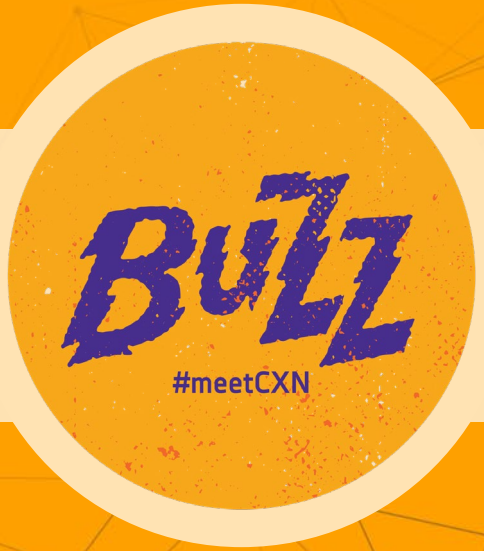
ChemAxon

Andras Volford



JCHEM FOCUS:
CHORAL, MICRO SERVICES

HISTORY

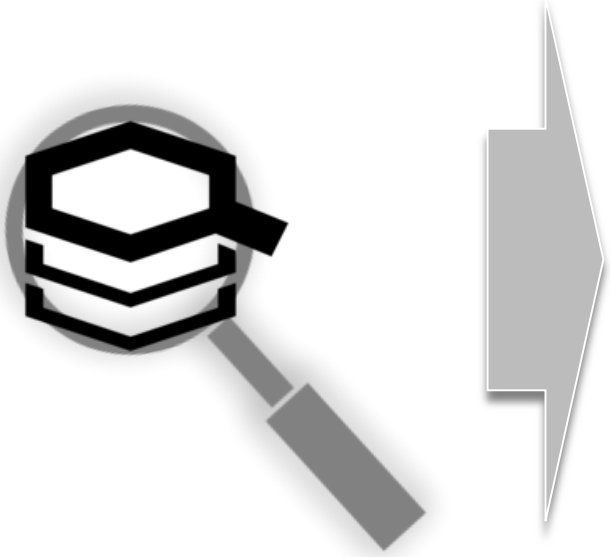
JChem Base

JChem Oracle Cartridge

new-search PostgreSQL Cartridge

new-search PostgreSQL engine

ChemAxon



JChem Coral

JChem Microservices

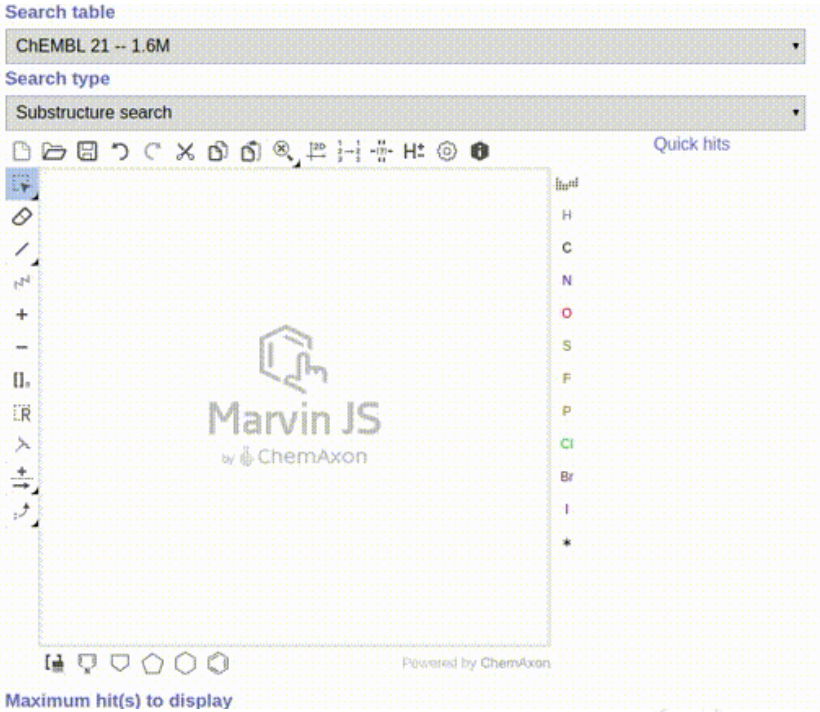
ChemAxon

JCHEM CHORAL

Hit as You Draw

Fast searches with hit limit

Hits ordered according to query similarity



ChemAxon

OPTIONS MISSING?

