

УДК 330.3

М. В. Бузмакова

Нижегородский государственный университет им. Н.И. Лобачевского,
Нижний Новгород, e-mail: margowifi@gmail.com

И. Н. Полушкина

Нижегородский государственный университет им. Н.И. Лобачевского,
Нижний Новгород, e-mail: poirinic@mail.ru

ВЛИЯНИЕ БОЛЬШИХ ДАННЫХ НА ЭКОНОМИЧЕСКУЮ ДЕЯТЕЛЬНОСТЬ И ПРОБЛЕМЫ ИХ СТОИМОСТНОЙ ОЦЕНКИ

Ключевые слова: термин «большие данные», терминология больших данных, эволюция термина Big Data, проблемы Big Data, цифровая экономика, СНС-2025, структура СНС, особенности определения стоимостной оценки данных, стоимость данных.

Актуальность статьи обоснована тем, что экономика больших данных (Big Data) на современном этапе развития представляет собой перспективную и непредсказуемую область. В статье рассмотрено довольно новое понятие «большие данные» (Big Data) и исследуются аспекты его появления, содержания и развития. Авторы рассмотрели современные возможности применения Big Data в различных сферах жизнедеятельности, рассмотрены источники их получения и направления использования, раскрыты сопутствующие при этом проблемы и угрозы. Особое внимание уделено проблеме стоимостной оценки потоков больших данных. Авторами проанализированы имеющиеся подходы к оценке стоимости данных, их учет в Системе национальных счетов, приведены особенности определения стоимостной оценки данных. В заключении авторы делают вывод, что характеристики данных возможно, не обеспечивают точной оценки их экономической ценности. Тем не менее, они необходимы для получения надежных и сопоставимых оценок, которые могут быть дополнительно улучшены с помощью экономического или эконометрического анализа. Также обновленная СНС-2025 предоставит более точную, а главное, общеустановленную методику оценки больших данных для экономики.

М. V. Buzmakova

Lobachevsky Nizhny Novgorod State University, Nizhny Novgorod,
e-mail: margowifi@gmail.com

I. N. Polushkina

Lobachevsky Nizhny Novgorod State University, Nizhny Novgorod,
e-mail: poirinic@mail.ru

THE IMPACT OF BIG DATA ON ECONOMIC ACTIVITY AND THE PROBLEMS OF ITS VALUATION

Keywords: the term “big data”, the terminology of big data, the evolution of the term Big Data, the problems of Big Data, the digital economy, the SNA-2025, the structure of the SNA, the specifics of determining the valuation of data, the cost of data.

The relevance of the article is justified by the fact that the economics of big data (Big Data) at the current stage of development is a promising and unpredictable area. The article examines a rather new concept of “Big Data” and examines aspects of its appearance, content and development. The authors examined the modern possibilities of using Big Data in various spheres of life, considered the sources of their receipt and areas of use, and revealed the associated problems and threats. Special attention is paid to the problem of valuation of big data flows. The authors analyzed the available approaches to estimating the value of data, their accounting in the System of National Accounts, and presented the specifics of determining the value of data. In conclusion, the authors conclude that the characteristics of the data may not provide an accurate assessment of their economic value. However, they are necessary to obtain reliable and comparable estimates that can be further improved through economic or econometric analysis. Also, the updated SNA-2025 will provide a more accurate, and most importantly, a generally established methodology for evaluating big data for the economy.

Введение

В современном мире большие данные играют ключевую роль в развитии инновационной экономики. Они позволяют компаниям и организациям принимать более эффективные решения, оптимизировать процессы и повышать эффективность работы. Однако с развитием больших данных возникают и новые проблемы, связанные с их обработкой и анализом.

Цель работы – совершенствование теоретических и практических научных методов анализа «больших данных» в условиях цифровой трансформации экономики и структурировать её отдельные звенья.

