

**Место и роль  
информационных моделей,  
программных инструментов и  
организационных решений  
в создании систем с неоднородными  
компонентами.  
Теория? Опыт реализации?**

*Зиндер Евгений Захарович,*

*Фонд ФОСТАС (дистанционное участие)*

*e-mail: [ezinder@fostas.ru](mailto:ezinder@fostas.ru)*

## Системы с неоднородными компонентами

*НАДО ОПРЕДЕЛЯТЬСЯ С ТЕМ, КАКИЕ СИСТЕМЫ РАССМАТРИВАЮТСЯ:*

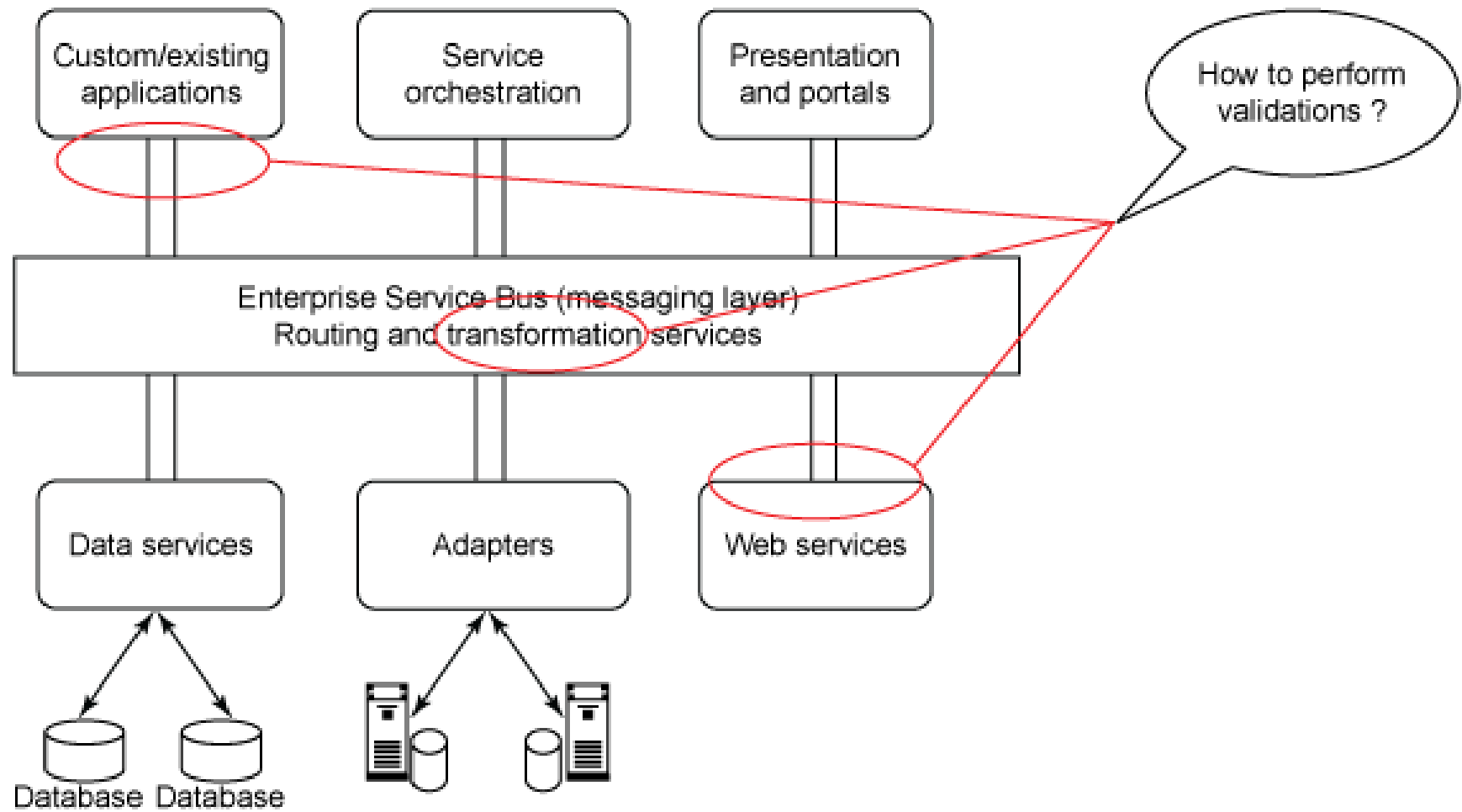
- Типа корпоративных ИС или
- т.н. Информационные пространства (ИП)

