



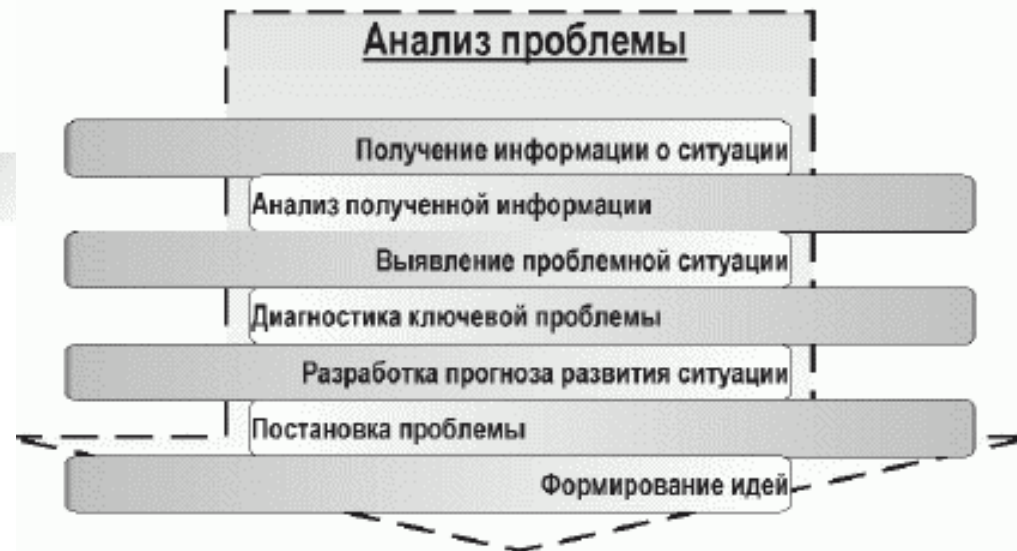
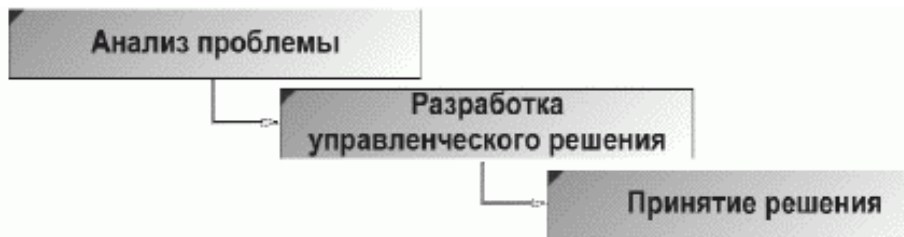
НАЦИОНАЛЬНЫЙ ИССЛЕДОВАТЕЛЬСКИЙ
УНИВЕРСИТЕТ

Современные информационные технологии в бизнесе

Тема V

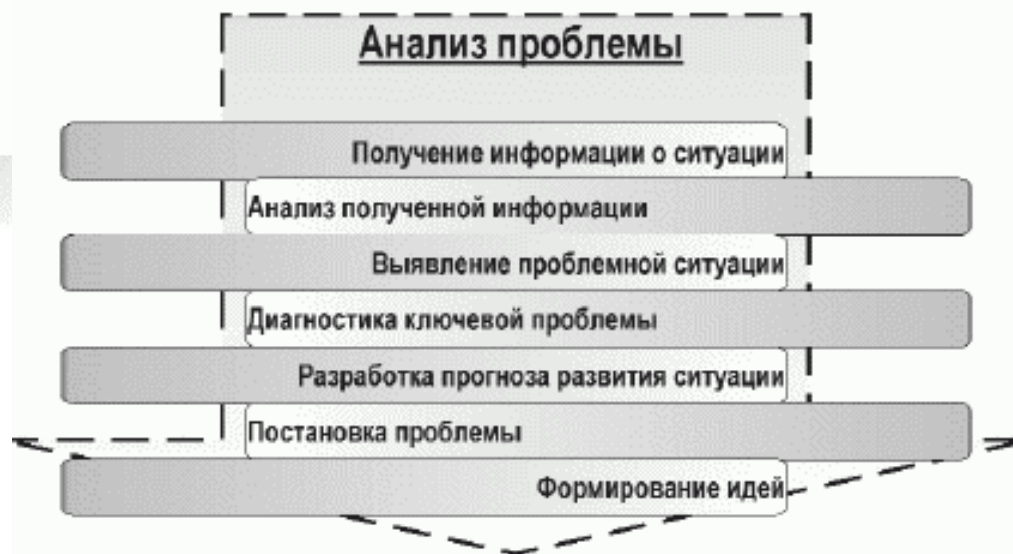
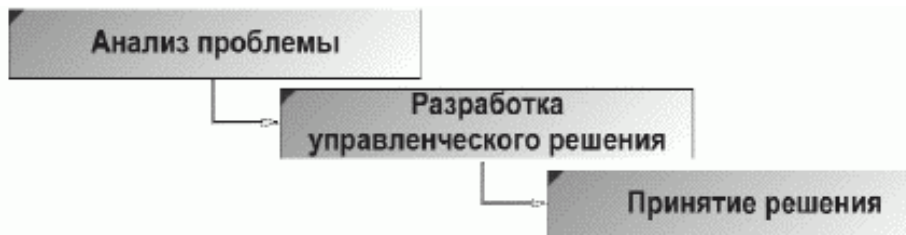
Аналитическая обработка данных.
Обзор BI-систем.

В современном мире успех компании на рынке напрямую зависит от того, как быстро менеджмент компании может распознать изменения динамики рынка и насколько своевременно может отреагировать на них с целью увеличения прибыли, исходя из существующих реалий рынка. Менеджеры компании должны отслеживать тенденции рынка, идентифицировать конкурентов и угрозы, оценивать риски, преобразовывать стратегию компании, оценивать свои ресурсы и т.д. Информация является необходимым производственным ресурсом для принятия эффективных управленческих решений.



Компании накопили значительные объемы данных и имеют доступ к еще большим объемам внешних данных. Менеджерам необходимо, чтобы эта информация была *преобразована*, *предварительно обработана* и соответствующим образом *организована* для быстрого доступа, анализа и принятия решений.

Ключевую роль в управлении организацией в целом и ее отдельными производственными функциями играет информация.



Данные, которые доступны менеджерам и аналитикам непосредственно из корпоративных информационных систем, не унифицированы, разрознены и в общем случае не готовы для анализа.

Возникает потребность в информационных системах аналитического анализа данных, способных превратить данные корпоративных информационных систем и данные из внешних источников в полезные для бизнеса информацию и знания, используемые в управлении, на основе которых можно принимать решения.



Что такое Business Intelligence?

Business Intelligence — это процесс анализа информации, выработки интуиции и понимания для улучшенного и неформального принятия решений бизнес-пользователями, а также инструменты для извлечения из данных значимой для бизнеса информации; процесс, технологии, методы и средства извлечения и представления знаний.

BI - это совокупность технологий, программного обеспечения и практик, направленных на достижение целей бизнеса путём наилучшего использования имеющихся данных.

Цель BI — интерпретировать большое количество данных, заостряя внимание лишь на ключевых факторах эффективности, моделируя исход различных вариантов действий, отслеживая результаты принятия решений.

BI наиболее эффективен, когда он объединяет данные, полученные из рынка, на котором работает компания (внешние данные), с данными из источников внутри компании, таких как финансовые и операции с данными (внутренние данные). В сочетании, внешние и внутренние данные дают полную картину бизнеса, которая, создает "интеллект" — быстрое понимание, которое не получить из простого набора данных.

Рождение термина "Business Intelligence" датируется 1958 годом, когда американский ученый Ханс Петер Лун (1896-1964) опубликовал в IBM System Journal статью «A Business Intelligence System». В ней он представил бизнес как набор различных видов деятельности в науке, технологиях, коммерции, индустрии и даже в законодательной сфере, а обеспечивающие его системы – системами, поддерживающими разумную деятельность (intelligence system).

Словом *intelligence* Лун обозначал способность устанавливать взаимосвязь между представлениями отдельных фактов и действиями в интересах решения поставленных задач и достижения намеченных целей.

В 1989 году аналитик из Gartner Ховард Дреснер дал BI расширительную трактовку, предложив использовать BI в качестве общего термина для различных технологий, предназначенных для поддержки принятия решений - "пользователецентрический процесс, включающий доступ и исследование информации, ее анализ, выработку интуиции и понимания, которые ведут к улучшенному и неформальному принятию решений".

