

Развитие Базы знаний ИЭОПП СО РАН как инструмента решения исследовательских задач¹

А.В. Костин

УДК 338.984

DOI: 10.30680/ECO0131-7652-2024-6-106-117

Аннотация. В статье на примере Института экономики и организации промышленного производства СО РАН рассматривается опыт разработки базы знаний научного учреждения с учётом поставленных исследовательских задач, демонстрируются точки её применения. Формируемая База знаний направлена на аккумулирование, анализ и структурирование первичных данных о ресурсах, объектах, технологиях, состоянии и развитии макрорегиона Азиатская Россия, в целях достижения качественно нового уровня понимания и представления происходящих в ней социально-экономических процессов. Помимо нескольких блоков данных в её состав входит модельный комплекс, ориентированный на расширение исследовательского инструментария для анализа, прогнозирования данных, их динамики и взаимосвязей. ГИС-интерфейс Базы знаний позволяет развивать методы ГИС-моделирования и прогнозирования социально-экономических показателей развития территорий.

Ключевые слова: База знаний; ГИС-моделирование; накопление знаний;
когнитивное моделирование; очистка данных

Введение

Еще 30 лет назад эмпирические исследования показали, как сильно экономика зависит от накопленных знаний [Foray, Lundvall, 1996]. Последние являются важным стратегическим активом и необходимым элементом успеха в высококонкурентной динамичной среде. С тех пор объем и значимость знаний многократно возросли. Но существенно прогрессировали и инструменты работы с ними, расширились исследовательские возможности. Развитие цифровых технологий и методов анализа больших данных способствовали повышению точности и скорости математических вычислений. Благодаря достижениям в области машинного обучения и искусственного интеллекта удалось автоматизировать решение множества задач.

В этих условиях концепция исследовательской работы претерпевает изменения, и применение цифровых технологий становится важным элементом не только в технических, но и в социальных науках. В частности, для экономистов и социологов важны владение инструментами сбора, хранения и структурирования данных, практические навыки эконометрии, математического и ГИС-моделирования.

Всё вышеперечисленное обусловило развитие новых подходов к анализу и структурированию данных в рамках разработки Базы знаний Института экономики

¹ Статья подготовлена в рамках выполнения проекта Государственного задания 5.6.3.2. (FWZF-2024-0001) «Экспертно-аналитические, организационные и методические составляющие системы индикативного планирования научно-технологического и сбалансированного пространственного развития России при реализации крупных инвестиционных проектов».

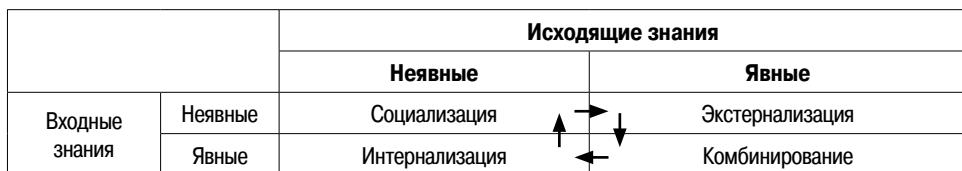
Развитие Базы знаний ИЭОПП СО РАН как инструмента решения исследовательских задач

и организации промышленного производства СО РАН с целью преобразования исходных данных о ресурсах, объектах, технологиях, состоянии и развитии макрорегиона Азиатская Россия, и достижения качественно нового уровня понимания происходящих в ней социально-экономических процессов.

В основе всех определений Базы знаний лежит понятие самого феномена «знание». В философии «знание есть такой результат познавательной деятельности, который обладает непреходящей истинностью, может быть логически или фактически обоснован и допускает эмпирическую или практическую проверку». В теории управления «знания – это комбинация данных и информации, к которой добавлены мнения, навыки и опыт экспертов, что дает в результате ценный актив, который может быть использован при принятии решений» [Тузовский, 2007].

Знания можно разделить на явные и неявные [Тузовский и др., 2005; Болбаков, 2015]. Первые логически организованы, формализованы, объективизированы и могут передаваться от одного субъекта другому, а вторые не сформулированы в явной форме, могут быть неосознанными, не допускают полной экстериоризации и потому трудно поддаются передаче. Развитие технологий обработки информации приводит к изменению границ между явными и неявными знаниями. Однако при этом возрастают значимость навыков и компетенций, связанных с отбором и эффективным использованием данных.

