

# Данные и информация

среда, 16 января 2019 г.

9:43

Информация – осмысленные данные.

Информация - (лат.) разъяснение, изложение. Термин "Достоверность" не характерно к информации.

"Достоверность" характерно для сигнала.

Шеннон – **информация** - снятая неопределенность в битах передаваемая по соответствующим каналам связи. ичеком носителя

Информация – концептуально связанные между собой данные, изменяющие представление о явлениях или объектах окружающего мира.

Данные – любые факты в окружающем мире, зафиксированные на материальном или физическом носителе.

Данные, необходимые для принятия решения - информация.

Данные, не использованные при принятии решений - сведения.

Сообщение – форма существования и передачи информации.

Сообщение может быть:

- Документальным (зафиксированное на мат носителе)
- Недокументальным (звуки, визуализация, жесты, мимика, обоняние, осязание, тактильно)

Знание – высшая форма информации, полученная в результате сочетания информации и опыта.

Свойства информации:

1. Атрибутивные
  - a. Неотрывность от физ или мат носителя
  - b. Языковая природа
  - c. Дискретность (прерывистость)
  - d. Непрерывность (связность)
2. Прагматические
  - a. Наличие смысла
  - b. Новизна
  - c. Ценность (полезность)
  - d. Кумулятивность (накапливаемость)
3. Динамические
  - a. Свойство роста
  - b. Свойство повторяемости
  - c. Свойство "старения" (не потеряла актуальность, но перестала активно использоваться)
  - d. Свойство рассеивания (Закон Бредфорда, не более, чем 1/3 информации)
  - e. Многократность использования

## Шеннон. Количественная теория информации.

четверг, 17 января 2019 г.

9:53

В 1941 поступил на работу BELL где занимался проблемами передачи информации по телефонному и телеграфному каналам.

В 1948 года выпустил книгу "Математическая теория передачи информации", в которой определяет понятие *информация*, вводит единицу его измерения - *бит* и разрабатывает механизм расчета пропускной способности канала связи.

Семантическая теория информации. (Бриллюэн)

В основе теории шеннона лежит используемая в физике формула расчета энтропии.

Энтропия в лингвистике степень неупорядоченности чередования символов. Текст с максимальной энтропией - текст с равновероятным распределением всех букв алфавита.

Информационные коммуникации – пути, обеспечивающие передачу сообщений от источника к потребителю;

- В зависимости от канала передачи выделяют:
  - Неформальные (напрямую, без посредников) скайп, рация, переписка в соцсетях и тд
  - Формальные (через специально организованного сотрудника) радио, тв, новости в интернет, блог

Виды информации

- Элементарная информация – присуща объектам неживой природы и изучается естественными науками.
- Биологическая информация – присуща объектам живой природы и изучается биологическими, медицинскими и с/х науками.
- Социальная информация – присуща человеческому обществу.
  - Обыденная
  - Экономическая
  - Эстетическая
  - Научно-техническая
  - тд

(политика, общественное знание, история.....)

Предметная область представляет собой часть реального мира, на основе которой строится структурированная модель.

## Представление информации.

среда, 23 января 2019 г.

9:47

Виды представления информации:

- В виде **образов** -> проходит через органы чувств при контакте с природой, объектами внешнего мира (танец, живопись, природный пейзаж, пение...) / неоднозначность смысла для каждого;
- В виде **знаков** -> материально или чувственно воспринимаемый предмет, явление или действие, служащие для обозначения другого предмета, свойства или отношения (буква, цифра, жест, пиктограмма);
- В виде **символов** (буква, цифра, математический, дорожный, товарный знак...);
- В виде **графического изображения** (геральдический знак, религиозный символ, пиктограмма);

Контекст- знаковая информация, как и образная, может иметь несколько значений/ в отличии от образной информации значение знаковой зависит от окружающего контекста.

В основе знаковой системы лежит понятие языка - определенная система символьного представления информации.

Основные понятия языка:

- Алфавит - множество используемых символов;
- Синтаксис - правила записи языковых конструкций;
- Семантика - смысловая сторона языковых конструкций;
- Прагматика - практические последствия применения текста на данном языке.

Язык как знаковая система:

- Естественные языки (разговорный, литературный, деловой) - отражает исторические и культурные традиции общества, психологические и образовательные особенности личности/ всегда эмоционален;
- Искусственный язык использует формализованные знаковые системы (нотный, математический язык, дорожные знаки, языки программирования), выполняющие задачу замещения многословных и неоднозначных выражений естественного языка.