К 1996 году содержание термина было уточнено, и "Business Intelligence" стал пониматься как "инструменты для анализа данных, построения отчетов и запросов, которые могут помочь бизнес-пользователям преодолеть море данных для того, чтобы помочь синтезировать из них значимую информацию".



BI как результат процесса извлечения знаний

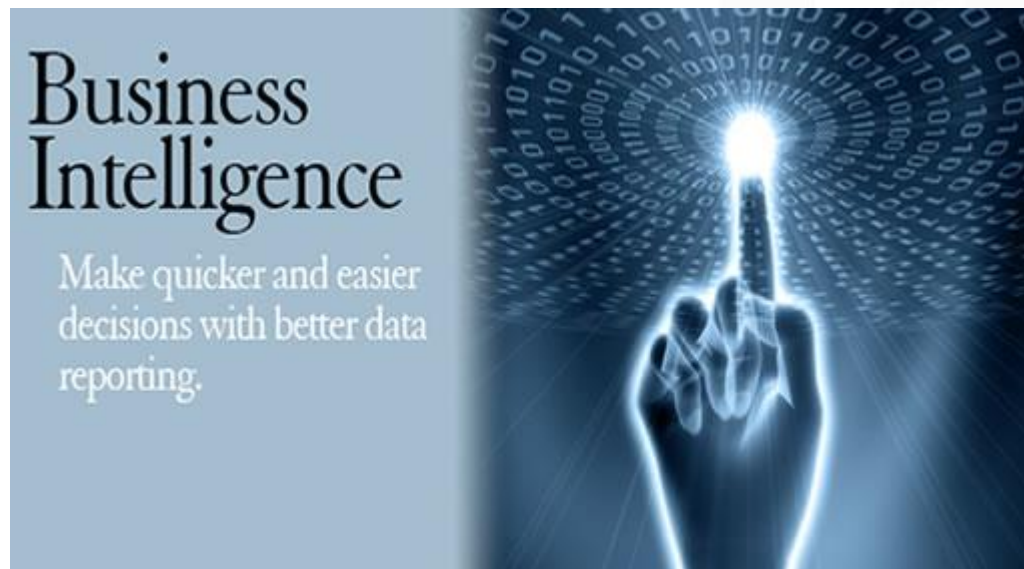
BI имеет отношение к процессу превращения данных в знания, добытые о бизнесе с использованием различных ИТ, а знаний - в действия для получения выгоды.

Инструменты BI — ИС, которое позволяют бизнес-пользователям видеть и использовать большое количество сложных данных. Знания, основанные на данных, (data-based knowledge) получаются из данных с использованием инструментов business intelligence и процесса создания и ведения хранилища данных (data warehousing).



Итак, Business Intelligence можно определить как:

- ❑ процесс превращения данных в информацию и знания о бизнесе для поддержки принятия улучшенных и неформальных решений;
- ❑ информационные технологии (методы и средства) сбора данных, консолидации информации и обеспечения доступа бизнес-пользователей к знаниям;
- ❑ знания о бизнесе, добытые в результате углубленного анализа детальных данных и консолидированной информации.



Что такое Business Intelligence?

Итак, Business Intelligence можно определить как:

- ❑ процесс превращения данных в информацию и знания о бизнесе для поддержки принятия улучшенных и неформальных решений;
- ❑ информационные технологии (методы и средства) сбора данных, консолидации информации и обеспечения доступа бизнес-пользователей к знаниям;
- ❑ знания о бизнесе, добытые в результате углубленного анализа детальных данных и консолидированной информации.

В русскоязычной литературе термин "Business Intelligence" переводится как "бизнес-аналитика", "бизнес-интеллект", "интеллектуальный анализ данных", "деловая осведомленность" или вводится просто как аббревиатура BI.

**«intelligence»: способность узнавать и понимать; готовность к пониманию; знания, переданные или приобретенные путем обучения, исследования или опыта; действие или состояние в процессе познания; разведка, разведывательные данные.*

В чем разница между BI и BA?

**Business
Analytics**



**Business
Intelligence**

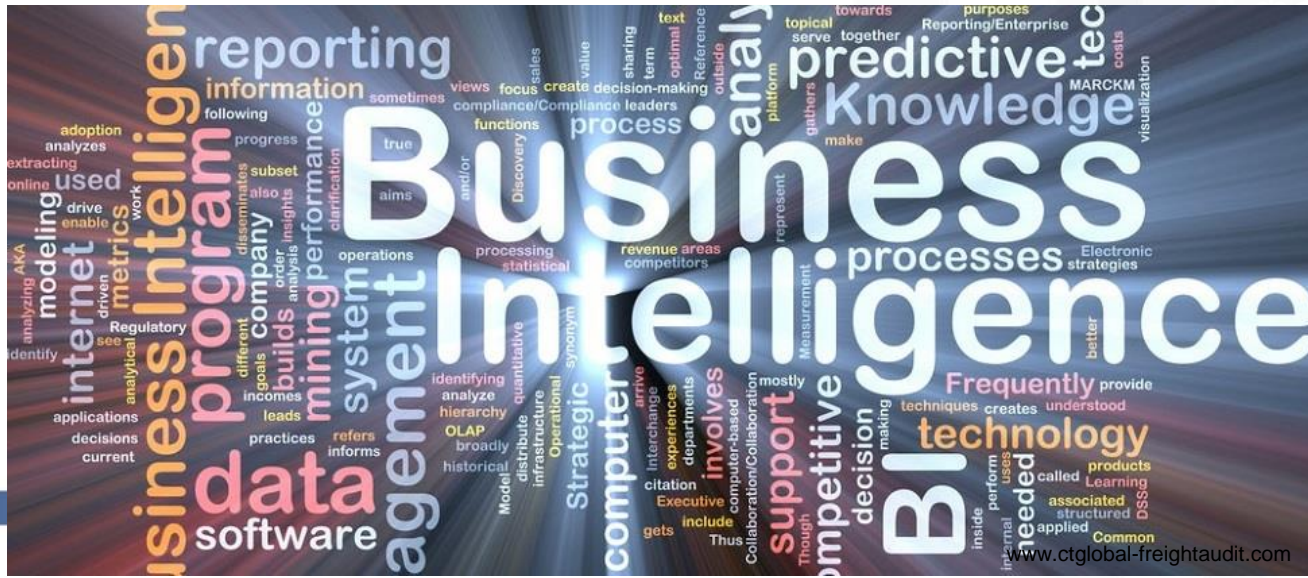
citia.co.uk

What's the difference between Business Analytics and Business Intelligence? The correct answer is: everybody has an opinion, but nobody knows, and you shouldn't care.

Тимо Эллиот, сотрудник компании SAP и эксперт в вопросах о бизнес-аналитике

BI-системы или системы бизнес-аналитики (Business Intelligence) - это аналитические системы, которые объединяют данные из любых различных источников информации, обрабатывают их и предоставляют удобный интерфейс для всестороннего изучения и оценки полученных сведений.

Данные полученные в результате такого анализа помогают достигать поставленных бизнес-целей с помощью оптимального использования имеющихся данных. Комплексный анализ данных по всем направлениям бизнеса позволяет повысить его эффективность и снизить издержки. BI-системы - это единый прозрачный источник данных о бизнесе компании для ее руководства.



Таким образом, основной целью систем бизнес-аналитики является обеспечение быстрого доступа к данным, выполнение анализа данных и информационная поддержка процесса принятия решений. Бизнес-аналитика помогает компании создавать знания из всей доступной информации для принятия эффективных управленческих решений и превращения этих решений в действие.

Система бизнес-аналитики является стержнем, вокруг которого формируются потоки стратегической бизнес-информации. Данный инструмент помогает компании принимать решения, которые будут основаны на корректной информации, полученной вовремя.

В условиях, когда рынок постоянно меняется, а конкуренция становится все жестче, руководителям крайне необходимо выявлять и анализировать имеющиеся у предприятия резервы, которые могут существенно расширить возможности бизнеса.

Предлагаемые решения в области бизнес-аналитики должны предоставлять возможность **оперативно анализировать тенденции рынка**, **осознавать движущие силы бизнеса** и, основываясь на объективной информации, **быстро реагировать на изменения рыночной ситуации** и **принимать верные решения**.

