует обратить внимание на доступность технологий и их функциональность. Среди возможных технологий для разработки платформы выделяется блокчейн как инструмент, положенный в основу некоторых сервисов для оборота нематериального [217]. Комментарий о возможностях этой технологии дан в Подразделе 5.6.3. Торговые ограничения, несомненно, создают препятствия для доступа к информационно-коммуникационным технологиям вообще и к электронной компонентной базе в частности. Хотя зависимость отечественной экономики от импортных поставок высока, производители конечной продукции в разных частях света зависят от внешних поставок

электронных компонент [127]. Однако наряду с этим следует отметить трудоёмкость создания онтологии для каталога объектов, которая существенно зависит от размера [106]. Указанное обстоятельство подсказывает естественный путь к повышению эффективности платформы для рынка данных. А именно, при разработке изначально следует ориентироваться на более широкий круг пользователей, не ограничиваясь национальным уровнем. Так, в России есть опыт создания информационной базы центров коллективного пользования и уникальных научных установок, доступной участникам СНГ, ЕАЭС, ШОС [167]. Этот опыт может оказаться полезным для подключения к предлагаемой платформе круга государств. Интерфейс платформы, в свою очередь, может включать в себя надстройку в виде большой языковой модели, обеспечивающей трансляцию каталога на языки стран-участниц платформы. Использование такой модели видится естественным ввиду успешной тематической классификации документов мультязычными моделями [74]. В целом, технологии интерфейса для подобной системы достаточно проработаны. Можно указать, например, на тексто-графическую систему представления знаний [143] и компактное мультязычное представление текстов в научном домене [146]. Поэтому разработка онтологии хотя и является затратным мероприятием, ложится на подготовленную почву из вспомогательных и дополняющих систем, которые оправдывают решение столь масштабной и трудоёмкой задачи.

5.6.1. Конкуренция с провайдерами услуг и разработчиками моделей

Доступ научного сообщества к собираемым корпорациями данным может натолкнуться на возражение следующего рода. Компании вкладывали собственные ресурсы для накопления этих данных, имеют право на монопольный доступ к ним, а расширение доступа создаст неравные условия по разработке сервисов на их основе, по выводу информационных продуктов на рынок, поскольку конкурент не понёс соответствующих издержек и, таким образом, не отягощён дополнительными обязательствами. Возражение выглядит

закономерным, если не вдаваться в детали данных и перечня лиц, которым открывается доступ.

Во-первых, рассмотрим суть накопленных данных. Они составляют информацию об экономическом и социальном поведении большой группы лиц – едва ли не всех граждан страны. Монопольный доступ к таким данным со стороны частных корпораций вызывает вопросы. Безусловно, существуют акты о доступе специальных служб и ведомств к таким данным. Но, как сказано выше, отечественная наука общественного профиля рискует застрять в неадекватной картине нашего общества. Для многих отраслей экономики доступны данные о выпуске продуктов – в разрезе моделей подвижного состава, например, или других характеристик. О потреблении продуктов и услуг в цифровой экономике таких детальных данных нет. Поэтому предлагается, реализуя поручения Президента, также продумать механизм получения этих данных.

Во-вторых, обратимся к конкуренции. Речь идёт именно о возможности научных исследований. Корпорации быстрее в выпуске продуктов, имеют профильные отделы, которые отвечают за улучшение пользовательского опыта, имеют несравнимый с научными организациями ресурс для продвижения своих продуктов. Это означает, что даже если какая-то из научных организаций задумает предложить свой коммерческий продукт на основе данных, он имеет низкие шансы на успех. Главный продукт науки – кадры высшей квалификации, научные публикации, научные мероприятия. Да, научные организации оказывают услуги (по проведению научных исследований) в рамках хозяйственных договоров. Но в целом, продукты и услуги показывают нацеленность на разные аудитории потребителей научного знания и информационных сервисов. В некотором узком сегменте деятельности корпораций – аналитических исследованиях – действительно может возникнуть конкуренция, но те, кто при этом обращается к научным организациям, идёт туда за брендом – за правом аргументировать свою позицию фразой «подтверждено научными исследованиями».

В-третьих, доступ к данным необязательно может быть свободным от обязательств. Известны различные типы лицензий на программные продукты – например, для некоммерческого использования частным лицом, академическая лицензия, лицензия для бизнеса и т.п. Каждая из них учитывает возможности и цели доступа к программным продуктам, накладывает соответствующие ограничения на функционал. Известны консорциумы организаций, которые, в том числе, разделяют общие цели и координируют свою деятельность. Полезным для научных исследований оказался бы такой механизм, который бы чётко устанавливал критерии доступа к социально-экономическим данным, хранящимся у корпораций.

5.6.2. Безопасность

Следующее возражение, которое можно ожидать в качестве препятствия к доступу, связано с вопросами безопасности – коммерческой инфраструктуры, коммерческого продукта, граждан и государства на всех уровнях управления. Безопасность финансовых сделок данных можно повысить путём их структурирования и разделения на этапы со своими атрибутами [252]. Однако, когда товаром выступают данные, важно учитывать риск несанкционированного доступа. Если данные открываются широкому кругу организаций, повышается риск их утечки и недобросовестного использования информации о покупках, передвижении, социальных взаимодействиях граждан. Ответ на это возражение также можно прорабатывать, имея в виду несколько пунктов.

Во-первых, практика обезличивания данных не является новой и уникальной. Здесь можно сослаться на результаты выборочных обследований. Каждая анкета фактически представляет конкретного индивида/ конкретное домохозяйство. Богатая и доступная коллекция таких данных собрана при повторяющемся обследовании RLMS¹⁴. Наборы данных включают доходы и

¹⁴ «Российский мониторинг экономического положения и здоровья населения НИУ ВШЭ (RLMS HSE)», проводимый Национальным исследовательским университетом "Высшая школа экономики" и ООО «Демоскоп» при участии Центра народонаселения Университета Северной Каролины в Чапел Хилле и Института социологии Федерального научно-исследовательского социологического центра РАН. (Сайты обследования RLMS HSE: <http://www.cpc.unc.edu/projects/rlms> и <http://www.hse.ru/rlms>)».

расходы, образование и занятость, медицинскую часть. И до сих пор неизвестны случаи деанонимизации респондентов и злоупотребления информацией.

Во-вторых, механизм, который позволил бы научным организациям получить доступ к данным, и не должен предполагать свободную передачу информации. Уже упомянутые данные RLMS доступны исследователям только после регистрации на информационном ресурсе проекта. Это самый мягкий барьер. Более сложные барьеры могут устанавливать строгие, но обоснованные требования к организациям, предполагающие, например, технологический аудит информационной инфраструктуры и политики обращения с конфиденциальной информацией у претендента. Здесь уместно напомнить об использовании медицинских данных для обучения искусственного интеллекта и об использовании соответствующих сервисов медицинскими учреждениями. Этот прецедент работы с чувствительными данными признаётся успешным, и он вселяет уверенность в том, что и для использования научным сообществом социально-экономических данных, собираемых корпорациями, также может быть найден подходящий механизм.

5.6.3. Блокчейн как возможное техническое решение для библиотеки

Анализ существующих и перспективных сервисов для управления интеллектуальными правами на основе блокчейн-инфраструктуры выявляет ограниченный масштаб их использования [217]. Можно говорить о локальных решениях, например, для обществ по коллективному управлению правами, которые стремятся повысить прозрачность своей деятельности. И подобные применения разумны лишь до тех пор, пока речь идёт об открытом блокчейн. Действительно, открытость всех операций с отсутствием координационного центра формируют доверие к информации, например, о размере и распределении лицензионного вознаграждения. Однако в случае необходимости редактирования блоков, системе должен назначаться администратор, и в случае сбора вознаграждения может потребоваться вмешательство, чтобы избежать

двойных платежей: авторскому обществу и автору в случае непосредственной договорённости лицензиата с последним. Закрытый блокчейн, в свою очередь, теряет свойство прозрачности и в этом смысле не отличается от решений на базе других технологий. Вопрос о снижении транзакционных издержек и росте доходов авторов также не является столь однозначным ввиду замещения одних посредников другими и высокой неравномерности в распределении зрительской аудитории. По причине существующих препятствий и противоречий эксперименты в данной области продолжаются, а уже реализуемые проекты находятся в поисках своей аудитории и совершенствовании своих сервисов.

5.7. Уточнённая архитектура

В заключении раздела следует привести информационную архитектуру цифровой платформы. При построении информационной архитектуры платформенного решения удобно отталкиваться от существующего уровня техники – известных информационных систем, в том числе, с приложениями к финансам. Необходимыми, отношения между которыми представлены на Рисунке 5.2, являются следующие.

- 1) Сервер приложений. Обеспечивает взаимодействие приложений с базами данных, координирует работу приложений.
- 2) Графический интерфейс. Обеспечивает визуальное представление данных, отчётов, обеспечивает взаимодействие пользователя с элементами платформенного решения.
- 3) Библиотека для онтологического представления объектов платформы.
- 4) База объектов. Содержит метаданные о доступных для распространения наборах данных, программных моделях.
- 5) Модуль для обеспечения краудсорсинговых работ на платформе. В качестве компонент содержит базу работников, список задач, протокол подтверждения результатов, интерфейс с базой объектов (для корректировки метаданных), сервис вознаграждения.
- 6) База пользователей. Содержит характеристики пользователей.

- 7) Ценовые сервисы для поддержки заключаемых пользователями платформы соглашений.
- 8) База данных для обучения. Выделенная база, которая используется для тестирования уязвимостей, области допустимых значений, доработки/ дообучения модели и т.п.
- 9) Тестируемая модель для обучения. Принятая на платформе технология обработки текстовых документов для обогащения метаданных объектов, которая находится в режиме доработки/ развития/ модификации.
- 10) Среда для обучения и разработки. Программная среда, которая позволяет вести разработку и тестирования и также отвечает требованиям доверия к получаемым с её помощью результатам.
- 11) Модуль машинной обработки документов. Принятая на платформе технология обработки текстовых документов для обогащения метаданных объектов. Обеспечивает функциональность работы с метаданными – не для тестирования и отладки.
- 12) Модуль подготовки отчётов. Программа, которая обеспечивает представление данных в установленных пользователем форматах.
- 13) Модуль передачи данных. Программа, которая обеспечивает чтение и запись данных в установленных форматах.

На данные элементы также рекомендуется накладывать требования безопасности (в частности, шифрование данных, разграничение прав доступа), соблюдение которых может реализовываться отдельным, не указанным на схеме элементом.

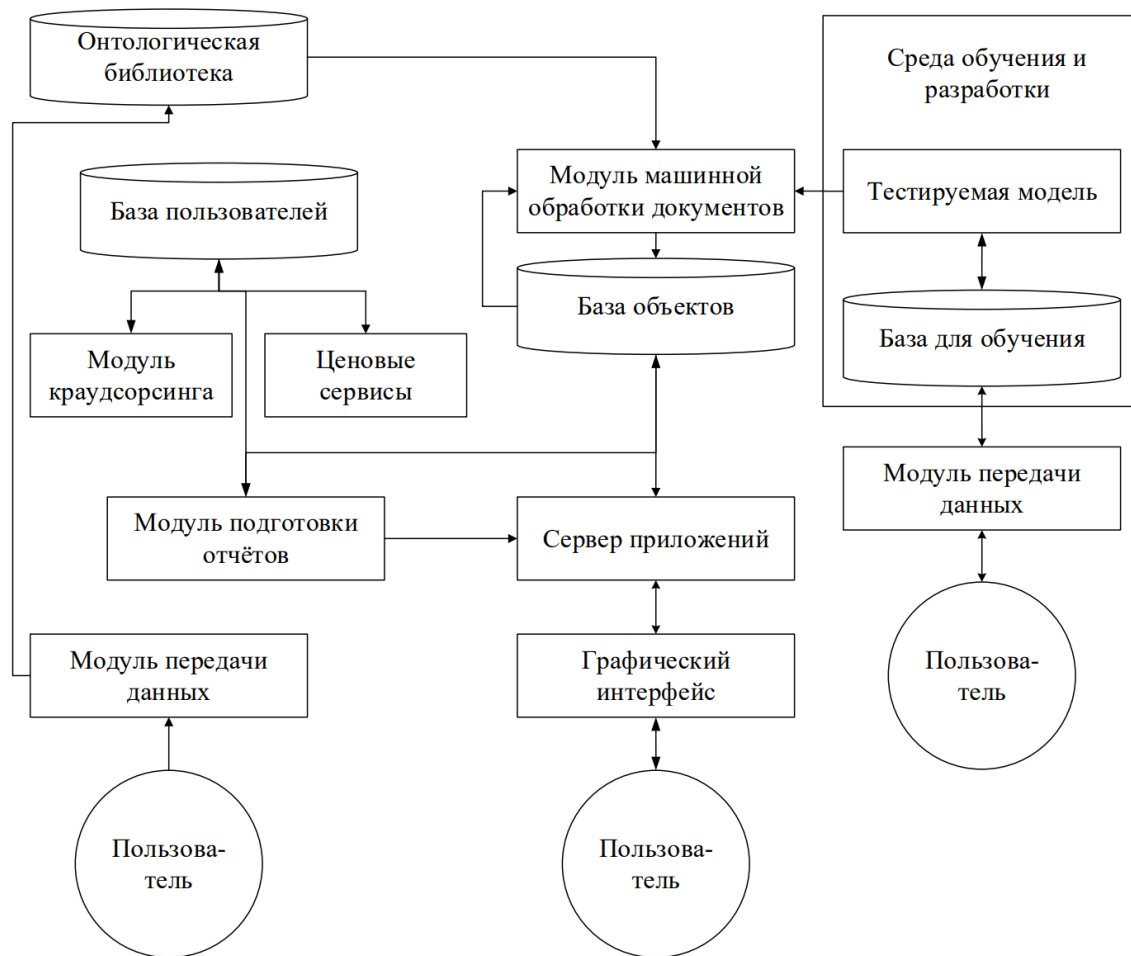


Рисунок 5.2 Схема информационной архитектуры платформенного решения.
Источник: Автор.

Выводы раздела

Выше, в Разделе 1, уже отмечалось, что знания, информация, данные помогают процессу принятия решений. Недостающие сведения можно приобрести по ценам, назначение которых возможно по методам, изложенным в Разделе 2, и даже адаптировано к конкретному случаю по процедуре из Раздела 3. Алгоритмический подход к количеству информации предполагает существование «наилучшей» программы для обработки данных. Если такая программа уже существует, возникает задача её поиска. Если нет – задача выбора подходящего коллектива для её разработки. Оставляя за скобками технический аспект задач (используемые методы программирования, критерии выбора из альтернатив и т.п.), можно указать набор программных сервисов, которые помогают существенно сократить область поиска. Применение

машинного обучения становится императивом для снижения транзакционных издержек при навигации в больших массивах информации (Милкова, 2021). Однако ими не ограничивается предлагаемый набор сервисов. Своё место находят ценовые алгоритмы, средства контроля качества данных, организационные механизмы для обеспечения юридической корректности сделок.

Раздел 5 заканчивает изложение теоретических основ для построения рынка информационных продуктов, и перед завершением исследования остаётся привести примеры продуктивного соединения данных и программного обеспечения для принятия решений, а также методов стоимостной оценки, лежащих в их основе неторгуемых на рынке активов – сути информационных продуктов.

6. Мотивированные решения на макро- и мезоуровнях по данным микроуровня

Как говорилось в Разделе 1, роль знаний состоит, в том числе, в расширении набора доступных ресурсов, как и в снижении ошибок принятия решений. Подраздел 1.1 содержит описание модели, которая иллюстрирует данные эффекты, и здесь, в Разделе 6, в конкретных примерах соединяются представленные выше результаты о рынке данных, стоимости продуктов на этом рынке, разделении стоимости между компонентами информационных продуктов и участии перечисленного в формировании ценности данных. Существенно в данном случае то, что данные выступают сигналами, аккуратная обработка которых с привлечением программных средств способствует информированности лиц, принимающих решения, и, таким образом, действиям, более адекватным сложившейся ситуации.

Представленные примеры относятся к сфере здравоохранения, интересной для данного исследования в двух отношениях. Во-первых, продолжающаяся цифровизация сферы здравоохранения, которая сопровождается накоплением данных, совершенствованием протоколов их передачи и созданием средств для их обработки [223]. Для поддержки цифровизации в этой области разрабатываются национальные стандарты, создан регистр научных разработок и применения цифровых продуктов [248]. Иными словами, цифровизация сферы здравоохранения находится на достаточно высоком уровне, что также подтверждается независимой оценкой [161]. Во-вторых, в составе коллектива исполнителей автор диссертации имеет практический опыт реализации проектов из области здравоохранения именно в части сбора и обработки данных, создания программно-аналитических комплексов. Этот опыт не ограничивается документами о конфиденциальности, и по этой причине может служить иллюстративным материалом для наглядного представления ситуаций, в которых возникают задачи стоимостной оценки данных, программ и распределения стоимости между компонентами информационных продуктов.

Начать следует с простого примера, в котором над приобретёнными данными проводится операция машинной обработки, и результат непосредственно влияет на принимаемое решение. Этому примеру посвящён Подраздел 6.1. В его основе лежит совместная работа с И.Д. Грачёвым [153]. Авторский вклад в цитируемую работу состоит в сборе и подготовке данных, в анализе полученных результатов.

В последующих подразделах приводится пример более сложной ситуации, где данные поступают из разнообразных источников и проходят несколько стадий обработки. Многослойность ситуации позволяет в подробностях обсудить сочетание ролей продавца и покупателя на рынке информационных продуктов, вопросы ценообразования, разделения стоимости между отдельными компонентами и оценки данных. Именно на этих вопросах фокусируется диссертационное исследование, но полнота изложения требует представить пример с разных сторон. В его основе – результаты исследовательского проекта целого коллектива под руководством А.Л. Машковой. Автор диссертации подчёркивает, что математические модели, программная реализация методов, содержательные выводы об анализируемой ситуации получены коллегами по исследовательскому проекту. Авторский вклад в совместную работу состоит в сборе, подготовке исходных данных, их анализе, в проведении патентных исследований, интерпретации содержательных выводов, которые непосредственно повлияли на их использование в компьютерных моделях. Изложение носит иллюстративный характер, и наряду с упоминаемыми результатами отмечается авторство конкретных исследователей. В контексте положений, выносимых на защиту, важен Подраздел 6.5., в котором представлен анализ исследовательского проекта с позиции профессионального оценщика. Стоимость инструмента поддержки принятия решений, который включает программные компоненты и наборы данных, оценена с позиции вытеснения других – альтернативных процедур принятия решений.

6.1. Дифференцированный подход к реализации мероприятий

Пример COVID-19 сегодня является понятным событием, которое мало кого обошло стороной. Быстрое распространение заболевания, с одной стороны, и неопределённость в эффективности протоколов терапии, с другой, остро поставили вопрос об ограничении контактов для снижения темпов роста количества инфицированных. Наблюдения за выявленными случаями и определение круга контактов являются широко известными и зарекомендовавшими себя приёмами. В частности, подробности этих приёмов изложены в работе о локализации и ликвидации очага оспы в Москве в 1959-1960 годах [246]. Однако мировой масштаб событий резко обострил вопрос об эффективности мер по сдерживанию эпидемий. Разнообразие видов реакции, в том числе, по строгости ограничительных мер, масштабах вакцинации и используемым препаратам, предоставили исследователям широкие возможности для сравнения национальных и региональных политик в этой области.

В действительности политика направлена на вмешательство в динамику числа инфицированных и выздоровевших. Стандартным инструментом при исследовании заболеваний являются SIR-модели и их модификации. В контексте данного диссертационного исследования принципиально то, что 1) такие модели реализуются программными средствами, и 2) данные для наполнения моделей о COVID-19 доступны в различных разрезах: города, регионы, страны. И современные средства наблюдения – смартфоны с функцией геолокации, городские камеры наблюдения, транспортные и банковские карты – позволяют собирать массивы детальной информации для количественного анализа тех или иных интервенций на государственном уровне, а высокотехнологичные компании открывают данные о мобильности населения¹⁵.

Говоря о выработке политики, можно указать на то, что снижение количества контактов – очевидное решение для сдерживания эпидемий, которое прямо следует из агентных моделей [52, 88]. В пределе, однако, получается изоляция граждан друг от друга, что на практике приводит к замиранию

¹⁵ Данными о мобильности населения делятся, например, Яндекс, Google, Apple

социальной и экономической жизни, которая лежит за рамками моделей распространения эпидемий. Тем не менее, возможность совместного исследования эпидемии и экономической активности сохраняется. В работах [151, 150, 38] сформулирована и частично решена задача оптимизации карантинных мероприятий на основе взвешенного по паре «ВВП – народосбережение» функционала с использованием реальных данных от сервисов Яндекс, Гугл и др. Для этой цели выстроена комбинированная система дискретных уравнений [150], в которой блок ВВП представлен авторской вероятностной моделью гибридных экономических систем [149], а блок «народосбережение» представлен классической SIR-моделью [150] эпидемий с возможным пересечением блоков.

Теперь, когда дано общее описание ситуации, следует интерпретировать её в контексте ценности данных для принятия информированных решений и показать в ней место для рынка информационных продуктов. Модельная ситуация при этом может выглядеть следующим образом. Пусть мэр города (далее – «принципал») преследует две цели: ограничить распространение заболевания и обеспечить некоторый заранее заданный уровень экономической активности. Обе цели имеют тип ограничения «не более чем», при этом в качестве первого показателя может выступать количество выявленных случаев за единицу времени, а в качестве второго – ВРП (города) за определённый промежуток времени. Пусть у принципала имеется компьютерная программа для исследования совместной динамики эпидемии и ВВП, которая реализует модель, представленную системой уравнений:

$$\frac{dS(t)}{dt} = -\beta(t) \times \frac{S(t) \times I(t)}{N}, \quad (6.1)$$

$$\frac{dI(t)}{dt} = \beta(t) \times \frac{S(t) \times I(t)}{N} - \gamma \times I(t) - \alpha \times I(t), \quad (6.2)$$

$$\frac{dR(t)}{dt} = \gamma \times I(t), \quad (6.3)$$

$$\frac{dD(t)}{dt} = \alpha \times I(t), \quad (6.4)$$

$$\bar{A}_{i+1} = \bar{A}_i + \text{diag}(\bar{\xi}_i \cdot (\beta(t)) \times \bar{A}_i - \langle \bar{\xi}_i, P \rangle \times \bar{A}_i), \quad (6.5)$$

где $S(t)$ – индивидуумы, подверженные заражению;

$I(t)$ – инфицированные и способные к передаче инфекции индивидуумы;

$R(t)$ – выздоровевшие индивидуумы;

$D(t)$ – умершие индивидуумы;

\bar{A}_i – вектор валовых производств (капитализаций) агентов экономической системы;

$\beta(t)$ – некоторая функция от времени, которая зависит от карантинных мероприятий власти и уточняется по накопленным экспериментальным данным (фактически – параметр интенсивности контактов между людьми);

Векторы P и $\bar{\xi}_i$ характеризуют цены в городе и индивидуальную экономическую эффективность агентов. Также для экономической оценки системы здравоохранения известны суточные затраты на лечение одного больного.

