



БАЗЫ ДАННЫХ

СИСТЕМЫ УПРАВЛЕНИЯ ИНФОРМАЦИЕЙ

Бычкова
Наталья Александровна

Директор по развитию GXGroup

к.т.н., доцент,
Кафедра Управления и
информатики в технических
системах ФГБОУ ВО МГТУ СТАНКИН

курс лекций
2020



*«Плохие программисты беспокоятся о коде.
Хорошие программисты беспокоятся
о структурах данных и их отношениях...», -*

Линус Торвальдс, создатель Linux

*«Bad programmers worry about the code.
Good programmers worry about data structures and
their relationships.»*



Содержание дисциплины:

1 семестр

- Введение:
 - Основные термины и понятия
 - Эволюция развития
 - Технологии и этапы разработки БД
- Классификации:
 - Объекты разработки (АИС)
 - Технологии разработки БД
 - Инструменты разработки (СУБД)
- Информационное моделирование:
 - Реляционная модель данных
 - Графовые модели данных
 - Постреляционные модели данных
- Банки данных
- Распределенные БД
- Клиент-серверные архитектуры БД

2 семестр

- Разработка БД:
 - Концептуальное проектирование
 - Инфологическое проектирование
 - Даталогическое проектирование
- Семантическое моделирование БД:
 - SADT-IDEF-ICAM, DFD, UML
- Стандарты разработки БД:
 - Международные, Отраслевые, Корпоративные
- SQL-стандарт
- Визуальная разработка БД: CASE
- Системы транзакционной и аналитической обработки данных:
 - On-Line Transaction Processing (OLTP)
 - Online Analytical Processing (OLAP)
- NoSQL БД:
 - Основные модельные представления
 - Сравнение с РМД

Структура курса:

1 семестр	
Занятия	Контроль
<ul style="list-style-type: none">• Лекции: 10	<ul style="list-style-type: none">• Контроль посещения (1 и 2 модуль)
<ul style="list-style-type: none">• Семинары: 4	<ul style="list-style-type: none">• Контрольная работа (2 модуль)
<ul style="list-style-type: none">• Лабораторные работы: 3	<ul style="list-style-type: none">• Тестирование знаний (2 модуль)
	<ul style="list-style-type: none">• Зачет



2 семестр	
Занятия	Контроль
<ul style="list-style-type: none">• Лекции: 10	<ul style="list-style-type: none">• Контроль посещения эссе
<ul style="list-style-type: none">• Семинары: 4	<ul style="list-style-type: none">• Контрольная работа (2 модуль)
<ul style="list-style-type: none">• Лабораторные работы: 3	<ul style="list-style-type: none">• Тестирование знаний (2 модуль)
	<ul style="list-style-type: none">• Экзамен



Литература:

Основная:

- Гущин, А.Н. Базы данных: учебно-методическое пособие / А.Н. Гущин. - 2-е изд., испр. Ии доп. - М. ; Берлин : Директ-Медиа, 2015. - 311 с: ил. - Библиогр.: с. 226-228. - ISBN 978-5-4475-3838-5,
- Гущин, А.Н. Базы данных: учебник / А.Н. Гущин. - М. : Директ-Медиа, 2014. - 266 с. :ил.,табл., схем. - ISBN 978-5-4458-5147-9,
- Карпова, Т.С. Базы данных: модели, разработка, реализация: учебное пособие /Т.С. Карпова. - 2-е изд., исправ. - М. : Национальный Открытый Университет«ИНТУИТ», 2016.-241 с: ил.,
- Клецова, Т.В. Базы данных / Т.В. Клецова, В.О. Наталья, И.В. Прохоров. - М.: МИФИ, 2008. - 132 с. - ISBN 978-5-7262-0899-2,
- Абросимова, М.А. Базы данных: проектирование и создание программного приложения в СУБД MS Access: практикум / М.А. Абросимова; Минобрнауки РФ,
- Быкова, В.В. Искусство создания базы данных в Microsoft Office Access 2007: учебное пособие / В.В.Быкова. - Красноярск: Сибирский федеральный университет, 2011. - 260 с: табл. - ISBN 978-5-7638-2355-4

