

**Комплексная совокупная
модель данных,
доступных в связанной форме во
всемирных референсных
био-медицинских БД**

Александр Осипов

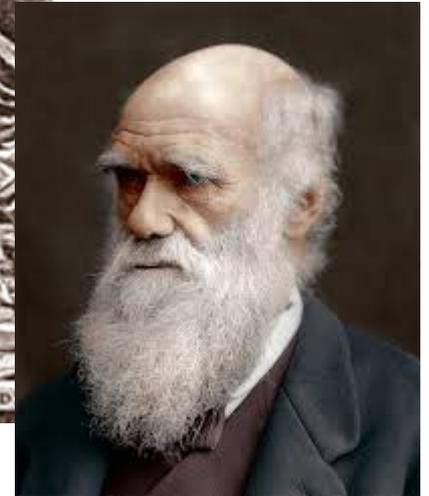
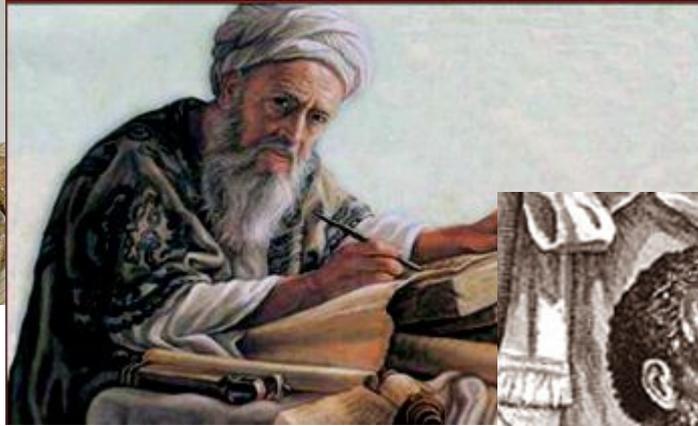
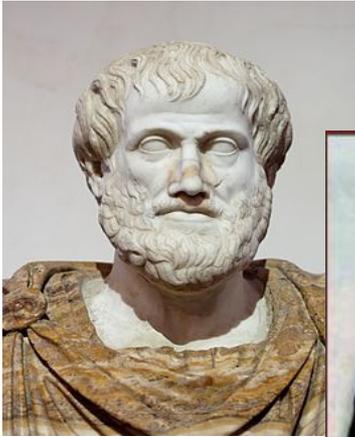
*к.б.н., с.н.с. Института Высшей Нервной Деятельности и
Нейрофизиологии РАН,*

*внештатный научный сотрудник
Института Теоретической и Экспериментальной Биофизики РАН*

