

IBM InfoSphere  
Information Governance Catalog

# Управление цифровыми знаниями в организации

# План

Все, что вы хотели узнать, но боялись спросить



## Зачем управлять информацией?

Современные вызовы. Признаки проблем с управлением. Что такое управление информацией



## Служба управления данными

Chief Data Officer. Процессы управления данными. Путь к зрелости



## IBM Information Governance Catalog

Концепция IBM IGC. Бизнес глоссарий. Информационные активы. Происхождение данных



## История успеха

Трансформация крупного розничного оператора



# Зачем управлять информацией?

Если и так все работает...

# Современные вызовы

В век информационных технологий и больших данных  
компании сталкиваются со множеством вызовов

01

ЭЛЕКТРОННЫХ ДАННЫХ  
СТАНОВИТСЯ СЛИШКОМ МНОГО

02

НЕПОНЯТНО КАК УПРАВЛЯТЬ  
ДАННЫМИ

03

МОЖНО ЛИ ДОВЕРЯТЬ  
ИМЕЮЩИМСЯ ДАННЫМ?

04

ПОНИМАЕМ ЛИ МЫ РИСКИ,  
СВЯЗАННЫЕ С ИНФОРМАЦИЕЙ?

К 2025 году объём цифровых данных вырастет  
с 33 000 000 000 Тб до

**175 000 000 000 Тб**

# Электронных данных становится слишком много



- Рост стоимости хранения данных в результате роста бизнеса и расширения автоматизации
- Повышение затрат, связанных с выполнением регуляторных требований к хранению и передаче данных
- Снижение производительности сотрудников при поиске необходимой информации, связанной с архитектурой хранения данных
- Значительное увеличение количества владельцев данных по множеству информационных систем

# Непонятно как управлять данными?

- Какими данными мы обладаем?
- Где они хранятся?
- Зачем они нам нужны?
- Где используется информация?
- Кто имеет к ним доступ?
- Как двигаются и трансформируются данные?



# Можно ли доверять имеющимся данным?

- Данные хранятся во множестве независимых систем
- В процессе загрузки, очистки и трансформации данные могут быть искажены или потеряны
- Актуальность хранимых данных
- Целостность данных



# Понимаем ли мы риски, связанные с информацией?



- Защищена ли информация?
- Множество независимых систем хранения информации повышают вероятность несанкционированного доступа к данным
- Обеспечение целостности данных
- Противоречивые данные в системах-источниках: определение, назначение и обогащение мастер-данных

# Управление информацией

это целостный подход к управлению корпоративными знаниями о данных путем внедрения процессов, ролей, элементов управления и метрик, которые рассматривают знания как ценный бизнес-актив

Цель управления информацией — предоставление менеджменту и сотрудникам данных, которым они могут доверять и которые легко доступны при принятии бизнес-решений





# Служба управления данными

Зачем, если есть IT-служба...

# Служба управления данными

Chief Data Officer (CDO) — бизнес-лидер, создающий и реализующий стратегии обработки и анализа данных для монетизации и повышения стоимости бизнеса

CDO отвечает за формирование, разработку и реализацию стратегий определяющих методики для получения, анализа и управления данными



# Процесс управления данными

## 01 Дизайн данных

- единый глоссарий бизнес-терминов организации
- единая бизнес-модель
- корпоративное хранилище данных как единый источник управленческой и регуляторной отчётности

## 02 Качество данных

- стандарты управления качеством
- процессы управления качеством и мониторинга качества данных
- определение владельцев данных
- инструментарий контроля качества данных

## 03 Происхождение данных

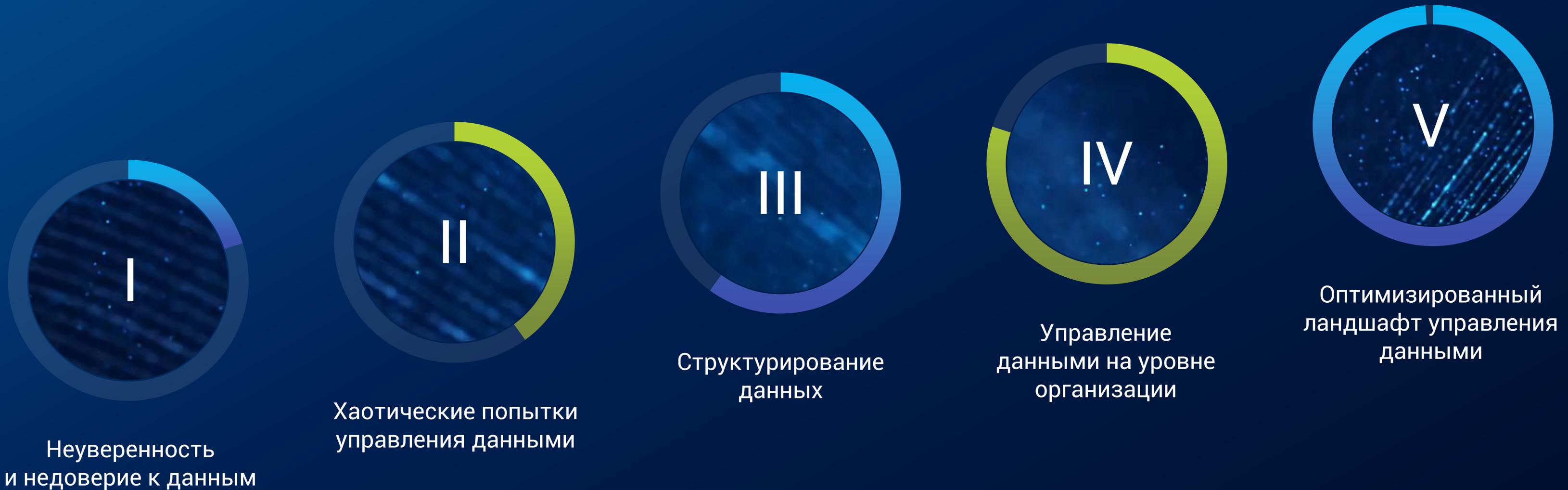
- единый инструментарий по анализу происхождения и преобразований данных
- анализ проблем с качеством данных
- анализ и выявление взаимных зависимостей и влияния

## 04 Управление стоимостью данных

- классификация данных исходя из их ценности
- управление стоимостью владения данными



