

УДК 62-004.62 + 004.75

## **ТЕНДЕНЦИИ ИСПОЛЬЗОВАНИЯ БОЛЬШИХ ДАННЫХ И МАШИННОГО ОБУЧЕНИЯ ПРИ ПОСТРОЕНИИ СОВРЕМЕННЫХ ИНФОРМАЦИОННЫХ СИСТЕМ В КРУПНЫХ КОРПОРАЦИЯХ**

*А.А. Шинкарев*

Представлен краткий обзор трендов при разработке современных информационных систем в крупных корпорациях. Определено место больших данных и машинного обучения при разработке сегодня и в ближайшей перспективе. Выявлены проблемы, стоящие сегодня перед специалистами по созданию программных систем, рынком труда, вузами, готовящими кадры, и бизнесом, создающим запрос на новые подходы к проектированию и созданию информационных систем.

Ключевые слова: корпоративные приложения, большие данные, машинное обучение, информационные системы, анализ.

Сегодня корпоративные системы создаются по-новому. Микросервисный подход в том или ином виде внедряется повсеместно. Во фронтенде через подход Micro Frontends [1], в бэкенде монолитные приложения дробятся на отдельные сервисы, а новые системы создаются изначально раздробленными на маленькие кусочки. С одной стороны, это дает гибкость, а с другой – вносит новую сложность. Новый подход обещает повысить скорость внедрения изменений в бизнес-процессы, но он в то же время и является причиной лавинообразного роста сложности системы, количества инвариантов поведения и точек выхода из строя.

Помимо дробления систем на модули высокого уровня, также большую силу приобрело сообщество разработчиков программного обеспечения с открытым исходным кодом (open source). Крупные корпорации уже не гнушаются использования открытых пакетов, а такие гиганты как Microsoft разрабатывают свой новый полностью открытый фреймворк .NET Core во многом силами сообщества разработчиков. Это течение, несомненно, будет и дальше набирать силу и в значительной степени влиять на направление вектора разработки в целом и корпоративных приложений в частности.

Место больших данных и машинного обучения в современном мире определено наличием огромных объемов информации о людях, пользователях систем, потенциальных покупателях, конкурентах. Эту информацию необходимо уметь сохранить, обработать и извлечь из нее выгоду. В этом ключе большие данные и машинное обучение приходят в массы

и становятся базовыми технологиями в арсенале современного разработчика и бизнес-аналитика.

Понятие больших данных еще не так давно было новым, заботящим сравнительно небольшую группу энтузиастов, неподъемным для большинства заказчиков и потребителей корпоративных приложений в сфере информационных технологий. Однако уже наступил тот переломный момент, когда можно с уверенностью сказать, что будущее в индустрии программных и программно-аппаратных продуктов именно за большими данными и умением извлечь из них разнообразную, доселе недоступную, информацию. И это актуально не только для крупных трансатлантических корпораций, но даже для малого бизнеса.

Было бы недальновидно говорить об отмирании необходимости в инструментах ведения бизнеса, которыми на сегодняшний день заполнен рынок. Они останутся и будут поддерживаться десятилетиями, но по-настоящему раскрыть потенциал смогут компании, которые оперативно перестроятся под новые реалии мира «Интернета вещей», мира, где объем информации удваивается каждые 2–4 года, а применение искусственного интеллекта и машинного обучения станет рутиной, а не сферой, к которой многие профессионалы со стажем не готовы.

Современные примеры систем поддержки принятия решений, таргетирования рекламы, прогнозирования необходимых размеров товарных запасов, начиная от ресторанов и заканчивая промышленными предприятиями, показывают, насколько новый вид анализа может быть распространен на разные сферы жизнедеятельности. Безусловно, говоря о больших данных как таковых, имеет смысл рассматривать их в связке с методами и моделями машинного обучения, математической статистики, теорией вероятностей, визуализацией данных. Эти инструменты необходимы для того, чтобы анализ массивов данных и извлечение из них важной информации были возможны.