Всевозможные типы отчетов

№	Имя	Описание	Бренд
0	Алгоритм	Алгоритм Календарный	Иконка
1	Алгоритм	Алгоритм Календарный	Иконка
2	Алгоритм	Алгоритм Календарный	Иконка
3	Алгоритм	Алгоритм Календарный	Иконка
4	Алгоритм	Алгоритм Календарный	Иконка
5	Алгоритм	Алгоритм Календарный	Иконка
6	Алгоритм	Алгоритм Календарный	Иконка
7	Алгоритм	Алгоритм Календарный	Иконка
8	Алгоритм	Алгоритм Календарный	Иконка
9	Алгоритм	Алгоритм Календарный	Иконка
10	Алгоритм	Алгоритм Календарный	Иконка
11	Алгоритм	Алгоритм Календарный	Иконка
12	Алгоритм	Алгоритм Календарный	Иконка

Единая таблица данных

Id	Name	Description	Brand
0	Алгоритм	Алгоритм Календарный	Иконка
1	Алгоритм	Алгоритм Календарный	Иконка
2	Алгоритм	Алгоритм Календарный	Иконка
3	Алгоритм	Алгоритм Календарный	Иконка
4	Алгоритм	Алгоритм Календарный	Иконка
5	Алгоритм	Алгоритм Календарный	Иконка
6	Алгоритм	Алгоритм Календарный	Иконка
7	Алгоритм	Алгоритм Календарный	Иконка
8	Алгоритм	Алгоритм Календарный	Иконка
9	Алгоритм	Алгоритм Календарный	Иконка
10	Алгоритм	Алгоритм Календарный	Иконка
11	Алгоритм	Алгоритм Календарный	Иконка
12	Алгоритм	Алгоритм Календарный	Иконка

Очистка данных

Id	Name	Mapping	Brand
0	Продукт 1	Имя продукта 1 в базе	Бренд 1
1	Продукт 2	Имя продукта 2 в базе	Бренд 2
2	Продукт 3	Имя продукта 3 в базе	Бренд 2
3	Продукт 4	Имя продукта 4 в базе	Бренд 2
4	Продукт 5	Имя продукта 5 в базе	Бренд 2
5	Продукт 6	Имя продукта 6 в базе	Бренд 2

Консолидация

В машино-читаемый формат

Очистка и сопоставление данных

Выгрузка данных в систему

Не-, полу-, разно-структурированные данные

Структурированные данные в едином формате

Агрегированные данные в машинном формате

Запрос обновленных данных

Принятие решения

Анализ данных

Изменение бизнес-процессов компании

Аналитический отчет

Для принятия управленческих решений в системах бизнес-аналитики реализованы следующие функциональные возможности:

- оперативный доступ к необходимой информации;
- интеграция, преобразование и хранение данных из различных информационных систем организации;
- конструирование многомерных хранилищ данных;
- углубленный анализ больших объемов информации;
- визуализация данных (построение наглядных отчетов);
- построение информативных отчетов различной сложности, в т.ч. регламентных отчетов;
- моделирование и прогнозирование ключевых показателей деятельности (KPI) для принятия решений (задачи «Что будет, если ..?» и «Что необходимо для ..?»).

Системы бизнес-аналитики позволяют руководителям принимать обоснованные управленческие решения при значительной экономии времени на поиск и анализ необходимой информации, а также при существенном снижении рутинной нагрузки на персонал.

Ключевыми преимуществами от внедрения BI-системы являются повышение скорости и точности составления отчетов, аналитики и планирования, повышение эффективности процессов принятия решений, увеличение удовлетворенности клиентов. Гораздо реже BI-системы позволяют добиться сокращения операционных расходов или увеличения объемом выручки.

Задачи, решаемые BI-системой

- ✓ поддержка развития бизнес-процессов и структурных изменений предприятия;
- ✓ возможность моделирования различных бизнес-ситуаций в единой информационной среде;
- ✓ проведение оперативного анализа по нестандартным запросам;
- ✓ снижение рутинной нагрузки на персонал и высвобождение времени для более глубокой аналитической работы;
- ✓ устойчивая работа при увеличении объема обрабатываемой информации, возможность масштабирования.

В части поддержки стратегического развития предприятия BI-системы обеспечивают

- ✓ оценку эффективности различных направлений бизнеса;
- ✓ оценку достижимости поставленных целей;
- ✓ оценку эффективности использования ресурсов, в том числе дочерними предприятиями;
- ✓ оценку эффективности операционной, инвестиционной и финансовой деятельности;
- ✓ бизнес-моделирование и оценку инвестиционных проектов;
- ✓ управление затратами, налоговое планирование, планирование капитальных вложений.



- Интерактивные информационные панели;
- Средства создания и выполнения нерегламентированных (ad-hoc) запросов пользователей на основе логической модели корпоративных данных;
- Подсистему регламентных отчетов (для печати);
- Инструменты пространственной аналитики (привязка к географии);
- Управление ключевыми показателями и стратегиями (scorecards);
- Активный бизнес-анализ (анализ и действия в едином замкнутом цикле, встраивание задач анализа в бизнес-процессы);
- Коллективную работу BI Collaboration (интеграция с WebCenter и порталами);
- Средства интеграции с MS Office;
- Поддержку мобильного доступа и многое другое.

Аналитики из Gartner считают, что в платформе для бизнес-анализа должны быть реализованы 15 ключевых возможностей, сгруппированных по трем основным категориям.

Возможности интеграции

VI-инфраструктура – все инструменты VI-платформы должны быть реализованы «в едином ключе», в том числе должны использовать общие данные, единую объектную модель, сквозную модель безопасности, администрирования, порталной интеграции, общий движок исполнения запросов.

Управление данными — все инструменты в рамках единой платформы должна скреплять не только модель данных, но и единые средства для поиска, получения, хранения, повторного использования и публикации объектов данных.

Разработка — VI-платформа должна предоставлять как пользователям, так и разработчикам набор специфических инструментов для создания VI-приложений, которые могут интегрироваться между собой и выстраиваться в сквозные бизнес-процессы, в том числе внедряемые в сторонние (внешние) бизнес-приложения.

Взаимодействие — VI-платформа должна содержать средства для обмена как самой информацией, так и мнениями по поводу полученных результатов. VI-приложения также должны иметь возможность на основании определенных бизнес-правил назначать задания конкретным пользователям и отслеживать их выполнение.

Представление информации

Отчетность — возможность создания форматированных и интерактивных отчетов, с развитыми механизмами для их распространения и обновления.

Контрольные (информационные) панели (dashboards) — это один из видов представления отчетности, способный представлять данные в наглядном, интуитивно понятном виде, при помощи различных шкал, показателей, индикаторов и т.п. Посредством таких контрольных панелей пользователи могут следить за текущим состоянием ключевых показателей и процессов и сравнивать их с намеченными, целевыми значениями. Такие панели позволяют извлекать операционную информацию из бизнес-приложений и делают ее доступной в реальном времени.

Произвольные (ad hoc) запросы — доступная для пользователей возможность самостоятельно (без привлечения ИТ-специалистов) создавать и выполнять уникальные, нетиповые запросы. Для реализации таких возможностей в BI-платформе должен присутствовать развитый семантический слой, позволяющий находить и извлекать нужную информацию из имеющихся источников.

Интеграция с приложениями Microsoft Office — в некоторых случаях BI-платформа используется как промежуточный инструмент для выполнения аналитических задач с соблюдением правил корректности и безопасности данных. При этом в качестве клиентской части BI-системы могут выступать продукты семейства Microsoft Office (в частности, Excel).

Представление информации

Возможности поиска - добавление поискового индекса как к структурированным, так и к неструктурированным источникам данных и объединение их в классифицированную структуру с определенными измерениями (зачастую использующими семантический слой VI) с тем, чтобы пользователи могли с легкостью ориентироваться и изучать данные посредством поискового (как у Google) интерфейса.

Мобильные возможности – этот функциональный блок объединяет инструментарий, позволяющий доставлять отчеты и контент аналитических панелей на мобильные устройства (такие как смартфоны и планшеты), а также использовать интерактивные возможности девайсов (например, нажатие) и другие опции, обычно не доступные на десктопах и ноутбуках (геолокация и другие).

