

## СТАНДАРТИЗАЦИЯ И ОБЕСПЕЧЕНИЕ ИНФОРМАЦИОННОЙ БЕЗОПАСНОСТИ В ОБРАЗОВАНИИ

**Байхожаева Б.У.**

baikhozhayeva\_bu@enu.kz

**Баймагамбетова Ж.М.**

baimagambetova.zm@gmail.com

Кафедра «Стандартизация, сертификация и метрология»  
Евразийский национальный университет им.Л.Н. Гумилева  
Нур-Султан, Казахстан

В современном мире уникальным ресурсом является информация, использование которой позволяет достичь максимальных высот во всех сферах жизнедеятельности человека. Количество информации превысило тот уровень, когда ее можно было обрабатывать традиционными способами, вследствие этого, данные формируются с промедлением, что усложняет, а в большинстве случаев не позволяет принимать оперативные управленческие решения.

На текущий момент одним из динамично развивающихся сегментов в информационной сфере является технология Больших данных. Данная технология является общепринятым признаком экономического и технологического развития. [1]

Анализ больших данных позволяет ускорить решение исследовательских и педагогических проблем. К примеру, это касается персонализации обучения, повышения качества преподавания, выбора будущей профессии, виртуального кампуса. Изучая статистику, можно работать и с индивидуальными траекториями, и с глобальными образовательными системами. [2]

Необходимо отметить, что с применением технологии Больших Данных появляется возможность отслеживать цифровой след каждого обучающегося: его участие в конкурсах, олимпиадах, качество выполняемых им учебных заданий, его проекты, изобретения, какими компетенциями владеет и т.д. [3]

Применение технологий Больших данных позволит снизить число проверок в организациях образования, возможность сделать их более точечными. В перспективе большие данные можно использовать и при совершенствовании системы аккредитации и оценки вузов. Другими словами, технологии Больших данных возможность совершенствовать оценку эффективности качества образования. В образование, как известно, вкладываются большие средства, и не всегда вложенные средства коррелируют с показателями его качества.

Управление системой образования, оперативный мониторинг текущих процессов, оценка влияния образования на экономический рост, оплату труда, занятость, заболеваемость, преступность в регионе, внутреннюю миграцию между регионами и утечку мозгов за рубеж, анализ вузовских программ, поиск тематически близких программ за рубежом, кластеризация и ранжирование образовательных учреждений по объёмам подготовки, мониторинг спроса на

отдельные профессии на рынке труда, улучшение качества прогнозов спроса на профессии на несколько лет вперед - это лишь несколько направлений улучшения системы образования с помощью технологий Больших данных. [3]

Одним из главных вопросов при использовании систем Больших Данных является безопасность. Системы Больших Данных сложны и неоднородны, и к безопасности всей системы необходимо подходить целостно. Кроме того, интеграция различных технологий создает новые проблемы безопасности, которые должны быть должным образом решены.

Цель работы - анализ нормативной базы в области информационной безопасности и разработка рекомендаций по ее совершенствованию.

Как известно, конкретные стандарты Больших Данных разрабатываются различными хорошо зарекомендовавшими себя организациями по разработке стандартов и отраслевыми консорциумами [3-6]

ISO/IEC JTC 1/SC 32 Управление данными и обмен ими, включая языки баз данных, управление мультимедийными объектами, управление метаданными и электронный бизнес. Стандарты электронного бизнеса, включая согласование ролей; метаданные репозитории, спецификации моделей, определения метамodelей; SQL; библиотеки объектов и пакеты приложений, построенные на (использующих) SQL.

ISO/IEC JTC 1/SC 38 Стандартизация для совместимых распределенных прикладных платформ и сервисов, включая веб-сервисы, сервис-ориентированную архитектуру (SOA) и облачные вычисления. Интерфейс управления облачными данными, открытый формат виртуализации, архитектура взаимодействия веб-служб (SOA) и возможности облачных вычислений

ITU-TSG13 Облачные вычисления для больших данных Облачные вычисления основываются на требованиях к большим данным, возможностях и вариантах использования

W3C Веб и семантические стандарты, связанные с разметкой, структурой, запросами, семантикой и обменом данными. Множество стандартов, включая стандарты спецификации онтологий, разметку данных, запросы, контроль доступа и обмен.