# Стадии зрелости процессов управления информацией



01



УМЕНЬШЕНИЕ  
ОБЪЕМА ДАННЫХ  
И СВЯЗАННЫХ  
С ЭТИМ РАСХОДОВ  
И РИСКОВ

БЫСТРАЯ РЕАКЦИЯ НА  
ПРЕДСТОЯЩИЕ ИЛИ  
ИЗМЕНЯЮЩИЕСЯ  
ТРЕБОВАНИЯ

02



03



УПРАВЛЕНИЕ  
ДАННЫМИ ПУТЁМ  
ВНЕДРЕНИЯ  
ПОЛИТИК И ПРАВИЛ  
ИХ ОРГАНИЗАЦИИ

СОКРАЩЕНИЕ  
ВРЕМЕНИ И ОБЪЁМОВ  
ДАННЫХ ДЛЯ  
ПОИСКА

04



05



ОБЕСПЕЧЕНИЕ  
УСТОЙЧИВОГО  
КОНКУРЕНТНОГО  
ПРЕИМУЩЕСТВА

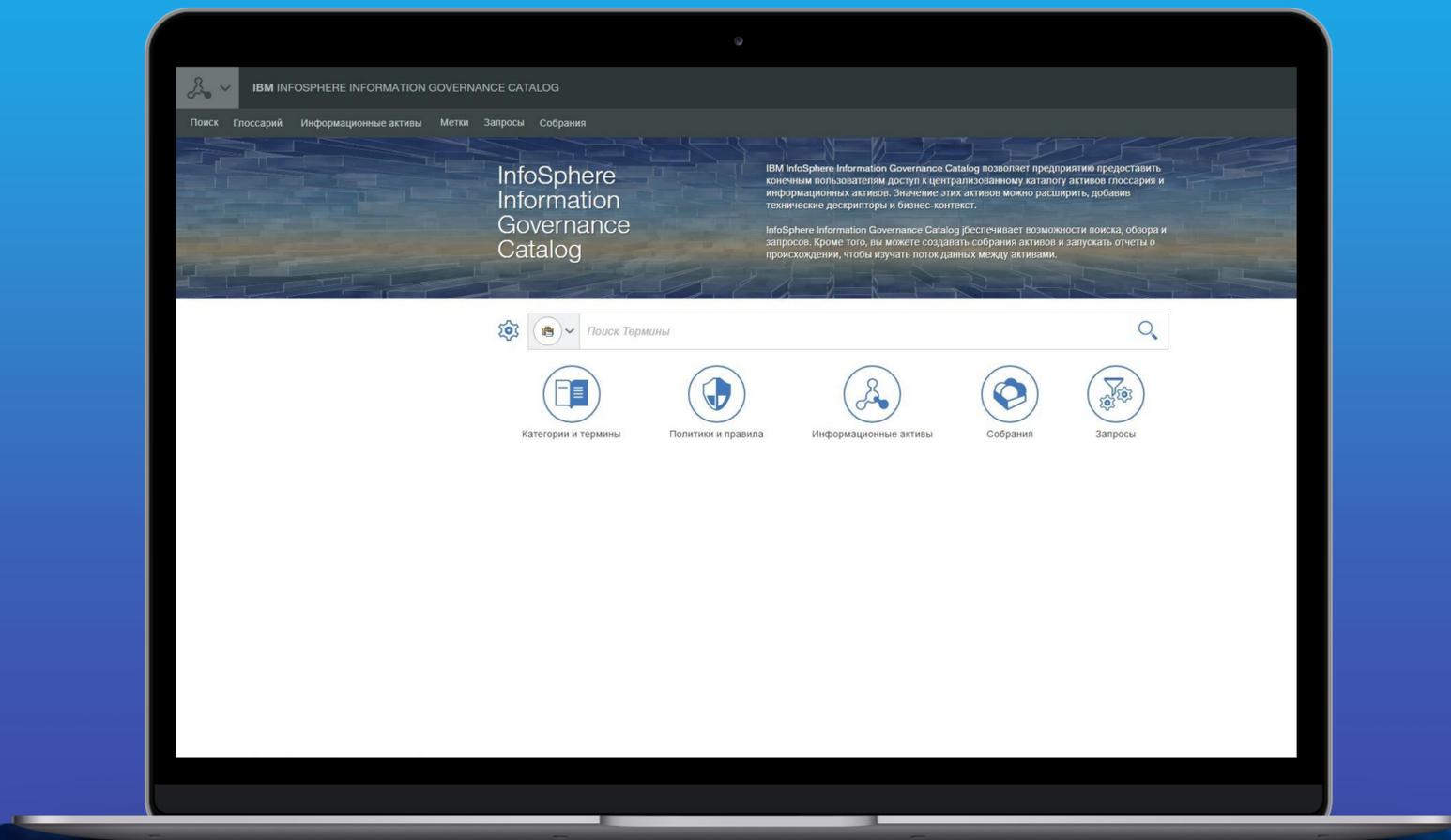
# 5 столпов управления информацией



# IBM Information Governance Catalog

Пока не понятно...

# Основа IBM Information Governance Catalog



КАТАЛОГИ  
ИНФОРМАЦИОННЫХ  
АКТИВОВ В БИЗНЕС-  
ТЕРМИНАХ

ОБЩИЙ  
РЕПОЗИТОРИЙ  
МЕТАДАННЫХ

СЕМАНТИЧЕСКИЕ  
ПОЛИТИКИ  
И ПРАВИЛА  
УПРАВЛЕНИЯ  
ИНФОРМАЦИЕЙ

ВЗАИМОСВЯЗИ  
МЕЖДУ СЛОВАРЯМИ И  
ИНФОРМАЦИОННЫМИ  
АКТИВАМИ

ОТВЕТСТВЕННОСТЬ  
ЗА ДАННЫЕ

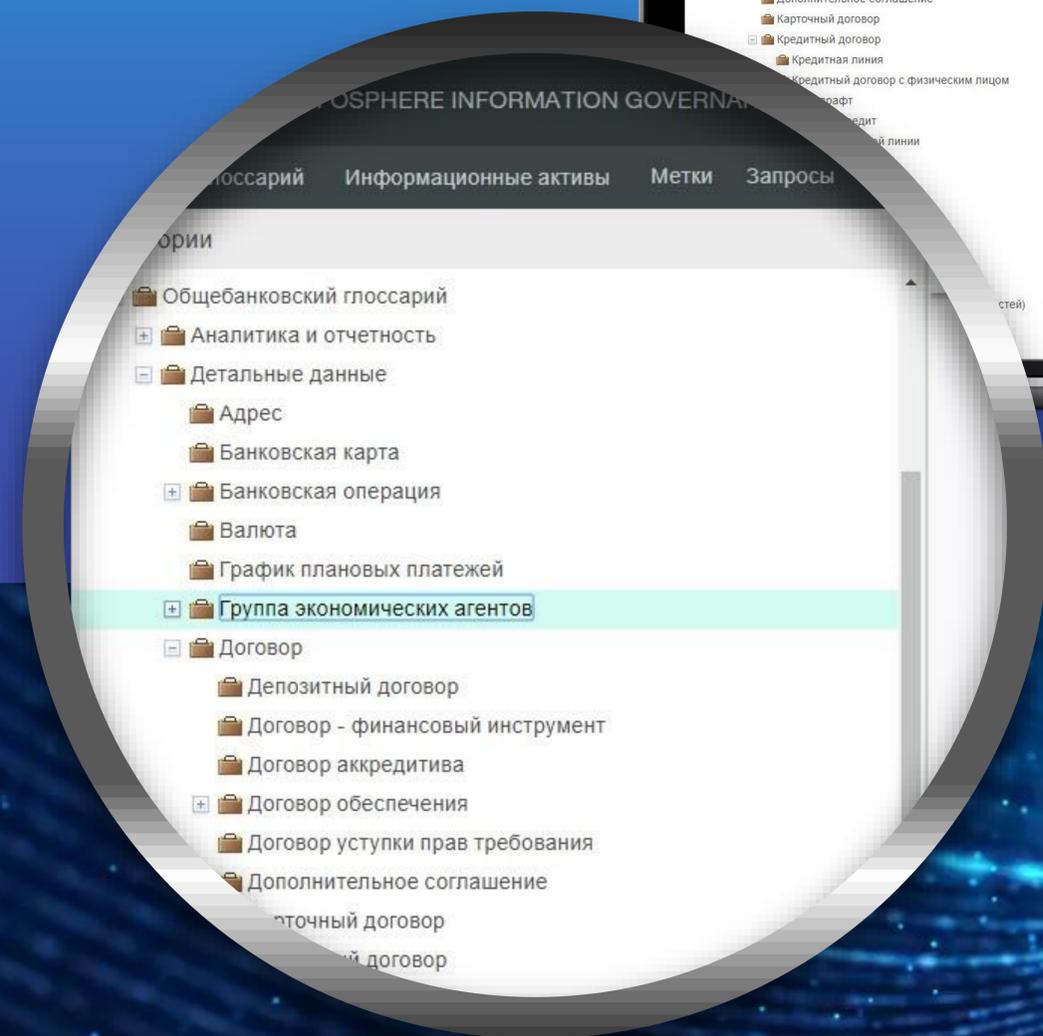
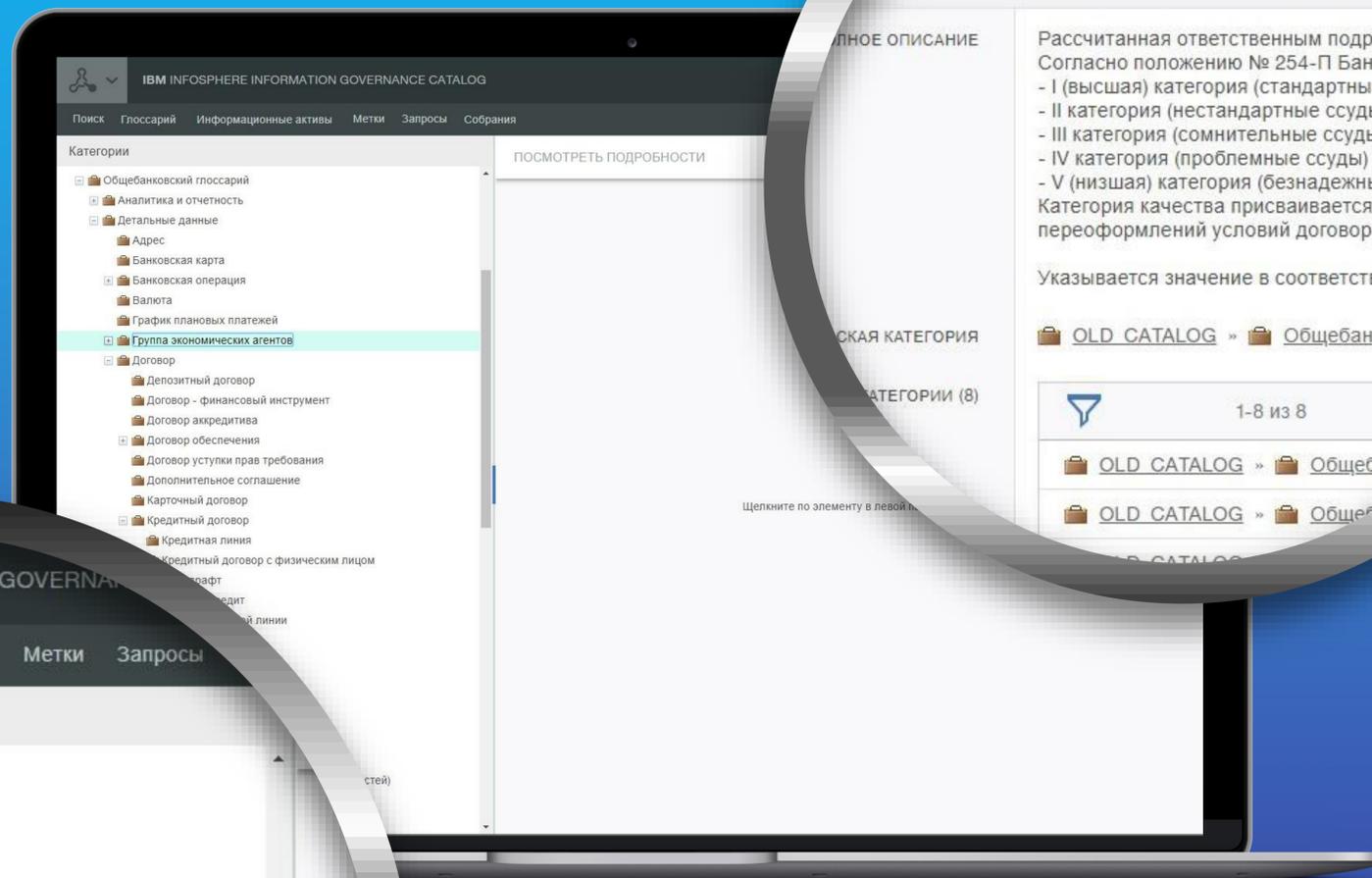
СОВМЕСТНАЯ  
РАБОТА  
С МЕТАДАНЫМ

