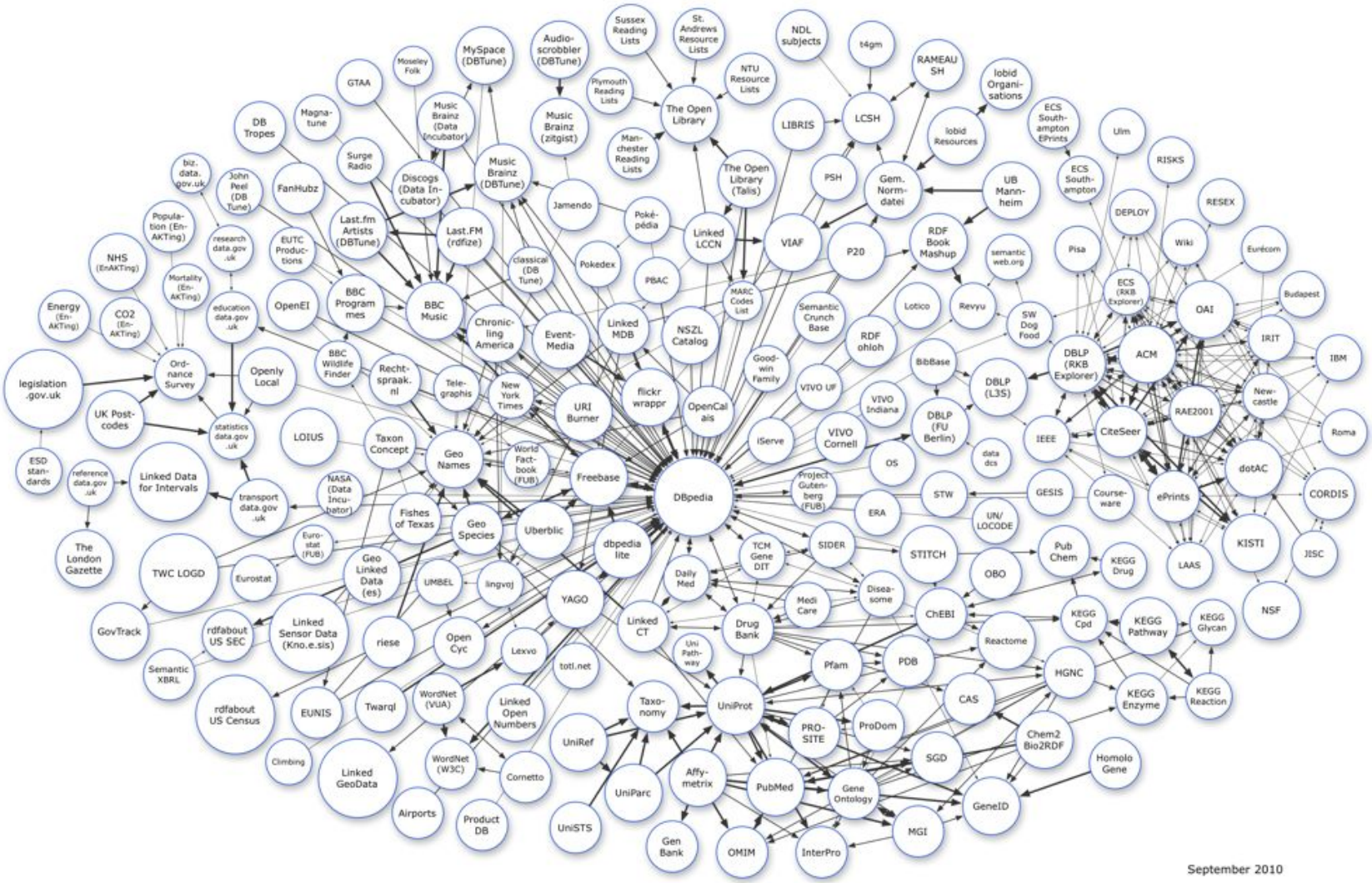
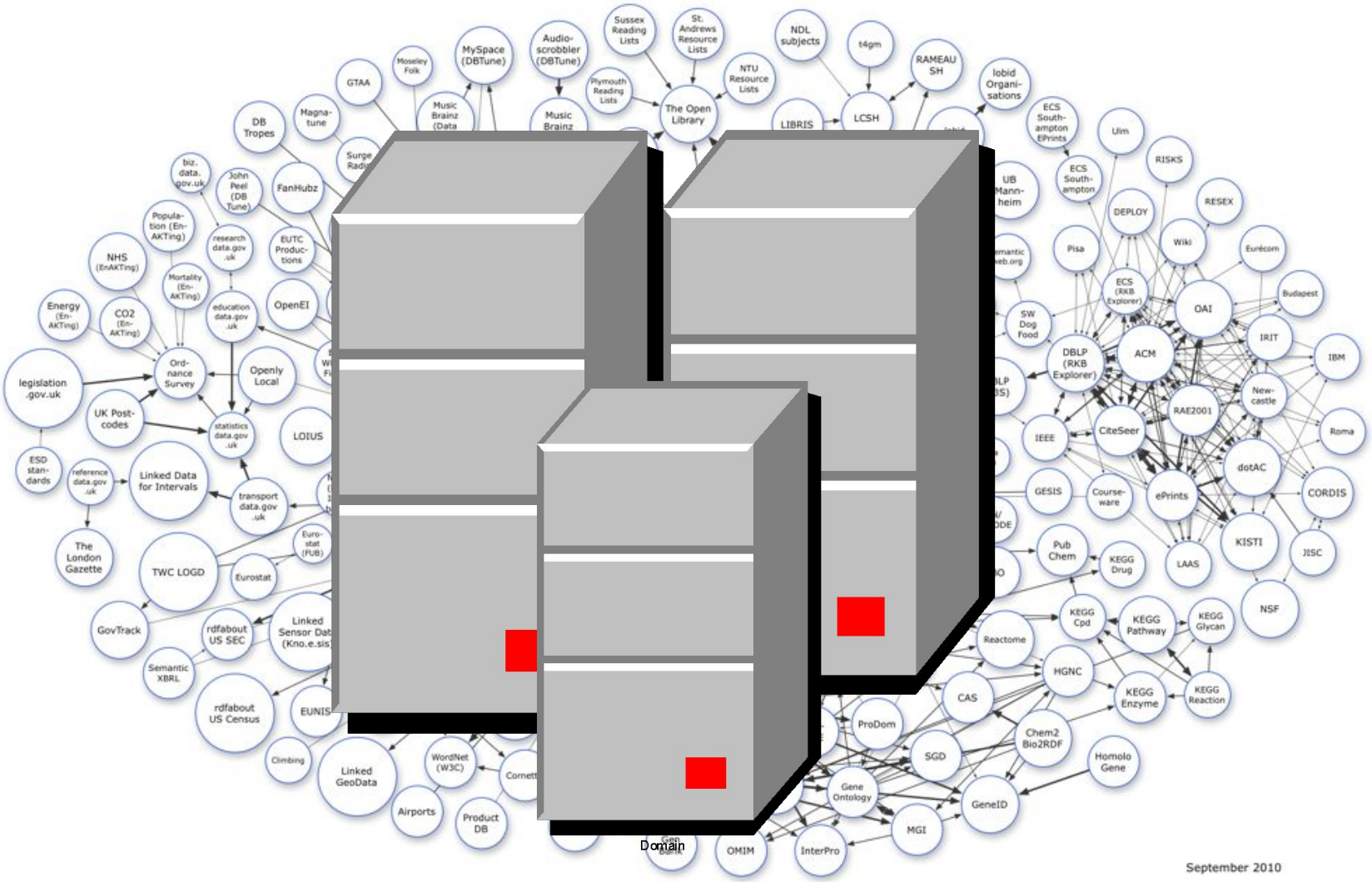
Semantic Web и продукционная модель знаний

