



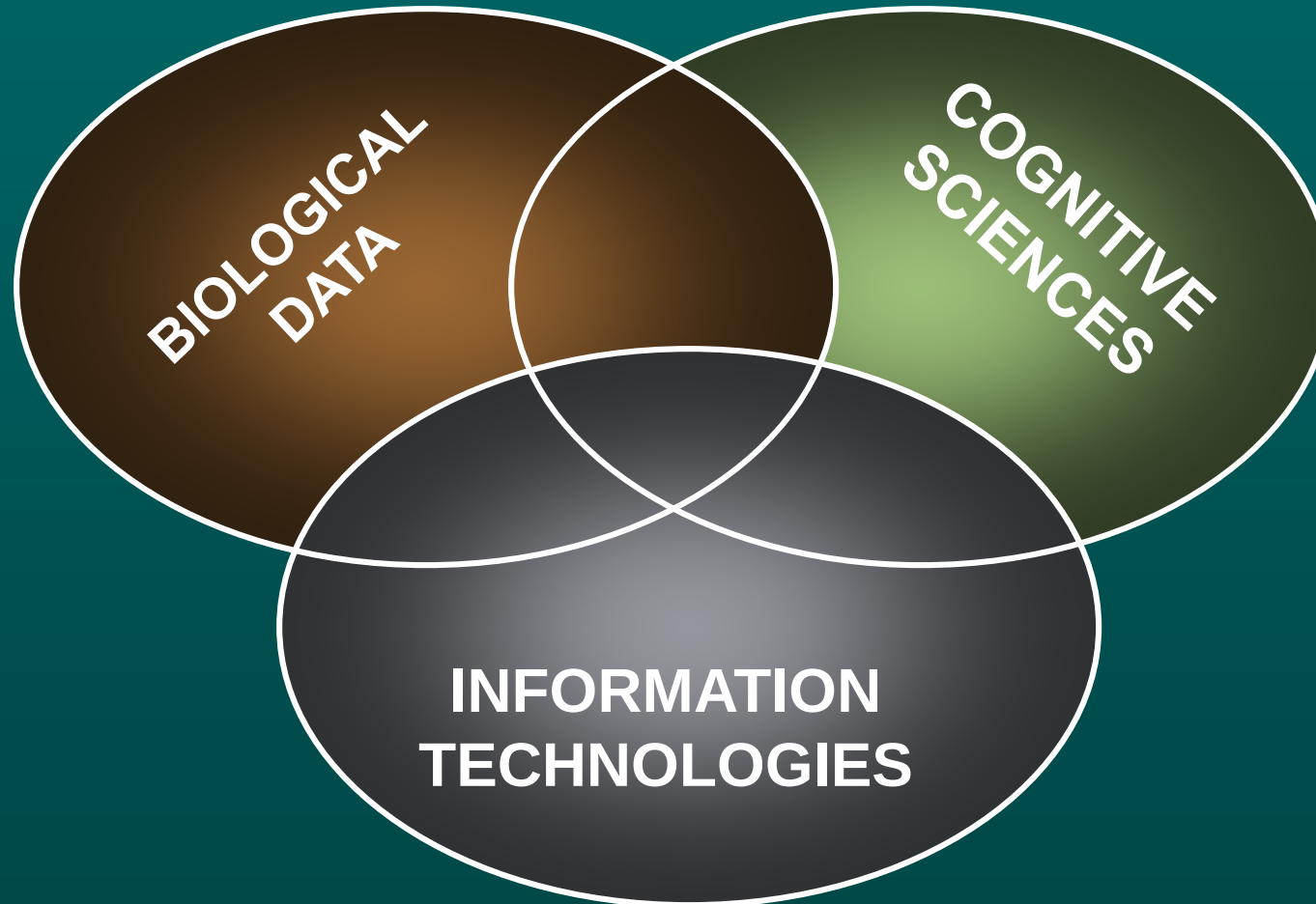
Biológiai adatbázisok integrációja

Pongor Sándor

az MTA és az Academia Europaea tagja



BIOINFORMATICS@ppke



Model, description and visualization



Molecular objects

- Genes, genomes
- Proteins, proteomes

Data collection:

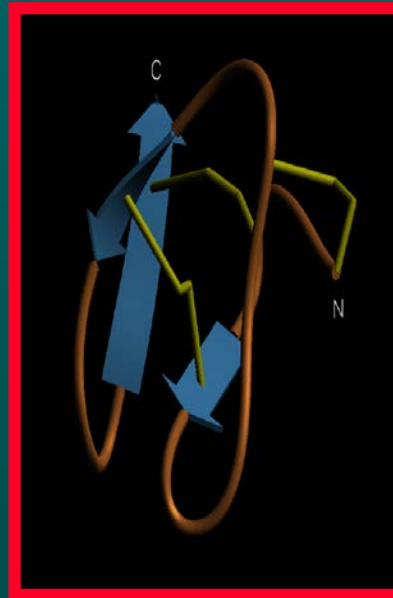
- Next generation sequencing
- Mass spectrometry

Core data-types

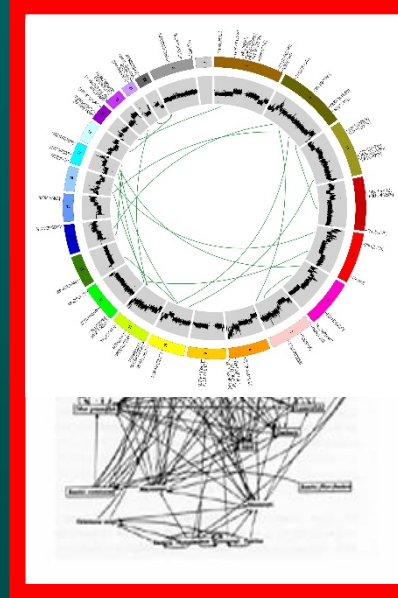
ALL HAVE SIMPLIFIED AND EXTENDED (ANNOTATED) VERSIONS

```
tassfvswvsasdtvsgfrvey  
elsegdepqyldlpstatsvni  
pdllpgrkytnvyeiseegeqn  
lilstsqttapdapdpdtdvd  
dtsivrvsrprapitgyrivys  
psvegsstelnlpetansvtlsd  
lqpgvqynitiyaveenqestpv  
fiqqettgvprsdkvpprdlqf  
vevtdvkitimwtppesvtygr  
vdivpvnlpgehgrlpvsrntf  
aevtglspgvtyhfkvfavnqgr  
esklptaqqatkldaptnlqfin  
etdttviviwtpprarivgyrlt  
vgltrggqpkqynvgpaasqypl  
rnlqpgseyavslvavkgnqqsp  
rvtgvfthlqplgsiphyntevt  
ettivitwtpaprigfklgvrps  
qggeaprevtsesgsivvsglt  
gveyvytisvlrdgqerdapivk
```