Объектом исследования являются массивы Big Data в условиях цифровой трансформации экономики. Предметом исследования является систематизация современных подходов к стоимостной оценке больших данных и анализ их влияния на экономические процессы.

Теоретико-методологической основой исследования является применение общих методов научно-теоретического анализа для изучения рассматриваемых явлений и концепций.

В целях повышения качества исследования авторы использовали в статье следующие методы: логический, исторический, метод научной абстракции, методы анализа и синтеза. С их помощью были выявлены сущность и эволюция исследуемого явления, а также причинно-следственные связи между экономическими категориями и компонентами нового цифрового пространства.

Big Data сегодня – это не только массив информации, измеряемый сотнями терабайт, но и инструмент для всестороннего анализа, способный выявить закономерности, недоступные традиционным методам обработки данных и человеческому интеллекту. Big Data – междисциплинарная концепция, объединяющая исследования из различных областей науки.

История вопроса

Словосочетание «концепция больших данных» может показаться новомодной тенденцией, однако их влияние и анализ основываются на теориях, которые существуют уже несколько веков.

Приведём определение Oracle: «Большие данные – это чрезвычайно большие и сложные наборы данных, которыми сложно управлять или анализировать с помощью

традиционных инструментов обработки данных, в частности электронных таблиц. Большие данные включают в себя структурированные данные, например, базу данных о товарах или список финансовых операций; неструктурированные данные, например, публикации в социальных сетях или видео; а также смешанные наборы данных, например, те, которые используются для обучения больших языковых моделей для ИИ. Эти наборы данных могут включать в себя что угодно: от произведений Шекспира до таблиц с бюджетом компании за последние 10 лет» [9].

Кроме того, это слишком ценная информация, чтобы оставлять её без анализа. Большие данные позволяют извлекать полезную информацию из этого обширного массива данных, чтобы помочь организациям и целым странам создать более эффективную промышленность, быстрее внедрять инновации, зарабатывать больше денег и в целом выигрывать.

История о том, как данные стали «большими», началась за много лет до того, как «большие данные» стали популярным термином. Более 80 лет назад человечество столкнулось с первыми попытками количественно оценить темпы роста объёма данных или то, что Гил Пресс называл «информационным взрывом» в публикации «Очень короткая история больших данных» на сайте Forbes.com [5]. А в 1971 году Артур Миллер в книге «Нападение на частную жизнь» пишет, что «слишком многие обработчики информации, похоже, оценивают человека по количеству битов памяти, которые займёт его досье».

Термин Big Data возник в 1990-х годах, когда общество начало проявлять интерес к этому новому сектору, связанному с обработкой и анализом больших объёмов данных.

В 1997 году Майкл Кокс и Дэвид Элсворт опубликовали статью «Управление подкачкой данных по требованию для визуализации вне ядра» в сборнике материалов 8-й конференции IEEE по визуализации. Это была первая статья в цифровой библиотеке ACM, в которой термин «большие данные» используется в современном контексте.

В 1998 году Джон Мэши, главный научный сотрудник SGI, представил на конференции USENIX доклад под названием «Большие данные... и новая волна инфраструктурных технологий». Джон Мэши использовал этот термин в своих выступле-

ниях, поэтому ему приписывают авторство термина «большие данные» [15].

В 2005 году Тим О'Рейли опубликовал свою инновационную статью «Что такое Web 2.0?» Автор намеренно применил термин «большие данные» для обозначения больших массивов данных, с которыми практически невозможно работать и которые практически невозможно обрабатывать с помощью традиционных инструментов бизнес-аналитики.

В 2010 году появились программные продукты, основанные на концепции применения больших данных. К 2011 году ведущие поставщики информационных технологий, такие как IBM, Oracle, Microsoft, Hewlett-Packard и EMC, уже применяли понятие Big Data.