Имеющиеся сведения о динамике эпидемии p позволяют принципалу выработать политику, ожидаемая ценность которой составляет $v(p)$. Примером таких сведений для построения априорных суждений о взвешенной паре «ВВП – народосбережение» являются первые данные о динамике эпидемии, сопровождаемой заведомо неоптимальным управлением. В рамках наиболее вероятных диапазонов значений параметров по этим данным возможно построить кривые динамики актуальных больных и соответствующего изменения общего выпуска. Так, в работе [150] это сделано для России и для города Москвы. Полученные результаты на интервале март-июнь 2020 года сравнивались с реализовавшимися и продемонстрировали совпадение по амплитуде и точке первой волны. Этот результат позволяет выполнить отбор оптимальных в смысле взвешенной пары «ВВП – народосбережение» кривых $\beta(t)$ при значениях параметров из диапазонов, обеспечивающих хорошее соответствие. При этом ценность данных $v(p)$ может выражаться следующим образом: «Наилучшее управление $\beta(t) \leq 0,7$ на интервале $t_0 < t < 360$ дней для России позволяет ограничить число больных величиной 600 тыс. чел., число умерших – величиной 30 тыс. чел., и спад ВВП при этом составит менее 2%» [153]. В данной формулировке присутствует не только функция выигрыша $v(p)$ (в

терминах числа умерших и снижения ВВП) от априорных суждений p (общеизвестная динамика в начале эпидемии), но и априорное распределение $\beta^p(t)$ (индекс p обозначает использование управления $\beta(t)$ при наблюдаемом сигнале p). Статистический анализ текущей ситуации допускает точечное оценивание параметров модели. Однако сужение указанных диапазонов значений в соответствии с данными о больных и умерших не считается целесообразным, так как изначально предположение состоит в появлении целого спектра штаммов, которые в каждом конкретном населённом пункте будут давать конкретные особенные значения параметров. Поэтому целесообразно говорить именно о распределениях параметров.

По мере развития эпидемии целенаправленно собирались данные об интенсивности контактов. Свои массивы подготовили и опубликовали крупные ИТ-компании – в частности, Яндекс и Google. Для исследований о динамике COVID-19 и для выработки реакции на будущие вспышки заболеваний (в том числе, сезонные) несомненный интерес представляет то, в какой мере приближения кривых подавления контактов согласованы с индексами самоизоляции Яндекса и индексами мобильности Google. Высокая степень согласованности позволила бы работать с параметризованными приближениями, что существенно технологичней в плане оптимизации карантинных мероприятий, несмотря на оговорки самих авторов этих индексов [153].

На Рисунке 6.1 представлены индекс самоизоляции Яндекс для Москвы и его возможное приближение в промежутке 2 марта 2020 г. – 2 марта 2021 г.

На Рисунке 6.2 представлены модельный и реальный графики числа актуальных больных при использовании в качестве $\beta(t)$ непосредственно нормированного к интервалу $(0,1)$ индекса самоизоляции от Яндекс после поправки, связанной с коррекцией длины волны. Подробности такой техники описаны в [152].

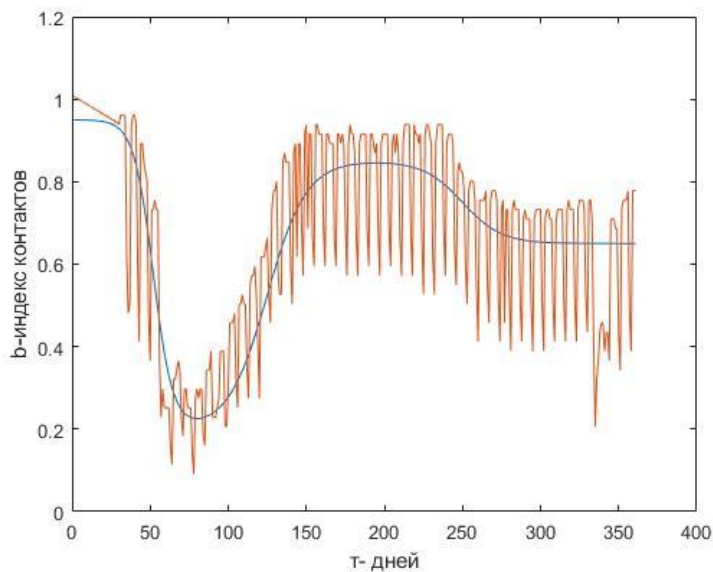


Рисунок 6.1. Индекс самоизоляции Яндекс для Москвы, нормированный на отрезке $[0,1]$ (красная кривая), и вычисленное по модели значение $\beta(t)$. Источник: [153]

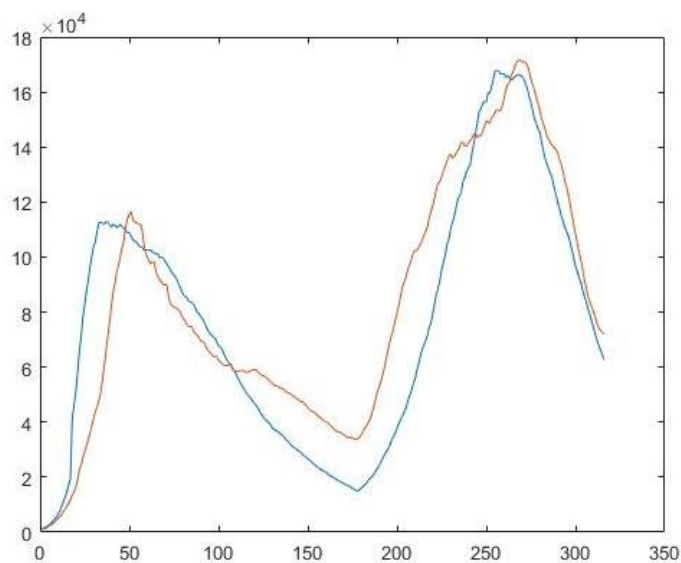


Рисунок 6.2. Реально наблюдаемое (красная кривая) числа инфицированных в Москве и его модельное приближение (синяя кривая) с учётом коррекции. Источник: [153]

Эти наблюдения показывают, что принципал получает в распоряжение текущий индикатор эффективности принимаемой политики – индексы ИТ-компаний. Однако интенсивность контактов можно дифференцировать, стремясь к минимальному подавлению трудовых контактов [38, 151, 150, 152]. Вся необходимую для этого информацию содержат базы данных по индексу

мобильности Google¹⁶. На Рисунке 6.3 представлены нормированные к интервалу (0,2) индексы контактов на работе и дома. Контакты этого типа неизбежно работают в противофазе. Уже этот рисунок позволяет оценить вклад этой пары контактов в управление $\beta(t)$. Фактически, эти данные можно принять за следующий уровень сигнала q , который позволяет перейти к постериорному (наблюдаемому по результатам обработки сигнала) распределению $\beta^q(t)$.

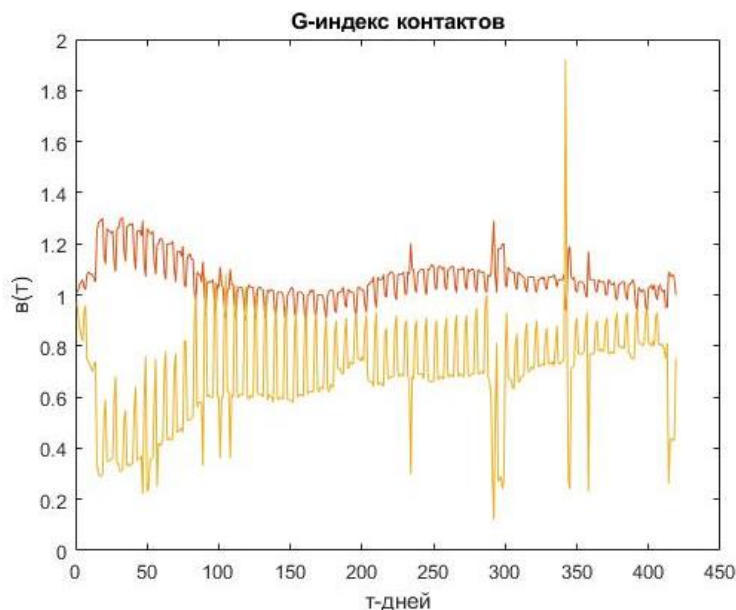


Рисунок 6.3. Значения $\beta(t)$ для трудовых контактов (жёлтая кривая) и домашних (красная кривая) для Москвы. Источник: [153]

На Рисунке 6.4 представлены также нормированные индексы мобильности Google для контактов в розничной торговле, общественном транспорте, на рабочем месте. Непосредственный анализ Рисунка 6.4 подтверждает приемлемое соответствие интервалов между волнами с интервалом между экстремумами транспортных контактов. При этом контакты в точках торговли коррелированы с общественным транспортом.

¹⁶ Данные по Москве, как и по другим территориям, с детализацией по типам контактов представлены на странице Google Community Mobility Reports по адресу <https://www.google.com/covid19/mobility/>

Как сообщается на самом сайте, данные доступны в течение ограниченного времени – до тех пор, пока официальные лица в системе здравоохранения находят их полезными в работе по предотвращению дальнейшего распространения COVID-19.

Обработывая результаты SIR-модели с детализацией контактов по типам, принципал неизбежно приходит к выводу о возможности (как минимум, для параметров эпидемии COVID-19, поскольку использованные данные несут сигнал именно об этом заболевании) обойтись почти без подавления трудовых контактов как второстепенных. Сосредоточиться следует на оптимальном подавлении контактов в общественном транспорте и в розничных точках продаж как экспериментально наиболее значимых для коэффициента интенсивности контактов индивидов с последующим инфицированием $\beta(t)$. Корреляцию между отдельными группами контактов и выявленными случаями инфицирования демонстрирует Рисунок 6.5, где дана динамика коэффициентов $\beta(t)$ в агрегированном виде (на основе индекса самоизоляции Яндекс) в системе с совмещённым началом координат и с декомпозицией по типам контактов (на основе индексов мобильности Гугл). Если в случае первой волны отдельные составляющие индекса мобильности имеют одинаковую динамику и, таким образом, в равной степени предсказывают количество инфицированных, то в случае второй волны контакты в общественном транспорте начинают играть более важную роль в распространении инфекции.

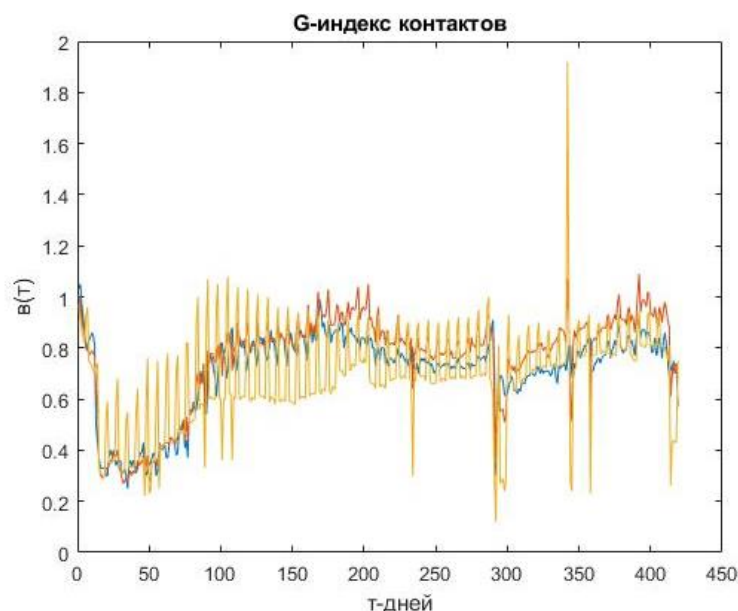


Рисунок 6.4. Значения $\beta(t)$ для трудовых контактов (жёлтая кривая), розничной торговли (синяя кривая) и общественного транспорта (красная кривая) для Москвы. Источник: [153]

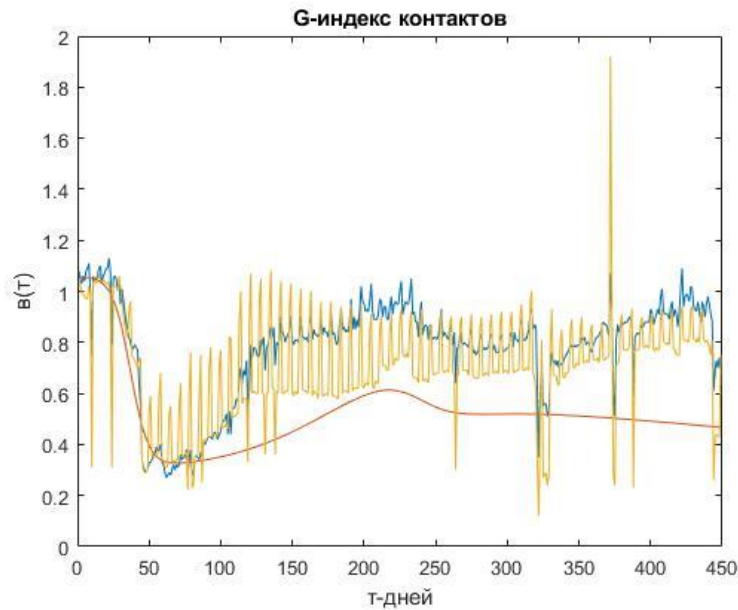


Рисунок 6.5. Значения $\beta(t)$ для трудовых контактов (жёлтая кривая), общественного транспорта (синяя кривая) и агрегированный коэффициент $\beta(t)$ по всем типам контактов (красная кривая) для Москвы. Источник: [153]

Транспорт оказывает значимое влияние на распространение эпидемии в Москве, как показывает анализ, и этот случай не является уникальным. С распространением COVID-19 многие страны стали придерживаться политики ограничительных мер в надежде на замедление темпов роста заболеваемости [125]. И транспорт оказался в фокусе внимания как место большого количества контактов. Логичным следствием анализа, который демонстрирует высокую роль именно общественного транспорта в том, что называют второй волной COVID-19 в Москве, является ограничение мобильности в метро и автобусах с одновременным стимулированием к использованию личных автомобилей. Это наблюдение показывает возможности принципа в выборе действий для реализации карантинных мероприятий. Априорным было бы решение общего карантина, ценность которого $v(p)$ оценена с помощью модели. Дополнительный сигнал q открывает возможности для дифференцированных ограничений — прежде всего, на транспорте — с оценкой ценности принимаемых решений в размере $v(q)$. И плата τ за этот сигнал остаётся оправданной до тех пор, пока $\sum_{q \in Q} v(q)x(q) - v(p) \geq \tau$, где Q — допустимое множество сигналов, $x(q)$ — распределение сигнала q . Причём сохраняется возможность

для уточнения информации о системе путём приобретения дополнительных данных (сигналов). Дело в том, что простое решение об уменьшении пассажиропотока на общественном транспорте хотя и является простым и очевидным, но зарубежный опыт говорит о слабой эффективности таких мер в системе, учитывающей общегородской дорожный трафик [10]. Например, простое ограничение на количество пассажиров оказывается бесполезным в часы пик. Переход на расписание выходного дня имеет ограниченный потенциал. А замещение автобусов на дорогах личными автомобилями чревато негативными последствиями уже в среднесрочной перспективе. Как показывают исследования, при жёстких ограничениях на общественный транспорт существенно возрастает риск увеличения парка частных автомобилей [10]. Опросы демонстрируют, что недоступность общественного транспорта повышает склонность к покупке автомобиля даже среди населения с низким уровнем дохода [10]. Результаты подобных опросов едва ли можно считать далёкими от реальности: в странах, смягчивших эпидемические ограничения, выросли продажи автомобилей, и всё больше граждан предпочитают добираться до места работы на личном транспорте [10, 125]. В перспективе это чревато замедлением общего автомобильного движения в городе, осложнением ситуации с парковкой и разорением частных автотранспортных предприятий даже после снятия эпидемических ограничений. Но эти нюансы не отражаются в представленной выше модели, поэтому экономический эффект (да и эффект для «народосбережения» через снижение аварийности и скорости прибытия бригад экстренной помощи) следует оценивать с помощью другой модели и, соответственно, компьютерной программы.

Анализ транспортной политики более уместно развивать в направлении не усложнения исходной модели, а использования другой в предположении о выборе главного направления – снижения контактов на транспорте. Необходимо отметить, что даже на этом этапе работа с более современными и сложными, чем SIR-модели, инструментами является нецелесообразной из-за больших случайных колебаний заболеваний, возможных систематических ошибок

и предполагаемой сильнейшей неоптимальности управления главным (а по факту – для нового вируса – единственным) параметром. Поэтому следует придерживаться разумности усилий в разработке программ (и моделей), восприимчивых к тем или иным сигналам о состоянии социально-экономической системы.

Разнообразие и роль явлений, которые возникают при вмешательстве в транспортные потоки, демонстрирует ряд исследований. С ростом автомобильного парка связывают повышенные риски ДТП, что влечёт за собой большее количество жертв и аварий, которые в экономических величинах проявляются через дополнительную нагрузку на систему здравоохранения и рынок страховых услуг (через возмещение затрат на ремонт). Оценку роли возросшего парка автомобилей на эти показатели после снятия ограничений можно выполнить сценарным расчётом, но снижение упомянутых величин вследствие уменьшения движения на дорогах уже отмечено в ряде исследований [14, 104]. Ограничения на мобильность граждан проявляются во многих исследованиях через снижение регистрируемых ДТП и их жертв. Другая сторона состоит в том, что возросшее количество автомобилистов может статистически значимо проявиться в увеличении происшествий на дорогах. Уже упомянутая склонность граждан пересесть с общественного транспорта на личный – особенно на фоне сюжетов о толчее в метро и автобусах – свидетельствует о правдоподобности такого сценария. Подобное уже наблюдалось и ярко описано, например, в работах [35, 36], где установлена связь между увеличением жертв ДТП в США из-за страха перед полётами после событий 11 сентября 2001 года. Статистически значимый рост количества ДТП, очищенный от других факторов, в том числе сезонности, наблюдается с октября 2001 года, и за несколько месяцев происшествия на дорогах из-за возросшего трафика унесли большее количество жизней, чем упомянутые террористические атаки.

Не рост частного автопарка, а, напротив, развитие городской сети с учётом разных видов транспорта – автобус, железнодорожный, метро, велосипеды, совместная аренда – в сочетании с динамическим управлением

пассажиропотоком показывает многообещающие результаты [10]. Новые маршруты и дополнительные возможности пересадки вносят некоторое разнообразие в контакты пассажиров друг с другом. Действительно, эпидемии быстро затухают при случайном характере контактов [122], в то время как общественный транспорт до известной степени предопределяет контакты: ежедневно, проходя по маршруту одни и те же районы, автобус собирает частых попутчиков – тех, кто спешит на работу к одному и тому же времени изо дня в день. В части «перемешивания» контактов на транспорте ряд городов может предъявить целую инфраструктуру с единым билетом, видами транспорта и городскими мобильными приложениями, городскими информационными табло. Некоторые стратегии повышения безопасности в общественном транспорте с одновременным фокусом на сохранении парка предусматривают вовлечение не только государства и перевозчиков, но и самих граждан [124]. В этом смысле системное решение проблемы естественным образом влечёт за собой проработку стимулов и мер поддержки на разных уровнях.

Таким образом, пример с карантинными мероприятиями показывает, что первоначальные данные о динамике системы позволяют получить предварительное представление и предпринять управление, которое наверняка окажется неоптимальным. По мере накопления данных появляются возможности для дифференцирования реакции и определения сферы, регулирование которой даст наибольший эффект. Однако детальная проработка регулирования требует дополнительных данных (сигналов) о состоянии системы и новых средств для их обработки, чтобы, в соответствии с алгоритмическим подходом, извлечь нужную информацию. Это позволяет говорить об этапности принятия решений, уточнении информации о системе и уточнении роли тех или иных данных (сигналов) в общем решении. В терминах теории дележей роль данных сводится к вопросу о вычислении вклада отдельных компонентов в общую стоимость, которая реализуется благодаря принятому решению. И это становится поводом рассмотреть вовлечение данных разной природы в анализ

социально-экономических процессов на более сложном примере, который излагается в следующих подразделах.

6.2. Управление на основе данных

Обратимся к более насыщенной данными ситуации, которая также касается принятия решений в области здравоохранения. Эта ситуация примечательна тем, что данные используются на разных этапах разработки системы для поддержки принятия решений: выбора стратегического направления решений, обеспечения методологической чистоты принимаемого подхода, построения функциональных зависимостей и обеспечения работы агентной модели. В последующем примере, однако, заболевание не распространяется контактным способом между индивидами. Дальнейшая речь идёт об анемии, и ситуация принятия решения связана с выработкой мер для её профилактики и лечения.

Анемия, согласно Всемирной организации здравоохранения, является состоянием с концентрацией красных кровяных телец, или гемоглобина, в крови ниже нормы¹⁷. Она проявляется через ухудшение общего самочувствия организма – головокружение, слабость, одышка – и, по некоторым данным, является причиной снижения качества жизни и распознаётся как один из факторов увеличения смертности. Одной из важнейших причин является дефицит железа в организме. Незаменимый микроэлемент, железо, необходим для поддержания структуры и функций клеток в организме, а также является компонентом таких белков, как гемоглобин, который жизненно важен для транспортировки кислорода в ткани организма [45]. Дефицит железа в детстве зачастую является следствием недостаточного питания, обусловленного постоянным отрицательным балансом железа, на который влияют потребление и усвоение железа: недостаточное потребление его с пищей, повышенная потребность в железе в период роста, а также потеря крови из-за кровотечений, малярии и гельминтов [119]. Экспериментальные и эпидемиологические данные

¹⁷ https://www.who.int/health-topics/anaemia#tab=tab_1

свидетельствуют о том, что, хотя нарушения могут возникать на ранних этапах жизни, они могут быть обратимыми благодаря вмешательствам системы здравоохранения, которые направлены на разнообразие рациона продуктами, богатыми микро- и макроэлементами и витаминами, а также на обогащение основных продуктов питания микроэлементами и биодобавками [23].

Такие заболевания, как анемия, которые напрямую не увеличивают показатели смертности, но способствуют общему ухудшению здоровья и потере продуктивных лет, получают меньше внимания [112]. Исследования подчеркивают макроэкономические последствия анемии, такие как повышенная нагрузка на системы здравоохранения [27, 79] и экономические потери из-за снижения когнитивных способностей и производительности [93], что акцентирует важность решения проблемы распространенности анемии. На индивидуальном уровне медицинские исследования показывают, что анемия, часто вызванная дефицитом железа, ослабляет иммунную систему [22, 26], увеличивает риск развития туберкулеза [34, 53] и приводит к худшим результатам у пациентов с черепно-мозговыми травмами [66]. Кроме того, анемия снижает эффективность распространенных методов лечения рака, таких как радиотерапия и химиолучевая терапия, из-за недостаточной оксигенации опухолей [103]. В совокупности эти результаты подчеркивают необходимость приоритизации профилактики и лечения анемии.

Анализ связей между социально-экономическими характеристиками и распространенностью заболеваний часто является сложным из-за междисциплинарного характера темы. Однако такой анализ бесценен для управления общественным здравоохранением. Например, результаты могут быть использованы для оценки количества лет, которые люди живут без хронических заболеваний, на основе их образа жизни [82]. Они также могут помочь выявить ключевые факторы риска в определенных группах населения, позволяя проводить целевые вмешательства для оптимизации результатов политики здравоохранения. С помощью современных инструментов планирования и доступа к надежным наборам данных исследователи могут моделировать потенциальное

воздействие государственных вмешательств на здравоохранение [72]. Большинство исследований в этой области сосредоточены на хронических заболеваниях [172, 234] и пропаганде здорового образа жизни [82].