Типы данных:

- Текстовый (символьный);
- Графический (изображение, видео)/ статичный, динамичный;
- Аудиальный (звуки, крики, музыка);
- Мультимедиа (2 канала восприятия) звук-тактильное, аудио-видео, видео-тактильное;
- Структурированные (организованы по определенной структуре) предназначены для поиска и анализа информации / список частный случай таблицы, графики, любые формы диаграмм/ признак: не предназначены для сплошного чтения. (словарь, энциклопедия - справочные издания).

## Структурированные данные

четверг, 24 января 2019 г.

9:49

База данных (БД) — это организованная структура, предназначенная для хранения, изменения и обработки взаимосвязанной информации, преимущественно больших объемов.

**Таблицы** - совокупность структурированных данных, предназначенных для представления, поиска и обработки./Форма организации базы данных.

Единица таблицы:

Поле

Виды:

- Вертикальные
- Горизонтальные

Поле содержит одну характеристику объекта и интерпретируется единым образом

Структура таблицы:

- **Столбец** вертикально, обозначается буквами латинского языка.
- **Строка** горизонтально, обозначаются арабскими цифрами.
- **Ячейка** обозначается номер столбца и номером строки.

По содержанию:

- Поле:

- Основной элемент таблицы
- Задаёт структуру таблицы;
- Содержит одну характеристику таблицы;
- Интерпретируются единым образом (один тип данных);
- Единица измерения указывается в названии поля (не содержится в ячейке).

Совокупность полей создаёт "шапку" таблицы

- Запись:
  - Формирует область значений;
  - Содержит сведения об одном экземпляре объекта.
- Ячейка:
  - Содержит одну характеристику одного объекта.

Фамилия	Группа	Пропуски в днях
Васечкин И	2	1
Иванов А	1	7
Катина М	2	31
Петров А	1	5
Сидорова Ю	1	3

Типы данных в таблице:

- Символьные/текстовые:
  - Носит относительный характер;
  - По левому краю;
  - Не участвует в вычислениях.
- Числовые:
  - Предназначены для ввода чисел;
  - По правому краю.
    - Абсолютное число
    - Вычисляемое поле
      - ☐ Формируется вводом формулы или формулы
      - ☐ Результат выводится в ячейке по правому краю
  - Дата
    - ☐ Особый тип данных предназначен для обработки дан
    - ☐ Выражается в виде порядкового номера дня от 01.01.1900.

Вычисления в таблице:

- Данные, используемые для вычислений:
  - Числовые данные;
  - Адреса ячеек;
  - Текстовые данные.
- Способы вычислений:
  - Формулы;
  - Функции.
- Вычисляемое поле начинается с =

- Курсор рабочей области ставится в результирующей ячейке (в которой записывается формула или функция)
- Исходная ячейка - которая содержит данные для формулы или функции адрес которой приводится в вычисляемом поле

Формула:

- Совокупность математических операторов, чисел, ссылок и функций
- Прямой порядок действий
- Состоит из:
  - Операндов (данных)
  - Операторов (действий)
    - Алгебраические
      - Арифметические + - \* / ^
      - Сравнения = > < >= <= <>
    - Логические not and or
- Формула начинается со знака =

Функция:

- Зависимость одной переменной у от одной или нескольких переменных
- Самостоятельная или входит в состав формулы
- Состоит из:
  - Имя - уникальное название функции (sum if...)
  - Аргумент - каждая независимая величина x
- Формат записи функции:  
= имя функции(аргумент1;аргумент2)
- Обозначение поля, записи или ячейки называется адресом
- Разделение ;
- Диапазон :

Функция вложением

Нельзя использовать функцию для действия

- Вычитания
- Деления

В качестве аргумента другая функция - вложенная

```
=sum(a1:b2)product(sum(d2;c4))
=product(c1:c3)sum(del(sum(a1;c4))
=Sum(a3;b4)del100produkt20
```

## Типы и виды моделей данных

среда, 30 января 2019 г.

18:30

Модель данных - совокупность данных и операций их обработки./ядро любой БД, которое представляет собой множество данных ограничений целостности и операций манипулирования данными.

Назначение:

Представление объектов предметной области и взаимосвязей между ними.

