

Моделирование данных в информационной архитектуре

Константин Валеев, февраль 2020



Привет!

Руководжу аналитиками и проектами в «Ростелеком ИТ» — софтверном IT-кластере, разрабатывающем цифровые продукты «Ростелекома»: системы видеонаблюдения, интерактивное телевидение, инфраструктурные и веб-проекты.

Про что будет рассказ

1. Какие бывают модели данных и зачем они нужны
2. Онтологии в философии и в IT
3. Инструменты и нотации для работы с онтологиями
4. От онтологии к объектной модели
5. Значение объектной модели в дизайне систем и интерфейсов

Модель данных

Модель данных



Реальность



Сознание



Артефакт



Сознание

Модель данных

- Нужна для общего понимания предметной области
- Формализация отдельных фактов
- Обеспечивает основу для проектирования

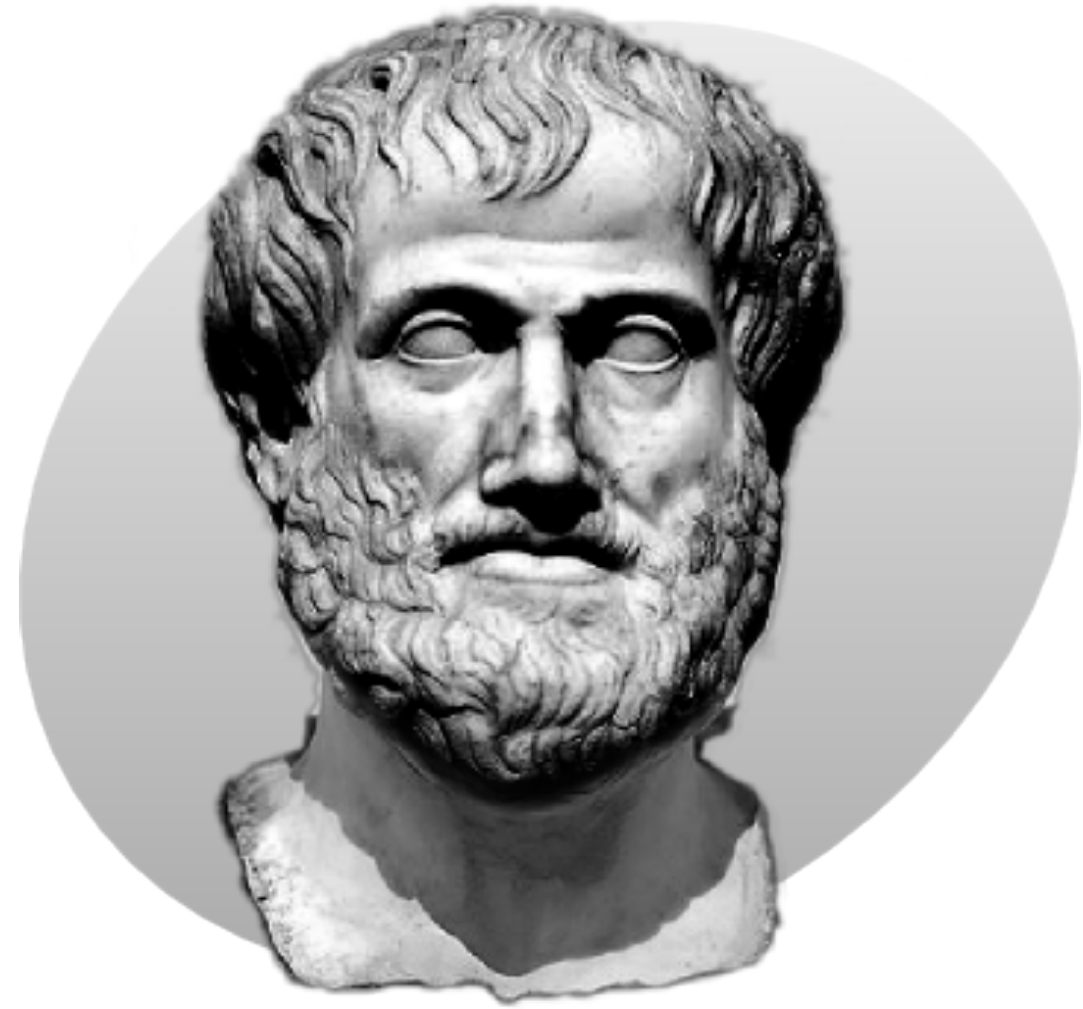
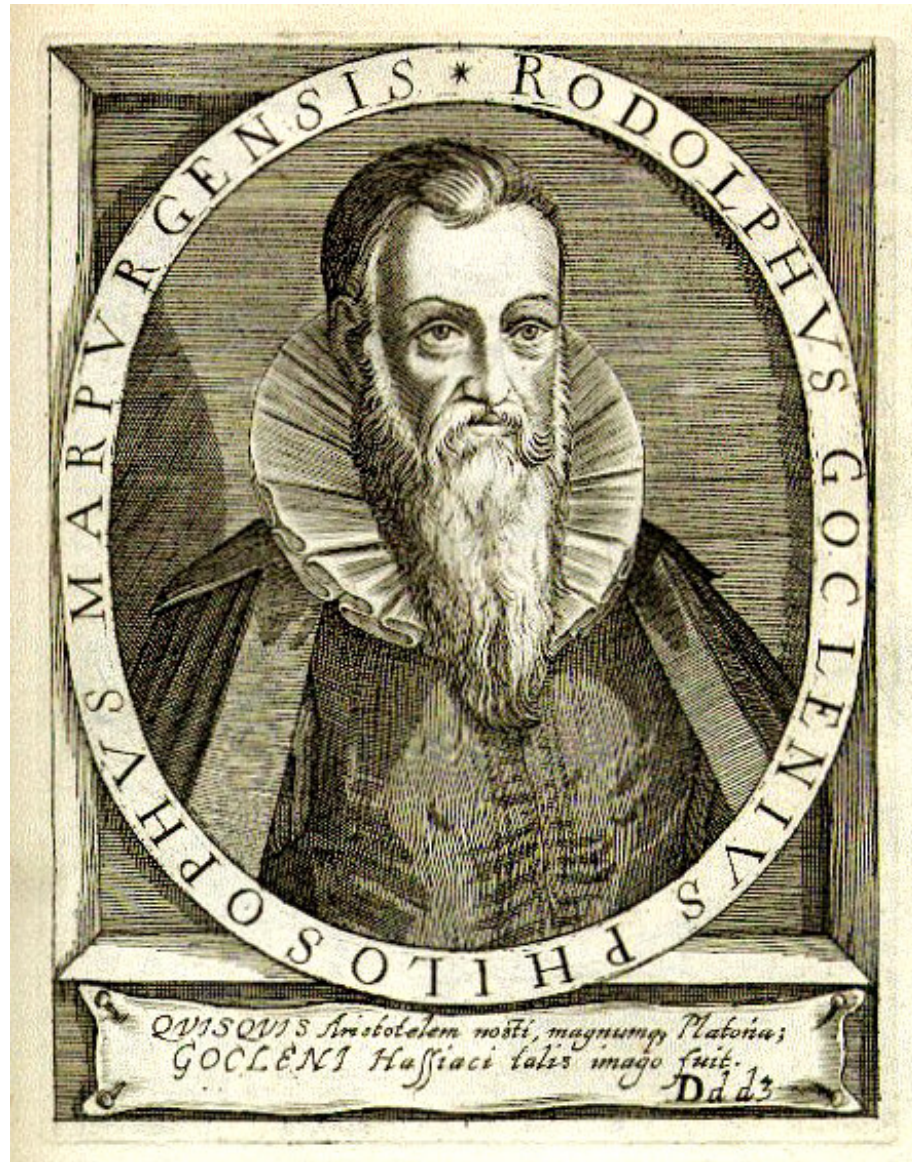
Передача смысла ~~без~~ с минимумом ошибок

Типы моделей данных

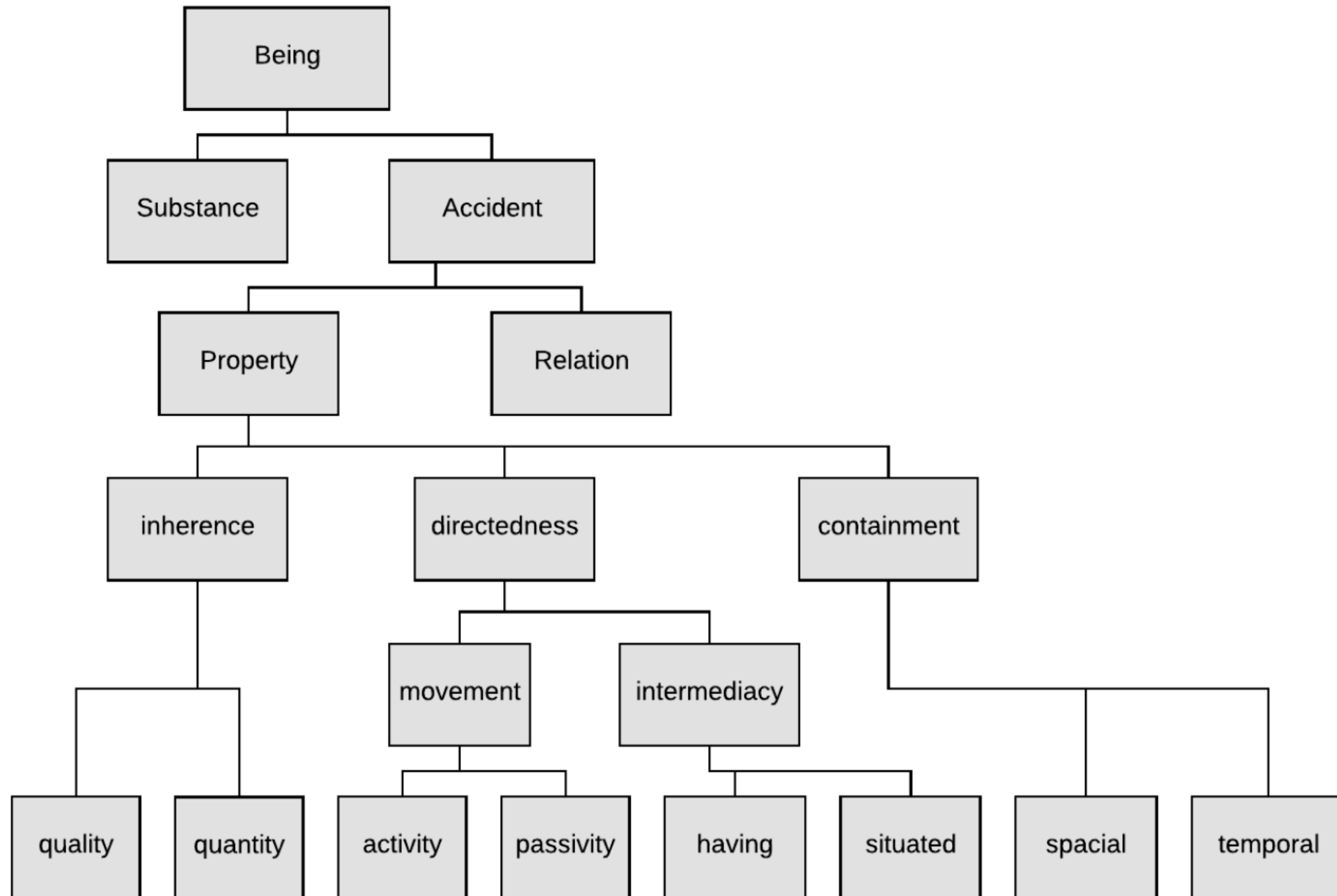
- Концептуальные
- Логические
- Физические

ОНТОЛОГИЯ

Ontology is the philosophical study of being... concerning what entities exist or may be said to exist and how such entities may be grouped, related within a hierarchy, and subdivided according to similarities and differences.



Denomination	Greek	Question	Example
Substance	ουσία (ousia)	What is something?	Man, horse
Quantity	ποσον (poson)	How much/big is something?	Two inches long
Qualities	ποιον (poion)	What are the features?	White, able to read
Relation	προσ τι (pros ti)	In what relationship is something (to something)?	double, half, bigger, daughter of, was born in
Location	που (pou)	Where is something?	On the table, in the swimming pool
Time	ποτέ (pote)	When is something?	Yesterday, in the future
Position/ orientation	κεισθηαι (keisthai)	What orientation does something have?	standing, sitting
Having	ἔχθειν (echein)	What does something have?	Has shoes on his feet, is armed
Doing	ποιεῖν (poiein)	What does something do?	Cut, burn
Experience	πάσχειν (paschein)	What experiences something?	Being cut, being burned



Формальная онтология

Philosophy	Computer Science
Substance	Class, entity, concept
Accidents	Attribute, property
Facts	Fact, statement, proposition, relation
Ontology	Base-ontology, foundational ontology

Термины,

важные в рамках предметной области...