Качественный переход с уровня корпоративных систем (шире объединения, отрасли и т.д.):

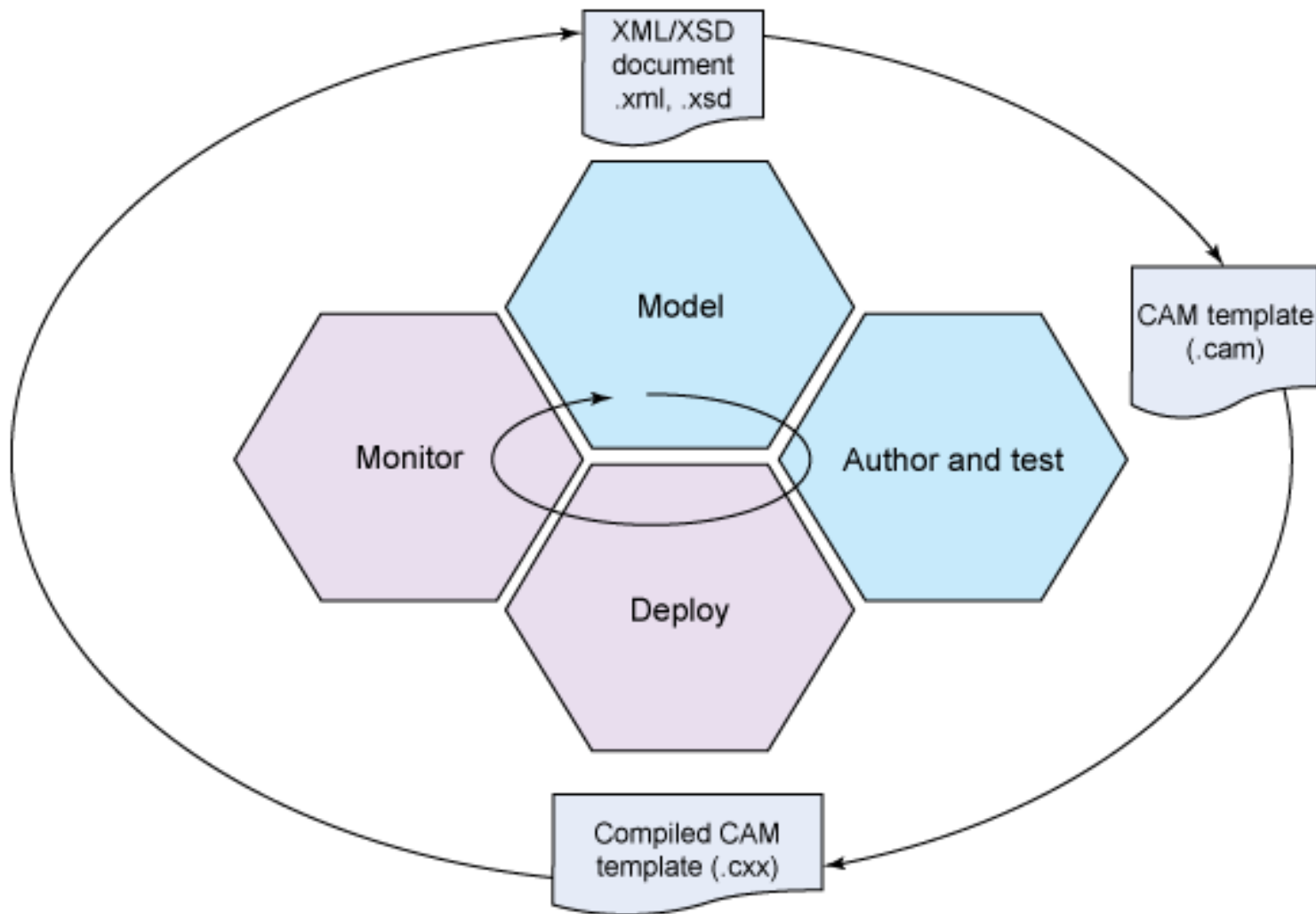
- Наличие определенного центрального субъекта (как сообщества) или **отсутствие**
- С поддержкой «Среды обитания» -- одной или нескольких и каких

и другие качественные отличия.

