

Тюкавкин Николай Михайлович

Самарский национальный исследовательский университет имени академика С.П.
Королева

Гнатышина Елизавета Игоревна

Самарский национальный исследовательский университет имени академика С.П.
Королева

**Методологический подход к форсайту инновационного развития регионального
промышленного комплекса на основе когнитивно-семантического анализа больших
данных**

Аннотация. Цель исследования — разработка методологического подхода к форсайту инновационного развития регионального промышленного комплекса на основе когнитивно-семантического анализа больших данных, интегрирующего анализ патентных ландшафтов и методы семантического моделирования для автоматизированного выявления перспективных технологических направлений и технологических разрывов. В работе использованы методы сбора и обработки больших данных, динамического тематического моделирования на основе латентного размещения Дирихле, а также методы количественной оценки семантической новизны текстовой информации. Систематизированы источники неструктурированных данных, релевантные для задач регионального форсайта: патентные базы, наукометрические порталы, информационные потоки, отражающие рыночный и общественный интерес к технологиям. Разработана концептуальная модель методологического подхода, включающая четыре этапа: сбор и предобработку разнородных данных, семантическую обработку и тематическое моделирование, построение динамической карты технологий и расчет интегральных показателей, поддержку принятия решений и выявление технологических разрывов. Предложена математическая модель расчета динамического индекса перспективности, объединяющего показатели патентной активности, наукометрической динамики и семантической энтропии. Формализован метод выявления технологических разрывов между общемировыми трендами и региональной специализацией, позволяющий объективно ранжировать направления для импортозамещения и опережающего развития. Предложенный подход создает основу для научно обоснованного инструментария поддержки принятия решений при определении приоритетов региональной промышленной политики, распределении ресурсов на НИОКР и формировании программ импортозамещения. Результаты могут быть использованы региональными органами власти, институтами развития и промышленными корпорациями.

Ключевые слова: форсайт, региональный промышленный комплекс, большие данные, патентный ландшафт, когнитивно-семантический анализ, инновационное развитие, технологическое прогнозирование.

Tyukavkin Nikolay Mikhailovich

Samara National Research University named after Academician S.P. Korolev

Gnatyshina Elizaveta Igorevna

Samara National Research University named after Academician S.P. Korolev

**Methodological approach to foresight of innovative development of the regional industrial
complex based on cognitive-semantic analysis of big data**

Abstract. The purpose of the study is to develop a methodological approach to foresight of innovative development of regional industrial complexes based on cognitive-semantic analysis of big data, integrating patent landscape analysis and semantic modeling methods for automated identification of promising technological areas and technological gaps. The work employs big data collection and processing, dynamic topic modeling based on latent Dirichlet allocation, and methods for quantifying semantic novelty of textual information. The study systematizes unstructured data sources relevant to regional foresight, including patent databases, scientometric portals, and information flows reflecting market and public interest in technologies. A conceptual model of the methodological approach has been developed, including four stages: collection and pre-processing of heterogeneous data, semantic processing and topic modeling, construction of a dynamic technology map and calculation of integral indicators, decision support and identification of technological gaps. A mathematical model for calculating a dynamic perspective index combining indicators of patent activity, scientometric dynamics and semantic entropy is proposed. The method of identifying technological gaps between global trends and regional specialization is formalized, allowing to objectively rank directions for import substitution and advanced development. The proposed approach enables development of scientifically grounded decision-support tools for prioritizing regional industrial policy, allocating R&D resources, and formulating import substitution programs. The results can be used by regional authorities, development institutions, and industrial corporations.

Keywords: foresight, regional industrial complex, big data, patent landscape, cognitive-semantic analysis, innovative development, technological forecasting.

Введение. В условиях продолжающегося экономического давления на Россию со стороны недружественных государств остро стоит вопрос необходимости форсированного импортозамещения, вследствие чего региональные промышленные комплексы сталкиваются с новым вызовом пересмотра стратегий инновационного развития [10]. Традиционные подходы к управлению уступают место потребности в проактивном предвидении технологических сдвигов. Однако существующий инструментарий форсайта и технологического прогнозирования, базирующийся преимущественно на экспертных оценках и экстраполяции ретроспективных данных, демонстрирует снижение своей эффективности [2]. Субъективизм экспертного мнения, инерционность мышления и невозможность оперативно обрабатывать большие объемы новой информации приводят к тому, что решения о поддержке тех или иных технологических направлений нередко принимаются с запозданием, когда потенциал роста уже реализован в других регионах или странах.

Дополнительно проблема усугубляется тем, что цифровая трансформация экономики порождает лавинообразный рост данных, которые могли бы служить основой для более точного прогнозирования, но остаются неиспользованными в силу отсутствия адекватных методологических подходов к их обработке и интерпретации. Патентные базы, наукометрические порталы, отраслевые новостные ленты и поисковые запросы содержат неструктурированную информацию о новых технологических трендах, однако традиционные методы регионального стратегирования не интегрируют эти источники в единую систему поддержки принятия решений. Возникает противоречие между потребностью региональных органов власти и промышленных предприятий в точном и оперативном инструменте выявления точек инновационного роста и ограниченностью существующего методологического аппарата, неспособного в полной мере использовать потенциал больших данных.