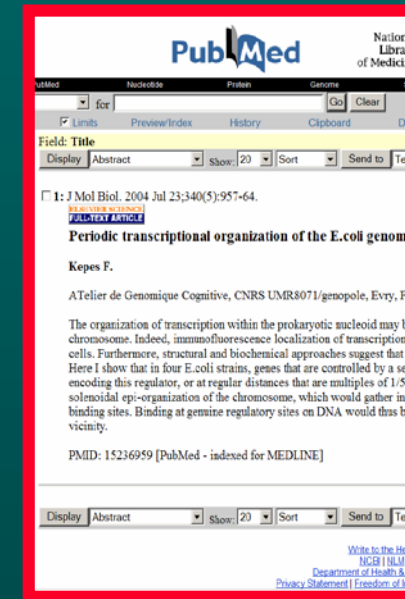
SEQUENCES



3-D



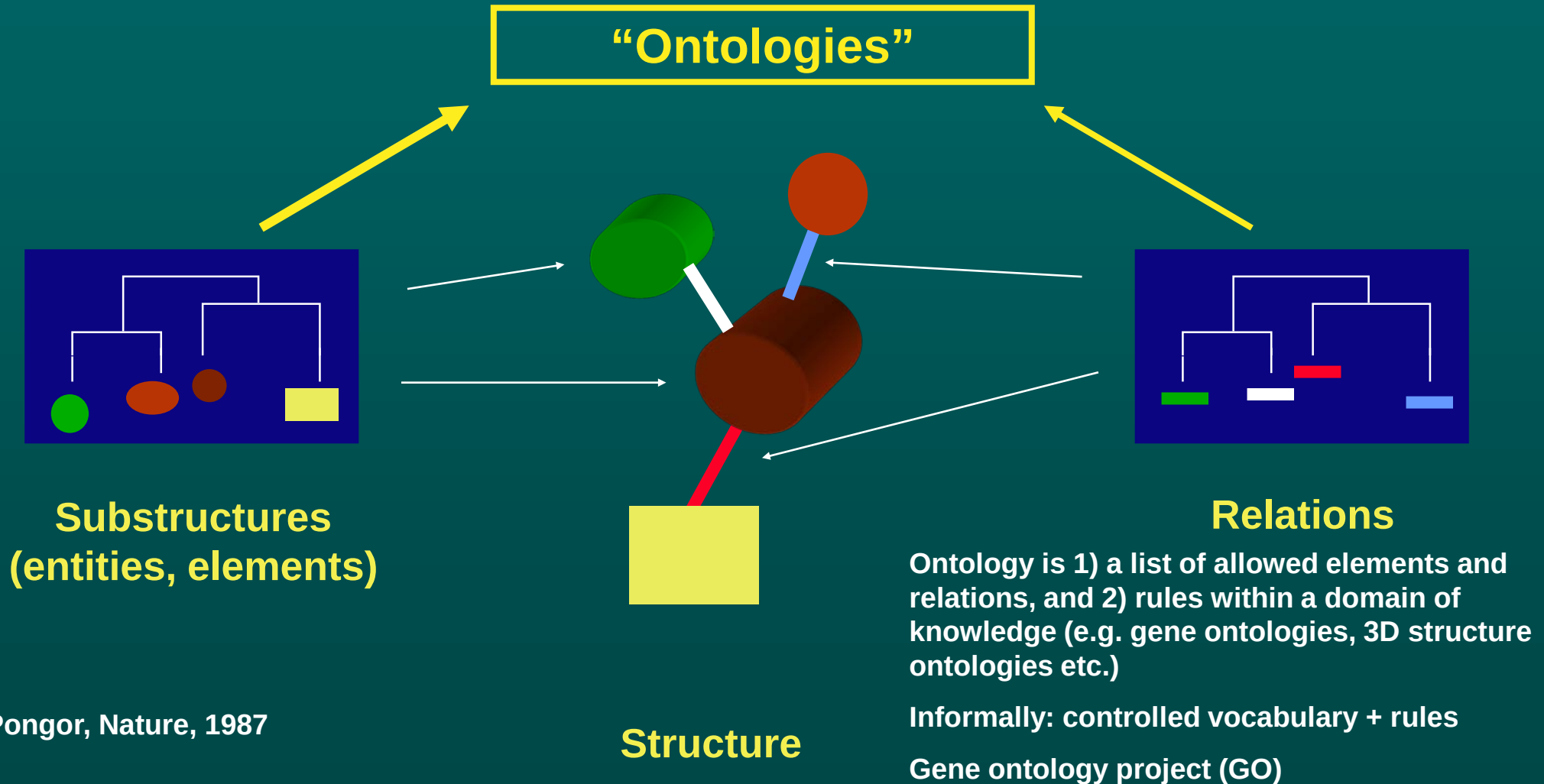
GENOMES
NETWORKS



TEXT

DEDICATED DATABASES

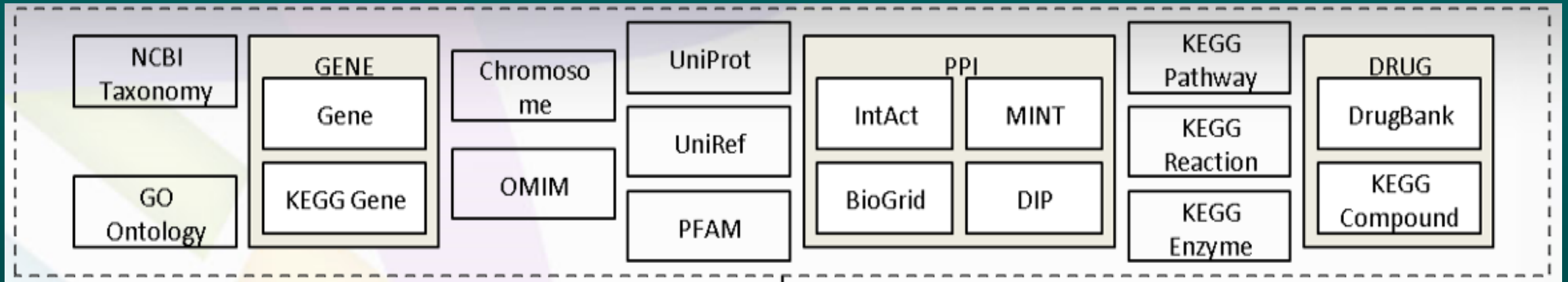
Common framework: substructures and relationships



