## Основные термины, понятия и определения

***База данных (БД)*** ‑ организованная в соответствии с определёнными правилами и поддерживаемая в памяти компьютера совокупность данных, характеризующая актуальное состояние некоторой предметной области и используемая для удовлетворения информационных потребностей пользователей.

***Система управления базами данных (СУБД) -*** совокупность программных, лингвистических и методических средств общего или специального назначения, обеспечивающих управление созданием и использованием БД.

***OLTP -*** *(Online Transaction Processing)*, Способ организации БД, при котором система работает с небольшими по размерам транзакциями, но идущими большим потоком, и при этом клиенту требуется от системы минимальное время отклика.

***OLAP (****Online Analytical Processing) -* Способ организации БД, заключающийся в подготовке суммарной (агрегированной) информации на основе больших массивов данных, структурированных по многомерному принципу. Эта БД предназначена для быстрой обработки аналитических запросов.

***Предметная область базы данных*** ‑ это та часть реального мира, о которой информационная система хранит, собирает и анализирует информацию.

***Приложение*** ‑ комплекс программ, автоматизирующий обработку данных для прикладной задачи.

***Сущность (****объект****)*** - предмет, человек или событие, о котором соби­рается и хранится информация.

***Запись*** – данные об одном конкретном экземпляре сущности - предмете, человеке или событии. Запись состоит из полей.

Пример. Сущность “СОТРУДНИК”, экземпляр сущности ‑ данные об Иванове И.

***Поле записи (****атрибут, столбец таблицы****)*** ‑ отдельная характеристика (свойство) сущности.

***Домен*** ‑ область значений одного или нескольких атрибутов.

***Таблица*** - совокупность записей с фиксированным числом полей. Таблица содержит данные о всех экземплярах сущности.

 Далее приведено соответствие терминов в различных областях знаний.

|  |  |  |
| --- | --- | --- |
| База данных | Реляционная алгебра | EXCEL |
| Таблица | Отношение | Таблица |
| Запись | Кортеж | Строка |
| Поле | Атрибут | Столбец |

***Ключ*** - поле или суперпозиция несколько полей и\или их фрагментов (называемых *ключевыми*), однозначно определяющая запись таблицы.

***Первичный ключ*** – поле однозначно определяющее запись (экземпляр сущности), используется для связи с другими таблицами. Значения этого поля должны быть уникальны в пределах таблицы. Такой ключ может быть только один на таблицу. В реальных БД в качестве первичных используются суррогатные ключи.

***Вторичный ключ (****кандидат****)*** – поле, содержащее значение уникального ключа, который может быть первичным, но не является оным, поскольку первичный ключ только один.

***Внешний ключ*** ‑ поле дочерней таблицы, которое содержит значение первичного ключа родительской таблицы (атрибут “Код подразделения” сущности “СОТРУДНИК” является внешним ключом, т. к. он содержит значение первичного ключа в родительской таблице “ПОДРАЗДЕЛЕНИЕ”).

***Родительская таблица*** *(сущность)* – таблица содержащая первичный ключ используемой в связи (см.), значение которого записывается в поле внешнего ключа дочерней таблицы.

***Дочерняя таблица*** *(сущность) -* таблица в которой присутствует внешний ключ используемой в связи (см.) и содержащий значение первичного ключа родительской таблицы.

***Связь*** между родительской и дочерней таблицами (сущностями) это сопоставление записей, которое производится по равенству значений внешнего и первичного ключей в обеих таблицах.

 ***Связь типа “Один-к-одному”, или бинарная связь (1:1)***. Одной записи (экземпляру) родительской таблицы (сущности) соответствует одна запись (экземпляр) дочерней таблицы (сущности).

 ***Связь типа “Один-ко-многим” (1:М)***. Одной записи родительской таблицы соответствует несколько записей дочерней таблицы. При этом внешний ключ участвующий в связи не уникален.

***Связь типа “Многие-ко-многим” (М:М)***. Одному экземпляру родительской сущности соответствуют несколько экземпляров дочерней сущности и наоборот одному экземпляру дочерней сущности соответствуют несколько экземпляров родительской сущности. Наличие такой связи между сущностями говорит о том, что, какая-то значимая сущность еще не выявлена. Именно эта сущность и преобразует связь типа “Многие-ко-многим” к двум связям типа “Один-ко-многим”.