Анализ данных

Оперативная аналитическая обработка данных (OLAP, OnLine Analytical Processing) — поддержка OLAP значительно ускоряет процессы обработки запросов и выполнения расчетов, обеспечивая анализ данных в различных срезах (такой стиль анализа обозначается термином «slicing and dicing», что дословно переводится как «нарезание на кубики и ломтики», то есть в продольном и поперечном направлении). Такие инструменты также позволяют осуществлять планирование, использовать анализ по типу "что, если...". OLAP при этом может быть реализован посредством различных архитектур данных и хранилищ, как дисковых, так и in-memory.

Интерактивная визуализация — максимально наглядное представление данных с использованием различных интерактивных изображений, схем и графиков (вместо привычных таблиц с обычными строками и столбцами).

Предикативное моделирование и data mining (интеллектуальный анализ данных) — эти инструменты призваны помочь компаниям классифицировать данные, формировать собственные номинальные и количественные шкалы, а также использовать для их анализа развитый математический аппарат.

Анализ данных

Карты показателей (scorecards) — позволяют на основании определенных ключевых показателей (отображаемых на контрольных панелях) контролировать ход выполнения стратегических планов и достижения ключевых показателей эффективности (KPI). Тем самым операционные показатели «привязываются» к целевым стратегическим индикаторам. Для дальнейшего, более детального анализа эти показатели могут расшифровываться при помощи дополнительных отчетов. Такие механизмы позволяют реализовывать различные методологии управления, такие как системы сбалансированных показателей (BSC).

Перспективное моделирование, симуляторы и оптимизация - инструментарий предназначен для поддержки принятия решений в условиях выбора правильного значения переменной как детерминированных, так и при моделировании случайных процессов.



Средства BI сегодня весьма разнообразны. Они включают:

- инструменты для построения реляционных хранилищ данных - специальным образом спроектированных баз данных, позволяющих с высокой скоростью выполнять запросы на выбор данных;
- серверные и клиентские средства построения многомерных хранилищ данных, содержащих в нереляционной структуре агрегатные данные (суммы, средние значения);
- клиентские приложения для предоставления пользовательских интерфейсов к реляционным и многомерным хранилищам данных;
- средства создания решений на основе подобных хранилищ, анализа многомерных и реляционных данных, генерации отчетов по многомерным и реляционным данным.

Основные возможности BI-систем развиваются по четырем основным направлениям: **хранение данных**, **интеграция данных**, **анализ данных** и **представление данных**.

Данные, используемые для бизнес-анализа, организуются в специальные хранилища (data warehouse, DW). Эти данные должны отражать текущую, реальную и полную картину бизнеса. Информация в хранилище данных (включая исторические данные) собирается из различных операционных (транзакционных) систем и структурируется специальным образом для более эффективного анализа и обработки запросов (в отличие от обычных баз данных, где информация организована таким образом, чтобы оптимизировать время обработки текущих транзакций).

Хранилища данных содержат огромные объемы информации, охватывающей все доступные стороны деятельности предприятия и позволяющие рассматривать все аспекты функционирования бизнеса в совокупности. Для решения более узких, конкретных задач из общего хранилища могут вычленяться подмножества данных – так называемые витрины данных (data marts).

Хранилища данных - это предметно-ориентированные, интегрированные, стабильные, поддерживающие хронологию наборы данных, используемые для поддержки принятия управленческих решений.

Данные в хранилище попадают из оперативных систем, а также из внешних источников. По аналогии с "материальными" хранилищами, хранилища данных предусматривают такие операции, как сбор данных ("приход материалов на склад"), хранение данных ("складской запас"), перемещение в витрины данных ("отгрузка товаров в розничную сеть").

Основные требования к хранилищам данных:

- поддержка высокой скорости получения данных из хранилища;
- поддержка внутренней непротиворечивости данных;
- возможность получения и сравнения так называемых срезов данных (slice and dice);
- наличие удобных утилит просмотра данных в хранилище;
- полнота и достоверность хранимых данных;
- поддержка качественного процесса пополнения данных.

Выполнить все перечисленные требования в рамках одного и того же продукта зачастую не удастся. Поэтому для реализации хранилищ данных обычно используется несколько продуктов, одни из которых представляют собой собственно средства хранения данных, другие - средства их извлечения и просмотра, третьи - средства их пополнения и т. д.

В отличие от так называемых оперативных баз данных, с которыми работают приложения, изменяющие данные, реляционные хранилища данных проектируются таким образом, чтобы добиться минимального времени выполнения запросов на чтение (у оперативных же баз данных чаще всего минимизируется время выполнения запросов на изменение данных). Обычно данные копируются в хранилище из оперативных баз данных согласно определенному расписанию. Типичная структура хранилища данных существенно отличается от структуры обычной реляционной СУБД. Как правило, эта структура денормализована (это повышает скорость выполнения запросов) и может допускать избыточность данных.

Для формирования и поддержания хранилищ данных используются так называемые **ETL-средства** – инструменты для извлечения данных (extract), преобразования данных (transform), то есть приведения их к необходимому формату, обработки в соответствии с определенными правилами, комбинировании с другими данными и т.п., а также для загрузки данных (load), записи данных в хранилище или в другую базу.

Инструменты для работы с SQL (structured query language), позволяющие пользователям напрямую обращаться к данным. В последнее время инструменты для формирования и обработки запросов стали более «дружественными», ориентированными на неподготовленных бизнес-пользователей (а не на квалифицированных ИТ-специалистов).

Корпоративные порталы, обеспечивающие взаимосвязь на уровне данных и бизнес-процессов. Такие порталы обеспечивают совместный доступ к информации. (EAI, enterprise application integration).

Для всестороннего анализа данных в современных BI используются **OLAP-инструменты** (online analytical processing). Они позволяют рассматривать различные срезы данных, в том числе временные, позволяющие выявлять различные тренды и зависимости (по регионам, продуктам, клиентам и т.п.).

В наиболее развитые BI-решения включены блоки для **глубокого исследования данных** (data mining). Иногда этот термин ошибочно используют для обозначения инструментов, позволяющих по-новому представить (отобразить) информацию, однако на самом деле эти инструменты призваны помочь в выявлении скрытых (неочевидных) закономерностей, моделей, составления прогнозов. Они основаны на сканировании и статистической обработке огромных массивов данных и в конечном итоге призваны облегчить принятие правильных и обоснованных стратегических решений благодаря анализу различных вариантов развития событий. В качестве инструментов используются нейронные сети, деревья решений.

Инструменты OLAP являются аналитическими инструментами, которые первоначально были основаны на многомерных базах данных (МБД). МБД — это базы данных, сконструированные специально для поддержки анализа количественных данных с множеством измерений, содержат данные в «чисто» многомерной форме. Большинство приложений включают измерение времени, другие измерения могут касаться географии, организационных единиц, клиентов, продуктов и др.

OLAP позволяет организовать измерения в виде иерархии. Данные представлены в виде гиперкубов (кубов) — логических и физических моделей показателей, коллективно использующих измерения, а также иерархии в этих измерениях. Некоторые данные предварительно агрегированы в БД, другие рассчитываются «на лету».

Средства OLAP позволяют исследовать данные по различным измерениям. Пользователи могут выбрать, какие показатели анализировать, какие измерения и как отображать в кросс-таблице, обменивать строки и столбцы «pivoting», затем сделать срезы и вырезки («slice&dice»), чтобы сконцентрироваться на определенной комбинации размерностей. Можно изменять детальность данных, двигаясь по уровням с помощью детализации и укрупнения «drill down/ roll up», а также кросс-детализации «drill across» через другие измерения.

Многомерные хранилища обычно содержат агрегатные данные (например, суммы, средние значения, количество значений) для различных выборок. Чаще всего такие агрегатные функции образуют многомерный набор данных, называемый кубом, оси которого (называемые измерениями) содержат параметры, а ячейки - зависящие от них агрегатные данные (иногда их называют мерами). Вдоль каждой оси данные могут быть организованы в иерархии, отражающие различные уровни их детализации.