Существуют исследования, показывающие процесс перехода неявных знаний в явные и наоборот [Нонака, Такеучи, 2011] (рис. 1).



Источник. Трофимова Л.А., Трофимов В.В. Управление знаниями: учебное пособие. СПб.: Изд-во СПбГУЭФ. 2012.

Rис. 1. Спираль знаний Нонака и Такеучи

В предложенной схеме *социализация* – это процесс передачи неявных знаний от одного человека другому в ходе взаимодействия и обсуждения; *экстернализация* – преобразование неявных индивидуальных знаний в явные путём их формулирования и фиксации; *комбинирование* (формирование системного знания) – это процесс накопления, структурирования и распространения существующих знаний; *интернализация* – это трансформация явных знаний в неявные (например, путём обучения на практике и формирования нового прикладного знания, или посредством непрерывного индивидуального и коллективного размышления, благодаря способности видеть связи и распознавать закономерности между разными фактами, идеями и концепциями и т.п.).

Процессы пополнения и трансформации явных и неявных знаний происходят постоянно, с нарастающей интенсивностью, что породило концепцию спирали знаний. Таким образом, Базу знаний организаций можно охарактеризовать как динамически развивающуюся, постоянно усложняющуюся систему. Для работы с ней используются различные формы ввода/вывода данных, инструменты и программные интерфейсы.

Полный жизненный цикл таких систем, как База знаний от создания до эксплуатации и обслуживания, включает ряд последовательных этапов [Felfernig, Wotawa, 2013; Hou et al., 2023].

1. Определение целей и задач базы и построение концептуальной схемы.
2. Сбор и анализ информации (данных) в соответствии с этими целями и задачами.
3. Структурирование собранных и вновь поступающих данных на основе концептуальной схемы, с учетом анализа информации.
4. Разработка пользовательского интерфейса для сбора, диспетчеризации, вывода данных и пр.
5. Тестирование и отладка моделей и методов обработки данных.
6. Внедрение, поддержка, масштабирование и развитие.

В рамках концепции спирали знаний предполагается, что база постоянно будет подвергаться доработке, расширению и включению новых элементов. Это позволит обогащать ее явные и неявные компоненты, обеспечивая эволюционное развитие всей системы и ее адаптацию к изменяющимся требованиям и условиям.

Анализ поставленных задач, текущих накопленных знаний и концептуальных схем существующих подходов позволил разработать общую схему, описывающую взаимосвязи и структуру Базы знаний ИЭОПП для научно-исследовательских задач (рис. 2).



Рис. 2. Структура Базы знаний, адаптированной для решения научно-исследовательских задач ИЭОПП СО РАН

Общая структура базы данных

Основой (костяком) разрабатываемой Базы знаний является база данных. Она разбита на несколько блоков, необходимых для решения задач, стоящих перед институтом: информационное обеспечение (компании, секторы, комплексы, регионы, кластеры); проекты – инвестиционные, программные, институциональные; полезные ископаемые (подземные воды, углеводородное сырье и твердые ПИ) и в этом же блоке – ресурсные кластеры, компании, участки недр, состояние геологоразведочных и добычных работ; технологии. Эти блоки планируется постоянно развивать и обновлять.

Указанные данные имеют разнородный характер (типы, форматы, логику), поскольку помимо объективных характеристик объектов накладывает отпечаток их происхождение – они поступают от разных организаций и ведомств, которые структурируют информацию исходя из собственных задач и потребностей. Поэтому для каждого блока была разработана оригинальная структура.

Входящая информация в блоках редактировалась экспертами. Во избежание ошибок и случайных утрат создана отдельная база данных с сохранением всех изменений для возможности восстановления истории и анализа экспертной деятельности. Охарактеризуем отдельные блоки системы. Техническое описание структуры данных приводится в работе автора [Костиц, 2023].

Блок данных о компаниях представляет собой комплекс информации из разных источников (СПАРК-Интерфакс, list-org, РБК-компании и т.д.). В отличие от официальной статистики Росстата такого рода сведения имеют высокую погрешность.

Так, по мнению экспертов ИЭОПП СО РАН, довольно часто встречаются ошибки, связанные с неверным вводом единиц измерения в первоисточник (что отражается, например, в резких колебаниях размеров выручки), нередки также расхождения в цифрах активов и пассивов, расчета прибыли и т.д. Для такого рода информации о компаниях была разработана специальная система верификации. Данные предварительно направляются для проверки экспертным группам, сопоставляются с информацией из альтернативных источников и лишь затем включаются в единую систему. Такой подход позволяет отсеивать недостоверную информацию, избегать накопления ошибок, способных исказить агрегированные показатели при проведении пространственной аналитики.