Заказ

Статус

Поисковый запрос

Избранное

Корзина

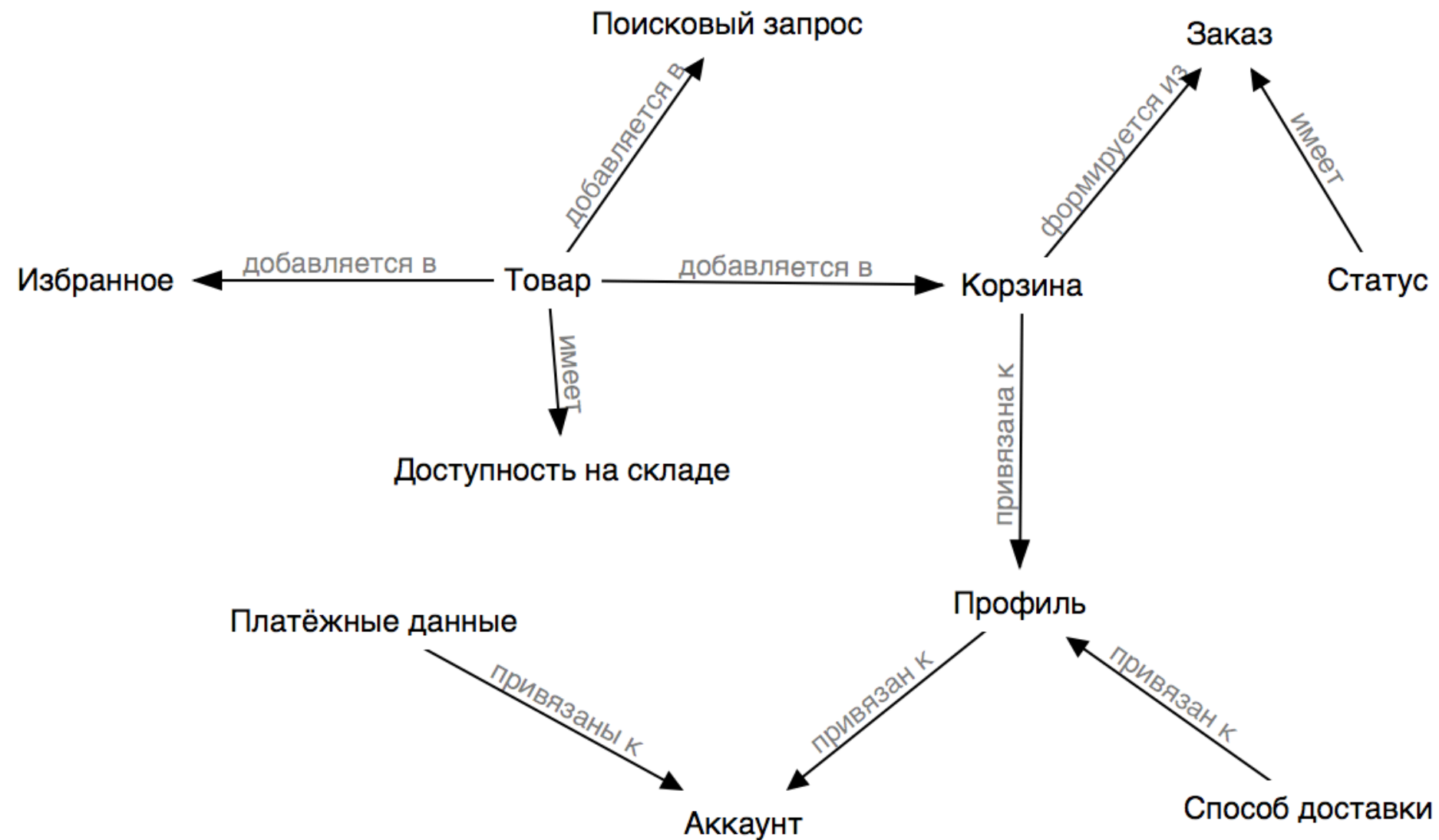
Способ доставки

Профиль

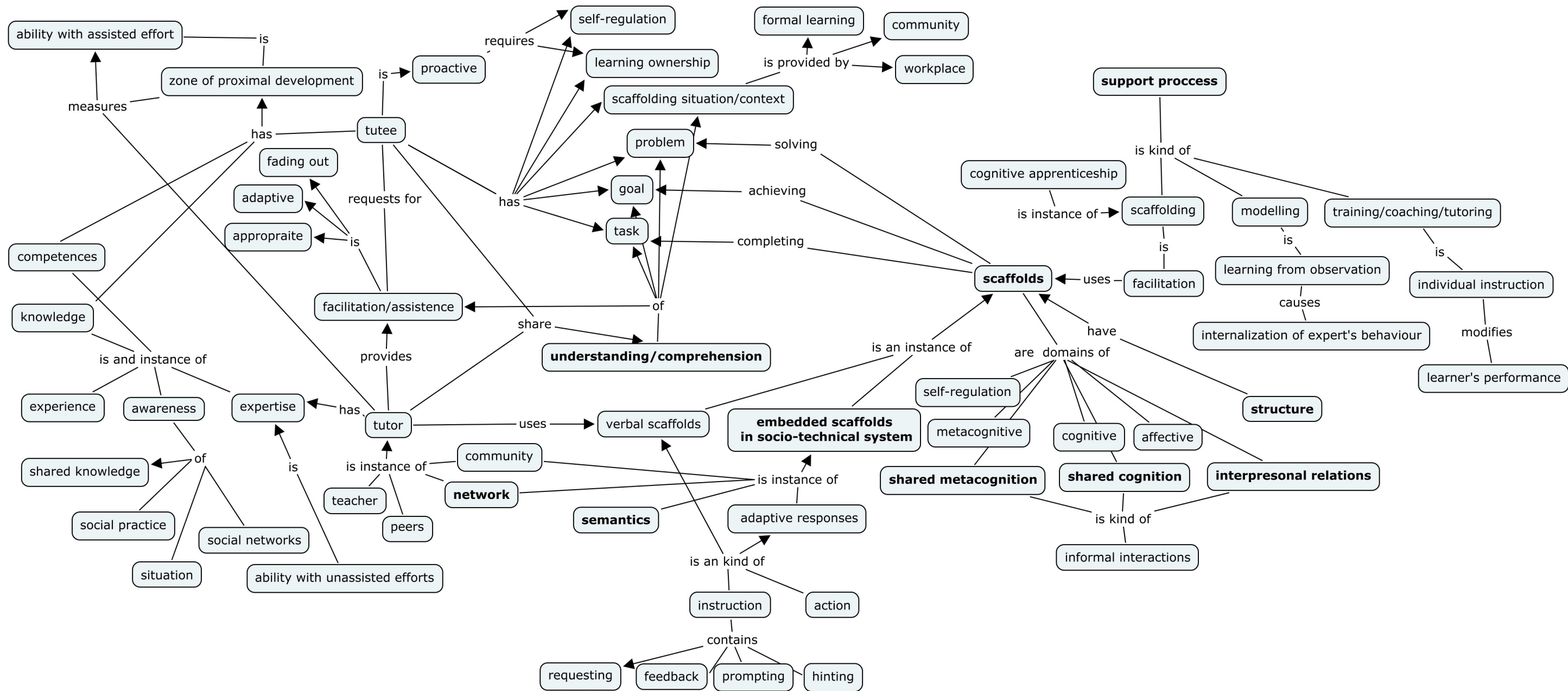
Аккаунт

Платёжные данные

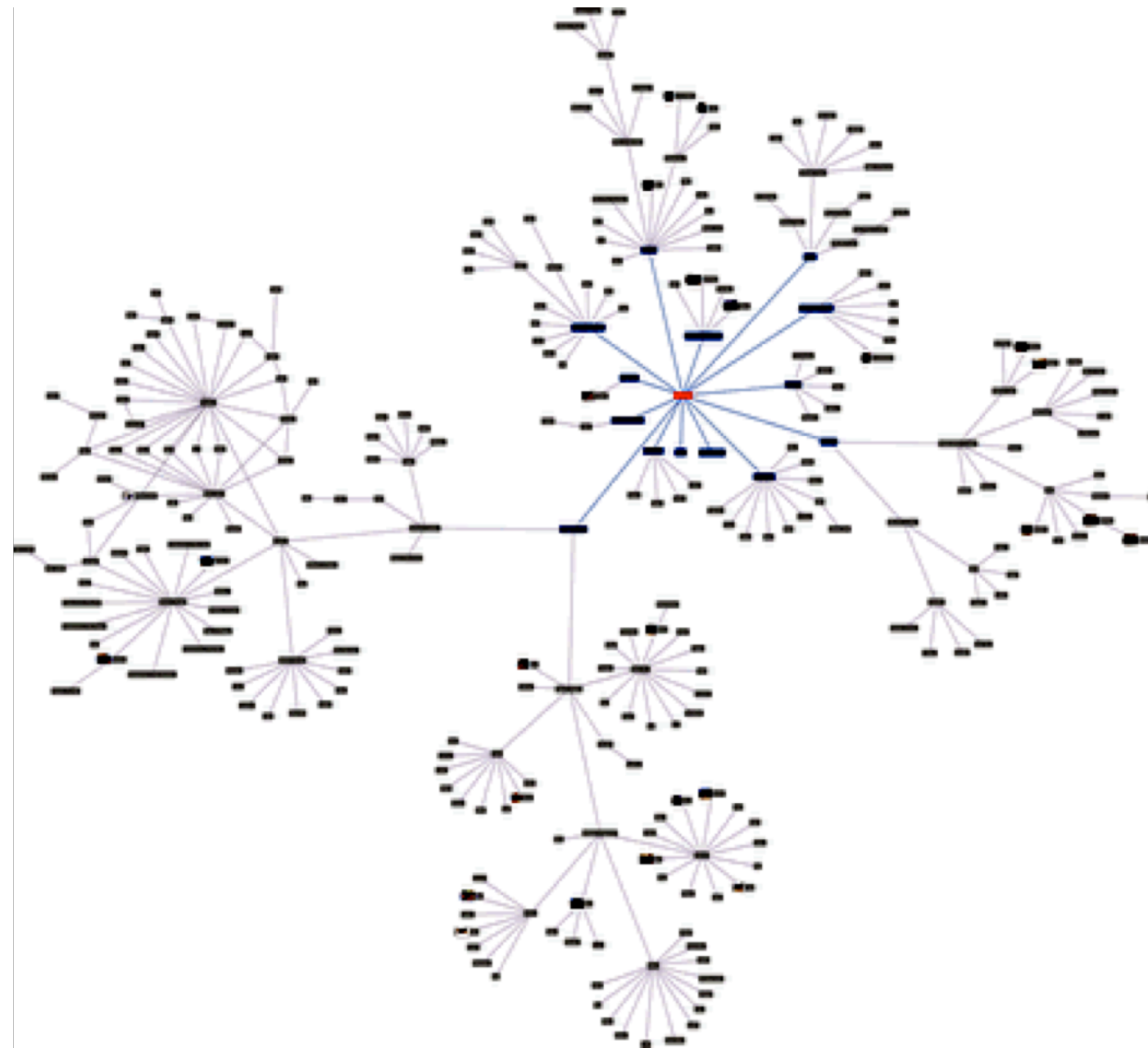
...И СВЯЗИ МЕЖДУ НИМИ



Много терминов и связей



Очень много терминов и связей



Уровни онтологий

- Метаонтология / верхняя онтология
- Онтология предметной области (домена)
- Онтология рассматриваемой системы

Уровни онтологий

- **Метаонтология / верхняя онтология**
- Онтология предметной области (домена)
- Онтология рассматриваемой системы