В рамках данного исследования под форсайтом понимается систематический процесс перспективного прогнозирования инновационно-технологического развития, основанный на использовании современного экономико-математического инструментария и когнитивно-семантических методов анализа больших данных, направленный на

формирование и выработку актуальных решений участниками регионального промышленного комплекса.

Целью данного исследования является разработка методологического подхода к форсайту инновационного развития регионального промышленного комплекса на основе когнитивно-семантического анализа больших данных, интегрирующего анализ патентных ландшафтов и методы семантического моделирования для автоматизированного выявления перспективных технологических направлений и технологических разрывов.

Для достижения поставленной цели в работе решаются следующие задачи: систематизировать источники больших данных, релевантные для прогнозирования инновационного развития регионального промышленного комплекса, и определить требования к их сбору и предобработке; разработать модель динамического тематического моделирования для идентификации скрытых технологических направлений на основе когнитивно-семантического анализа текстов патентных документов, научных публикаций и информационных потоков, отражающих рыночный и общественный интерес; предложить математическую модель расчета интегрального индекса перспективности технологического направления, объединяющего показатели патентной активности, наукометрической динамики и семантической энтропии; формализовать метод выявления технологических разрывов между общемировыми трендами и текущей специализацией региона, позволяющий объективно ранжировать направления для импортозамещения и опережающего развития; определить направления практического применения разработанного методологического подхода в деятельности региональных органов власти, институтов развития и промышленных предприятий.

Научная новизна исследования заключается в разработке методологического подхода к форсайту инновационного развития регионального промышленного комплекса на основе интеграции разнородных источников больших данных и методов когнитивно-семантического моделирования. В отличие от существующих подходов, предлагаемый подход включает экономико-математическую модель расчета динамического индекса перспективности, комбинирующего показатели патентной активности, наукометрической динамики и семантической энтропии, а также формализованный метод выявления технологических разрывов на региональном уровне с применением когнитивно-семантических инструментов анализа [2].

Обзор литературы и исследований. Фундаментальные основы технологического прогнозирования и формирования инновационной политики заложены в работах академика Глазьева С.Ю., разработавшего подход к теории смены технологических укладов и методологии долгосрочного технико-экономического развития. Значительный вклад в развитие методов форсайта внесен коллективом Института статистических исследований и экономики знаний НИУ ВШЭ под руководством Гохберга Л.М., где была адаптирована российская практика на основе международных подходов и предложены методики проведения форсайт-исследований на национальном и отраслевом уровнях.

Проблематика регионального инновационного развития и вопросы «умной специализации» территорий глубоко проработаны в трудах Земцова С.П., Бариновой В.А. и Клеповой Т.А., акцентирующих внимание на необходимости учета локальных компетенций и существующей инфраструктуры при выборе приоритетов технологического развития [3].

Также следует отметить работы российских исследователей для целей прогнозирования в сфере анализа больших данных и патентной аналитики, активно развивающих данное направление. Исследования Ена О.В. и Коцемира М.Н. посвящены использованию методов интеллектуального анализа текстов для формирования технологических дорожных карт и выявления перспективных направлений исследований. В работах Абдрахмановой Г.И. рассматриваются возможности применения патентных ландшафтов для оценки технологического потенциала отраслей и регионов [5].

Цифровое стратегирование промышленных систем, основанное на устойчивых экоинновационных и циркулярных бизнес-моделях, предложенное Глуховым В.В., Бабкиным А.В. и Шкарупетой Е.В., обосновывает внедрение моделей, построенных по принципу цикла «проектирование – производство – потребление – переработка» [6]. Куликова Т.А. и Балахонова Е.В. рассматривают применение современных методов прогнозирования для разработки сценариев будущего развития и формирования устойчивой бизнес-стратегии промышленных предприятий, что особенно актуально в условиях повышенной неопределенности [8].

Существенный вклад в развитие методов тематического моделирования применительно к анализу научно-технической информации внесен российскими специалистами в области компьютерной лингвистики и анализа данных, в частности Воронцовым К.В., разработавшим теоретические основы аддитивной регуляризации тематических моделей, что позволяет повышать интерпретируемость выделяемых тем. Проблематика цифровой трансформации промышленности и формирования киберфизических систем на региональном уровне исследуется в трудах Бабкина А.В. и Плотникова В.А., которые обосновывают необходимость интеграции цифровых платформ в контур управления инновационным развитием.