# «Корпоративный» подход (до 2007 года)



# Схема проверок (и возможных преобразований) в САМ от IBM: постоянство и тяга к централизации



# Модели, инструменты и организация

**Нужное** в проектировании ИБД и ИП известно после 1975:

- **типы моделей** (т.к. компьютерные семантические сети разработаны ранее)
- **как и типы переходов между моделями.**

**К сожалению:**

- виды моделей и переходов И ДАЖЕ «что такое семантика» часто определяют не сутью, а инструментами (((
- инструменты (языки в т.ч.) и процедуры имеют ограничения, часто непреодолимые (NP-полнота, ограничения на виды семантических связей, ...)

**Инструменты** (включая эффективные преобразователи, “Data Handlers”) **есть не для всех моделей и переходов**, что-то пишется at hoc

Без организационных структур и усилий

- не будут построены модели (и в Wikipedia есть организация!)
- не будут заполнены позиции в процессе взаимодействий и интеграции, где нет подходящего инструмента

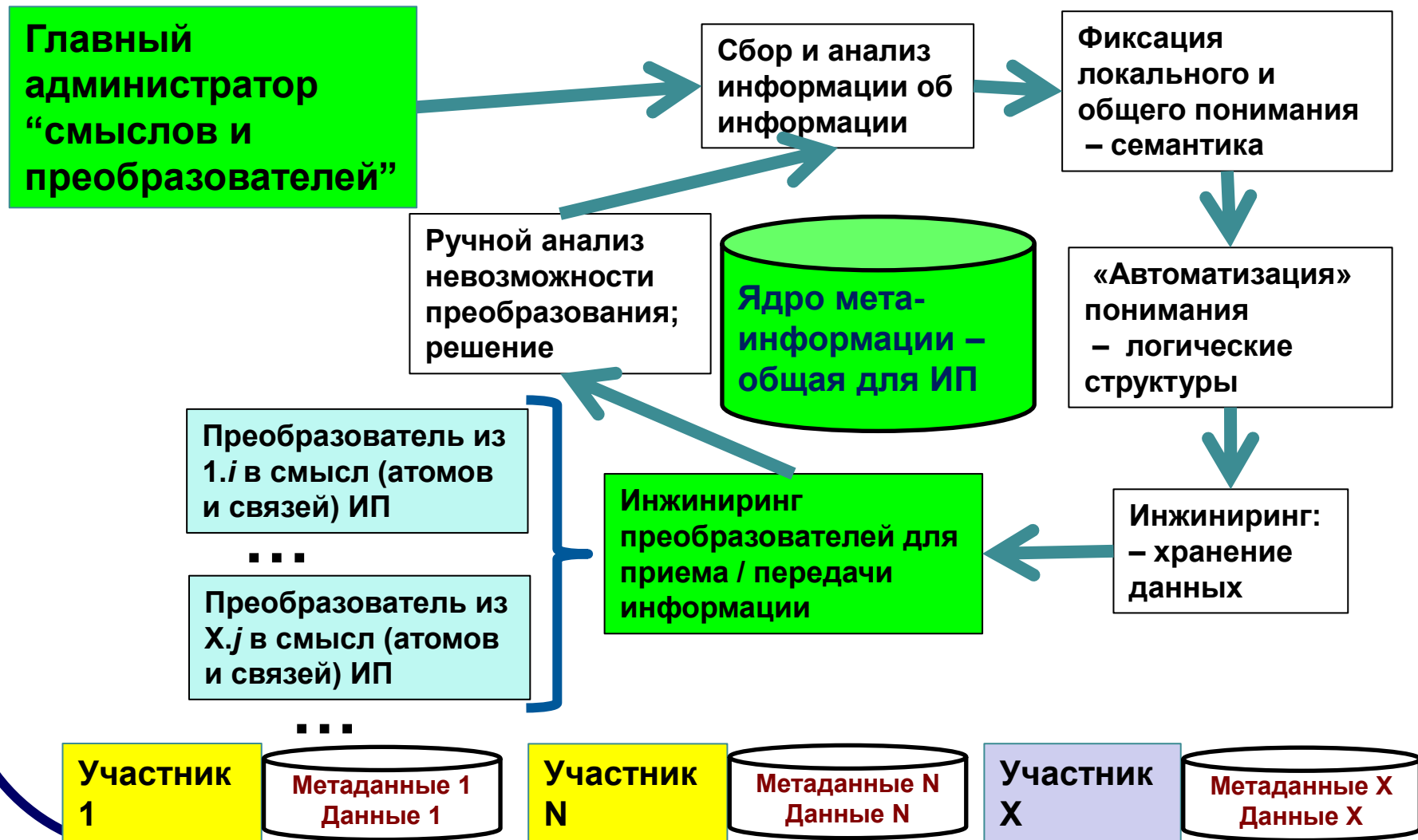
# Информационные модели ИП – есть варианты подходов и архитектур