**Basic Formal Ontology
(BFO)**

Top-level ontology to promote interoperability of domain ontologies

34 categories, 8 relations

**Business Objects Reference Ontology
(BORO)**

Top ontology as well as a process for constructing the ontology

Three main categories, 4D constructs

**Unified Foundational Ontology
(UFO)**

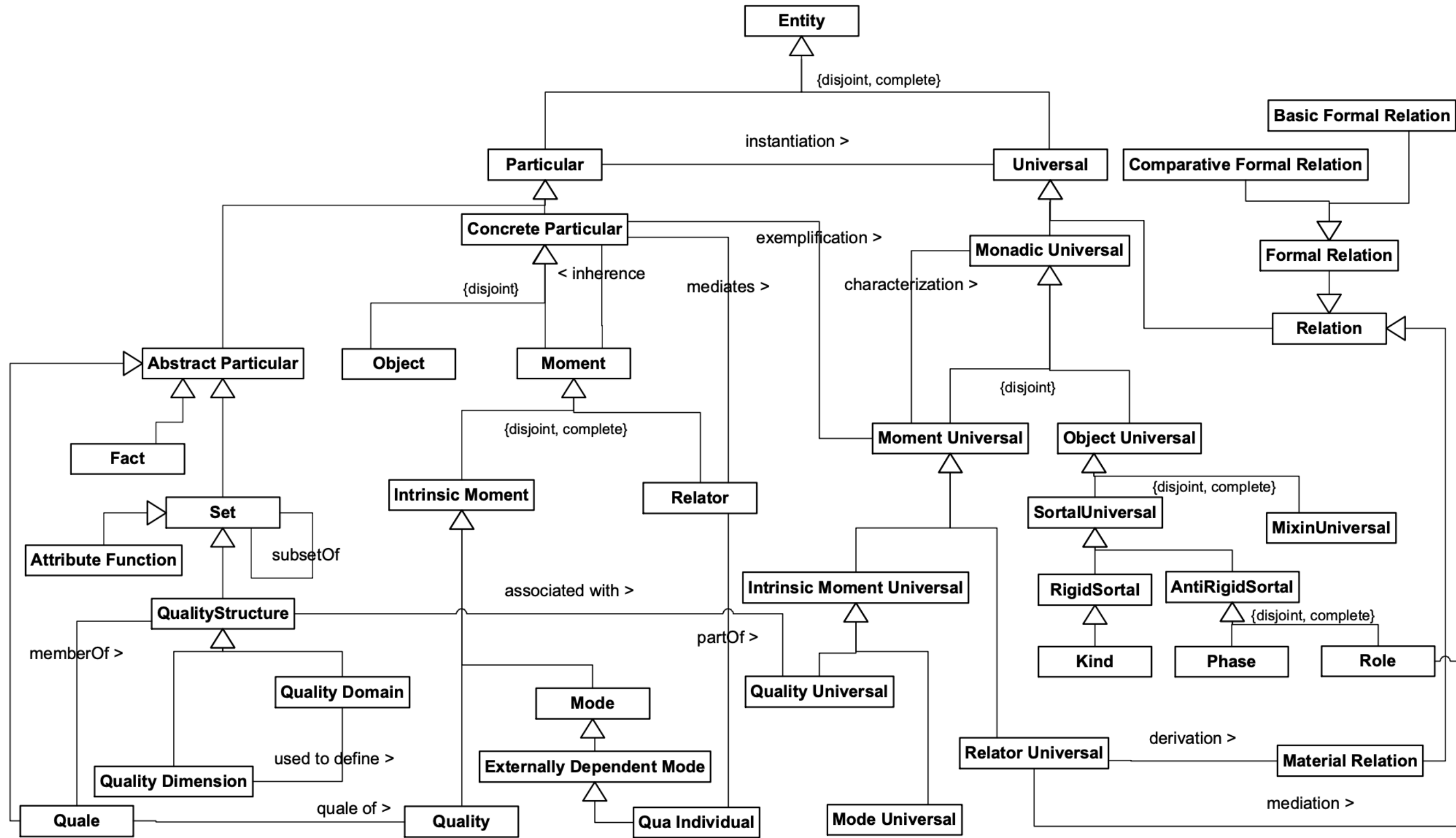
Top-level ontology as basis for domain ontologies

Three core modules, 3D constructs

Cyc

Everyday common sense knowledge, with the goal of enabling AI applications to perform human-like reasoning

12,000 WordNet synsets



Верхняя онтология

- Уходит в философию и метафизические понятия
- Невозможно договориться об аксиоматике
- Сложна для понимания
- Практическая польза зачастую неочевидна

Уровни онтологий

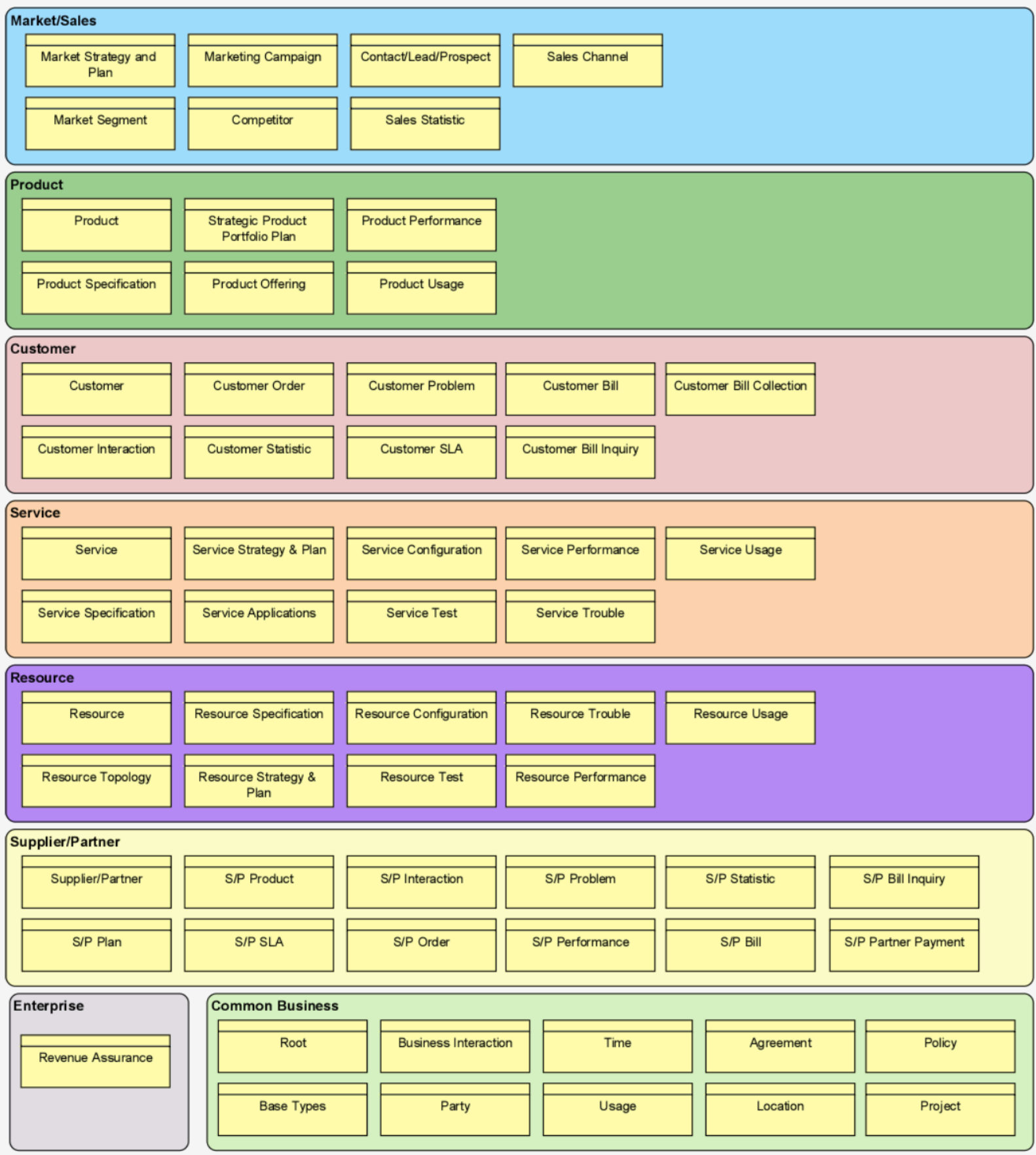
- Метаонтология / верхняя онтология
- **Онтология предметной области (домена)**
- Онтология рассматриваемой системы

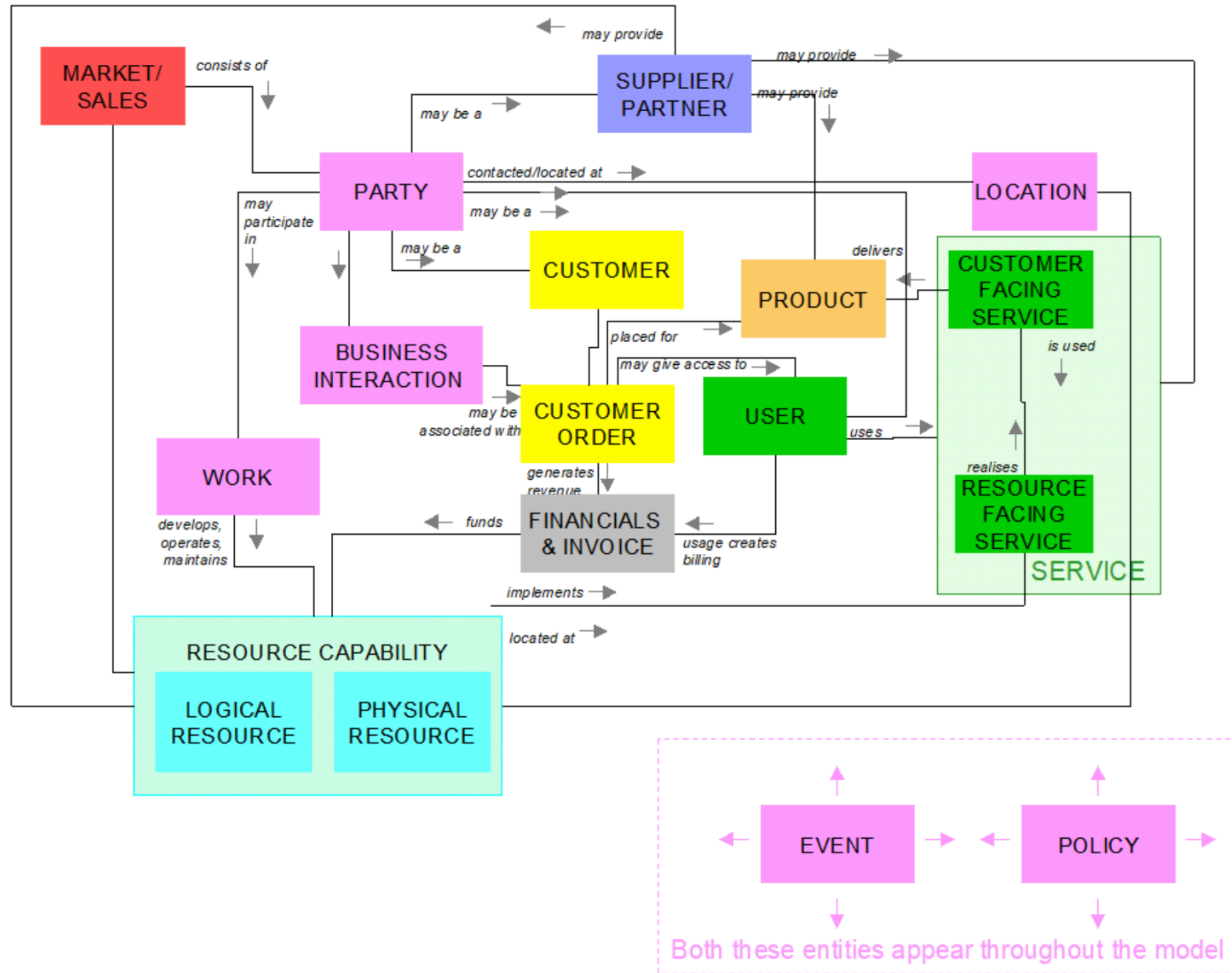
Отраслевые онтологии

- The Organization Ontology — бизнес-организации
- The Music Ontology — музыкальная индустрия
- OMG: MOF, ODM, UML — моделирование систем и ПО
- HL7 FHIR — здравоохранение
- Disease Ontology (DO) — медицина
- ISO 15926 — нефтегаз
- Gene Ontology (GO) — генетика
- TM Forum SID — телеком