*Влияние Больших Данных
на отдельные экономические аспекты*

Компания Gartner в 2011 году отмечала Большие Данные как один из ключевых трендов в области информационно-технологической инфраструктуры [6]. Согласно прогнозам компании, внедрение данной технологии в будущем должно оказать значительное влияние на информационные технологии в различных отраслях, включая государственное управление, образование, здравоохранение, производство, торговлю и другие сферы, где активно используются информационные ресурсы.

В мае того же года Джеймс Маниака, Майкл Чуи, Брэд Браун, Жак Бюген, Ричард Доббс, Чарльз Роксбург и Анджела Хунг Байерс из Глобального института McKinsey публикуют статью «Большие данные: следующий рубеж для инноваций, конкуренции и производительности». В апреле 2012 года Международный журнал коммуникаций публикует специальный раздел под названием «Информационная ёмкость», посвящённый методологиям и результатам различных исследований, измеряющих объём информации.

С 2013 года в ведущих вузовских программах по науке о данных (Data Science) в качестве академической дисциплины начали изучать большие данные. С 2014 года к сбору и анализу подключились крупнейшие IT-корпорации. В настоящее время крупные компании из различных сфер деятельности, а также государственные структуры активно применяют технологии работы с большими данными.

Проблема взаимодействия официальной статистики и больших данных уже несколько лет серьёзно рассматривается в мире. В числе основных ведущих мировых центров, занимающихся этими вопросами, можно выделить следующие:

- Статистическая комиссия ООН [17];
- Глобальная рабочая группа по большим данным в официальной статистике (Big Data UN Working global group) [2] – является лабораторией Генерального секретаря по инновациям и инклюзии.

- Проект ООН «Глобальный пульс» (Global Pulse) [13]. Является инновационной лабораторией Генерального секретаря, поддерживает ответственные и инклюзивные инновации во всей системе ООН. Его работа охватывает широкий спектр инноваций, использующих данные, цифровые технологии, поведенческие науки, методы стратегического прогнозирования – от расширения доступа женщин к цифровым технологиям в Индонезии до прогнозирования последствий стихийного бедствия на Филиппинах или разработки национальной стратегии работы с данными в Уганде.

В 2017 году в РФ утверждена правительственная программа «Цифровая экономика Российской Федерации», в которой большие данные включены в перечень девяти ключевых цифровых технологий. В настоящее время На сегодняшний день базовые структурированные государственные данные аккумулируются на Портале открытых данных Российской Федерации [16].

Big Data представляет собой новое измерение в контексте современного развития, открывая новые возможности для планирования и управления в различных сферах жизни общества. Однако, с точки зрения безопасности, экономика всегда функционирует в условиях внутренних и внешних угроз. И этот факт порождает новые угрозы применения Big Data во всех областях человеческой деятельности, включая экономическую безопасность.

Например, государство с помощью цифровых данных пытается регулировать все сферы жизни человека. На наших глазах свобода заменяется надзором со стороны силовых структур государства в целях увеличения безопасности граждан и предотвращения противоправных действий. Мы уже можем наблюдать, как в Китае внедряется Система Социального Кредита (доверия) – SCS [14].