aosypov@gmail.com

Мировая наука и региональные науки

Язык науки



Ελληνική γλώσσα

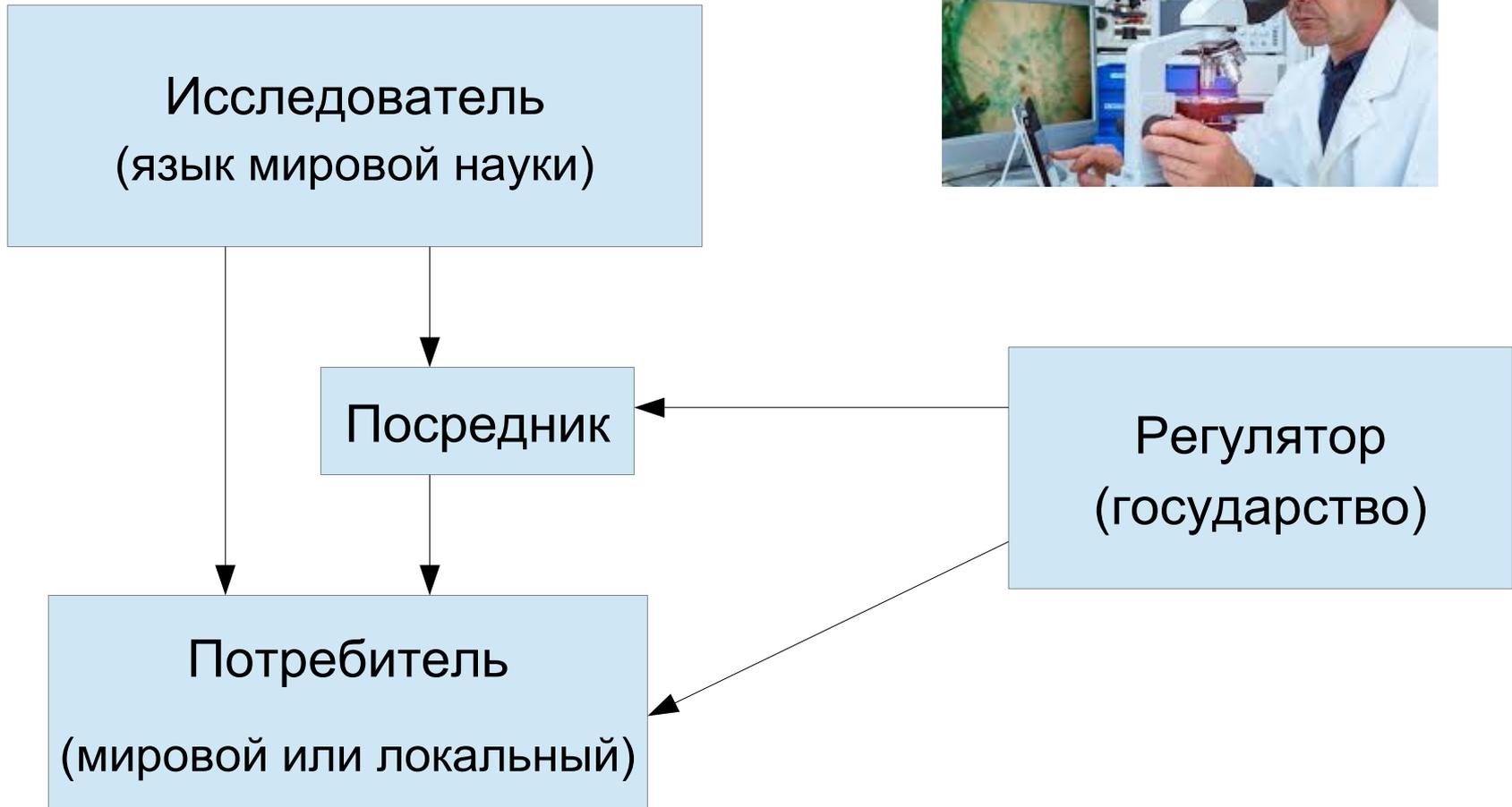
العربية

Lingua Latina

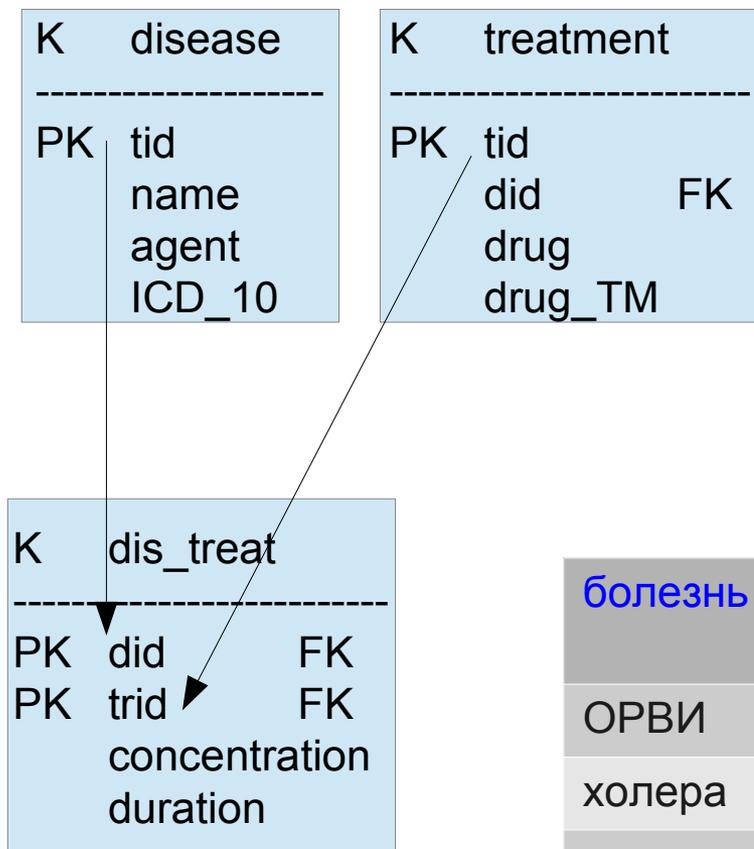
English

Мировая наука и региональные науки

Язык науки



Что такое СВЯЗНЫЕ данные



```
SELECT d.name `болезнь`, d.agent `возбудитель`,  
concat(t.drug, dt.duration) `лекарство`  
FROM disease d, treatment t, dis_treat dt  
WHERE  
dt.did = d.tid AND dt.trid = t.tid  
AND d.name IN ('ОРВИ', 'холера', 'чума')  
ORDER BY d.name
```

болезнь	возбудитель	лекарство
ОРВИ	вирусы всякие	отдых и питье, неделя
холера	Vibrio cholerae	суперпуперин-2, месяц
чума	Yersinia pestis	могила, навсегда

Entity-relation (ER schema)