Отраслевые онтологии

- The Organization Ontology — бизнес-организации
- The Music Ontology — музыкальная индустрия
- OMG: MOF, ODM, UML — моделирование систем и ПО
- HL7 FHIR — здравоохранение
- Disease Ontology (DO) — медицина
- ISO 15926 — нефтегаз
- Gene Ontology (GO) — генетика
- **TM Forum SID — телеком**





```
1 {
2   "version": "2.0",
3   "ProductOffering": [{
4     "id": 400955271890,
5     "instId": "400955271890-1571294344971-1266803410550",
6     "name": "Конвергент_Для безлимита везде_основное (B2C Юг)",
7     "version": "1.0",
8     "isAdditional": false,
9     "isBundle": true,
10    "isSellable": true,
11    "bundledProductOffering": [{
12      "id": 400955298890,
13      "instId": "400955298890-1571294344992-380760235130",
14      "name": "Конвергент_ШПД_Для безлимита везде (B2C Юг)",
15      "version": "1.0",
16      "isAdditional": true,
17      "isSellable": false,
18      "products": [{
19        "id": 3146475471,
20        "codeName": "SHPD",
21        "name": "Проводной интернет"
22      }],
23      "properties": [{
24        "propertyReference": {
25          "id": 201808909173,
26          "codeName": "DISTRIBUTION_CHANNEL",
27          "isList": true,
28          "propertyType": ["TECH", "COMM"],
29          "reference": "DISTRIBUTION_CHANNELS",
30          "valueType": "STRING"
31        },
32        "values": [{
33          "id": 3146686203,
34          "valueCode": "48",
```


Доменные онтологии

- Основа для собственных онтологий
- Общие термины и смыслы
- Упрощает обмен данными между системами

Формальный язык для онтологий

RDF

```
<?xml version="1.0" encoding="utf-8"?>
<rdf:RDF xmlns:contact="http://www.w3.org/2000/10/swap/pim/contact#"
xmlns:eric="http://www.w3.org/People/EM/contact#"
xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#">
  <rdf:Description rdf:about="http://www.w3.org/People/EM/contact#me">
    <contact:fullName>Eric Miller</contact:fullName>
  </rdf:Description>
  <rdf:Description rdf:about="http://www.w3.org/People/EM/contact#me">
    <contact:mailbox rdf:resource="mailto:e.miller123(at)example"/>
  </rdf:Description>
  <rdf:Description rdf:about="http://www.w3.org/People/EM/contact#me">
    <contact:personalTitle>Dr.</contact:personalTitle>
  </rdf:Description>
  <rdf:Description rdf:about="http://www.w3.org/People/EM/contact#me">
    <rdf:type rdf:resource="http://www.w3.org/2000/10/swap/pim/contact#Person"/>
  </rdf:Description>
</rdf:RDF>
```

RDF

```
@prefix eric:    <http://www.w3.org/People/EM/contact#> .
@prefix contact: <http://www.w3.org/2000/10/swap/pim/contact#> .
@prefix rdf:     <http://www.w3.org/1999/02/22-rdf-syntax-ns#> .

eric:me contact:fullName "Eric Miller" .
eric:me contact:mailbox  <mailto:e.miller123(at)example> .
eric:me contact:personalTitle "Dr." .
eric:me rdf:type contact:Person .
```

wine (http://www.w3.org/TR/2003/PR-owl-guide-20031209/wine) : [/Users/user/Downloads/wine.rdf]

wine (http://www.w3.org/TR/2003/PR-owl-guide-20031209/wine) Search...

Active Ontology x Entities x Individuals by class x DL Query x OntoGraf x

Class hierarchy: WhiteNonSweetWine Asserted

Search: contains Search Clear

No Reasoner set. Select a reasoner from the Reasoner menu Show Inferences

Запросы: SPARQL

```
PREFIX ex: <http://example.com/exampleOntology#>
SELECT ?capital
      ?country
WHERE
{
  ?x ex:cityname      ?capital ;
     ex:isCapitalOf  ?y .
  ?y ex:countryname  ?country ;
     ex:isInContinent ex:Africa .
}
```

Правила: Semantic Web Rule Language (SWRL)

$$\text{hasParent}(\text{?x1}, \text{?x2}) \wedge \text{hasBrother}(\text{?x2}, \text{?x3}) \Rightarrow \text{hasUncle}(\text{?x1}, \text{?x3})$$



Grakn

define

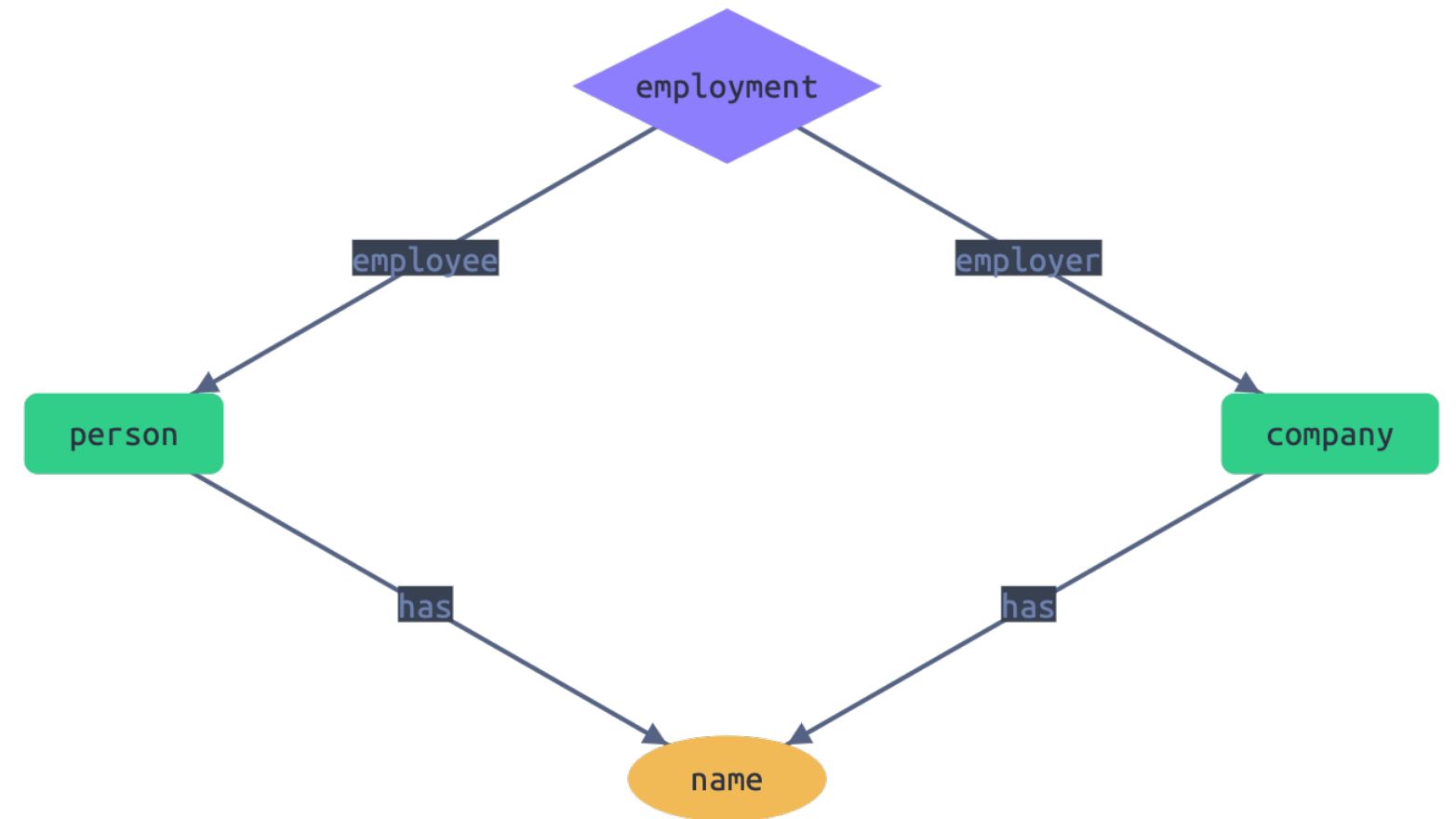
```
person sub entity,  
  has name,  
  plays employee;
```

```
company sub entity,  
  has name,  
  plays employer;
```

```
employment sub relation,  
  relates employee,  
  relates employer;
```

```
name sub attribute,  
  datatype string;
```

commit



Grakn

```
define
```

```
transitive-location sub rule,
```

```
when {
```

```
  (located: $x, locating: $y);
```

```
  (located: $y, locating: $z);
```

