

Технологии хранения, обработки и анализа данных

Информация как ресурс

Документирование информации – выделение единичной смысловой части информации в общей ее массе и придание ей самостоятельной роли (имя, статус, реквизиты).

Информация

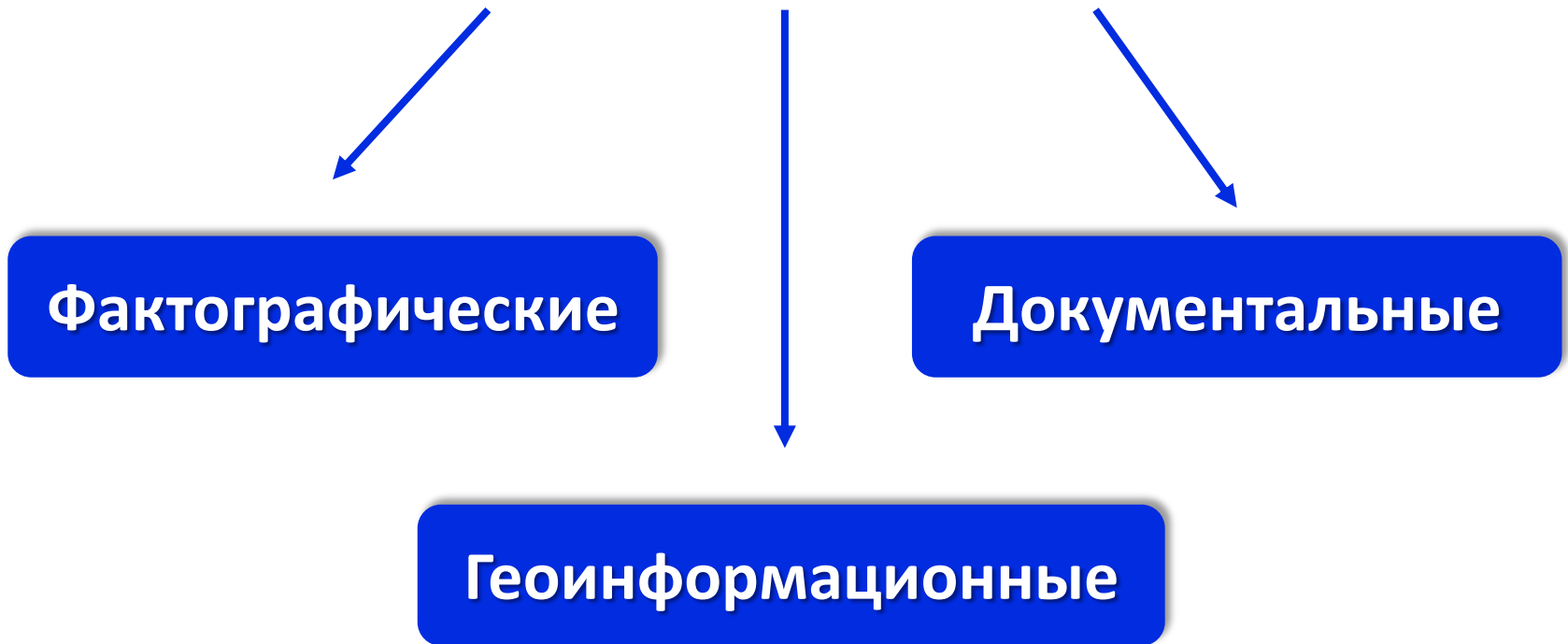
Документирование

Информационные
ресурсы

Информационная система

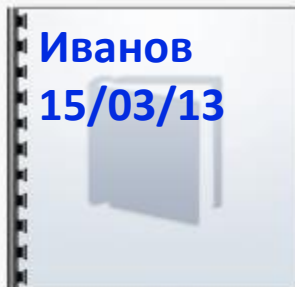
организованная совокупность программно-технических и других вспомогательных средств, технологических процессов и функционально-определенных групп работников, обеспечивающих **сбор, представление и накопление** информационных ресурсов в определенной предметной области, **поиск и выдачу** сведений, необходимых для удовлетворения информационных потребностей установленного контингента пользователей - абонентов системы.