Открытый Геопространственный консорциум (OGK) Геопространственные стандарты для спецификации, структуры, запросов и обработки данных, связанных с местоположением. Множество стандартов, связанных с кодированием, обработкой, запросами и контролем доступа к геопространственным данным.

Организация по продвижению структурированных информационных стандартов (OASIS) Доступ к информации и обмен ею Набор протоколов для взаимодействия со структурированным контентом данных, таким как OData стандарты на безопасность, облачная вычислительная техника, SOA, веб-сервисы, интеллектуальная сеть, Электронная публикация, управление чрезвычайными ситуациями и другие области

Совет по эффективности обработки транзакций Бенчмаркинг для систем больших данных. Спецификация TPCEXPRESSBenchmark™ для системы NADOOR и соответствующего комплекта.

TMForum Обмен опытом для решения важнейших бизнес-задач, включая трансформацию ИТ, оптимизацию бизнес-процессов, анализ больших данных, управление облаком и кибербезопасность.

В Казахстане активно осуществляется деятельность по разработке стандартов в области информационных технологий Больших Данных, но тем не менее, четко виден недостаток стандартов в данной сфере, а также основополагающих требований к обеспечению безопасности и конфиденциальности больших объемов данных [7-15]

Одним из основополагающих стандартов является СТ РК 2927-2016 «Информационные технологии. Большие Данные. Требования и возможности облачных вычислений».

Стандарт разработан на основе рекомендации ITU Y.3600 (11/2015) Bigdata. Cloudcomputingbasedrequirementsandcapabilities (Большие данные. Облачные вычисления на основе требований и возможности) [6].

Данный стандарт представляет собой обзор требований к использованию Больших Данных, а также затрагивает основные вопросы определения Больших Данных, таких как:

- Введение в Большие Данные;
- Экосистемы и роли Больших Данных;
- Отношения между облачными вычислениями и Большими Данными;
- Облачные вычисления, основанные на системном контексте Больших Данных и их преимущества;
- Возможности облачных вычислений, основанных на Больших Данных.

В стандарте СТ РК 2927-2016 «Информационные технологии. Большие Данные. Требования и возможности облачных вычислений» определены следующие термины, касающиеся технологии Больших Данных:

- Большие Данные (bigdata) - система понятий для обеспечения сбора, хранения, управления, анализа и визуализации данных, потенциально при ограничениях в режиме реального времени, обширных наборов данных с разнородными характеристиками;
- Большие Данные как услуга (БДкУ) (bigdataasaservice (BDaaS)) - категория облачных услуг, в которой возможности, предоставляемые клиенту облачных услуг являются сбором, хранением, анализом, визуализацией и управлением данными с использованием Больших Данных;
  - Сторона (party) - физическое или юридическое лицо;
  - Действие (activity) - указанные занятия или набор задач;
  - Облачные вычисления (cloudcomputing) - система понятий для обеспечения сетевого доступа к масштабируемому и эластичному объединению общих физических или виртуальных ресурсов с самообслуживания, выделения ресурсов и управления по требованию;
  - Клиент облачных услуг (cloudservicecustomer) лицо или организация, которая потребляет предоставленную облачную услугу в рамках договора с поставщиком облачных услуг;
  - Партнер облачных услуг (cloudservicepartner) - лицо или организация, которая обеспечивает поддержку услуг, предоставляемых провайдером облачных услуг (например, интеграция услуг);
  - Поставщик облачных услуг (cloudserviceprovider) - организация, которая обеспечивает и поддерживает предоставленную облачную услугу;
  - Метаданные (metadata) - данные, описывающие другие данные;

- Роль (role) - набор деятельности, который служит общей целью;
- Подчиненная роль (sub-role) - подмножество деятельности данной роли.

Стандарт также рассматривает вопросы облачных вычислений, основанных на системном контексте Больших Данных, выдвигая требования к сбору, хранению, предварительной обработке и анализу данных [6].