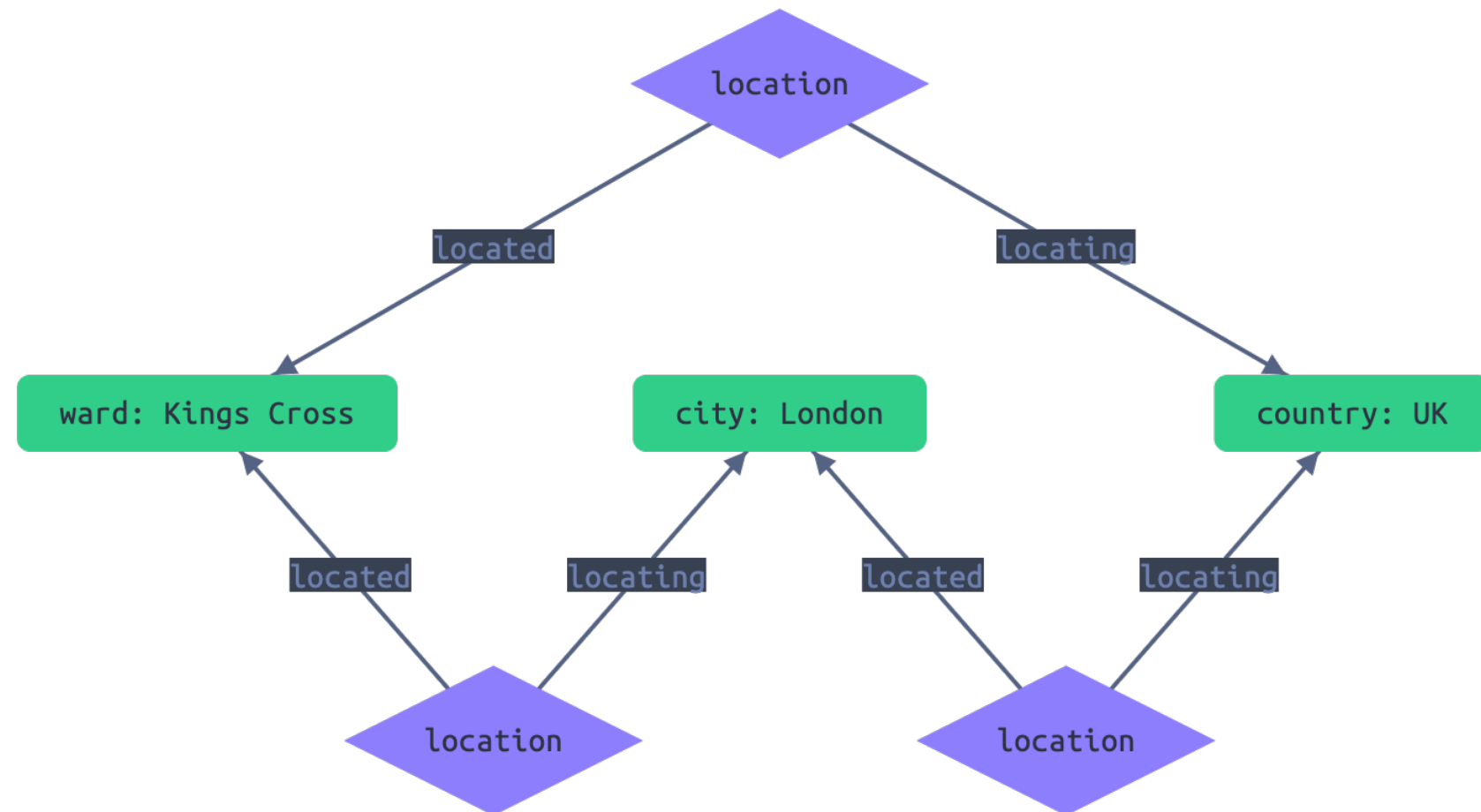
```
},
```

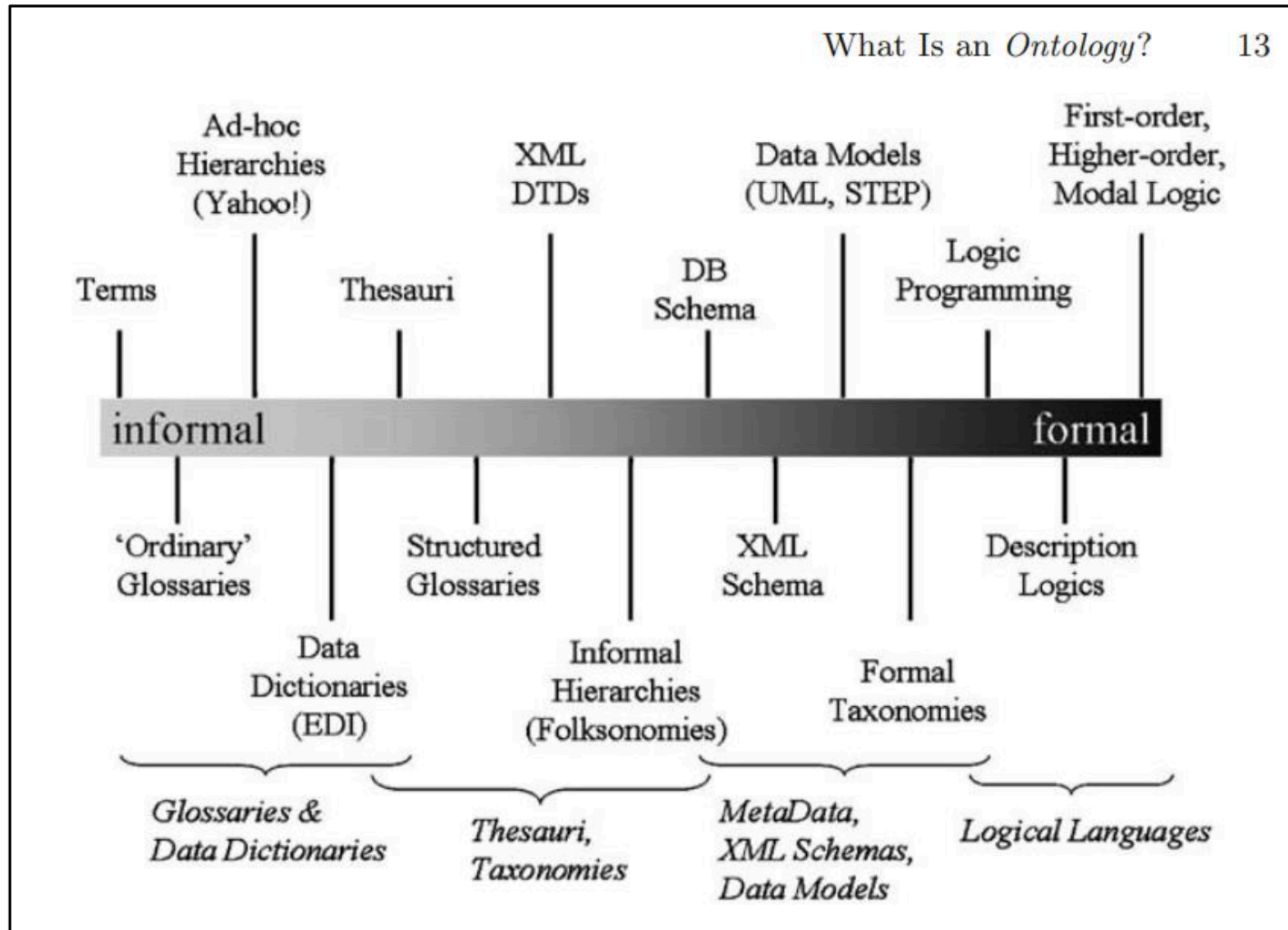
```
then {
```

```
  (located: $x, locating: $z);
```