***Индекс –*** объект БД (*таблица*), содержащий записи, каждая из которых содержит два значения: индекса (ключа индексируемой таблицы) и адреса записи индексируемой таблицы со значением данного индекса.

***Словарь*** - централизованное хранилище сведений об объектах БД: таблицах, связях, форматах и диапазонах значений полей, представлениях, хранимых процедурах и т.д. Словарь — это БД содержащая всю служебную информацию СУБД.

***Запрос -*** это выражение, определяющее выборку нужных данных из одной или нескольких таблиц базы, удаление или изменение этих данных, создание новых таблиц. Все реляционные СУБД понимают структурированный язык запросов SQL.

***Согласованность (целостность) базы данных*** - соответствие имеющейся в [базе данных](https://ru.wikipedia.org/wiki/%D0%91%D0%B0%D0%B7%D0%B0_%D0%B4%D0%B0%D0%BD%D0%BD%D1%8B%D1%85) информации её внутренней логике, структуре и всем явно заданным правилам. Каждое правило, налагающее некоторое ограничение на возможное состояние базы данных, называется [ограничением целостности](https://ru.wikipedia.org/w/index.php?title=%D0%9E%D0%B3%D1%80%D0%B0%D0%BD%D0%B8%D1%87%D0%B5%D0%BD%D0%B8%D0%B5_%D1%86%D0%B5%D0%BB%D0%BE%D1%81%D1%82%D0%BD%D0%BE%D1%81%D1%82%D0%B8&action=edit&redlink=1).

Примеры правил: каждый студент должен состоять к учебной группе; количество знаков в телефонном номере не должно превышать 15; баланс банка должен сходится; число проданных билетов не превышает число посадочных мест и т. д.

***Транзакция*** - процесс, который переводит базу данных из одного согласованного состояния, в другое согласованное состояние. Допускается, что в процессе работы транзакции согласованность может нарушаться, но извне транзакции этого не видно. По сути это группа последовательных операций с базой данных, обычно записанная на языке SQL, которая представляет собой логическую неделимую единицу работы с данными. Транзакция может быть выполнена либо целиком и успешно, соблюдая согласованность данных и независимо от параллельно идущих других транзакций, либо не выполнена вообще и тогда она не должна произвести никакого эффекта. Транзакции обрабатываются СУБД и другими транзакционными системами, в процессе работы которых создаётся история транзакций (журналы транзакций). Транзакции необходимы для поддержания согласованности базы данных.

***Блокировка записей –*** процесс предотвращения доступа к данным изменяемым транзакцией, пока она не будет закончена. Средства блокировки обеспечивают безопасность данных в случае параллельной обработки данных.

***Триггер* -** хранимая процедура особого типа, которую пользователь не вызывает непосредственно, а исполнение которой обусловлено действием по модификации данных: добавлением INSERT, удалением DELETE строки в заданной таблице, или изменением UPDATE данных в определенном столбце заданной таблицы реляционной базы данных. Триггеры применяются для обеспечения согласованности данных и реализации сложной бизнес-логики. Триггер запускается сервером автоматически при попытке изменения данных в таблице, с которой он связан. Все производимые им модификации данных рассматриваются как выполняемые в транзакции, в которой выполнено действие, вызвавшее срабатывание триггера. Соответственно, в случае обнаружения ошибки или нарушения согласованности данных в результате работы триггера может произойти откат этой транзакции.

***Нормализация*** – это разбиение таблицы на две или более, обладающие лучшими свойствами при включении, изменении и удалении данных. Окончательная цель нормализации сводится к получению такого проекта базы данных, в котором каждый атрибут появляется лишь в одном месте, т.е. исключена избыточность информации. Это делается не столько с целью экономии памяти, сколько для исключения возможной противоречивости хранимых данных и предсказуемости поведения системы во время эксплуатации. Применялось в прошлом веке при проектировании БД.

***Реляционная база данных*** представляет собой набор взаимосвязанных двухмерных таблиц. Эта модель предложена сотрудником фирмы “IBM” Эдгаром Коддом в 1970 году.

