

УДК 338.512

DOI 10.18413/2687-0932-2020-47-3-491-500

Монетизация больших данных: технико-экономический анализ драйверов роста и издержек

Ю.П. Шальнова

ПАО Сбербанк, Россия, 603000, г. Нижний Новгород, ул. Октябрьская, д. 35

Email: julia.shalnova@gmail.com

Аннотация

В статье на примерах банковской деятельности показано, что внедрение инновационной технологии больших данных является одним из IT-трендов развития мировой и отечественной экономик. Акцентировано внимание на монетизации технологии как необходимом условии внедрения. Предложен не требующий количественных оценок подход повышения инвестиционных результатов, основанный на качественной постановке задач обработки больших данных и формировании множества релевантных им данных. Исследованы основные характеристики больших данных, влияющие на издержки внедрения технологии. Сделан вывод о необходимости приоритетной разработки методик инвестиционного анализа технологии больших данных. Отмечен пилотный характер исследования.

Ключевые слова: большие данные, монетизация, факторы роста, характеристики больших данных, издержки внедрения и эксплуатации.

Для цитирования: Шальнова Ю.П. 2020. Монетизация больших данных: технико-экономический анализ драйверов роста и издержек. Экономика. Информатика. 47 (3): 491–500. DOI: 10.18413/2687-0932-2020-47-3-491-500.

Big data monetization: qualitative technical and economic analysis of drivers of growth and costs

J.P. Shalnova

Sberbank, 35 Oktyabrskaya St, Nizhny Novgorod, 603000, Russia

Email: julia.shalnova@gmail.com

Abstract

Using examples from the banking industry, the article demonstrates that introduction of the big data technologies has become one of the main IT-trends in the development of domestic, as well as world economies. The author highlights that high efficiency of this innovative technology is taken for granted, since in open sources she found none of the studies dedicated to evaluation of investment into such innovation. She claims that one of the key requirements for big data technology implementation in various subject domains is the potential to monetize it. The author puts forward a qualitative approach for result maximization, which is based on a more robust objective setting for big data processing and subsequently forming datasets necessary to achieve the goals of big data tasks at hand. Difficulties of setting big data processing tasks are defined. The author examines distinctive features of big data which determine the size of costs associated with big data technology implementation. She concludes that further analysis of investment into big data technology and development of special techniques for such studies should be prioritized. The pilot studies, therefore, are iterative in nature.

Keywords: big data, monetization, growth factors, big data characteristics, implementation and running costs.

For citation: Shalnova J.P. 2020. Big data monetization: qualitative technical and economic analysis of drivers of growth and costs. Economics. Information technologies. 47 (3): 491–500 (in Russian). DOI: 10.18413/2687-0932-2020-47-3-491-500.

Введение

Внедрение цифровых технологий в экономике и социальной сфере – одна из целей национального развития Российской Федерации [Указ, 2018]. Для ее достижения принята Национальная программа «Цифровая экономика». Федеральный проект этой программы

«Цифровые технологии» предусматривает создание инновационных «сквозных» цифровых технологий, в том числе технологию *больших данных*.

Устремления России в области больших данных соответствуют современной экономической тенденции. Так, согласно исследованию International Data Corporation (IDC), мировые доходы от решений для больших данных и бизнес-аналитики вырастут с \$166 млрд в 2018 году до \$260 млрд в 2022 году [Release, 2018]. При этом флагманом внедрения больших данных являются крупные банки. В США 76 % крупнейших банков используют большие данные для привлечения клиентов, улучшения коммуникаций и повышения лояльности. [McKinsey, 2017], а по оценкам исследования Gartner еще в 2013 году 34 % банков в мире инвестировали в развитие этих технологий [Cnews, 2013]. Поэтому тему, вынесенную в заглавие настоящей статьи, рассмотрим в контексте банковских задач.

В Интернете и популярной литературе широко представлены описания новых возможностей клиентов банков, реализуемых на основе технологий больших данных. Внимание автора привлек следующий фрагмент на сайте BBC: если клиент «собирается погасить свою ипотеку, банк может устроить для него «момент конфетти» – отправить поздравительную открытку, цветы и воздушные шарики, попутно предложив полезные идеи о том, как распорядиться высвобождающимися дополнительными денежными средствами, например, сформировать индивидуальный финансовый план. Прогнозирование жизненных событий и предстоящих финансовых трудностей может помочь банкам укрепить лояльность клиентов и удовлетворять их реальные потребности» [BBC].