Основные типы:

- Иерархическая - совокупность элементов, связанных между собой по определенным правилам. (древовидная структура, формирующая ориентированный граф - только один путь от корневого узла).
  - Уровень, узел - совокупность атрибутов данных, описывающих некоторый объект;
  - Связь (дуги).
- Сетевая - совокупность элементов, при котором каждый элемент может быть связан с любым другим элементом. (Наборы соединяются с помощью записей связей, образуя цепочки итд).
  - Уровень, узел - совокупность атрибутов данных описывающих некоторый объект;
  - Связь (ребра).
- Реляционная - простые двумерные таблицы, соединенные связями.
- + таблица

Виды:

- Инфологическая модель (выполненное с использованием естественного языка, мат формул, таблиц, графиков и других средств)
- Даталогическая модель (отображение логических связей между элементами данных независимо от содержания и среды хранения)ф]
- Физическая модель (для отыскивания данных на внешних запоминающих устройствах)

## Гипертекстовые системы

вторник, 5 февраля 2019 г.

12:20

Гипертекст - форма нелинейной организации текстового материала как система явно указанных связей между фактами.

- Нелинейность
- Явные связи
- Факты (нет трактовок)

Прим. Использование в качестве факта различных форм подачи информации приводит к формированию гиперсистемы, гипермедиа и тд.

История гипертекста:

- Ванневер Буш "как мы возможно думаем" (1945) - машина тетех
- Теодор Нельсон "информационные системы будущего" (1965)
- Даглас Энгельбард (программа с компьютерной мышью, 1968)

Структура гиперсистемы:

- **Тезис - единица гипертекста** (каждый факт в гиперсистеме)
  - **Номер (или заголовок)**, которые используются при организации гиперссылки на этот тезис;
  - **Ссылка на другой тезис (прямая или ассоциативная);**
  - **Ссылка на первоисточник**, откуда извлечен данный факт.

Достоинства гиперсистем:

- Возможность комбинировать структуру линейного документа
- Возможность применения при автоматизированном обучении
- Возможность навигации в больших массивах данных
- Возможность объединения различных фактов (универсальность)

Недостатки гиперсистем:

- Теряется привычный навык работы с линейными текстами
- Возникает интеллектуальная перегрузка при отслеживании многочисленных связей и их трактовке
- Дезорганизация пользователя в сети

## Реляционная модель базы данных

вторник, 19 февраля 2019 г.

11:40

Эдгар Кодд предложил новую модель, в которой соединил удобство табличной формы и возможность использования формального аппарата алгебры отношений (логические выражения) и реляционных исчислений.

Данные вносятся вручную однократно.

Идеи:

- Автоматизированные системы (не нужно использовать неавтоматизированные)

Свойства двумерной таблицы в реляционной базе данных:

- 1 элемент данных - 1 элемент таблицы - ячейка;
- Каждое поле таблицы однородно и содержит один тип данных;
- Одинаковые строки в таблице отсутствуют (данные не повторяются);
- Каждое поле в таблице обладает уникальным именем.

Между двумерными таблицами устанавливаются связи -> связи устанавливаются при помощи ключа.

Виды ключей:

- Первичный ключ - возникает в таблице от которой далее пойдет связь; 1 атрибут - ключевой (идентифицирует как уникальный)
- Вторичный ключ - связь от первичного ко вторичному (внешний - отражение области первичного внутри другой таблицы).

Каждая таблица посвящена только одной сущности.

Поле или сочетание полей однозначно идентифицирующее каждую запись таблицы - каждую сущность.

Поле - Атрибуты посвящены описанию

Остальные атрибуты - описательные

Правило описание сущности: каждый описательный атрибут должен зависеть от ключевого.

Ключевой может быть в несколько полей (редко);

Совокупность нескольких полей - составной;

Чтобы связать 2 реляционные таблицы необходимо ключ первой таблицы отразить атрибутом во вторую таблицу.

Типы данных первичного и вторичного ключей совпадают по первичному ключу.

Связи между таблицами и структура таблицы, т.е. выделение описательных и ключевых атрибутов организуются различными способами, при этом связь должна быть рациональной и минимизировать дублирование данных и кол-во процедур.

Кодд ввел понятие нормализации отношений и выделил 3 базовые нормальные формы:

- Первая - образуется в том случае, если все атрибуты таблицы простые и неделимые; Стремление к первой нормальной форме может привести к увеличению кол-ва полей в таблице, что является неоправданным для некоторых описательных атрибутов. (Кодд считал несовершенной);
- Вторая - образуется в случае, когда все описательные атрибуты простые и находятся в полной функциональной зависимости от ключевого атрибута;
- Третья - основано на понятии нетранзитивная зависимость - не должно быть. Для ликвидации транзитивной сущности производится расщепление сущности.

#### Типы Связи

— «один-ко-многим», имеет место, когда одной записи родительской таблицы может соответствовать несколько записей в дочерней таблице.

Связь "один-ко-многим" является самой распространенной для реляционных баз данных.