Анализ научной литературы позволяет выявить методологический пробел, заключающийся в отсутствии комплексных исследований, интегрирующих разнородные источники больших данных в единую методологию форсайта на основе когнитивно-семантических методов, адаптированную под специфику регионального промышленного комплекса. Большинство существующих работ либо ограничиваются анализом патентной статистики на макроуровне, либо фокусируются на инструментарии без привязки к задачам регионального стратегического планирования.

Материалы и методы. Предлагаемый подход базируется на трех концепциях, отражающих технологическую динамику, природу знаний и территориальную специфику инноваций. Их интеграция формирует теоретический базис для анализа региональных технологических систем.

Первая концепция – эволюционная теория. Технологическое развитие нелинейно и является результатом взаимодействия множества агентов в условиях неопределенности. Как показано в работах Глазьева и Хирооки, изменения носят кумулятивный характер, проходя фазы зарождения, диффузии и замещения [1,12]. Смена технологических укладов ведет к структурной перестройке экономики. Для регионального форсайта отсюда следуют два вывода: анализ должен быть динамическим и учитывать технологическую зависимость прошлых решений.

Вторая концепция — управление знаниями. В постиндустриальной экономике знания являются основным фактором производства. Патенты и статьи выступают носителями явного знания. Методы когнитивно-семантического анализа текстов позволяют работать с большими массивами данных, выявляя латентные тренды, в отличие от субъективных экспертных оценок [4, 7]. Концепция обосновывает использование патентных и наукометрических баз, а также различение явного и неявного знания, определяя границы применимости методологии.

Третья концепция – умная специализация территорий. Вместо дублирования глобальных приоритетов регион должен развивать уникальные направления исходя из своих компетенций. Концепция выполняет три функции: обоснование выявления технологических разрывов между мировыми трендами и региональной специализацией; учет ресурсных ограничений (кадры, инфраструктура); задание критерия практической значимости результатов.

Эти три концепции образуют взаимосвязанный каркас. Эволюционная теория объясняет нелинейность, управление знаниями дает инструменты измерения, умная специализация – нормативные критерии. Подход дополняет экспертные оценки количественными методами, повышая прозрачность решений.

Реализация концепций в этапах предлагаемого подхода: управление знаниями определяет первый этап, а именно выбор патентных и наукометрических источников; эволюционная теория лежит в основе второго этапа — динамического тематического моделирования; умная специализация задает логику четвертого этапа — выявления технологических разрывов с использованием когнитивно-семантических инструментов.

Подход основан на пяти принципах. Принцип проактивности: выявление трендов на ранних стадиях через динамическое тематическое моделирование. Принцип объективности: анализ данных патентных баз и наукометрики, а не экспертных мнений. Принцип гетерогенности источников: объединение патентов, публикаций и новостей для сбалансированной оценки. Принцип динамичности: циклический сбор данных для корректировки приоритетов. Принцип интерпретируемости: визуализация карт технологий и интегральных показателей для лиц, принимающих решения [11].

Информационную базу составили пять групп источников: патенты (ФИПС, WIPO), научные публикации (РИНЦ, Scopus, Web of Science), новости, данные поисковых систем (Google Trends, Яндекс.WordStat) и стратегические документы.

В работе использованы методы обработки неструктурированных текстов (токенизация, лемматизация), динамическое тематическое моделирование на основе латентного размещения Дирихле, оценка семантической новизны (расстояние Кульбака-Лейблера, индекс Жаккара), метод анализа иерархий Саати для весовых коэффициентов и визуализация динамических карт технологий.

Ход исследования. Разработанный методологический подход включает четыре взаимосвязанных этапа (рисунок 1).