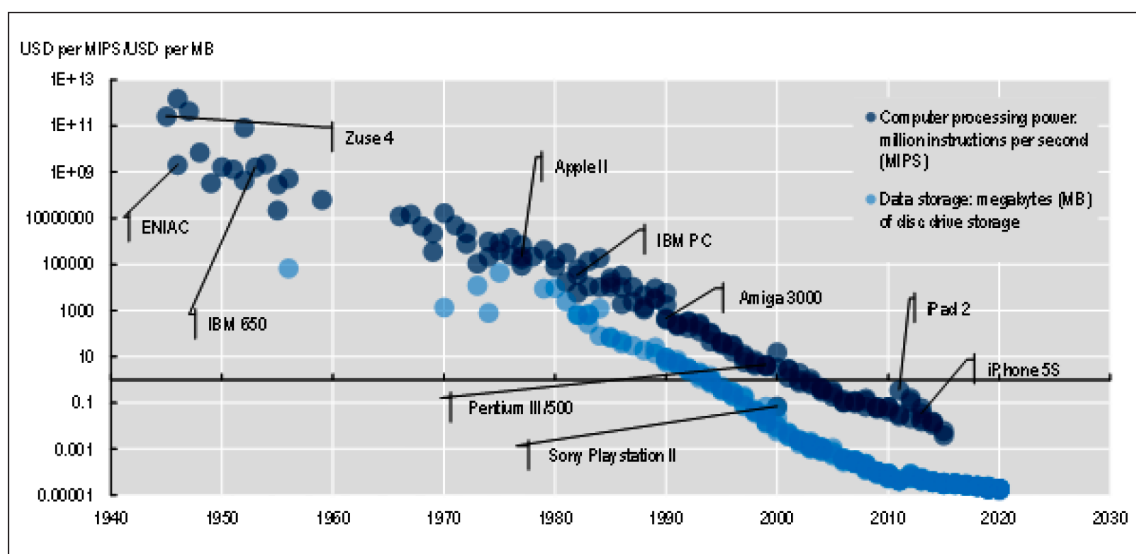


Рис. 1. Стоимость вычислительной стоимости компьютера и хранения данных [3]

Технологические достижения привели к быстрому росту важности данных в экономической деятельности, которая выражается в следующем:

- снижение затрат на обработку и хранение данных;
- рост активности и внедрения цифровых технологий; быстрые достижения в области искусственного интеллекта;
- снижение затрат на обработку и хранение данных

Достижения в области цифрового оборудования позволили снизить затраты на сбор, обработку и хранение данных (рис. 1).

Реальные затраты как на вычислительную мощность компьютеров, так и на хранение данных сокращались вдвое каждые 15 месяцев на протяжении второй половины прошлого столетия. Между тем, с начала 1980-х годов цены на программное обеспечение начали падать гораздо более умеренными темпами – примерно на 2% в год [3].

Число пользователей Интернета, создающих массивы данных, существенно выросло. В 2000 году в среднем только 25% населения стран ОЭСР пользовалось Интернетом, к 2020 году эта доля увеличилась почти до 90%. Возросло не только число пользователей, но и частота, и интенсивность использования Интернета.

Возможности подключения Интернета сначала смещаются от персональных компьютеров в офисах и домах к смартфонам. А в последнее время происходит подключение и других устройств со встроенными датчиками – развивается Интернет вещей

(IoT). По ряду оценок, к 2019 году в мире насчитывалось 20 миллиардов подключённых устройств по сравнению с 8,8 миллиарда в 2010 году, половина этой суммы приходится на устройства Интернета вещей.

Особенности определения стоимостной оценки данных

Как экономический ресурс, данные часто сравнивают с нефтью или солнечным светом. Они также имеют некоторые общие характеристики с идеями или знаниями. Однако, несмотря на некоторое сходство, данные обладают специфическими характеристиками, которые отличают их от всего вышеперечисленного. Среди особенностей надо отметить следующие.

Неконкурентность и исключаемость. Данные неконкурентны: разные субъекты могут использовать их снова и снова без ограничений [7]. С другой стороны, данные, как правило, являются исключаемыми, и в большинстве случаев организации могут препятствовать доступу других людей или учреждений к своим данным. В то время как неконкурентность присуща данным, исключаемость зависит от правовой базы, которая регулирует доступ к данным и право собственности на них [10]

Переливы и внешние эффекты. Исключаемость тесно связана с понятием побочных эффектов, которые возникают, когда субъекты, не участвующие в сборе данных, могут извлечь из них выгоду или понести расходы. Если эти выгоды (расходы) не учитываются в ценах, они создают положитель-

ные (отрицательные) внешние эффекты. Хотя данные в принципе можно исключить, побочные эффекты всё равно могут возникать, например, в случае утечки данных. В сочетании с отсутствием конкуренции положительные внешние эффекты служат веским основанием для широкого распространения данных, например, для улучшения производственных процессов и повышения производительности. Однако в случае с персональными данными возможны негативные внешние эффекты. Например, публикация персональных данных может раскрыть информацию о других людях [1].