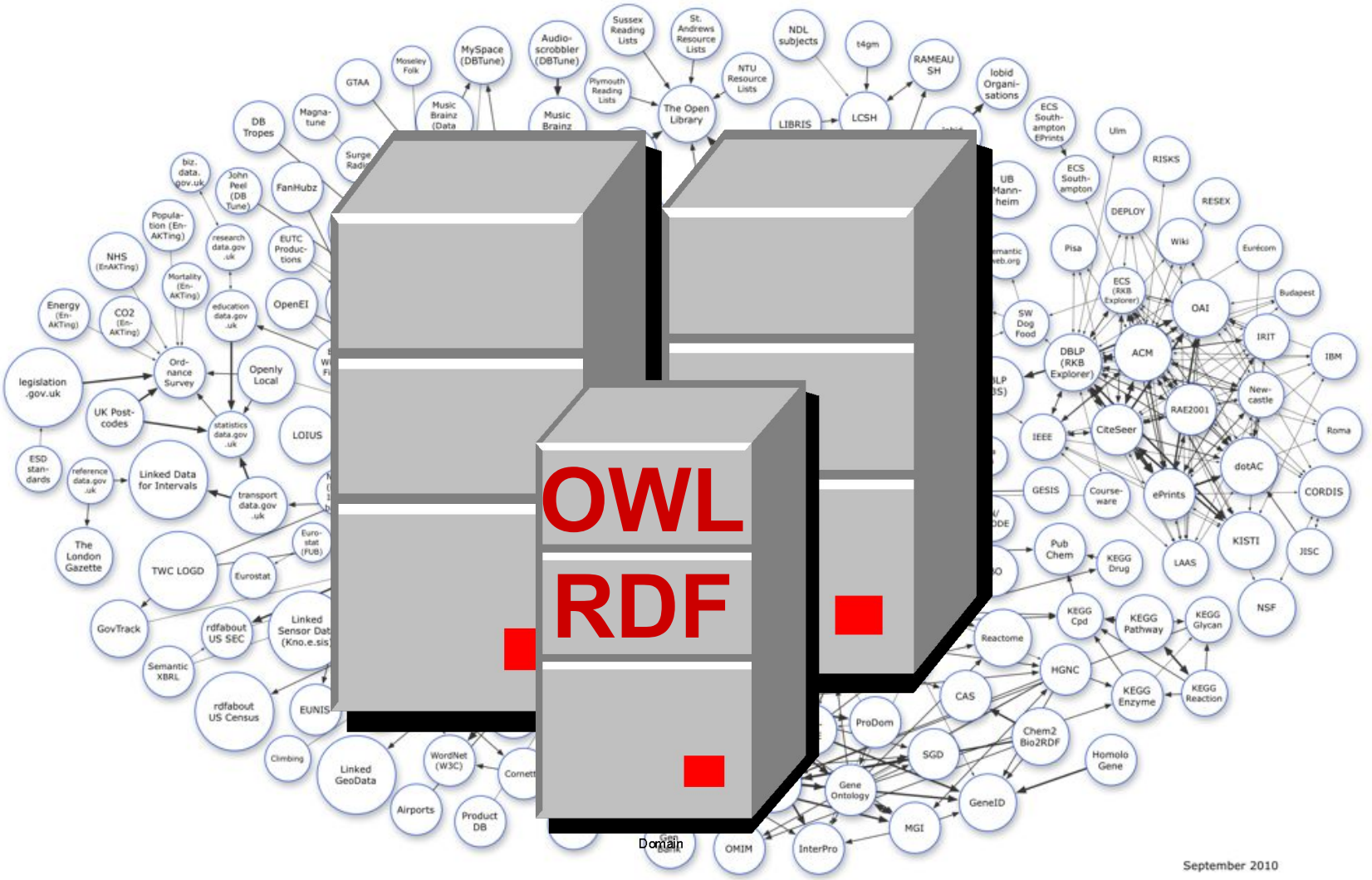
Катериненко Р. С.