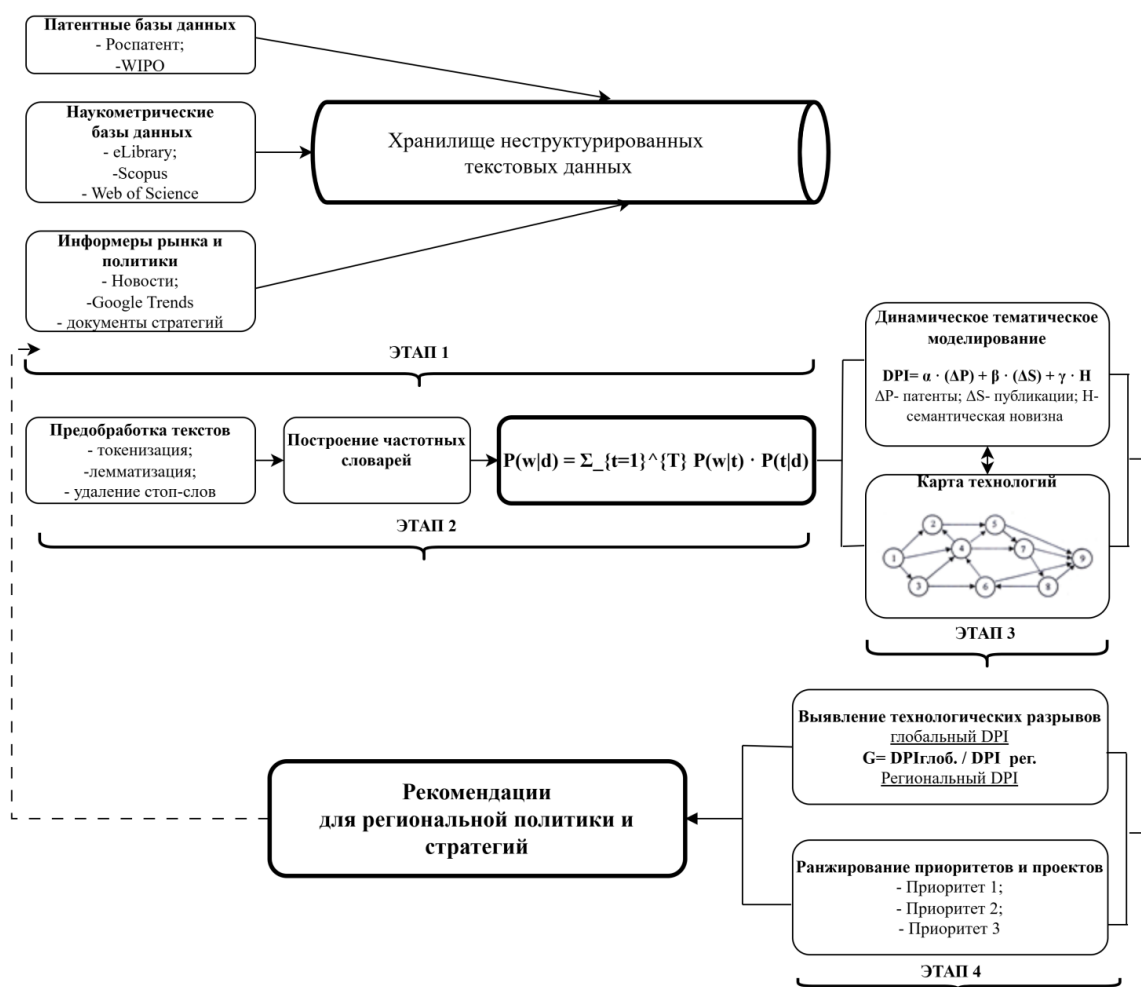


Рисунок 1 – Структурно-логическая схема методологического подхода к форсайту инновационного развития регионального промышленного комплекса на основе анализа больших данных.

Первый этап методологического подхода направлен на создание репрезентативной эмпирической базы, достаточной для последующего когнитивно-семантического анализа и тематического моделирования. Источники информации на данном этапе структурируются в три функциональных потока в соответствии с их природой и ролью в прогностической модели.

Первый поток представлен патентными данными. Патентная информация рассматривается в качестве наиболее надежного индикатора результатов научно-исследовательских и опытно-конструкторских работ, обладающих потенциальной коммерческой ценностью. Патентные документы проходят экспертизу на мировую новизну и содержат формализованные описания технических решений, что обеспечивает высокий уровень стандартизации и достоверности данных. Источниками патентной информации выступают базы Федерального института промышленной собственности, аккумулирующие сведения о российских патентах, а также международная система PATENTSCOPE Всемирной организации интеллектуальной собственности.

Второй поток – наукометрические данные. Это научные публикации в рецензируемых изданиях. Патенты фиксируют прикладные разработки на стадии готовности к коммерциализации. Научные статьи, напротив, отражают более ранние этапы – фундаментальные и поисковые исследования. Публикационная активность обычно предшествует патентной. По этой причине наукометрические данные служат важным источником для выявления формирующихся трендов. В данном исследовании использованы три основные базы. Российский индекс научного цитирования дает наиболее полный охват отечественных публикаций. Международные базы Scopus и Web of Science позволяют сопоставлять региональную динамику с глобальными трендами.

Третий поток составляют данные о рыночном и общественном интересе к технологиям, а также об институциональных условиях их развития. Этот поток объединяет три группы источников. Первая группа – новостные сообщения отраслевых информационных агентств и специализированных средств массовой информации. Они отражают текущую конъюнктуру и информационную повестку. Вторая группа – данные поисковых систем, которые агрегируют сервисы Google Trends и Яндекс.WordStat. Эти данные позволяют количественно оценить динамику запросов по технологическим терминам и зафиксировать уровень общественного внимания к отдельным направлениям. Третья группа – документы стратегического планирования федерального и регионального уровня: стратегии социально-экономического развития, государственные программы, отраслевые планы импортозамещения.

Сбор данных из перечисленных источников осуществляется в автоматизированном режиме. Для патентных и наукометрических баз используются программные интерфейсы приложений (API), предоставляемые операторами соответствующих платформ. Для новостных лент и открытых веб-ресурсов применяются технологии веб-скрапинга. По каждому документу фиксируются три обязательных элемента: полное текстовое содержание, дата публикации или регистрации, а также метаданные (источник происхождения, тип документа, сведения об авторах и аффилиации, классификационные индексы при наличии).