# Каталоги бизнес-терминов

Бизнес-словарь — структурированный каталог бизнес-атрибутов организации, их определения и методики формирования

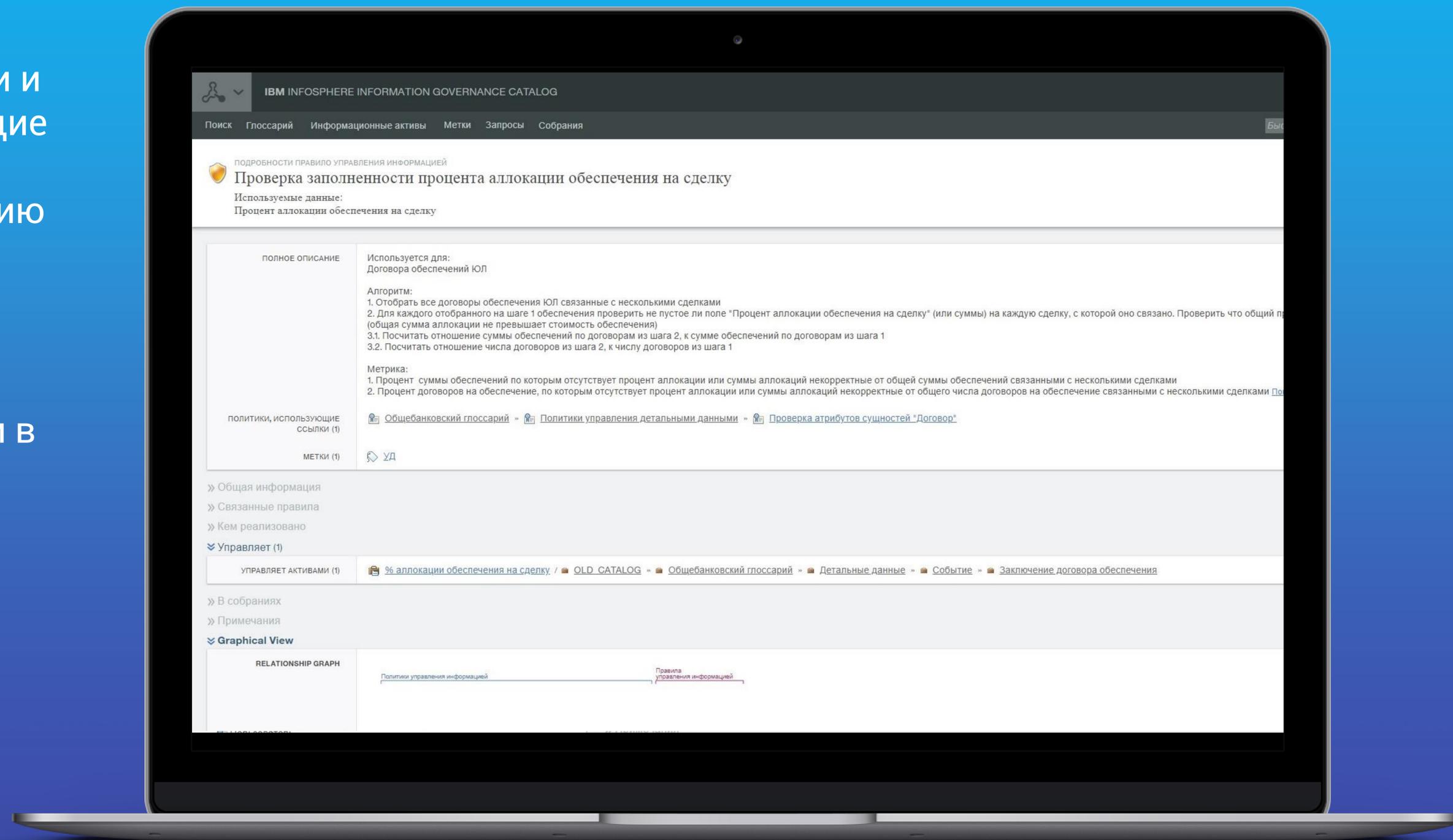
Содержит:

- Описание
- Категории
- Статус
- Правила управления
- Синонимы
- Информационные активы
- Ответственные за термин и правила
- Чат по термину