Экономия за счет масштаба. При производстве данных наблюдается экономия за счет масштаба, т.е. стоимость создания набора данных высока по сравнению с стоимостью создания дополнительных копий, которая незначительна. Эффект масштаба влияет на ценообразование: когда предельные издержки близки к нулю, цена полностью определяется спросом, то есть ценностью, которую пользователи приписывают данным. Поскольку потребители могут по-разному оценивать набор данных, ценообразование, основанное на ценности, естественно приводит к дифференцированному ценообразованию.

Увеличение отдачи от масштабирования. Данные могут привести к увеличению отдачи от масштаба производства, т.е. увеличение экономической активности порождает больше данных, которые, в свою очередь, порождают больше экономической деятельности и так далее. Если данные будут передаваться между фирмами, то увеличение отдачи от масштаба может также действовать на уровне экономики, демонстрируя растущую отдачу от масштабирования в качестве вклада в машинное обучение [11].

Низкая специфичность. По сравнению с другими продуктами данные обладают меньшей степенью специфичности, т.е. их можно использовать в более широком диапазоне производственных операций. Часто данные повторно используются для целей, отличных от предполагаемых. Например, анонимизированные записи данных о мобильных звонках поставщиков телекоммуникационных услуг были повторно использованы для мониторинга и контроля во время распространения COVID-19.

Синергетический эффект. Данные демонстрируют синергию (или взаимодополняемость) тремя способами. Во-первых,

ценность данных возрастает в присутствии других данных того же типа. Например, отдельные сведения о каком-либо человеке могут иметь неочевидную ценность. Однако в сочетании с данными об этом же человеке за длительный период или о других людях с теми же сведениями, можно выявить тенденции или закономерности. Во-вторых, ценность данных возрастает при наличии других данных с взаимодополняющими характеристиками. Например, официальная экономическая статистика репрезентативна на уровне экономики, но публикуется со значительным временным лагом. Когда эти статистические данные сочетаются с высокочастотными данными частного сектора с цифровых платформ, они могут выявить закономерности. В-третьих, ценность данных возрастает при наличии других, не связанных с данными, факторов, особенно нематериальных. В качестве примеров можно привести конкретные технологии, например, информационно-коммуникационные технологии, аппаратное и программное обеспечение, датчики и навыки. Кроме того, поскольку машинное обучение критически зависит от входных данных, рост популярности машинного обучения может значительно повысить ценность данных в будущем [8].

Оценка Стоимости Больших Данных

Значительные объёмы данных, собранных частными организациями, не продаются на рынках. Например, в среднем 15% предприятий в Европе анализировали большие данные в 2019 году, однако только 1,3% из них покупали данные, а 0,6% продавали данные другим экономическим единицам (рис. 2). Подчеркнем и тот факт, что крупные фирмы, как правило, чаще торгуют данными. Доли тех, кто покупает и продает данные, в среднем составляли менее 5% и 2% соответственно. Это справедливо даже для предприятий с численностью 250 сотрудников и более.

Эти же статистические данные также свидетельствуют о том, что торговля данными неравномерно распределена по различным отраслям. Доля фирм, покупающих и продающих данные, выше в информационно-коммуникационном секторе и ниже в обрабатывающей промышленности и строительстве, а также в сфере услуг «синих воротничков», таких как гостиничный сектор [4].