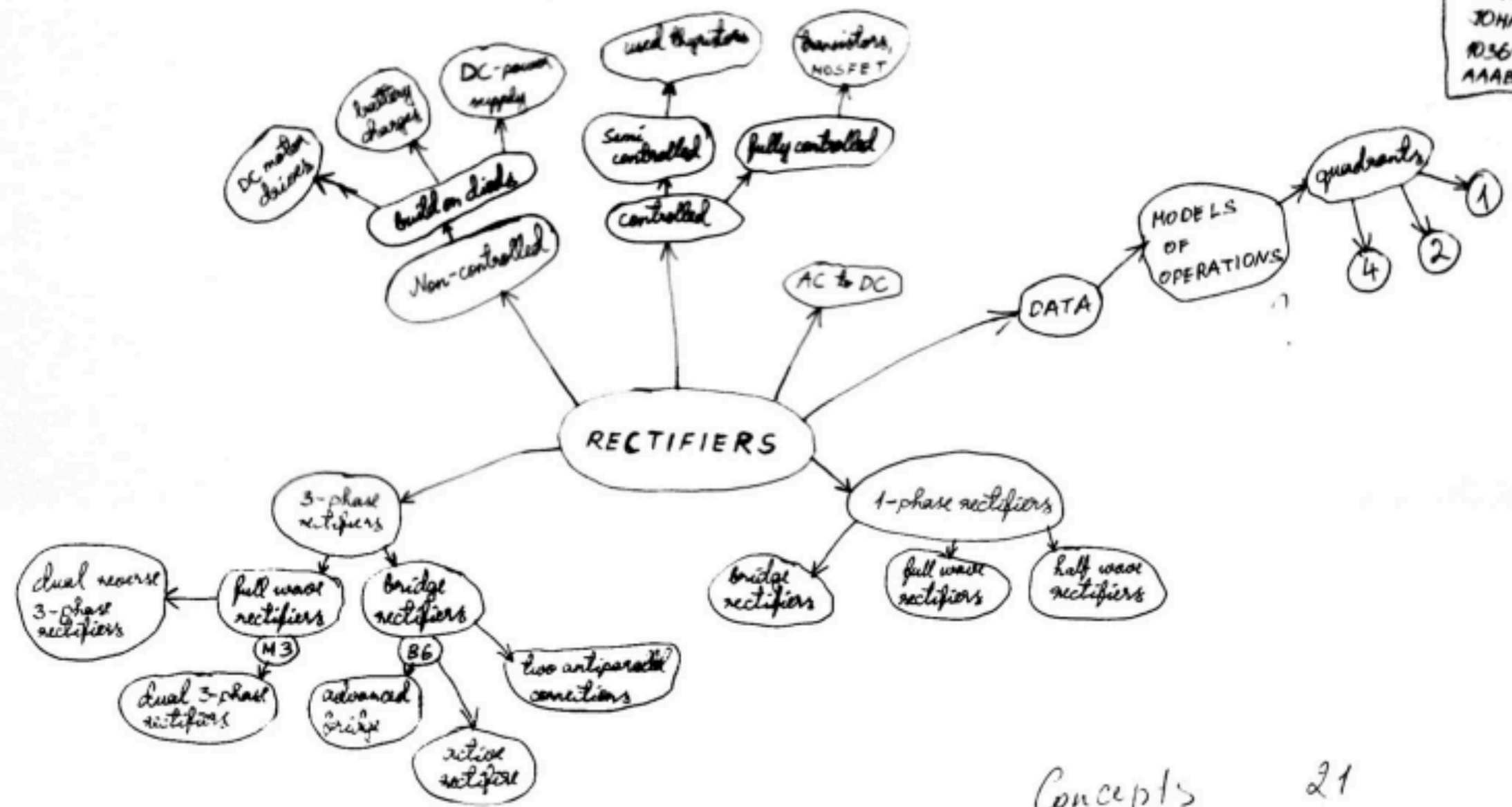
```
};
```

```
commit
```





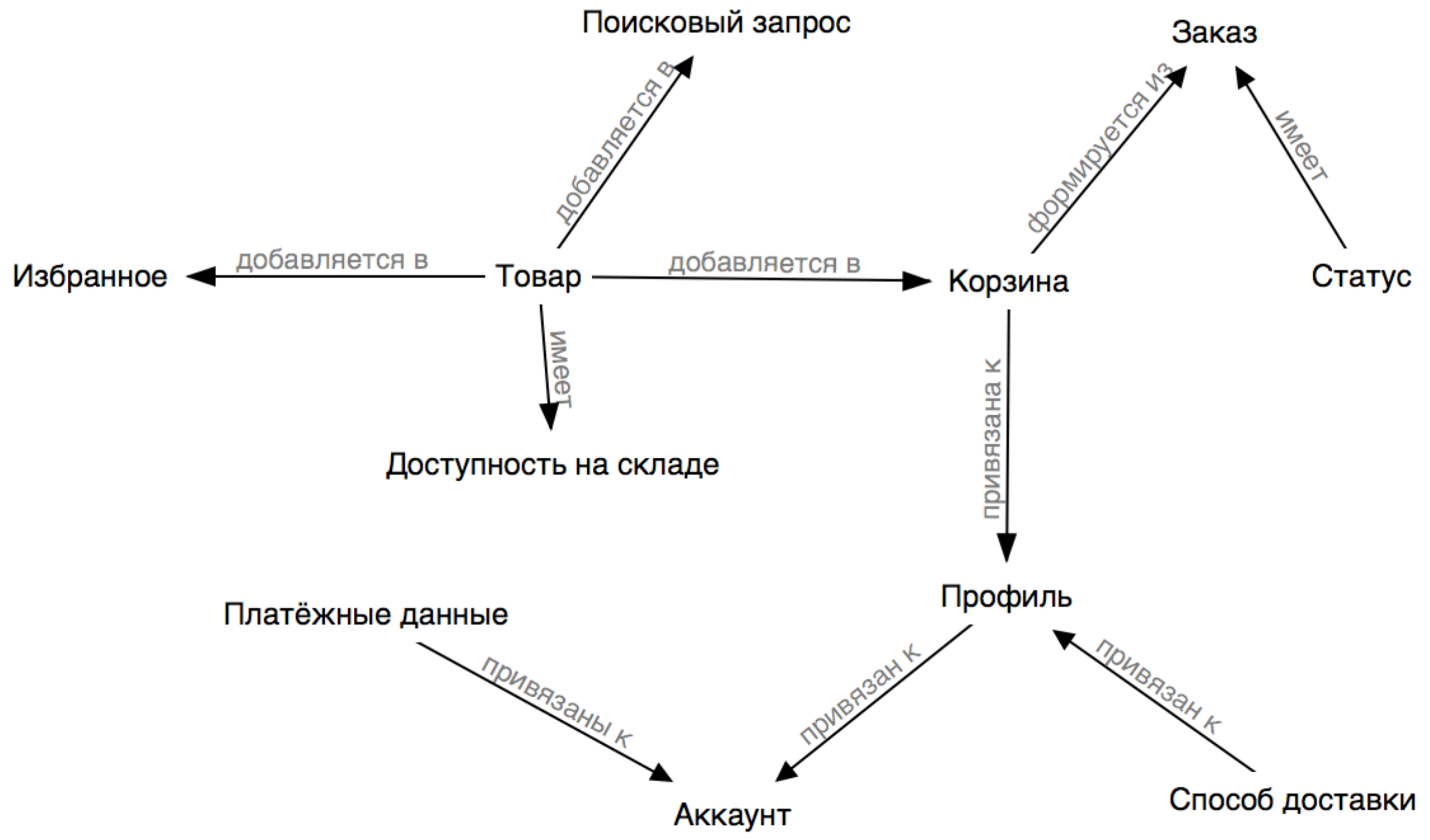
RAGNAR
JOHANSON
R3607
AAAB.51

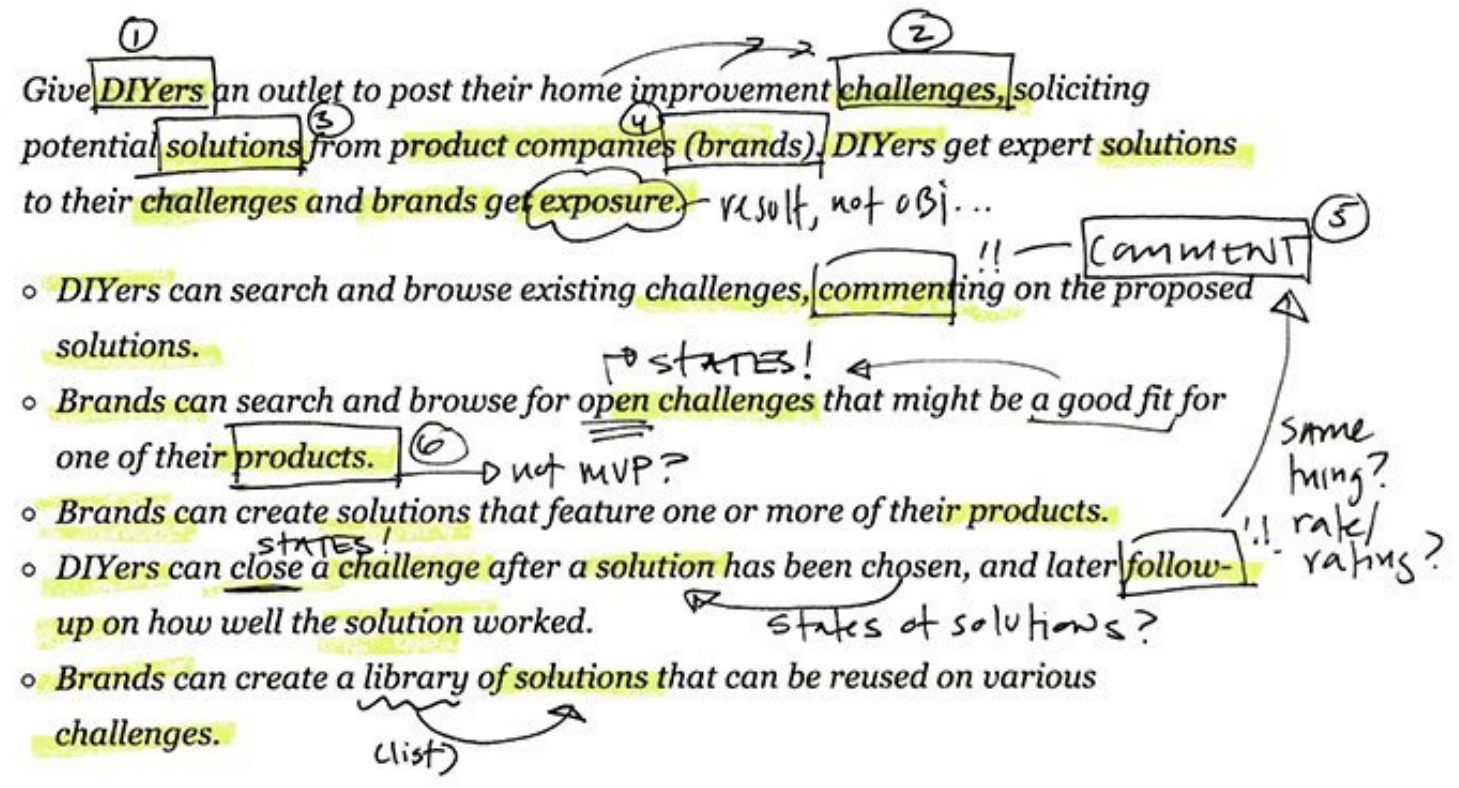


Concepts 21
 Linking verbs 0
 Max hierarchical levels 3

 Total 24

От терминов к сущностям





Noun in Combination with a Genitive

Examples: the date of the order, the diameter of the circle, the color of the car

Sentence Construction with: <class> has <attribute>

Example: a person has a date of birth; an address has a postal code; the process has a transition time of...

Adjective in Combination with a Noun

Example: a fast car; a large display; a huge bank account; a red car; a black list

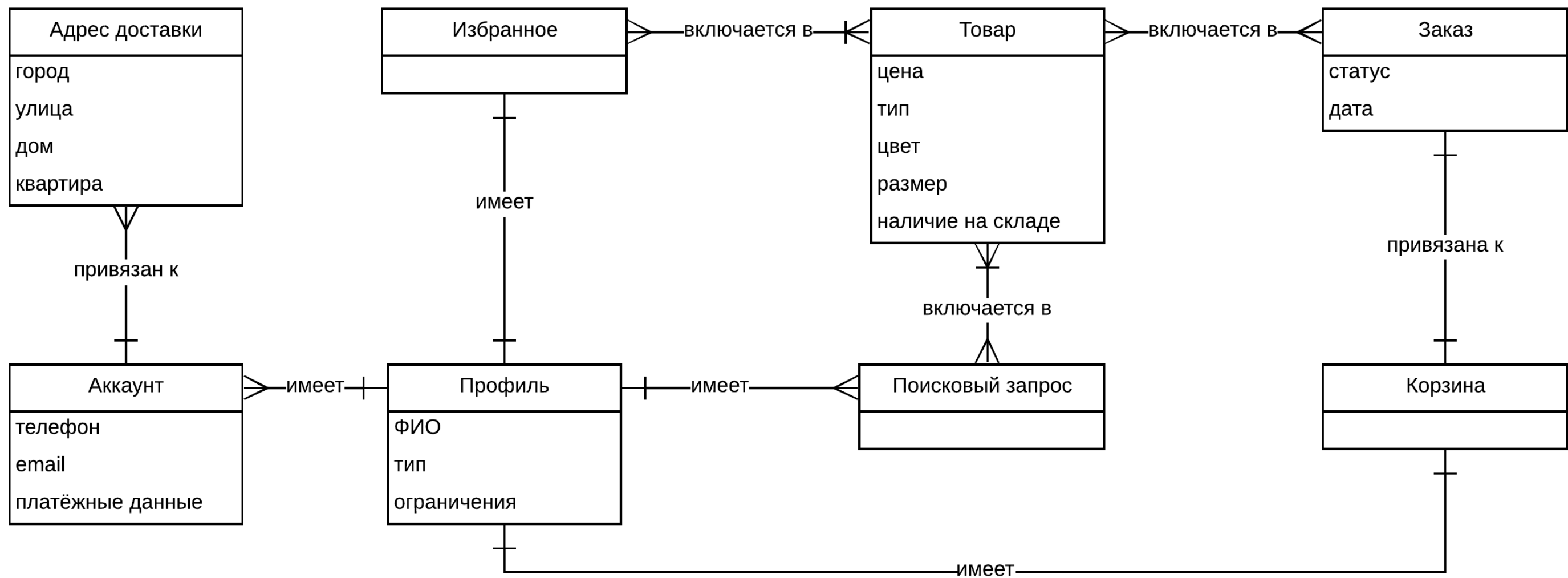
Sentence Structures with: <class> is <attribute value>

Example: If the person is an adult; if the application is approved; ...

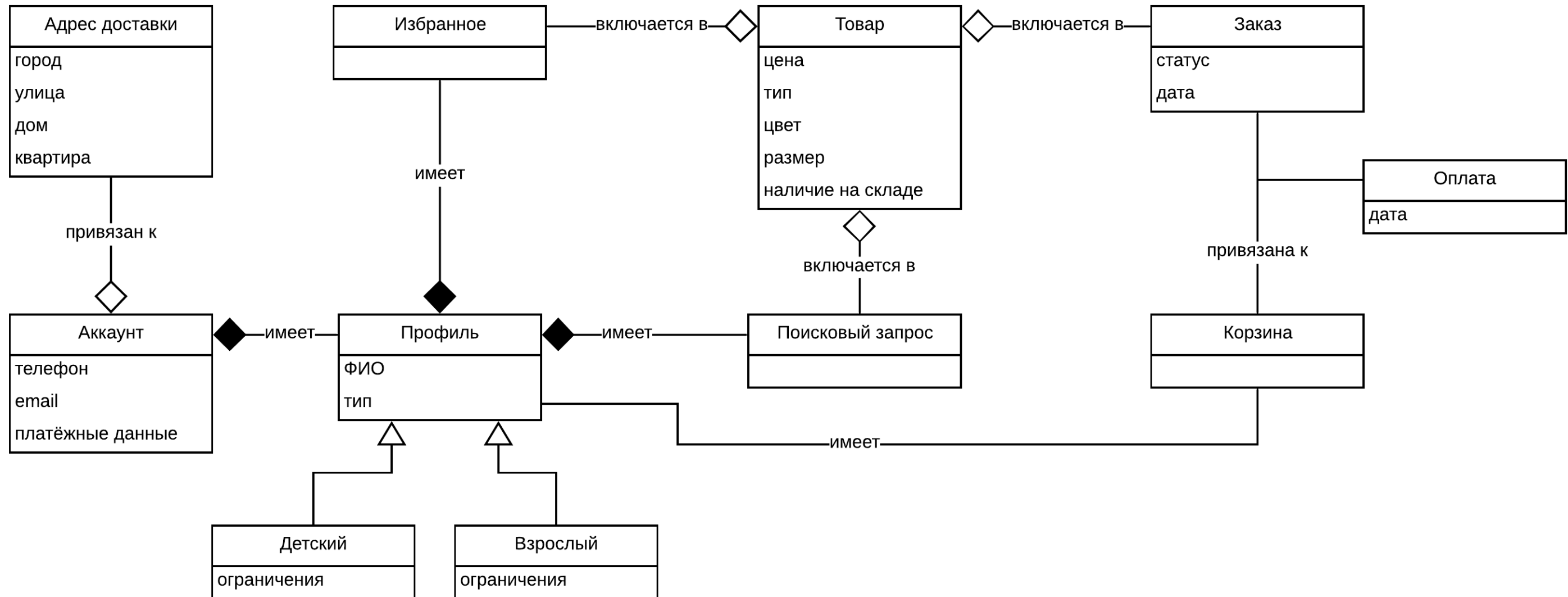
Differentiating Objects

Example: Modeling the object Peter Schulz with only two attributes (first name, last name) may not be sufficient to distinguish it from another person with the same name. If the class person also has the date of birth as an attribute, its objects may be clearly distinguishable (i.e., another person with the same name but born on a different day).