В действительности прагматика накладывает существенные ограничения на описанные перспективы. Это могут быть технические, организационные (например, оппортунистическое поведение персонала, «кадровый голод» на Data Scientist), правовые ограничения. И, самое главное, должно выполняться условие монетизации инновационных IT-решений.

Монетизация как основное условие внедрения технологии больших данных

Монетизация данных – это современный бизнес-процесс, при котором генерируемые организацией и внешние данные используются для получения измеримой экономической выгоды (создания реальной стоимости из данных).

В материалах Всемирного экономического форума монетизация данных рассматривается как один из факторов, которые потенциально могут полностью изменить ландшафт финансового рынка [World Economic Forum, 2017]. Перспективность монетизации данных отражается в большом количестве прогнозов. В частности компания Reports and Data прогнозирует рост глобального рынка монетизации данных к 2026 году до \$4,75 млрд [Reports and Data, 2019].

В общем случае монетизация возможна при использовании данных:

1. Во внутреннем контуре организации для:
 - увеличения доходов;
 - снижения затрат;
 - создания новых услуг и реализации иных инновационных возможностей (например, для снижения оттока клиентов, расширения клиентской базы, в том числе в новых направлениях развития, сохранения и увеличения доли рынка);
 - минимизации рисков.

2. Во внешнем контуре организации путем продажи данных третьим лицам или передачи (обмена) их в рамках контрагентских отношений. Однако в настоящее время в силу правовых ограничений (банковская и коммерческая тайны, персональные данные) эти возможности кредитные организации не используют, хотя продажа обезличенной агрегированной информации, аналитики, которая не нарушает закон, не запрещена.

На обязательность выполнения условия монетизации в немногочисленных интервью указывают и представители российских банков, курирующие соответствующие IT-направления [Плас, 2018; Big Data, 2019].

По их мнению, основной целью внедрения инновационных IT является увеличение прибыли через улучшение бизнес-процессов. В случае больших данных – это извлечение

коммерческой ценности из огромных массивов данных, который формируют frontend и, добавим, backend информационные системы.

Для выполнения указанного условия необходимо уметь измерять процессы. Важно понимать потенциальные уровни роста доходности и/или снижения издержек по сравнению с состоянием до внедрения IT-проекта.

Практически банки монетизируют данные, интерпретируя ВВС, путем повышения клиентоориентированности. Для этого анализируют демографию, индивидуальные потребности, исторические покупки и взаимодействия, поведение покупателей и ключевые события и уже на этой основе предлагают персонализированные продукты и услуги на всех стадиях взаимодействия клиент-банк.

Это интуитивно понятный подход. Однако, во-первых, он реализует лишь одно направление монетизации (повышение эффективности продаж). Во-вторых, не отвечает на главный, по нашему мнению, вопрос исследуемого бизнес-процесса: как измерить полученные результаты монетизации (эффект конфетти)? В специальной литературе также не описываются даже общие схемы его разрешения, а экономическая эффективность работы с большими данными просто постулируется.

Как показывают прикладные исследования, лидерами монетизации данных могут считаться компании, у которых рост доходов (снижение расходов) как следствие обработки данных за последние три года составляет не менее 6 %. При значении указанного показателя менее 1 % организация относится к группе аутсайдеров в области обработки данных [McKinsey, 2017]. Приведенные значения можно принять в качестве критериальных и не более того, поскольку McKinsey не раскрывает ноу-хау своих методик.

Принципиальным недостатком всех вышеприведенных результатов исследований является отсутствие упоминания о каком-либо инвестиционном анализе проектов больших данных, под которым в простейших случаях понимают сравнение полученной прибыли и инвестиционных затрат. Иначе: не рассматриваются и не оцениваются драйверы роста прибыли и факторы возникновения издержек при монетизации больших данных.

При этом «попытки использования традиционных методов для инвестиционного анализа IT-проектов сопряжены с большими трудностями количественной оценки элементов формулы ROI (*Return On Investment* – курсив наш), т.к. расчет прибыли и инвестиций является самостоятельной нетривиальной задачей. Эта сложность имеет институциональный характер ...» [Карпычев, 2010].