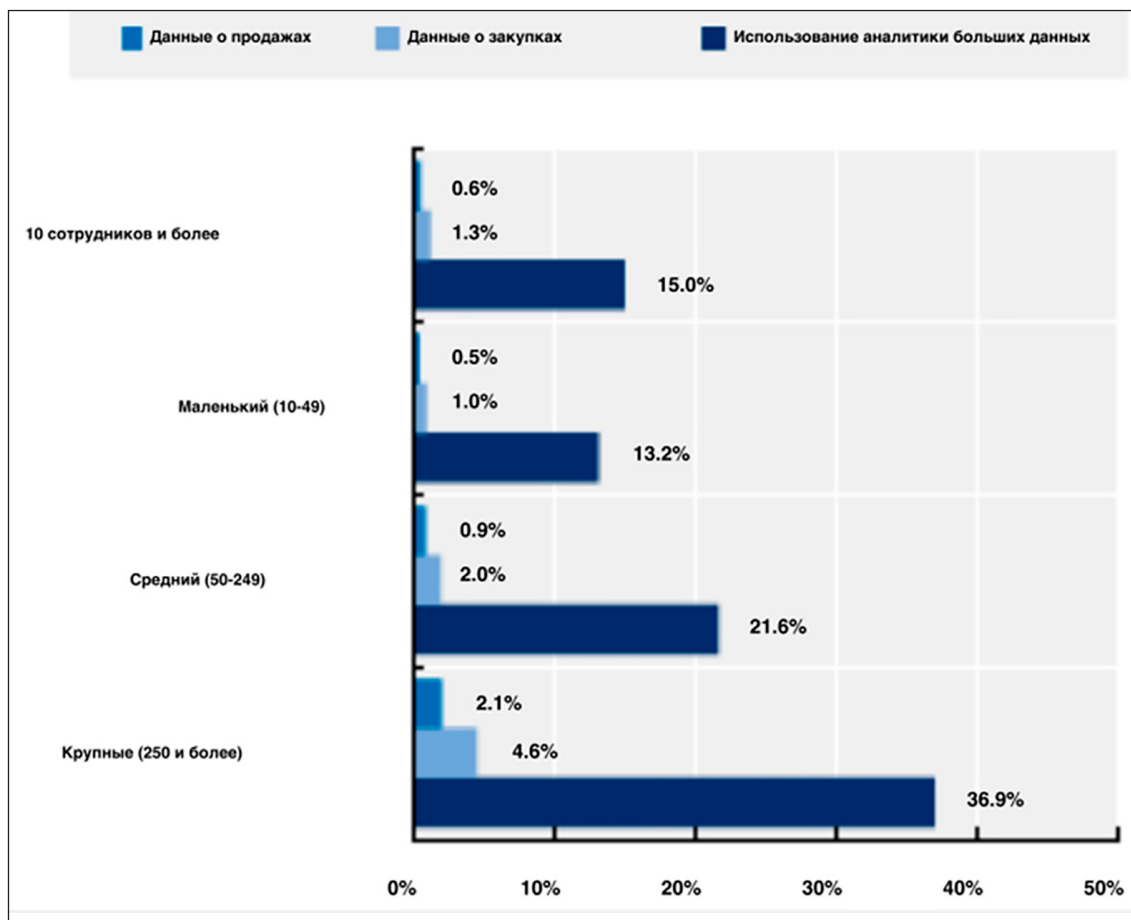


Рис. 2. Доля европейских предприятий, использующих большие данные о покупках и продажах за 2019 год [4].

Однако при ограниченности продаж массивов данных лишь небольшая часть их стоимости может быть измерена на основе рыночной статистики. Эта часть включает в себя доходы фирм, международную торговлю и рыночную оценку. В США доходы от прямых продаж данных в 2019 году оцениваются в 33,3 миллиарда долларов США. В том же году экспорт услуг по передаче данных из Европейского Союза и из США были равны 18,6 миллиарда долларов США и 6,7 миллиарда долларов США, соответственно.