Интернет-ресурсы:

- Электронно-библиотечная система «Университетская библиотека онлайн» <http://biblioclub.ru/index.php?page=main> ub red
- Электронная библиотека МГТУ «СТАНКИН» в ЭБС «Университетская библиотека онлайн» - http://biblioclub.ru/index.php?page=izdat_coll nd&id=I 0957
- Электронно-библиотечная система издательства «ЮРАЙТ» - <http://vAvw.biblio-online.ru>
- Портал [SQL.ru](http://sql.ru) - <http://sql.ru>
- .Култыгин, О. П. Администрирование баз данных. СУБД MS SQL Server [Электронный ресурс]: учеб. пособие / О.П. Култыгин. - М.: МФПА, 2012. - 232 с. <http://znanium.com/bookread2.php?book=451114>
- ЭБС «Университетская библиотека онлайн» издательства «Директ-Медиа» // Режим доступа URL: www.biblioclub.ru.
- Электронные издания научно-технической библиотеки, размещенные в разделе университета в ЭБС «Университетская библиотека онлайн».
- Электронная образовательная среда в сети Интернет // Режим доступа URL: <http://edu.stankin.ru/>.
- Электронная библиотека научных публикаций «Российский индекс научного цитирования» // Режим доступа URL: elibrary.ru/.



ЛЕКЦИЯ 1: УПРАВЛЕНИЕ ИНФОРМАЦИЕЙ

Информационный взрыв

Информационный взрыв — постоянное увеличение скорости и объёмов публикаций (объёма информации) в масштабах планеты. Лавинообразное нарастание массы разнообразной информации в современном обществе получило название «информационного взрыва» в 1975 году (Урсул А. Д.)

- Цивилизационные тренды развития информационного общества характеризуются тем, что в 2002 г. человечеством было произведено информации $18 * 10^{18}$ байт (18 Эксабайт).
- За пять предыдущих лет человечеством было произведено информации больше, чем за всю предшествующую историю.
- Объём информации в мире возрастает ежегодно на 30 %.
- В среднем на человека в год в мире производится $2,5 * 10^8$ байт.

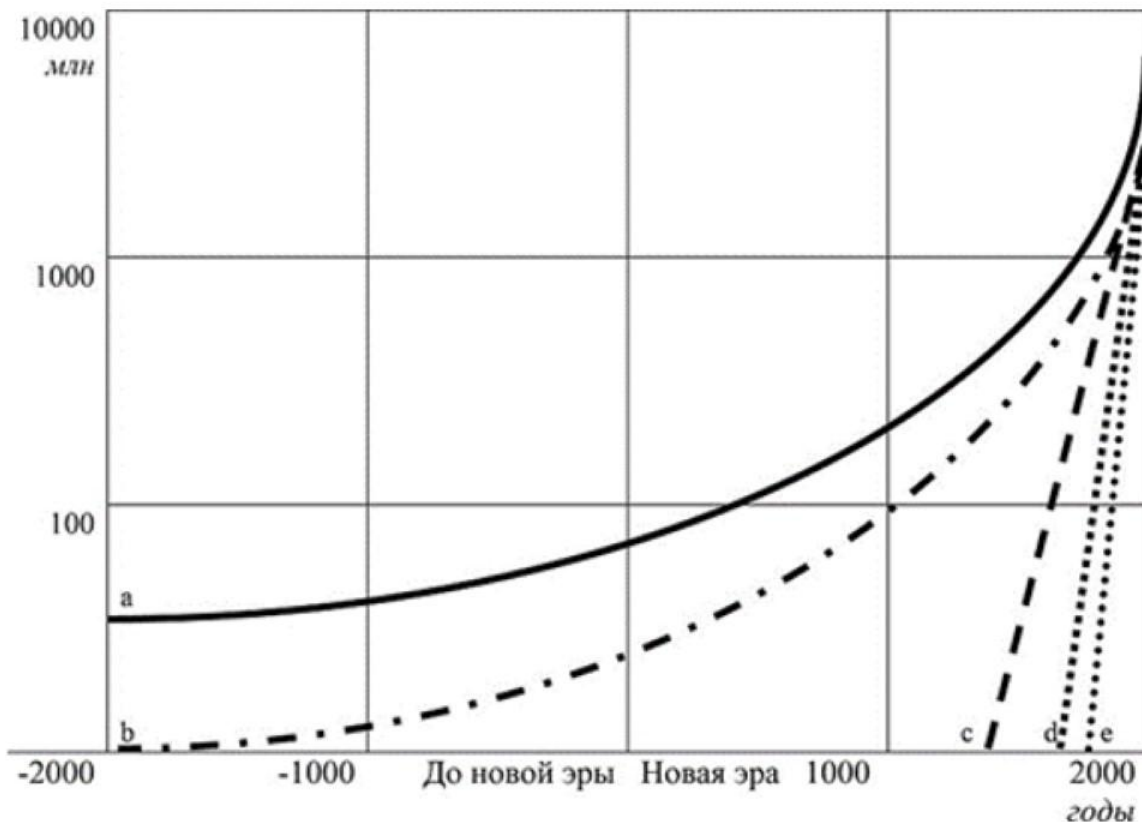
Современная революция в информационных технологиях характеризуется тем, что на 7 млрд человек приходится:

- 6 млрд телефонов (данные компании «Ericsson», 2016),
- 6 млрд телевизионных установок («Guinness Today», 2015),
- 2,8 млрд компьютеров («Gartner», 2015),
- 2,5 млрд интернет-пользователей («Internet World Stats», 2012).



ЛЕКЦИЯ 1: УПРАВЛЕНИЕ ИНФОРМАЦИЕЙ

Динамика информационного обмена



Распространение новых функционально-поведенческих признаков человека в виде навыков информационного взаимодействия.

- a — население Земли 7 млрд;
- b — грамотность;
- c — чтение-печатание — доступность всем грамотным;
- d — получение радио-, телевизионной информации;
- e — информационная связь через телефоны, компьютеры, интернет



ИНФОРМАЦИОННЫЕ РЕВОЛЮЦИИ

1 РЕВОЛЮЦИЯ

письменность. Появилась возможность фиксировать знания на материальном носителе, тем самым отчуждать их от производителя и передавать от поколения к поколению через ее фиксацию в знаках;

2 РЕВОЛЮЦИЯ

- **книгопечатание**, расширила доступ к информации широким слоям населения благодаря тиражированию знаний.

3 РЕВОЛЮЦИЯ

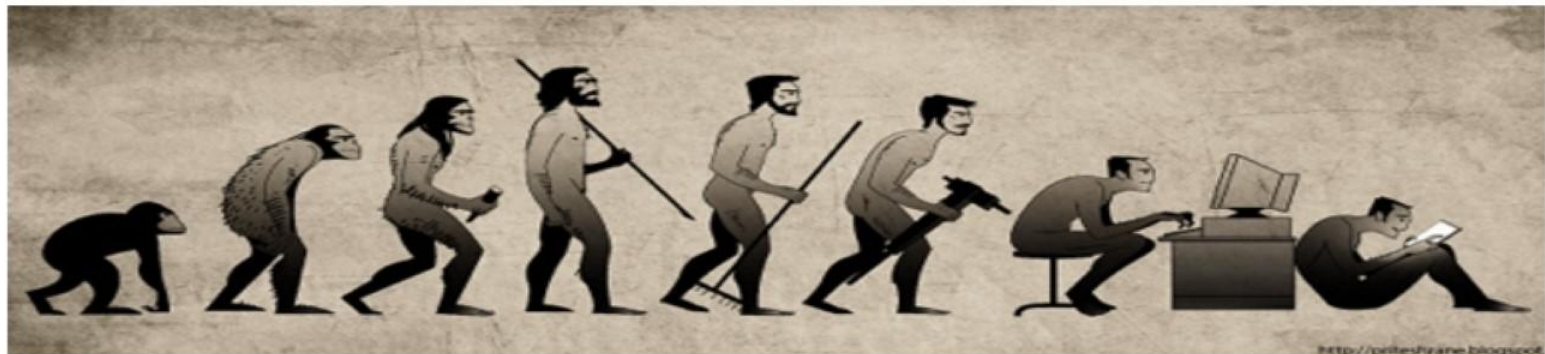
- изобретение **телеграфа, телефона, радио, телевидения**, что позволяло оперативно, в больших объемах передавать и накапливать информацию, передавать звуковые и визуальные образы на большие расстояния - эффект «сжатия пространства»;