2003			
2004			
2005	VISA	Mastercard	Cash
Argentina	11806.28	9190.48	1263.9
Austria			4039.5
Belgium	1745.42	1207.28	14924.12
Brazil			5208.28
Canada	2952.4		
Denmark	1739.76	1376	
Finland	5470.98	3538.92	2328.46
France	11927.48	9823.43	11052.28
Germany	2208.62	1739.6	4681.16
Ireland		330.9	608
Italy	2139.1		1357.6
Mexico			786
Norway	459		
Poland	1268.3	716.72	285.12
Portugal	236.5	220.3	2235.8
Spain	3021.23	2380	1488.8
Sweden	2490.5		1628.32
Switzerland	5094.88	1520.8	901.2

Как правило, агрегатные данные получают путем выполнения серии запросов на группировку данных при этом довольно часто в качестве источника данных для подобных запросов выступают реляционные хранилища данных.

Цель хранения агрегатных данных - сократить время выполнения запросов, поскольку в большинстве случаев для анализа и прогнозов интересны не детальные, а суммарные данные.

Как исходные, так и агрегатные данные могут храниться либо в реляционных, либо в многомерных структурах. Поэтому в настоящее время применяются три способа хранения данных:

- MOLAP (Multidimensional OLAP) - исходные и агрегатные данные хранятся в многомерной базе данных;
- ROLAP (Relational OLAP) - исходные данные остаются в той же реляционной базе данных, где они изначально находились, агрегатные же данные помещают в специально созданные для их хранения служебные таблицы в той же базе данных;
- HOLAP (Hybrid OLAP) - исходные данные остаются в той же реляционной базе данных, где они изначально находились, а агрегатные хранятся в многомерной базе данных.

Некоторые OLAP-средства поддерживают хранение данных только в реляционных структурах, некоторые - только в многомерных. Однако большинство современных серверных OLAP-средств поддерживают все три способа хранения данных. Выбор способа хранения зависит от объема и структуры исходных данных, требований к скорости выполнения запросов и частоты обновления OLAP-кубов.

Выпущенные в течение последних лет СУБД ведущих производителей - IBM, Microsoft, Oracle, содержат средства для создания многомерных хранилищ данных. Существуют и отдельные продукты для создания OLAP-хранилищ - их выпускают компании Hyperion, Sybase, Business Objects и некоторые другие.

Интеллектуальный анализ данных (data mining) представляет собой процесс обнаружения корреляции, тенденций, шаблонов, связей и категорий. Она выполняется путем тщательного исследования данных с использованием технологий распознавания шаблонов, а также статистических и математических методов. При этом многократно выполняются различные операции и преобразования над сырыми данными (отбор признаков, стратификация, кластеризация, визуализация и регрессия), которые предназначены:

- 1) для нахождения представлений, которые являются интуитивно понятными для людей, которые, в свою очередь, лучше понимают бизнес-процессы, лежащие в основе их деятельности;
- 2) для нахождения моделей, которые могут предсказать результат или значение определенных ситуаций, используя исторические или субъективные данные.

В отличие от использования OLAP разведка данных в значительно меньшей степени направляется пользователем, вместо этого полагается на специализированные алгоритмы, которые устанавливают соотношение информации и помогают распознать важные (и ранее неизвестные) тенденции, свободные от предвзятости и предположений пользователя.

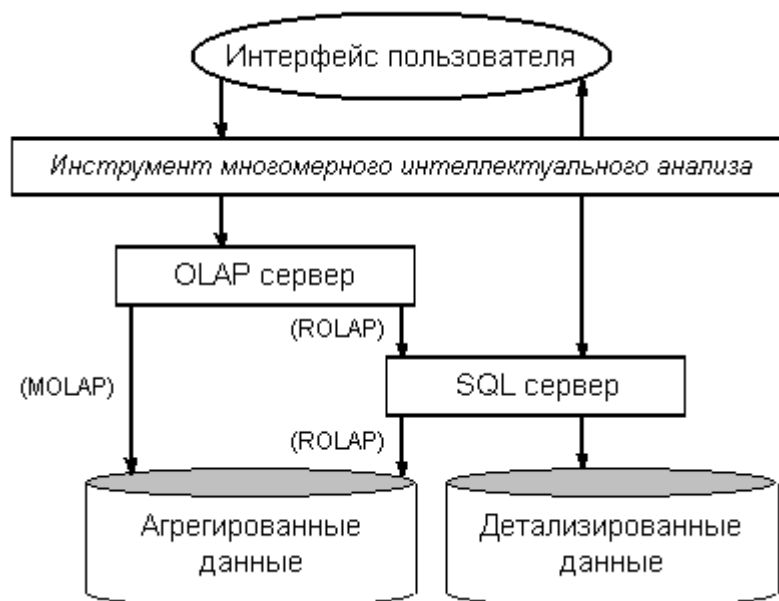
В общем случае процесс Data Mining состоит из трёх стадий :

- 1) выявление закономерностей (свободный поиск);
- 2) использование выявленных закономерностей для предсказания неизвестных значений (прогностическое моделирование);
- 3) анализ исключений, предназначенный для выявления и толкования аномалий в найденных закономерностях.

Иногда в явном виде выделяют промежуточную стадию проверки достоверности найденных закономерностей между их нахождением и использованием (стадия валидации).

Примерами искомой информации могут служить сведения о том, какие категории покупателей чаще всего приобретают тот или иной товар, какая часть покупателей одного конкретного товара приобретает другой конкретный товар, какая категория клиентов чаще всего вовремя не выплачивает предоставленный кредит. Подобного рода информация обычно используется при прогнозировании, стратегическом планировании, анализе рисков, и ценность ее для предприятия очень высока. Если при статистическом анализе или при применении OLAP обычно формулируются вопросы типа "Каково среднее число клиентов банка, не вернувших вовремя кредит, среди неженатых мужчин от 40 до 50 лет?", то применение Data Mining, как правило, подразумевает ответы на вопросы типа "Существует ли типичная категория клиентов, не возвращающих вовремя кредиты?". При этом именно ответ на второй вопрос нередко обеспечивает принятие успешного бизнес-решения.

Оперативная аналитическая обработка и интеллектуальный анализ данных - две составные части процесса поддержки принятия решений. Но сегодня большинство систем OLAP заостряет внимание только на обеспечении доступа к многомерным данным, а большинство средств data mining, работающих в сфере закономерностей, имеют дело с одномерными перспективами данных. Эти два вида анализа должны быть тесно объединены, то есть системы OLAP должны фокусироваться не только на доступе, но и на поиске закономерностей.



Отчет представляет собой документ, содержимое которого динамически формируется на основе информации, содержащейся в базе данных.

Генераторы запросов и отчетов — типично «настольные» инструменты, предоставляющие пользователям доступ к базам данных, выполняющие некоторый анализ и формирующие отчеты.

Запросы могут быть как незапланированными (ad hoc), так и иметь регламентный характер.

На рынке ПО сейчас представлено немало средств создания отчетов: как отдельных продуктов, так и входящих в состав средств разработки приложений или СУБД, и реализованных в виде либо серверных служб, либо клиентских приложений. Как правило, средства создания отчетов поддерживают широкий спектр универсальных механизмов доступа к данным (ODBC, OLE DB, ADO.NET), нередко - средства прямого доступа к наиболее популярным СУБД с помощью их клиентских API, содержат средства деловой графики, интегрируются с офисными приложениями, позволяют публиковать отчеты в Интернете, включают классы или компоненты, предназначенные для создания приложений, реализующих (наряду с другими возможностями) генерацию отчетов.