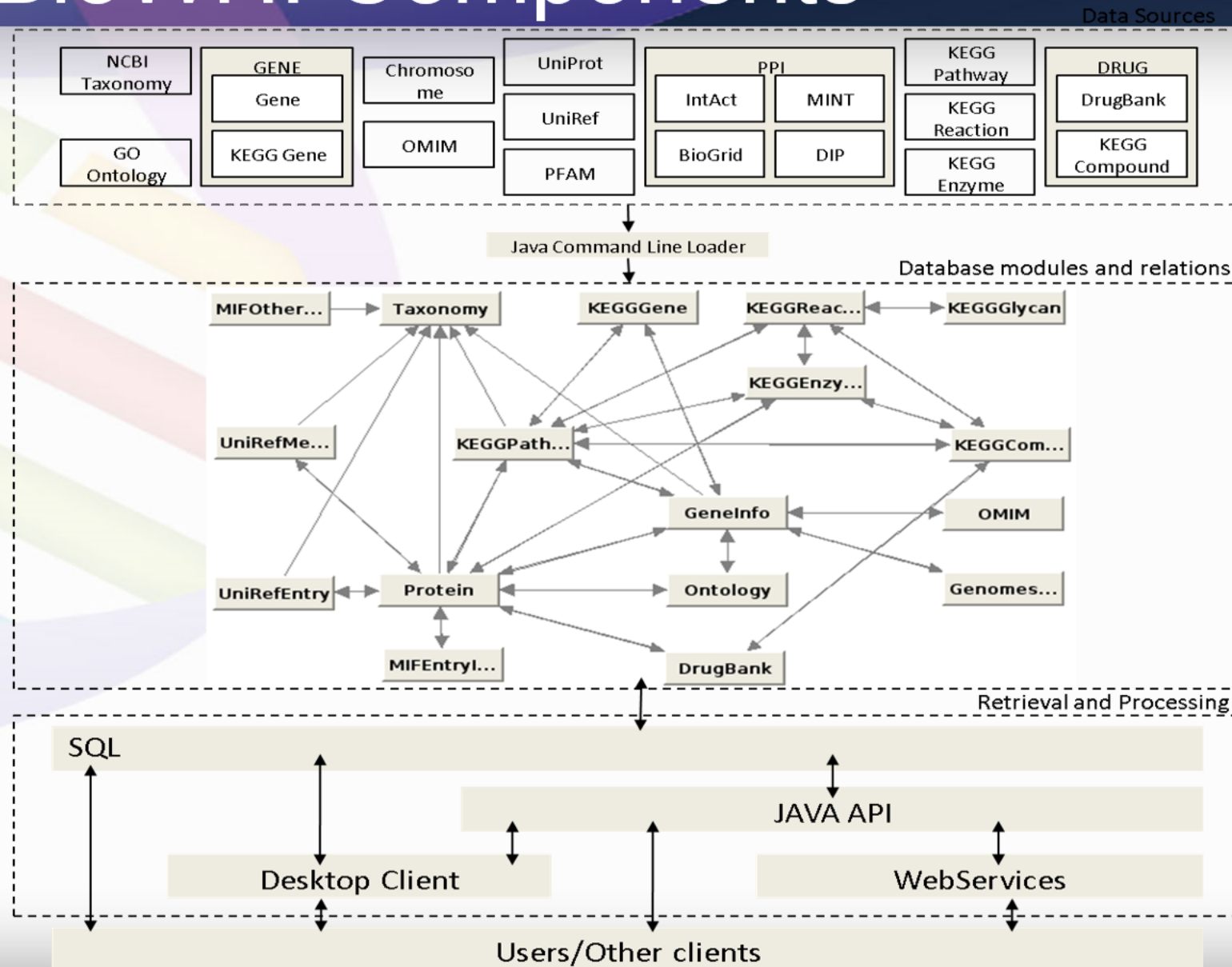


BIOINFORMATICS@ppke