аспирантура ИТМО ВТ


Sussex Reading Lists
St. Andrews Resource Lists
NDL subjects
t4gm
RAMEAU SH
lobid Organi-sations
Audio-scrobbler (DBTune)
MySpace (DBTune)
Moseley Folk
GTAA
Plymouth Reading Lists
NTU Resource Lists
The Open Library
LIBRIS
LCSH
lobid Resources
ECS South-ampton EPrints
Ulm
RISKS
DB Tropes
Magna-tune
Music Brainz (Data Incubator)
Music Brainz (zitgist)
Man-chester Reading Lists
The Open Library (Talis)
PSH
Gem. Norm-datei
UB Mann-heim
ECS South-ampton
RESEX
DEPLOY
Surge Radio
Discogs (Data In-cubator)
Music Brainz (DBTune)
Jamendo
Poké-pédia
Linked LCCN
VIAF
P20
RDF Book Mashup
Wiki
Eurécom
biz. data. gov.uk
John Peel (DB Tune)
FanHubz
Last.fm Artists (DBTune)
Last.FM (rdfize)
classical (DB Tune)
Pokedex
PBAC
MARC Codes List
semantic web.org
Pisa
Popula-tion (En-AKTing)
research data.gov .uk
EUTC Produc-tions
NHS (EnAKTing)
Mortality (En-AKTing)
BBC Programmes
BBC Music
Chronic-ling America
Event-Media
Linked MDB
NSZL Catalog
Semantic Crunch Base
RDF ohloh
Lotico
Revyu
SW Dog Food
ECS (RKB Explorer)
OAI
Budapest
IRIT
Energy (En-AKTing)
CO2 (En-AKTing)
education data.gov .uk
OpenEI
BBC Wildlife Finder
Recht-spraak. nl
Tele-graphis
New York Times
URI Burner
flickr wrappr
OpenCal ais
Good-win Family
VIVO UF
VIVO Indiana
BibBase
DBLP (RKB Explorer)
ACM
IBM
legislation .gov.uk
Ord-nance Survey
Openly Local
VIVO Cornell
DBLP (L3S)
New-castle
UK Post-codes
statistics data.gov .uk
LOIUS
Taxon Concept
Geo Names
World Fact-book (FUB)
Freebase
DBpedia
iServe
DBLP (FU Berlin)
RAE2001
CiteSeer
Roma
ESD stan-dards
reference data.gov .uk
Linked Data for Intervals
transport data.gov .uk
NASA (Data Incu-bator)
Fishes of Texas
Geo Species
Uberblic
dbpedia lite
OS
data dcs
IEEE
dotAC
Project Guten-berg (FUB)
STW
GESIS
Course-ware
ePrints
CORDIS
ERA
UN/ LOCODE
The London Gazette
TWC LOGD
Euro-stat (FUB)
Eurostat
Geo Linked Data (es)
UMBEL
lingvoj
YAGO
Daily Med
TCM Gene DIT
SIDER
STITCH
Pub Chem
KEGG Drug
LAAS
KISTI
JISC
OBO
GovTrack
rdfabout US SEC
Linked Sensor Data (Kno.e.sis)
riese
Open Cyc
Lexvo
Linked CT
Uni Path-way
Drug Bank
Medi Care
Disea-some
ChEBI
Reactome
KEGG Cpd
KEGG Pathway
KEGG Glycan
NSF
Semantic XBRL
totl.net
Pfam
PDB
HGNC
rdfabout US Census
EUNIS
Twarql
WordNet (VUA)
Linked Open Numbers
Taxo-nomy
UniProt
CAS
KEGG Enzyme
KEGG Reaction
Climbing
Linked GeoData
WordNet (W3C)
Cornetto
UniRef
PRO-SITE
ProDom
Chem2 Bio2RDF
SGD
Homolo Gene
Airports
Product DB
UniParc
Affy-metrix
PubMed
Gene Ontology
GeneID
UniSTS
Gen Bank
OMIM
InterPro
MGI
September 2010


Domain
September 2010

OWL
RDF



RDF

- Hayes

RDF Semantics

W3C Recommendation 10 February 2004

This Version:
http://www.w3.org/TR/2004/REC-rdf-mt-20040210/
Latest Version:
http://www.w3.org/TR/rdf-mt/
Previous Version:
http://www.w3.org/TR/2003/PR-rdf-mt-20031215/
Editor:
Patrick Hayes (IHMC)< phayes@ihmc.us>

- Horst, H. J. *Combining RDF and Part of OWL with Rules: Semantics, Decidability, Complexity.*
- etc