Также очевидно, что в любом IT-проекте можно выделить ряд частных системно связанных задач, решение которых направлено на увеличение ROI. Применительно к большим данным такой частной задачей является минимизация издержек формирования стоимости больших данных и технологий их обработки.

Рассмотрим основные факторы и технические характеристики, формирующие стоимость больших данных. Для этого, прежде всего, определим понятие «*большие данные* – это большой объем, высокая скорость и/или большое разнообразие информационных активов, которые требуют *экономически эффективных* (курсив наш), инновационных форм обработки информации, позволяющих улучшить понимание, принятие решений и автоматизацию процессов» [Glossary].

Это определение фиксирует основные сугубо технические характеристики больших данных: объем (Volume), скорость (Velocity), разнообразие (Variety), 3V, которые следует учитывать при технико-экономическом проектировании систем больших данных. Но оно не охватывает ряд важнейших для экономического анализа факторов, непосредственно влияющих на его результаты.

Цели обработки больших данных

Как и любая сложная деятельность, монетизация технологий больших данных может быть представлена развитым деревом целей. Корневая (основная) цель – получение различных

выгод в стоимостном исчислении – сформулирована нами выше. Эта цель достигается посредством постановки и достижения целей более низкого уровня.

В качестве таких целей применения технологии больших данных банковское сообщество в настоящее время рассматривает:

- определение и анализ структуры доходов и расходов клиентов;
- выявление основных каналов транзакций (через отделение банка или разнообразные дистанционные каналы обслуживания) и характерных для них действий клиента (оплата картой, безналичные переводы, оформление вкладов или кредитов и т. д.).
- сегментирование клиентов в соответствии с их профилями;
- анализ рентабельности промоакций, новых услуг и совместных предприятий;
- кросс-продажи банковских продуктов на основе сегментации клиентов;
- противодействие мошенничеству (anti-fraud) и иным противоправным деяниям;
- оценку риска, соответствие требованиям безопасности и отчетности перед регулятором;
- анализ отзывов клиентов о работе банков.

Приведенный перечень не является закрытым. Более того, ведутся постоянные исследования по его расширению, например, как следует из приведенной цитаты ВВС, прорабатываются решения по созданию персонального «финансового советника».

В общем случае известно несколько направлений использования больших данных [McKinsey], в том числе для получения *скрытой* информации (закономерностей, взаимосвязей и тенденций), содержащейся в больших данных (*hidden pattern*), и, что важно, изначально не предназначенных для ее добычи и использования.

В этой связи необходимо упомянуть о технологии Data Mining (раскопка данных), которая вкуче с технологией больших данных позволяет добывать *скрытую* в данных большого объема информацию (закономерности, взаимосвязи и тенденции). *Data mining* – это нетривиальное извлечение из данных неявной, ранее неизвестной и потенциально полезной информации [Frawley et al, 1992]. Именно это направление является перспективой широкого применения больших данных.

Для целей дальнейшего изложения представим задачу создания системы больших данных простым выражением

$$G=F(D), \quad (1)$$

где G – требуемый результат (цель), D – набор исходных данных, F – функционал информационной системы, преобразующей исходные данные в требуемый результат.

Условие качественной постановки задачи (1) является определяющим фактором в успешности проекта (максимизации ROI) больших данных. Поясним понятие качественной постановки задачи. Прежде всего, это четкая постановка цели, которая, пожалуй, является самой сложной для формулирования, формализации и интерпретации процедурой. Некоторые эксперты считают, что многие проекты больших данных оканчиваются неудачей (в контексте статьи – низкой монетизацией) из-за некорректно поставленных целей [Marx, 2015].

Обобщенная вербализация этой стадии проекта больших данных достаточно сложна. Поэтому проиллюстрируем изложенные тезисы частным примером. Пусть в терминах бизнес-процесса сформулирована проблема: «клиент ушел в отток», что означает снижение денежных поступлений от клиентской базы. Непосредственными причинами этой проблемы могут быть приостановка, расторжение договоров, неиспользование услуг в течение определенного промежутка времени, несвоевременная оплата счетов и др.

В свою очередь, эти причины могут быть вызваны обстоятельствами более высокого порядка, например, макроэкономическими и политическими. Естественно, ни в одном хранилище или источнике данных не заложены бизнес-правила, устанавливающие связи снижения денежных поступлений (например, закрытие депозитных договоров) и этих факторов.