Автоматизированные информационные системы (АИС)

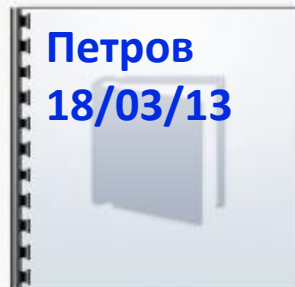


Документальные АИС

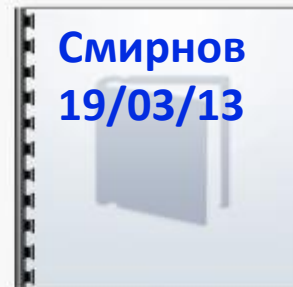
Сведения о статьях,
присланных на конференцию



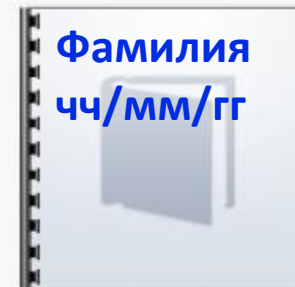
SPE-142337-MS-P



SPE-144338-MS-P



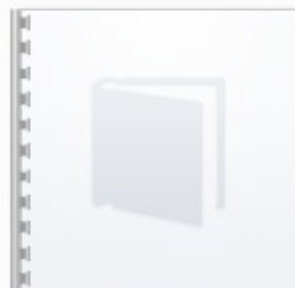
SPE-144365-MS-P



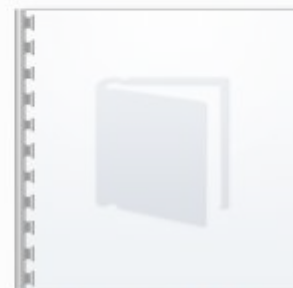
SPE-145634-MS-P



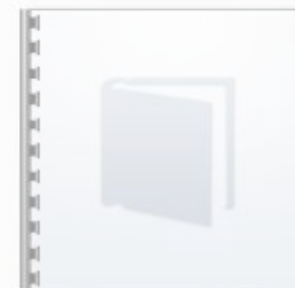
SPE-149832-MS-P



SPE-149833-MS-P



SPE-149924-RU



SPE-150002-MS-P

Геоинформационные АИС



<http://www.tomsk.ru09.ru/map#>

Фактографические АИС

Сведения о товарах в магазине бытовой техники

Тип товара	Холодильник	→	Категория	Многокамерный
Бренд	LG		Управление	Электронное
Модель	GA-B409BLQA		Габариты	190*60*65 см
Наличие	Да			
Цена	20 000			

Фактографические АИС

Сведения о товарах в магазине бытовой техники

Тип товара	Стиральная машина	→	Категория	Узкая
Бренд	Indesit		Загрузка	4 кг
Модель	WIUN 102		Габариты	190*60*65 см
Наличие	Да		Скоро отжима	1000
Цена	10 000			

Файловые системы

Недостатки файловых систем

Избыточность данных

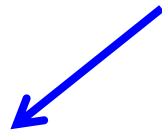
Слабый контроль данных

Недостаточные возможности
управления данными

Большие затраты труда

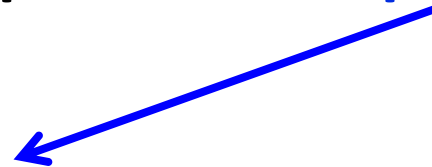
База данных

это совокупность данных,
организованная с определенной целью



Означает, что **данные**:

- сохраняются
- отформатированы
- доступны
- репрезентативны



Означает, что:

- нет информации, не имеющей отношения к задаче;
- есть вся информации, касающаяся задачи

Пример базы данных

Телефонный справочник ТПУ

Телефон:

Фамилия: Имя: Отчество:

Подразделение:

№	ФИО	Подразделение, должность	Расположение	Гор. тел.	Вн. тел.
1.	Гвоздев Николай Иванович	Институт социально-гуманитарных технологий, Заместитель директора по учебной работе	Научно-техническая библиотека, 162	8 (3822) 56-38-09, 56-45-15	1212
2.	Турнаев Валерий Иванович	Институт социально-гуманитарных технологий, Заместитель директора по научной работе	Научно-техническая библиотека, 354	8 (3822) 70-17-77	1211
3.	Чайковский Денис Витольдович	Институт социально-гуманитарных технологий, Проректор-директор	Научно-техническая библиотека, 353	8 (3822) 56-37-25	2099

<http://portal.tpu.ru/www/phone>

Виды баз данных

По сферам

в деловой сфере

в научной сфере

в военной сфере

в медицинской
сфере

По формату информации

текстовая

числовая

изображения

видеоинформация

Виды баз данных по размеру

За счет интеграции

БД уровня
рабочей группы

БД уровня
подразделения

БД уровня
предприятия

Банк данных

За счет большего периода накопления информации

малая



Сверхбольшая БД
(измеряются в Пб)

Система базы данных

База данных

Оборудования

**Система
управления БД**

Люди

Система управления базой данных

СУБД – это совокупность программ и языковых средств, предназначенных для управления данными в БД, ведения БД и обеспечения ее взаимодействия с прикладными программами

Преимущества СУБД

- **Контроль за избыточностью данных**
- **Непротиворечивость данных**
- **Применение стандартов**
- **Совместное использование данных**
- **Поддержка целостности данных**
- **Повышенная безопасность**
- **Улучшенное управление параллельной работой**

Недостатки СУБД

- **Сложность**
- **Размер**
- **Стоимость**
- **Дополнительные затраты на аппаратное обеспечение**
- **Затраты на преобразование**
- **Производительность**
- **Более серьезные последствия при выходе системы из строя**

Словарь данных СУБД

Содержит сведения обо всем, что хранится в базе данных

Наименования

Тип данных

Размещение данных



Метаданные

Безопасность данных в СУБД

- **Обязательная регистрация пользователей**
- **Распознавание пользователей до предоставления им доступа к данным**
- **Разделение пользователей на группы, с различными привилегиями доступа**
- **Поддержка профилей**

Целостность данных в СУБД

Целостность подразумевает взаимную **согласованность** отдельных фрагментов данных и их **корректность**.

Способы обеспечения целостности

- блокирование постороннего доступа к данным в процессе их обработки
- тиражирование изменений фрагмента данных
- отслеживание вводимых данных на соответствие требованиям

Многопользовательский доступ

data	USD	gold	corp_obl	mni_obl	CLOSE	MosPrime	MMVB
04.07.2009	31,2481	936,8400	154,32	109,28	115,05	10,07	1278.53
07.07.2009	31,4143	930,7100	154,48	109,31	115,11	10,03	1308.14
08.07.2009	31,4695	936,6400	154,52	109,63	115,11	9,94	1357.44
09.07.2009	31,7819	940,8300	154,89	109,32	115,13	9,93	1320.7
10.07.2009	31,8878	937,8200	154,86	109,44	115,14	9,97	1362.86
11.07.2009	32,0353	937,2600	154,82	109,07	115,14	10,97	1341.64
14.07.2009	33,0597	965,6400	154,91	109,43	115,13	10,72	1317.07
15.07.2009	32,5072	963,3500	155,11	109,94	115,21	10,41	1365.24
16.07.2009	32,0470	958,2100	155,3	109,68	115,46	10,26	1361.75
17.07.2009	31,6943	953,0100	155,3	109,89	115,48	10,16	1373.2
18.07.2009	31,7837	954,9400	155,23	110,72	115,61	10,08	1359.06
21.07.2009	31,3733	960,5100	155,38	110,42	115,75	10,14	1365.09
22.07.2009	31,1791	950,0500	155,86	110,67	115,76	10,1	1358.31
23.07.2009	31,0785	944,7400	155,9	110,88	115,89	10,17	1333.32
24.07.2009	31,0760	954,4100	155,99	111,22	115,81	10,12	1253.35
25.07.2009	31,1372	950,7800	155,8	111,5	115,8	10,1	1291.99
28.07.2009	30,7457	945,0000	156,2	111,88	115,87	10,03	1297.18
29.07.2009	30,6431	940,8600	156,34	112,21	115,88	10,14	1266.8
30.07.2009	31,4162	944,9100	156,15	112,17	115,87	10,04	1218.42
31.07.2009	31,7555	951,5400	156,43	112,63	115,96	10	1258.5