Блок инвестиционных проектов содержит информацию об инвестиционных проектах различного масштаба, находящихся на разных стадиях реализации (планирование, предпроектные проработки, проектирование, подготовка к строительству, строительство, модернизация), располагающихся как в азиатской, так и в европейской частях России. Основными источниками данных послужили платформы «Инвестиционные проекты»², B2b Global и «Росинфра».

Наличие нескольких источников может приводить к дублированию, несогласованности информации, кроме того, здесь также нередко встречаются ошибки

² Право использования ПО «Инвестиционные проекты» предоставлено ООО «ПКР Аналитика» по Лицензионному договору № 119–11/21, заключенному при финансовой поддержке Минобрнауки РФ в рамках выполнения работ по научному проекту «Социально-экономическое развитие Азиатской России на основе синергии транспортной доступности, системных знаний о природно-ресурсном потенциале, расширяющегося пространства межрегиональных взаимодействий» (регистрационный номер заявки 2020–1902–01–377).

и неточности. Поэтому в структуру блока добавлен идентификатор источника данных, и пользователь получает возможность фильтровки сведений по нему. В дальнейшем предполагается введение процедуры экспертного согласования проектов.

В блоке технологий собирается информация об особенностях технологических процессов нефте-, газопереработки и переработки твердых полезных ископаемых. При этом обязательно учитывается экономическая составляющая технологических процессов. Ключевые элементы структуризации данных – входные (сырье) и выходные продукты и компании, в которых реализуется технология. Для каждой технологии добавлено техническое описание. Этот блок пополняется отраслевыми специалистами и является ключевым для оценки крупных инвестиционных проектов.

Модельно-аналитическая составляющая Базы знаний

Развитие Базы знаний происходит не только с накоплением новых данных, но также с развитием модельного и аналитического аппарата, позволяющего генерировать и фиксировать новые знания, которые за счет свойства кумулятивности могут лечь в основу создания последующих слоев знаний.

Инкорпорирование модельного комплекса в общую структуру Базы знаний предполагает в первую очередь перенос в нее модельного аппарата, уже используемого в исследованиях ИЭОПП СО РАН, в частности – моделей КАМИН (комплексный анализ межотраслевой информации) и ОМВЕАР (оптимизационная модель взаимодействия Европейской и Азиатской России). Но вводится и новый для института инструментарий, позволяющий увеличить возможности исследователей и ускорить процесс базового анализа данных. Это, в частности, системы когнитивного и ГИС-моделирования, прогнозные модели на основе методов машинного обучения и нейронных сетей.

Так, был создан блок решения задач **когнитивного моделирования**, вычислительная схема которого включает два этапа. На первом из них в результате анализа данных³ происходят выделение информативной системы факторов и построение когнитивной схемы в виде ориентированного графа (знакового или взвешенного). Это позволяет оценить взаимосвязи разных факторов между собой. На втором этапе осуществляется импульсное воздействие на факторы путем задания значения в какой-либо вершине графа и отслеживаются сопутствующие изменения в других его вершинах. Задачи этого этапа уже не аналитические, а прогнозические – исследователь определяет, как изменится сложная система при изменении одного или нескольких управляющих факторов. Когнитивное моделирование применяется при формировании сценариев развития хозяйственных систем, оценок влияния инвестиционных импульсов на экономический рост, а также решения других исследовательских задач.

В рамках развития Базы знаний началось формирование системы анализа и прогнозирования с использованием **методов машинного обучения на основе**

³ Анализ исходных данных проводится методом главных компонент (PCA, Principal Component Analysis), при котором исходные данные проецируются на гиперплоскость меньшей размерности. Сильной стороной метода можно считать возможность выявления неочевидных и скрытых закономерностей и взаимосвязей между переменными. Кроме того, он позволяет понижать размерность пространства данных и строить ортогональный базис факторов.

Развитие Базы знаний ИЭОПП СО РАН как инструмента решения исследовательских задач

анализа больших данных. В частности, на этих принципах была создана модель первичных преобразований данных для анализа и прогнозирования межрегиональных и межотраслевых взаимодействий.