Options:

- …
- exactQueryAtomMatching
- radical
- isotope
- charge
- valence
- doubleBondStereo
- vagueBond
- HCountMatching
- implicitHMatching
- exactStereoMatching
- ...

Hits ordered according to query similarity.

Transformations:

- full fragment
- double bond markedonly

JCHEM CHORAL

Fast joined / combined search:

chemical search + non-chemical search

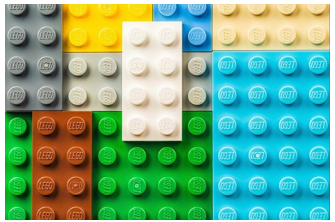
Handle large databases

Amazon RDS

ORACLE DATABASE 12c

ChemAxon

JCHEM MICROSERVICES



Modular

Scalable

Easily manageable

Highly Available

MODULES

Stateless

Stateful

IO

Structure manipulation

Base

Chemical property calculations

DB Search

Search stateless

Search

Markush stateless

Markush DB

Markush

Structure Checker

Reactor DB

Reactor

Reaction handling

Cartridge

JOC query

JPC query

Choral query

PROJECTS, EXPERIMENTS

PROJECT HAYSTACK

ChEMBL
1.7M

Corporate
Compound
Repository
1.5M – 3M

MolPort
All Stock
~7M

Enamine REAL
~680M

SureChEMBL
~18M

Namiki
Shoji
~6M

PubChem
Compounds
~95M

Screening
library
250k – 1M

eMolecules
~18M

BindingDB
651K

ChemAxon

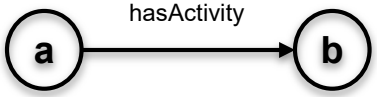
GRAPH DB

Nodes : entities, structures, etc.

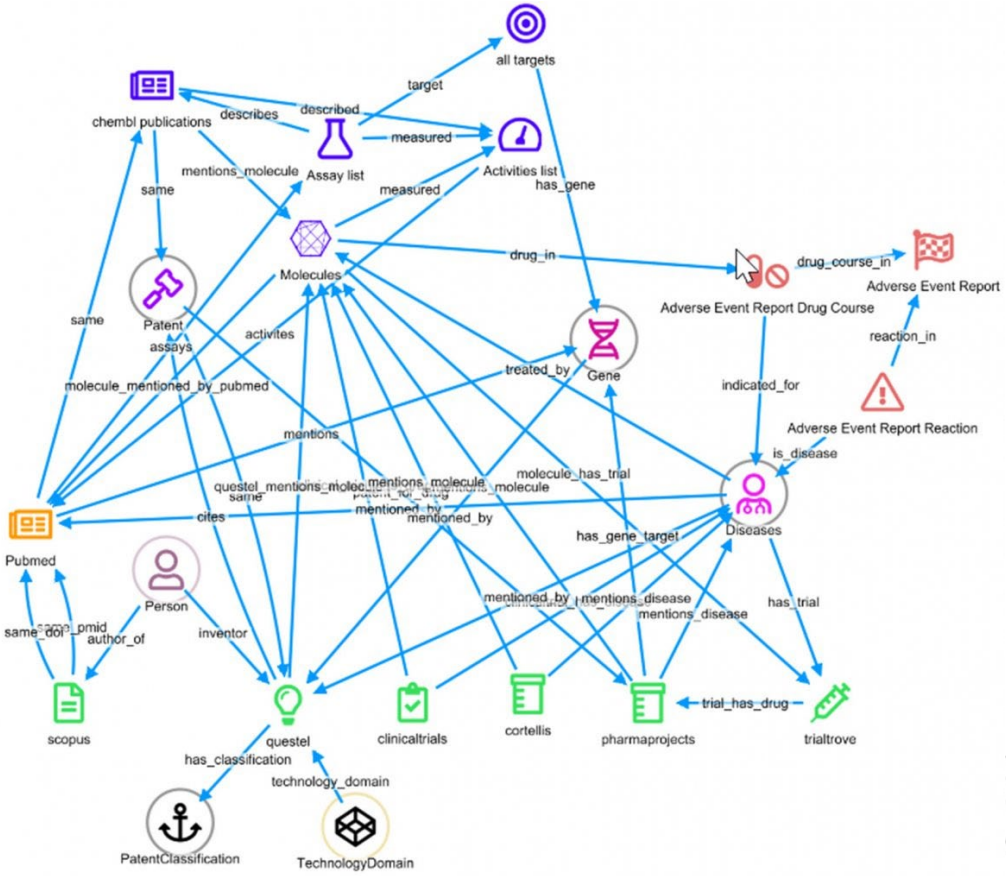
Relations : connection between the nodes.