4 РЕВОЛЮЦИЯ

- **микроспроцессорная** технология, RISC-архитектура. Персонализация данных.

5 РЕВОЛЮЦИЯ

- формирование и развитие трансграничных глобальных информационно-телекоммуникационных сетей. Интеграция в едином информационном пространстве программно-технических средств, средств связи и телекоммуникаций, информационных запасов или запасов знаний как **единой информационной телекоммуникационной инфраструктуры**.



В сфере электронных коммуникаций возник взрывообразный рост разнообразных частных публикаций во Всемирной сети - **количество блогов удваивается каждые 6 месяцев**



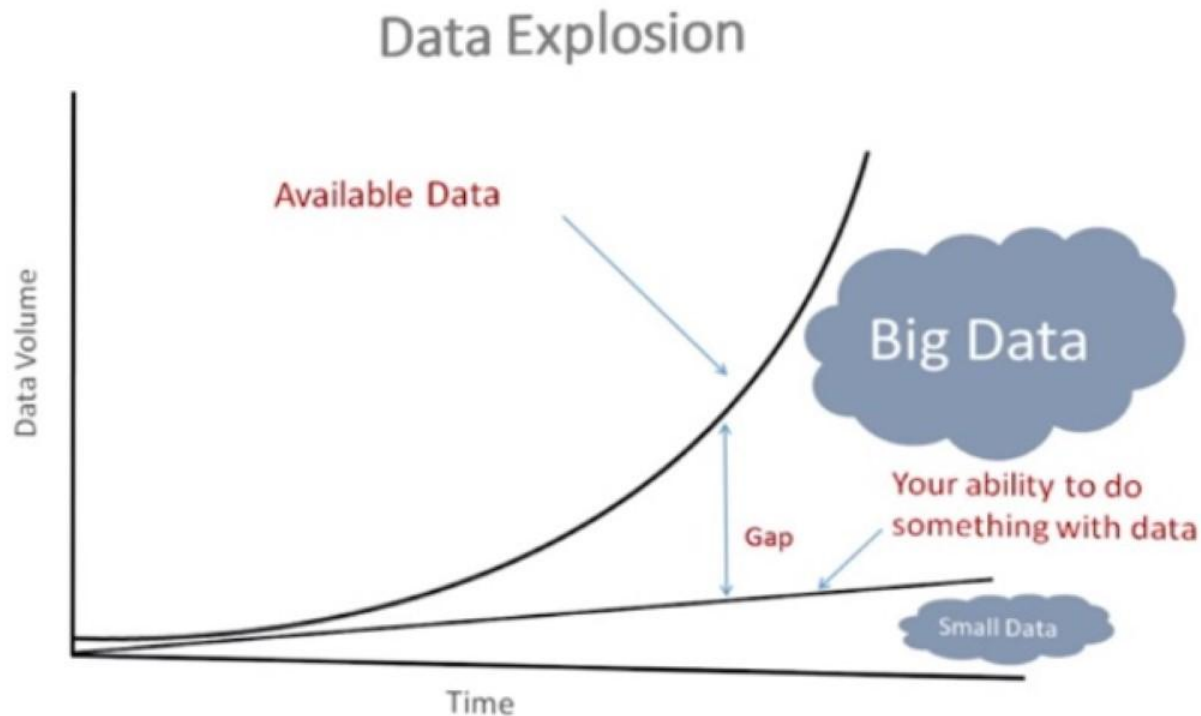
По статистике **объём цифровой информации удваивается каждые 18 месяцев.**