Использование методов машинного обучения значительно расширяет результативность экономических исследований. Однако в большинстве случаев для корректного и эффективного решения задач требуется провести тщательную настройку параметров и алгоритмов машинного обучения. Так, необходимо учитывать, что большинство алгоритмов последнего построены на основе временной статистики, и нестационарность данных может привести к неправильной параметризации модели. В ходе обучения данные разделяются на две подвыборки: тестовую и тренировочную. При наличии нестационарности в данных распределение в них может существенно различаться, что значительно снижает качество моделей и точность прогнозов. Кроме того, в машинном обучении при поиске взаимосвязей в случае гетероскедастичности некоторые наблюдения могут получить больший вес, чем другие, что приведет к неверным выводам. Потому для этого инструмента был разработан блок анализа данных на стационарность, гетероскедастичность и алгоритмы решения этих проблем.

Важный этап при выборе факторов для прогнозных моделей – первичное выявление долгосрочных взаимосвязей показателей (коинтегрированности параметров). Для этого используется тест Педрони на коинтеграцию, а также матрица связанных показателей на основе теста Гренджера, в которой оцениваются причинно-следственные зависимости факторов. В результате формируются коинтеграционные и корреляционные матрицы взаимосвязей показателей.

Следующим шагом стало создание комплекса моделирования и прогнозирования показателей на основе алгоритмов машинного обучения и нейронных сетей. На первом этапе в модельный пул были включены три модели машинного обучения и нейронных сетей: рекуррентная нейронная сеть (RNN); рекуррентная нейронная сеть с долгой краткосрочной памятью (LSTM); градиентный бустинг (Gradient boosting).

В результате была разработана система автоматического поиска взаимосвязей и прогнозирования, реализованная в виде оригинального программного продукта. Она обладает высокой степенью автоматизации и позволяет эффективно находить взаимосвязи между различными переменными и предсказывать прогнозные значения на основе имеющихся данных.

ГИС-интерфейс Базы знаний

Для анализа цепочек создания стоимости, основанных на крупных инвестиционных проектах, современная наука применяет пространственный анализ территорий, что особенно актуально для Азиатской России. Поэтому все исходные и аналитические данные, имеющиеся в информационных блоках, привязываются к ГИС-координатам. Это значительно расширяет модельный аппарат пространственного анализа, но требует создания ГИС-интерфейса в рамках Базы знаний. В ИЭОПП СО РАН ГИС-интерфейс является частью веб-интерфейса и реализуется на базе пакета Leaflet и системы NextGis, которые позволяют создавать собственные слои данных, используя стандартные методы моделирования ГИС-информации

На сегодня ГИС-интерфейс имеет следующие слои, накладываемые на растровую карту территории: инфраструктура; твердые полезные ископаемые (на рис. 3, в качестве примера показаны минералогические районы и узлы Азиатской России); углеводородное сырье; инвестиционные проекты; компании (на рис. 4 в качестве примера показано размещение компаний, относящихся по основной деятельности к ОКВЭД 72.1 Научные исследования и разработки в области естественных и технических наук); статистические показатели.