Национальный регулятор заинтересован в снижении распространённости анемии, стремясь снизить социально-экономические издержки, связанные с заболеванием, и, в то же время, действуя в условиях ограниченных ресурсов. Таким образом, возникает задача выбора таких действий, эффект от которых по соотношению между результатом и приложенными усилиями, будет максимальным. Конечно, обращаясь к проблеме из сегодняшнего дня, можно отметить большой практический опыт, накопленный в данной области. Его копирование, однако, без достаточного количественного анализа, может оказаться опрометчивым. Например, хорошие результаты показывает введение в рацион обогащённых железом продуктов. В то же время, их потребление должно быть ограничено, чтобы не вызвать отравление организма железом. Поэтому, ориентируясь на данную меру, регулятору необходимо представлять, какие продукты до какой степени обогащать железом и как организовать их распространение, чтобы избежать негативных последствий. Как видно, проблема оказывается такой комплексной, что затрагивает не только население, но также сферы производства и распространения продовольствия. Здесь можно усмотреть мезоэкономический уровень проблемы. Во-первых, помимо населения в поле зрения оказываются предприятия и даже отрасли (производство продуктов питания, транспорт, торговля, химическая промышленность в части производства необходимых добавок). Во-вторых, обращение к социально-экономической проблеме такого рода требует взаимодействия государства и частного сектора. В-третьих, неоднородность распространения заболевания по регионам неизбежно требует дифференциации мер воздействия по территориям (субъектам Федерации, районам и т.п.). Дифференциация мер также может быть оправдана по той причине, что разные группы населения по-разному подвержены риску заболевания анемией. Распространённость заболеваний часто связана с определенными группами риска — сегментами

населения с характеристиками, которые повышают их уязвимость к угрозам здоровью. Эти характеристики могут включать биологические факторы (например, возраст, индекс массы тела, артериальное давление), образ жизни (например, физическая неактивность, пищевые привычки) и социально-экономические показатели (например, образование, занятость, общее благосостояние).

Многие исследования и сравнения между странами показывают, что такие факторы, как уровень дохода и системы здравоохранения, влияют на распространенность анемии. Исследования подтверждают, что социально-экономические факторы, включая доход, образование, проживание в сельской местности или в городе, а также качество медицинской помощи, значительно влияют на риск анемии [54, 59, 76, 121]. Определенные социально-экономические группы связаны с основными причинами анемии, такими как недостаточное или несбалансированное питание, хроническое воспаление, потеря крови и снижение мышечной массы тела [55, 75].

В такой ситуации оказывается регулятор: необходимо выработать эффективные меры в условиях известных сведений о проблеме, понимании широкого круга лиц, которые могут быть затронуты разрабатываемой политикой. Системный анализ социально-экономических процессов осложняется обилием связей между разнообразными по своей сути факторами. Пока предмет состоит из одного-двух объектов и модель их взаимодействия ограничивается параметрами одной природы – например, при анализе финансового состояния двух конкурирующих фирм, – зачастую удаётся найти аналитическое решение задачи, которое позволяет получить количественную оценку без необходимости в численном моделировании. Однако при увеличении количества объектов и наблюдении факторов различной природы человеку становится непросто отслеживать согласованное изменение величин в модели. К подобным ситуациям относится анализ последствий экономических решений в социальной сфере или более конкретно – в сфере здравоохранения, которую затрагивает данная работа.

Предположим, государство меняет правила медицинского страхования, корректирует экологические требования или устанавливает специальный налоговый режим, например, для производителей пищевых добавок – и эти нововведения находят отражение в доступности товаров и услуг, в появлении/исчезновении новых предприятий, в изменении уровня благосостояния, в распространении отдельных видов заболеваний. Имея различные начальные условия в регионах по перечисленным выше параметрам, государство столкнётся с различной динамикой отдельных индикаторов в региональном разрезе: не только в части темпов изменения, но даже в направлении тренда. Изменения проявляются на разных субъектах – фирмах и гражданах, в различных сферах – экономике и здравоохранении, различных территориях – регионах и типах населённых пунктов. И эти изменения описываются множеством индикаторов, часть из которых не является экономическими по своей сути. Описание сущностей, индикаторов и связей между ними – уже кропотливая работа, не говоря об исследовании их согласованного изменения во времени. Если учесть, что нормативным регулированием занимаются государственные ведомства, действующие в жёстких рамках процедур взаимодействия, а их сотрудникам приходится одновременно заниматься множеством вопросов, то принятие системных решений кажется практически нереализуемым. Слишком много этапов согласования должны проходить решения, и слишком поверхностным оказывается рассмотрение конкретных вопросов без дополнительной техники для систематизации данных. Специальные инструменты, которые, с одной стороны, сводят воедино различные требования для выработки управляющего воздействия и, с другой стороны, позволяют оценить эффект от предлагаемых мер, могли бы стать частью процедуры принятия системных решений.

Опираясь на обширную литературу по анемии, исследователи разработали математические и компьютерные модели для оценки распространенности анемии. Эти модели могут использоваться государственными учреждениями для планирования вмешательств, направленных на борьбу с этим

заболеванием. Например, мы ссылаемся на агентную модель, которая имитирует производство продуктов питания, потребление и распространенность анемии в регионах России [72]. Насколько известно автору, в настоящее время нет других моделей такого типа. Хотя результаты прогнозирования анемии с использованием социально-экономических или медицинских профилей людей можно найти в литературе [4, 115, 123], не удалось определить какую-либо другую комплексную модель, которая прогнозирует анемию на основе данных о потреблении питания и производстве продуктов питания на региональном уровне страны. Рассматриваемая агентная модель заселяет отдельных агентов, используя официальную российскую демографическую статистику. Авторы включают дополнительные параметры из российского национального исследования, чтобы дифференцировать отдельных агентов на основе потребления продуктов питания. Однако есть возможность улучшить моделирование, введя группы населения, уязвимые к анемии. Поскольку модель работает с социальными и экономическими данными, описание таких групп в социально-экономических терминах было бы естественным расширением.

Агент-ориентированное моделирование (АОМ) – известный инструмент для изучения влияния новых правил в социально-экономических системах. Важно, что агентами в модели являются необязательно люди – это могут быть фирмы, государственные учреждения, транспортные средства и другие объекты [193]. Агентное моделирование позволяет проводить анализ сценариев. Этот анализ ограничен набором параметров модели. Таким образом, лица, принимающие решения, получают инструмент для оценки перспективных преобразований и их влияния на общество.

В течение последних десятилетий агент-ориентированное моделирование (АОМ) становится все более популярным подходом для изучения социальных систем. В то же время, его использование в сфере здравоохранения находится в зачаточном состоянии. Уже разработаны модели в таких разных областях, как эпидемиология, незаконный оборот лекарственных средств и физическая активность. В государственном здравоохранении моделирование

на основе агентов исторически использовалось почти исключительно для моделирования передачи инфекционных заболеваний и борьбы с ними. АОМ являются естественным подходом для моделирования передачи инфекционных заболеваний, потому что взаимодействия между индивидами и влияние окружающей среды часто приводят к популяционным моделям заболеваемости.

Что касается вопросов политики, можно отметить модели, которые исследуют распространение информации и эффективность профилактических мер [9] или систем питания [64]. В [9] описана АОМ, в которой агенты (люди) реагируют на коммуникационные сообщения в соответствии с их собственным отношением. Связь между агентами осуществляется в виде набора «сообщений». Каждое сообщение состоит из заранее определенных вариантов выбора в течение промежутка времени (например, каждые 10 дней), предполагает нацеленность на аудиторию (медиаканал и группа населения) и контент. Агенты отвечают на полученные сообщения в соответствии с их содержанием, которое может принимать следующие значения: описание преимуществ; предоставление информации об эпидемическом статусе; акцентирование норм. Таким образом, сообщение влияет на принятие решения и выбор поведения, изменяя факторы принятия решения.

В [64] авторы пытаются ответить на вопрос, может ли гипотетическая кампания в средствах массовой информации и просвещении по вопросам питания привести к увеличению потребления фруктов и овощей в Нью-Йорке. Предыдущие исследования показали, что такие факторы, как социальный статус, благосостояние, местные пищевые привычки и социальные факторы (например, социальные нормы), напрямую связаны с потреблением фруктов и овощей (например, [37, 92]). Традиционные статистические модели имеют ограниченную способность прогнозировать пищевое поведение, потому что они не в полной мере отражают сложные взаимодействия между людьми и распространяются только на эффекты, вызванные этими взаимодействиями. Для решения комплексных проблем привлекается агент-ориентированная модель (АОМ), которая учитывает индивидуальные факторы (например, возраст,

пол, образование, пищевая среда), а также соседство с другими агентами, их взаимодействие для прогнозирования пищевых пристрастий на замкнутой территории.

АОМ также можно использовать, анализируя поведение потребителя в контексте конкуренции за его внимание. В случае с питанием поведение может заключаться в выборе между здоровой и нездоровой пищей. В [62] авторы предлагают программное обеспечение для моделирования поведения потребителей на конкурентном рынке с несколькими брендами. Программа генерирует виртуальное население, которое воспроизводит реальные решения о покупке.

Итак, в рассматриваемом примере регулятор привлекает агентное моделирование для количественной оценки вариантов политики для борьбы с распространением анемии. Этот инструмент выбирается, поскольку система, о которой идёт речь, состоит из большого количества объектов и связей между ними. Трудно формализуемый, статистический характер связей свидетельствует о возможных препятствиях на пути к аналитическому решению, и в этой связи имитационное моделирование в виде взаимодействующих агентов представляется оправданным. Дополнительным аргументом в пользу выбранного инструментария является возможность изменения уровня анализа – от макроэкономического (в масштабах государства) через мезоэкономический (в масштабах отрасли, региона, группы связанных агентов) к микроэкономическому (на уровне индивида – его структуры доходов и потребления; на уровне предприятия с его разнообразной активностью – в производстве, на финансовом рынке, на рынке труда и т.п.). Следующим этапом на данном пути являются формирование перечня рассматриваемых сценариев и формализация модели. Как показывает краткий обзор работ по агентному моделированию, сценарии влияют на правила взаимодействия агентов. Какой природы могут быть эти правила? Какие действия доступны регулятору? Какие продукты/ решения он может поддержать в производстве? При ответе на эти вопросы следует отталкиваться от практики. Её источником могут служить уже принятые

политики. Ретроспектива, однако, исключает внимание к новым решениям, а также требует внимательного отношения к экспертизе таких решений – в том числе, с учётом национальных и культурных особенностей. Другой путь более соответствует основной теме работы – соединению программ с данными и формированию рынка данных. Он состоит в обращении к патентной библиотеке как к собранию решений – как ретроспективных, так и перспективных – по борьбе с анемией. Патентная библиотека хороша тем, что содержит как методы (способы), так и конкретные препараты (вещества) или приспособления (устройства) для детальной проработки рассматриваемой регулятором политики. Эти вопросы – исследование актуальной реакции на возникновение и распространение анемии, а также формирование контуров агентной модели – подробно рассматриваются в следующих подразделах.

6.3. Патентная аналитика и типология подходов к реагированию на заболевания

Патенты на изобретения представляют собой подробное описание технических решений, связанных с продуктом (устройством, веществом) или способом (методом производства, способом приготовления смесей и растворов, обращения с изделиями и т.п.), в объёме, достаточном для воспроизводства охраняемого результата интеллектуальной деятельности. Поэтому обзор патентной библиотеки позволяет выбрать концептуальные направления, в которых регулятор (в контексте обсуждаемого примера о борьбе с анемией) может посчитать целесообразным поиск конкретных мер для разрабатываемой политики [58]. Однако, при работе с коллекцией документов неизбежно возникает проблема обработки больших данных. На сегодняшний день в мире подаются миллионы заявок на изобретения. Это при том, что прежние документы никуда не исчезают – они продолжают определять уровень техники. Обработать такое количество документов вручную не представляется возможным. Патентная система многих стран является строго регламентированной, и усилия государств по охране интеллектуальной собственности координируются на международном уровне. Благодаря этому современный патент на изобретение (как

и заявка) является структурированным документом, который вписан в иерархическое дерево классификации изобретений. Но, несмотря на усилия по упорядочиванию патентной литературы, поиск изобретений может быть затруднён. В качестве примера можно указать на область лучевой терапии. Методы лечения относятся к разделу «удовлетворение человеческих потребностей», в то время как получение препаратов – к разделу «физика». Говоря в уже принятых терминах, патентная библиотека является строкой символов с высокой колмогоровской сложностью. Обработка таких строк (данных) сегодня осуществляется специализированными информационными системами – каждая отличается внутренними алгоритмами анализа текстов и показателями, рассчитываемыми как для отдельных документов, так и для их наборов. Соответственно, обращение за помощью к той или иной системе позволяет получить более упорядоченную строку, с которой связана меньшая сложность. Операция анализа посредством специализированной системы, настроенной поверх патентной библиотеки, повышает информативность данных. В этом подразделе раскрывается суть этой процедуры в контексте выбора вариантов политики по борьбе с анемией.

Поскольку анемия развивается при недостатке питательных веществ, для поиска изобретений выбраны ключевые слова “anemia”, “iron”, “vitamin C”, “folate”, “folic acid”, “iron fortification”, “iron rich biofortified crop”, “iron supplementation”, “iron rich food product” в названии, реферате и формулах изобретения с самой ранней датой приоритета между 01-01-1930 и 07-07-2020. На основе систематических обзоров питания железом были выбраны эти ключевые слова¹⁸. Самый ранний год приоритета выбирается как индикатор даты, когда была разработана первая идея концепции или изобретения. Другие даты (год подачи заявки, год приоритета, год публикации) не отражают этот момент надлежащим образом, особенно когда речь идет о патентных семействах.

¹⁸ Список ключевых слов согласован с экспертами в области профилактики и лечения анемии с медицинским образованием.

В результате обработки запроса с указанными ключевыми словами патентная информационная система PatSeer выдала 2 868 патентных семей. Тот же запрос в патентных документах выявил 8 033 записи. Поиск был выполнен 7 августа 2020 года.

Рост патентов и патентных заявок на анемию, связанной с недостаточным питанием, имеет общие черты с ростом изобретательской активности в других областях. В литературе сообщается как минимум о двух периодах всплеска. Несмотря на то, что разные области изобретений могут расти с разной скоростью, существуют определенные общие факторы роста, такие как экономические (увеличение инвестиций) и политические (последующее заполнение из-за глобализации) [30]. Поскольку Китай несет ответственность за общий рост изобретений во всем мире, исследователи изучают основные факторы, лежащие в основе этого роста [50, 117]. И Китай также входит в число лидеров в выборке изобретений анемии, вызванной недостаточным питанием.

Отчеты, подготовленные Всемирной организацией интеллектуальной собственности, помогают понять фактические темпы роста изобретений [116, 118]. Ежегодное число новых патентных семейств выросло с примерно 780 тысяч в 2001 году до 1,56 миллиона в 2015 году и 1,98 миллиона в 2017 году. Средний темп роста в мировой патентной базе данных составляет 7,1% в 2001-2015 годах и 9,6% в 2001-2017 годах. Что касается исследуемой выборки, то с 2016 года наблюдается снижение темпов роста. В 2001-2015 годах средний темп роста составил 11,7%, что намного выше значения по всем темам. Но за период 2001-2017 годов средний рост изобретений в области анемии, вызванной недостаточным питанием, снизился до 7,7%. Хотя мировые усилия в разработке изобретений демонстрируют устойчивый рост, следует указать на спад в рассматриваемой теме. В случае общей выборки патентов с ключевым словом «анемия» из Раздела 1 цифры умеренные – средний рост 5,4% в 2001-2015 гг. и 3,3% в 2001-2017 гг. Следуя технологическому соответствию ВОИС МПК, следует сравнить вышеуказанные темпы со средним темпом роста фармацевтических препаратов (А61К) и полимеров макромолекулярной химии

(A61P), т. е. двух ведущих подклассов для общей выборки из Раздела 1. ВОИС сообщает о темпе роста 3% для фармацевтических препаратов и 6,6% для полимеров макромолекулярной химии в 2008-2018 гг. [118]. С другой стороны, общая выборка, использованная в этом исследовании, демонстрирует средний рост 5,5% за тот же период 2008-2018 гг. Таким образом, темпы роста весьма схожи для роста патентных заявок в фармацевтике и макромолекулярной химии, полимерах, о которых сообщает ВОИС.

Патентная система достаточно структурирована, чтобы своевременно отслеживать технологии с использованием формальной классификации без ключевых слов. Всемирная организация интеллектуальной собственности поддерживает усилия по разработке иерархической международной классификации, принятой большинством стран (также существуют параллельно некоторые национальные системы классификации, наиболее известными из которых являются США и Япония). Выборка демонстрирует довольно стабильную динамику изобретений в ведущих подклассах с ростом A61K, A61P, A23L и A23K, в то время как остальные демонстрируют некоторую волновую динамику. A61K и A61P также являются наиболее популярными подклассами для общей выборки патентов, связанных с анемией, хотя A23L и A23K выглядят как специфические для изобретений, связанных с пищевой анемией.

Распределение изобретений по темам, согласно кодам МПК, предполагает ведущую роль в изобретениях следующих медицинских направлений: диагностика, терапия, лекарственные составы. Хотя технологии для пищевой промышленности занимают второе место в рейтинге по номинальным показателям. Но это – обзор изобретений по кодам МПК – даёт лишь общее представление о ситуации. Существует, однако, возможность интерпретации ситуации в терминах колмогоровской сложности. Она вовлекает машинную обработку текстов изобретений – рефератов и подробного описания. Автоматизированный подход интеллектуального анализа текста необходим именно из-за наличия большого массива документов. Алгоритмы информационной системы PatSeer выделяют ключевые слова в описании изобретений с помощью

n-грамм. После нормализации и удаления незначимых слов (предлогов, вводных слов и т. д.) остается 1500 ключевых слов, характеризующих выборку. Анализ ключевых слов методом k-средних дает 3 тематических класса изобретений.

Результаты анализа ключевых слов основаны на терминах, извлеченных из названий, рефератов и заявок на выбранные патенты. Концептуальная структурная карта используется для демонстрации использования результатов исследований, поскольку патенты больше связаны с коммерческими аспектами технологии (Рисунок 6.6). Этот метод кластеризует похожие ключевые слова, преобразуя высокоразмерные данные в легко представимое двумерное пространство, фиксируя максимальную вариацию в данных. Метод анализа множественных соответствий [32] используется для кластеризации изобретений, основанных на топ-100 ключевых слов, в три группы. Этот метод приводит к картине патентов, связанных с химическими и медицинскими соединениями (C2), рецептами блюд и аспектами питания (C1) и изобретениями, связанными с иммунитетом, с анемией в качестве побочного эффекта (C3). Первые два измерения, представленные как Dim1 и Dim2, объяснили почти 83% общей вариации. Иммунные аспекты проявляются как хронические заболевания, включая иммунные, вызывающие анемию. Терапевтические аспекты в выборке также следует отметить. Важно напомнить, что ключевые слова патентов имеют иное происхождение, чем ключевые слова статей. В случае исследовательской публикации авторы предоставляют список ключевых слов, тогда как в случае патентов ключевые слова получаются из заголовка, аннотации и формул с помощью методов текстового анализа. Следует отметить сходства и различия в темах, но надо быть осторожным при сравнении представленных результатов с библиометрическим анализом исследовательских работ по анемии.

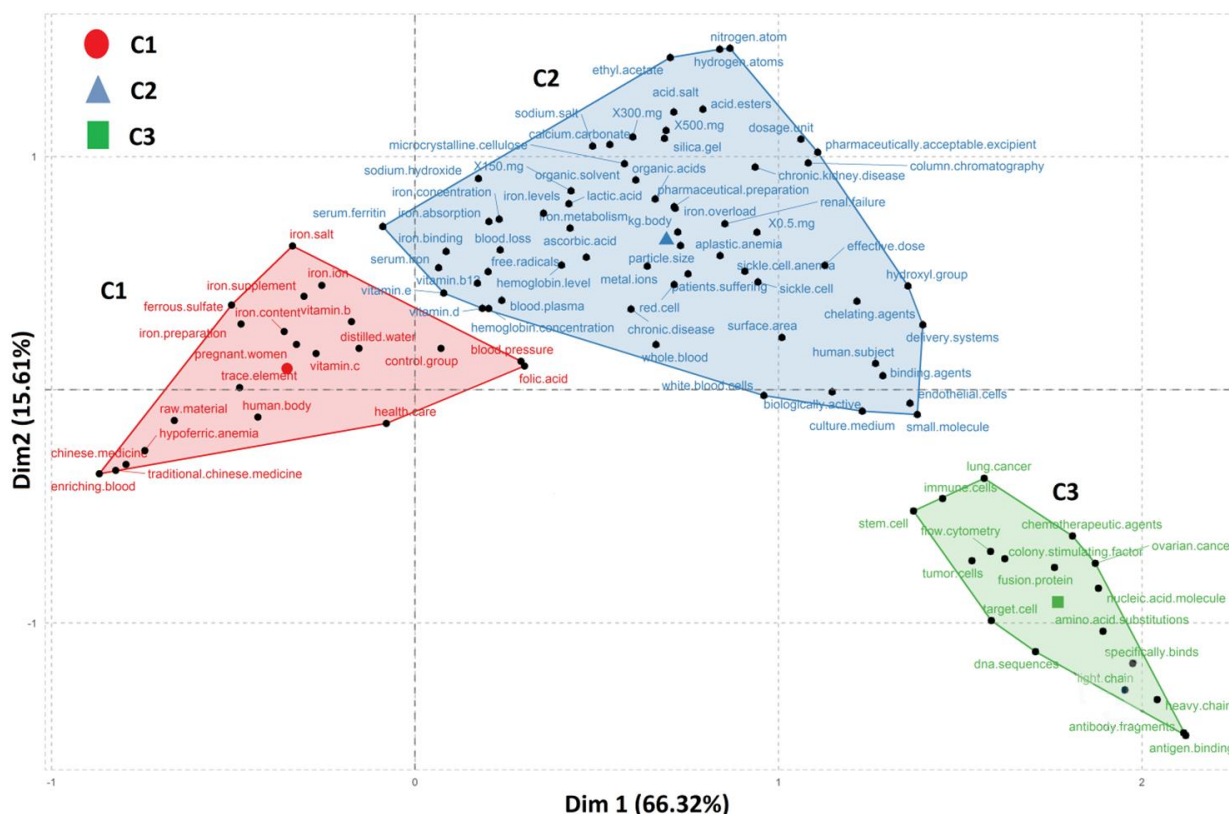


Рисунок 6.6. Схема концепций патентов по теме пищевой анемии. На плоскости представлены три кластера изобретений, а именно С1 (красный), С2 (синий) и С3 (зеленый). Измерения выделены методом главных компонент: Dim1 (объясняет 66,3% вариации в семействах патентов по ключевым словам) и Dim2 (объясняет 15,6% вариации). Источник: [58].

Рассмотрим 10 стран, являющихся основными источниками изобретений из выборки (Таблица 6.1). Здесь приоритетная страна обычно интерпретируется как место происхождения, поскольку необходимо подать заявку на национальный патент до получения правовой защиты за рубежом. Международные корпорации не всегда следуют этому правилу, поскольку у них есть отделы НИОКР по всему миру, и они выбирают приоритетную страну в качестве места первой заявки в соответствии со своей бизнес-стратегией. В Таблице 6.1 перечислены страны в соответствии с количеством патентных семейств, впервые поданных в них, причем лидируют США, а Китай немного отстает. Каждая из этих стран объясняет около трети общей выборки. Каждая другая страна являлась источником во много раз меньшего количества изобретений. Необходимо также обратить внимание на отсутствие Канады и

Австралии среди стран, входящих в первую десятку по изобретениям, связанным с анемией, вызванной недостаточным питанием. Напомним, что эти страны замыкают список из 5 стран, являющихся основными источниками изобретений по ключевому слову «анемия».

MCI (изобретения из нескольких стран) и SCI (изобретения из одной страны) различают изобретения на основе гражданства соавторов. Патентное семейство принадлежит MCI в случае международного сотрудничества изобретателей. Происхождение изобретателя и правопреемника не всегда присутствует в патентном документе. В этом случае не удастся различить изобретение между MCI и SCI, помещая их в категорию NA (не известно). Таким образом, большинство китайских и японских патентов не могут быть классифицированы автоматически. В случае американцев, россиян и немцев изобретения не имеют никакого международного сотрудничества соавторов.