Primary objects

Taxon

Genome

Organism location

Compound

Metabolic (regulatory) Pathway

Disease

Article

Ontology

Genome
- elements
- products
= variations

Genome element
- id
- sequence
- coordinates
- type
- function
- names

Genome element product (RNA/protein)
- id
- sequence, structure
- type
- names
- biological properties
- modifications

Example databases

= NCBI Taxonomy

= INSDC – Genbank - RefSeq
Gene Expression Omnibus
UniProt, Brenda

= Visible Human Project,
DBSubLoc

= PubChem, FDA, DrugBank

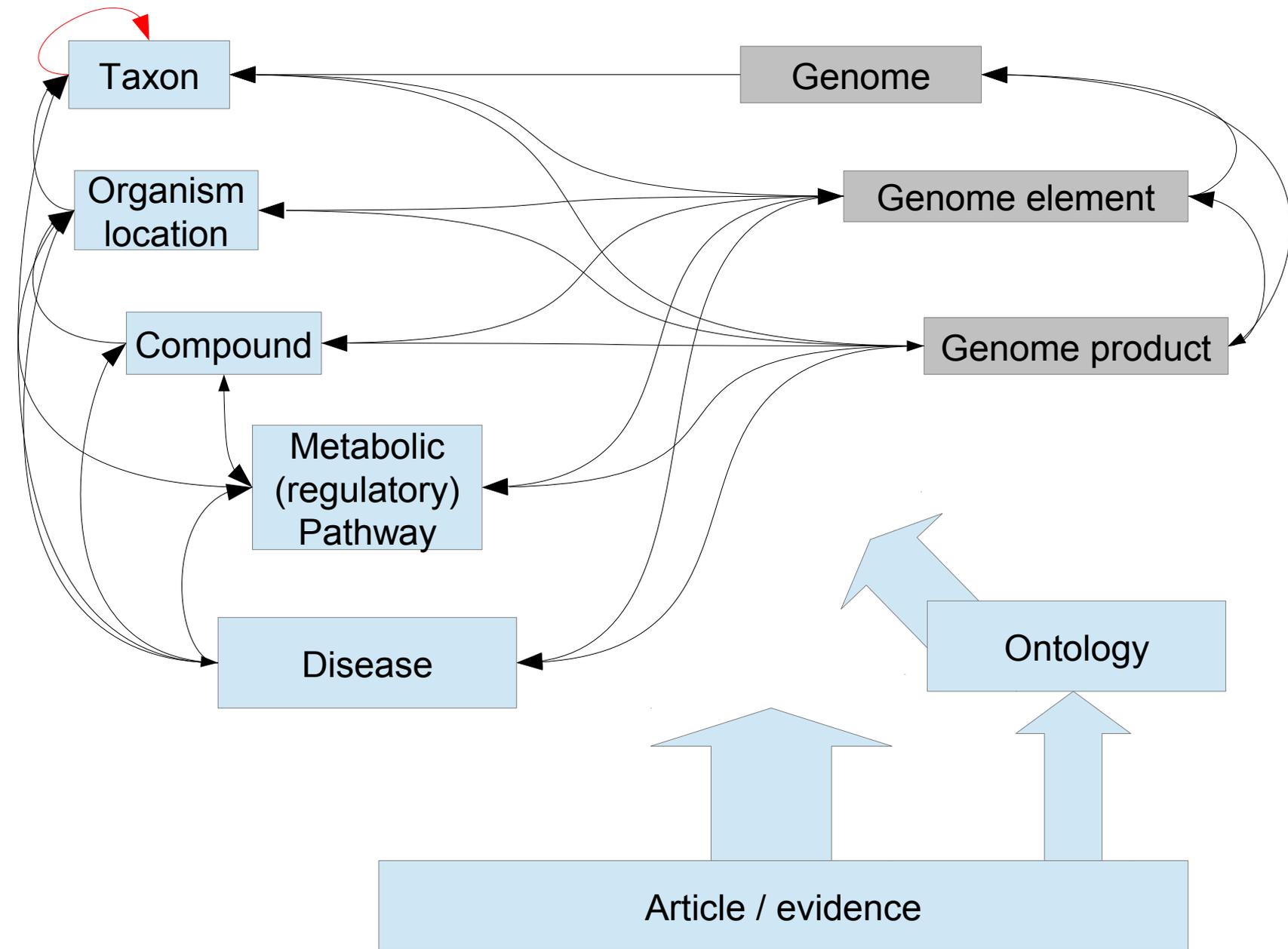
= KEGG,
BioCyc,
EMP

= OMIM, ICD-10

= Pubmed

= GeneOntology

Objects relations



Роль связанных структурированных данных

- Сохранение, автоматическая обработка и извлечение информации
- Субстрат словарного перевода с мирового на язык локального регулятора
- Основа для (полу)автоматической переработки человеческих текстов

dixi,
thank you for your attention!