Предобработка текстовых данных включает три последовательные операции. Первая операция – токенизация, то есть разбиение сплошного текста на отдельные лексические единицы – токены. Вторая операция – лемматизация, обеспечивающая приведение каждого слова к его нормальной словарной форме. Данная операция имеет критическое значение для русскоязычных текстов в силу развитой морфологии русского языка, предполагающей многообразие словоформ. Третья операция – удаление стоп-слов, то есть исключение единиц, не несущих смысловой нагрузки: предлогов, союзов,

местоимений, а также общеупотребительных слов, встречающихся практически во всех документах и не обладающих дискриминативной способностью.

Результатом первого этапа является структурированное хранилище предобработанных текстов. Каждая текстовая единица в хранилище снабжена временной меткой, атрибутами источника и результатами выполненных операций нормализации. Данное хранилище служит входным массивом для второго этапа — когнитивно-семантической обработки и тематического моделирования.

Второй этап связан с извлечением из текстов скрытых тематических структур. Центральным элементом выступает модифицированная модель латентного размещения Дирихле (LDA), адаптированная для работы с динамическими потоками данных [7].

Математическая формализация тематического моделирования исходит из следующего допущения. Каждый документ из множества D есть случайная смесь латентных тем. Каждая тема задается распределением вероятностей по словам из множества V . Вероятность встретить слово w в документе d вычисляется по формуле:

$$P(w|d) = \sum_1^T P(w|t) \cdot P(t|d) \quad (1)$$

Параметры модели находят методом максимизации правдоподобия для наблюдаемой коллекции документов. Каждую выделенную тему интерпретируют как технологическое направление.

Чтобы учесть изменение тематик во времени, применяют динамическое тематическое моделирование. Весь временной интервал разбивают на дискретные промежутки. Для каждого промежутка строят отдельную тематическую модель. Параметры моделей соседних срезов связывают вероятностными зависимостями.

Количественную оценку перспективности технологических направлений дает интегральный показатель – динамический индекс перспективности (DPI):

$$DPI_{\{i,j\}} = \alpha \cdot \left(\frac{\Delta P_{\{i,j\}}}{\bar{P}} \right) + \beta \cdot \left(\frac{\Delta S_{\{i,j\}}}{\bar{S}} \right) + \gamma \cdot H_{\{i,j\}} \quad (2)$$

Где: $\Delta P_{\{i\}}$ – относительный темп прироста патентной активности по направлению i ; $\Delta S_{\{i\}}$ – относительный темп прироста числа научных публикаций по направлению i ; $H_{\{i\}}$ – индекс семантической новизны; α , β и γ – весовые коэффициенты, причем $\alpha + \beta + \gamma = 1$.

Определение весовых коэффициентов α , β и γ осуществляется с использованием метода анализа иерархий Т. Саати. Для этого формируется экспертная группа численностью не менее трех человек, представляющих научное сообщество, региональные органы власти и промышленный бизнес. Эксперты выполняют попарное сравнение значимости трех компонент интегрального индекса: патентной активности, публикационной динамики и семантической новизны. На основе матриц парных сравнений рассчитываются индивидуальные векторы приоритетов. Итоговые весовые коэффициенты определяются как среднее арифметическое этих векторов при условии согласованности экспертных оценок. Согласованность считается приемлемой при значении коэффициента консистентности, не превышающем 0,1.

Для количественной оценки технологического разрыва между глобальным и региональным уровнями предлагается индекс разрыва:

$$G_i = \frac{DPI_{\{global,i\}}}{(DPI_{\{local,i\}} + \varepsilon)} \quad (3)$$

В данной формуле используется константа ε , принимающая значение 0,001. Она вводится для исключения ситуации деления на ноль в случае отсутствия региональных значений показателя по отдельным технологическим направлениям. Значение индекса G_i больше единицы фиксирует наличие технологического разрыва, то есть отставание региона от глобальных трендов. Значение G_i , не превышающее единицу, свидетельствует о том, что регион находится на уровне мировых тенденций или опережает их.

В табл. 1 представлено сравнение разработанного подхода с традиционными методами форсайта [2].

Таблица 1 – Сравнительная характеристика традиционного и предлагаемого подходов к форсайту

Аспект	Традиционный подход	Предлагаемый подход
Источники данных	Экспертные оценки, узкий круг статистических данных	Разнородные массивы неструктурированных данных
Методы анализа	Экспертные панели, экстраполяция трендов	Когнитивно-семантический анализ, динамическое тематическое моделирование
Учет динамики	Статичные карты технологий	Непрерывный мониторинг, динамические траектории
Выявление приоритетов	Субъективные оценки экспертов	Интегральный индекс перспективности, индекс разрыва

Источник: разработано авторами

Для демонстрации практической реализуемости предложенного методологического подхода проведена апробация на условном примере. Необходимо подчеркнуть, что приводимые ниже числовые значения имеют исключительно иллюстративный характер. Они не являются результатом полномасштабного эмпирического исследования и не могут служить основанием для управленческих решений без дополнительной калибровки параметров, расширения временного горизонта и углубленной фильтрации данных. Цель данной апробации – показать порядок применения математического аппарата и логику интерпретации получаемых результатов.