Предварительный анализ небольшой выборки из 59 изобретений привёл к гипотезе о разнице в направленности изобретений между восточными и западными правообладателями. А именно: Восток больше внимания уделяет факторам питания, в то время как Запад уделяет больше внимания диагностике и лекарственным методам лечения. США и Китай являются яркими представителями этих областей: каждая страна имеет более 900 патентных семейств (примерно треть выборки из 2 868 патентных семейств). Рассмотрим их изобретения в контексте подклассов МПК. Таблица 6.2 представляет распределение частот патентных семейств по подклассам для общей выборки в столбце «Выборка» таблицы. Соответствующие столбцы хранят распределение частот патентных семейств для Китая и США. Наблюдается та же разница, отмеченная выше: Китай в среднем патентует относительно меньше в химических или медицинских подклассах и больше в пищевых продуктах. Соединенные Штаты, наоборот, патентуют больше по медицинским или химическим темам.

Таблица 6.1. Список стран, упорядоченных по количеству изобретений по теме анемии, вызванной особенностями питания (патентных семей). Источник: [58]

Юрисдикция приоритета	Всего	Доля	SCI	MCI	NA	Доля MCI
США	960	0,334	615	103	242	0,107
Китай	947	0,330	171	2	774	0,002
Япония	158	0,055	31	0	127	0,000
Россия	144	0,050	113	0	31	0,000
Европейский патент	98	0,034	53	24	21	0,245
Республика Корея	80	0,028	53	3	24	0,038
Великобритания	54	0,019	25	5	24	0,093
Украина	54	0,019	50	0	4	0,000
Германия	51	0,018	40	1	10	0,020
Италия	31	0,011	12	0	19	0,000

Таблица 6.2. Частотное распределение по подклассам МПК. Источник: [58].

Доля патентных семей по кодам патентной классификации. Названия кодов классификации – неофициальный перевод официальной классификации с https://www.wipo.int/classifications/ipc/en/			
	Выборка изобретений		
	Весь мир	Китай	США
C12P - ФЕРМЕНТАЦИЯ ИЛИ ФЕРМЕНТАТИВНЫЕ ПРОЦЕССЫ ДЛЯ СИНТЕЗА ЖЕЛАЕМОГО ХИМИЧЕСКОГО СОЕДИНЕНИЯ ИЛИ КОМПОЗИЦИИ ИЛИ ДЛЯ РАЗДЕЛЕНИЯ ОПТИЧЕСКИХ ИЗОМЕРОВ ИЗ РАЦЕМИЧЕСКОЙ СМЕСИ	2.5%	1.3%	3.4%
C12Q - ИЗМЕРИТЕЛЬНЫЕ ИЛИ ИСПЫТУЕМЫЕ ПРОЦЕССЫ С ИСПОЛЬЗОВАНИЕМ ФЕРМЕНТОВ; НУКЛЕИНОВЫХ КИСЛОТ ИЛИ МИКРООРГАНИЗМОВ	2.8%	0.8%	4.2%
A23K - КОРМА, СПЕЦИАЛЬНО АДАПТИРОВАННЫЕ ДЛЯ ЖИВОТНЫХ; МЕТОДЫ, СПЕЦИАЛЬНО	3.2%	5.1%	2.2%

АДАПТИРОВАННЫЕ ДЛЯ ИХ ПРОИЗВОДСТВА			
G01N - ИССЛЕДОВАНИЕ ИЛИ АНАЛИЗ МАТЕРИАЛОВ ПУТЕМ ОПРЕДЕЛЕНИЯ ИХ ХИМИЧЕСКИХ ИЛИ ФИЗИЧЕСКИХ СВОЙСТВ	7.1%	1.7%	9.8%
C12N - МИКРООРГАНИЗМЫ ИЛИ ФЕРМЕНТЫ; ИХ КОМПОЗИЦИИ	7.8%	1.9%	11.7%
C07D - ГЕТЕРОЦИКЛИЧЕСКИЕ СОЕДИНЕНИЯ	10.0%	2.5%	14.6%
C07K - ПЕПТИДЫ	11.1%	1.5%	17.2%
A23L - ПИЩЕВЫЕ ПРОДУКТЫ ИЛИ БЕЗАЛКОГОЛЬНЫЕ НАПИТКИ; НЕ ОТНЕСЕННЫЕ К ПОДКЛАССАМ	17.6%	34.4%	9.4%
A61P - СПЕЦИФИЧЕСКАЯ ТЕРАПЕВТИЧЕСКАЯ АКТИВНОСТЬ ХИМИЧЕСКИХ СОЕДИНЕНИЙ ИЛИ ЛЕКАРСТВЕННЫХ ПРЕПАРАТОВ	45.3%	49.1%	43.7%
A61K - ПРЕПАРАТЫ ДЛЯ МЕДИЦИНСКОГО, СТОМАТОЛОГИЧЕСКОГО И ТУАЛЕТНОГО ПРИМЕНЕНИЯ	69.3%	53.5%	77.8%

Темпы роста числа изобретений по теме «анемия», связанных с факторами питания, в среднем были выше, чем в мире. Однако в последнее время этот рост замедляется. Исследуемая выборка демонстрирует разницу в тематике изобретений. Их можно разделить на две группы: одна относится к химии и медицине, вторая — к пищевым технологиям. Эта дифференциация проявляется как на уровне стран, так и на уровне правообладателей. Западные страны, предприятия и учреждения ближе к первой группе, тогда как восточные страны и отдельные лица ближе ко второй группе. Можно было бы предположить, что основным фактором являются капиталовложения. Однако присутствие Китая, Японии и Республики Корея во второй группе отвергает эту гипотезу. Эти страны способны мобилизовать ресурсы для масштабных исследований, и они фигурируют в географических рейтингах инноваций наряду с западными странами [65]. Что касается отдельных лиц среди правообладателей, то их активность относится к довольно короткому периоду времени. В данной работе в основном отмечаются эти различия, и для выявления причины различий требуются дополнительные исследования. Ссылки поступают из

различных областей изобретательской деятельности — не только из области исследования анемии, вызванной недостаточным питанием. Анализ цитирования также показывает, что патенты западных стран играют более высокую роль в определении уровня техники в иностранных юрисдикциях. Восточные изобретения больше влияют на уровень техники в национальных юрисдикциях. Возможно, это является результатом разницы тем между Востоком и Западом.

Как и любое другое библиометрическое исследование, этот анализ также имеет определенные ограничения [1]. Например, патент считается эффективным на основе количества ссылок, полученных из других патентов. Может быть, что другие патенты цитируются только для критики. Более того, результаты основаны на выборке патентов, связанных с анемией в результате недостаточного питания, извлеченной из базы данных PatSeer с использованием поискового запроса на основе ключевых слов. Немного более подробный анализ, по сравнению с кодами МПК, показывает, что выборка также содержит изобретения, которые косвенно связаны с анемией, а именно, группа С3. Выполняя анализ программными средствами, можно достаточно быстро выделить три группы изобретений, исключить одну из них (С3) как наименее релевантную и ясно увидеть две стратегии в борьбе с анемией: восполнение дефицита питательных веществ лекарственными средствами или организация сбалансированного питания в качестве профилактической меры. Полученные результаты оказываются полезными в смысле ценности информации по Харкевичу, поскольку проведенный анализ снизил количество возможных маршрутов к желаемому состоянию и, следовательно, повысил вероятность достижения цели.

Руководствуясь целью профилактики анемии и важностью сбалансированного питания в этой области, А.Л. Машкова предложила схему сценарного анализа ситуаций и архитектуру агентной модели [200, 199]. Пусть регулятором принято решение о выработке мер по борьбе с анемией, связанных с профилактикой заболевания. Соответственно, возникает задача обеспечения

населения сбалансированным питанием. Агенты компьютерной модели, их действия и возможное воздействие со стороны регулятора принимают более чёткие очертания. В базовой постановке присутствуют агенты-потребители, агенты-производители и государство, которое устанавливает правила игры. Непосредственное регулирование производителей продовольствия (и норм потребления питательных веществ) косвенно влияют на динамику очагов заболеваний. Так, банкротство производителей того или иного продукта может создать определённый дефицит продовольствия нужного качества (при неспособности конкурентов нарастить сбыт для удовлетворения потребителей в прежней степени) и, как следствие, сместить баланс микроэлементов, поступающих в организм вместе с пищей. В простейшем случае предприятия на продовольственном рынке сталкиваются с конкуренцией за каналы сбыта (место на прилавках ограничено) и действуют в условиях ограничений – организационных (санитарные правила и налоговое законодательство) и естественных (природные закономерности выращивания продукции, её хранения и переработки). Граждане, в свою очередь, получают доход, участвуя на рынке труда, и часть имеющихся в распоряжении средств направляют на потребление продуктов питания. На уровне агентов-индивидов внимание уделяется диете. Последняя зависит от наличия продуктов, а также от моделей потребления: имеют значение кулинарные традиции, религиозные предпочтения, информация о здоровом питании и т.д. Показано, что окружение индивида, как и его доход, в значительной степени влияет на потребление [57]¹⁹. Следовательно, имея данные о ключевых параметрах агентов, можно формализовать динамику спроса.

Модель с дискретным временем реализует в общем случае следующую последовательность действий. Регулятор выпускает нормативные документы. В ответ остальные агенты следуют правилам производства и потребления

¹⁹ В литературе особое место занимает вопрос потребления овощей и фруктов, поскольку их связывают с уменьшением риска некоторых заболеваний. Поэтому исследования паттернов потребления часто проводятся именно на этой категории продуктов – овощах и фруктах. Будем предполагать, однако, что доход и окружение влияют на рацион в целом.

товаров, чтобы максимизировать прибыль и удовлетворение от потребления продуктов. Со стороны предложения основными показателями являются цена, качество продуктов питания и количество производителей. Потребляя продукты питания, агенты-индивиды меняют свои состояния здоровья («больной» или «здоровый») в соответствии с функцией изменения состояния. Регулятор делает ход лишь в начале моделирования, и далее агенты-предприятия взаимодействуют с агентами-индивидами в соответствии с принятыми правилами.

Уточнение деталей в поведении агентов поднимает новый пласт задач, связанных с анализом данных, которые раскрываются в следующем подразделе. Итак, ядро описываемой модели составляют агенты следующих типов: государство, производители продуктов питания, потребители. Отслеживаемыми параметрами являются: общее состояние экономики (в том числе, благосостояние населения), демографические показатели, динамика производства и распространения продуктов питания, правила ведения бизнеса и нормативы пищевой промышленности, установленные государством. Искомая переменная – количество граждан, подверженных высокому риску диагностирования анемии.

Пользователь компьютерной модели имитирует действия агента-государства – устанавливает значения параметров экономики. Базовые из них – налоговые ставки и льготы, ограничение цен на продукты питания и наценок торговых агентов, рекомендуемый рацион в образовательных учреждениях, общий сценарий развития экономики (в обобщённом виде через темп роста). Эти параметры неизбежно влияют на выживаемость производителей, на благосостояние населения, определяя, в конечном итоге, баланс спроса и предложения. Соответствие объёма потребления известным нормам – содержание микроэлементов в пище известно – позволяет оценить вероятность развития анемии. Общая логика взаимодействия различных сущностей внутри модели изображена на Рисунке 6.7.



Рисунок 6.7. Сущности модели и связи между ними. Источник: [200]

Агенты-предприятия оценивают собственную финансовую модель, принимая решение – остаться на рынке или прекратить деятельность. Примером модели для оценки выживаемости предприятий является работа [147]. Модель описывает совместную динамику рынков и развивает теорию самосогласованного поля в экономике. Ключевым её достоинством для агентной модели в контексте динамики производства является эндогенное формирование границы отсечения – чувствует ли фирма себя в среднем лучше или хуже рынка. Соответственно, положение фирмы непосредственно влияет на её привлекательность перед потенциальными кредиторами и, таким образом, на её выживаемость.

Агенты-индивиды получают доход от основной деятельности и трансферты от государства. Исходя из состава семьи, совокупный доход распределяется, в том числе, на потребление продуктов питания, выбирая из нескольких вариантов диет (разной сбалансированности и общей цены). По результатам потребления для агента-индивида вычисляется состояние здоровья [71, 160].

Перечень параметров, охваченных моделью, достаточно обширный и требует большого количества данных для корректной симуляции. Это количество включает открытые наборы (блоки «экономика», «демография») и закрытые (например, блок «цены»). Последние могут приобретаться с «рынка»,

концепция которого описана выше в Разделе 5, но также могут быть получены по результатам специальных обследований. Следует особенно отметить, что данные «как есть» не всегда позволяют корректно отобразить динамику модели. В этой связи они могут потребовать предварительной обработки или уточнения в ходе специального исследования. В конкретном примере об анемии существенной является функция изменения состояния агентов-индивидов, сведения о которой ранее не были известны и появились именно при разработке агентной модели. Это указывает на то, что в проектах подобного масштаба присутствует скрытый слой обработки данных: их очистка, фильтрация и подготовка к использованию в информационном продукте – компьютерной программе, обрабатывающей форматированные наборы данных в соответствии с predetermined процедурой для получения нужного пользователю результата. В окончательном продукте работа этого слоя может оцениваться по затратам разработчиков, поскольку выполняемые процедуры являются исключительно техническими, а применение результата ограничено конкретным приложением. Работе это скрытого слоя посвящён следующий подраздел.

6.4. Согласование данных

Когда сформирована общая рамка модели и разработаны требования к данным, начинается этап их сбора и подготовки. Фокус на производстве и потреблении продуктов питания в контексте регионального распространения анемии позволяет выделить три основных блока данных: о производителе, о продуктах и о населении. Производитель характеризуется объёмом производства и ценой продукции. Некоторое приближение к его портрету даёт совместный анализ отрасли на макро- (по публикациям Росстата) и микроуровнях (по данным СПАРК-Интерфакс). Регион производителя известен достоверно, в то время как продукт – лишь до некоторой степени, в которой это позволяют сделать коды статистики. Для производителя вводится индикатор его присутствия на рынке – является ли предприятие активным или прекратившим деятельность. Индикатор вычисляется по кредитному рейтингу и текущей задолженности предприятия. Последняя доступна из бухгалтерской отчётности (по

данным СПАРК). Кредитный рейтинг определяется на каждом шаге в результате моделирования: действует лучше рынка – займы доступны, хуже рынка – нельзя рассчитывать на кредит, и предприятию грозит банкротство.

Натуральный выпуск отдельных продуктов питания в годовом разрезе доступен из официальной статистики, и анализ потребления микроэлементов требует именно этого измерения. Но выживаемость предприятий – оценка потенциала производителей – требует финансового анализа на микроуровне (физические ограничения пока игнорируются). Подходящим для этого источником информации служит информационная система СПАРК-Интерфакс. И именно её данные – бухгалтерская отчётность – позволяет построить приближённое распределение валового выпуска по предприятиям и по регионам (коды отраслей и специализация предприятий присутствуют в СПАРК). Однако разнесение общего выпуска продукта по отдельным агентам – не единственная трудность. Отчётность в СПАРК фрагментарна, и выборка предприятий – 9 183 юридических лица – содержит как пропущенные данные, так и выбросы. Распределения по выручке и по основным средствам имеют всплески после 1 млрд. руб. (см. Рисунок 6.8), хотя в последнем случае выброс не является таким существенным.

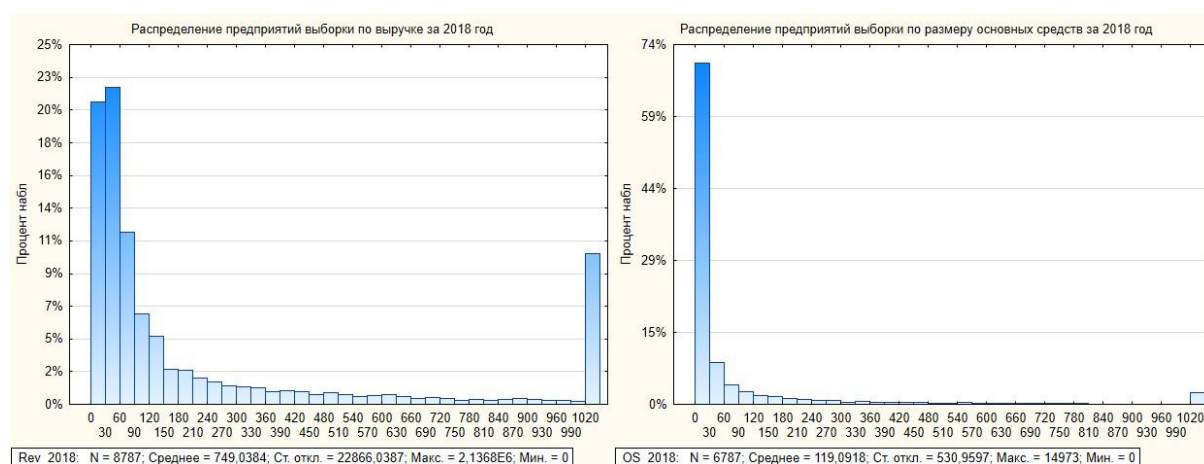


Рисунок 6.8. Распределение предприятий выборки по выручке и основным средствам по состоянию на 2018 год. Горизонтальная ось – значения в млн. руб. Источник: [200]

Совместный анализ данных выявляет выручку, несоразмерную основным средствам, и, соответственно, чрезмерно высокую рентабельность

активов для некоторых предприятий. В Таблице 6.3 предложено описание выборки по показателям выручки в 2017-2018 годах (соответственно, строки Rev_2017 и Rev_2018), по затратам в тот же период (строки Cost_2017 и Cost_2018), по основным средствам и кредиторской задолженности в период 2014-2018 (строки, начинающиеся с OS и KZ), по рентабельности продаж и активов в период 2014-2018 (строки, начинающиеся с ROS и ROA) После исключения явного недоразумения – выручки в триллионы рублей – разброс в данных всё ещё остаётся значительным, и аномалии следует обрабатывать специальными методами, прежде чем записывать в базу данных для корректного их использования в модели. О поиске аномалий см., например, [257]. В общем можно сделать вывод о доступности данных для инициализации агентов-предприятий, хотя они и требуют специальной обработки.

Таблица 6.3. Характеристика выборки по основным переменным. Денежные значения выручки, затрат, основных средств и кредиторской задолженности в млн. руб. Источник: [200]

Переменная	Описательные статистики (СПАРК_Выборка_компаний_20200221_Пищевое)					
	N набл.	Среднее	Медиана	Минимум	Максим.	Ст.откл.
Rev_2017	8477	530,6754	68,00000	0,0	1969911	21406,22
Rev_2018	8787	749,0384	75,00000	0,0	2136761	22866,04
Cost_2017	8477	517,7301	65,00000	0,0	1939732	21077,58
Cost_2018	8787	726,2628	73,00000	0,0	2123620	22721,17
OS_2014	5028	98,2019	10,00000	0,0	14129	377,05
OS_2015	5565	100,8138	10,00000	0,0	13035	396,69
OS_2016	6135	103,5068	9,00000	0,0	12677	414,23
OS_2017	6729	107,0309	8,00000	0,0	13420	459,68
OS_2018	6787	119,0918	9,00000	0,0	14973	530,96
KZ_2014	5908	83,3091	10,00000	-10,0	6344	289,76
KZ_2015	6791	90,3770	11,00000	-10,0	20079	397,20
KZ_2016	7557	94,0589	11,00000	-11,0	16471	410,04
KZ_2017	8264	94,6475	13,00000	0,0	14373	409,96
KZ_2018	8187	105,1871	15,00000	-5,0	12696	450,33
ROS_2014	5796	-0,0984	0,04330	-372,8	2	5,34
ROS_2015	6662	-1,3592	0,04190	-8459,5	1	103,70
ROS_2016	7462	-0,1365	0,03915	-279,9	1	4,54
ROS_2017	8211	-0,9764	0,03610	-4551,0	1	56,58
ROS_2018	8132	-3,5503	0,03420	-24786,3	12	275,62
ROA_2014	5728	0,0499	0,03460	-54,4	57	1,56
ROA_2015	6643	-0,0890	0,04210	-1072,0	21	13,18
ROA_2016	7453	0,0854	0,04130	-43,5	14	0,90
ROA_2017	8213	0,1033	0,04220	-177,6	247	3,65
ROA_2018	8222	0,4721	0,03505	-545,3	1713	25,65

В отличие от случая с предприятиями, данные об агентах-потребителях на момент записи в базу считаются выверенными. Потребление детализируется до уровня домохозяйства и индивида, а подходящим источником в контексте исследования являются наблюдения RLMS. Разделы о регионе проживания, доходах, месте проживания, потребляемых продуктах при гарантии репрезентативности выборки в национальном масштабе (от кураторов RLMS) позволяют построить потребление домохозяйств в региональном разрезе (и количество микроэлементов в рационе). Сообщение об анемии, как показывает обследование RLMS, слышало очень небольшое количество респондентов (Рисунок 6.9). Обследование населения, о котором идёт речь, не является медицинским, и поэтому представленная оценка является нижней границей распространённости заболевания. Специальные биохимические анализы не проводились, и вопрос респонденту буквально звучал в формулировке, представленной на Рисунке 6.9. В целом данные об анемии следуют общемировой закономерности: преобладание заболевания среди женщин и среди старшего поколения (Рисунок 6.10). Дети не охвачены наблюдением индивидов, но они также находятся в группе риска. Обследование также включает вопрос о приёме биологически активных добавок, витаминно-минеральных комплексов. Любопытно, что всего 15% респондентов потребляют добавки и комплексы, тогда как среди тех, кому говорили о наличии анемии, таких уже 31%. Тем не менее, потребление БАД остаётся за рамками модели. Во-первых, ввиду малости этого сегмента на фоне пищевой промышленности. Во-вторых, по причине слишком общей формулировки в опросных листах RLMS. Более ранние волны мониторинга выделяли такие важные в контексте модели элементы, как витамины В12 и А.



Рисунок 6.9. Распределение ответов на вопрос об анемии среди всех респондентов репрезентативной выборки 27-й волны RLMS. Отсутствие ответов или затруднение с ответом исключены при построении диаграммы. Источник: [200]



Рисунок 6.10. Распределение ответов на вопрос об анемии среди всех респондентов репрезентативной выборки 27-й волны RLMS по возрастным группам и полу. Отсутствие ответов или затруднение с ответом исключены при построении диаграммы. Источник: [200]

«Потребитель» является составным в том смысле, что для мониторинга распространения анемии учитываются индивиды, а в экономике участвуют домохозяйства – объединения индивидов. Основные характеристики индивидов – пол, возраст, доход, регион и местность проживания (город/ село), наличие заболевания – генерируются по данным RLMS. Характеристики домохозяйств

– состав, потребление (рацион), доход – совместно на основе данных Росстата и RLMS.

В контексте выработки мер для противодействия заболеванию «подверженность анемии» рассматривается в двух разрезах. Во-первых, анемичное состояние индивида на момент заполнения опросного листа RLMS. Во время собеседования респондент отвечает на вопрос, диагностировал ли врач анемию в течение последних 12 месяцев. Итак, «анемия», о которой сообщает респондент, не выявлена в результате медицинского обследования.

Во-вторых, формализации требуют переходы между анемичным и неанемичным состояниями. На данных RLMS это может быть сделано путём наблюдения одного и того же человека в разные моменты времени в соответствии с волнами мониторинга. Эти переходы могут быть исследованы статистическими методами. Например, с построением логит-моделей, где изменённое состояние (заболевание или выздоровление) является зависимой переменной, в то время как изменения в количестве потребляемого индивидом продукта питания (в килограммах или литрах в неделю) является независимой переменной. В список продуктов питания, подлежащих тестированию, включены те, употребление которых часто связывают с факторами развития или профилактики анемии.