Состав информации:

- до 95 % этот поток состоит из **неструктурированных данных**
- 5 % составляют различные базы данных — тем или иным образом **структурированная информация**

ЛЕКЦИЯ 1: УПРАВЛЕНИЕ ИНФОРМАЦИЕЙ

ПРОБЛЕМЫ: ТЕХНОЛОГИЧЕСКИЙ РАЗРЫВ

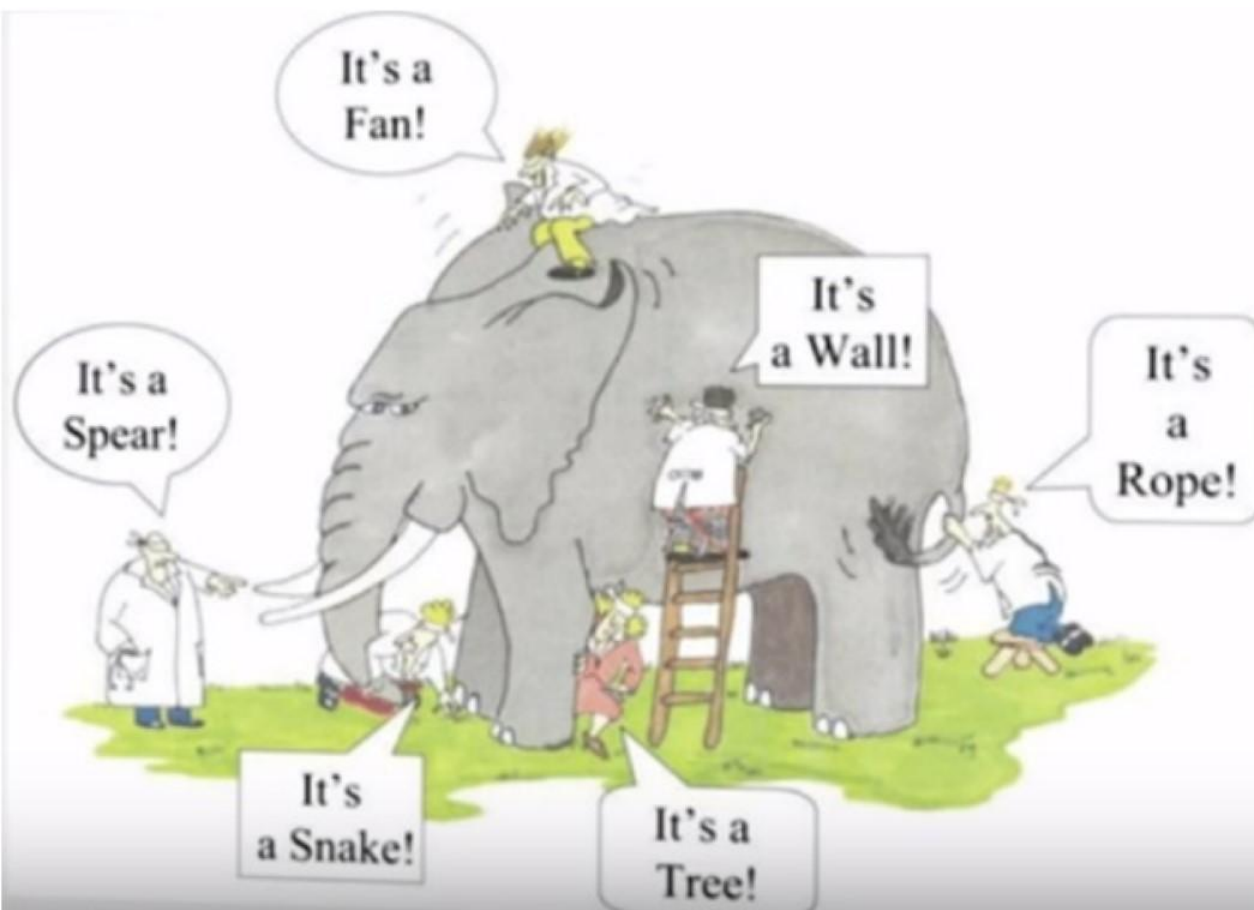


- Резкий рост объема данных
- Разнородность данных
- Разрозненность данных
- Противоречие между объемом человеческих когнитивных возможностей и потребностями в хранении и обработке информации
- Ограниченные возможности традиционных методов обработки структурированных данных
- Задержка во времени при обработке данных