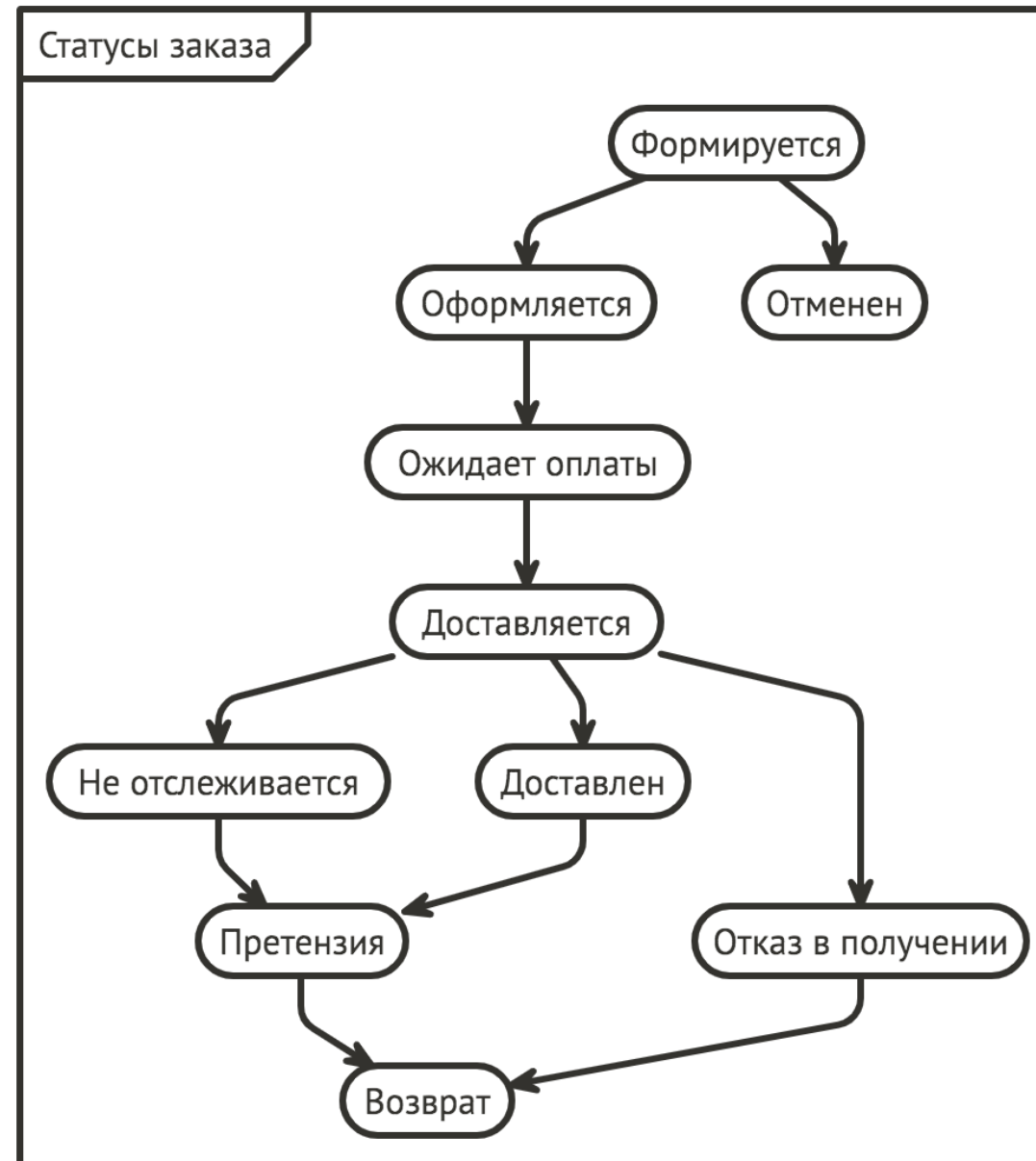
Сущности и атрибуты



Классы и атрибуты



Состояния



На что это всё влияет

Консистентный интерфейс

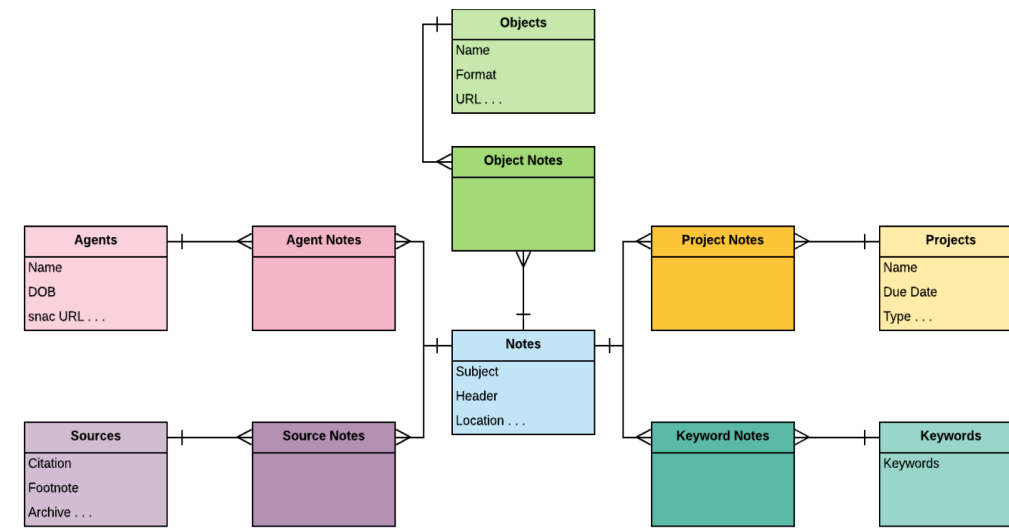
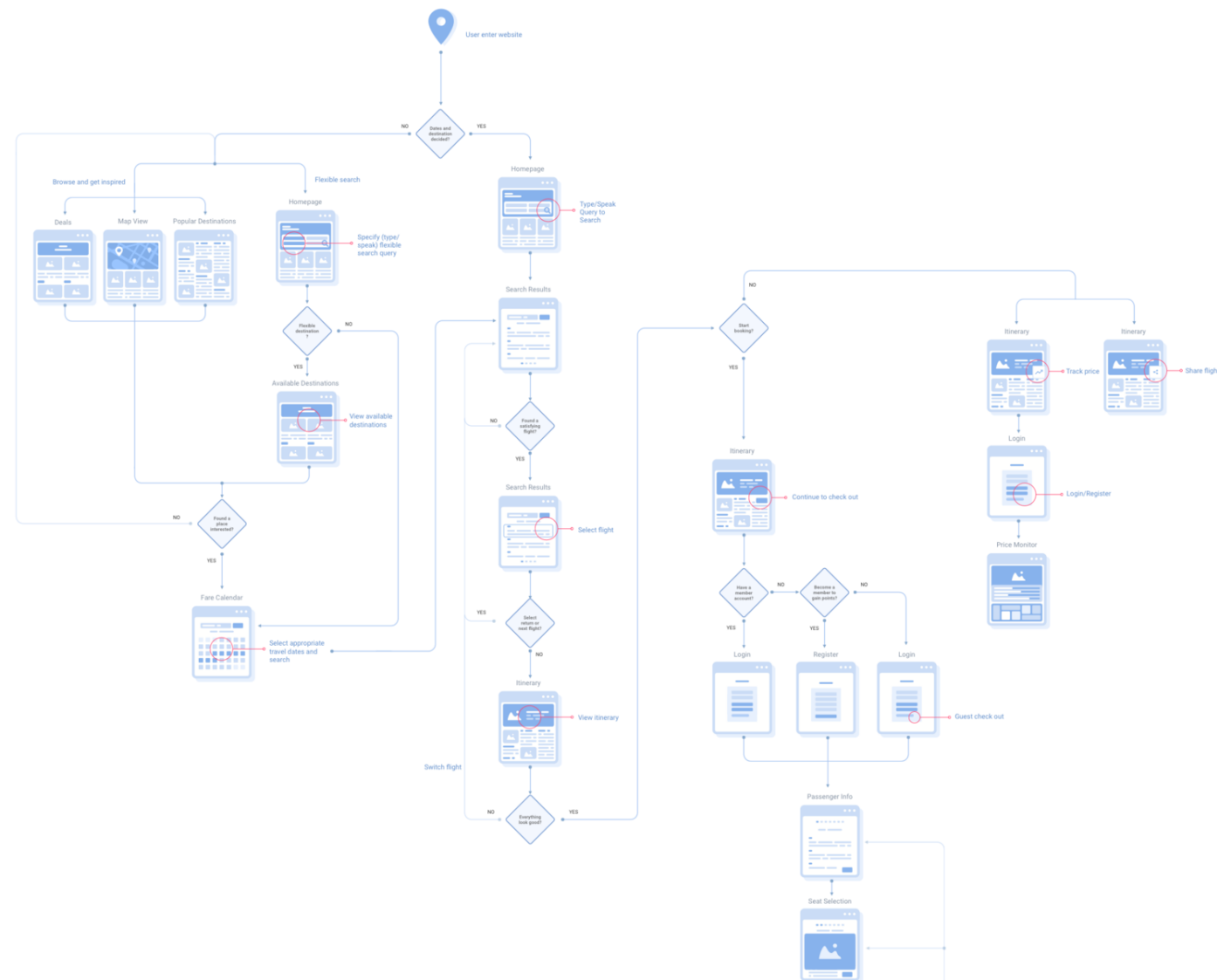
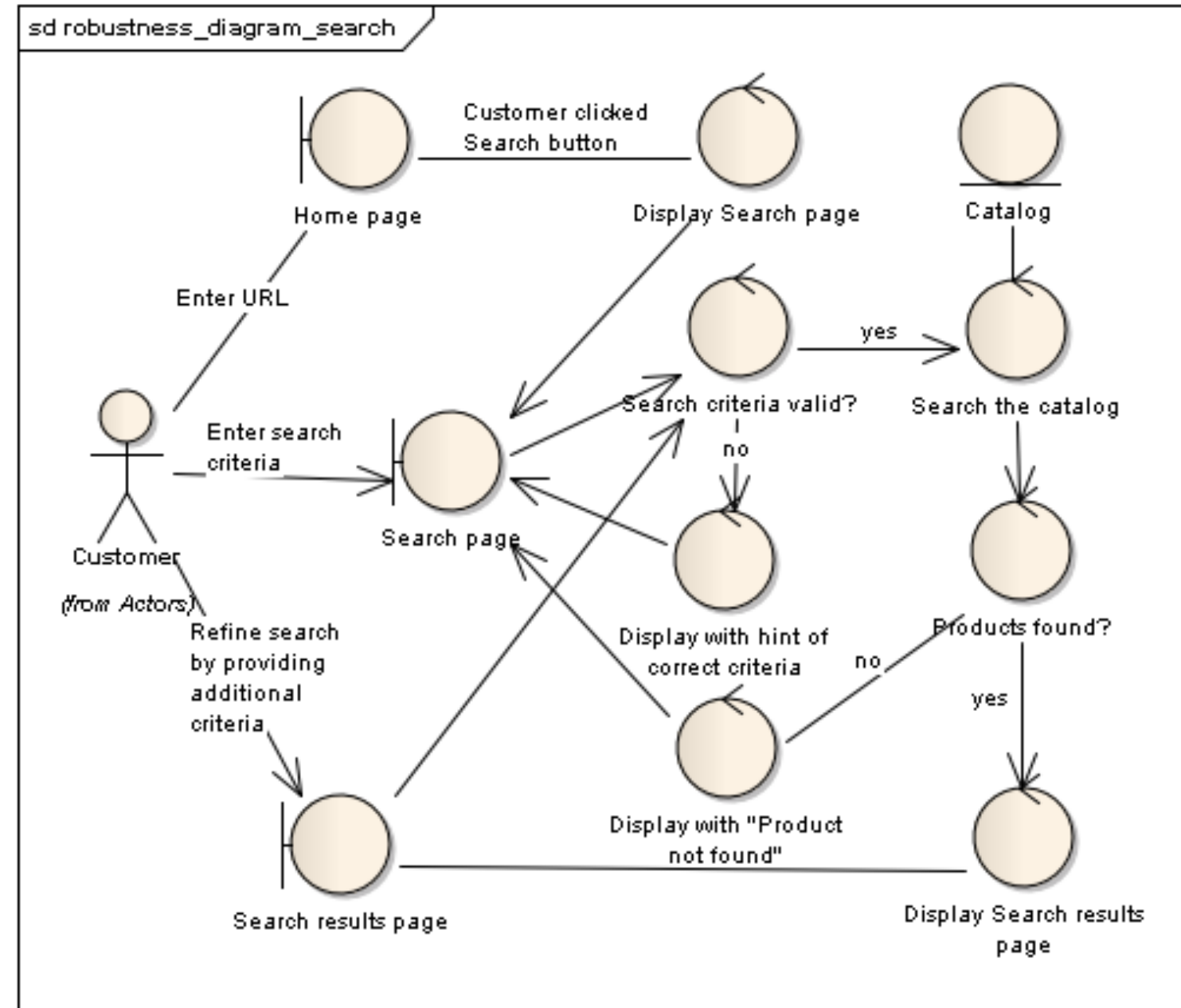
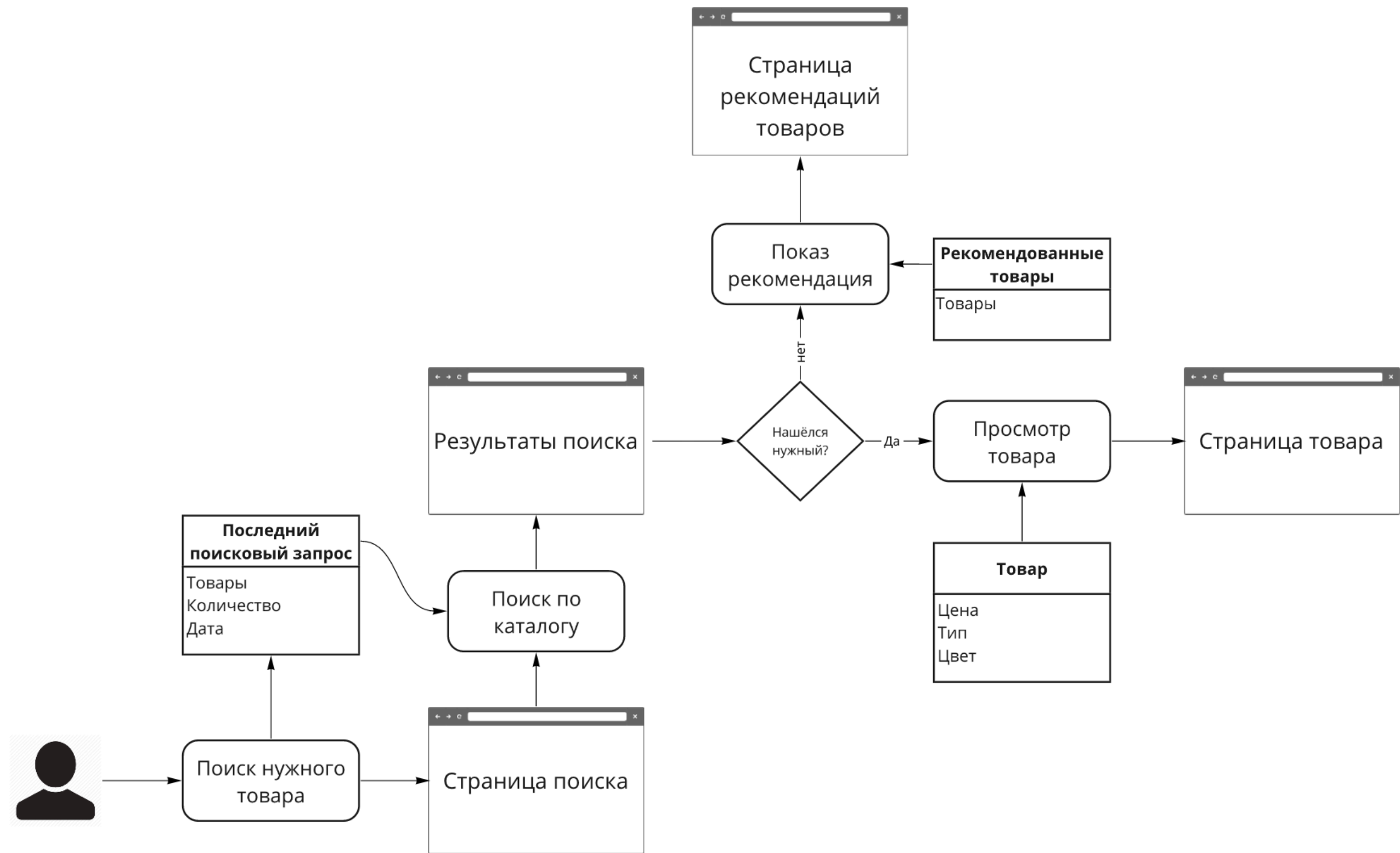