Однако надёжность полного набора анкет за десятки лет наблюдений как источника данных для выявления закономерностей между питанием и анемией может быть поставлена под сомнение. Действительно, поскольку об анемичных и неанемичных состояниях сообщают сами респонденты, а не медицинские организации на основе тщательного обследования, у некоторых из тех, кто отрицает анемию, на самом деле может быть низкий уровень гемоглобина в крови. Можно предположить, что люди с хотя бы одним положительным ответом об анемии в течение многих лет наблюдения более внимательны к проблемам здоровья и следят за показателем гемоглобина в своих анализах. Те же тесты можно выполнить на этой подвыборке, явно предполагая, что

негативные ответы на вопросы об анемии среди этой группы респондентов подтверждаются клиническими исследованиями.

Анализ показывает, что потребление продуктов питания, согласно данным RLMS, частично подтверждает связь анемичного состояния с определённым рационом, хотя и не по широкому списку продуктов. Красное мясо, свежие фрукты, птица и рыба являются факторами профилактики анемии в точном соответствии с теорией.

Пример с функцией изменения состояний показывает, что, хотя RLMS является обширным источником разнородных данных о населении на микроуровне, его использование в вопросах здравоохранения может быть сомнительным из-за самооценок состояния здоровья. А ведь этот источник позволяет сэкономить ресурсы на проведения дополнительных обследований, так необходимых, в том числе, и для выявления групп риска по социально-экономическим факторам. Поэтому научная строгость моделирования требует специальной валидации данных RLMS [18].

Для обоснования корректности использования RLMS необходимо сопоставление с результатами обследования, которое включает как клиническую диагностику анемии, так и социологический профиль наблюдаемых. Доступным обследованием с требуемым уровнем детализации является южноафриканское наблюдение SANHANES (South African National Health and Nutrition Examination Survey). Вообще говоря, доступ к результатам наблюдения также может выступать товаром на рынке данных с соответствующими условиями о режиме доступа. В конкретном сквозном примере об исследовании анемии доступ получен в рамках научного сотрудничества.

Как говорится в отчете об исследовании SANHANES, «уникальность [исследования] заключается в его способности интегрировать результаты личных интервью со стандартизированными физическими осмотрами, диагностическими процедурами и различными лабораторными тестами» [105]. Метод выборки, используемый в южноафриканском исследовании, аналогичен методу RLMS. Южноафриканское исследование дает более подробную картину

распространенности анемии на уровне страны и использует оценки анемии на основе биомаркеров крови. Однако выбранные для сравнения страны – Россия и ЮАР – отличаются по уровню распространённости анемии. В Таблице 6.4 представлены оценки Всемирной организации здравоохранения о распространённости анемии среди возрастных групп для выбранных стран. Также известны оценки о распространённости анемии в России и Южной Африке из Глобального исследования бремени болезней 2021 г. [33]. Согласно этому исследованию, общая оценка для России составляет 15,1% от общей численности населения (95% CI 13,7–16,6%), тогда как для Южной Африки она составляет 25,7% (95% CI 24,4–27,2%) по состоянию на 2021 г.

Таблица 6.4. Распространённость анемии среди групп населения в России и ЮАР. Источник: WHO²⁰

2019, %	Россия	ЮАР
Анемия у детей в возрасте 6-59 месяцев	21,9	44,4
Распространённость анемии у женщин репродуктивного возраста (15-49 лет)	21,1	30,5
Распространённость анемии у небеременных женщин (15-49 лет)	21,1	30,5
Распространённость анемии у беременных женщин (15-49 лет)	23,4	30,8

Методическая корректность требует описания двух баз перед их сопоставлением и последующей проверки значимости общих переменных, характеризующих социально-экономическое положение индивидов, в наборах данных для России и для ЮАР.

Российский продолжающийся мониторинг здоровья населения (RLMS), курируемый Высшей школой экономики, представляет собой исследование, основанное на самооценке здоровья населения. Это продолжающееся исследование началось в 1994 году и структурно состоит из анкеты домохозяйства, анкеты взрослого и анкеты ребенка. Всего анкеты RLMS для домохозяйств и взрослых охватывают 2474 переменных. Поскольку референтное

²⁰ https://www.who.int/data/gho/data/themes/topics/anaemia_in_women_and_children (Accessed 13 January 2025)

исследование SANHANES проводилось в 2012 году, мы используем данные RLMS только за 2012 год.

Анкета домохозяйства определяет место жительства (регион и жилищные условия), состав семьи, имеющееся имущество, участие в сельском хозяйстве и животноводстве, питание (включая источники продуктов, их количество и цены), расходы на товары и услуги, внешнюю помощь в ведении домашнего хозяйства, доход (источники и сумма), сбережения и долги. В индивидуальной анкете взрослого человека фиксируются личные характеристики (пол, возраст, национальность, вероисповедание), семейное положение, идентификатор домохозяйства, участие на рынке труда (статус занятости, рабочее место, отрасль, должность, заработок, обязанности и условия труда), уровень образования, использование ИТ, состояние здоровья и полученная медицинская помощь, образ жизни (употребление психоактивных веществ, спорт, регулярное питание) и общее отношение к другим. В детской анкете фиксируются сведения о воспитании детей и образовании, состоянии здоровья и медицинской помощи. Для этого исследования мы используем данные о взрослых в возрасте 18 лет и старше. Эти участники отбираются случайным образом, и выборка является репрезентативной для России. Краткое описание выборки, состоящей из 13 735 человек, приведено в Таблице 6.5. Большинство участников — женщины, проживают в городских районах, и 43,4% выборки старше 50 лет. RLMS различает четыре типа населенных пунктов: областной центр, город, поселок городского типа и сельская местность. В данном исследовании первые два считаются городскими, а последние два классифицируются как сельские.

Таблица 6.5. Демографические характеристики выборки RLMS для волны наблюдения в 2012 году. Источник: [18]

	Количество	Частота, %
Пол		
Мужской	5663	41,2
Женский	8072	58,8

Место проживания (населённый пункт)		
Город	9345	68,0
Сельская местность	4390	32,0
Возраст		
18-30	3169	23,0
31-40	2401	17,5
41-50	2204	16,0
51-60	2534	18,5
61 и старше	3427	25,0

SANHANES — это референтное исследование с медицинскими осмотрами, используемое для сравнения результатов с RLMS. Оно было проведено в 2012 году Советом по исследованиям в области гуманитарных наук (Human Sciences Research Council) Южно-Африканской Республики. Часть исследования также включает опросный лист, который фиксирует характеристики домохозяйства и описывает каждого человека, включенного в исследование. Домохозяйство характеризуется местом проживания (местоположение и жилищные условия, доступность социальной инфраструктуры), членами семьи, имеющимся имуществом, участием в сельском хозяйстве и животноводстве, расходами на питание, регулярным рационом питания, основным языком общения, уровнем образования и занятости членов семьи, источниками дохода, употреблением психоактивных веществ, доступом к медицинской помощи и предполагаемой стоимостью жизни. Индивидуальные характеристики в первую очередь описываются с точки зрения показателей здоровья: артериальное давление, объективные измерения состава крови (включая гемоглобин и ферритин, которые важны в контексте анемии), курение и употребление алкоголя, пищевые привычки, рост и вес, возраст и раса. В целом, данные опроса, использованные в этом исследовании, охватывают 396 переменных. И для целей агентной модели анализируются данные только для взрослого населения. Краткое резюме выборки, состоящей из 14 794 человек, представлено в Таблице 6.6. Как и в российской выборке, большинство участников — женщины

и живут в городских районах. Однако южноафриканская выборка значительно моложе по сравнению с российской: менее 30% выборки старше 50 лет.

Таблица 6.6. Демографические характеристики выборки SANHANES для 2012 года. Источник: [18]

	Количество	Частота., %
Пол		
Мужской	5978	40,4
Женский	8816	59,6
Место проживания (населённый пункт)		
Город	9869	66,7
Село	4925	33,3
Возраст		
18-30	5188	35,0
31-40	2745	18,6
41-50	2546	17,2
51-60	2154	14,6
61 и старше	2161	14,6

Анализ, выполненный совместно с М.А. Бурилиной, Н. Дукхи, А.Л. Машковой и Р. Сьюпол, позволил получить важные методические выводы [18]. Сравнение анкет показывает значительное совпадение вопросов об образе жизни, питании, доходах и расходах. Важно отметить, что вопросы, связанные с питанием, имуществом и участием в сельскохозяйственной деятельности, фиксируются на уровне домохозяйств, в то время как состояние здоровья регистрируется на индивидуальном уровне. Это сходство подчеркивает методологическую согласованность опросов. Статистический анализ переменных, характеризующих социально-экономическое положение индивидов, выявляет общее влияние образа жизни на здоровье, но также отмечает различия в наборах благ, которые описывают благополучие домохозяйств.

Уровень образования не проявил себя в качестве фактора анемии, как и уровень дохода или место жительства. Общими для российской и

южноафриканской выборки является отрицательная связь анемии с курением и употреблением алкоголя: статистически, лица, которые курят и употребляют алкоголь, имеют более низкий риск анемии. Мы связываем эту связь с общим состоянием здоровья. Лица со слабым здоровьем и подверженностью различным рискам, как правило, отказываются от нездоровых привычек в пользу более здорового образа жизни. Подобная связь ранее наблюдалась у индийских мужчин [59], где курение также отрицательно коррелировало с риском анемии. Однако эта связь не была более подробно рассмотрена в цитируемом исследовании.

Следующим значимым фактором в России является практика обращения за медицинской помощью вместо самолечения. Лица, придерживающиеся этой практики, имеют более низкий риск анемии. Эта связь связана с выявлением заболевания независимо от первоначальной причины посещения больницы. Анемию можно определить с помощью анализа крови, который является стандартным компонентом первичной медицинской помощи в России.

Однако было обнаружено, что прокси переменных благосостояния домохозяйства — таких, как владение имуществом — статистически значимы. В контексте Южной Африки наличие электрической или газовой духовки, стиральной машины и отсутствие участия в выращивании фруктов и овощей показали положительное значение для здоровья. Мы связываем отсутствие анемии со снижением усилий в домашних делах, что указывает на большую сосредоточенность на осознанном выборе питания. В российской выборке наличие персонального компьютера, DVD-плеера и автомобиля продемонстрировало значительное влияние. Исследования связывают рост доходов в России с возросшим спросом на добавки, содержащие витамины и минералы [17]. Мы не видим причин опровергать этот вывод и склонны связывать снижение анемии среди более обеспеченных людей с их доступом к более качественным продуктам и пищевым добавкам.

Поскольку литература по анемии в России ограничена, наши результаты способствуют расширению доступных источников информации для

исследования заболевания в этом регионе. Сходство факторов риска в России с факторами риска в других странах позволяет нам с осторожностью предложить использовать набор данных RLMS, включающий около 2 500 вопросов, в качестве источника для создания гипотез, связанных с анемией. Затем эти гипотезы можно было бы подвергнуть более строгой проверке посредством медицинских обследований. Кроме того, наши результаты могут оказаться ценными для разработки компьютерных моделей распространенности анемии в России, что потенциально приведет к более точным симуляциям и анализам сценариев.

Мы также должны признать ограничения нашего исследования. Во-первых, мы сравнили южноафриканское исследование, которое включало три стадии анемии, с российским исследованием, которое было сосредоточено только на одной стадии. Во-вторых, мы проанализировали устаревшие данные, поскольку данные SANHANES относятся к 2012 году. В-третьих, поскольку наш анализ основан на анкетах, результаты, касающиеся факторов риска, применимы в первую очередь к исследованиям с похожими формулировками вопросов или не зависят от различий в формулировках вопросов. В-четвертых, наше исследование не учитывает региональные и национальные нюансы, которые имеют решающее значение для такой большой и разнообразной страны, как Россия. Это ограничение возникает из-за того, что RLMS является репрезентативным на национальном уровне исследованием и не предоставляет данные на региональном уровне.

В начале сравнительного исследования сохранялась неопределённость относительно корректности использования самооценок для исследования динамики анемии с помощью компьютерного моделирования социально-экономической системы. Своей проверки требовало предположение о том, что опросники могут дать результаты, столь же ценные, как и медицинские измерения. Как выборки, исследованные с медицинской точки зрения (SANHANES), так и выборки, составленные с помощью самооценки (RLMS), демонстрируют сходство в факторах здорового образа жизни, в то время как

социально-экономические факторы демонстрируют различия между выборками. Мы приписываем эти различия культурным особенностям Южной Африки и России. Предыдущие медицинские исследования сообщали о расходящихся выводах относительно схожих факторов, при этом некоторые исследования подчеркивали роль социально-экономических факторов, а другие — нет [60]. Поэтому мы признаем обоснованность самооценки при моделировании распространенности заболеваний, учитывая культурные нюансы моделируемой популяции.

Подраздел о подготовке данных для использования в агентной модели иллюстрирует трудности, с которыми сталкивается покупатель данных на рынке информационных продуктов. Во-первых, наборы могут содержать некорректные или ошибочные данные. Во-вторых, разрозненные данные требуют от разработчика программы исследования вопроса о согласованности предположений, в рамках которых подготовлены приобретённые наборы. В-третьих, размер наборов не позволяет выполнить необходимые проверки вручную. В этой связи привлекаются дополнительные программные средства — как разработанные самостоятельно, так и приобретённые. Существенно в данном случае то, что усилия на подготовку данных оцениваются по затратам — стоимости разработки или приобретённой лицензии на программное обеспечение. В этом отношении существует техническая ясность в их калькуляции, и эти затраты можно отнести к одному из составных блоков итогового информационного продукта — агентной модели. В следующем подразделе предлагается метод стоимостной оценки созданного информационного продукта — агентной модели, функционирование и надёжность выводов которой существенно зависят от состава используемых данных и программных средств, обеспечивающих встраивание данных в модель.

6.5. Цена информационного продукта

Выше подробно описана общая реализация схемы для поддержки принятия решения о предотвращении распространения анемии. Патентные исследования показывают два принципиальных направления усилий: профилактика

с помощью питания и применение лекарственных препаратов. Причём эти направления можно детализировать до конкретных изобретений. Этот результат получен применением патентной информационной системы (программного обеспечения 1 – ПО-1), настроенной поверх массива патентных документов (набор данных 1 – Д-1). С выбором конкретного направления – профилактики – возникает задача обеспечения населения сбалансированным рационом. Комплексная экономическая задача требует внимания к производству и распространению продовольствия, обеспечению его доступности населению. И принимаемые меры могут быть точечными – затрагивать группы производителей, группы поставщиков, группы населения. Взаимность влияний на продовольственном рынке подсказывает инструмент количественного анализа – агент-ориентированное моделирование. Разработанная модель (программное обеспечение 2 – ПО-2) оперирует данными о производстве, распространении и потреблении продуктов питания, данными о населении (набор данных 2 – Д-2). Однако агентная модель нуждается в правилах формирования сценариев и в функциях изменения состояний агентов-индивидов. Сценарии формируются в полуавтоматическом режиме. Из анализа нормативно-правовой базы известны основные документы, имеющие отношение к регулированию рынка продовольствия, и новые сценарии могут представлять из себя проекты правовых актов, которые вносят изменения в существующую базу. Фактически, речь идёт о том, что перечень параметров известен и зафиксирован действующими документами, а сценарии представляют собой значения этих параметров. Блок формирования сценариев дополняется модулем чтения и распознавания электронного документа, поиска значений и записи их в базу данных агентной модели. Этот модуль можно не выделять в качестве самостоятельного, а полностью атрибутировать его из агентной модели.

Функции изменения состояний агентов-индивидов получаются в результате анализа данных о социальном благополучии населения (набор данных 3 – Д-3) статистическими методами с помощью программных пакетов (программное обеспечение 3 – ПО-3). Это же относится и к группам риска. В

разбираемом примере группы риска выявлены в результате исследования Д-3, но другого их среза. Это, однако, потребовало специальной процедуры валидации и привлечения дополнительных данных. В этой связи без ограничения общности можно считать, что задачи выявления и описания групп риска, а также построения функции изменения состояния могут решаться на разных наборах данных, и к ним заведомо применяются различные методы обработки. Поэтому можно считать, что группы риска выявляются по результатам обработки набора данных 4 (Д-4) программным обеспечением 4 (ПО-4).

Как видно, решение в модельной ситуации требует участия разных наборов данных и соответствующих им программных средств обработки. Продуктом – инструментом поддержки принятия решений – в этом случае становится агентная модель ПО-2, данные для которой Д-2 обновляются достаточно регулярно (в соответствии с целями моделирования). Следует обратить внимание на то, что агентная модель ПО-2 является результатом научного проекта, и разработчик – академический институт – не является её непосредственным пользователем. Программа помогает в анализе сценариев на этапе выработки политики регулятором. Он и является основным пользователем и потребителем продукта. При этом ПО-2 и Д-2 являются дополняющими друг друга продуктами. Набор данных Д-2 включает данные официальной статистики, выгрузку из системы СПАРК-интерфакс, реально наблюдаемые цены в различных точках продаж, результаты наблюдений РЛМС. Доступ к ним оценивается по цене лицензий (на момент написания работы данные официальной статистики и результаты РЛМС доступны без оплаты). В то же время, результаты ПО-3, Д-3, ПО-4, Д-4 необходимы для корректной работы агентной модели. Именно в контексте этой тесной связи информационных продуктов следует отвечать на вопрос о стоимости неторгуемого на рынке продукта ПО-2.

Далее рассматриваются две ситуации стоимостной оценки продукта ПО-2: первая предполагает использование продукта наряду с другими правилами принятия решений, вторая – использование продукта взамен других правил. Первая ситуация выглядит следующим образом. Регулятор использует три

процедуры принятия решений: по результатам экспертной оценки, по результатам информирующего моделирования, по результатам моделирования с привлечением продукта ПО-2. Каждая процедура опирается на данные, сформированные наборами Д-2 (общие сведения об объекте) и Д-4 (группы риска). Данные Д-3 являются специфичными, могут обрабатываться только продуктом ПО-2 и совершенно бесполезны для процедур принятия решений, не предполагающих использование ПО-2. Ожидаемый эффект от решения, основанного на экспертных оценках, принимается за единицу. Применение модельных расчётов позволяет уточнить видение ситуации и улучшить эффект на 10%. Это можно интерпретировать так: результатом применения вычислительных процедур является получение объектов меньшей колмогоровской сложности и, соответственно, извлечения большего количества информации – в размере, достаточном для снижения ошибок принятия решений. Пусть по результатам каждой процедуры принятия решений регулятор воздействует на разные (не пересекающиеся) наборы параметров. Например, первая процедура может повлиять на изменение стандартов питания в общественных учреждениях (детских садах, школах и т.п.), вторая процедура – на величины налоговых льгот для производителей, третья процедура – на величину субсидий для обеспечения доступности сбалансированной диеты. Возникает вопрос: какой может быть стоимость продукта ПО-2, разработанного научной организацией, в таком контексте? Ответ на него может быть найден с привлечением теории де-лежей – разнесения стоимости портфеля прав между его компонентами по схеме вычисления Вектора Шепли.

В соответствии с процедурой поиска компонентов Вектора Шепли выделяются группы результатов: данные и программы для их обработки. В соответствии с методикой разбиения стоимости портфеля прав по отдельным компонентам следует обозначить эффекты «синергии» и «каннибализма». Последние можно исключить в случае рассмотрения не баз данных как таковых, а срезов данных. Действительно, на этапе согласования передаваемых наборов можно исключить дублирование полей, поступающих из разных источников.

Сложнее обстоят дела в случае приобретения доступа к информационным системам, поскольку они не предполагают фрагментированный доступ к данным. Например, данные официальной статистики могут пересекаться с данными СПАРК-интерфакс. Так, агрегирование компаний по выручке и отраслям в коммерческой системе не должно отличаться от выпуска отраслей по официальной статистике. В общем случае, конечно, эта ситуация неразрешима, но каждый частный случай может быть исследован на предмет проявления эффекта «каннибализма» с учётом конкретных деталей. В рассматриваемом случае агентной модели случаи каннибализма не обнаруживаются, поскольку на коммерческой основе получен доступ к срезам данных, которые невозможно получить из открытых источников. И эти срезы не пересекаются между собой.

Эффект «синергии» просматривается совершенно чётко в связках «данные – программа для их обработки». Как показано выше, каждая такая связка решает свою задачу, и набор этих задач не пересекается. Соответственно, «каннибализм» между такими связками не проявляет себя. Синергетические эффекты связок с результатами других проектов могут иметь место. Однако, поскольку в рассмотрении находится единственный проект – прогнозирование распространённости анемии в результате принятия тех или иных мер, – следует иметь в виду ненулевую оценку каждой связки «данные – программное обеспечение», предполагая экзогенность такой оценки.

Рассматриваемые объекты имеют одинаковую природу и, соответственно, не отличаются по силе правовой охраны. Все они участвуют в общем проекте, связанном с исследованием распространённости анемии, и группировка по участию в других проектах не прослеживается. Единственное, что учитывается в оценке, как сказано, ненулевые оценки каждой связки «данные – программное обеспечение». К назначению этих оценок и следует перейти, следуя процедуре поиска компонентов Вектора Шепли.

Для дальнейшего введём следующие обозначения:

1 – процедура уточнения сведений о ситуации для принятия информированного решения с использованием агентной модели (ПО-2);

2 – процедура уточнения сведений о ситуации для принятия информированного решения с привлечением модельного инструментария, отличного от ПО-2;

3 – процедура уточнения сведений о ситуации для принятия информированного решения путём обработки экспертных оценок;

4 – набор данных для принятия информированного решения, включающий Д-2 и Д-4;

5 – набор данных для реализации функции изменения состояний агентов-индивидов Д-3.

Следует обратить внимание на то, что каждая из процедур 1-3 опирается на данные 4, и только данные 5 являются специфичными – их использует только процедура 1. Как сказано выше, принимается допущение о том, что моделирование улучшает ожидаемый эффект на 10%. Таким образом, для сочетания {3,4}, принимая равенство {3,4}=1, можно получить {2,4}=1,1 и {1,4}=1,1. Последнее предполагает использование в агентной модели реальных данных для инициализации социально-экономических параметров, описывающих систему, но условные значения в функции изменения состояний. Сочетание {1,4,5}, как предполагается, приводит более адекватному моделированию, и поэтому {1,4,5}=1,2. Вновь обращаясь к фундаментальному положению о том, что данные без процедуры их обработки не имеют ценности, как и процедуры – без данных, можно выписать ненулевые наборы, используемые для вычисления компонентов Вектора Шепли:

Таблица 6.7. Наборы продуктов с ненулевыми оценками ожидаемых эффектов. Источник: Автор.

Число продуктов в наборе	Состав набора	Ожидаемый эффект от набора
2	{1,4}	1
	{2,4}	1,1
	{3,4}	1

3	{1,2,4}	1,1
	{1,3,4}	1,1
	{1,4,5}	1,2
	{2,3,4}	1,1
	{2,4,5}	1,1
	{3,4,5}	1
4	{1,2,3,4}	1,1
	{1,2,4,5}	1,2
	{1,3,4,5}	1,2
	{2,3,4,5}	1,1
5	{1,2,3,4,5}	1,2

Используя выражения для вычисления компонентов Вектора Шепли, можно получить: $\varphi(1) = 0,13$, $\varphi(2) = 0,10$, $\varphi(3) = 0,08$, $\varphi(4) = 0,85$, $\varphi(5) = 0,04$. Оценки показывают, что ПО-2 может оцениваться на 60% больше стоимости процедуры 3 или на 33% больше стоимости процедуры 2.