## НА ЧЕМ ОСНОВЫВАТЬ?

- Идти
  - от семантических моделей («моделей знаний», ЕЯ, и т.п.)
  - или от обобщения теории интегрированных баз данных (общие метаданные, совмещение представлений, и др.)
- **Вариант с централизацией: да, пока так** (м.б. центральный субъект, но всегда явно централизованная семантика)
- Вариант без централизации (пока и еще долго – вряд ли!)
- Множество семантических моделей (СМ) рассматривается в любом случае (?!)
- Инструменты поддержки СМ и логических преобразований – существуют разные, но не исчерпывающие
- Не все инструменты практически пригодны для всех трех уровней (по ANSI SPARC 75) общения в ИП

# Цикл работ в ИП:

Где начало? Кто инициатор? Кто ответственный?  
Как определяется ядро метаинформации? Какова автоматическая часть контроля совместимости? Насколько автоматически формируются преобразователи (Data Handlers)?



# Цикл работ в ИИ и лакуны практики

	Инф. модели	Прогр. инструменты	Орг. решения
<b>Сбор и анализ информации об информации</b>	Анкеты, Словари синонимов, омонимов... Классификаторы, кодификаторы	SW-обработчики анкет BPMN-моделлеры?! ER-моделлеры?!	Политические усилия и административные решения о сборе (ответах) и анализе
<b>Синтез и фиксация понимания – семантика</b>	Глоссарии, Тезаурусы, Семантические сети, онтологии, Классификаторы («коды»)	Редакторы OWL?! RDF?! Gellish!!!	Административные решения → Затраты на инструменты и орг. единицы
<b>НЕОДНОЗНАЧНЫЕ ПРЕОБРАЗОВАНИЯ</b>			
<b>– логические структуры</b>	SQL-схемы, XML-схемы	Алгоритмы получения?! DBMS (SQL) CAM Components?!	
<b>– хранение данных</b>	SQL...	DBMS	
<b>Транзакции: Передача и приём сообщений и информации в них</b>	Представления («view») Шаблоны и правила соответствий Схемы SQL, XML Алгоритмы at hoc?!	Запросы SQL Скрипты CAM Templates CAM content validation Data services (std)	Затраты на инструменты и орг. единицы
<b>Ручной анализ невозможности преобразования и авт. решения</b>	Классификаторы и шаблоны несоответствий, область ручных действий	Запросы SQL ETL Скрипты Data services (at hoc)	Затраты на инструменты и орг. единицы Политики. Мотивация



# Что делается «мозгом вручную»?

Сбор и анализ информации об информации

Анкеты, Словари синонимов, омонимов...  
Классификаторы, кодификаторы

## НЕОДНОЗНАЧНЫЕ ПРЕОБРАЗОВАНИЯ

Фиксация понимания  
– семантика

Глоссарии, Тезаурусы  
Классификаторы («коды»)

## НЕОДНОЗНАЧНЫЕ ПРЕОБРАЗОВАНИЯ

– формальное  
представление  
интегрированной, «единой»  
семантики

Плоскости абстракции и агрегации  
Разновидности семантических связей и их  
модальности  
Темпоральные связи  
Семантическая близость и условные связи  
Семантические сети, онтологии  
RDF-схемы ?? Онтологии Gellish ??