— «один-к-одному», имеет место, когда одной записи в родительской таблице соответствует одна запись в дочерней таблице.

— «многие-ко-многим» имеет место, когда:

а) записи в родительской таблице может соответствовать больше одной записи в дочерней таблице;

б) записи в дочерней таблице может соответствовать больше одной записи в родительской таблице.

Нормализация отношений - формальный аппарат ограничений на формирование отношений (таблиц), который позволяет устранить дублирование, обеспечивает непротиворечивость данных, уменьшает трудозатраты на создание и введение баз данных.

Функциональная зависимость - зависимость, при которой определенному значению ключевого атрибута соответствуют значения описательного атрибута.

Транзитивная зависимость - возникновение описательного атрибута, зависящего от другого описательного атрибута, но не от ключа.



Объекты Access:

- Таблица
- Запрос
- Форма
- Отчёт

Режимы :

Конструктор

Просмотр

Макет

SQL

Индекс – это внутренняя таблица, имеющая 2 столбца: значение индексного поля (или выражения из нескольких полей) и местоположение (адрес) записи с этим значением индекса.

Индексируемое - Поле используемое при поиске (смысловые поля)

Обязательное - обязательно заполненное но не обязательно индексируемое

**Правила подстановки:**

- **Типы данных одинаковы/указать источник**
- **Ширина столбца**
- **Присоединение идёт по ключевому столбцу(полю)**
- **Выбор из списка/вручную**

**Типы данных полей подстановки:**

- **Числовой**
- **Текстовый**

Ключевое поле — это одно или несколько полей, комбинация значений которых однозначно определяет каждую запись в таблице. Если для таблицы определены ключевые поля, то Microsoft Access предотвращает дублирование или ввод пустых значений в ключевое поле. Ключевые поля используются для быстрого поиска и связи данных из разных таблиц при помощи запросов, форм и отчетов.

## Организация базы данных

вторник, 5 марта 2019 г.

11:40

Возможна организация тремя способами(внедрение объектов):

1. Создание таблицы в режиме конструктора;
2. Импорт объекта из внешнего источника;
3. Установление связи с внешним объектом при помощи ярлыка.

Схемы данных - представляют собой отражение объектов и установление связей между ними.  
(Только таблицы и запросы)

Схемы данных нужны:

- Наглядность
- Установление связей

Связь между первичным ключом и его отражением в другой таблице (внешней/вторичный), при этом типы данных должны совпадать по типу данных первичного ключа.

Для контроля работы ставится обеспечение целостности данных. Искл.: обеспечение целостность невозможно поставить для внешнего объекта (ярлыка).

Типы связей:

- Один к одному - каждому экземпляру объекта А соответствует один экземпляр Б и наоборот (<->)
- Один к многим - каждому экземпляру А соответствует 0,1,2 и более экземпляра объекта Б, но не наоборот (<->>)
- Многие ко многим - каждому экземпляру объекта А соответствует 0,1,2 и более экземпляра объекта Б и наоборот (<<->>)

При изменении структуры сначала надо разрывать связи с этим объектом, потом поменять и восстановить.

## Организация запросов

среда, 27 марта 2019 г.

10:08

Запрос предназначен для выборки записей из источника. (1 шаг-выбрать источник)

Источником для создания запроса может быть:

- Таблица
- Другой запрос

Создание запроса в access

- Запрос QBE (создаётся вручную пользователем)
- SQL

Типы запросов:

- Запрос на выборку !
- Запрос на добавление;
- Запрос на обновление (актуализация данных в уже имеющейся строке);
- Запрос на удаление;
- Запрос на создание таблицы;
- Перекрёстный запрос;
- Запрос на объединение.

QBE шаблон в режиме конструктора

1 панель - выбор источника; если в блоке источник выбрано две таблицы, между которыми отсутствует связь, система делает выборку по каждой таблице отдельно, даже если в схеме данных эти таблицы связаны между собой через другую-> для организации корректной выборки: две независимые таблицы могут быть связаны между собой; в бланк запросы добавлены таблицы, связывающие две необходимые.

2 панель - работа с блоком; мин. 4 поля: поле, имя таблицы, условие отбора; или.

Поле таблицы может быть только имя собственное или полное имя, которые включает в себя имя таблицы.имя поля. Имя собственное поля должно быть уникально в пределах Таблицы, полное имя должно быть уникально в базе данных.

В шаблон поиска вносятся поля которые должны быть видны и по которым ищем.

Условия отбора задаются с помощью операторов:

- Логические (соединение двух и более условий поиска, и или не)
- Математические (используются внутри условия)