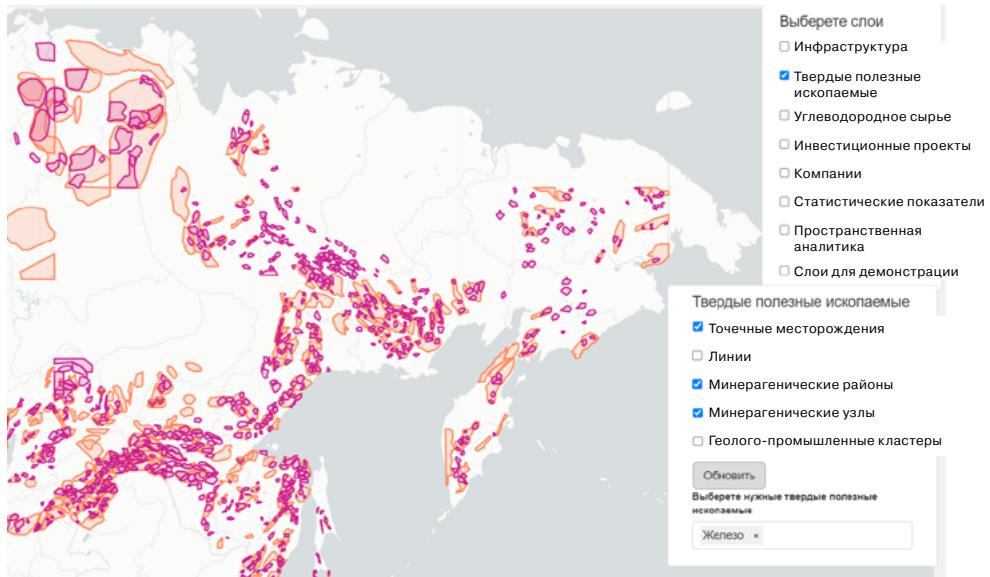


Рис. 3. Слой ГИС-интерфейса: твердые полезные ископаемые

Отображение большого количества точек в виде ГИС-слоёв требовало много серверных ресурсов, поэтому для оптимизации представления данных, относящихся к блокам инвестиционных проектов и компаний, была использована кластеризация, предварительно объединяющая множество близлежащих точек в один кластер, который и отображается на карте.

Пока для пользователей доступны только указанные слои, использующие загруженную в Базу знаний информацию, но в настоящее время разрабатывается интерфейс создания экспертических баз ГИС-данных, которые позволят специалистам самостоятельно рисовать карты, добавлять объекты и пояснения, в зависимости от исследовательских задач.

Разработка новых модулей ГИС-моделирования позволила создать системы пространственного анализа. Теперь исследователь может получить агрегированную информацию по выбранному участку территории по таким показателям, как инвестиционные проекты, компании (их специализация, объем выручки и пр.), полезные ископаемые, инфраструктура и пр., что дает новые возможности в анализе цепочек инвестиционных проектов.

Развитие Базы знаний ИЭОПП СО РАН как инструмента решения исследовательских задач

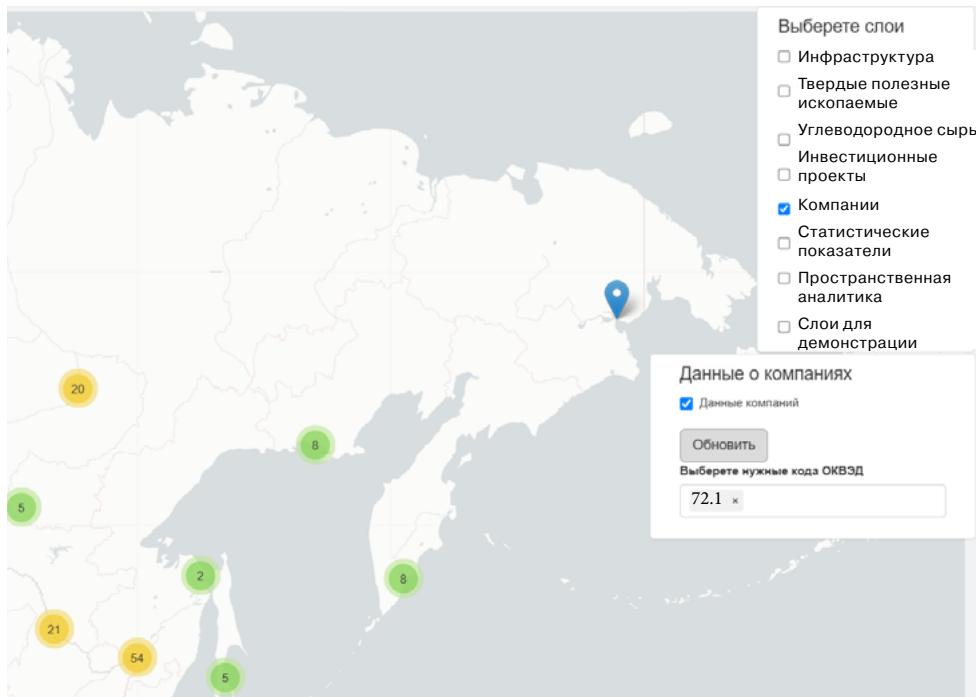


Рис. 4. Слой ГИС-интерфейса: данные о компаниях, ОКВЭД 72.1

На этой основе создается новая база данных, привязанная к территории, включающая информацию о ее инфраструктурном обустройстве, размещении и объеме природных богатств, концентрации компаний и инвестиционных проектов, а также социально-экономической статистики в разрезе макро-, мезо- и микроуровней. Такая база представляет собой очередной шаг на пути от анализа регионов или агломераций к детальному изучению конкретных территорий.