Вторая ситуация стоимостной оценки продукта ПО-2 предполагает другой контекст. С одной стороны, использование различных процедур для выработки мер, направленных на решение конкретной проблемы, может найти своё отражение в реальности, когда регулятор представляется разными ведомствами. Действительно, за нормирование рациона отвечает Роспотребнадзор, за налоговые льготы – Министерство финансов России, за субсидии – Министерство труда и социальной политики России, Министерство экономического развития России, Пенсионный Фонд России (существуют варианты того, по какой линии выплачивать субсидии населению). И каждый орган вправе использовать те инструменты принятия решений, которые посчитает нужным. Обратимся, однако, к ситуации, когда решения принимаются согласованно, и для этого выбирается одна из трёх конкурирующих процедур. В этом случае решения 1, 2 и 3 являются конкурирующими, и каждое из них принимается с некоторой вероятностью. Аппарат для оценки компонентов Вектора Шепли с вероятностным образованием коалиций из имеющегося набора уже существует, и остаётся лишь назначить вероятности принятия той или иной

процедуры. Предполагается, что процедура 3 (обработка экспертных оценок) уже используется регулятором, и ситуация предполагает варианты: оставить, как есть, выбрать процедуру 2, выбрать процедуру 3. Допуская инерцию в системе принятия решений, принимается вероятность сохранения процедуры 3 на уровне 0,5, переход на процедуру 2 – с вероятностью 0,25, переход на процедуру 3 – с вероятностью 0,25. Следует заметить, что полученная в результате стоимостная оценка процедур справедлива на дату оценки – на момент сохранения неопределённости в выборе той или иной процедуры для поддержки принятия решения.

Контекст выбора одной из процедур для поддержки принятия решений оставляет место для использования платформы рынка данных, речь о которой шла выше. Руководствуясь описанием доступных программных моделей, пользователь выбирает те из них, которые соответствуют стоящей перед ним задаче. Далее, имея набор вариантов (процедуры 1 и 2), пользователь находит данные для наполнения моделей. В описываемом случае предполагается, что процедура 1 нуждается в большем наборе данных (лучший результат достигается с реальной, а не модельной функцией изменения состояний), который также может быть найден посредством упомянутой платформы.

При тех же значениях оценок ненулевых наборов получаются невзвешенные на вероятность оценки $\hat{\varphi}(1) = 0,58$, $\hat{\varphi}(4) = 0,58$ в случае принятия процедуры 1, $\hat{\varphi}(2) = 0,55$, $\hat{\varphi}(4) = 0,55$ в случае принятия процедуры 2 и $\hat{\varphi}(3) = 0,50$, $\hat{\varphi}(4) = 0,50$ в случае принятия процедуры 3. И тогда взвешенные на вероятности исходов оценки принимают вид $\varphi(1) = 0,15$, $\varphi(2) = 0,14$, $\varphi(3) = 0,25$, $\varphi(4) = 0,53$. Стоимость процедур 1 и 2 превышают стоимость процедуры на основе экспертной оценки лишь при значении $p(3) = 0,35$ (в предположении равновероятных исходов для 1 и 2) со значениями $\varphi(1) = 0,19$, $\varphi(2) = 0,18$, $\varphi(3) = 0,18$, $\varphi(4) = 0,54$.

Выводы раздела

Имитационное моделирование, к которому относится и агентное, часто используется в ситуациях принятия решений. На примере разработки

конкретной агентной модели показаны разнообразие и масштаб данных, обрабатываемых моделью и, соответственно, используемых при этом программных средств. Часть из них являются открытыми и бесплатными. Часть – проприетарными, распространяемыми за плату. Одни методы являются более сложными, и их программные реализации приобретаются на рынке. Другие доступны для программной реализации собственными силами и оцениваются по затратам. В результате сочетания всех обозначенных компонентов реализуется агентная модель – неторгуемый на рынке актив, стоимость которого поддается оценке методами, изложенными в данном диссертационном исследовании.

Стоит ещё раз отметить: агентная модель – это результат усилий целого коллектива. Вклад автора в общую работу состоит в сборе, приобретении данных, их фильтрации и передаче далее по цепочке разработки конечного информационного продукта. Иными словами, роль автора диссертационного исследования заключалась в активном участии на рынке данных. Также среди авторского вклада в проект отмечается патентная аналитика (Подраздел 6.3).

Заключение

Дополнительные сведения – важный фактор в принятии информированных решений, и неважно, идёт ли речь о технических или социально-экономических системах. Существенная на сегодняшний день особенность информированных решений отличает ситуацию от уже накопленного опыта объёмом обрабатываемых данных. Эксперт или даже группа экспертов уже не в состоянии обзреть и обработать их самостоятельно, и в этой связи растёт спрос на инструментальные средства обработки данных. Следовательно, вопрос о распространении данных, о формировании рынка данных тесно связан с программным обеспечением. Причём эта связь имеет, как минимум, два аспекта. Во-первых, программные сервисы облегчают навигацию среди массивов данных и поиск нужной информации. Во-вторых, данные и программы для их обработки становятся комплементарными продуктами, что затрудняет их стоимостную оценку. Однако программы касаются лишь технической стороны распространения данных. Возможные препятствия для формирования рынка данных организационного характера связаны с ожиданием упущенной выгоды поставщика в результате появления конкурента, с наличием ясных методов ценообразования, с соблюдением правовых режимов доступа к данным.

Вышесказанное явным образом высвечивает ценность данных для принятия информированных решений, необходимость технических средств для поддержки рынка данных, нетривиальность ценообразования на этом рынке и ограничения, накладываемые правовыми режимами. Внимательный анализ этих пунктов, выявление связей между ними, устранение барьеров для совершения сделок и разработка решений (как технических, так и организационно-экономических), как представляется, способны повысить эффективность реализации национального проекта «Экономика данных» и способствовать росту национальной экономики за счёт 1) более эффективного использования имеющихся ресурсов по результатам анализа данных; 2) расширения рынка данных и, соответственно, углубления разделения труда.

Сегодня рынок данных можно охарактеризовать двумя большими направлениями. Первое – данные как есть. Поставщики предлагают именно наборы данных: размеченные датасеты для машинного обучения, сведения о поведении пользователей/ потребителей/ клиентов, список транзакций, наблюдения цен и т.п. Второе – лицензия на использование информационной системы, которая обрабатывает большие данные. Такие системы могут обрабатывать не только общедоступные данные. Однако поставщиков подобных услуг (доступа к информационным системам) следует отнести к участникам рынка данных не только по той причине, что они приобретают данные для наполнения систем. Дело в том, что системы имеют встроенный аналитический функционал, который предлагает лицензиату (покупателю лицензии, пользователю) доступ к уникальным индикаторам и показателям, рассчитанным правообладателем. Есть и чисто формальная причина для того, чтобы считать лицензиара (продавца лицензии на доступ к аналитической информационной системе) поставщиком на рынке данных. В соответствии с российским законодательством об интеллектуальной собственности, базы данных охраняются как составные произведения – аналогично словарям, сборникам, энциклопедиям. Такая юридическая позиция находится в общей линии с общими рекомендациями ВОИС и общемировой практикой, поэтому сказанное можно не ограничивать национальным уровнем. Но, если лицензиар собирает данные из различных источников (в том числе, и на коммерческой основе), упорядочивает их в соответствии со своими целями использования, устанавливает дополнительные связи между данными, – одного этого достаточно для идентификации имеющегося у него набора в качестве самостоятельного произведения, а в случае России – ещё и регистрации в качестве базы данных²¹.

²¹ Имеется в виду правовой объект «база данных». Он относится к объектам авторского права, правовая охрана которых исчисляется с момента создания, и государственная регистрация таких объектов не является обязательной. Однако в России программы для ЭВМ и базы данных принимаются к регистрации Федеральной службой по интеллектуальной собственности. Эта регистрация не является обязательной именно с точки зрения правовой охраны, однако, получение соответствующего документа – свидетельства о регистрации – может оказаться удобным в ряде случаев. Например, при постановке результата на баланс в качестве нематериального актива свидетельство о регистрации рассматривается в качестве правоустанавливающего документа.

Оба направления отталкиваются в своём ценообразовании скорее от эмпирических наблюдений. Если поставщик предлагает данные в наборе с программным сервисом для их обработки, оцениваются затраты на приобретение данных, разработку сервисов, поддержку ИКТ-инфраструктуры и текущие расходы. Общие расходы делятся на потенциальный объём рынка, и удельные затраты, увеличенные на премию поставщика, составляют цену лицензии. Если поставщик предлагает лишь наборы данных, или срезы своей базы данных, то цена может определяться по аналогии (в соответствии с ценой на близкие по характеристикам наборы) или в результате переговоров (торга). Формирование рынка для аналитических информационных систем – процесс, аналогичный тем, что протекают в случае программных продуктов. Он включает целенаправленное воздействие на потенциальную аудиторию и систематическое информирование о продукте. Однако распространение наборов данных имеет свою специфику.

Поставщик данных может предлагать различные выгрузки из доступной ему базы. Эти выгрузки могут различаться как по составу полей (характеристикам объектов), так и по составу строк (наблюдений, каждое из которых описывает тот или иной объект). Различные комбинации полей и строк базы данных далее называется «срезом данных». Накапливаемые сегодня данные, как уже сказано, являются обширными и самыми разнообразными. Потенциальные покупатели могут просто не представлять себе срезы, которые можно приобрести. В случае B2B, а поставка данных предполагается именно на таком рынке, можно получить лишь некоторое представление о спросе – например, через торговые площадки, на которых размещаются сведения о закупках. По признанию самих операторов данных, они обладают такими наборами, применение которых не понимают в полной мере. Часть данных собирается, но остаётся невостребованной. Не только на рынке, участники которого, как сказано выше, могут и не знать о существовании таких наборов, но и самим оператором. В этой связи актуален вопрос о поиске потребителей данных – срезов, которые прежде не предлагались на рынке, – и о назначении цен на такие

товары. Причём ценообразование в таком случае должно быть дифференцированным – не только по срезам, но и по потребителям. В соответствии с этим положением, которое имеет под собой теоретическое обоснование, разрабатываются алгоритмические процедуры продажи информации.

Следуя цели вовлечения в оборот большего количества информационных продуктов (данных и программ), в ходе выполнения диссертационного исследования получены следующие результаты.

1. Построена и исследована экономико-математическая модель для анализа совместной динамики отраслей экономики с выделенной отраслью производства знаний. Известные работы представляют выгоду от вложений в науку и технологии через изменение коэффициентов производственной функции. В предложенной модели отрасли могут детализироваться до отдельной фирмы, и их решения состоят в выборе размера инвестиций в производство знаний. Последние, в свою очередь, влияют на доступ к ресурсам (как отдельной отрасли или фирмы, так и экономики в целом), а также на снижение ошибок принятия решений.

2. В экономический дискурс введена колмогоровская сложность как объективная мера для различения потребителей и поставщиков на рынке информационных продуктов. Ценовая дискриминация на рынке информационных продуктов является известным теоретически обоснованным результатом. Для различения потребителей – на практике и в теоретических моделях – опираются на типы. При этом возникает проблема вычисления данного типа. Колмогоровская сложность даёт объективную меру для 1) дифференциации программ обработки данных и 2) дифференциации потребителей на рынке данных по тому, доступом к какой программе они обладают. В случае отсутствия ограничения на ресурсы, такая типологизация является однозначной в том смысле, что существует «наилучшая» программа для обработки данных. В случае колмогоровской сложности с ограничением на ресурсы оказывается, что пользователи разных программ могут быть отнесены к одному типу, если они обрабатывают строку одинаковым образом. Типологизация по колмогоровской

сложности имеет потенциал для применения к оценке стоимости программного обеспечения в привязке к аналогу с известной стоимостью и разнице в индивидуальных ценах.

3. Предложена модель рынка данных, поддерживаемая цифровой платформой с онтологическим представлением цифровых объектов, обеспечивающей соединение спроса и предложения с использованием ценовых алгоритмов. Полных аналогов платформ, которые позволили бы соединить спрос и предложение на рынке данных, не обнаружено. Некоторым приближением могут служить электронные библиотеки – публикаций, программного обеспечения, поскольку они помогают распространять информационные продукты, поддерживают их повторное использование. Однако при соединении спроса и предложения используются рекомендательные сервисы, основанные на профайлинге пользователей (анализе профиля пользователя по индивидуальным характеристикам и истории активности на платформе). Также электронные библиотеки не предлагают сервисы для оптимизации цены сделки. При анализе предложенной платформы исследованы подходы к построению сервисов для хранения и обработки цифровых объектов, их распространения, учитывающие правовые ограничения и технические возможности.

4. Путём обобщения известной процедуры назначения цен за информацию на случай повторяющихся взаимодействий развит математический аппарат — основа сервиса ценообразования на цифровой платформе. Существует протокол продажи информации с использованием информирующих сигналов для потребителей разных типов. Теоретически обосновано, что при однократном взаимодействии этот протокол, который полагается на ценовую дискриминацию, позволяет полностью изъять потребительский излишек при соблюдении условий индивидуальной рациональности и совместимости со стимулами, исключая, таким образом, возможность арбитража со стороны потребителя. На рынке данных возможны ситуации повторяющихся покупок с некоторой периодичностью, что позволяет потребителю актуализировать доступный ему аналитический сервис. В этой связи предложена модификация

условий индивидуальной рациональности и совместимости со стимулами в протоколе передачи информации для расширения на случай повторяющихся взаимодействий.

5. Формализована математическая процедура оптимизации двухкомпонентных цен на рынке данных с использованием алгоритма достижения Парето-оптимальных состояний в задаче разделения ресурсов. Цена, назначаемая протоколом передачи информации из предыдущего пункта, может выступать параметром функции, описывающей двухкомпонентные цены (тарифы). Такие цены, с одной стороны, теоретически оправданы в случае возрастающей отдачи, к которому относятся высокие технологии, и, с другой стороны, позволяют вычислить равновесное для рынка состояние. Известный алгоритм распределения ресурсов для достижения Парето-оптимального состояния, который автором уже применялся в контексте лицензионной торговли, адаптирован к двухкомпонентным ценам.

6. Предложена процедура стоимостной оценки неторгуемого на рынке нематериального актива, основанная на аппарате теории игр. Потребитель на рынке данных обладает средствами их обработки – той или иной программой, которая характеризуется колмогоровской сложностью. Цены на данные можно наблюдать на рынке. Программа является уникальным активом, стоимость которого заранее не известна, но может быть полезна для ряда задач: использования для целей амортизации, вклада в уставный капитал вновь создаваемого общества, использования в качестве залога. Стоимость такой программы устанавливается по известной цене данных с поправкой на синергетические эффекты по известной процедуре расчёта компонентов Вектора Шепли.

7. Теоретические положения о построении рынка данных приложены к примеру разработки многокомпонентной системы принятия решений, которая предполагает несколько уровней обработки данных. Модельная ситуация о выработке политики в области здравоохранения предполагает использование разнообразного программного инструментария и неоднородных данных. И

инструменты, и данные характеризуются различными условиями доступа: одни являются открытыми и бесплатными, другие оцениваются по затратам, третьи – по рыночным ценам. Все эти компоненты важны для корректности работы верхнеуровневой агентной модели, которая представляется неторгуемым на рынке активом. Стоимость компьютерной программы, реализующей модель, оценивается по описанной выше процедуре.

Полученные результаты, объединённые цифровой платформой для поддержки рынка данных, приобретают особенное значение в контексте проблемы обеспечения технологического суверенитета. Последняя широко обсуждается в последние годы, и бизнес ищет возможности для установления долгосрочного, предсказуемого сотрудничества [138]. В качестве возможного инструмента такого сотрудничества предлагается СДС-консорциум (аббревиатура от «сети добавленной стоимости») [231]. Для координации участников такого консорциума предлагается использовать инструмент поддержки. В процитированной работе он не детализируется, однако, отмечается роль цифровых платформ в создании и функционировании сетей. Отмечается также, что «при высоком уровне доверия платформа консорциума становится коллаборативной, обеспечивая сотрудничество между его членами» [231]. Представляется, что описанная в исследовании концепция цифровой платформы с набором инструментов может претендовать на роль технического средства для поддержки СДС-консорциума. При этом обеспечиваются несколько типов доверия. Во-первых, «рациональное (rational) доверие, основанное на расчете ожидаемых выгод и рисков», поскольку инструменты обеспечивают согласование со стимулами и учитывают стремление участников к выгоде. Во-вторых, «процедурное (procedural) доверие, возникающее вследствие воспринимаемой легитимности, прозрачности и справедливости процедур взаимодействия, которые позволяют прогнозировать поведение других агентов» в силу строгости и неманипулируемости лежащего в основе аппарата.

Поскольку СДС-консорциумы включают «научные, проектные, производственные, сбытовые и посреднические организации» [231], оператор

платформы для рынка данных может не только предоставлять инструмент для работы консорциума, но и являться его участником. Вопрос об операторе платформы для поддержки рынка данных следует отнести к дальнейшему развитию изложенных в диссертации результатов. Встраивание цифровой платформы в рынок становится поводом для разделения власти и стоимости в цепочках добавленной стоимости [158]. Как отмечается в процитированной работе, «можно заключить, что, по крайней мере, некоторые цифровые платформы способны не только снижать одни риски, но и генерировать другие, оказывая влияние на распределение добавленной стоимости». Этот вопрос – перераспределение рисков и связанных с ними платежей в результате появления в цепочке добавленной стоимости нового участника – заслуживает специального рассмотрения за рамками настоящего диссертационного исследования, которое фокусируется на механизмах работы платформы и их экономического обоснования.

Список сокращений и условных обозначений

API	– application programming interface, англ. «интерфейс программирования приложения», интерфейс взаимодействия программ между собой
CI	– confidence interval
CRM	– customer relationship management, англ. «система управления взаимодействия с клиентами»
DRP	– demand-revealing process, англ. «процедура выявления спроса»
NFT	– non-fungible token, англ. «невзаимозаменяемый токен», уникальный токен.
RLMS	– The Russia Longitudinal Monitoring Survey, Российский мониторинг экономического положения и здоровья населения НИУ ВШЭ
SaaS	– software as a service, англ. «программное обеспечение как услуга», программное обеспечение по требованию
SANHANES	– South African National Health and Nutrition Examination Survey
SIR-модели	– динамические модели инфицирования, название которых образовано первыми буквами S
AOM	– одновременно агент-ориентированное моделирование и агент-ориентированная модель в зависимости от контекста
БАД	– биологически активная добавка
B2B	– business-to-business, англ. «бизнес для бизнеса», взаимодействие коммерческих юридических лиц
B2C	– business-to-consumer, англ. «бизнес для потребителя», взаимодействие коммерческих юридических лиц с конечными потребителями
ВОИС	– Всемирная организация интеллектуальной собственности
ВЦКП	– Вычислительный центр коллективного пользования

ГИС	– государственная информационная система
ЗПР	– задача принятия решения
ИКТ	– информационно-коммуникационные технологии
ИС	– интеллектуальная собственность
ИТ	– информационные технологии
КСИРИН	– корпоративная сеть изобретательства, рационализации, инноваций
ЛПР	– лицо принимающее решение
МПК	– международная патентная классификация
НИОКР	– научно-исследовательские и опытно-конструкторские работы
НМА	– нематериальные активы
ОГАС	– Общегосударственная автоматизированная система учёта и обработки информации
P2P	– peer-to-peer, англ. «равный к равному», здесь понимается в значении взаимодействия на уровне потребителей без посредничества третьих лиц
СДС	– сети добавленной стоимости
ССИРИН	– социальная сеть изобретательства, рационализации, инноваций
ЦОД	– центр обработки данных
ЭВМ	– электронная вычислительная машина
ЭДО	– электронный документооборот

Литература

1. Abdi. 2007. Multiple correspondence analysis. *Encycl. Meas. Stat.* 2 (4): 651.
2. Adams R., McCormick K. 1993. The Traditional Distinction between Public and Private Goods Needs to Be Expanded, not Abandoned. *Journal of Theoretical Politics* 5 (1): 109-116.
3. Aghion P., Howitt P. A Model of Growth Through Creative Destruction // *Econometrica*, Vol. 60, No. 2 (Mar., 1992), pp. 323-351
4. Aguree S, Owora A, Hawkins M, Reddy MB. Iron deficiency and iron deficiency anemia in women with and without obesity: NHANES 2001–2006. *Nutrients*. 2023;15(10):2272. DOI: 10.3390/nu15102272
5. Akerlof G. The Market for "Lemons": Quality Uncertainty and the Market Mechanism // *The Quarterly Journal of Economics*, v.84, August 1970, p.488-500
6. Arrow K. J. The Economic Implications of Learning by Doing // *The Review of Economic Studies*, Vol. 29, No. 3 (Jun., 1962), pp. 155-173
7. Babaioff M., Kleinberg R., Leme R. P (2012): "Optimal Mechanisms for Selling Information", in *Proceedings of the 13th ACM Conference on Electronic Commerce, EC '12*, pp. 92-109.
8. Baniak A., Dubina I. Innovation analysis and game theory: A review // *Innovation: Management, Policy & Practice*. - 2012. - Vol. 14. - Issue 2. - pp. 178-191.
9. Barbrook-Johnson P., Badham J., Gilbert N. Uses of Agent-Based Modeling for Health Communication: the TELL ME Case Study // *Health Communication*. - 2016. - Vol.32. - №8 - p. 939-944.
10. Basu R., Ferreira J. Sustainable mobility in auto-dominated Metro Boston: Challenges and opportunities post-COVID-19 // *Transport Policy*. - 2021 - Vol. 103. - p. 197-210.
11. Beebe J. 1977. Institutional Structure and Program Choice in Television Markets. *Quarterly Journal of Economics* 91: 15-37.