Модели данных на основе записей

1

Реляционная модель данных

2

Сетевая модель данных

3

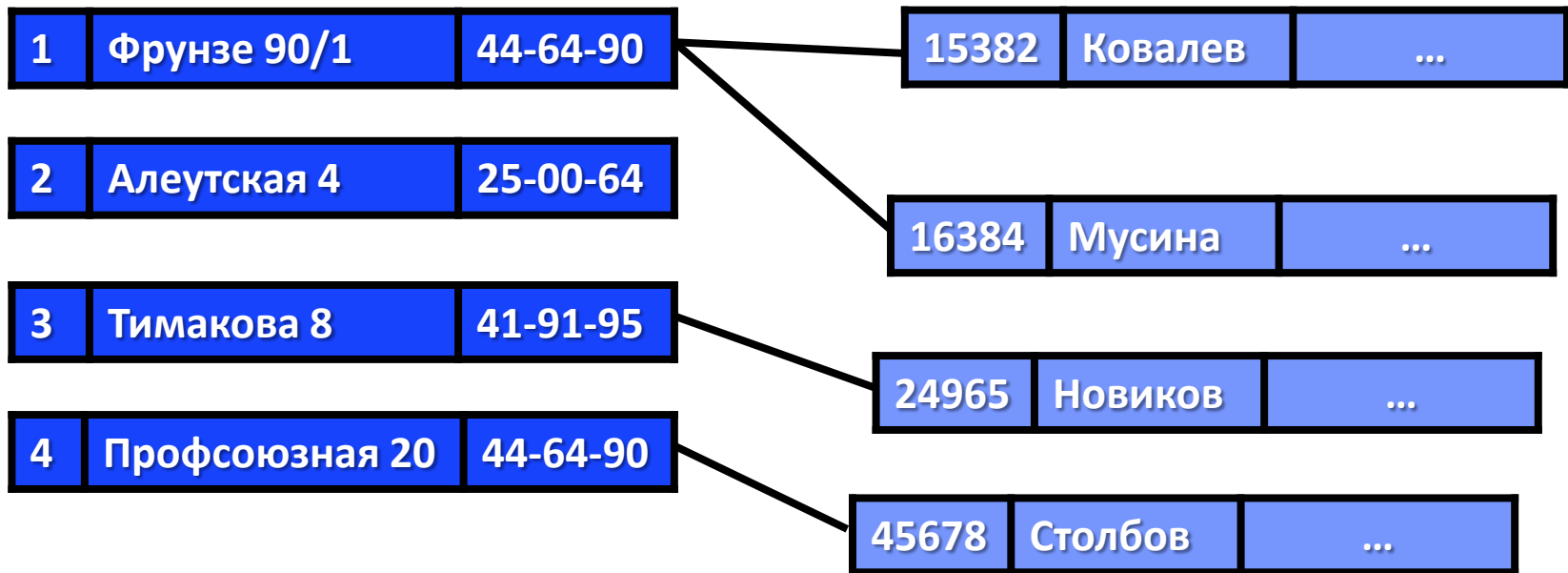
Иерархическая модель данных

Реляционная модель данных

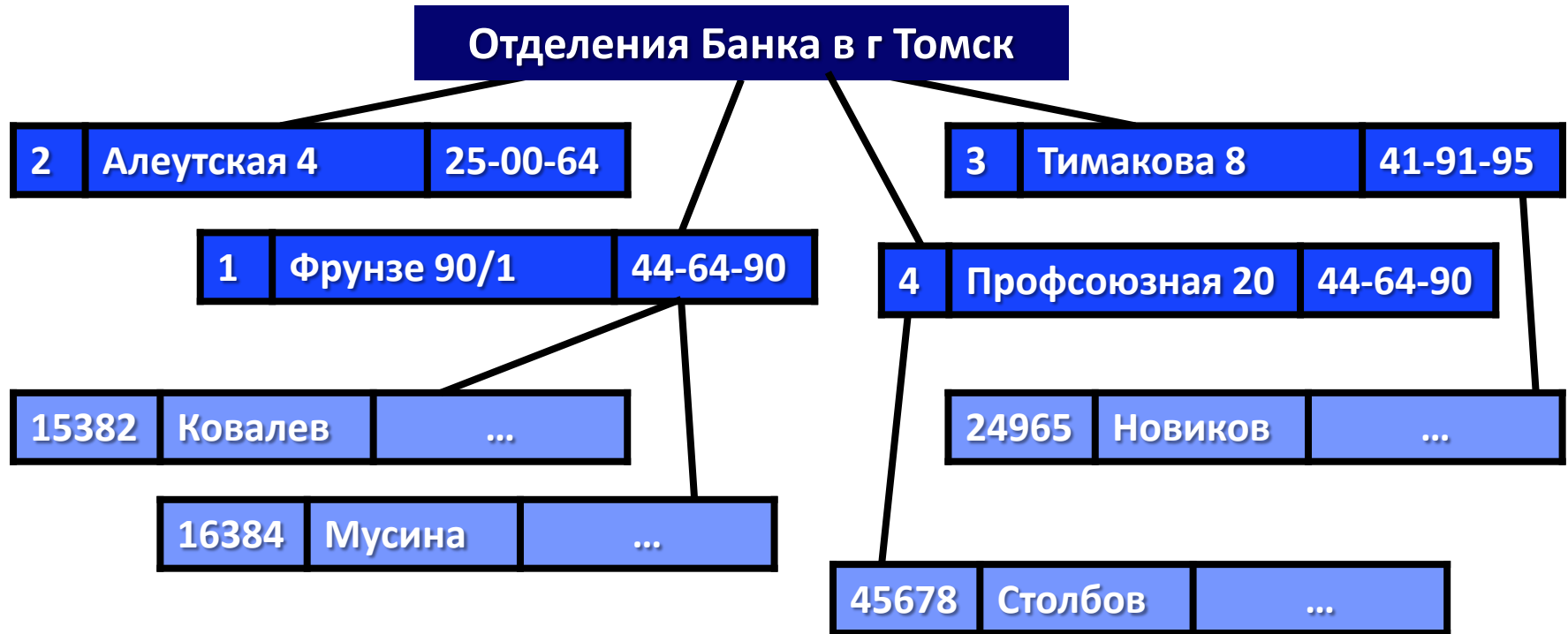
Office No	Address	Phone Number
1	Фрунзе 90/1	44-64-90
2	Алеутская 4	25-00-64
3	Тимакова 8	41-91-95
4	Профсоюзная 20	46-13-50

ID	Last Name	Name	Patronymic	Office No
15382	Ковалев	Артем	Викторович	1
16384	Мусина	Лидия	Степановна	1
24965	Новикова	Юлия	Викторовна	3
45678	Столбов	Михаил	Иванович	4