RDF

- Hayes

RDF Semantics

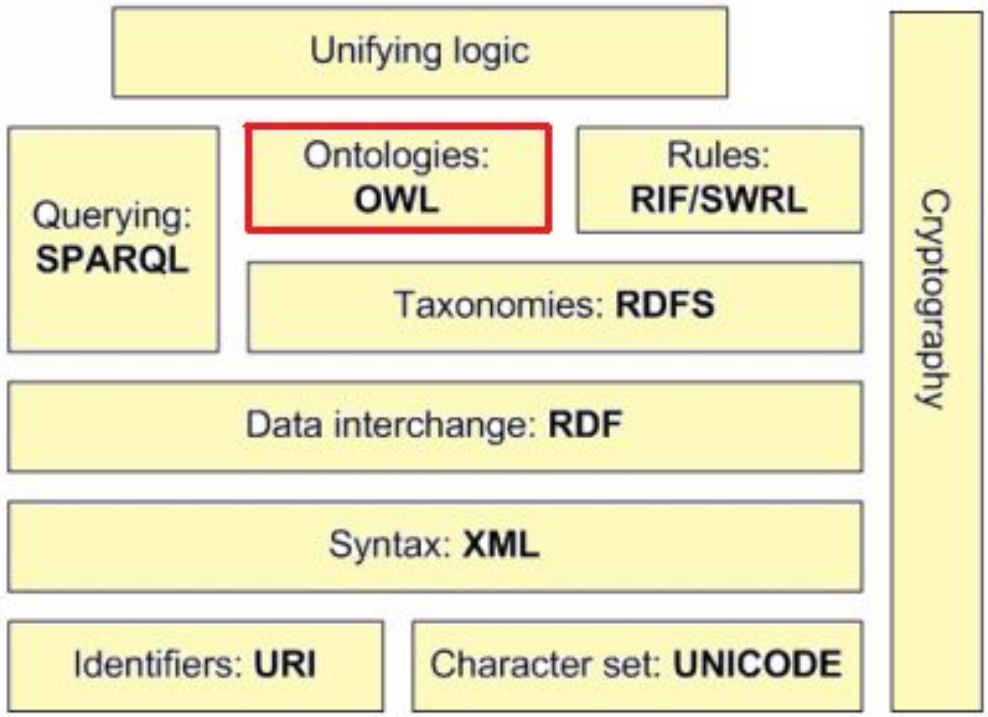
W3C Recommendation 10 February 2004

This Version:
http://www.w3.org/TR/2004/REC-rdf-mt-20040210/
Latest Version:
http://www.w3.org/TR/rdf-mt/
Previous Version:
http://www.w3.org/TR/2003/PR-rdf-mt-20031215/
Editor:
Patrick Hayes (IHMC)< phayes@ihmc.us>