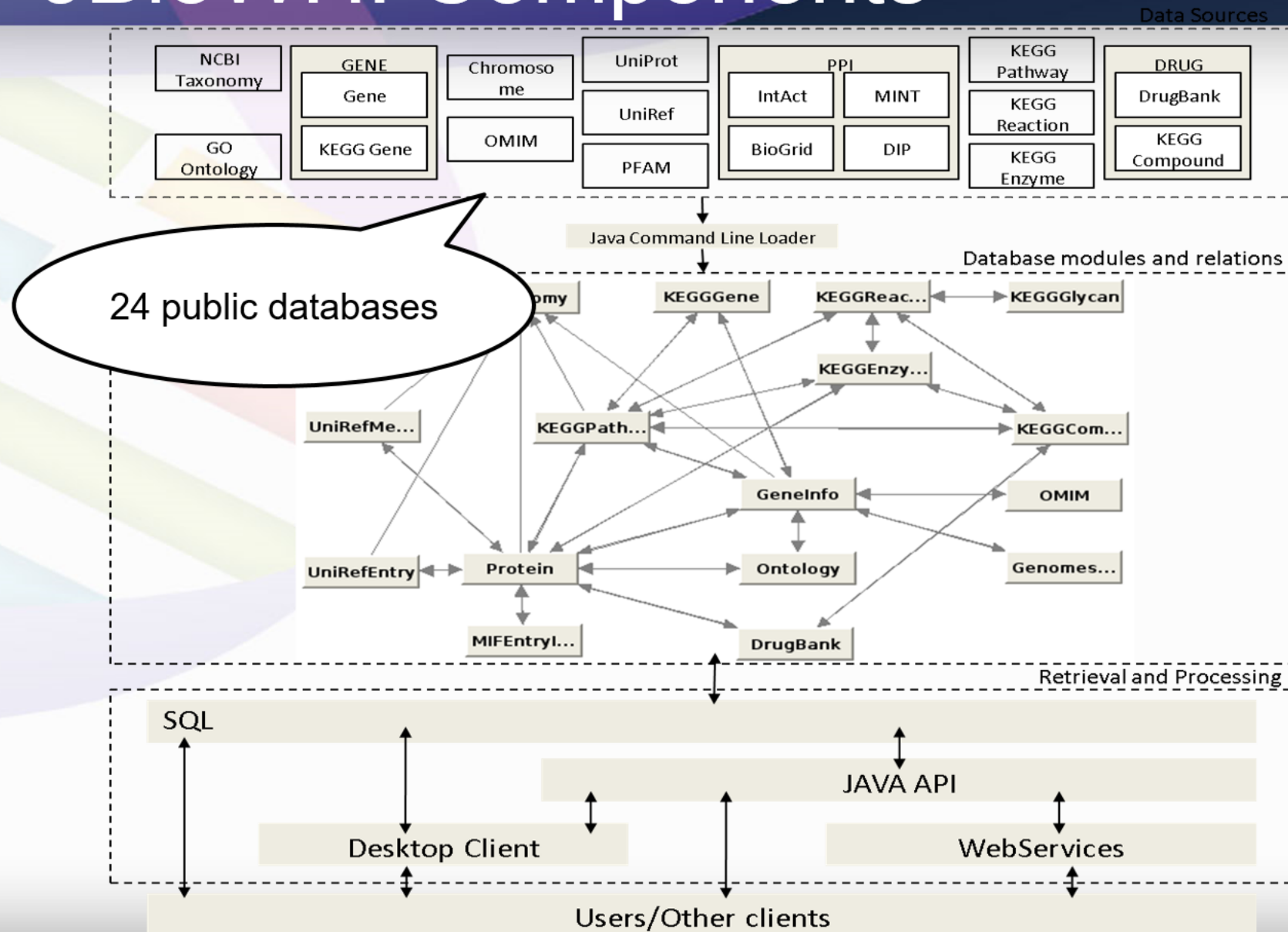
Databases by content