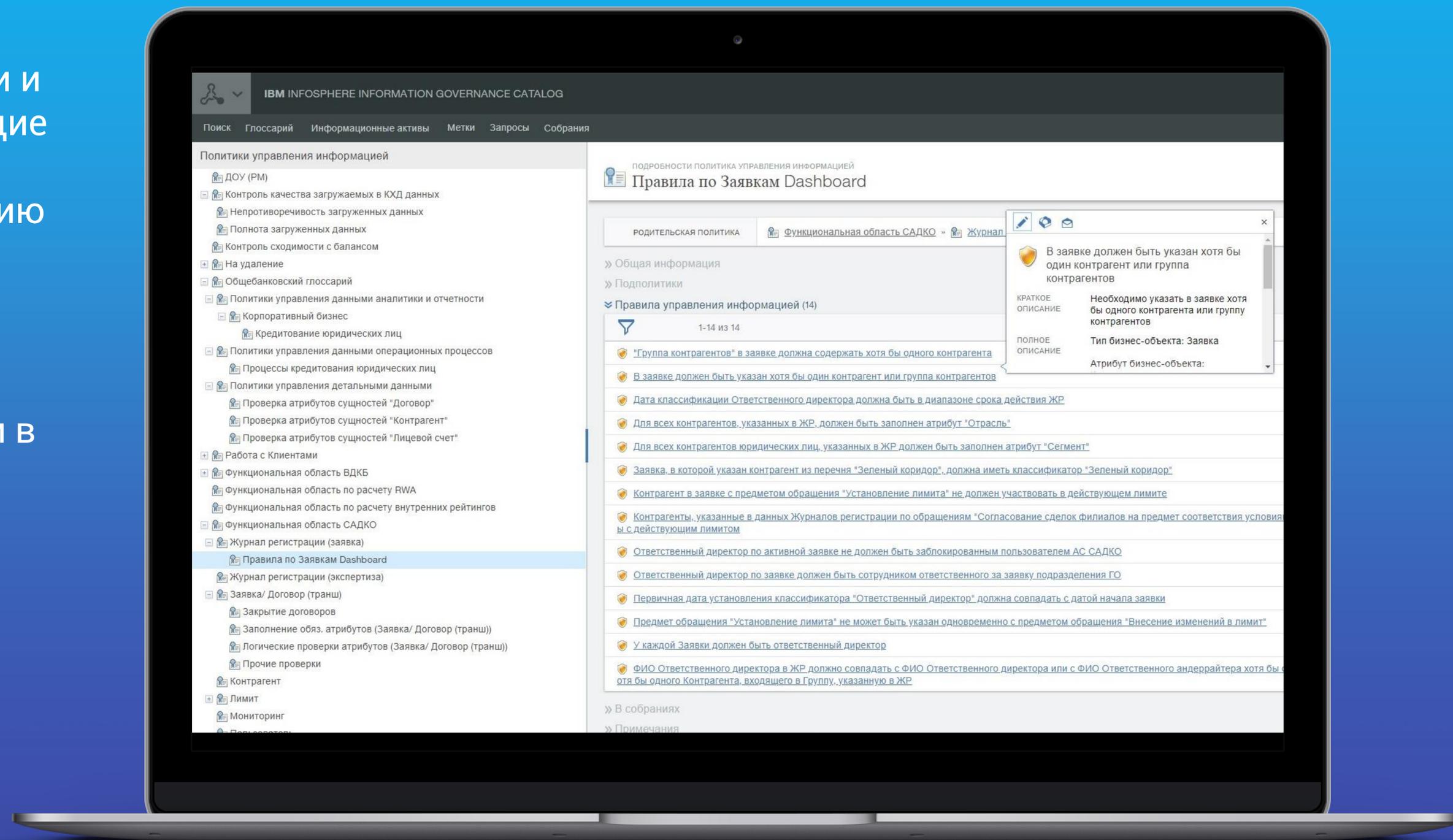
# Управление бизнес-терминами

- Семантические политики и правила, поддерживающие точную передачу требований по управлению информацией
- Общий язык, рационализирующий разработку информации в соответствии с требованиями бизнеса



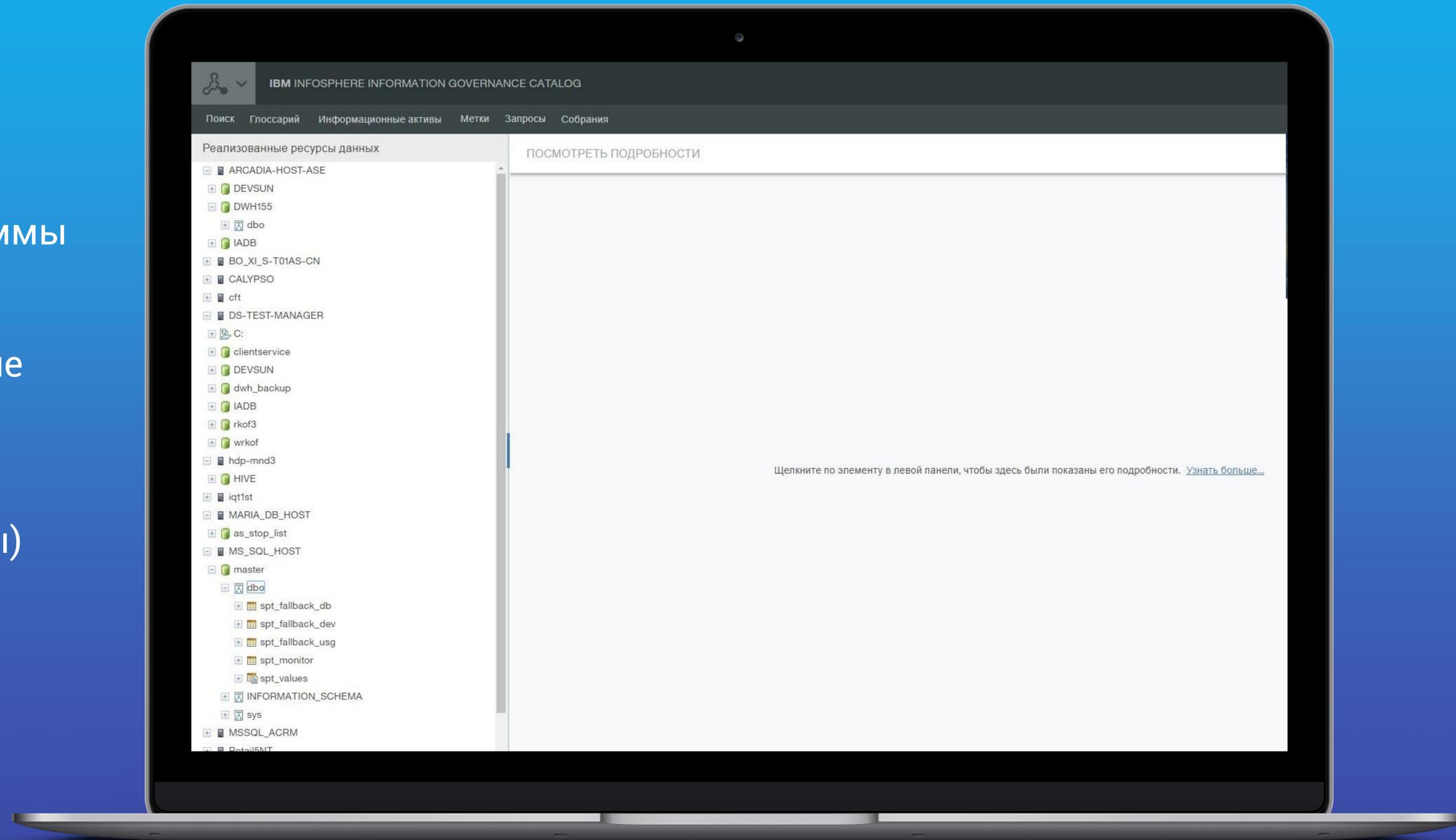
# Управление бизнес-терминами

- Семантические политики и правила, поддерживающие точную передачу требований по управлению информацией
- Общий язык, рационализирующий разработку информации в соответствии с требованиями бизнеса