Хотя сами данные важны, а инструментарий необходим, без специалистов предметной области для каждого конкретного случая обойтись весьма сложно. Да, безусловно, современные модели машинного обучения, например, на основе нейронных сетей глубокого обучения, могут демонстрировать впечатляющие результаты, но не везде можно избежать участия человека как для разметки данных, так и для анализа результатов работы моделей.

Разрыв между потребностями предприятий и тем, кого готовят университеты сегодня, не позволяет эффективно решать задачи, поставленные указом Президента РФ от 09.05.2017 № 203 «О Стратегии развития информационного общества в Российской Федерации на 2017–2030 годы» [2]. Необходимо наладить более тесную кооперацию между вузами, государством и представителями бизнеса, чтобы готовить востребованных

рынком труда специалистов, в первую очередь владеющих практически навыками, которых так не хватает сегодняшним выпускникам.

Можно сколько угодно говорить о впечатляющих достижениях в области интеллектуального анализа данных, которые есть уже сейчас, однако для руководителей крупных организаций, для лиц, принимающих решения и несущих за них ответственность, важны мощные и в тоже время не перегруженные инструменты представления результатов анализа, взвешивающие и сравнивающие различные стратегии поведения. Важен конечный этап процесса работы с большими данными, а именно этап визуализации результата, который играет ключевую роль, а качество его исполнения определяет, будут ли достигнутые вами результаты поняты и адекватно восприняты конечными потребителями информации.

Несмотря на обилие терминов и клишированных фраз, таких как большие данные, машинное обучение, интернет вещей, асинхронность, параллельность, распределенность, их объединяют между собой общие базовые концепции, идеи и проблемы, которые были известны еще в двадцатом веке. Например, такая абстракция, как подтверждение (acknowledgement) или же проблема состояния гонок в автоматах (Race Condition Automaton). существует в схемотехнике достаточно давно, но современные информационные технологии продолжают их переоткрывать вновь и вновь при появлении новых модных технологий или парадигм разработки.

В этом контексте современные «прорывные» или «хайповые» направления видятся отчасти переосмыслением уже существующего багажа знаний и оформлением его в новом сверкающем варианте. Отсюда и отсутствие той «серебряной пули», которая бы позволяла решать все проблемы и задачи в определенной области легко и непринужденно. Ведь создание такого инструмента может потребовать значительного переосмысления подходов к построению информационных систем.

Из-за сложностей идеологического, методологического и технического характера в подавляющем большинстве компаний построение корпоративных систем происходит с использованием подходов, которые апробировались последнее десятилетие.

Можно сказать, что существует целый ряд стандартных задач, таких как документооборот, аутентификация, авторизация, CRUD-операции для различных бизнес-сущностей (Create, Read, Update, Delete), построение отчетов и аналитики и т.д. И решение этих задач тоже уже во многом устоялось, хоть и существует множество вариаций с точки зрения не только технологического стека, но и качества исполнения конечной реализации.

Выход на сцену такого направления подготовки специалистов IT-сферы, как Data Science (Наука о данных), вполне вероятно поменяет, а точнее уже меняет, направление вектора разработки информационных систем, подготовки специалистов по их созданию и сопровождению,

а главное, те задачи, решение которых позволяет сделать бизнес более конкурентоспособным.

Можно однозначно сказать, что Big Data, Data Science и развитие направлений масштабируемости и отказоустойчивости информационных систем предопределили развитие информационных технологий в долгосрочной перспективе. Компании, которые не воспользуются всем спектром предоставляемых возможностей, могут остаться на плаву, но максимальную выгоду и лидирующие позиции обеспечат себе лишь те, что инвестируют ресурсы и смогут перестроиться под стремительно меняющиеся реалии.

#### Библиографический список

1. Geers, M. Micro Frontends: Extending the Microservice Idea to Frontend Development [Электронный ресурс] / M. Geers. – URL: <https://micro-frontends.org> (дата обращения: 22.04.2019).
2. О Стратегии развития информационного общества в Российской Федерации на 2017–2030 годы [Электронный ресурс]: указ Президента РФ от 09.05.2017 № 203. – URL: <http://www.kremlin.ru/acts/bank/41919> (дата обращения: 22.04.2019).

[К содержанию](#)