## НЕОДНОЗНАЧНЫЕ ПРЕОБРАЗОВАНИЯ

– логические структуры

SQL-схемы, XML-схемы

## NIEM NDR 9.2 Usage of English

The name of any XML Schema component defined by the schema SHALL be composed of words from the English language, using the prevalent U.S. spelling, as provided by **[OED]**.

**Rationale** The English language has many spelling variations for the same word. For example, American English “program” has a corresponding British spelling “programme.” **This variation has the potential to cause interoperability problems** when XML components are exchanged because of the different names used by the same elements. Providing users with a dictionary standard for spelling will mitigate this potential interoperability issue.

# Какого типа преобразования в ИП нужны?

## Для каждого разумен и нужен свой набор средств

### 1) Распознавать «заголовки» и адресовать сообщения?

→ распознавать минимум элементов и их кодов, обеспечить общий набор отображаемых символов, оставить распознавание реципиенту-человеку

### 2) Распознавать «заголовки», адресовать сообщения, контролировать выполнение заявок по простым критериям?

→ м.б. распознавать большее число типов элементов (атомов) и их кодов, но все еще в пределах специфических расширений дублинского ядра

### 3) Распознавать все сообщения полностью и включать их элементы (атомы) во внутреннее ИП получающей организации?

- производить отображение не только «имён», но имен со всеми релевантными семантическими связями в схему получателя,
- проверять эквивалентность концептуальных схем и границы их смысловой совместимости,
- корректно преобразовывать структуру данных для помещения в «свою» схему без потери информации

## Если для ИШ ЭП принять достаточность:

2) Распознавать «заголовки», адресовать сообщения, контролировать выполнение заявок по простым критериям?

→ м.б. распознавать большее число типов элементов (атомов) и их кодов, но все еще в пределах специфических расширений дублинского ядра

### **то и в этом случае**

- **«ручная работа» все еще требуется на первых шагах этапа анализа и формирования семантики: анализ, фиксация понимания, формальное представление**
- **и для следующих шагов рекомендуется ограничивать схемы онтологий упрощенными графами (например, без транзитивных связей), допускающими простые проверки соответствий и эквивалентности.**

## Вопросы для обсуждения и ответы : №1

- **Вопрос:** Являются ли вопросы информационного моделирования межведомственного / межкорпоративного взаимодействия более важными по сравнению с вопросами технологического и организационного характера?
- **Ответ:** **Нет**

## Вопросы для обсуждения и ответы : №2

- **Вопрос:** Что должно являться информационной основой межведомственного/межкорпоративного обмена: модели предметной области или модели более высокого уровня с описыванием метаданных потоков данных, набора данных?
- **Ответ:**
  - (а) **и то, и другое, и третье**, но:
  - (б) каков измеритель «высоты уровня»?! В каком «дискурсе» упомянуты метаданные?! (см. **NIEM NDR 9.8 Object-Class Term** )