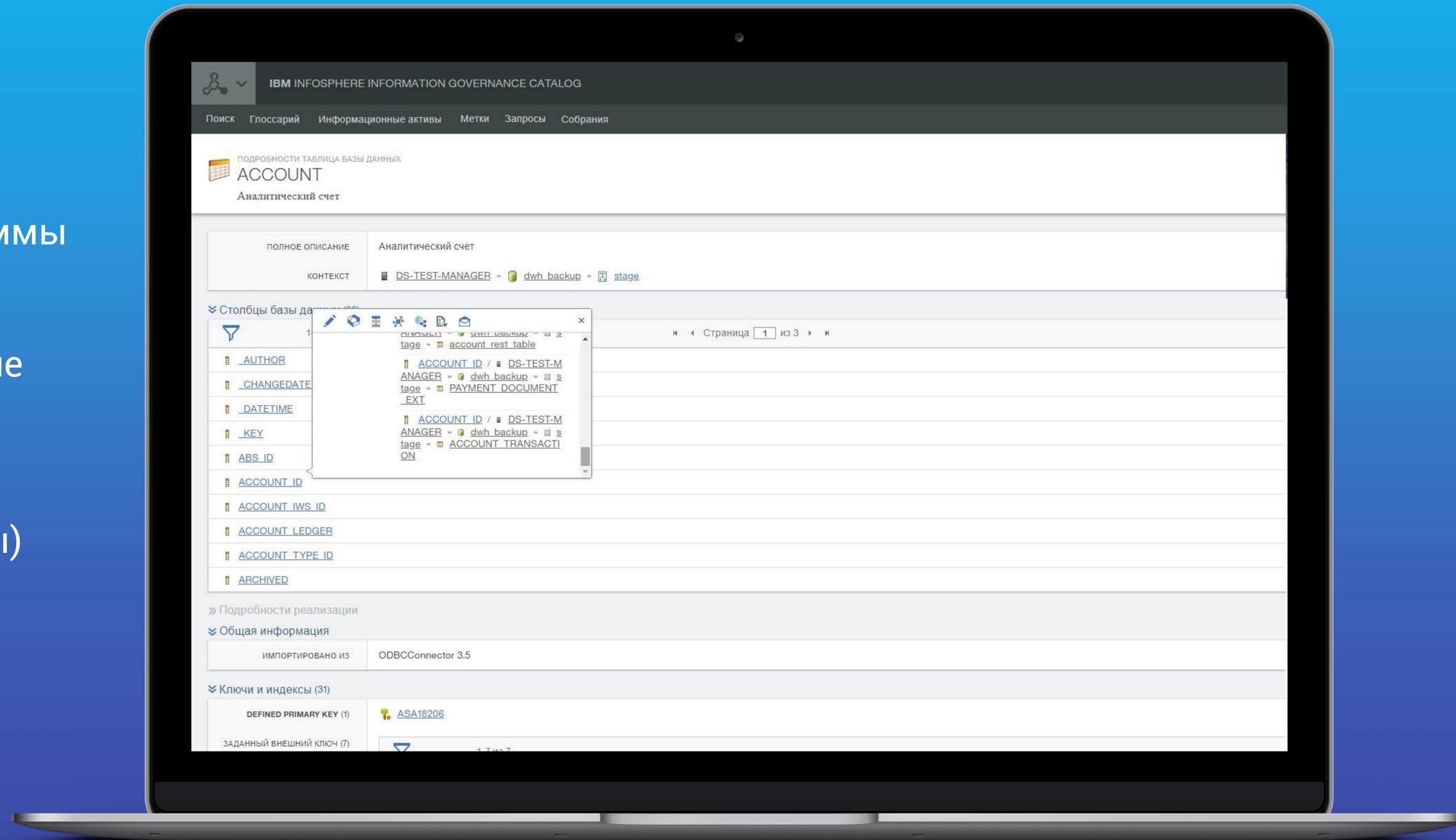
# Информационные активы

- Базы данных
- ETL-процедуры (программы преобразования)
- Логические и физические модели данных (CASE-модели)
- Приложения (программы)
- BI-приложения



# Информационные активы

- Базы данных
- ETL-процедуры (программы преобразования)
- Логические и физические модели данных (CASE-модели)
- Приложения (программы)
- BI-приложения



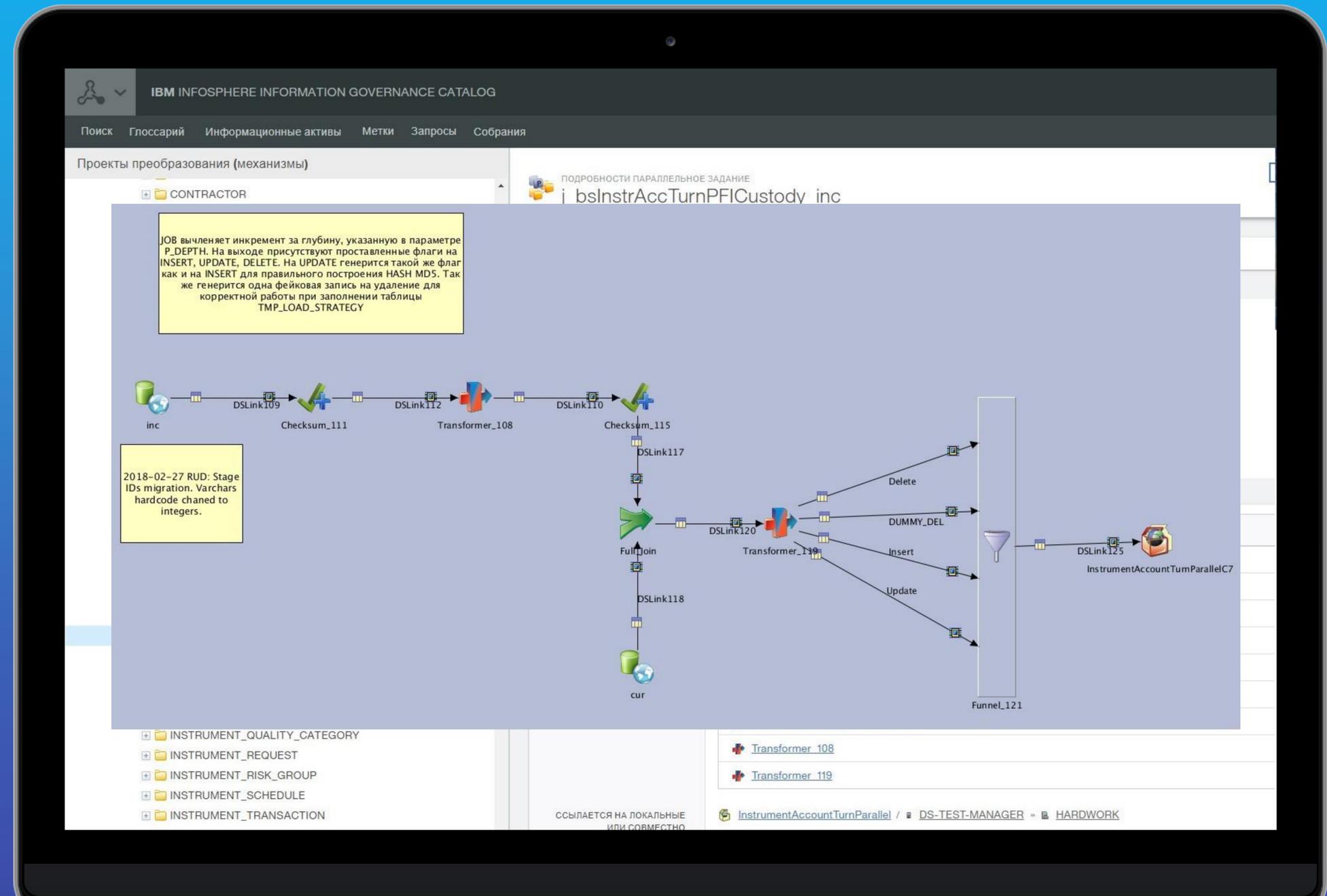
# Информационные активы

- Базы данных
- ETL-процедуры (программы преобразования)
- Логические и физические модели данных (CASE-модели)
- Приложения (программы)
- BI-приложения

The screenshot displays the IBM Information Governance Catalog interface. The main navigation bar includes 'Поиск', 'Глоссарий', 'Информационные активы', 'Метки', 'Запросы', and 'Собрания'. The left sidebar shows a tree view of 'Проекты преобразования (механизмы)' with various categories like CONTRACTOR, COUNTRY, and INSTRUMENT. The selected item is 'j\_bslnstrAccTurnPFICustody\_inc'. The main content area shows the details of this project, including a context menu with 'DS-TEST-MANAGER' and 'HARDWORK', a diagram titled 'Изображение' showing a flowchart, and a list of stages or containers titled 'Содержит этапы или контейнеры (10)'. The list includes items like 'Checksum\_111', 'Checksum\_115', 'cur', 'Full\_Join', 'Funnel\_121', 'inc', 'InstrumentAccountTurnParallelC7', 'Transformer\_108', and 'Transformer\_119'. The bottom of the interface indicates 'ССЫЛАЕТСЯ НА ЛОКАЛЬНЫЕ ИЛИ СОВМЕСТНО' and provides a breadcrumb trail: 'InstrumentAccountTurnParallel / DS-TEST-MANAGER / HARDWORK'.