В результате проведенного сравнительного анализа современного состояния стандартизации и оценки соответствия в сфере информационной безопасности Больших Данных в странах ЕС и РК, было выявлено, что в национальном стандарте СТ РК 2927-2016 «Информационные технологии. Большие Данные. Требования и возможности облачных вычислений» недостаточно терминов и определений, дающих подробный обзор об информационной безопасности Больших Данных, что может повлечь за собой недопонимания между поставщиками и клиентами Больших Данных [16]. В этой связи, в научной работе, предложены следующие термины для внедрения:

- алгоритм - это эффективная категориальная спецификация того, как решить сложную задачу и как выполнить анализ данных. Он состоит из нескольких шагов, чтобы применить операции над данными для решения конкретной проблемы;
- данные реального времени - данные, которые создаются, обрабатываются, хранятся, анализируются и визуализируются в течение миллисекунд;
- анализ рисков - применение статистических методов на одном или нескольких наборах данных для определения вероятного риска проекта, действия или решения;
- искусственный интеллект (ИИ) - это развитие компьютерных систем для выполнения задач, обычно имеющих человеческий интеллект, таких как распознавание речи, визуальное восприятие, принятие решений и т. д.;
- автоматическая идентификация и сбор данных (AIDC) - это термин Больших Данных, который относится к методу автоматической идентификации и сбора объектов данных с помощью вычислительного алгоритма и последующего хранения их в компьютере. Например, радиочастотная идентификация, штрих-коды, биометрия, оптическое распознавание символов, магнитные полосы - все это включает алгоритмы идентификации захваченных объектов данных;
- поведенческая аналитика - это недавний прогресс в области бизнес-аналитики, который представляет собой новое понимание поведения клиента на платформах электронной коммерции, веб- и мобильных приложениях, онлайн-играх и т. д. Это позволяет маркетологам делать правильные предложения нужным клиентам в нужное время;
- бизнес-аналитика - это набор инструментов и методологий, которые могут анализировать, управлять и предоставлять информацию, имеющую отношение к бизнесу. Он включает в себя инструменты отчетов/запросов и панель мониторинга;
- BigDataScientist - это человек, который может взять структурированные и неструктурированные точки данных и использовать свои огромные навыки в статистике, математике и программировании, чтобы организовать их. Он применяет всю свою аналитическую мощь (контекстуальное понимание, знание отрасли и понимание существующих предположений), чтобы раскрыть скрытые решения для развития бизнеса;

- Cassandra - это распределенная и открытая система управления базами данных NoSQL. Она предназначена для управления большим объемом распределенных данных на товарных серверах, поскольку обеспечивает высокую доступность услуг без каких-либо сбоев;

- детализация звонков (CDR) и анализ CDR содержит метаданные, т. е. данные о данных, которые телекоммуникационная компания собирает о телефонных звонках, таких как продолжительность и время вызова. Анализ CDR предоставляет предприятиям точные сведения о том, когда, где и как осуществляются звонки для целей выставления счетов и отчетности;

- облачные вычисления - это один из самых известных терминов Больших Данных. Это новая парадигмальная вычислительная система, которая предлагает визуализацию вычислительных ресурсов для запуска на стандартном удаленном сервере для хранения данных. Облачные вычисления предоставляют ИТ-ресурсы, такие как инфраструктура, программное обеспечение, платформа, база данных, хранилище и так далее в качестве услуг. Гибкое масштабирование, быстрая эластичность, объединение ресурсов, а также самообслуживание по требованию;

- кластерный анализ - это термин Больших Данных, относящийся к процессу группировки объектов, сходных между собой в общую группу (кластер). Это делается для того, чтобы понять сходства и различия между ними. Важная задача исследовательского интеллектуального анализа данных и общих стратегий для анализа статистических данных в различных областях, таких как анализ изображений, распознавание образов, машинное обучение, компьютерная графика, сжатие данных и т.д.;

- комплексная обработка событий - это процесс анализа и идентификации данных, а затем их объединения для вывода событий, способных предложить решения для сложных обстоятельств. Основная задача - выявлять/отслеживать значимые события и реагировать на них как можно скорее;

- архитектура и дизайн данных - в ИТ-индустрии архитектура данных состоит из моделей, стандартов политик или правил, которые управляют тем, какие данные агрегируются, а также тем, как они организуются, хранятся, интегрируются и используются в системах данных. Она состоит из трех фаз: концептуальное представление хозяйствующих субъектов; логическое представление отношений между хозяйствующими субъектами; физическое построение системы функционального обеспечения;

- администратор базы данных - это термин Больших Данных, относящийся к роли, которая включает в себя планирование емкости, конфигурацию, проектирование баз данных, мониторинг производительности, миграцию, устранение неполадок, безопасность, резервное копирование и восстановление данных. Администратор баз данных несет ответственность за поддержание и поддержание правильности содержания и структуры базы данных;

- модель данных является начальной фазой проектирования базы данных и обычно состоит из атрибутов, типов сущностей, правил целостности, отношений и определений объектов;

- моделирование данных - это процесс создания модели данных для информационной системы с использованием определенных формальных методов. Моделирование данных используется для определения и анализа потребности в данных для поддержки бизнес-процессов;

- система управления базами данных (СУБД) - это программное обеспечение, которое собирает данные и обеспечивает доступ к ним в организованном виде. Он создает базу данных и управляет ею.