Очевидно, что качественное формулирование цели определяется знаниями предметной области и IT весьма дорогостоящих специалистов, известных как Data Scientist. Это означает,

что на стадии формирования целей экспертно должна быть решена классическая оптимизационная задача минимизации затрат на Data Scientist при допустимом уровне риска некорректного формирования целей.

Цели последовательно определяют данные и их источники, а также методы анализа данных и, соответственно, стоимостные характеристики проектов больших данных.

Формирование множества релевантных данных

Сформулированная цель G обработки больших данных детерминирует формирование предметно-ориентированного множества больших данных D (перечень и семантику), т. е. эту задачу можно представить выражением

$$G \rightarrow D = \{d_1, \dots, d_m\}, \quad (2)$$

где d_i – элемент множества релевантных цели G данных.

С предметной точки зрения элемент d_i – фактор, каким-либо образом влияющий на достижение цели G .

Основными критериями включения элемента d_i в множество D являются:

- семантическая принадлежность (релевантность) элемента к предметной области;
- важность (вес) элемента d_i .

Важность – показатель, характеризующий степень влияния (вклад) на достижение цели G . Важность некоторых данных может быть неочевидна для конкретной цели G . Так, продолжая вышеприведенный пример (отток клиентов), следует определить необходимость использования соответствующих высказываний клиентов (данных) на специализированных площадках (banki.ru) и в социальных сетях (источников).

Постановка задачи (2) не является тривиальной. Большинство предметных специалистов, принимающих решение, затрудняются:

- в оценке объемов и качества скрытой информации, которую содержат данные;
- в определении источников данных.

Действительно, «найти данные для анализа – это отчасти наука, отчасти исследовательская работа и отчасти предположения» [Ohlhorst, 2013].

Методы формирования множества D – экспертные, результаты – уникальные и определяются целью G и квалификацией эксперта. Стоимостная характеристика процедуры формирования множества D и стоимость входящих в него данных в конкретных проектах имеют широчайшую вариативность. С точки зрения минимизации издержек на внедрение и последующую эксплуатацию информационной системы больших данных множество данных D должно отвечать условию семантической необходимости и достаточности для достижения цели G . Кроме этого, на стоимость технологии больших данных влияет еще и ряд рассматриваемых ниже характеристик.

Основные характеристики, влияющие на стоимость больших данных

Объем больших данных (Volume) – технический параметр, значения которого не вполне определены, в том числе и потому, что научно-технический прогресс непрерывно расширяет их диапазон. С экономической точки зрения объем данных прямо, но не всегда линейно влияет на стоимостные характеристики процесса обработки (генерация, передача, преобразование, хранение) данных. При этом стоимость обработки единицы данных существенным образом зависит от схемы и технологии обработки данных. Возможны варианты технологических решений, например, хранение данных в облаке, удешевляющее процедуру.

Кроме того, следует указать, что известны публикации, в которых показано, что чрезмерно большой объем данных также негативно влияет на результаты работы систем больших данных (ошибки в работе), как и их недостаток [Calude, Longo G, 2017].

Таким образом, в технико-экономическом обосновании инвестиционного проекта больших данных, очевидно, должен присутствовать раздел, аргументирующий объем данных для поставленных задач.

Семантика и синтаксис больших данных. Сущность работы с большими данными заключается в автоматизированной обработке семантики данных: постановку цели G , формирование множества D и методы обработки осуществляет эксперт, непосредственную обработку данных – ЭВМ.

Автоматизированная обработка семантической информации возможна только при ее формализации. В настоящее время это одна из самых трудных и дорогостоящих задач в IT-сфере. По этому критерию в теории больших данных выделяют структурированные, полуструктурированные и неструктурированные данные [Vozábal, 2016].

Структурирование – это выделение семантически важных элементов в информации и установление связей между ними. В структурированных данных четко и однозначно определены семантические единицы и связи между ними (например, реляционные данные). Обработка структурированных данных имеет наименьшую стоимость.

В неструктурированных данных семантические единицы и связи между ними, соответствующие предметной задаче, в явном виде не определены. Неструктурированные данные: текстовая (электронные письма, твиты, форумы, блоги), числовая (временные ряды данных финансовых рынков), графическая информация (рисунок, фото), аудио (массивы записей человеческой речи) и видеофайлы.