Примеры использования платформенных решений Базы знаний ИЭОПП

Хотя База знаний ИЭОПП находится на этапе разработки, она уже применяется для решения исследовательских задач. Ее описание и примеры использования представлены в работах сотрудников института [Новый импульс..., 2023; Модели..., 2023; Анализ..., 2024]. Так, автор принимал участие в работе, посвященной оценке инвестиционного импульса с использованием когнитивной модели [Ягольницер и др., 2022. С. 51–65]. С помощью факторного анализа инвестиционные проекты были распределены по кластерам «Качество жизни», «Добыча и переработка полезных ископаемых, машиностроение», «Энергетика и химическая промышленность» и «Металло- и деревообработка», для которых была построена когнитивная карта. При помощи импульсного моделирования было показано, что наибольший

положительный эффект⁴ для кластера «Качество жизни» дает импульс в фактор «Металло- и деревообработка», а минимальный – в фактор «Добыча и переработка полезных ископаемых, машиностроение» (рис 5).

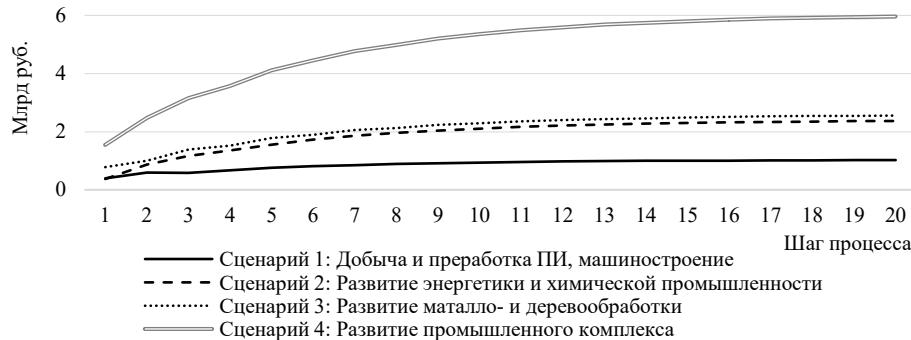


Рис. 5. Влияние различных сценариев распределения прироста инвестиций на прирост инвестиций в проекты качества жизни и строительства

В другом исследовании [Костина, Костин, 2023] База знаний института (конкретно – блок данных компаний) была использована для оценки взаимного влияния высокотехнологичных и наукоемких компаний и цифровизации городской среды. Работу осложняла необходимость аналитического среза высокотехнологичных и наукоемких компаний в пределе конкретных городов. С этой задачей помогли справиться аналитические модули, позволившие сформировать выборку по городам. Последующие расчеты подтвердили значимость гипотетически предполагавшейся зависимости между цифровизацией городской среды и успешной деятельностью высокотехнологичных и наукоемких компаний, в большей степени – для крупнейших городов.

Есть пример проведения первичного анализа масштабного инфраструктурного инвестпроекта с помощью системы ГИС-моделирования. Речь идет о проекте строительства Северо-Сибирской магистрали [Новый импульс..., 2023. С. 344–412], которая должна соединить железнодорожную сеть Ханты-Мансийского АО с Байкало-Амурской магистралью. С помощью системы ГИС-моделирования были определены характеристики минерально-сырьевой базы в зоне влияния будущей дороги, которые становятся доступными для освоения, охарактеризованы крупные компании и инвестпроекты, находящиеся на территории, что позволяет в общих чертах оценить потенциальный грузопоток, вероятную синергию для развития территории (рис. 6).

⁴ Подобные системы оценивают кумулятивный эффект инвестиционного импульса, когда вложения в одну отрасль по цепочке подтягивают за собой развитие взаимосвязанных отраслей и инфраструктуры.

Развитие Базы знаний ИЭОПП СО РАН как инструмента решения исследовательских задач

243 Инвестиционных проекта.

Суммарный объем инвестиций – 788 млрд руб.

Основные отрасли реализации:

обработка древесины и цветная металлургия



Rис. 6. Основные инвестиционные проекты в зоне влияния Северо-Сибирской магистрали и их потенциал

Заключение

Создаваемая в ИЭОПП СО РАН База знаний представляет собой инструмент исследования больших массивов структурированных данных на основе современных методов программирования и моделирования.