- очистка данных - это процесс пересмотра данных для удаления неправильного написания, дубликатов записей, добавления недостающих данных и обеспечения согласованности. Это необходимо, так как неверные данные могут привести к плохому анализу и неправильным выводам;

- документооборот - это программное обеспечение, которое используется для отслеживания, хранения и управления электронными документами и электронным изображением бумаги через сканер. Это один из основных терминов Больших Данных, которые вы должны знать, чтобы начать карьеру в области Больших Данных;

- визуализация данных - это представление данных в графическом формате, предназначенном для передачи информации или получения смысла. Он позволяет пользователям / лицам, принимающим решения, видеть аналитику визуально, чтобы они могли разработать новые концепции.

Нами разработаны рекомендации по систематизации основных требований, терминов и понятий в сфере информационной безопасности Больших Данных для того, чтобы внедрять и использовать данную технологию. Введение систематизированной терминологии, повысит качество работы при обеспечении информационной безопасности, а также упростит процессы взаимодействия между клиентами и поставщиками Больших Данных.

Необходимо внедрять стандарты информационной безопасности Больших Данных в сферу образования. Адаптация или создание стандартов в данной сфере поможет предоставлять потребителям более качественные услуги. Для этого органы стандартизации должны создавать отраслевые группы (поставщики и пользователи Больших Данных), а также включать регуляторов из заинтересованных секторов и согласовывать общие стандарты и сертификаты.

#### **Список использованной литературы**

- 1 ISO/IEC JTC 1 Information technology. Big data Preliminary Report - 2014. – 5с.
2. <http://www.edutainme.ru/post/big-data-edu/>
3. <http://www.unkniga.ru/vishee/9614-bolshie-dannye-v-obrazovanii.html>
4. BSI Big Data and standards market research, January 2016 – с. 14-26.
5. IDC, Big Data Predictions 2014: Beyond Irrational Exuberance - Opportunities in the Big Data and Analytics Markets Web Conference Dec 11, 2013. Framingham, MAUnitedStates – с. 34-36.
6. DRAFT NIST Big Data Interoperability: Volume 4: NIST Big Data Security and Privacy Requirements, Version 1.0 April 23, 2014, –
7. Большие данные: международный опыт и тенденции развития в Казахстане. <https://zerde.gov.kz/activity/ict/publication/2225/>
8. СТ РК 2927-2016 Информационные технологии. Большие Данные. Требования и возможности облачных вычислений.
9. ISO/IEC 27000 Информационная безопасность системы менеджмента. Обзор и словарь.
10. ISO/IEC 27001 Информационные технологии. Методы обеспечения безопасности. Системы менеджмента информационной безопасности. Требования.
11. ISO/IEC 27002 Свод правил по средствам управления защитой информации.

- 12.ISO/IEC 27003 Руководство по внедрению системы менеджмента информационной безопасности.
- 13.ISO/IEC 27004 Менеджмент информационной безопасности. Измерение.
- 14.ISO/IEC 27005 Менеджмент риска информационной безопасности.
- 15.ISO/IEC 27006 Требования к органам, осуществляющим аудит и сертификацию систем менеджмента информационной безопасности.
- 16.ISO/IEC 27017 Свод правил для информационной безопасности на основе ISO/IEC 27002.
- 17.ISO/IEC 27018 Информационные технологии. Методы обеспечения защиты. Свод правил по защите персональных данных (PII) в среде публичного облака, действующего как PII процессор
- 18.101 Big Data Terms You Should Know <https://www.whizlabs.com/blog/big-data-terms/>
- 19.Automation of harmonization, analysis and evaluation of information security requirements// Dmitrij OLIFER// Vilnius Gediminas Technical University – с.36-42.