Работа с неструктурированными данными – технически сложная задача. Для их семантической обработки необходимы сложные математико-алгоритмические решения, дорогостоящее аппаратно-программное обеспечение с высокими техническими характеристиками. Техническая сложность обуславливает повышенную стоимость решений при разработке, эксплуатации и сопровождении систем больших данных. Это должно быть учтено при принятии решения о включении (важности) конкретного неструктурированного элемента d_i в множество D .

Кроме семантики в описании больших данных должен быть отражен *синтаксис*, под которым обычно понимают формально-структурные характеристики информации, не затрагивающие ее смыслового содержания. Физически в электронном виде формально-структурные характеристики данных реализуются в конкретных форматах файлов.

Корреспондируя и расширяя это понятие до технологии работы с данными, мы выявляем различия и, как следствие, ограничения (по производителям, платформам, версиям, форматам и др.) на аппаратно-программные средства. Экономический аспект этого явления заключается в необходимости введения в бизнес-процесс обработки данных дополнительной затратной процедуры согласования форматов данных. Простейший пример необходимости такой процедуры – согласование форматов файлов xls и xlsx табличного процессора Excel.

Интеграция данных. Ключевой в 3V-определении больших данных показатель *разнообразие* (Variety) предполагает сбор данных из двух и более источников данных с различными структурно-параметрическими характеристиками (синтаксис и семантика). Это означает обязательность процесса интеграции – объединения данных разнородных источников. Данные могут быть неоднородны как в части программной реализации (форматы), так и в части модели их представления (структуры). Интеграция обеспечивает единое представление (формирование полного и непротиворечивого набора данных) и оперирование этими данными [Dong, Srivastava, 2015].

В качестве примера интеграции приведём задачу «Аналитика телефонных переговоров» для контактных центров банков. Целями задачи могут быть:

- поиск каких-либо трендов на множестве телефонных коммуникаций (структурированная информация);
- распознавание определенного содержания (неструктурированная информация);
- увязка структурированных и неструктурированных данных с выделением новой (скрытой) информации.

Интеграция данных также весьма сложная техническая задача, решения которой находятся в начальной стадии. Выполнение этой обязательной для больших данных процедуры предполагает дополнительные затраты на аппаратно-программное обеспечение,

квалифицированный персонал и связано с дополнительными административно-хозяйственными расходами и временными издержками.

Техническая сложность задачи также сопряжена с повышенными рисками ошибок, отказами технических средств и иными нештатными ситуациями, прямо влияющими на экономические результаты организации (снижение ROI, рост TCO – полной стоимости владения информационной системой, total cost of ownership). Это означает, что наибольшие инвестиционные и эксплуатационные издержки технологии больших данных существуют при использовании (включении в D) всех видов неструктурированной информации.

Иные экономически значимые параметры больших данных. Для исследования экономической эффективности технологии больших данных целесообразно учитывать еще ряд параметров.

Режим обработки данных – способ выполнения задач, непосредственно влияющий на время, стоимость и другие характеристики. При обработке больших данных могут быть использованы режимы: пакетный; реального масштаба времени; разделения времени; регламентный; телеобработки, однопрограммный; мультипрограммный и др. [Шепелев и др., 2019].

Выбор режима обработки данных обусловлен условиями задачи (1). Так, для обработки транзакционных данных в anti-fraud задаче необходим режим реального времени, для определения динамики расходов клиентов возможно применение пакетного режима и т. д.

Инвестиционные и эксплуатационные расходы на реализацию режима обработки данных относятся к прямым издержкам и рассчитываются обычным калькулированием затрат.

Для некоторых задач имеют значение *темпоральные* (временные) *характеристики* данных: момент, интервал, период, время хранения [Jensen, 2000], обуславливающие особенности, в том числе экономические, их обработки. Для финансовых организаций это могут быть:

- разрыв между моментами генерации данных источником и окончанием их обработки приемником (режим торгов на Московской бирже T+2);
- асинхронное время работы источников и пользователей данных (мировые финансовые рынки);
- различные интервалы времени обработки одних и тех же данных при использовании их несколькими процессами или задачами (подготовка отчетности в разных разрезах и глубины детализации);
- загрузка данных «задним числом» (пакетная передача темпорально значимых данных) и др.

Для банков имеет значение время (стоимость) хранения больших данных. Например, платежные транзакции, актуальные в моменте, приобретают временное значение при анализе истории операций клиента.