## Вопросы для обсуждения и ответы : №3

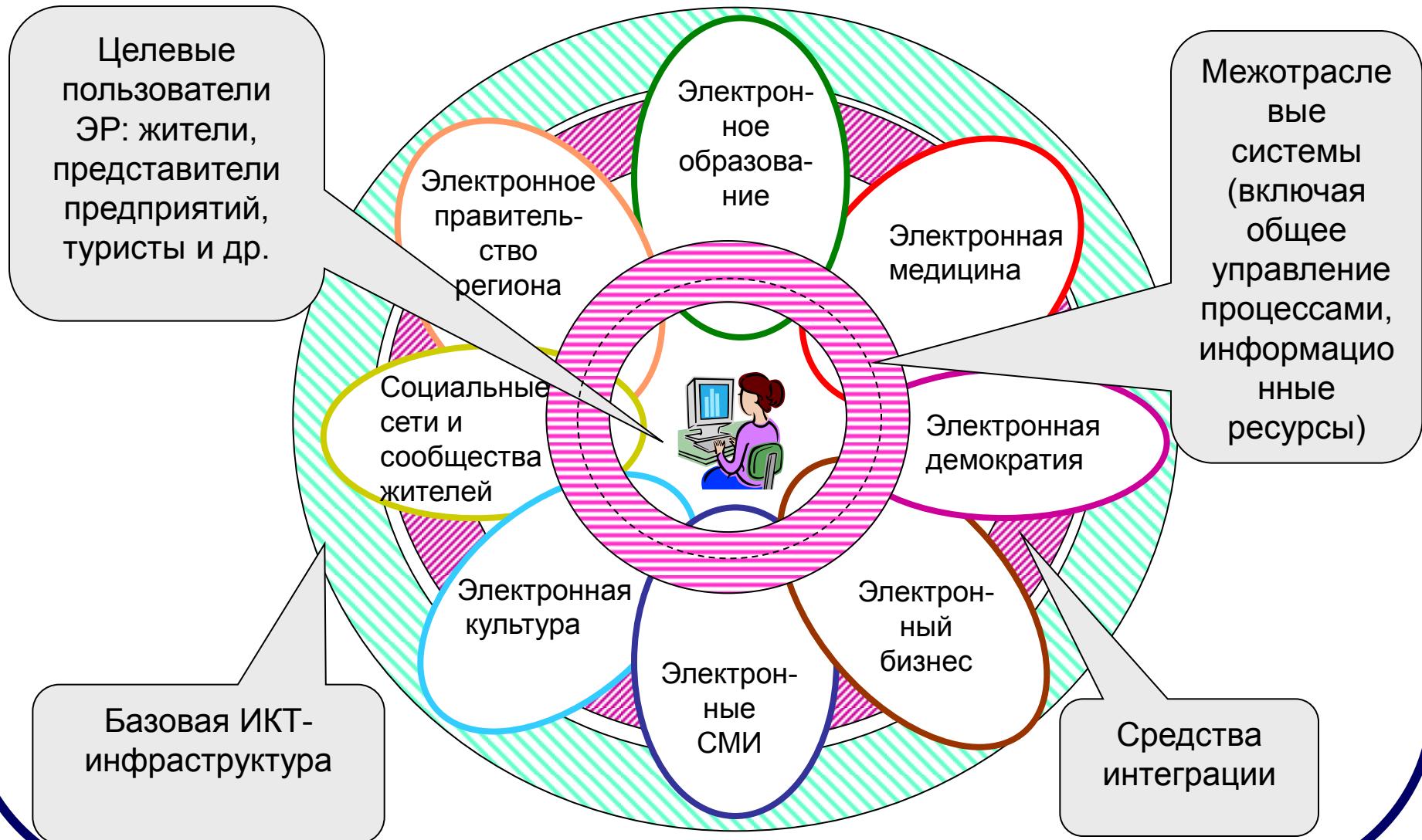
- **Вопрос:** Какие элементы инфраструктуры являются критическими для информационного обмена:
  - единые списки кодов,
  - словари метаданных,
  - категоризация предметных областей,
  - другие?
- **Ответ:**
  - (а) **и то, и другое, и третье с четвертым**, но:
  - (б) в какой терминологии задан вопрос? В каком «дискурсе»?!