No tables. Extra nodes and relations
can be added without corrupting the db.

query language (Cypher QL in Neo4j)

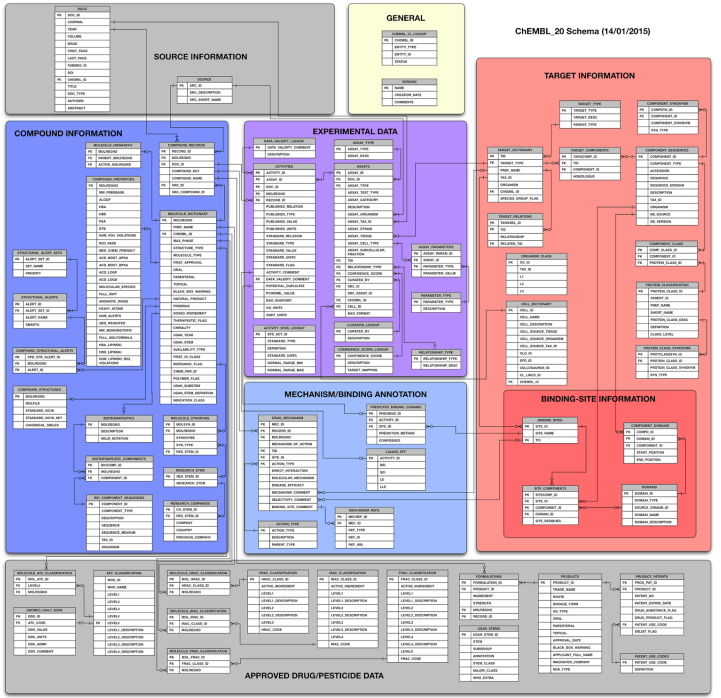


(a) - [:hasActivity] - (b)



WHY GRAPH DATABASE?

- Relationships > Structural data
- *SELECT TableA.*, TableB.*, TableC.*, TableD.*
  FROM TableA
  JOIN TableB
  ON TableB.aID = TableA.aID
  JOIN TableC
  ON TableC.cID = TableB.cID
  JOIN TableD
  ON TableD.dID = TableA.dID
  WHERE ...



ChemAxon

JCHEM NEO4J CARTRIDGE

Chemical search:

substructure search

similarity

on set of nodes

Kind of an index mode search
(no functional mode search yet)

Create an index set

```
call jchem.createdb('test','sample')
```

Index nodes:

```
match (n) with collect(n) as nodes call
jchem.addBatch('test',nodes,'molString') yield
responseCode return responseCode
```

Search for nodes:

```
call jchem.search('test','c1ccccc1',10)
```

Create trigger

```
call jchem.createTrigger('molecule','molString','test')
```

ChemAxon

Choral	*Zsuzsa Szabó* Developer Workshop
	Akos Tarcsay How fast is ChemAxon RDMS search
JPC, JMS	*Iván Solt* Haystack Project
New search engine	*Tamás Varga* Predictive search and drawing
Neo4J	*Andras Volford* Graph Database at Sanofi
Cartridge	*Zsuzsa Szabo* Developer Workshop

Thank you

ChemAxon