Таблица базы данных соответствует одной сущности предметной области и состоит из фиксированного числа полей, собранных в записи, каждая запись соответствует экземпляру сущности.

Большинство современных СУБД являются реляционными.

Все реляционные СУБД поддерживают язык SQL.

**SQL** (*structured query language* - «язык структурированных запросов») - формальный непроцедурный язык программирования, применяемый для создания, модификации и управления данными в произвольной реляционной базе данных, управляемой соответствующей системой управления базами данных (СУБД).

***Постреляционная модель*** ‑ это реляционная модель, допускающая многозначные поля (атрибуты), т.е. само поле может быть таблицей.

***Многомерная модель*** ‑ узкоспециализированная модель, предназначен­ная для хранения данных в виде многомерного массива (гиперкуба), исполь­зуемых OLAP (см.).

***Декомпозиция*** — разделение целого на части. Также декомпозиция — это научный [метод](https://ru.wikipedia.org/wiki/%D0%9C%D0%B5%D1%82%D0%BE%D0%B4), использующий структуру [задачи](https://ru.wikipedia.org/wiki/%D0%97%D0%B0%D0%B4%D0%B0%D1%87%D0%B0) и позволяющий заменить решение одной большой задачи решением серии меньших задач, пусть и взаимосвязанных, но более простых.

Декомпозиция, как процесс расчленения, позволяет рассматривать любую исследуемую систему как сложную, состоящую из отдельных взаимосвязанных подсистем, которые, в свою очередь, также могут быть расчленены на части. В качестве систем могут выступать не только материальные объекты, но и процессы, явления и понятия.

***Агрегируемость*** ***данных*** наличие различных уровней обобщения информации, подразумевает привязку данных к какому-либо параметру (например, ко времени, географическому положению, структурным единицам и т.д.).

***Срез*** ‑ подмножество гиперкуба (см.), полученное в результате фиксации одного или нескольких измерений.

***Вращение*** ‑ вращение гиперкуба (см.) при этом местоположение отдельных осей меняются местами.

***Агрегация/детализация*** ‑ переход к более общему/детальному представлению информации по параметрам агрегируемости (см.).

***Распределенная база данных*** - это совокупность логически взаимосвязанных баз данных, распределенных в компьютерной сети. Распределенная база данных может объединять базы данных, поддерживающие любые модели (иерархические, сетевые, реляционные и объектно-ориентированные базы данных) в рамках единой глобальной схемы. Подобная конфигурация должна обеспечивать для всех приложений прозрачный доступ к любым данным независимо от их местоположения и формата.

***Репликация*** (replication) — механизм [синхронизации](https://ru.wikipedia.org/wiki/%D0%A1%D0%B8%D0%BD%D1%85%D1%80%D0%BE%D0%BD%D0%B8%D0%B7%D0%B0%D1%86%D0%B8%D1%8F_%28%D0%B8%D0%BD%D1%84%D0%BE%D1%80%D0%BC%D0%B0%D1%82%D0%B8%D0%BA%D0%B0%29) содержимого нескольких копий БД. Репликация — это процесс, под которым понимается копирование данных из одной БД в другую и поддержание копий в актуальном и согласованном состоянии.

При репликации изменения, сделанные в одной копии БД, распространяются в другие копии.

***Тиражирование данных*** предполагает создание копии (репликаций) фрагментов базы в узлах сети. Синхронизация копий и базы осуществляется специальной программой-репликатором, при этом передаются только изменения или транзакции.

***Монопольный доступ*** используется обычно при массовых операциях с базой (реорганизация или восстановление базы и др.) и блокирует доступ ко всей БД другим пользователям.

***Коллективный доступ*** дает возможность одновременной работы с БД многим пользователям. Для корректной работы при одновременной корректировке отдельных записей возможны блокировки записей.