# Информационные активы

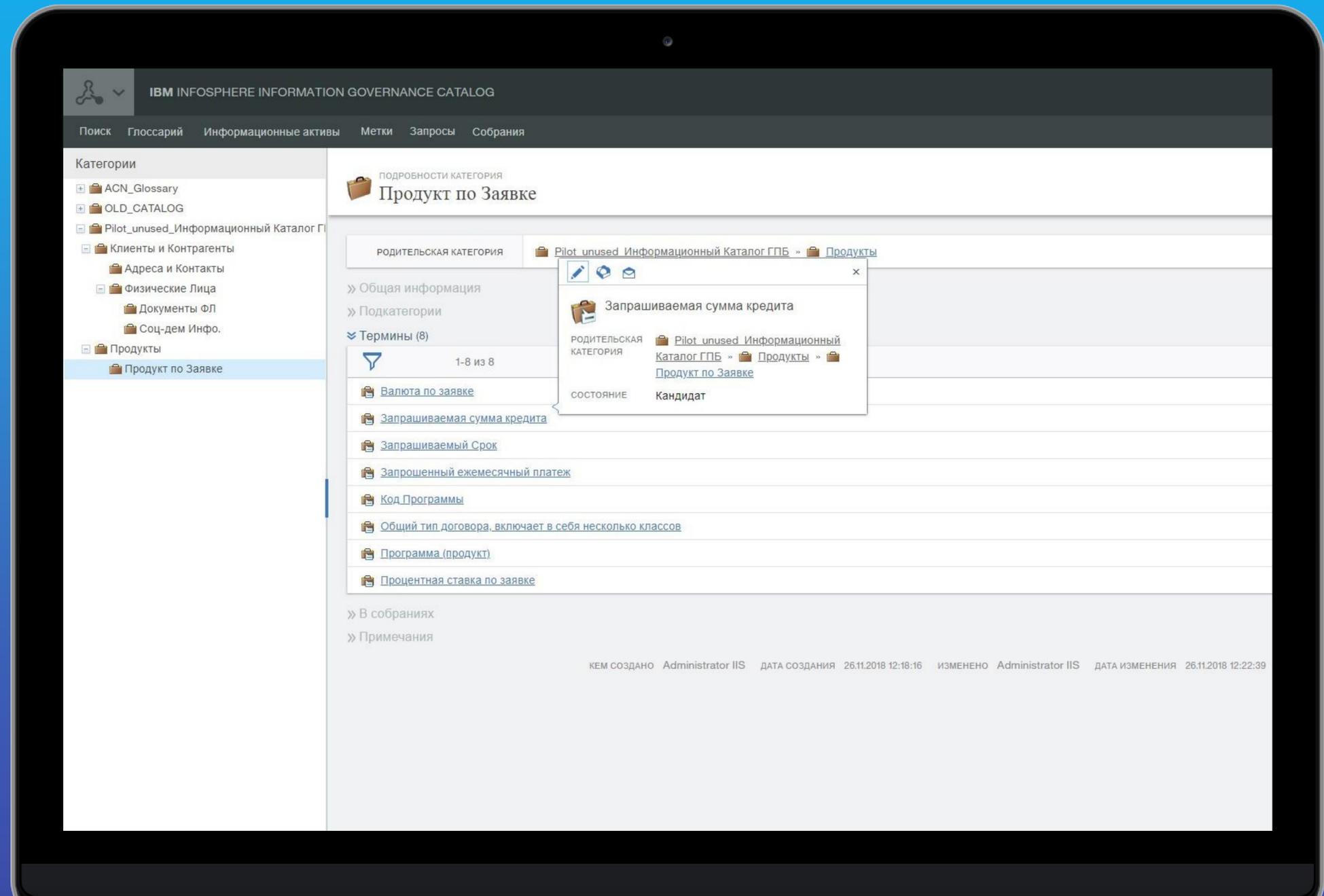
- Базы данных
- ETL-процедуры (программы преобразования)
- Логические и физические модели данных (CASE-модели)
- Приложения (программы)
- BI-приложения



# Происхождение данных

IBM Information Governance Catalog предоставляет удобный интерфейс для исследования и документирования потоков данных

- Определяет взаимосвязи между информационными активами и их происхождение:
  - от физического места хранения
  - через процедуры трансформации
  - до бизнес-термина
- Анализ влияния изменений



# Происхождение данных

IBM Information Governance Catalog предоставляет удобный интерфейс для исследования и документирования потоков данных

- Определяет взаимосвязи между информационными активами и их происхождение:
  - от физического места хранения
  - через процедуры трансформации
  - до бизнес-термина
- Анализ влияния изменений

The screenshot shows the IBM Information Governance Catalog interface. The main page is titled "Запрашиваемая сумма кредита" (Requested credit amount). It displays the following information:

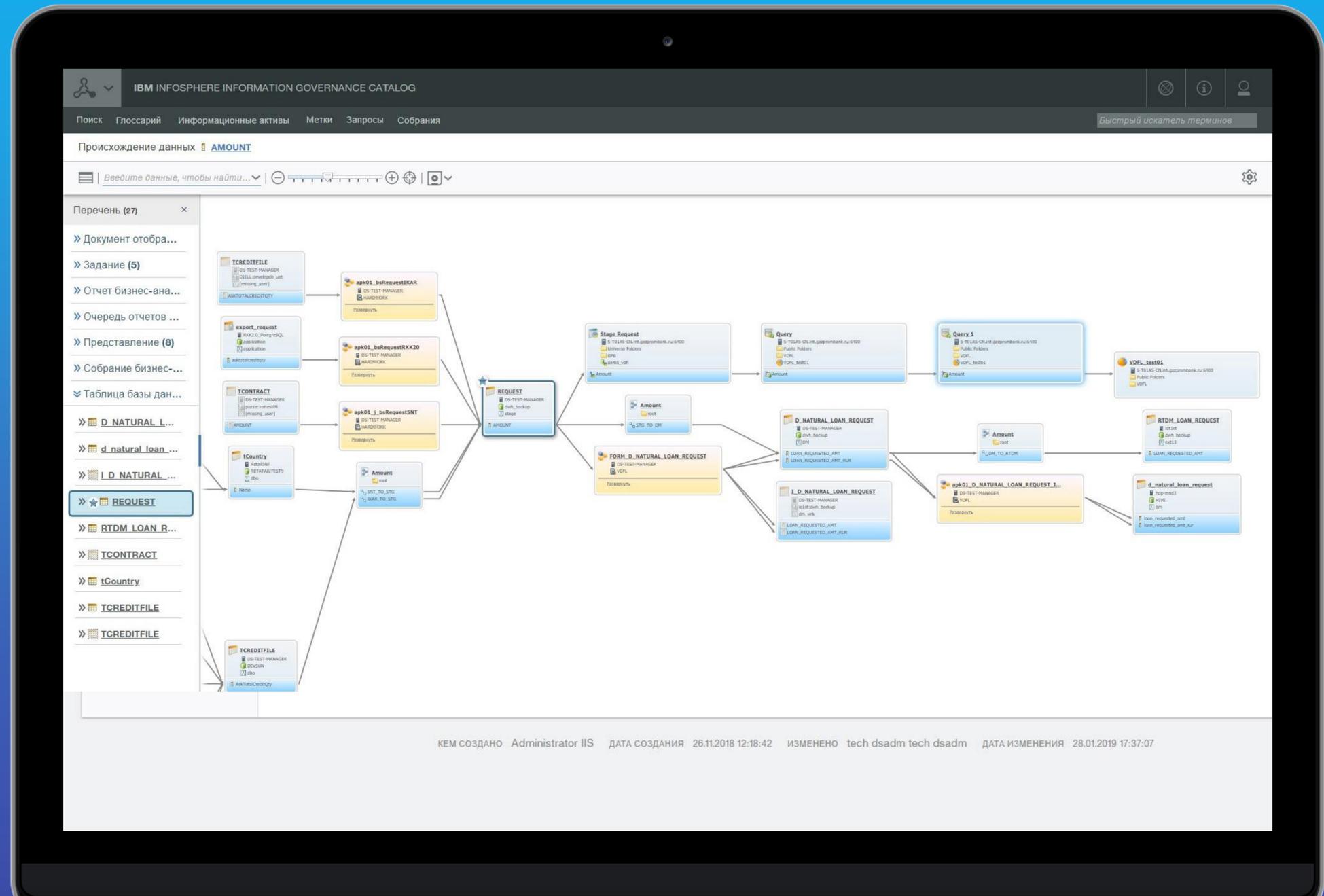
- Родительская категория:** Pilot\_unused Информационный Каталог ГПБ > Продукты > Продукт по Заявке
- Состояние:** Кандидат
- Общая информация:**
  - Модификатор: False
  - Тип: Нет
- Связанные термины:** (None listed)
- Назначенные активы (3):**
  - AMOUNT / DS-TEST-MANAGER > dwh\_backup > stage > REQUEST
  - amount / hdp-mnd3 > HIVE > v\_stage > request
  - AskTotalCreditQty / DS-TEST-MANAGER > DEVSUN > dbo > ICREDITFILE
- В собраниях:** (None listed)
- Примечания:** (None listed)
- Graphical View:** RELATIONSHIP GRAPH (Нет доступных данных)