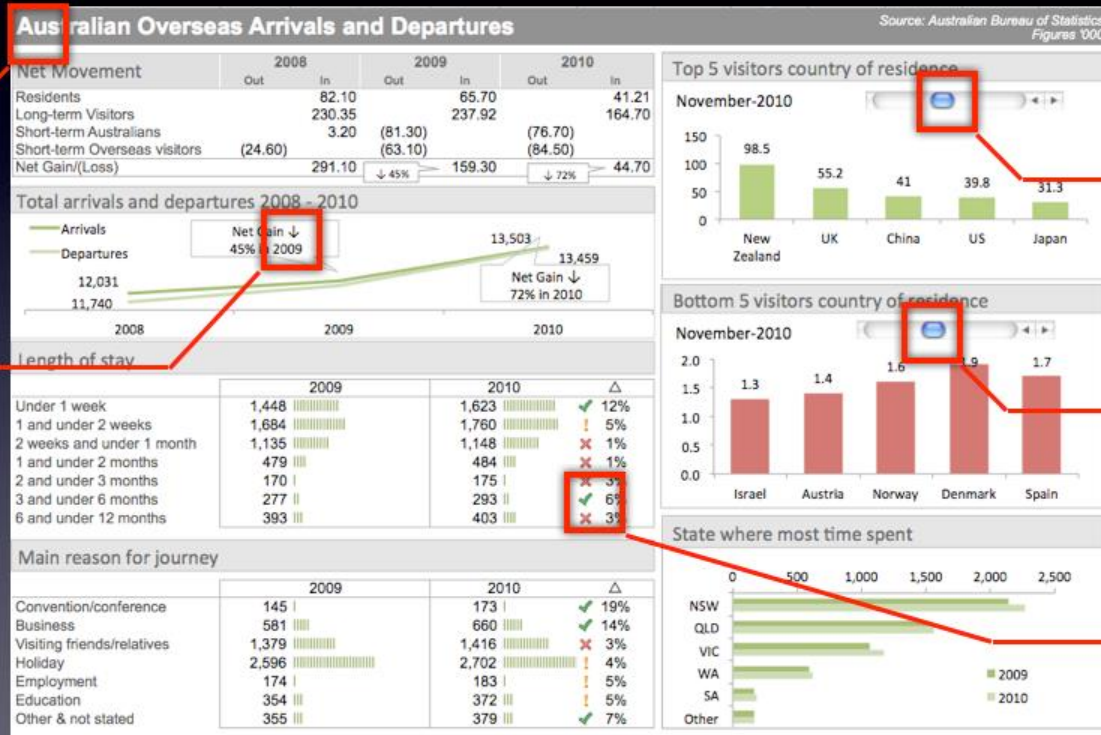
Для представления данных используются различные графические средства – отчеты, графики, диаграммы, настраиваемые при помощи параметров.

Общепринятым средством визуализации данных в современных BI-решениях являются информационные (контрольные, приборные) панели (dashboards), на которых результаты отображаются в виде шкал и индикаторов, позволяющих контролировать текущие значения выбранных показателей, сравнивать их с критическими (минимально\максимально допустимыми) значениями и таким образом выявлять потенциальные угрозы для бизнеса.

Контрольные панели считаются одним из наиболее удобных способов представления информации о «состоянии здоровья» бизнеса. Они позволяют уместить на экране всю важнейшую информацию о текущих операциях, выявленных и потенциальных проблемах.

Контрольные панели, как и карты показателей (scorecards), основаны на анализе ключевых показателей эффективности (KPIs). Однако, как правило, контрольные панели отображают текущее состояние общих показателей, а карты показателей предназначены для сравнения текущих показателей с плановыми, целевыми, и отображают динамику изменения этих показателей во времени. Карты показателей обычно бывают более персонализированными, настраиваются в зависимости от ролей и задач конкретного пользователя (финансовое управление, снабжение, продажи и т.п.). При необходимости все эти показатели могут быть детализированы при помощи дополнительных отчетов, графиков и диаграмм.

Dashboard Example



Fits Screen

- One page report
- Gives overview
- Inter-relationships

Charts

- Visual presentation
- In-chart comments
- Compare trends

Slider 1

- Interactive element
- Mouse controlled

Slider 2

- Interactive element
- Mouse controlled

Checkmarks

- Visual indicators

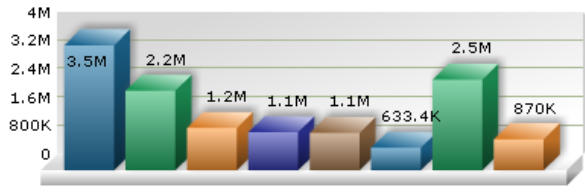


Business Intelligence. Представление данных

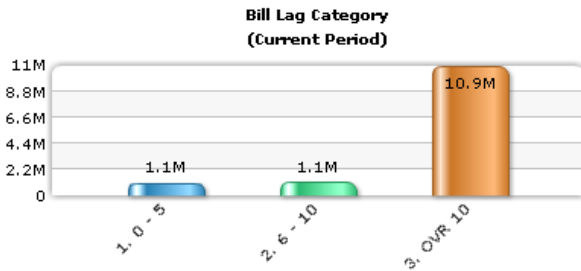
Finance Healthcare Legal Manufacturing Marketing Projects Sales Service Management

Add a dashboard item:

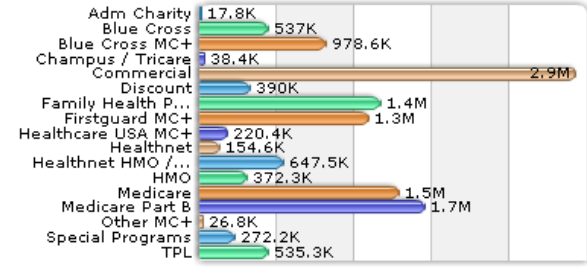
Hospital Aging Summary



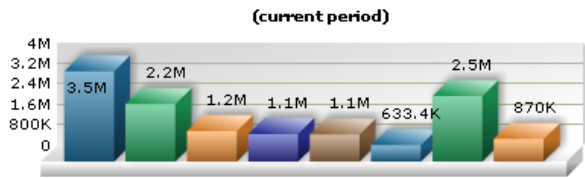
Hospital Bill Lag Category Summary



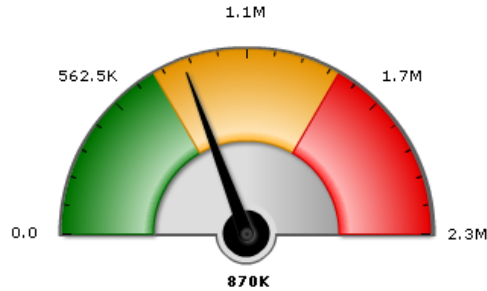
Hospital Payer Summary



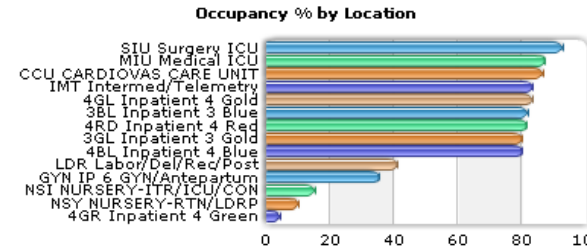
McKesson ATB Aging Summary



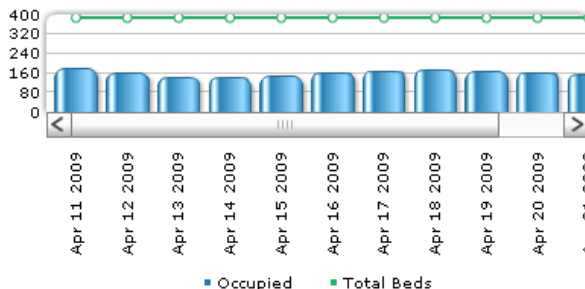
McKesson ATB Aging (Over 360 Days)



McKesson Occupancy Percentage



McKesson Occupancy Trend

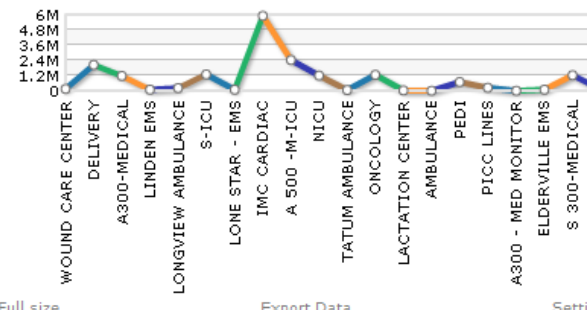


Meditech Monthly Overtime by Dept

Showing 1 to 88 of 88

Dept Number	Dept Name	Overtime Hours
16710	E.R.	1593.3
16721	LONGVIEW AMBULANCE	1486.4
158450	HOUSEKEEPING	1094.0
16730	GSMC - ALS EMS	895.6
16727	HARLETON EMS	714.4
17420	PULMONARY SERVICES	661.9
16745	ORE CITY EMS	661.4
16290	A 600 IMC	648.6
16732	LINDEN EMS	648.3
16011	REHAB UNIT	603.1
16009	SHORT STAY UNIT	585.5
16280	IMC CARDIAC	557.8
16729	LONE STAR - EMS	519.8
16728	DAINGERFIELD EMS	434.8
17713	LAB	422.0

Meditech MTD Revenue by Dept (Top 20)