Данные могут использоваться как разово при промежуточном потреблении, так и многократно в течение некоторого времени в качестве ресурса для производства товаров и услуг. Однако, в соответствии с Системой Национальных Счетов (СНС), данные явно не идентифицируются в качестве входного ресурса для производства или как самостоятельный актив в макроэкономической статистике [12]. Хотя в СНС-2008 суще-

ствует актив, называемый «базы данных», затраты на приобретение или создание данных не учитываются явно при расчёте его стоимости. Кроме того, способ учёта затрат на активы в виде данных искажает, на наш взгляд, вклад данных в производство. Явное включение данных в СНС обеспечило бы сопоставимый подход к измерению данных в разных странах. Все страны могли бы единообразно составлять отчёты об инвестициях в данные и их запасах. Затем эти данные можно было бы использовать для различных целей, в том числе для анализа производительности.

Структура СНС сейчас обновляется и будет выпущена в 2025 году. Ключевым приоритетом этого обновления является явное включение производства и использования данных предприятиями в национальные счета. СНС-2025 будет содержать 39 глав (СНС 2008 – 29 глав). Одна из новых глав, предполагается, будет называться «Цифровизация». Также разрабатывается согласованная

методология сбора данных в макроэкономической статистике и возможность учета данных в качестве актива. Однако пока среди разработчиков нет единства по методологии оценки стоимости данных.

В СНС-2008 разъясняется, что активы должны учитываться «по стоимости, по которой они могли бы быть куплены на рынках на момент проведения оценки» и что «в конечном итоге следует использовать рыночную стоимость, наблюдаемую или рассчитанную исходя из наблюдаемой рыночной стоимости» [12]. Но несколько факторов ограничивают объем данных, которыми можно торговать на традиционном рынке. В результате большинство данных, используемых в производстве, создаются за собственный счет, то есть той же фирмой, которая их использует. Следовательно, часто требуются альтернативные методы оценки активов. Когда рыночные цены недоступны, СНС 2008 предоставляет два метода оценки стоимости актива: либо чистую приведенную стоимость актива, либо сумму производственных затрат.

Чистая приведенная стоимость (NPV) использует потенциальный будущий доход, который может быть получен от актива, в качестве альтернативного подхода к оценке текущей стоимости актива. Подход NPV уже используется в некоторых областях СНС, например, для оценки природных ресурсов. Теоретически NPV может служить точным показателем стоимости информационных активов. На практике повторное использование данных может быть выше, чем для других активов, из-за отсутствия физического износа. Поэтому подход NPV требует многочисленных допущений, которые может быть трудно обосновать.

Подход, основанный на сумме затрат, заключается в измерении ценности продукции путем суммирования издержек производства: промежуточное потребление, оплата труда работников, потребление основного

капитала, используемого в производстве, чистая прибыль на основной капитал, используемый в производстве (также известная как «наценка»), и налоги (за вычетом субсидий) на производство.

Вывод

Снижение затрат на обработку и хранение данных, расширение возможностей подключения и внедрение цифровых технологий, а также недавние достижения в области искусственного интеллекта привели к быстрому повышению важности данных в экономической деятельности. Данные обладают определенной комбинацией экономических характеристик, которые отличают их от других производственных ресурсов и влияют на измерение их стоимости. Неконкурентоспособность данных и их потенциальная исключаемость подразумевают, что создание ценности на основе данных зависит от той степени, в которой данные передаются, используются первично и повторно во всей экономике, т. е. от набора инструментов и институтов, регулирующих открытость данных и их совместное использование. Негативные внешние эффекты, связанные с персональными данными, могут привести к слишком широкому обмену данными по слишком низкой цене. Следовательно, информация о ценах и объемах может неадекватно отражать оценку пользователями своих данных. В целом, характеристики данных подразумевают, что затраты и цены могут не обеспечивать полностью точной оценки их экономической ценности. Тем не менее, они необходимы для получения надежных и сопоставимых оценок, которые могут быть дополнительно улучшены с помощью экономического или эконометрического анализа. Также обновленная СНС-2025 предоставит более точную, а главное, общеустановленную методику оценки больших данных для экономики.