Планируемые виды анализа данных. При постановке задачи (1) желательно определить вид анализа (Types of Data Analytics) больших данных, иначе – функционал информационной системы F . Как правило, вид (метод) анализа определяется требуемым результатом.

В терминах *Data Mining* виды анализа представляются двумя классами моделей: *описательными* (Descriptive) и *предсказательными* (Predictive). Виды моделей, составляющих каждый класс, формализуют задачу (1), например, выявление закономерностей между связанными событиями (ассоциация).

Решение задачи обеспечивается применением одного из методов Data Mining, соответствующего модели. В свою очередь, применение конкретного метода Data Mining накладывает определенные требования на характеристики данных D , например, вид данных: числовой, категориальный и др. [Барсегян и др., 2007; Kantardzic, 2020].

Экономические аспекты выбора метода анализа больших данных заключаются в сравнительной стоимостной оценке типового программного инструментария выполнения функционала F или стоимостной оценке вновь разрабатываемого инструмента.

Принадлежность данных. Информационная потребность организации в конкретных данных всегда сопряжена с вопросом о принадлежности прав владения, пользования, распоряжения данными. По этому критерию выделим внутренние (собственные) и внешние данные. Соответственно, можно говорить о внутренних и внешних источниках данных.

Внутренние данные (first party data) – это существующие, ранее созданные, полученные иным образом приобретенные для целей предметной деятельности и размещенные в хранилищах организации данные. Права на такие данные в полном объеме принадлежат этой организации, что означает отсутствие издержек на реализацию права их использования в технологии больших данных.

Внешние данные – данные, правообладателем которых являются третьи лица, сторонние для заинтересованной организации. В этом случае данные рассматриваются как объект экономических отношений (объект обмена, имеющий ценность, товар).

По критерию возмездности доступа (использования) к внешним данным выделим две категории: условно-бесплатные (shareware) и проприетарные данные (данные как услуга, data as a service). Обе схемы предполагают приобретение анонимизированных или агрегированных данных. Так, банк «УралСиб» обращается к размещаемым в свободном доступе данным ФНС РФ [Big Data, 2019], а АО «Национальное бюро кредитных историй» приобретало у одного из стартапов открытую информацию для оценки кредитоспособности пользователей социальной сети «ВКонтакте» (судебное дело № А40-18827/2017). Весьма интересными для таргетированных контактов банка могут быть агрегированные и анонимизированные данные геолокации клиентов.

В контексте настоящей статьи обращение к внешним источникам данных означает введение в функционально-стоимостную модель технологии больших данных дополнительных транзакционных издержек.

Заключение

Проведенное исследование выявило факторы роста монетизации и снижения издержек технологии больших данных. Полученные результаты полезны для обобщенного понимания стоимостной природы проекта и могут быть использованы при пилотном (предпроектном) технико-экономическом анализе систем больших данных, а также для экспертной поддержки принятия решений при невозможности количественной оценки результатов проекта.

Для более точной постановки задач анализа больших данных необходим учет не рассмотренных в настоящем исследовании уникальных факторов: географии, отрасли и размера организации [Release, 2018]. И, наконец, еще одним направлением дальнейших научных исследований является разработка подходов (методик) количественной оценки экономически значимых характеристик больших данных.

Список литературы

1. Барсегян А.А., Куприянов М.С., Степаненко В.В., Холод И.И.. 2007. Технологии анализа данных: Data Mining, Visual Mining, Text Mining, OLAP. СПб.: БХВ-Петербург, 384.
2. Карпычев В.Ю. 2010. Инвестирование в информационные технологии: проблемы и решения, Экономический анализ, 25: 2–8.
3. Плас. 2018. Big Data в банкинге: универсальных рецептов нет [Электронный ресурс]. URL: https://plusworld.ru/journal/section_2018/plus-2-2018/big-data-v-bankinge-universalnyh-retseptov-net/ (Дата обращения: 03.06.2020).
4. Указ Президента Российской Федерации № 204. 2018. О национальных целях и стратегических задачах развития РФ на период до 2024 года. [Электронный ресурс]. URL: <http://kremlin.ru/acts/bank/43027> (Дата обращения: 14.06.2020).
5. Шепелев К.В., Суркова Н.Е., Шувалова И.С. 2019. Анализ режимов автоматизированной обработки данных. Промышленные АСУ и контроллеры, 12: 48–53.
6. BBC. Demystifying Big Data in banking. [Электронный ресурс]. URL: <http://www.bbc.com/storyworks/banking-on-innovation/bigdata> (Дата обращения: 10.05.2020).