12. Beliga S., Martincic-Ipsic S., Meštrović A. (2015). An overview of graph-based keyword extraction methods and approaches. *Journal of Information and Organizational Sciences*, no. 39 (1).
13. Blei, D., Ng, A., Jordan, M. Latent Dirichlet allocation // *Journal of Machine Learning Research*. 2003. № 3, pp. 993-1022.
14. Brodeur A., Cook N., Wright T. On the effects of COVID-19 safer-at-home policies on social distancing, car crashes and pollution // *Journal of environmental economics and management*. - 2021 - Vol. 106. - p. 102427.
15. Brynjolfsson E., Hitt L. M. Beyond computation: Information technology, organizational transformation and business performance // *Journal of Economic perspectives*. – 2000. – T. 14. – №. 4. – C. 23-48.
16. Bui H., Venkatesh S., Kieronska D. Learning other agents preferences in multi-agent negotiation using the Bayesian classifier. - *Int. J. Coop. Inform. Syst.*, 1999, v. 8, 4, p. 275-294.
17. Burggraf C, Teuber R, Brosig S, Glauben T. Economic growth and the demand for dietary quality: evidence from Russia during transition. *Econ Hum Biology*. 2015;19:184–203. DOI: 10.1016/j.ehb.2015.08.002
18. Burilina M.A., Dukhi N., Mashkova A.L., Nevolin I.V., Sewpaul R. Factors that influence anemia prevalence: a comparative study of datasets from Russia and South Africa // *BMC Public Health*. – 2025. – Vol. 25, No. 1. – P. 1170. – DOI 10.1186/s12889-025-22363-6.
19. Chamberlin E. 1956. *The Theory of Monopolistic Competition*. Cambridge: Harvard University Press.
20. Clarke E. 1971. Multipart Pricing of Public Goods. *Public Choice* 11: 17-33.
21. Coase R. H. The marginal cost controversy // *Economica, New Series*, 1946, Vol. 13, No. 51, pp. 169-182.
22. Dallman PR. Iron deficiency and the immune response. *Am J Clin Nutr*. 1987;46(2):329–34. DOI: 10.1093/ajcn/46.2.329

23. DeMaeyer E.M., Dallman P., Gurney J.M., Hallberg L., Sood S.K., Srikantia S.G. Preventing and controlling iron deficiency anaemia through primary health care // World Health Organization: Geneva. - 1989. - 58p.
24. Demsetz H. 1970. The Private Production of Public Goods. *Journal of Law and Economics* 13: 293-305.
25. Detering D. 2001. *Ökonomie der Medieninhalte. Allokative Effizienz und soziale Chancengleichheit in den Neuen Medien*. Herausgegeben von Klaus Backhaus, Heinz-Lothar Grob, Bernd Holznagel, Wolfram-Manfred Lippe und Gerhard W. Wittkämper. *Telekommunikation und Multimedia*. Band 6. Münster: LIT. Zugl.: Münster (Westf.), Univ., Diss., 1999.
26. Ekiz C, Agaoglu L, Karakas Z, Gurel N, Yalcin I. The effect of iron deficiency anemia on the function of the immune system. *Hematol J*. 2005;5(7):579–83. DOI: 10.1038/sj.thj.6200574
27. Ershler WB, Chen K, Reyes EB, Dubois R. Economic burden of patients with anemia in selected diseases. *Value Health*. 2005;8(6):629–38. DOI: 10.1111/j.1524-4733.2005.00058.x
28. Erutku C., Richelle Y. Licensing a new product with non-linear contracts // *Canadian Journal of Economics/Revue canadienne d'économique*. – 2006. – Vol. 39. – №. 3. – C. 932-947.
29. Essegai S., Gupta S., Zhang Z. J. Pricing access services // *Marketing Science*. – 2002. – Vol. 21. – №. 2. – C. 139-159.
30. Fink C., Khan M., Zhou H. 2015. Exploring the worldwide patent surge. *Economics of Innovation and New Technology* 25 (2): 114–142. DOI: 10.1080/10438599.2015.1055088
31. Firoozeh N., Nazarenko A., Alizon F., Daille B. (2019). Keyword extraction: Issues and methods. *Natural Language Engineering*, no. 26 (3), pp. 259-291.
32. Fisch C., Sandner P., Regner L. 2017. The value of Chinese patents: An empirical investigation of citation lags. *China Economic Review* 45: 22–34. DOI: 10.1016/j.chieco.2017.05.011

33. GBD 2021 Anaemia Collaborators. Prevalence, years lived with disability, and trends in anaemia burden by severity and cause, 1990–2021: findings from the global burden of disease study 2021. *Lancet Haematol.* 2023;10(9):e713–34. DOI: 10.1016/S2352-3026(23)00160-6
34. Gelaw Y, Getaneh Z, Melku M. Anemia as a risk factor for tuberculosis: a systematic review and meta-analysis. *Environ Health Prev Med.* 2021;26:1–15. DOI: 10.1186/s12199-020-00931-z
35. Gigerenzer G. Dread risk, September 11, and fatal traffic accidents // *Psychological science.* - 2004. - Vol. 15. - Issue 4. - p. 286-287.
36. Gigerenzer G. Out of the frying pan into the fire: Behavioral reactions to terrorist attacks // *Risk Analysis: An International Journal.* - 2006 - Vol. 26. - Issue 2. - p. 347-351.
37. Glanz K, Sallis J.F., Saelens B.E., Frank L.D. Healthy nutrition environments: concepts and measures // *American journal of health promotion.* - 2005. - Vol. 19. - №. 5. - p. 330-333.
38. Grachev I. D., Larin S. N., Noakk N. V. Development of a comprehensive digital model of cost-effective quarantine measures against the COVID-19 epidemic // *Materials of the International Conference "Scientific research of the SCO countries: synergy and integration"* - 2020. - p. 10-16.
39. Granstrand O., Holgersson M., Opedal A. Towards fair pricing in technology trade and licensing // *Stockholm Intellectual Property Law Review.* - 2020. - Vol. 3. - Issue 1. - pp. 6-15.
40. Green J., Laffont J.-J. 1977. Imperfect Personal Information and the Demand Revealing Process. A Sampling Approach. *Public Choice* 29 (2/special issue): 79-94.
41. Grubb, Michael D. 2009. Selling to Overconfident Consumers // *American Economic Review* 99 (5): 1770–1807.
42. Gul F., Pesendorfer W. Self-control, revealed preference and consumption choice // *Review of Economic Dynamics* – 2004 – Vol. 7 – Issue 2 – pp. 243–264

- 43.Hartley R.V.L. Transmission of information. Bell System Techn. J., 1928, v.7
- 44.Hayek F.A. The Use of Knowledge in Society // The American Economic Review, Vol. 35, No. 4. (Sep., 1945), pp. 519-530
- 45.Hlatswayo B.P.S., Ntshangase S., de Villiers F.P.R. The effects of iron deficiency and anaemia on primary school learners' scholastic performance // South African Journal of Child Health. - 2016. - Vol. 10. - №. 2. - p. 111-115.
- 46.Hofmann T. Probabilistic Latent Semantic Analysis // Proceedings of the Fifteenth conference on Uncertainty in artificial intelligence, Stockholm, 1999, pp. 289–296.
- 47.Hotelling H. 1929. Stability in Competition. Economic Journal 39 (March): 41-57.
- 48.Hotelling H. The general welfare in relation to problems of taxation and of railway and utility rates // Econometrica, 1938. Vol. 6, No. 3, pp. 242-269.
- 49.Howard R.A. Information Value Theory // IEEE Transactions on Systems Science and Cybernetics. 1966. vol. 2. no. 1. pp. 22–26.
- 50.Hu A.G., Jefferson G.H. 2009. A great wall of patents: What is behind China's recent patent explosion? Journal of Development Economics 90 (1): 57–68. DOI: 10.1016/j.jdeveco.2008.11.004
- 51.Hu D., Yan J. D. Profit Allocation and Realization in Regional Technology Transfer Alliance Based on Game Theory // Advanced Materials Research. - 2014. - Vols. 834-836. - pp. 2011-2016.
- 52.Hunter E., Namee B.M., Kelleher J.D. A taxonomy for agent-based models in human infectious disease epidemiology // Journal of Artificial Societies and Social Simulation. - 2017. - Vol. 20. - Issue 3. - p.2.
- 53.Isanaka S, Mugusi F, Urassa W, Willett WC, Bosch RJ, Villamor E, et al. Iron deficiency and anemia predict mortality in patients with tuberculosis. J Nutr. 2012;142(2):350–7. DOI: 10.3945/jn.111.144287
- 54.Ismail IM, Kahkashan A, Antony A, Sobhith VK. Role of socio-demographic and cultural factors on anemia in a tribal population of North Kerala, India.

- Int J Community Med Public Health. 2016;3(5):1183–8. DOI: 10.18203/2394-6040.ijcmph20161381
55. Jackson AA. Anemia in severe undernutrition (malnutrition). In: Kraemer K, Zimmermann B, editors. Nutritional anemia. Basel: Sight and Life; 2007. p.215–
56. Jiang X., Wang L., Cao B., Fan X. Benefit distribution and stability analysis of enterprises' technological innovation cooperation alliance // Computers & Industrial Engineering. - 2021. - Vol. 161. - p. 107637.
57. Kamphuis C.B., Giskes K., de Bruijn G.J., Wendel-Vos W., Brug J., Van Lenthe F.J. Environmental determinants of fruit and vegetable consumption among adults: a systematic review // British journal of nutrition. - 2006. - Vol.96. - №4. - p.620-635.
58. Kaur R., Mishra S., Nevolin I.V., Roy Choudhury D., Singh M. Nutritional anemia: Patent landscape // World Patent Information. – 2022. – Vol. 70. – P. 102123. – DOI 10.1016/j.wpi.2022.102123.
59. Kumar P, Sharma H, Sinha D. Socio-economic inequality in anaemia among men in India: a study based on cross-sectional data. BMC Public Health. 2021;21:1345. DOI: 10.1186/s12889-021-11393-5
60. Kumar V, Jain M, Shukla U, Swarnkar M, Gupta P, Saini P. Prevalence of anemia and its determinants among pregnant women in a rural community of Jhalawar, Rajasthan. Natl J Community Med. 2019;10(04):207–11.
61. Lambrecht A., Skiera B. (2006). Paying Too Much and Being Happy About It: Existence, Causes, and Consequences of Tariff-Choice Biases // Journal of Marketing Research, 43(2), 212–223.
62. Lamjed B.S., Drogoul A., Bouron T. Multi-agent based simulation of consumer behaviour: Towards a new marketing approach // Proceedings of the International Congress on Modelling and Simulation (MODSIM 2001), Canberra, Australia. 2001.
63. Ley E. On the private provision of public goods: a diagrammatic exposition // Investigaciones Economicas, 20 (1996), pp. 105-123

64. Li Y., Zhang D., Pagán J. A. Social norms and the consumption of fruits and vegetables across New York City neighborhoods // *Journal of Urban Health*. - 2016. - Vol.93. - №2. - p. 244-255.
65. List J. 2010. An A to X of patent citations for searching. *World Patent Information* 32 (4): 306–312. DOI: 10.1016/j.wpi.2010.01.004
66. Litofsky NS, Martin S, Diaz J, Ge B, Petroski G, Miller DC, et al. The negative impact of anemia in outcome from traumatic brain injury. *World Neurosurg*. 2016;90:82–90. DOI: 10.1016/j.wneu.2016.02.076
67. Loise V., Stevens A.J. The Bayh-Dole Act Turns 30 // *les Nouvelles*. - 2010. – Vol 45. - № 4. – pp. 185-194.
68. Lucas Jr R. E. On the mechanics of economic development // *Journal of monetary economics*. – 1988. – T. 22. – №. 1. – C. 3-42.
69. Luo L., Lu R. Application of the Shapley value in the game model of the cooperation of enterprises, colleges and institutes // *Soft Science* - 2001 - Vol.15 - Issue 2 - pp.17-19.
70. Machlup F. *The production and distribution of knowledge in the United States* / Princeton university press, 1962.
71. Mashkova A.L., Savina O.A., Mashkov E.A., Nevolin I.V., Buriлина M.A. Generating Social Environment for Agent-Based Models of Computational Economy // *Communications in Computer and Information Science*. – 2020. – Vol. 1349. – P. 291-305. – DOI 10.1007/978-3-030-67238-6_21. – EDN JKDVZR.
72. Mashkova AL, Dukhi N, Nevolin IV, Savina OA. Predictive assessment of the availability of a balanced diet for residents of Russian regions: an agent-based approach. *Economic and social changes: facts, trends, forecast*. 2021;14(6):107-25. DOI: 10.15838/esc.2021.6.78.6
73. Matschke M.J., Brösel G., Matschke X.. "Fundamentals of Functional Business Valuation," *Journal of Business Valuation and Economic Loss Analysis*: Vol. 5 : Iss. 1, Article 7, 2010.

74. Mimno D., Wallach H. M., Naradowsky J., Smith D. A., McCallum A. Polylingual topic models // Proceedings of the 2009 Conference on Empirical Methods in Natural Language Processing: Volume 2 - Volume 2. — EMNLP '09. — Stroudsburg, PA, USA: Association for Computational Linguistics, 2009. — Pp. 880–889
75. Montoro-Huguet MA, Santolaria-Piedrafita S, Cañamares-Orbis P, García-Erce JA. Iron deficiency in Celiac disease: prevalence, health impact, and clinical management. *Nutrients*. 2021;13(10):3437. DOI: 10.3390/nu13103437
76. Morsy N, Alhady S. Nutritional status and socio-economic conditions influencing prevalence of anaemia in pregnant women. *Int J Sci Technologyresearch*. 2014;3(7):54–60.
77. Nevolin I. Crowdsourcing Opportunities for Research Information Systems // *Procedia Computer Science: 13th, Scotland, 09–11 июня 2016 года*. – Scotland, 2017. – P. 19-24. – DOI 10.1016/j.procs.2017.03.004.
78. Nevolin I., Kozyrev A. Developing CRIS Module for Technology // *Procedia Computer Science, Rome, 13–15 мая 2014 года*. – Rome, 2014. – P. 158-162. – DOI 10.1016/j.procs.2014.06.026.
79. Nissenon AR, Wade S, Goodnough T, Knight K, Dubois RW. Economic burden of anemia in an insured population. *J Managed Care Pharm*. 2005;11(7):565–74. DOI: 10.18553/jmcp.2005.11.7.565
80. Nitzan S., Ostrom E. (1992) 'The Nature and Severity of Collective-Action Problems - The Voluntary, Collective Provision of Mixed Goods Approach', Mimeo, Indiana University.
81. Nonaka I., Takeuchi H. *The Knowledge-Creating Company: How Japanese Companies Create the Dynamics of Innovation* / Oxford University Press, 1995.
82. Nyberg ST, Singh-Manoux A, Pentti J, Madsen IEH, Sabia S, Alfredsson L, et al. Association of healthy lifestyle with years lived without major chronic

- diseases. *JAMA Intern Med.* 2020;180(5):760–8. DOI: 10.1001/jamaintern-med.2020.0618
83. Oakland W. 1974. Public Goods, Perfect Competition, and Underproduction. *Journal of Political Economy* 82 (Sept./Oct.): 927-939.
84. Oi W.Y. A Disneyland dilemma: Two-part tariffs for a Mickey Mouse monopoly // *The Quarterly Journal of Economics.* – 1971. – Vol. 85. – №. 1. – pp. 77-96.
85. Parinov S. Citation contexts as a data source for evaluation of scholarly consumption // *Scientometrics.* – 2021. – vol. 126 – pp.9249-9265
86. Parinov S. Towards a semantic segment of a research e-infrastructure: Necessary information objects, tools and services // *International Journal of Metadata, Semantics and Ontologies.* – 2013. – Vol. 8, No. 4. – P. 322-331
87. Pazienza M.T., Pennacchiotti M., Zanzotto F.M. (2006). Terminology Extraction: An analysis of linguistic and statistical approaches. Sirmakessis S. (ed.) *Knowledge Mining. Studies in Fuzziness and Soft Computing*, vol. 185. Berlin: Springer, Heidelberg. P. 255-279.
88. Perez L., Dragicevic S. An agent-based approach for modeling dynamics of contagious disease spread // *International journal of health geographics.* - 2009. - Vol. 8. - p. 50.
89. Pollok R. Forever Minus a Day? Theory and Empirics Of Optimal Copyright Term // *MPRA Paper No. 8887.* - 2008. - 28p.
90. Quastler H. *The emergence of biological organization.* — New Haven: Yale University Press, 1964.
91. Romer P.M. Endogenous Technological Change, *Journal of Political Economy*, Vol. 98, No. 5, Part 2, 1990 pp. S71-S102
92. Rose D., Richards R. Food store access and household fruit and vegetable use among participants in the US Food Stamp Program // *Public health nutrition.* - 2004. - Vol. 7. - №. 8. - C. 1081-1088.
93. Ross J, Horton S. *Economic consequences of iron deficiency.* Ottawa: Micro-nutrient Initiative, IDRC; 1998.

94. Rothenberg J. 1962. Consumer Sovereignty and the Economics of TV Programming. *Studies in Public Communication* 4: 45-54.
95. Roy B. *Multicriteria Methodology for Decision Aiding*. – Dordrecht: Kluwer Academic Publisher, 1996.
96. Saaty T.L. *Decision Making with Dependence and Feedback. The Analytic Network Process*. Pittsburgh: PWS Publications, 1996, 370 p.
97. Saaty T.L. *Fundamentals of Decision making and Priority Theory with the Analytic Hierarchy Process*. – Pittsburgh: RWS Publications, 1994, 527 p.
98. Salop S. 1979. Monopolistic Competition with Outside Goods. *Bell Journal of Economics* 10: 141-156.
99. Salton G., Buckley C. (1988). Term-weighting approaches in automatic text retrieval. *Information Processing and Management*, no. 5 (24), pp. 513-523.
100. Samuelson P. 1954. The Pure Theory of Public Expenditure. *Review of Economics and Statistics* 36 (November): 387-389.
101. Schmalensee R. Monopolistic two-part pricing arrangements // *The Bell Journal of Economics*. – 1981. – Vol.12. – №. 2. – pp. 445-466.
102. Shannon C.E. A mathematical theory of communication. *Bell System Techn. J.*, 1948, v. 27, N 3
103. Shasha D. The negative impact of anemia on radiotherapy and chemotherapy outcomes. *Semin Hematol.* 2001;38(3):8–15. DOI: 10.1016/S0037-1963(01)90125-8
104. Shilling F., Waetjen D. Special report (update): Impact of COVID19 mitigation on numbers and costs of California traffic crashes. - 2020.
105. Shisana O, Labadarios D, Rehle T, Simbayi L, Zuma K, Dhansay A, et al. South African National health and nutrition examination survey (SANHANES-1): 2014 edition. Cape Town: HSRC; 2014.
106. Simperl E., Bürger T., Hangl S., Wörgl S., Popov I. ONTOCOM: A reliable cost estimation method for ontology development projects // *Journal of Web Semantics*. – 2012. – Vol. 16. – pp. 1-16.

107. Solow R. M. A Contribution to the Theory of Economic Growth // The Quarterly Journal of Economics. — 1956. —vol. 70, № 1. — P. 65—94.
108. Spence M., Owen B. 1977. Television Programming, Monopolistic Competition and Welfare. Quarterly Journal of Economics 91: 103-126.
109. Steiner P. 1952. Program Patterns and Preferences, and the Workability of Competition in Radio Broadcasting. Quarterly Journal of Economics 66: 194-223.
110. Thompson E. 1968. The Perfectly Competitive Production of Collective Goods. The Review of Economics and Statistics 50: 1-12.
111. Trigeorgis L. Real options: Managerial flexibility and strategy in resource allocation / Cambridge: MIT press, 1996. – 427pp.
112. UNICEF, WFP, WHO. The state of food security and nutrition in the world 2017. Building resilience for peace and food security. Rome: Food and Agriculture Organization; 2017.
113. Varian H (2019), Artificial Intelligence, Economics, and Industrial Organization // National Bureau of Economic Research, Volume Title: The Economics of Artificial Intelligence, p. 399 – 419
114. Varian, Hal R. 1998. Markets for Information Goods.
115. Wang C, Wang Y. Trends in prevalence and treatment rate of anemia in the U.S. Population: cross-sectional study using data from NHANES 2005–2018. Hematology. 2022;27(1):881–8. DOI: 10.1080/16078454.2022.
116. WIPO 2018. World Intellectual Property Indicators.
117. WIPO 2019. World Intellectual Property Indicators.
118. WIPO 2020. World Intellectual Property Indicators.
119. World Health Organization (WHO). Serum ferritin concentrations for the assessment of iron status and iron deficiency in populations. Vitamin and Mineral Nutrition Information System / Geneva: World Health Organization (WHO) (WHO/NMH/NHD/MNM/11.2) - 2011. URL: http://www.who.int/vmnis/indicators/serum_ferritin.pdf (Дата обращения 20.09.2019).

120. Wu C., Marchese M., Jiang J., Ivanyukovich A., Liang Y. (2007). Machine learning-based keywords extraction for scientific literature. *Journal of Universal Computer Science*, vol. 13, no. 10, pp. 1471-1483.
121. Xu X, Liu S, Rao Y, Shi Z, Wang L, Sharma M, et al. Prevalence and sociodemographic and lifestyle determinants of anemia during pregnancy: A Cross-Sectional study of pregnant women in China. *Int J Environ Res Public Health*. 2016;13:908. DOI: 10.3390/ijerph13090908
122. Yang X.H., Wang B., Chen S.Y., Wang W.L. Epidemic dynamics behavior in some bus transport networks // *Physica A: Statistical Mechanics and its Applications*. - 2012. - Vol. 391. - Issue 3. - p. 917-924.
123. Zemariam AB, Yimer A, Abebe GK, Wondie WT, Abate BB, Alamaw AW et al. Employing supervised machine learning algorithms for classification and prediction of anemia among youth girls in Ethiopia. *Scientific Reports*. 2024;14: 9080(2024). DOI: 10.1038/s41598-024-60027-4
124. Zhang J. Transport policymaking that accounts for COVID-19 and future public health threats: A PASS approach // *Transport policy*. - 2020. - Vol. 99. - p. 405-418.
125. Zhang J., Hayashi Y., Frank L.D. COVID-19 and transport: Findings from a world-wide expert survey // *Transport policy*. - 2021 - Vol. 103. - p. 68-85.
126. Zuse K. Verfahren zur selbsttätigen Durchführung von Rechnungen mit Hilfe von Rechen-maschinen. Patentanmeldung Z 23 139 / GMD Nr. 005/021. – 1936
127. Абдулов Р.Э. Экономическая безопасность отраслей электронной промышленности и ИТ: Научный доклад. – М.: Институт экономики РАН, 2024 – 52 с.
128. Аверкин А.И., Батыршин И.З., Блишин А.Ф., Силов В.Б., Тарасов В.Б. Нечеткие множества в моделях управления и искусственного интеллекта. – М.: Наука, Физматлит, 1986, 312 с.

129. Аджемоглу Д., Робинсон Дж. А. Почему одни страны богатые, а другие бедные. Происхождение власти, процветания и нищеты / М.: АСТ, 2016. — 693 с.
130. Андрейчиков А.В., Андрейчикова О.Н. Анализ, синтез, планирование решений в экономике. – М.: Финансы и статистика, 2000, 368 с.
131. Андрейчиков А.В., Андрейчикова О.Н. Интеллектуальные информационные системы. – М.: Финансы и статистика, 2006, 424 с.
132. Андрейчиков А.В., Андрейчикова О.Н. Наука и искусство принятия решений. Кн. 1: Принятие решений: Условия определённости. Условия риска. / М.: ЛЕНАНД, 2021. – 240с.
133. Андрейчиков А.В., Тевелева О.В., Неволин И.В. [и др.] Методика проведения поисковых исследований по выявлению возможностей импортозамещения высокотехнологичной продукции на основе мировых патентных и финансовых информационных ресурсов // Экономика и предпринимательство. – 2019. – № 4(105). – С. 157-167.
134. Багриновский К.А. Модели и методы экономической кибернетики / М.: Экономика, 1973. -206 с.
135. Багриновский К.А., Егорова Н.Е. Имитационные системы в планировании экономических объектов / М.: Наука, 1980. - 238 с.
136. Балычев С.Ю., Батьковский А.М., Батьковский М.А., Белоусов Ф.А., Неволин И.В. Совместное использование ресурсов в индустрии 4.0 // Цифровая экономика. – 2023. – № 3(24). – С. 89-96. – DOI 10.34706/DE-2023-03-11.
137. Бахтизин А.Р., Варшавский А.Е., Варшавский Л.Е., Козырев А.Н., Комкина Т.А., Макаров В.Л., Неволин И.В., Паринов С.И. Проблемы оценки и измерения человеческого капитала в образовании и науке: Коллективная монография / Москва, Санкт-Петербург: Издательство Нестор-История, 2014. – 240 с.

138. Белоусов Ф.А., Иванова А.К., Неволин И.В. Технологический суверенитет и глобальная конкуренция // Цифровая экономика. – 2024. – № 4(30). – С. 24-33. – DOI 10.34706/DE-2024-04-03.
139. Борисов А.Н., Алексеев А.В. и др. Обработка нечеткой информации в системах принятия решений. – М.: Радио и связь, 1989, 304 с.
140. Борисов А.Н., Крумберг О.А., Федоров И.П. Принятие решений на основе нечетких моделей. – Рига: Зинатне, 1990, 184 с.
141. Вентцель Е.С. Исследование операций: задачи, принципы, методология / М.: Наука. – 1988. — 208 с.
142. Верещагин Н.К., Семёнов А.Л., Шень А.Х. Последнее открытие Колмогорова? (Колмогоров и алгоритмическая статистика), Теория вероятн. и ее примен., 2023, том 68, выпуск 4, 719–750
143. Воронцов К.В. Антропоцентрическая система представления знаний в будущей человеко-машинной цивилизации // Цифровая дидактика: Сборник статей по материалам Всероссийского педагогического форума, Пермь, 28–29 марта 2024 года. – Пермь: Пермский государственный гуманитарно-педагогический университет, 2024. – С. 29-33.
144. Воротникова Д.В., Моисеев Н.А., Трындина Н.С. Анализ реакции цен отрасли на влияние внешних факторов на основе модели межотраслевого баланса // Вестник Российского экономического университета имени Г.В. Плеханова. – 2023. – Т. 20, № 1(127). – С. 109-115.
145. Вэриан Х. Р. (1997) Микроэкономика. Промежуточный уровень. Современный подход. Учебник для вузов/Пер. с англ, под ред. Н.Л. Фроловой. - М.: ЮНИТИ, 1997. - 767 с.
146. Герасименко Н.А., Ватолин А.С., Янина А.О., Воронцов К.В. SciRus: легкий и мощный мультязычный энкодер для научных текстов // Доклады Российской академии наук. Математика, информатика, процессы управления. – 2024. – Т. 520, № S2. – С. 216-227. – DOI 10.31857/S2686954324700589.