Уже на текущем этапе она позволяет проводить комплексный анализ территорий с учётом существующей инфраструктуры, действующих компаний, запасов природных ресурсов и потенциальных инвестиционных проектов, а также прогнозировать эффекты от реализации потенциальных инвестиционных проектов.

Последующее развитие Базы знаний, выраженное в увеличении и очистке базы данных, прогрессе системы ГИС-моделирования и сетевых моделей, реализации новых прогнозных моделей как на макро-, мезо- и микроуровне даст исследователям новые инструменты для аналитической работы и поможет увеличить точность оценки эффектов реализации инвестиционных проектов как части комплекса пространственно-распределенных цепочек создания стоимости.

Литература/References

Анализ и оценка процессов создания и развития в Азиатской России транспортной магистральной сети различного назначения / Под ред. А.А. Ширрова, О.В. Тарасовой. Новосибирск: Изд-во ИЭОПП СО РАН, 2024. 484 с.

Analysis and assessment of the processes of creation and development of a transport backbone network for various purposes in Asian Russia (2024). Ed. by A.A. Shirov, O.V. Tarasova. Novosibirsk: Publishing house of IEIE SB RAS. 484 p. (In Russ.).

Болбаков Р.Г. Отношение между явным и неявным знанием // Перспективы науки и образования. 2015. № 1 (13). С. 10–16.

- Bolbakov, R.G. (2015). The relation between explicit and implicit knowledge. *Perspectives of Science and Education*. No. 1 (13). Pp. 10–16. (In Russ.).
- Костин А.В. База знаний как инструментарий решения задач в экономической, социальной и производственной сферах / Под ред. М.А. Ягольнициера. Новосибирск: Изд-во ИЭОПП СО РАН, 2023. 146 с.
- Kostin, A.V. (2023). Knowledge base as a tool for solving problems in the economic, social and industrial spheres / ed. by M.A. Yagolnitzer. Novosibirsk. Publishing house of IEOPP SB RAS. 146 p. (In Russ.).
- Костина Е.А., Костин А.В. Умный город как фактор развития высокотехнологичных компаний // Регион: экономика и социология. 2023. № 3 (119). С. 84–110. DOI: 10.15372/REG20230304
- Kostina, E.A., Kostin, A.V. (2023) Smart city as a factor in the development of high-tech companies. *Region: Economics and Sociology*. No. (119). Pp. 84–110. (In Russ.). DOI: 10.15372/REG20230304
- Модели и методы прогнозирования: Азиатская Россия в экономике страны / Под ред. А.О. Баранова, В.И. Суслова; ИЭОПП СО РАН. Новосибирск: Изд-во ИЭОПП СО РАН, 2023. 436 с. DOI: 10.36264/978-5-89665-376-9-2023-012-436
- Models and forecasting methods: Asian Russia in the country's economy* (2023). Ed. by A.O. Baranov, V.I. Suslov; IEIE SB RAS. Novosibirsk: Publishing House of IEOPP SB RAS. 436 p. (In Russ.). DOI: 10.36264/978-5-89665-376-9-2023-012-436.
- Нонака И., Такеучи Х. Компания – создатель знания: зарождение и развитие инноваций в японских фирмах / Пер. с англ. А. Трактинского. М.: ЗАО «Олимп-Бизнес», 2011. 366 с.
- Nonaka, I., Takeuchi, H. (2011). *The company – creator of knowledge: the origin and development of innovations in Japanese firms* [trans. from the English by A. Traktinsky]. Moscow. Olymp-Business CJSC. 366 p. (In Russ.).
- Новый импульс Азиатской России: источники и средства развития. В 2-х т. Т. 1 / Под ред. В.А. Крюкова, Н.И. Суслова; ИЭОПП СО РАН. Новосибирск: Изд-во ИЭОПП СО РАН, 2023. 418 с. DOI: 10.36264/978-5-89665-375-2-2023-011-418
- A new impulse of Asian Russia: sources and means of development*. In 2 vol. Vol. 1 (2023). Ed. by V.A. Kryukov, N.I. Suslov; IEIE SB RAS. Novosibirsk: Publishing House of IEOPP SB RAS. 418 p. (In Russ.). DOI: 10.36264/978-5-89665-375-2-2023-011-418
- Тузовский А.Ф. Разработка систем управления знаниями на основе единой онтологической базы знаний // Известия Томского политехнического университета. Инжиниринг георесурсов. 2007. Т. 310. № 2. С. 182–185.
- Tuzovskiy, A.F. (2007). Development of knowledge management systems based on a unified ontological knowledge base. *Proceedings of Tomsk Polytechnic University. Georesource engineering*. Vol. 310. No. 2. Pp. 182–185. (In Russ.).
- Тузовский А.Ф., Чириков С.В., Ямпольский В.З. Системы управления знаниями (методы и технологии). Томск: Изд-во НТЛ, 2005. 260 с.
- Tuzovsky, A.F., Chirikov, S.V., Yampolsky, V.Z. (2005). *Knowledge management systems (methods and technologies)*. Tomsk: Publishing House of NTL. 260 p. (In Russ.).
- Ягольницер М.А., Овсянникова М.А., Костин А.В. Синергия инвестиционных проектов: когнитивный подход // Мир экономики и управления. 2022. Т. 22. № 3. С. 51–65.
- Yagolnitzer, M.A., Ovsyannikova, M.A., Kostin, A.V. (2022) Synergy of investment projects: a cognitive approach. *The World of Economics and Management*. Vol. 22. No. 3. Pp. 51–65. (In Russ.).
- Felfernig, A., Wotawa, F. (2013). Intelligent engineering techniques for knowledge bases. *AI Commun.* T. 26. No. 1. Pp. 1–2. (In Russ.).