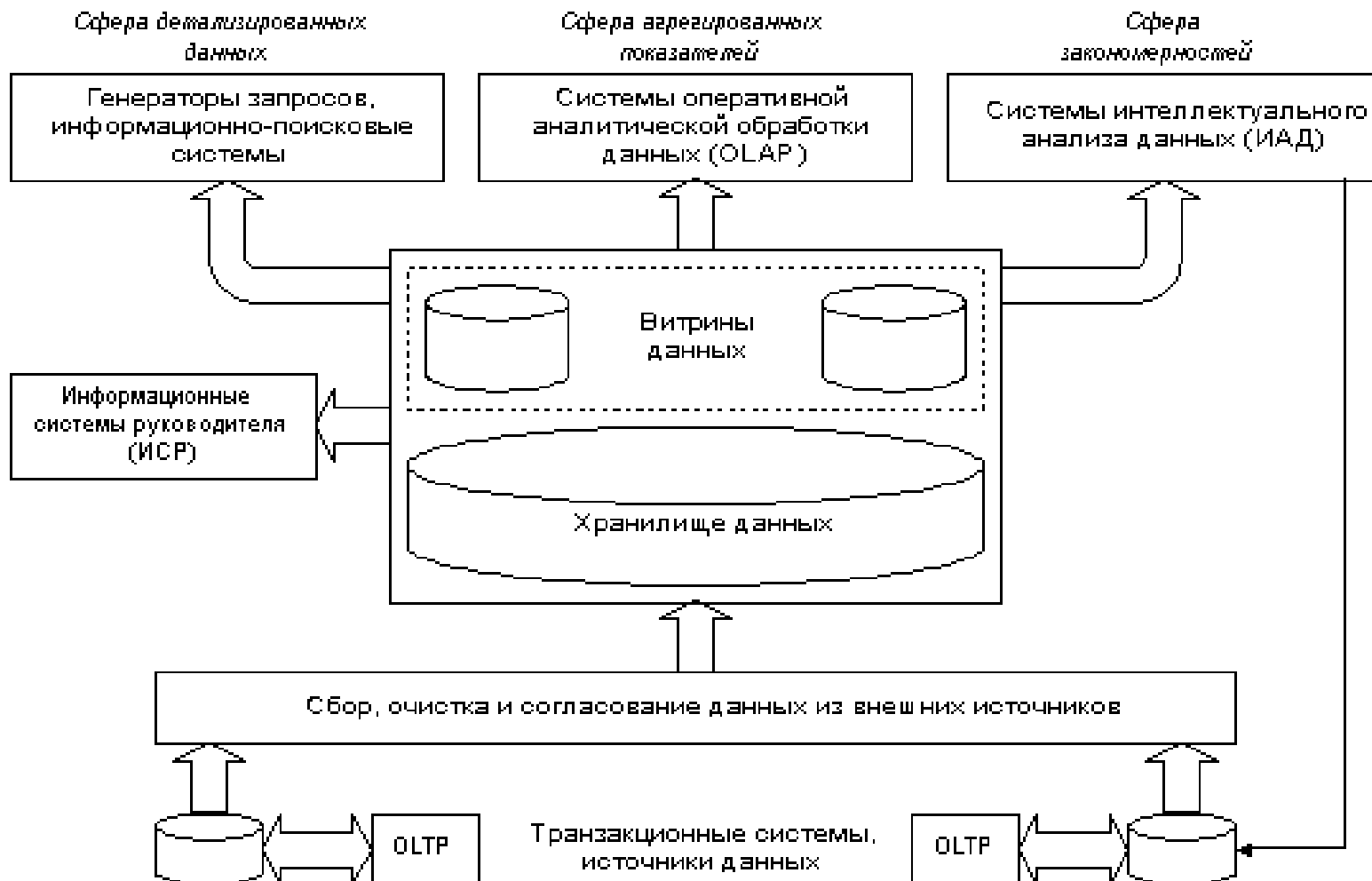




В информационную среду бизнес-аналитики поступает первичный материал — данные, которые затем перерабатываются в автоматизированных системах и информационных продуктах.

В процессе переработки происходит переход от данных к информации. Хранилище данных извлекает данные из множества транзакционных или оперативных систем, а затем интегрирует и хранит данные в специализированной БД. Например, в хранилище данных могут приводиться в соответствие и объединяться пользовательские записи из четырех оперативных систем (приложений для обработки заказов, обслуживания, продаж и поставок). Такой процесс извлечения и интеграции преобразует данные в новый информационный продукт — информацию.

Затем пользователи, работающие с аналитическими инструментами (например, для создания запросов, отчетов, OLAP-анализа и выполнения операций интеллектуального анализа данных), обращаются к данным из хранилища данных и анализируют ее. Таким образом, определяются тенденции, структуры и исключения. Аналитические инструменты помогают пользователям преобразовать информацию в знания.



Business Performance Management (BPM) - это методология, направленная на оптимизацию реализации стратегии и состоящая из набора интегрированных циклических аналитических процессов, которые поддерживаются соответствующими технологиями и имеют отношение как к финансовой, так и к операционной информации. BPM позволяет предприятию определять, измерять и управлять эффективностью своей деятельности, направленной на достижение стратегических целей. Ключевые финансовые и операционные процессы BPM включают планирование, консолидацию и отчетность, анализ ключевых показателей эффективности и их распространение в рамках организации.

- управление эффективностью бизнеса (Business Performance Management, BPM);
- управление эффективностью деятельности предприятия (Enterprise Performance Management, EPM);
- управление эффективностью деятельности корпорации (Corporate Performance Management, CPM);
- стратегическое управление предприятием (Strategic Enterprise Management, SEM).

Методологию сбалансированных систем показателей (Balanced Scorecard, BSC), определяют как инструмент, позволяющий трансформировать миссию и стратегию организации в исчерпывающий набор показателей эффективности, которые служат основой для системы стратегического управления и контроля.

Система Balanced Scorecard позволяет выстроить сквозную связь между стратегией и тактикой организации, в результате чего задача трансформации стратегии в реальные действия оказывается решена. Кроме того, такая система позволяет не только формализовать стратегию, но и контролировать успешность ее реализации за счет измерителей и значений целевых показателей.

Balanced Scorecard по праву можно назвать наиболее популярной среди методик стратегического управления. Но это не означает отсутствия других методов и подходов, многие из которых также получили достаточно широкое распространение и признание. Примерами таких разработок могут служить методика управления стоимостью компании (Value Based Management, VBM).