В качестве объектов анализа выбраны три технологических направления, актуальных для многих региональных промышленных комплексов: «искусственный интеллект в промышленности», «аддитивные технологии» и «водородная энергетика». Выбор этих направлений обусловлен их присутствием в публичном дискурсе, доступностью открытых данных и различиями в стадиях жизненного цикла – от зарождающихся до относительно зрелых.

Сбор данных осуществлялся из открытых источников: патентной базы ФИПС, наукометрической платформы eLibrary.ru и сервиса поисковых запросов Яндекс.WordStat. Временной интервал охватил период 2021–2025 годов. Для каждого направления рассчитаны темпы прироста патентной активности и наукометрической динамики, а также индекс семантической новизны на основе когнитивно-семантического анализа ключевых терминов. Весовые коэффициенты в формуле DPI приняты равными $\alpha = 0,3$, $\beta = 0,3$, $\gamma = 0,4$. Данные значения получены экспертным путем с использованием метода анализа иерархий и отражают приоритет семантической новизны как наиболее раннего индикатора зарождающихся трендов. Результаты расчетов сведены в таблицу 2.

Таблица 2 – Результаты расчета динамического индекса перспективности (2025 г.)

Технологическое направление	$\Delta P/\bar{P}$	$\Delta S/\bar{S}$	H	DPI
Искусственный интеллект в промышленности	0,42	0,38	0,51	0,44
Аддитивные технологии	0,28	0,31	0,35	0,32
Водородная энергетика	0,35	0,33	0,42	0,37

Источник: рассчитано авторами

Как следует из таблицы, наибольшее значение динамического индекса перспективности (0,44) зафиксировано для направления «искусственный интеллект в промышленности». Это обусловлено высокой патентной активностью, значительным ростом числа публикаций и высокой семантической новизной, отражающей активное формирование нового понятийного аппарата. Для направления «аддитивные технологии»

получено наименьшее значение – 0,32. Это указывает на возможный переход данной области из фазы активного роста в фазу зрелости.

Далее выполнен расчет индекса технологического разрыва для условного региона с традиционной промышленной специализацией (машиностроение, металлургия). Глобальные и региональные значения DPI сопоставлены. По направлению «водородная энергетика» индекс разрыва превысил пороговый уровень. Данный факт трактуется как наличие технологического разрыва, который требует управленческих решений. Примеры таких решений – включение этого направления в программы импортозамещения или поддержка соответствующих НИОКР [10].

Проведенная апробация, хотя и носит демонстрационный характер, показывает, что предложенный инструментарий работает. Для условного региона получено следующее ранжирование технологических направлений по убыванию динамического индекса перспективности: искусственный интеллект в промышленности, затем водородная энергетика, затем аддитивные технологии. На основе этих результатов могут быть предложены следующие управленческие ориентиры. Приоритетная поддержка должна быть направлена на искусственный интеллект в промышленности, что предполагает создание региональных центров компетенций. Водородная энергетика, несмотря на более низкий индекс, требует федерального софинансирования в силу ее стратегического значения и текущего отставания региона. Аддитивные технологии, занявшие последнее место, нуждаются преимущественно в мониторинге и точечной поддержке при наличии конкретных запросов со стороны промышленности.

Чтобы использовать данный методологический подход на практике для конкретного региона, необходимо выполнить несколько условий. Нужно расширить временной горизонт анализа (как минимум до 10–15 лет). Патентные документы следует фильтровать по классам МПК, соответствующим отраслевой специализации региона. Весовые коэффициенты требуется калибровать с участием региональных экспертов. Также необходимо учитывать ресурсные ограничения – кадровые, финансовые, инфраструктурные.

Полученные количественные оценки не могут рассматриваться как готовые рекомендации без дополнительной калибровки.

Разработанный подход ориентирован на решение проблемы запаздывания управленческих решений. Высокое значение динамического индекса перспективности DPI не может рассматриваться как безусловная рекомендация к реализации соответствующего технологического направления. Окончательное определение приоритетов должно осуществляться с учетом кадровых, инфраструктурных и финансовых ограничений региона. Технологические направления, характеризующиеся высоким DPI, но низкой ресурсной обеспеченностью, целесообразно относить к зоне стратегического партнерства, предполагающей федеральную поддержку, а не к зоне немедленного импортозамещения. Традиционные подходы фиксируют уже свершившиеся изменения, тогда как предлагаемый методологический подход на основе когнитивно-семантического анализа обеспечивает обработку информации в режиме, близком к реальному времени.