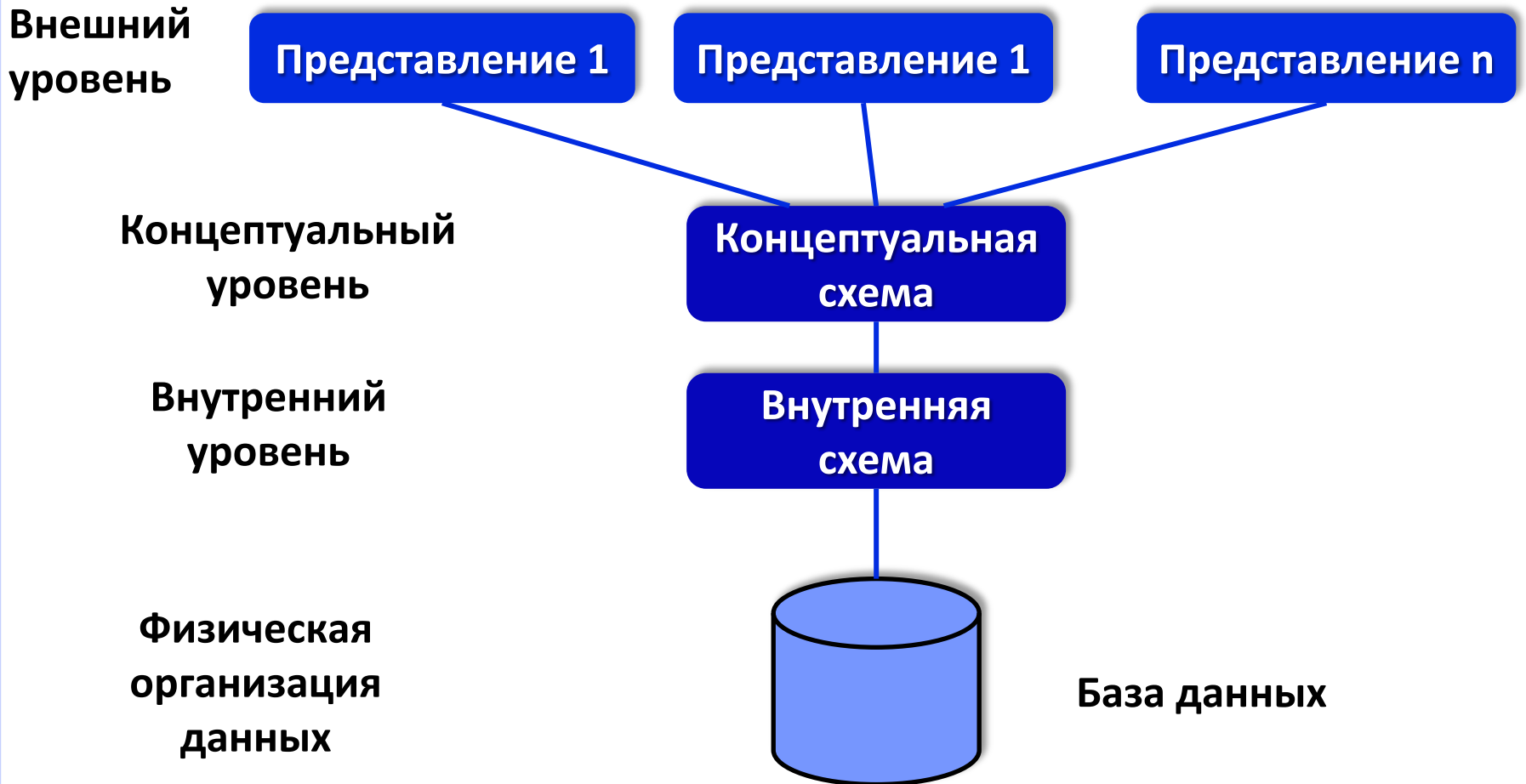
Сетевая модель данных



Иерархическая модель данных



Трёхуровневая архитектура БД



Внешний уровень

- представление БД с точки зрения пользователей

Отображение данных, интересных
для отдельного пользователя

Формат даты:
20 марта 2013
20/03/13

Предоставление производных или
вычисляемых данных

Возраст
сотрудника

Различное отображение одних и
тех же данных

Концептуальный уровень

- какие данные хранятся в БД и связи между ними

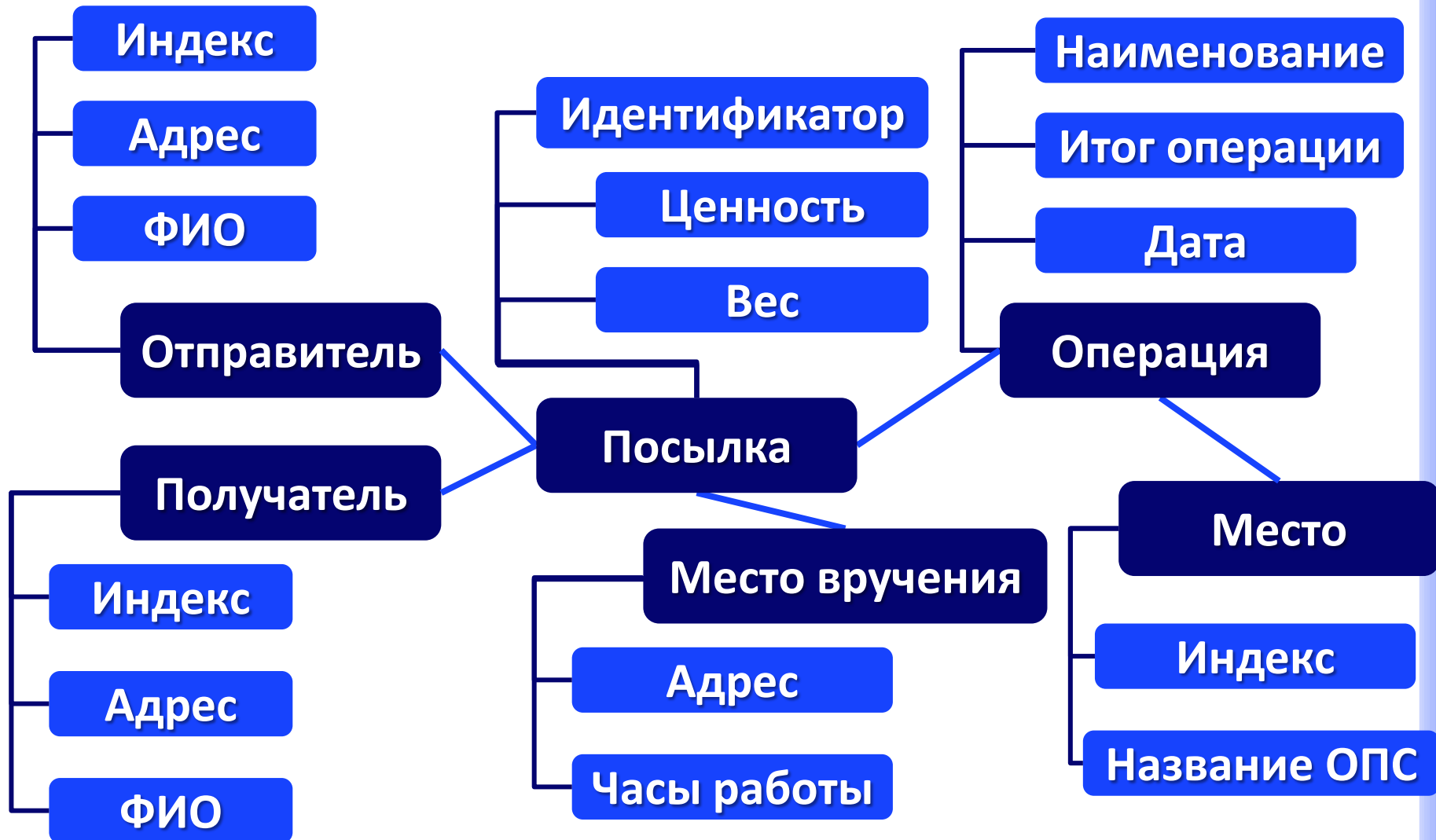
Все сущности и их атрибуты

Накладываемые ограничения

Семантическая информация

Человек –
сущность
ФИО – атрибут

Семантическая информация



Внутренний уровень

- как информация хранятся в БД

Распределение дискового
пространства

Описание подробностей
сохранения записей

Сведения о сжатии данных и
выбранных методах шифрования

Этапы проектирования БД

- 1. Концептуальное проектирование**
- 2. Логическое проектирование**
- 3. Физическое проектирование**
- 4. Проектирование транзакций**

Проектирование транзакций

Транзакция- это одно или несколько действий, выполняемых одним и тем же пользователем, которые получают доступ к БД или изменяют ее содержимое.

Пример Перевод денег с одного счета на другой

Виды транзакций

- Транзакции извлечения
- Транзакции обновления
- Смешанные транзакции

Спасибо за внимание!