CHALLENGERS

LEADERS

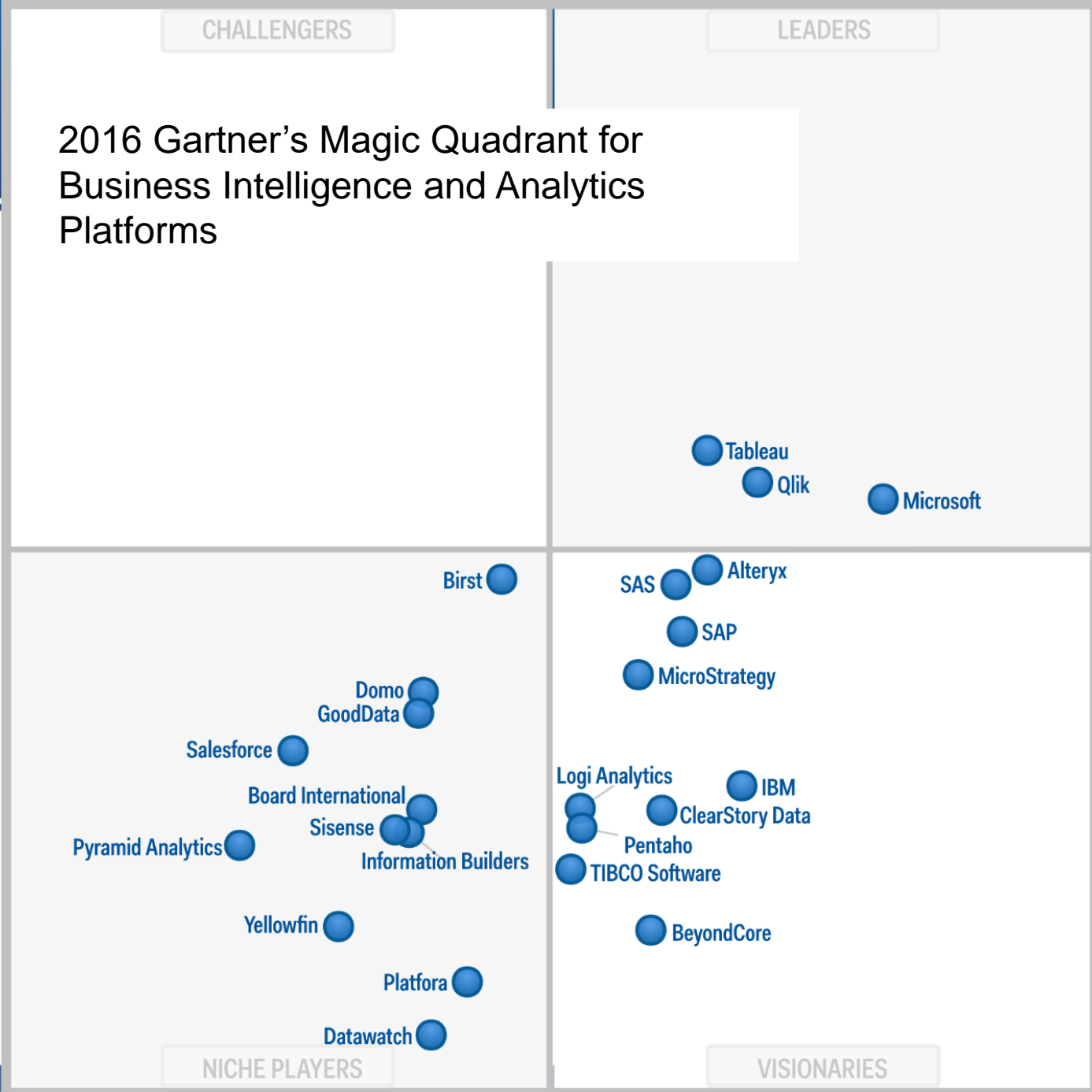
2016 Gartner's Magic Quadrant for Business Intelligence and Analytics Platforms

↑
ABILITY TO EXECUTE

COMPLETENESS OF VISION →

NICHE PLAYERS

VISIONARIES



As of February 2016

Tableau, Qlik, Microsoft

- Интуитивно понятный пользовательский интерфейс
- Предоставление некоторых версий своих продуктов в бесплатное и неограниченное использование
- Доступ к бесплатным учебникам и профессиональному форуму, поощрение совместного сотрудничества и обмена знаниями
- Упрощение процесса развёртывания

According to Gartner, the enterprise analytics industry has surpassed the tipping point, and has nearly finished its evolution from a market based heavily on reporting to one that is more business-centric and user friendly, offering self-service analytics and allowing more users to get their hands on tools that drive insights. The majority of solution buying in BI and analytics now comes in the form of modern platforms which focus on user engagement, and this fact has essentially reorganization the vendor landscape.

Centralized provisioning and tightly governed platforms are becoming a thing of the past, and are being counterbalanced and replaced by tools that promote analytical agility and business user autonomy.

<https://techdemic.com/tableau-public>

<https://techdemic.com/qlikview-architecture>

- Развитие хранилищ данных
- Неструктурированные данные
- Self-service аналитика
- Работа с новыми типами данных
- Рост рынка Cloud BI
- Развитие визуальной аналитики
- Предикативный анализ
- Повсеместное использование мобильного BI
- Совместная работа с данными
- Развитие аналитических инструментов

Целый ряд факторов будет способствовать дальнейшему росту рынка аутсорсинговых BI услуг. Во-первых, повсеместно отмечается нехватка квалифицированных специалистов, которые требуются для реализации BI-инициатив, в том числе математиков, непосредственно бизнес-аналитиков, специалистов по моделям данных, статистиков и научных сотрудников различного профиля.

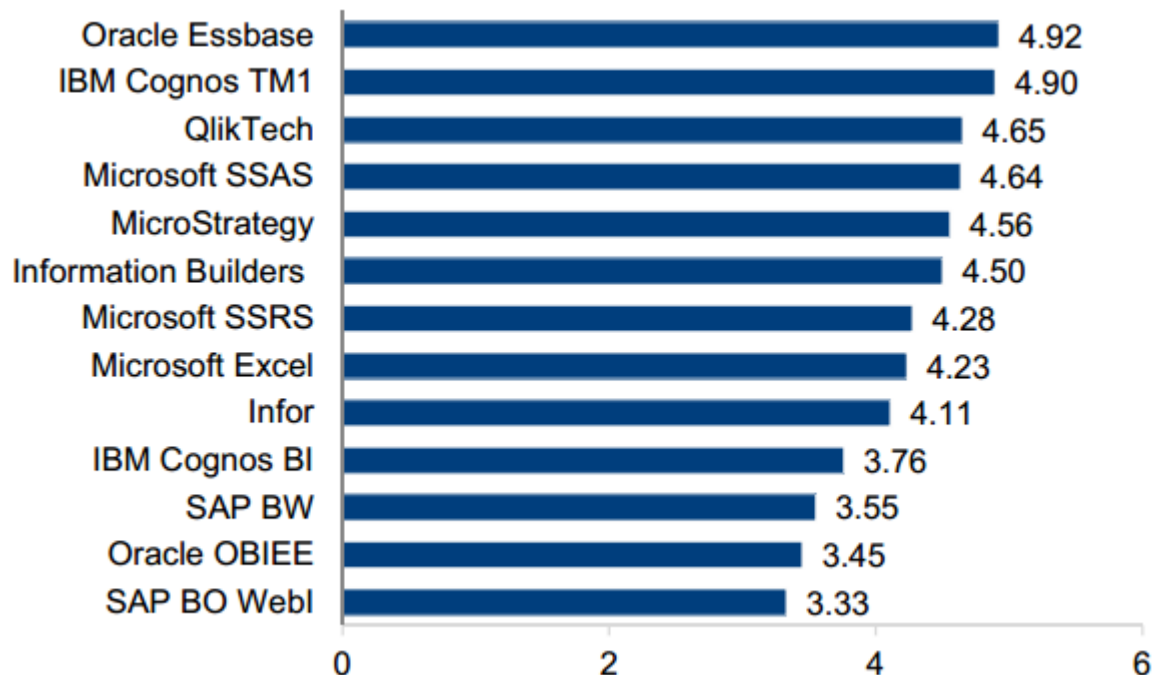
С другой стороны, развитие новых технологий и тесная интеграция BI с социальными сетями и мобильными платформами делают инструменты бизнес-анализа доступными для широких категорий конечных пользователей. Однако, конечным пользователям сложно организовывать команды внедрения, а затраты на необходимую инфраструктуру BI-систем зачастую слишком велики для этой категории клиентов.

<http://www.itproportal.com/2015/11/19/top-10-business-intelligence-trends-for-2016/>

Основные критерии выбора BI-систем - как организации должны ранжировать критерии выбора BI-системы и то, как это происходит на самом деле.

Criterion	How buyers should chose		Difference	How buyers choose
Fast performance	1	↓	-3	4
Proof of concept faster or better	2	↓	-4	6
Ease of use for application builders	3	↓	-4	7
Ease of use for end-users	4	↑	+2	2
Range of server platforms supported	5	↓	-13	18
Functionality/product features	6	↑	+5	1
Large data handling capacity	7	↓	-4	11
Web architecture	8	↓	-7	15
Availability of local support	9	↓	-4	13
User scalability	10	↑	+1	9
Low price	11	↑	+6	5
Integrates with other products in use	12	↑	+9	3
Vendor relationship/reputation	13	↓	-3	16
Chosen vendor did a better sales job	14	↓	-3	17
Product reputation	15	↑	+3	12
Bundled with another product	16	↑	+6	10
Corporate standard	17	↑	+9	8
Other	18	↑	+4	14

Критерии функциональности и производительности системы должны ставиться при выборе решения для бизнес-аналитики во главу угла. Выбирая или обновляя систему для бизнес-анализа, следует продумать способы хранения и интеграции данных, средства визуализации и аналитики.



Степень удовлетворенности пользователей крупных BI-платформ



НАЦИОНАЛЬНЫЙ ИССЛЕДОВАТЕЛЬСКИЙ
УНИВЕРСИТЕТ

Спасибо за внимание!

101000, Россия, Москва, Мясницкая ул., д. 20

Тел.: (495) 621-7983, факс: (495) 628-7931

www.hse.ru