С практической точки зрения предложенный инструментарий может быть использован для объективизации выбора приоритетов региональной промышленной политики, мониторинга появления новых рыночных и технологических ниш, а также для оптимизации распределения ресурсов поддержки инновационной деятельности [8].

Разработанный методологический подход имеет ряд ограничений. Качество результатов зависит от полноты и качества исходных данных: не все разработки проходят патентование, не все исследования отражаются в публикациях. Требуется настройка весовых коэффициентов α , β и γ под специфику региона и отрасли, что сохраняет долю субъективизма. Интерпретируемость результатов когнитивно-семантического моделирования может требовать привлечения экспертов для верификации. Кроме того, недостаточно учтены ресурсные ограничения регионального промышленного комплекса,

что может приводить к формированию рекомендаций по труднодостижимым технологиям.

Направления дальнейших исследований: интеграция нейросетевых моделей, в частности специализированных русскоязычных моделей типа RuBERT на основе архитектуры трансформеров, что позволит повысить качество когнитивно-семантического анализа за счёт учёта контекстуальных связей и глубинных смысловых паттернов в текстах патентных документов и научных публикаций; разработка цифровой платформы визуализации в виде интерактивного дашборда, обеспечивающего наглядное отображение динамической карты технологических трендов, фильтрацию направлений по отраслевым признакам и автоматическую генерацию рекомендаций для лиц, принимающих решения; расширение источников данных за счёт включения информации из социальных сетей, профессиональных форумов, экспертных блогов и краудфандинговых платформ, что позволяет фиксировать зарождающиеся тренды на более ранних стадиях по сравнению с официальными патентными и наукометрическими базами; учёт качества патентных документов и научных публикаций при расчёте индекса перспективности на основе таких критериев, как частота цитирования, наличие зарубежного патентования, импакт-фактор журналов и принадлежность авторов к ведущим исследовательским группам; сопряжение результатов форсайта с анализом ресурсной базы региона, включающим оценку кадрового потенциала, инфраструктурных возможностей, финансовых ограничений и институциональной среды для обоснованного ранжирования технологических направлений с учётом их реализуемости в конкретных региональных условиях [9].

Выводы

В результате проведенного исследования разработан методологический подход к форсайту инновационного развития регионального промышленного комплекса на основе когнитивно-семантического анализа больших данных. Отличительной особенностью предложенного подхода является интеграция трех взаимодополняющих инструментальных блоков: анализа больших данных, патентных ландшафтов и методов когнитивно-семантического моделирования. Такое сочетание позволяет преодолеть ограничения традиционных форсайт-методик, ориентированных преимущественно на экспертные оценки или анализ изолированных источников информации [2].

Первым научным результатом, полученным в ходе исследования, следует считать разработку архитектуры сбора и предобработки разнородных данных. Данная архитектура адаптирована к задачам регионального технологического прогнозирования и обеспечивает релевантность исходной информации для последующих этапов когнитивно-семантического анализа. В рамках теоретического обоснования методологического подхода определен ее концептуальный базис, включающий эволюционную экономическую теорию, концепцию управления знаниями и теорию умной специализации территорий [1, 3, 12], а также сформулирована система принципов.

Вторым результатом является модифицированная модель динамического тематического моделирования. В отличие от статических аналогов, данная модель позволяет не только идентифицировать скрытые технологические направления, но и прослеживать траектории их развития во времени. Возможность фиксации зарождающихся трендов на ранних стадиях достигается за счет разбиения временного интервала на дискретные промежутки и установления вероятностных связей между параметрами моделей соседних срезов [7].

Третьим результатом следует назвать введение интегрального показателя – динамического индекса перспективности (DPI) – и формализацию метода выявления технологических разрывов. Индекс DPI объединяет три компоненты: относительный темп прироста патентной активности, относительный темп прироста публикационной активности и индекс семантической новизны. Метод выявления технологических разрывов основан на сопоставлении глобальных и региональных значений DPI. Такое сопоставление дает возможность объективно ранжировать технологические направления

по степени отставания региона от мирового уровня и выделять зоны потенциального импортозамещения, с одной стороны, и зоны опережающего развития – с другой [4,10].

Четвертым результатом стала апробация предложенного инструментария на условном примере. Несмотря на демонстрационный характер расчетов, апробация подтвердила работоспособность математического аппарата и логическую непротиворечивость процедур, предусмотренных предлагаемым подходом. Полученные в ходе апробации количественные оценки, в частности, ранжирование трех технологических направлений по величине DPI, соответствуют экспертным представлениям о различных стадиях жизненного цикла анализируемых технологий, что косвенно свидетельствует о валидности предложенного подхода.