Развитие Базы знаний ИЭОПП СО РАН как инструмента решения исследовательских задач

- Foray, D., Lundvall, B. (1996). *The knowledge-based economy: from the economics of knowledge to the learning economy*. OECD Documents: Employment and Growth in the Knowledge-Based Economy. Pp. 11–32. DOI:10.1016/B978-0-7506-7009-8.50011-2
- Hou, X. et al. (2023). A semantic data-driven knowledge base construction method to assist designers in design inspiration based on traditional motifs. *Advanced Engineering Informatics*. T. 56. P. 101987.

Статья поступила 01.11.2024
Статья принята к публикации 07.11.2024

Для цитирования: Костин А.В. Развитие Базы знаний ИЭОПП СО РАН как инструмента решения исследовательских задач // ЭКО. 2024. № 6. С. 106–117. DOI: 10.30680/ECO0131–7652–2024–6–106–117

Информация об авторе

Костин Андрей Владимирович (Новосибирск) – кандидат экономических наук, доцент. Институт экономики и организации промышленного производства СО РАН; Новосибирский национальный исследовательский государственный университет.

E-mail: Andrey.v.kostin@gmail.com; ORCID: 0000-0002-9337-3978

Summary

A.V. Kostin

Development of IEOPP SB RAS Knowledge Base as a Tool for Resolving Research Problems

Abstract. Progress in the field of modern technologies significantly expands the horizons of research activities, but it is associated with increasing requirements for both researchers and research organizations. This work is aimed at analyzing the experience of developing a knowledge base for a scientific institution, taking into account research tasks set, and demonstrating its application. The Knowledge Base of the Institute of Economics and Industrial Production Organization SB RAS is given as an example of such a structure. The purpose of this knowledge base is to transform the initial data on resources, facilities, technologies, state and development of the macro-region Asian Russia, and achieving a qualitatively new level of understanding of the socio-economic processes taking place in it. The developed Knowledge Base includes an updated database that provides researchers with access to structured data on socio-economic indicators, companies, investment projects, and technologies. In addition, it includes a model package focused on expanding research tools for analyzing, predicting data and data relationships. A key task of the Knowledge Base is analysis of the economic space, so it was decided to link the data to GIS coordinates and develop methods of GIS modeling and forecasting of socio-economic indicators of territorial development.

Keywords: Knowledge base; GIS-modeling; knowledge accumulation; cognitive modeling; data cleaning

For citation: Kostin, A.V. (2024). Development of IEOPP SB RAS Knowledge Base as a Tool for Resolving Research Problems. *ECO*. No. 6. Pp. 106–117. (In Russ.). DOI: 10.30680/ECO0131–7652–2024–6–106–117

Information about the authors

Костин, Андрей Владимирович (Новосибирск) – Кандидат экономических наук, Аспирант. Институт экономики и промышленной инженерии, СО РАН; Новосибирский государственный университет.

E-mail: Andrey.v.kostin@gmail.com; ORCID: 0000-0002-9337-3978