147. Грачев И. Д., Ноакк Н. В.. Комбинированная модель реально-виртуальной экономики//Аудит и финансовый анализ. 2011. № 1. С. 92-101.
148. Грачев И.Д. Самосогласованное поле рыночных стоимостей // Стратегическое планирование и развитие предприятий: Материалы XX Всероссийского симпозиума, Москва, 09–10 апреля 2019 года / Под редакцией Г.Б. Клейнера. – Москва: Центральный экономико-математический институт РАН, 2019. – С. 187-189.
149. Грачёв И.Д. Статфизическая модель экономического прогресса // Электронное научно-практическое периодическое издание "Современные научные исследования и разработки" - 2018. - №10(27). - с. 267-269.
150. Грачёв И.Д., Грачёв Д.И., Ларин С.Н., Ноакк Н.В. (2020b) Цифровая модель экономически оптимальных противоэпидемических ограничений // Экономика и предпринимательство - 2020. - Вып.14. - №2(115). - с.869-872.
151. Грачёв И.Д., Грачёв Д.И., Ларин С.Н., Ноакк Н.В. (2020a) Оценка экономических результатов различных вариантов карантинных с использованием комбинированной цифровой экономико-эпидемической модели // Экономика и предпринимательство. - 2020. - Вып.14. - №2(115). - с.902-909.
152. Грачёв И.Д., Ларин С.Н., Ноакк Н.В. Оптимизация управления противоэпидемическими мероприятиями на основе цифровой социально-экономической модели // Экономика и предпринимательство - 2021. - Вып.15. - №3(128). - С.67-74.
153. Грачев И.Д., Неволин И.В. Дифференцированный подход к карантинным мероприятиям // Цифровая экономика. – 2021. – № 2(14). – С. 16-23.
154. Грачев И.Д., Неволин И.В. Инновационная активность и экономический рост // Инновации. – 2019. – № 8(250). – С. 3-8.
155. Грачев И.Д., Неволин И.В. Модель для анализа отраслевой динамики // Системное моделирование социально-экономических процессов:

- труды 43-ой Международной научной школы- семинара, Воронеж, 13–18 октября 2020 года / Под редакцией В.Г. Гребенникова, И.Н. Щепиной. – Воронеж: Воронежский государственный университет, 2020. – С. 427-432.
156. Губанов Д.А., Кузнецов О.П., Курако Е.А., Лемтюжникова Д.В., Новиков Д.А., Чхартишвили А.Г. Информационная система анализа научной деятельности (ИСАНД) в области теории управления // Проблемы управления. – 2024. – № 3. – С. 42-65. – DOI 10.25728/ru.2024.3.4.
157. Дементьев В.Е., Светлов Н.М. Эффект стратегического подхода при динамическом ценообразовании на сетевые блага // Экономика и математические методы. – 2020. – Т. 56, № 2. – С. 20-31. – DOI 10.31857/S042473880009218-8
158. Дементьев В.Е., Устюжанина Е.В., Евсюков С.Г. Цифровая трансформация цепочек создания ценности: "улыбка" может оказаться "хмурой" // Журнал институциональных исследований. – 2018. – Т. 10, № 4. – С. 58-77. – DOI 10.17835/2076-6297.2018.10.4.058-077.
159. Дубов Ю.А., Травкин С.И., Якимец В.Н. Многокритериальные модели формирования и выбора вариантов систем. – М.: Наука, 1986, 296 с.
160. Дукхи Н., Каур Р., Машкова А.Л., Неволин И.В. Информационное обеспечение агент-ориентированной модели динамики распространения анемии в странах БРИКС // Информационные системы и технологии. – 2022. – № 1(129). – С. 36-44.
161. Ершов П.С., Катин А.В., Малахов В.А., Паджев В.В., Хохлов Ю.Е., Шапошник С.Б., Янышен А.А. Оценка уровня цифрового развития отдельных сфер деятельности в Российской Федерации: пилотный рейтинг // Информационное общество. – 2024. – № S7. – С. 2-20.
162. Жуковин В.Е. Нечеткие многокритериальные модели принятия решений. – Тбилиси: Мецниереба, 1988, 70 с.

163. Завьялова Е.А., Кобылко А.А. Применение маркетинговых инструментов в деятельности научных организаций экономического профиля: результаты бенчмаркинга // Экономическая наука современной России. – 2025. – Т. 28, № 2. – С. 70-84. – DOI 10.33293/1609-1442-2025-28(2)-70-84.
164. Заде Л. Понятие лингвистической переменной и его применение к принятию приближенных решений. – М. Мир, 1986, 165 с.
165. Заде Л.А. Основы нового подхода к анализу сложных систем и процессов принятия решений. В кн.: «Математика сегодня». – М.: Знание, 1974, с. 5-49.
166. Ильин Н.И. Интеллектуализация ситуационных центров развития // Актуальные проблемы глобальных исследований: Россия в глобализирующемся мире: Сборник научных трудов участников VI Всероссийской научно-практической конференции с международным участием, Москва, 04–06 июня 2019 года / Под редакцией И.В. Ильина. – Москва: Межрегиональная общественная организация содействия изучению, пропаганде научного наследия Н.Д. Кондратьева, 2019. – С. 124-127.
167. Ильина И.Е., Гагиев Н.Н., Белкин Ю.Д. Научно-технологическая инфраструктура как инструмент коммуникации на пространстве СНГ, ЕАЭС, ШОС // Международное сотрудничество вузов государств - участников Содружества Независимых Государств. Роль технических университетов в формировании единого научно-технологического и образовательного пространства СНГ: Сборник научных статей. – Москва: Ассоциация технических университетов, 2025. – С. 40-47.
168. Ильина И.Е., Дегтярев А.Н. СНТР-2.0. Триггеры и драйверы технологического суверенитета России // Управление наукой и наукометрия. – 2024. – Т. 19, № 3. – С. 486-499.
169. Казакова Е. А., Сандомирская М. С., Суворов А. Д., Хажгериева А. И., Шавшин Р. К. (2023). Платформы, онлайн–рынки труда и

- краудсорсинг. Часть 2. Краудсорсинг // Журнал Новой экономической ассоциации. № 4 (61). С. 128–144.
170. Канеман Д. Думай медленно... решай быстро. — М.: АСТ, 2013. — 625 с.
171. Капустина О.В. Моделирование ценообразования на рынке информационной продукции: специальность 08.00.13 "Математические и инструментальные методы экономики" : автореферат диссертации на соискание ученой степени кандидата экономических наук / Капустина Ольга Владимировна. — Москва, 2011. — 25 с.
172. Каравай А.В. Динамика состояния здоровья россиян в 2003–2018 гг.: взаимосвязь объективных и субъективных характеристик // Социологическая наука и социальная практика, 2020. Т. 8. № 4. С. 69-87. DOI: <https://doi.org/10.19181/snsp.2020.8.4.7657>
173. Качалов Р.М., Кобылко А.А., Слепцова Ю.А., Чернявский С.В.. Маркетинг в деятельности научной организации: функции и основные направления работ / Москва : Центральный экономико-математический институт РАН, 2023. — 152 с. — ISBN 978-5-8211-0820-3. — DOI 10.33276/978-5-8211-0820-3.
174. Когаловский М.Р., Паринов С.И. Научные коммуникации в среде семантически обогащаемых электронных библиотек // Программная инженерия. — 2015. — № 4. — С. 31-38.
175. Козырев А. Н. (2020а) Параллели - Анатолий Китов и Конрад Цузе // Цифровая экономика. — 2020. — № 3(11). — С. 60-72.
176. Козырев А. Н. (2020b) Современное состояние исследований в области торговли информацией // Цифровая экономика. — 2020. — № 1(9). — С. 63-75
177. Козырев А. Н. Оптимальные двухкомпонентные цены в экономиках с возрастающей отдачей // Цифровая экономика. — 2023. — № 1(22). — С. 54-64.

178. Козырев А. Н. Оценка интеллектуальной собственности: Функциональный подход и математические методы / Екатеринбург: Общество с ограниченной ответственностью «Издательские решения», 2016. – 350 с.
179. Козырев А. Н., Макаров В.Л. Оценка стоимости нематериальных активов и интеллектуальной собственности / Москва : РИЦ ГШ ВС РФ, 2003. – 368 с.
180. Козырев А.Н. Моделирование НТП, упорядоченность и цифровая экономика // Экономика и математические методы. - 2011. - Т. 47, № 4. - С. 131-142.
181. Козырев А.Н. Оптимизация распределения ресурсов в системе линейных моделей производства. Оптимизация 16 (33), 1975. с. 62-72.
182. Козырев А.Н. Сети изобретательства и инноваций на основе смарт-контрактов // Цифровая экономика. – 2021. – № 2(14). – С. 5-15.
183. Козырев А.Н. Цифровизация, математические методы и системный кризис экономической науки // Цифровая экономика. - 2019. - № 4(8). - С. 5-20.
184. Колмогоров А.Н. (1965) Три подхода к определению понятия “количество информации”. Проблемы передачи информации. Т.1, вып.1, сс.3-11
185. Копкин Е. В., Кобзарев И. М. Использование меры ценности информации Стратоновича для оптимизации гибких программ диагностирования технических объектов // Труды СПИИРАН. – 2019. – Т. 18. – № 6. – С. 1434-1461.
186. Костин А.В., Неволин И.В. Стоимость права использования товарного знака в составе группы средств индивидуализации // Цифровая экономика. – 2023. – № 3(24). – С. 23-28.
187. Кофман А. Введение в теорию нечетких множеств. – М.: Радио и связь, 1982, 432 с.

188. Ларичев О.И., Браун Р. Количественный и вербальный анализ решений: сравнительное исследование возможностей и ограничений // Экономика и математические методы. – 1998. – Т.34. – вып.4 – с. 97-107.
189. Ларичев О.И., Мошкович Е.М. Качественные методы принятия решений. Вербальный анализ решений / М.: Наука, Физматлит. – 1996. – 208 с.
190. Лессиг Л. (2007) Свободная культура / Пер. с англ. – М.: Прагматика Культуры, 2007.
191. Лорьер Ж.-Л. Системы искусственного интеллекта. – М.: Мир, 1991, 568 с.
192. Лугачев М. И. Россия на пути к цифровой экономике: аспект времени и пространства // Цифровая экономика. – 2019. – № 2(6). – С. 12-18.
193. Макаров В. Л., Бахтизин А. Р., Сушко Е. Д., Агеева А. Ф. Имитация социально-экономической системы Евразийского континента с помощью агент-ориентированных моделей // Прикладная эконометрика. - 2017. - Т.48. - с.122-139.
194. Макаров В.Л., Бахтизин А. Р., Россошанская Е. А., Дорошенко Т.А., Самсонова Н.А. Проблемы стандартизации описания агент-ориентированных моделей и возможные пути их решения // Вестник Российской академии наук. – 2023. – Т. 93, № 4. – С. 362-372.
195. Макаров В.Л., Клейнер Г.Б. Микроэкономика знаний / М.: ЗАО «Издательство «Экономика», 2007. – 204 с.
196. Макаров В.Л., Маршак В.Д. Имитация взаимодействия отраслей АСУ и АСПР Госплана СССР в процессе составления плана отрасли // Применение имитационных методов к анализу процедур планирования и управления в экономике. - М.: изд. ЦЭМИ АН СССР, 1977. - С. 6-10.
197. Макаров В.Л., Маршак В.Д. Модели оптимального функционирования отраслевых систем. - М.: Экономика, 1979.- 160 с.

198. Макаров В.Л., Маршак В.Д. О соотношении оптимального имитационного подходов при построении и анализе экономических моделей. // Оптимизация. - Новосибирск: изд. ИМ СО АН СССР, 1974. - №13 (30). - С. 12-13.
199. Машкова А.Л., Дрипта Р.Ч., Ришемжит К., Неволин И.В. Агент-ориентированная модель взаимосвязи доступности продуктов питания и динамики распространения анемии // Искусственные общества. – 2021. – Т. 16, № 1. – DOI 10.18254/S207751800013573-9.
200. Машкова А.Л., Милкова М.А., Неволин И.В. Система мониторинга заболеваний в региональном разрезе // Цифровая экономика. – 2020. – № 1(9). – С. 30-39. – DOI 10.34706/DE-2020-01-03.
201. Мелихов А.Н., Бернштейн Л.С., Коровин С.Я. Ситуационные советующие системы с нечеткой логикой. – М.: Наука, 1990, 272 с.
202. Милкова М.А. Разработка и экспериментальная апробация методики построения тематической модели при работе с научной информацией: специальность 08.00.13 «Математические и инструментальные методы экономики»: диссертация на соискание ученой степени кандидата экономических наук / Милкова Мария Александровна, 2021. – 172 с.
203. Милкова М.А., Андрейчикова О.Н. Multichoice // Свидетельство о государственной регистрации программы для ЭВМ № 2016616698. – 2016.
204. Милкова М.А., Неволин И.В., Пигорев Д.П. Современные методы извлечения ключевой информации из нормативных документов // Экономическая наука современной России. – 2021. – № 2(93). – С. 101-114. – DOI 10.33293/1609-1442-2021-2(93)-101-114.
205. Мильнер Б.З. Управление знаниями / М.: ИНФРА-М, 2003
206. Миркин Б.Г. Проблема группового выбора. – М.: Наука, 1974, 256 с.
207. Моисеев Н.А., Внуков И.А., Сокерин П.О. Оценка эффектов различных вариантов импортозамещения методом «затраты-выпуск» на

- примере Российской Федерации // Экономика и математические методы. – 2023. – Т. 59, № 1. – С. 30-47.
208. Неволин И. В. Оптимизация ставок роялти в лицензионных договорах на основе распределенных вычислений и сетевых технологий : специальность 08.00.13 "Математические и инструментальные методы экономики" : диссертация на соискание ученой степени кандидата экономических наук / Москва, 2012. – 143 с.
209. Неволин И.В. (2023а) Искусственные общества: технологии построения и сферы применения // Экономическая наука современной России. – 2023. – № 4(103). – С. 85-93
210. Неволин И.В. (2023b) Разделение стоимости портфеля прав на средства индивидуализации между его компонентами // Цифровая экономика. – 2023. – № 3(24). – С. 29-33.
211. Неволин И.В. (2023с) Вознаграждение за передачу технологий: методы расчёта и границы применимости // Цифровая экономика. – 2023. – № 2(23). – С. 31-35.
212. Неволин И.В. (2023d) Экономические модели и недискриминируемый доступ к данным // Цифровая экономика. - 2023. - № S5(26). - С. 12-17.
213. Неволин И.В. (2023е) Перспективы развития модуля CRIS для передачи технологий // Цифровая экономика. – 2023. – № 4(25). – С. 19-22. – DOI 10.34706/DE-2023-04-02.
214. Неволин И.В. (2024а) Алгоритмический подход к информации в стоимостной оценке нематериального // Современные тенденции развития математики и её прикладные аспекты -2024 : Материалы XIII Международной научно-практической интернет-конференции, Донецк, 24 мая 2024 года. – Донецк: ФГБОУ ВО «Донецкий национальный университет экономики и торговли имени Михаила Туган-Барановского», 2024. – С. 8-12.

215. Неволин И.В. (2024b) Интеграция технологий искусственного интеллекта в информационные системы финансового // Цифровая экономика. – 2024. – № 3(29). – С. 21-31. – DOI 10.34706/DE-2024-03-03.
216. Неволин И.В. (2024c) Концептуальный подход к построению платформы для рынка данных // Цифровая экономика. – 2024. – № 2(28). – С. 31-37. – DOI 10.33276/DE-2024-02-03.
217. Неволин И.В. Сервисы для управления интеллектуальными правами на базе блокчейн-инфраструктуры // Цифровая экономика. – 2021. – № 1(13). – С. 27-33. – DOI 10.34706/DE-2021-01-02.
218. Неволин И.В., Грачев И.Д. Модель для анализа инновационной активности // Системное моделирование социально-экономических процессов : труды 42-ой Международной научной школы-семинара, Ростов-на-Дону, 01–06 октября 2019 года. – Воронеж: Воронежский государственный университет, 2019. – С. 312-315.
219. Неволин И.В., Соколов Н.А. Назначение цен на информирующие сигналы при передаче ноу-хау // Обзорение прикладной и промышленной математики. – 2013. – Т. 20, № 2. – С. 181-182.
220. Нейлор К. Как построить свою экспертную систему. – М.: Энергоатомиздат, 1991, 286 с.
221. Оптимизация качества. Сложные продукты и процессы. / Калинина Э.В., Лапига А.Г., Поляков В.В. и др. – М.: Химия, 1989, 256 с.
222. Орлов А.И. Организационно-экономическое моделирование: теория принятия решений. – М.: КНОРУС, 2010, 568 с.
223. Орлова Е.Р., Бочарова И.Е., Козенко Т.Е., Федотова А.М. Основные направления цифровизации в области здравоохранения // Информационные технологии и вычислительные системы. – 2023. – № 2. – С. 18-26. – DOI 10.14357/20718632230202
224. Орлова Е.Р., Тищенко А.В. Новые подходы к организации инновационных процессов в России на современном этапе // Труды Института

- системного анализа Российской академии наук. – 2024. – Т. 74, № 4. – С. 86-93. – DOI 10.14357/20790279240410.
225. Орловский С.А. Проблемы принятия решений при нечеткой входной информации. – М.: Наука, 1981, 208 с.
226. Оуэн Г. Теория игр / М.: Мир, 1971. – 230 с.
227. Первозванская Т.П., Первозваницкий А.А. Распределение централизованных ресурсов между многими предприятиями // Экономика и математические методы - 1966. - Т.2. - №5. - С. 682-689.
228. Первозваницкий А.А., Гайцгори В.Г. Декомпозиция, агрегирование и приближенная оптимизация. - М.: Наука, 1979. - 342с
229. Планкетт Л., Хейл Г. Выработка и принятие управленческих решений. – М.: Экономика, 1984, 187 с.
230. Подиновский В.В. Лексикографические задачи оптимизации М.: 1972.
231. Полтерович В. М. Формирование отечественных сетей добавленной стоимости // Журнал Новой экономической ассоциации. – 2024. – № 3(64). – С. 251-257. – DOI 10.31737/22212264_2024_3_251-257.
232. Поспелов Д.А. Моделирование рассуждений. – М.: Радио и связь, 1989, 184 с.
233. Райфа Г. Анализ решений / М.: Наука. – 1977. – 408с.
234. Рамонов А.В. Ожидаемая продолжительность здоровой жизни как интегральная оценка здоровья россиян // Экономический журнал Высшей школы экономики. – 2011. – Т. 15, № 4. – С. 497-518.
235. Розенмюллер И. Кооперативные игры и рынки / М.: Мир, 1974. 159с.
236. РСТ (2017) Защита информации. Разработка безопасного программного обеспечения. Общие требования. ГОСТ Р 56939-2016.
237. РСТ (2021) Системы искусственного интеллекта. Способы обеспечения доверия. Общие положения. ГОСТ Р 59276-2020.

238. РСТ (2022а) Информационные технологии. Большие данные. Техническое задание. Требования к содержанию и оформлению. ГОСТ Р 59925-2021.
239. РСТ (2022с) Информационные технологии. Комплекс стандартов на автоматизированные системы. Автоматизированные системы. Требования к содержанию документов. ГОСТ Р 59795-2021.
240. РСТ (2023) Информационные технологии. Искусственный интеллект. Структура управления процессами аналитики больших данных. ГОСТ Р ИСО/МЭК 24668-2022.
241. Руа Б. К общей методологии выработки и принятия решений. / В сб.: Статистические модели и многокритериальные задачи принятия решений. – М.: Статистика, 1979, с.123-167.
242. Руа Б. Классификация и выбор при наличии нескольких критериев (метод ЭЛЕКТРА). – В кн.: Вопросы анализа и процедуры принятия решений. – М.: Мир, 1976, с. 80-107.
243. Румянцев А. С. Модели формирования цены программных продуктов : специальность 08.00.13 "Математические и инструментальные методы экономики" : диссертация на соискание ученой степени кандидата экономических наук / Румянцев Алексей Сергеевич. – Санкт-Петербург, 2010. – 147 с.
244. Саати Т. Л. Принятие решений при зависимостях и обратных связях: Аналитические сети. — М.: Издательство ЛКИ, 2008, 360 с.
245. Саати Т.Л. Принятие решений. Метод анализа иерархий. – М.: Радио и связь, 1989, 316 с.
246. Серенко А.Ф. Заносные вспышки оспы / М.: Государственное издательство медицинской литературы. – 1962. – 120с.
247. Сотникова М.В., Белкин Ю.Д., Сотникова О.И., Казакова М.А. Закупка научного оборудования из средств грантов для центров коллективного пользования и уникальных научных установок // Экономика науки. – 2022. – Т. 8, № 3-4. – С. 186-201.

248. Стародубов В.И., Богданова Т.Г., Ходакова О.В., Люцко В.В., Котовский М.Ю., Армашевская О.В. Регистр научных разработок и применения цифровых продуктов, технологий в рамках специальности «Организация здравоохранения и общественное здоровье» // Свидетельство о государственной регистрации базы данных № 2024624568 Российская Федерация. – 2024
249. Таха Х. Исследование операций. В 2-х кн. – М.: Мир, 1985.
250. Тевелева О.В., Неволин И.В. Малые инновационные предприятия. 11 лет практики // Инновации. – 2021. – № 3(269). – С. 16-29. – DOI 10.26310/2071-3010.2021.269.3.003.
251. Трахтенгерц Э.А. Компьютерная поддержка принятия решений. – М.: СИНТЕГ, 1998, 376 с.
252. Федотова Г.В., Орлова Е.Р., Бочарова И.Е. Вопросы кибербезопасности цифровых финансовых сервисов // Информационные технологии и вычислительные системы. – 2022. – № 2. – С. 37-45. – DOI 10.14357/20718632220205.
253. Харкевич А. А. Теория информации. Оpozнание образов. / Избранные труды в трех томах. Т. III. М., «Наука», 1973 г., стр. 524.
254. Хейес-Рот Ф., Уотерман Д., Ленат Д. Построение экспертных систем. – М.: Мир, 1987, 441 с.
255. Чайнов А.В. К вопросу теории некапиталистических систем хозяйств / В кн. Чайнов А.В. Крестьянское хозяйство. Избранные труды. – М.: Экономика, 1989. сс.114-143.
256. Чернавская О. Д. Динамическая теория информации как базис естественно-конструктивистского подхода к моделированию мышления // Компьютерные исследования и моделирование. – 2017. – Т. 9. – №. 3. – С. 433-447.
257. Чесноков М. Ю. Поиск аномалий в задаче повышения качества открытых данных. Проблемы управления. 2019;3(0):53-60.

258. Шень А.Х. Понятие (α, β) -стохастичности по Колмогорову и его свойства, Докл. АН СССР, 1983, том 271, номер 6, 1337–1340
259. Эддоус М., Стэндфилд Р. Методы принятия решений / М.: Аудит, ЮНИТИ. – 1997. – 590 с.
260. Ягер, Рональд. Нечеткие множества и теория возможностей. Последние достижения. – М.: Радио и связь, 1986, 408 с.