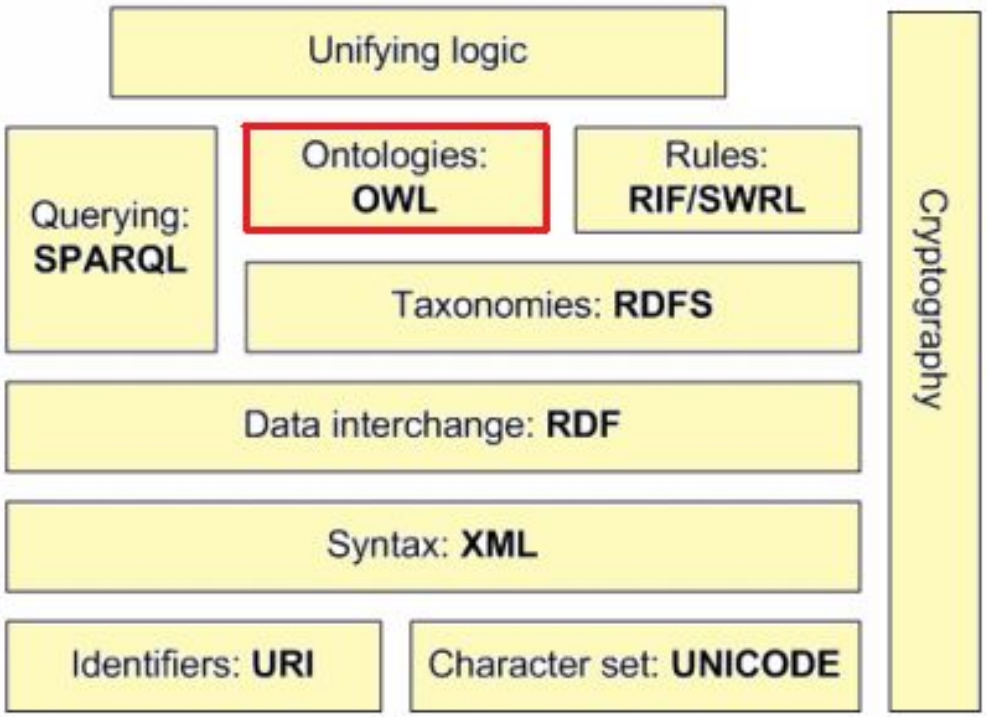
- Sesame
- Virtuoso

- Horst, H. J. *Combining RDF and Part of OWL with Rules: Semantics, Decidability, Complexity.*
- etc

Логическая модель



Логическая модель



- Pallet
- Hawk
- Racer
- Fact

Логическое программирование

Логическое программирование – общее название для языков программирования, программы которых состоят из продукций.

- Prolog
- Datalog

Числа Фибоначчи:

```
F(0,0)
F(1,1)
F(X+Y,N+2) :- F(X,N), F(Y,N+1), N>1

?F(X,100)
```

Стратегии вычисления продукций

- **Forward-chaining, (bottom-up)**

  Последовательное применение продукций к фактам, а затем и к комбинациям фактов и выведенных фактов

  Достоинства:при ответе на запрос не требуется никакого логического вывода
  Недостатки: генерация избыточного количества фактов
  (попробуйте посчитать числа Фибоначчи)

- **Backward-chaining, (top-down)**

Сначала целью вывод объявляется запрос пользователя.
Ризонер пытается найти факты для текущей цели, если их нет,
то ищется продукции, которые могут быть использованы для текущей цели.
Найденная продукция объявляется текущей целью вывода и процесс повторяется.

Преимущества: нет порождения избыточных фактов
Недостатки: для сложных графов вычисления не эффективны,
некоторые цели могут проходится много раз.

Что еще можно делать с продукциями?

Мы используем комбинированную стратегию. При этом применяются продукции к фактам, но мы ограничиваем множество всех продукций, только теми, которые могут быть действительно необходимы для ответа на запрос.

Существуют разные способы это сделать:

- Алгоритм magic set
- Стратификация
- Алгоритм Rete
- Переупорядочивание продукций

Мы используем зависимость продукций по предикатам

Граф зависимости продукций

Запрос:
?a(x)

Продукции:
R1: a(x) :- p(x), q(x)
R2: p(y) :- p1(y)
R3: p(k) :- p1(k)
R4: q(z) :- q1(z)

Генерация SQL

- Предикат *p(x)* опрделеляет множество :
  Select * from F where F.s like “P”

- Join
  a(x) :- p(x), q(x)
  p(x) – S1
  q(x) – S2
  Select * from S1 join S2
- Union
  p(y) :- p1(y) //S3
  p(k) :- p2(k) //S4
  Select * from S3 union S4

- Вложенный подзапрос:
  S1 = S3 union S4 = Select * from (Select * from S3 union S4)

Генерация продукций

Продукции => SQL

+ Совмещение SQL запросов с семантическими к реляционной СУБД

Генерация продукций

SPARQL => Продукции => SQL

+ Совмещение SQL запросов с семантическими к реляционной СУБД

+ Инфраструктура для семантических приложений

Генерация продукций

+ Быстрая инфраструктура для семантических приложений:
    на 20% быстрее, чем Jess (Rete)

SPARQL => Продукции => SQL

+ Совмещение SQL запросов с семантическими к реляционной СУБД

+ Инфраструктура для семантических приложений


Спасибо!