ЛЕКЦИЯ 1: УПРАВЛЕНИЕ ИНФОРМАЦИЕЙ

РЕШЕНИЯ: ИНТЕГРАЦИЯ ИНФОРМАЦИИ



- Интеграция стандартов интеграции ☺
- Структуризация неструктурированных данных ☺
- Единое представление об объекте
 - CALS
- Новые модели представления данных
 - Пост-реляционные МД
 - Объектно-ориентированные МД
 - Кубы данных
 - NoSQL
- Новые технологии обработки данных
- Новые технологии анализа данных



ЛЕКЦИЯ 1: УПРАВЛЕНИЕ ИНФОРМАЦИЕЙ

Предметная область.

Информация, данные, знания

Рискну утверждать, что появление БД стало одним из самых важных достижений в области программного обеспечения.

Сей факт послужил основой для крупномасштабных разработок коммуникационных систем и телефонии, транспорта и логистики, финансового менеджмента и законодательства, систем БД и методов доступа к научной литературе, а также большого количества общегражданских и военных приложений.

Системы БД можно встретить во всех сферах человеческой деятельности: в магазине, в банке, в туристическом агентстве, в библиотеке, в здравоохранении и образовании, наконец, в управлении 😊

Предметная область — это область реального мира, использующая конкретную информационную систему.

Информация - абсолютно любые сведения о каком-либо событии, сущности, процессе или явлении, характерные для рассматриваемой предметной области.



ЛЕКЦИЯ 1: УПРАВЛЕНИЕ ИНФОРМАЦИЕЙ

Предметная область.


Информация, данные, знания

Перед тем, как определить понятие «данные», представим следующую абстрактную ситуацию:

*Имеется некоторая **система**, информация о состоянии которой представляет особый интерес для некоторого **наблюдателя**, способного воспринимать состояния системы и в определенной форме фиксировать их в своей **памяти**.*

*В этом случае говорят, что в **памяти наблюдателя** находятся «данные», описывающие состояние системы.*

В общем случае, в качестве таких вот наблюдателей выступают информационные системы.



Информация - абсолютно любые сведения о каком-либо событии, сущности, процессе или явлении, характерные для рассматриваемой предметной области.

Данные – это информация, зафиксированная в определенной форме и пригодная для последующего хранения, обработки и передачи.



ЛЕКЦИЯ 1: УПРАВЛЕНИЕ ИНФОРМАЦИЕЙ

Предметная область.

Информация, данные, знания

Знания основаны на данных, полученных эмпирическим путем.

*Знания – это хорошо структурированные данные, или данные о данных, или **метаданные**.*

*Знания организуются в **базы знаний**, являющиеся фундаментом всех интеллектуальных систем, предназначенных для решения неформализованных задач, для которых невозможно представить алгоритм их решения заранее.*

Информация – абсолютно любые сведения о каком-либо событии, сущности, процессе или явлении, характерные для рассматриваемой предметной области.

Данные – это информация, зафиксированная в определенной форме и пригодная для последующего хранения, обработки и передачи.

Знания – это закономерности предметной области (принципы, связи, законы), позволяющие ставить и решать задачи в этой области.