Диаграмма надёжности

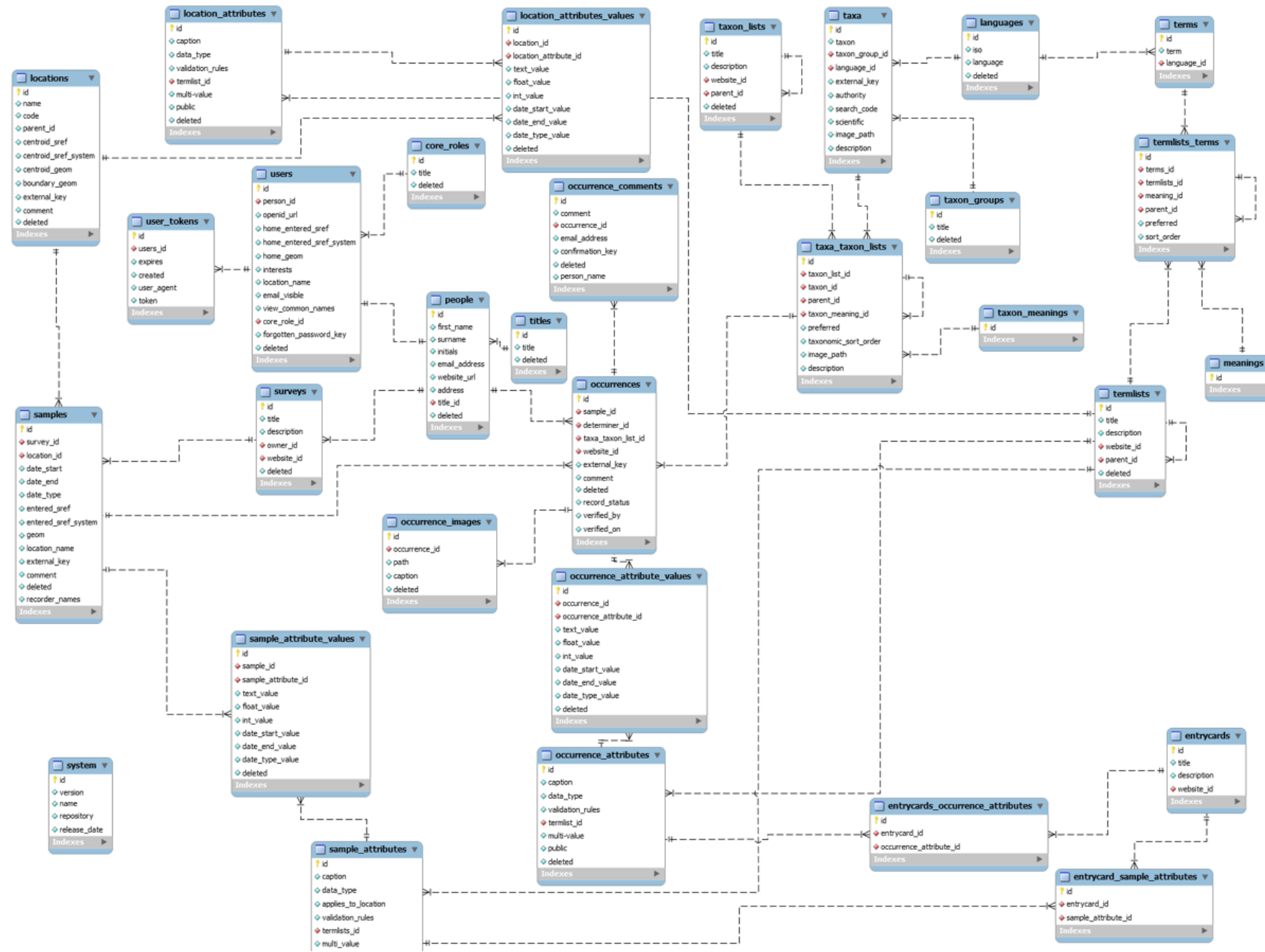




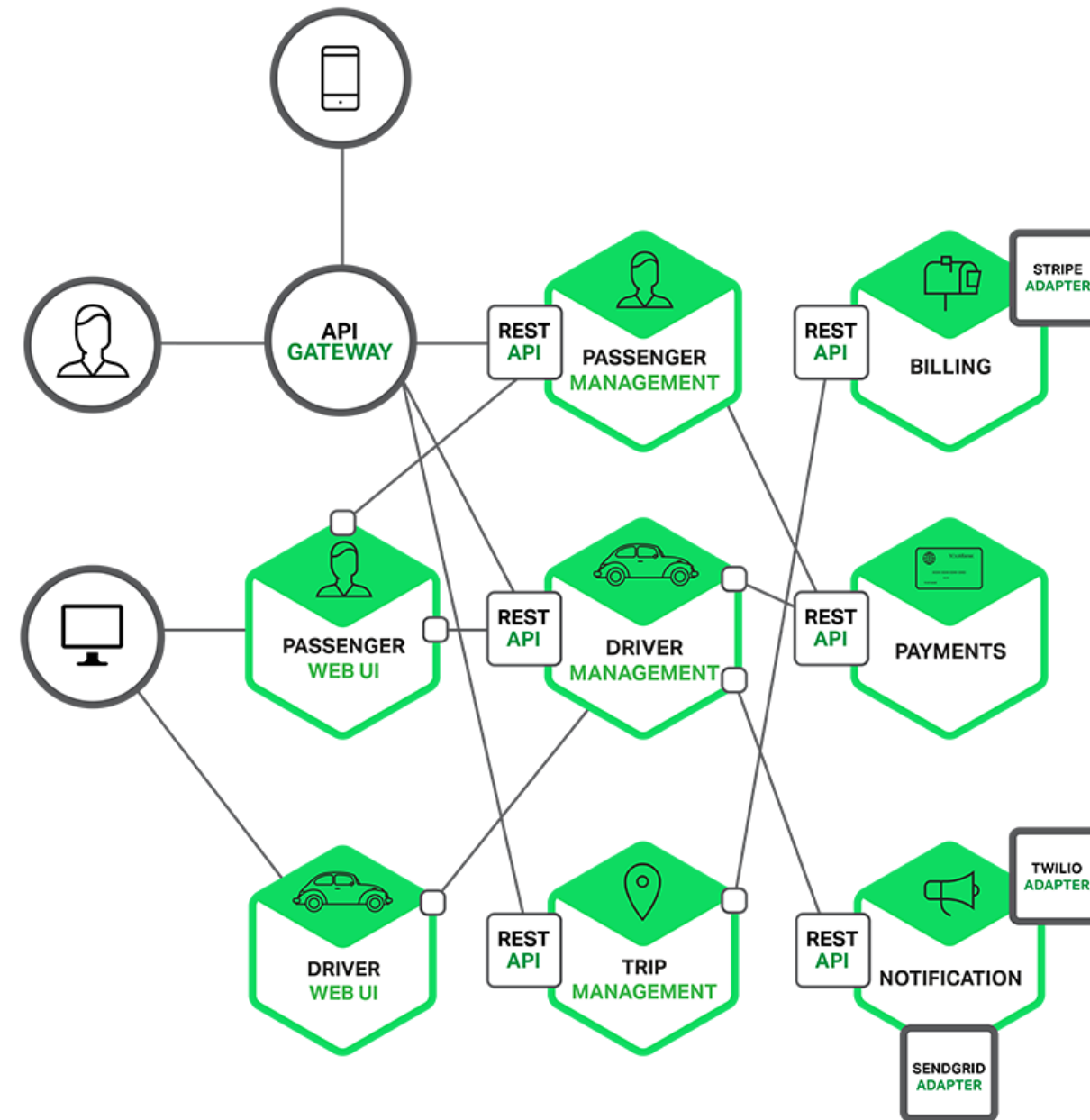
OOUX

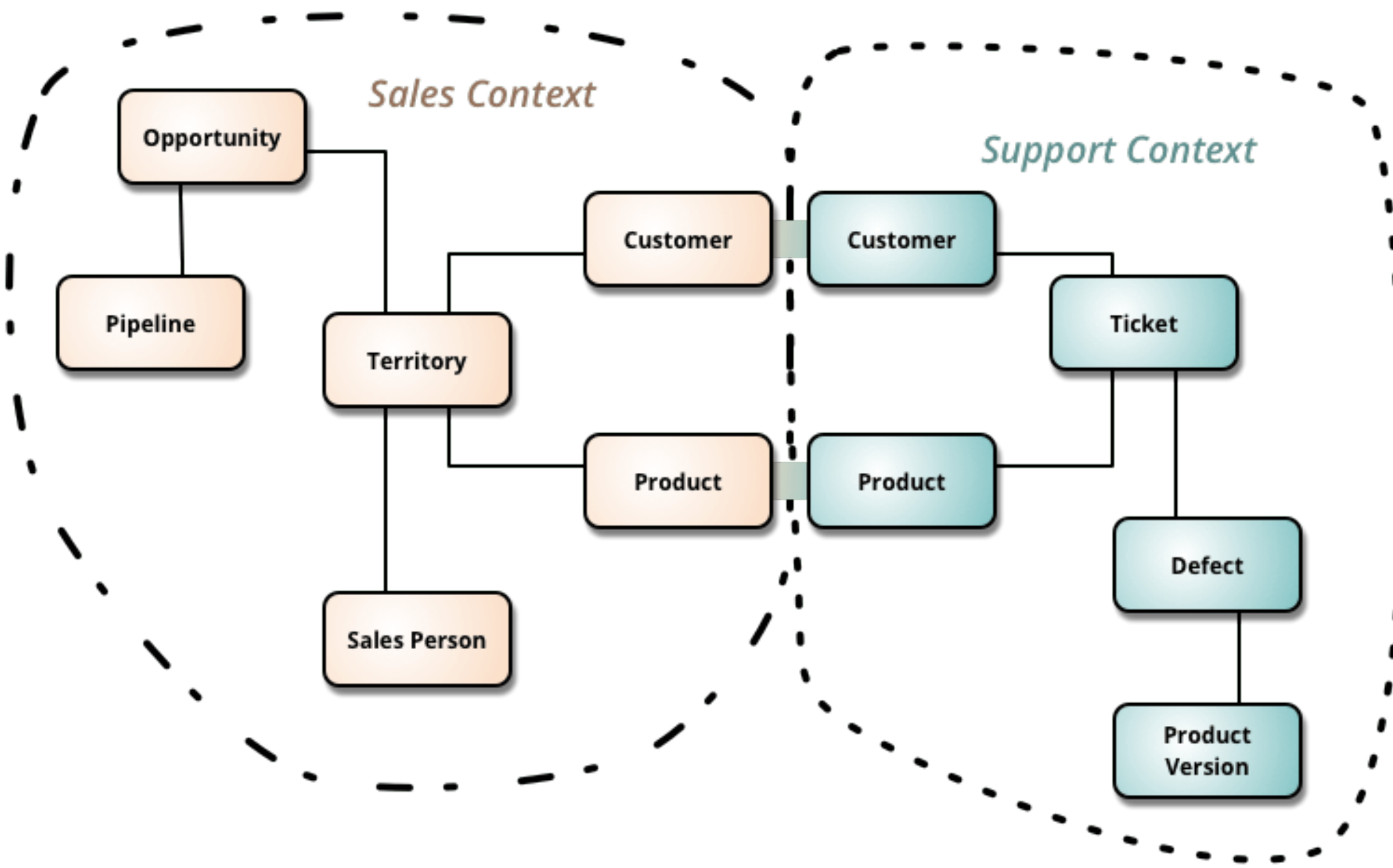


Хранение данных



Системная архитектура





Вместо заключения

- Разберитесь в предметной области
- Сформируйте онтологию, хотя бы в виде глоссария
- Вычлените сущности и атрибуты
- Начните проектирование отталкиваясь от них

Спасибо!

hello@kvaleev.me

@kvaleev 