***Администратор базы данных*** ‑ лицо или группа лиц, ответствен­ная за проектирование и эффективное использование БД.

*Основные функции администратора базы данных:*

1. Участие в разработке проекта по созданию БД.
2. Обеспечение согласованности БД (копирование, репликация, восстановление).
3. Обучение пользователей работе с БД (вход в БД, ввод паролей, работа в нормальной и экстремальной ситуациях и др.).
4. Отслеживание трафика загрузки БД, сбоев, характеристик (время доступа, общее время обработки запросов и т.д.).
5. Реорганизация БД с целью улучшения характеристик базы.
6. Реализация многопользовательского режима работы с БД (защита от несанкционированного доступа, шифрование данных, разделение доступа и др.).

***Администратор СУБД (DBA) -*** лицо или группа лиц, ответствен­ная за развертывание, бесперебойную эксплуатацию и эффективное использование СУБД.

***Концептуальная модель БД (****концептуальная схема*) - это совокупность сущностей и их взаимосвязей вне зависимости от конкретной СУБД.

***Логическая модель БД*** (*логическая схема*) - это концептуальная модель в терминах конкретной СУБД.

 ***Локальная архитектура СУБД*** ‑ приложение и база находятся на одном компьютере.

***Файл‑серверная архитектура СУБД*** – приложение находятся на одном компьютере, а база данных - на другом. Приложение обращается за информацией к файлу‑серверу, который возвращает приложению файлы, содержащие необходимые данные. В настоящее время не применяется.

***Клиент‑серверная архитектура СУБД*** ‑ от­личается от предыдущей модели тем, что запросы от приложения выполняются на сер­вере БД, а приложению передаются выбранные данные.

***Трехзвенная архитектура СУБД*** ‑ к серверам баз добавляются сер­веры приложений (Application Server), на которых выполняются приложения клиентов.Клиент формирует исходную информацию для расчета, посылает запрос на выполнение расчета на сервер приложения, где он и выполняется. При необходимости сервер приложения формирует запрос к серверу БД.

***Тонкий клиент*** (“Клиент‑Интернет”) - клиентская часть, реализуемая на браузере.

***Толстый клиент***- клиентская часть, реализуемая в виде приложения.

***Интерфейс ODBC*** (Open Database Connectivity) является посредником между приложением и СУБД; обеспечивает доступ из приложения к базам с различными СУБД. В состав ODBC входят драйверы (для каждой СУБД один драйвер, который преоб­разует форматы данных и команды приложения в форматы и команды СУБД и обратно) и диспетчер драйверов, который подключает нужный драйвер. Разработан и поддерживается компанией Microsoft.

***Конфиденциальная информация*** (sensitive information) — информация, которая требует защиты.

***Доступ к информации*** (access to information) — ознакомление с информацией, ее обработка (в частности, копирование), модификация, уничтожение.

***Субъект доступа*** (access subject) — лицо или процесс, действия которого регламентируются правилами разграничения доступа.

***Объект доступа*** (access object) — единица информации автоматизированной системы, доступ к которой регламентируется правилами разграничения доступа. Объектами доступа (контроля) в СУБД является практически все, что содержит конечную информацию: таблицы (базовые или виртуальные), представления, а также более мелкие элементы данных: столбцы и строки таблиц и даже поля строк (значения). Таблицы базы данных и представления имеют владельца или создателя.

***Правила разграничения доступа*** (security policy) — совокупность правил, регламентирующих права субъектов доступа к объектам доступа.

***Санкционированный доступ*** (authorized access to information) — доступ к информации, который не нарушает правил разграничения доступа.

***Несанкционированный доступ*** (unauthorized access to information) — доступ к информации, который нарушает правила разграничения доступа с использованием штатных средств, предоставляемых средствами вычислительной техники или автоматизированными системами.

***Идентификатор доступа*** (access identifier) — уникальный признак объекта или субъекта доступа.

***Идентификация*** (identification) — присвоение объектам и субъектам доступа идентификатора и (или) сравнение предъявляемого идентификатора с перечнем присвоенных идентификаторов.

***Пароль*** (password) — идентификатор субъекта, который является его секретом.

***Аутентификация*** (authentification) — проверка принадлежности субъекту доступа предъявленного им идентификатора, подтверждение подлинности.

***Дискреционное управление доступам*** (discretionary access control) — разграничение доступа между поименованными субъектами и поименованными объектами. Субъект с определенным правом доступа может передать это право любому другому субъекту.