Пятым результатом определены направления дальнейшего совершенствования методологического подхода. В их число входят: учет ресурсных ограничений региона (кадровых, финансовых, инфраструктурных) при формировании итоговых рекомендаций; интеграция нейросетевых моделей типа Transformer, включая специализированные русскоязычные модели, такие как RuBERT, для повышения качества когнитивно-семантического анализа; разработка цифровой платформы в виде интерактивного дашборда для визуализации карты технологических трендов; расширение круга анализируемых источников за счет включения информации из социальных сетей и профессиональных форумов; развитие методов учета качества патентных документов и научных публикаций при расчете индекса перспективности [9].

Предложенный подход требует наличия экспертной группы для калибровки весовых коэффициентов, доступа к патентным и наукометрическим базам данных, а также обязательного последующего сопоставления полученных результатов с ресурсной обеспеченностью региона. Разработанный подход не претендует на статус окончательного и завершеного решения проблемы технологического прогнозирования на региональном уровне. Он представляет собой открытую систему, допускающую дальнейшее развитие и адаптацию по мере появления новых методов анализа данных, расширения доступных источников информации и накопления опыта практического применения. Вместе с тем уже на текущем этапе он обладает достаточной степенью готовности для использования в деятельности региональных органов власти, институтов развития и промышленных корпораций, заинтересованных в повышении обоснованности стратегических решений в сфере инновационного развития.

Практическая значимость исследования заключается в создании теоретической и методической базы для разработки прикладного инструментария на основе когнитивно-семантических методов. Использование данного инструментария позволяет повысить обоснованность решений при распределении ресурсов на научно-исследовательские и опытно-конструкторские работы, формировании программ импортозамещения, а также при определении приоритетов региональной промышленной политики [10]. В отличие от традиционных подходов, опирающихся на экспертные оценки и ретроспективный анализ, предложенный методологический подход на основе когнитивно-семантического анализа больших данных обеспечивает переход к количественно обоснованному, динамически обновляемому и адаптируемому к региональной специфике инструменту поддержки принятия решений.

Список источников

1. Глазьев С.Ю., Косакян Д.Л. Состояние и перспективы формирования 6-го технологического уклада в российской экономике // Экономика науки. 2024. Т. 10, № 2. С. 11–29.
2. Гохберг Л.М. Новые тенденции в российской практике Форсайт-исследований // Форсайт. 2009. Т. 3, № 3. С. 5.
3. Земцов С.П., Баринаева В.А., Семенова Р.И. Государственная поддержка высоких технологий и инноваций в России // Инновации. 2019. № 3 (245). С. 33–44.

4. Ена О.В. Радары для многопараметрической оценки технологий. Сценарии патентной аналитики // Вестник ФИПС. 2024. Т. 3, № 1 (7). С. 12–29.
5. Абдрахманова Г.И., Ковалева Г.Г. Тенденции развития информационных и коммуникационных технологий // Форсайт. 2009. Т. 3, № 4. С. 44–55.
6. Глухов В.В., Бабкин А.В., Шкарупета Е.В. Цифровое стратегирование промышленных систем на основе устойчивых эоинновационных и циркулярных бизнес-моделей в условиях перехода к Индустрии 5.0 // Экономика и управление. 2022. Т. 28, № 10. С. 1006–1020.
7. Ирхин И.А., Воронцов К.В. Сходимость алгоритма аддитивной регуляризации тематических моделей // Труды института математики и механики УрО РАН. 2020. Т. 26, № 3. С. 56–68.
8. Куликова Т.А., Балахонова Е.В. Форсайт в управлении развитием промышленных предприятий // Модели, системы, сети в экономике, технике, природе и обществе. 2023. № 2 (46). С. 42–61.
9. Бабкин А.В., Михайлов П.А. Цифровые платформы в экономике: понятие, сущность, классификация // Вестник Академии знаний. 2023. № 1 (54). С. 25–36.
10. Фролов И.Э., Борисов В.Н., Ганичев Н.А. Проблемы перехода к инновационно-модерному развитию российской экономики в условиях форсированного импортозамещения // Проблемы прогнозирования. 2023. № 4 (199). С. 67–81.
11. Porter A.L., Cunningham S.W. Tech mining: Exploiting new technologies for competitive advantage. New York: Wiley, 2004. 400 p.
12. Hirooka M. Innovation Dynamism and Economic Growth. A Nonlinear Perspective. Cheltenham: Edward Elgar, 2006. 448 p.

Сведения об авторах

Тюкавкин Николай Михайлович, заведующий кафедрой экономики инноваций, Самарский национальный исследовательский университет имени академика С.П. Королева, Самара, Россия.

Гнатышина Елизавета Игоревна, доцент кафедры экономики инноваций, Самарский национальный исследовательский университет имени академика С.П. Королева, Самара, Россия.

Information about the author

Tyukavkin Nikolay Mikhailovich, Head of the Department of Innovation Economics, Samara National Research University named after Academician S.P. Korolev, Samara, Russia.

Gnatyshina Elizaveta Igorevna, Associate Professor of the Department of Innovation Economics, Samara National Research University named after Academician S.P. Korolev, Samara, Russia.