At the bottom, it shows the creator information: КЕМ СОЗДАНО Administrator IIS, ДАТА СОЗДАНИЯ 26.11.2018 12:18:42, ИЗМЕНЕНО tech dsadm tech dsadm, ДАТА ИЗМЕНЕНИЯ 28.01.2019 17:37:07.

# Происхождение данных

IBM Information Governance Catalog предоставляет удобный интерфейс для исследования и документирования потоков данных

- Определяет взаимосвязи между информационными активами и их происхождение:
  - от физического места хранения
  - через процедуры трансформации
  - до бизнес-термина
- Анализ влияния изменений



# Использование IBM Information Governance Catalog

01

ПОДДЕРЖИВАТЬ  
ОТВЕТСТВЕННОСТЬ  
И ОТЧЕТНОСТЬ

- Связь между бизнес-метаданными и техническими метаданными для обеспечения соответствия
- Назначение ответственных за данные для обеспечения единой точки контактов

# Использование IBM Information Governance Catalog

01

ПОДДЕРЖИВАТЬ  
ОТВЕТСТВЕННОСТЬ  
И ОТЧЕТНОСТЬ

02

ПОВЫШАТЬ ДОСТУПНОСТЬ  
ИНФОРМАЦИИ

- Унификация и единый подход к бизнес-терминологии описания данных
- Подробное описание контекста информации
- Совместное использование аннотаций в отношении данных

# Использование IBM Information Governance Catalog

01

ПОДДЕРЖИВАТЬ  
ОТВЕТСТВЕННОСТЬ  
И ОТЧЕТНОСТЬ

02

ПОВЫШАТЬ ДОСТУПНОСТЬ  
ИНФОРМАЦИИ

03

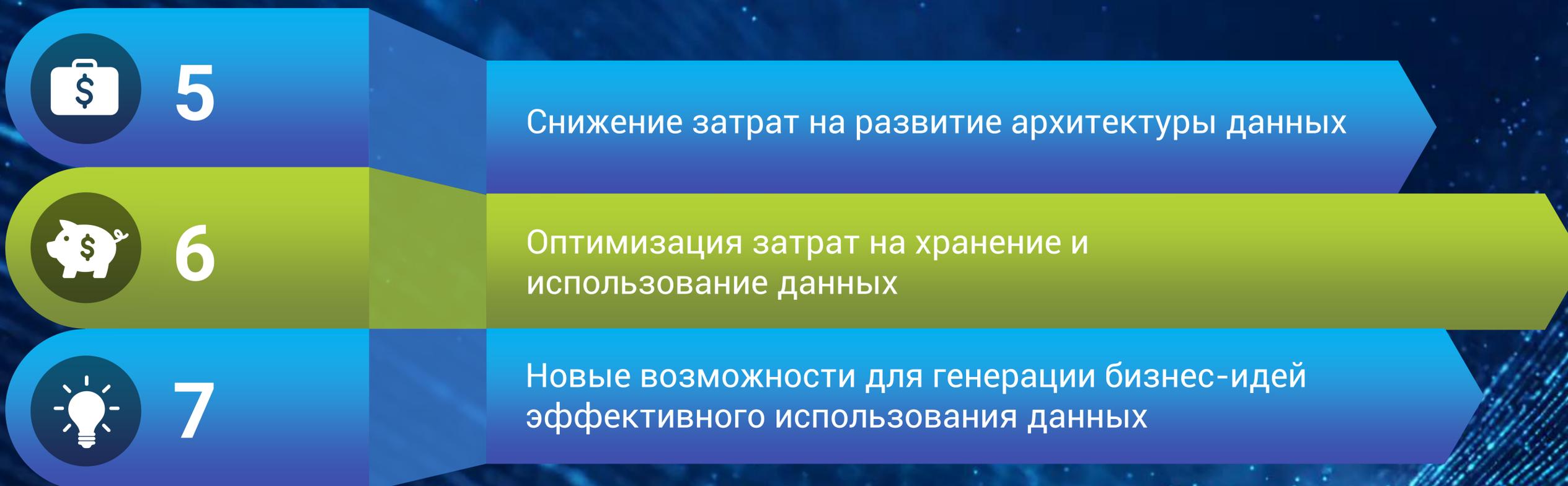
ОБЕСПЕЧИВАТЬ  
КОЛЛЕКТИВНУЮ РАБОТУ

- Оперативный доступ к корпоративной информации о данных
- Гибкая настройка инструмента на потребности пользователей
- Поддержка коллективного чата обсуждения и совместного использования бизнес-терминов

# Преимущества использования IBM Information Governance Catalog

-  **1** Прозрачность происхождения данных от источника информации до конечного потребителя
-  **2** Единое централизованное хранилище знаний обо всех данных организации
-  **3** Наглядное представление осуществляемых изменений
-  **4** Внедрение политик и правил управления качеством данных

# Преимущества использования IBM Information Governance Catalog





# История успеха

А что, так можно было?!...