7. Big Data и ИИ в банках: тренд или реальный инструмент? 2019. [Электронный ресурс]. URL: <https://mcs.mail.ru/blog/big-data-i-ii-v-bankah-trend-ili-real-instrument> (Дата обращения: 02.06.2020).
8. Calude C.S., Longo G. 2017. The Deluge of Spurious Correlations in Big Data. *Foundations of Science*, Volume 22: 595–612.
9. Cnews. 2013. Как крупнейшие банки используют большие данные. [Электронный ресурс]. URL: https://cnews.ru/articles/kak_kрупnejshie_banki_ispolzuyut_bolshie (Дата обращения: 10.06.2020).
10. Dong X.L., Srivastava D. 2015. *Big Data Integration*. Morgan & Claypool. 178.
11. Frawley W., Piatetsky-Shapiro G., Matheus C. 1992. Knowledge Discovery in Databases: An Overview. *AI Magazine*: 213–228.
12. Glossary Gartner. [Электронный ресурс]. URL: <https://www.gartner.com/en/information-technology/glossary/big-data> (Дата обращения: 12.02.2020).
13. Jensen C.S. 2000. *Temporal Database Management*. Aalborg University. 1323.
14. Kantardzic M. 2020. *Data mining: Concepts, Models, Methods, and Algorithms*. Wiley. Hoboken. 661.
15. Marr B. 2015. Where Big Data Projects Fail. [Электронный ресурс]. URL: <https://www.forbes.com/sites/bernardmarr/2015/03/17/where-big-data-projects-fail/#c8ef3f6239f6> (Дата обращения: 14.06.2020).
16. McKinsey. 2017. Fueling growth through data monetization. [Электронный ресурс]. URL: <https://www.mckinsey.com/business-functions/mckinsey-analytics/our-insights/fueling-growth-through-data-monetization> (Дата обращения: 09.06.2020).
17. McKinsey Global Institute. Big data: The next frontier for innovation, competition and productivity. [Электронный ресурс]. URL: <https://www.mckinsey.com/business-functions/mckinsey-digital/our-insights/big-data-the-next-frontier-for-innovation> (Дата обращения: 12.02.2020).
18. Ohlhorst F. 2013. *Big Data Analytics Turning Big Data into Big Money*. Wiley. 176.
19. Reports and Data, 2019. Report: Data Monetization Market By Data Type, By Component, By Organization Size, By Deployment Mode (Cloud and On-premises), By End Use, By Industry Vertical, And Segment Forecasts, 2016-2026. [Электронный ресурс]. URL: <https://www.reportsanddata.com/report-detail/data-monetization-market> (Дата обращения: 08.06.2020).
20. Release Summary. 2018. Сайт BusinessWire. [Электронный ресурс]. URL: <https://www.businesswire.com/news/home/20180815005095/en/Revenues-Big-Data-Business-Analytics-Solutions-Forecast> (Дата обращения: 08.06.2020)
21. Vozábal M. 2016. *Tools and Methods for Big Data Analysis*. Pilsen. University of West Bohemia. 134.
22. World Economic Forum. 2017. Beyond Fintech: A Pragmatic Assessment Of Disruptive Potential In Financial Services [Электронный ресурс]. URL: http://www3.weforum.org/docs/Beyond_Fintech_-_A_Pragmatic_Assessment_of_Disruptive_Potential_in_Financial_Services.pdf (Дата обращения: 10.06.2020).