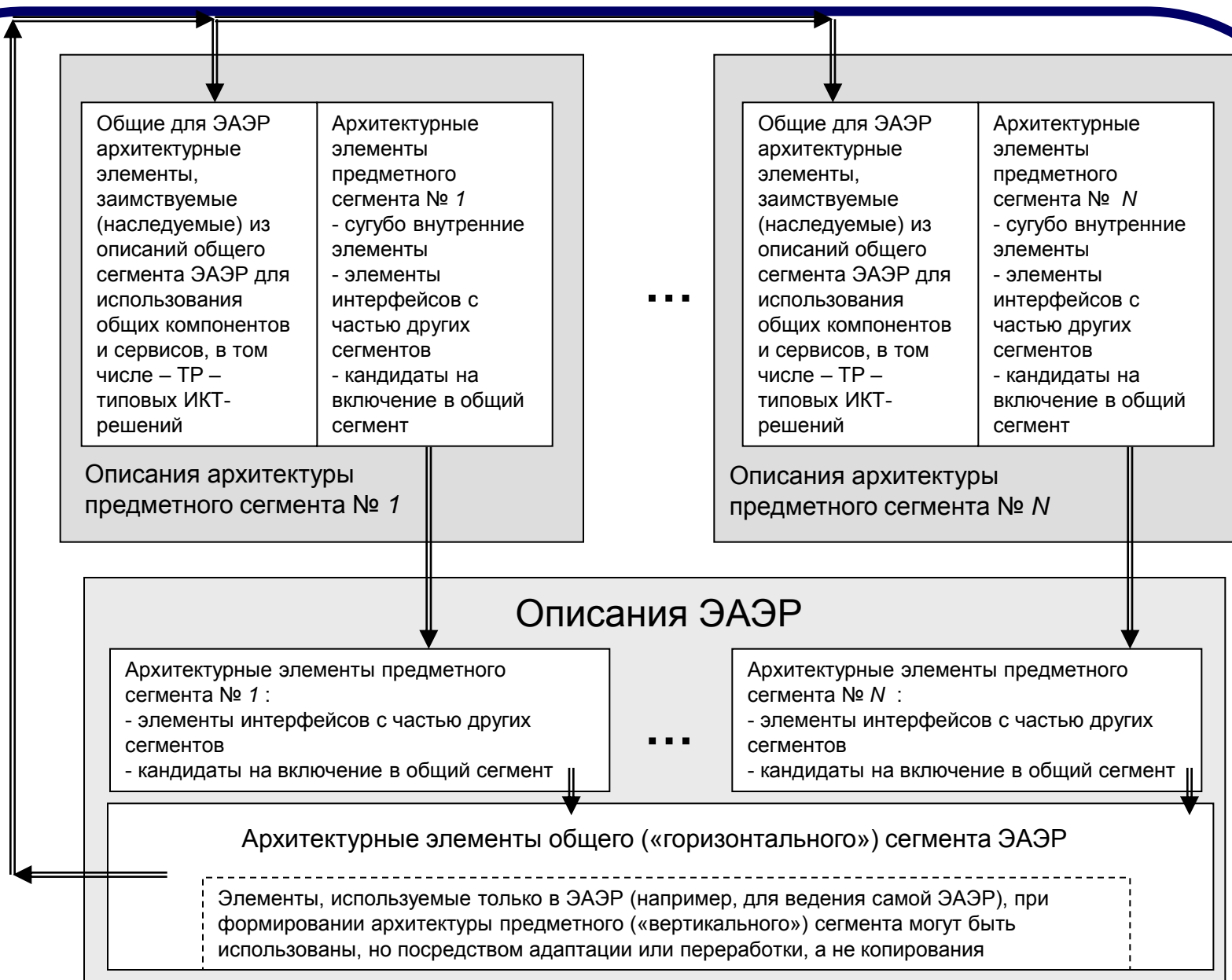
**Проект 2006-2007 г.г.:**

**схема-ромашка поддержки  
электронного региона  
архитектурой ведения репозиториев (ядра и  
доменов) и сервисов данных  
(аналогична NIEM)**



# «Вертикальные» и «горизонтальный» сегменты эл. региона





**Рис. XX.XX. Взаимосвязи описаний ЭАЭР и архитектур, разработанных для предметных сегментов электронного региона (при сочетании централизованного и децентрализованного управления развитием ИКТ в субъекте РФ).** *Примечание:* стрелки показывают потоки передачи описаний из ЭАЭР в архитектуры предметных сегментов и из последних в ЭАЭР в процессе постоянного развития архитектуры электронного региона.

**Спасибо за внимание!**

**Ваши вопросы?**

*Зиндер Е. З., НО Фонд ФОСТАС*

*E-mail: [ezinder@fostas.org](mailto:ezinder@fostas.org)*