# Крупная европейская розничная сеть

- 60 стран мира
- 40 тыс. человек персонала
- Розничная торговля и финансовые услуги



Развитие бизнеса

Взрывной рост объёмов данных



Беспрецедентные проблемы в аналитике и отчётности

# Беспрецедентные проблемы в аналитике и отчётности

## ЦЕЛИ КОМПАНИИ

Способность предложить клиентам правильные товары в правильное время

## ПРОБЛЕМЫ

- Поддержкой сотен ETL-сценариев
- Качество, управление и происхождение данных
- Внесение изменений в исходные данные

## ОТДЕЛЬНЫЕ ЗАДАЧИ ЗАНИМАЛИ СЛИШКОМ МНОГО ВРЕМЕНИ

Анализ влияния изменений на IT-ландшафт занимала до 3-х недель с негарантированным результатом



Проблемы с аналитикой

01

Правильные товары в правильное время

02

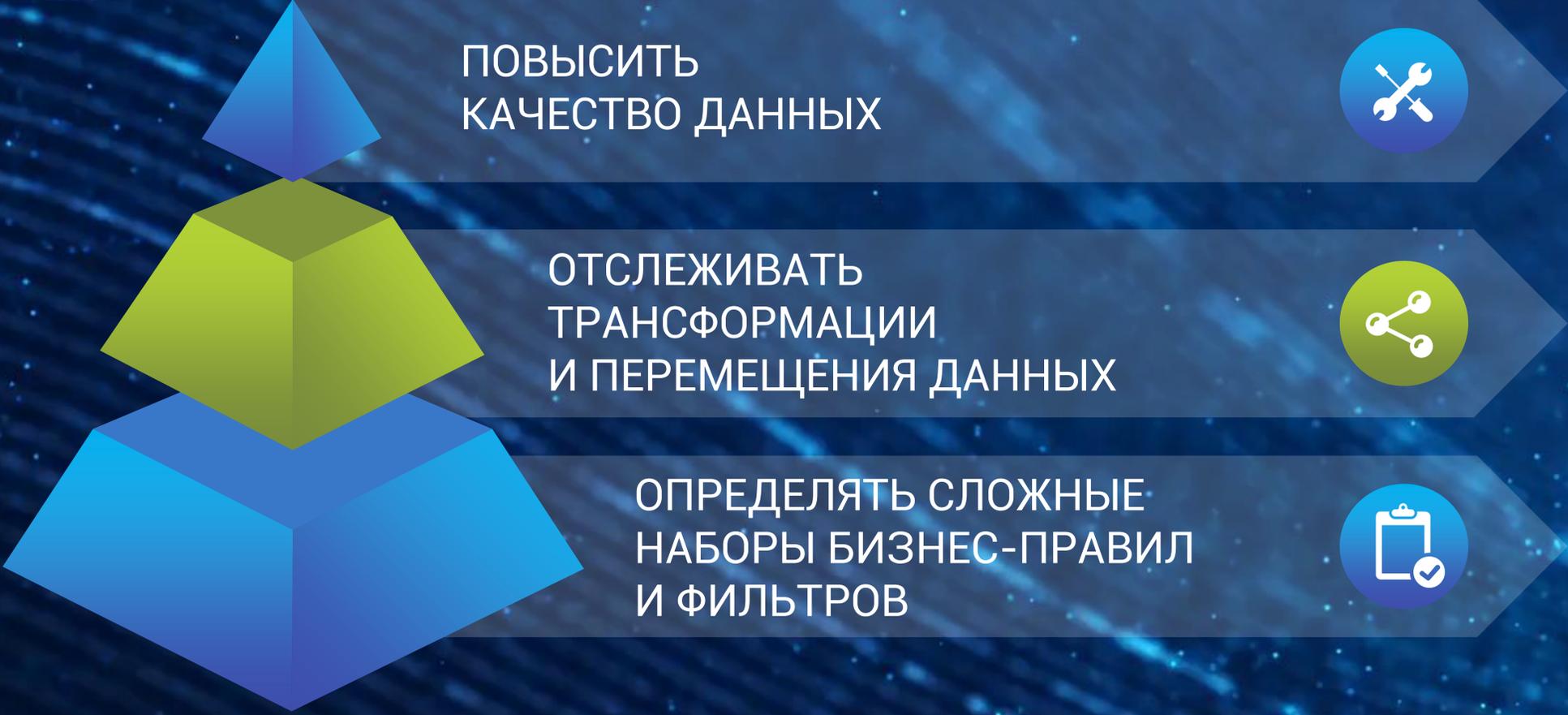
Сотни ETL-процессов  
Качество данных

03

Отдельные задачи занимали слишком много времени

# Внедрение процессов управления данными

Для решения возникших перед компанией задач необходимо было выбрать инструмент, позволяющий:



ПОВЫСИТЬ  
КАЧЕСТВО ДАННЫХ



ОТСЛЕЖИВАТЬ  
ТРАНСФОРМАЦИИ  
И ПЕРЕМЕЩЕНИЯ ДАННЫХ



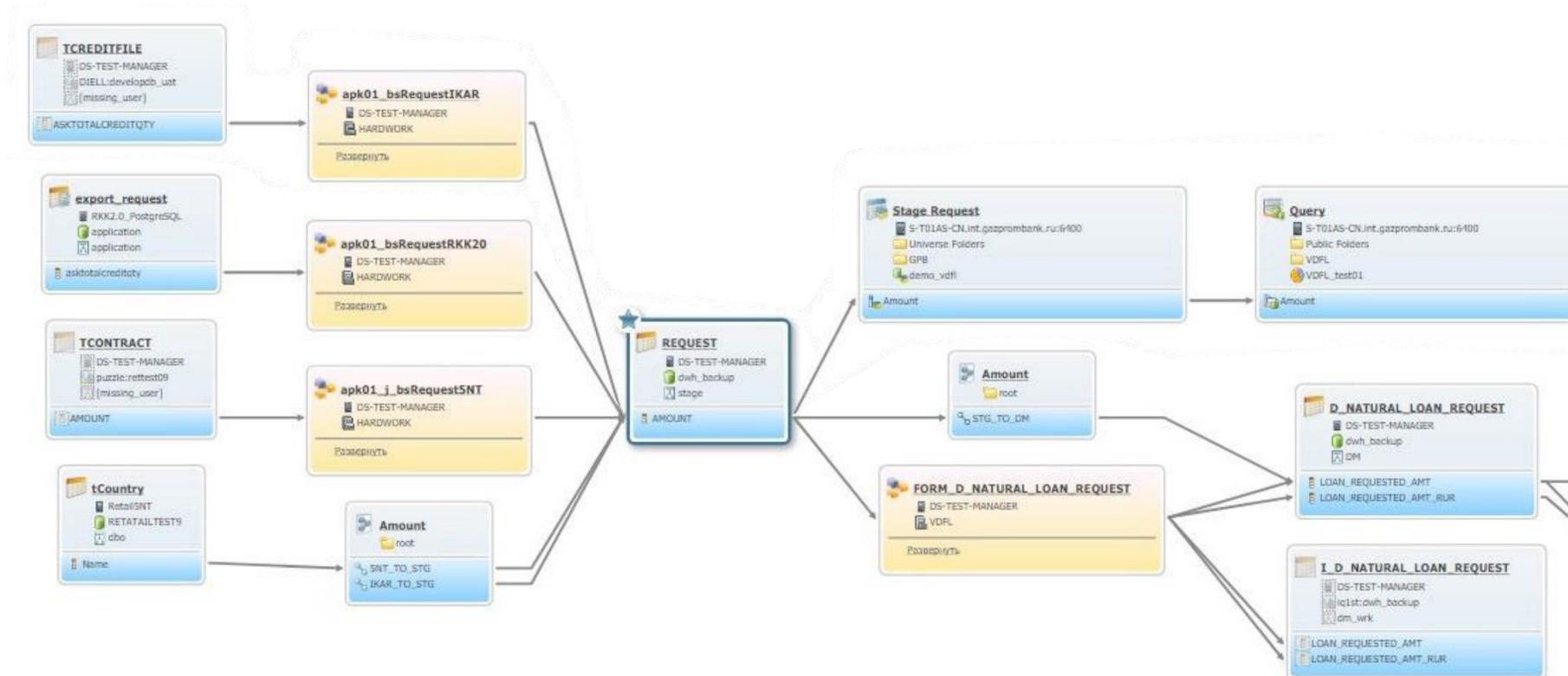
ОПРЕДЕЛЯТЬ СЛОЖНЫЕ  
НАБОРЫ БИЗНЕС-ПРАВИЛ  
И ФИЛЬТРОВ



Все эти задачи позволил  
решить

**IBM Information  
Governance Catalog**

# Результаты внедрения процессов управления данными



В результате внедрения процессов управления данными посредством IBM InfoGuard в ресурсе DataGuard стало возможным решать

- обеспечить выполнение задач - именованных процедурных отчётов
- обеспечить большую гибкость при выполнении операций или интеграции с изменениями архитектуры данных
- осуществлять оценку влияния производимых изменений на IT-ландшафт за минуты



# Спасибо за внимание

Удачи