References

1. Barsegyan A.A., Kupriyanov M.S., Stepanenko V.V., Kholod I.I. 2007. *Tekhnologii analiza dannykh: Data Mining, Visual Mining, Text Mining, OLAP [Data analysis technologies: Data Mining, Visual Mining, Text Mining, OLAP]*. Saint-Petersburg: BKhV-Peterburg, 384.
2. Karpychev V.Yu. 2010. Investing in IT: problems and solutions. *Economic analysis*, 25: 2–8. (in Russian)
3. Plus. 2018. Big Data in banking: there are no universal recipes. Available at: https://plusworld.ru/journal/section_2018/plus-2-2018/big-data-v-bankinge-universalnyh-retseptov-net/ (accessed 03 June 2020). (in Russian)
4. President of the Russian Federation, edict № 204. 2018. Regarding national goals and and strategic development of the Russian Federation until 2024. Available at: <http://kremlin.ru/acts/bank/43027> (accessed 14 June 2020). (in Russian)
5. Shepelev K.V. Shepelev K.V., Surkova N.E., Shuvalova I.S. 2019. Analysis of automated data processing modes. *Industrial ACS and controllers*, 12: 48–53. (in Russian)
6. Big Data and AI in banks: a trend or a real tool? Available at: <https://mcs.mail.ru/blog/big-data-i-ii-v-bankah-trend-ili-real-instrument> (accessed 02 June 2020). (in Russian)
7. BBC. Demystifying Big Data in banking. Available at: <http://www.bbc.com/storyworks/banking-on-innovation/bigdata> (accessed 10 May 2020).

8. Calude C.S., Longo G. 2017. The Deluge of Spurious Correlations in Big Data. *Foundations of Science*, Volume 22: 595-612.
9. Cnews. 2013. How major banks use big data. [Электронный ресурс]. URL: https://cnews.ru/articles/kak_krupnejshie_banki_ispolzuyut_bolshie ((accessed 10 June 2020).
10. Frawley W., Piatetsky-Shapiro G., Matheus C. 1992. Knowledge Discovery in Databases: An Overview. *AI Magazine*: 213–228.
11. Glossary Gartner. Available at: <https://www.gartner.com/en/information-technology/glossary/big-data> (accessed 12 February 2020).
12. Dong X. L., Srivastava D. 2015. *Big Data Integration*. Morgan & Claypool. 178.
13. Jensen C. S. 2000. *Temporal Database Management*. Aalborg University. 1323.
14. Kantardzic M. 2020. *Data mining: Concepts, Models, Methods, and Algorithms*. Wiley. Hoboken. 661.
15. Marr B. 2015. Where Big Data Projects Fail. Available at: <https://www.forbes.com/sites/bernardmarr/2015/03/17/where-big-data-projects-fail/#c8ef3f6239f6> (accessed 10 June 2020).
16. McKinsey. 2017. Fueling growth through data monetization. Available at: <https://www.mckinsey.com/business-functions/mckinsey-analytics/our-insights/fueling-growth-through-data-monetization> (accessed 09 June 2020).
17. McKinsey Global Institute. Big data: The next frontier for innovation, competition and productivity. Available at: <https://www.mckinsey.com/business-functions/mckinsey-digital/our-insights/big-data-the-next-frontier-for-innovation> (accessed 12 February 2020).
18. Ohlhorst F. 2013. *Big Data Analytics Turning Big Data into Big Money*. Wiley. 176.
19. Reports and Data. 2019. Report: Data Monetization Market By Data Type, By Component, By Organization Size, By Deployment Mode (Cloud and On-premises), By End Use, By Industry Vertical, And Segment Forecasts, 2016-2026. [Электронный ресурс]. URL: <https://www.reportsanddata.com/report-detail/data-monetization-market> (accessed 08 June 2020).
20. Release Summary. 2018. Available at: <https://www.businesswire.com/news/home/20180815005095/en/Revenues-Big-Data-Business-Analytics-Solutions-Forecast> (accessed 08 June 2020).
21. Vozábal M. 2016. *Tools and Methods for Big Data Analysis*. Pilsen. University of West Bohemia. 134.
22. World Economic Forum. 2017. Beyond Fintech: A Pragmatic Assessment Of Disruptive Potential In Financial Services. Available at: http://www3.weforum.org/docs/Beyond_Fintech_-_A_Pragmatic_Assessment_of_Disruptive_Potential_in_Financial_Services.pdf (accessed 10 June 2020).

ИНФОРМАЦИЯ ОБ АВТОРЕ

Шальнова Юлия Павловна, ведущий экономист отдела экономической оценки Центра компетенций «Финансово-экономическая оценка» Службы финансового менеджмента ПАО Сбербанк, Нижний Новгород, Россия

INFORMATION ABOUT THE AUTHOR

Julia P. Shalnova, Leading Economist of the Economic Assessment Department of the Financial and Economic Assessment Competence Center of the Financial Management Service of Sberbank, Nizhny Novgorod, Russia