ЛЕКЦИЯ 1: УПРАВЛЕНИЕ ИНФОРМАЦИЕЙ ЗНАНИЯ-ИНФОРМАЦИЯ-ДАННЫЕ



*По мотивам модели Акроффа

- **Данные:**
 - это факты, измерения, сигналы, имеющие отношение к определенному событию.
 - это фактическое описание любого явления, которое представляет ценность
 - это форма представления сведений в процессе их хранения и обработки



ЛЕКЦИЯ 1: УПРАВЛЕНИЕ ИНФОРМАЦИЕЙ ЗНАНИЯ-ИНФОРМАЦИЯ-ДАННЫЕ



- **Информация:**
 - это структурированные данные, имеющие смысл и представленные в определенном контексте.
 - при использовании данных в процессе решения конкретных задач появляется информация.
 - это содержание сведений об объектах, явлениях и событиях реального мира.
 - это приращение знаний, которая может быть выведена на основе данных.
($I = -\Delta \mathcal{E}$)
 - это отрицательное приращение энтропии (хаос)





ЛЕКЦИЯ 1: УПРАВЛЕНИЕ ИНФОРМАЦИЕЙ ЗНАНИЯ-ИНФОРМАЦИЯ-ДАННЫЕ

- **Знание:**
 - *это информация, которую индивид или организация могут трансформировать в действия.*
 - Это множество информации, достаточное для принятия управленческого решения.
 - Это целостная систематизированная совокупность научных понятий о закономерностях природы, общества и сознания (мышления), накопленные человечеством в процессе активной образовательной деятельности и направленная на дальнейшее познание и преобразование окружающего мира.
 - Результаты решения задач, истинная, проверенная информация (сведения), обобщенная в виде законов, теорий, совокупностей взглядов и представлений



ЗНАНИЯ – ИНФОРМАЦИЯ – ДАННЫЕ

ЛЕКЦИЯ 1: УПРАВЛЕНИЕ

ИНФОРМАЦИЕЙ

ЦЕННОСТЬ

СИСТЕМНОСТЬ

ГЛОБАЛЬНОСТЬ

ЗНАНИЯ

ИНФОРМАЦИЯ

ДАННЫЕ

СУБЪЕКТИВНОСТЬ

- Система знаний

- Общее

- Информационная система

- Специфическое

- Не системны

- Единичное

*По мотивам модели Акроффа

Универсальным средством представления данных, информации и знаний является естественный язык (ЕЯ).

В ЕЯ данные и их интерпретация осуществляется совместно.

В вычислительной среде данные хранятся на логических носителях, а их интерпретация возлагается на человека.



*Семиотический подход

ЗНАНИЯ – ИНФОРМАЦИЯ – ДАННЫЕ

ЛЕКЦИЯ 1: УПРАВЛЕНИЕ ИНФОРМАЦИЕЙ



**Семиотический подход*





Виктор Беспалов, вице-президент, генеральный директор Siemens PLM Software в России и СНГ

Сегодня под «цифровым производством» понимается, прежде всего, использование технологий цифрового моделирования и проектирования как самих продуктов и изделий, так и производственных процессов на всем протяжении жизненного цикла.

По сути, речь идет о создании **цифровых двойников продукта и процессов его производства**. Изменения в современной промышленности, которые «цифровое производство» подразумевает, будут происходить по следующим ключевым направлениям:

ЛЕКЦИЯ 1: УПРАВЛЕНИЕ ИНФОРМАЦИЕЙ

Цифровое производство

- ❑ **Цифровое моделирование** – развитие получает концепция цифрового двойника, то есть изготовление изделия в виртуальной модели, включающей в себя оборудование, производственный процесс и персонал предприятия.
- ❑ **«Большие данные» (big data)** и бизнес-аналитика, которые возникают в процессе производства.
- ❑ **Автономные роботы**, которые получают большую промышленную функциональность, независимость, гибкость и исполнительность по сравнению с предыдущим поколением.
- ❑ **Горизонтальная и вертикальная интеграция систем** – большая часть из огромного количества используемых в настоящее время информационных систем интегрировано, но необходимо наладить более тесное взаимодействие на различных уровнях внутри предприятия, а также между различными предприятиями.
- ❑ **Промышленный интернет вещей**, когда поступающая с производства информация с большого количества датчиков и оборудования объединяется в единую сеть.