Библиографический список

1. Acemoglu Daron, Makhdoumi Ali, Malekian Azarakhsh, Ozdaglar Asuman. Too much data: Prices and inefficiencies in data markets // NBER Working Paper. 2019. No. 26296. National Bureau of Economic Research, Cambridge, MA. DOI: 10.3386/w26296.
2. Big Data UN Working global group. URL: <https://unstats.un.org/bigdata/> (дата обращения: 21.12.2024).
3. Bureau of Economic Analysis “Nonresidential: Intellectual property products: Software”, Private Fixed Investment, Chain-type Price Index, Federal Reserve Bank of St. Louis (FRED). 2022. URL: <https://fred.stlouisfed.org/series/B985RG3A086NBEA> (дата обращения: 26.12.2024).

4. Digital Economy and Society (database). URL: <https://ec.europa.eu/eurostat/web/digitaleconomy-and-society/data/database> (дата обращения: 07.02.2025).
5. FORBES. URL: <https://www.forbes.com/sites/gilpress/2013/05/09/a-very-short-history-of-big-data/> (дата обращения: 15.02.2025).
6. Gartner Glossary. Big Data. 2025. URL: <https://www.gartner.com/en/information-technology/glossary/big-data> (дата обращения: 07.02.2025).
7. Jones C., Tonetti C. Nonrivalry and the economics of data // American Economic Review. 2020. Vol. 110/9. P. 2819–2858. DOI: 10.1257/aer.20191330.
8. OECD. 2022 // Data shaping firms and markets // OECD Digital Economy Papers. No. 344. OECD Publishing, Paris. URL: https://www.oecd.org/en/publications/data-shaping-firms-and-markets_7b1a2d70-en.html (дата обращения: 12.02.2025).
9. ORACLE. URL: <https://www.oracle.com/big-data/what-is-big-data/> (дата обращения: 10.12.2024).
10. Ostrom E. Beyond markets and states: Polycentric governance of complex economic systems // American Economic Review. 2010. Vol. 100/3. P. 641-672. URL: <http://www.aeaweb.org/articles.php?doi=10.1257/aer.100.3.641> (дата обращения: 15.02.2025).
11. Posner Eric A., Weyl Eric G. Radical Markets: Uprooting Capitalism and Democracy for a Just Society, Princeton University Press, Princeton & Oxford. 2018. DOI: 10.1080/09538259.2019.1596564.
12. System of National Accounts 2025, United Nations Statistics Division 2025, Department of Economic and Social Affairs. URL: https://unstats.un.org/unsd/nationalaccount/SNAUpdate/2025/2025_SNA_Combined.pdf (дата обращения: 15.02.2025).
13. UN Global Pulse (Глобальный пульс ООН). 2024. URL: <https://www.unglobalpulse.org/> (дата обращения: 15.02.2025).
14. Бузмакова М.В., Былинская А.А., Глушич Н.Г. Некоторые аспекты роли государства в условиях цифровой экономики // Инновации и инвестиции. 2020. № 12. С. 266-269.
15. Полушкина И.Н. К вопросу о возникновении и содержании понятия Big Data // Социально-экономические предпосылки и результаты развития новых технологий в современной экономике: Материалы VI Международной научной конференции, Нижний Новгород, 14 февраля 2024 года. Нижний Новгород: Национальный исследовательский Нижегородский государственный университет им. Н.И. Лобачевского, 2024. С. 149-153.
16. Портал открытых данных РФ. URL: <https://data.gov.ru/> (дата обращения: 05.12.2024).
17. Статистическая комиссия ООН. URL: <https://unstats.un.org/UNSDWebsite/statcom/> (дата обращения: 21.12.2024).