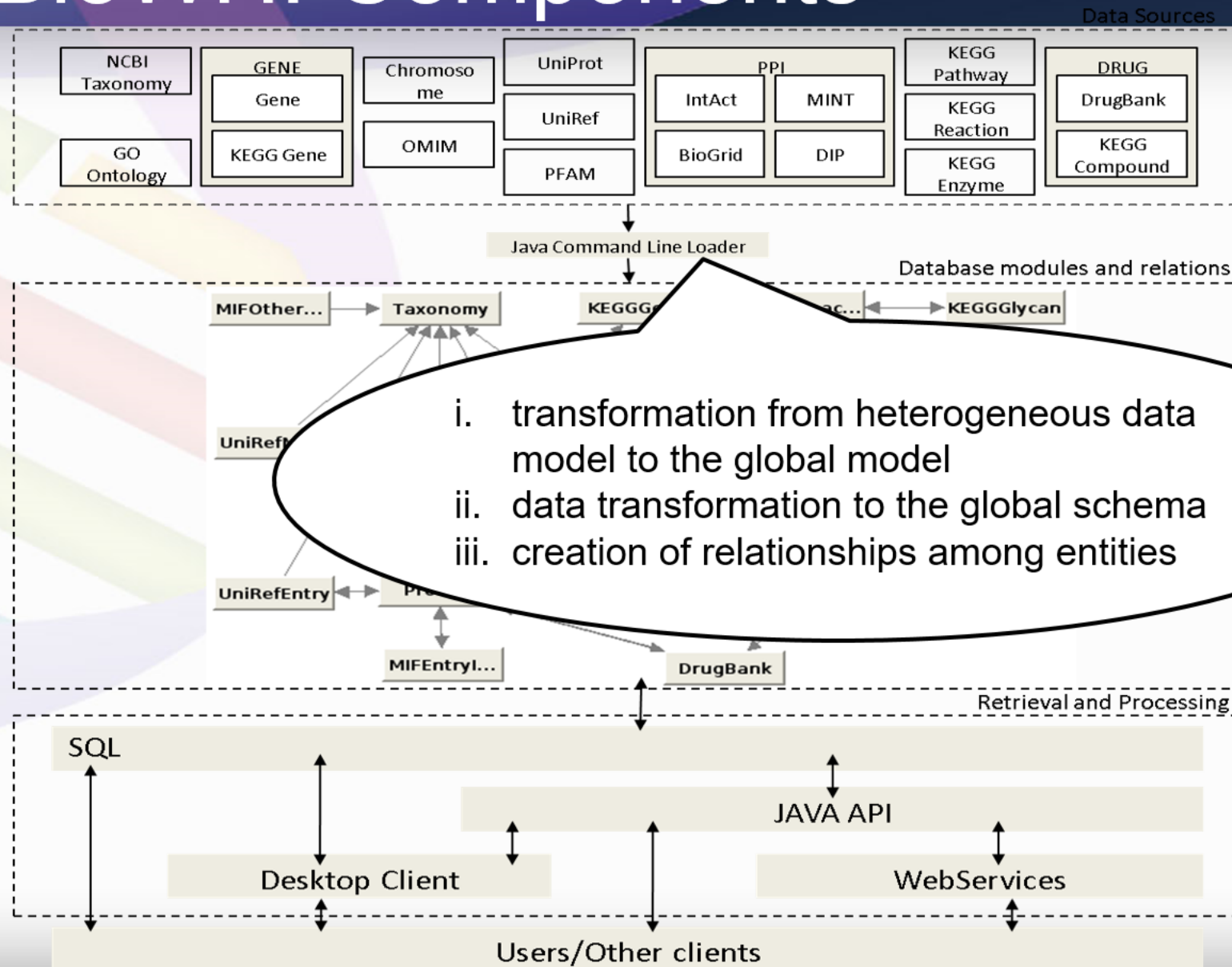
JBioWH: Components



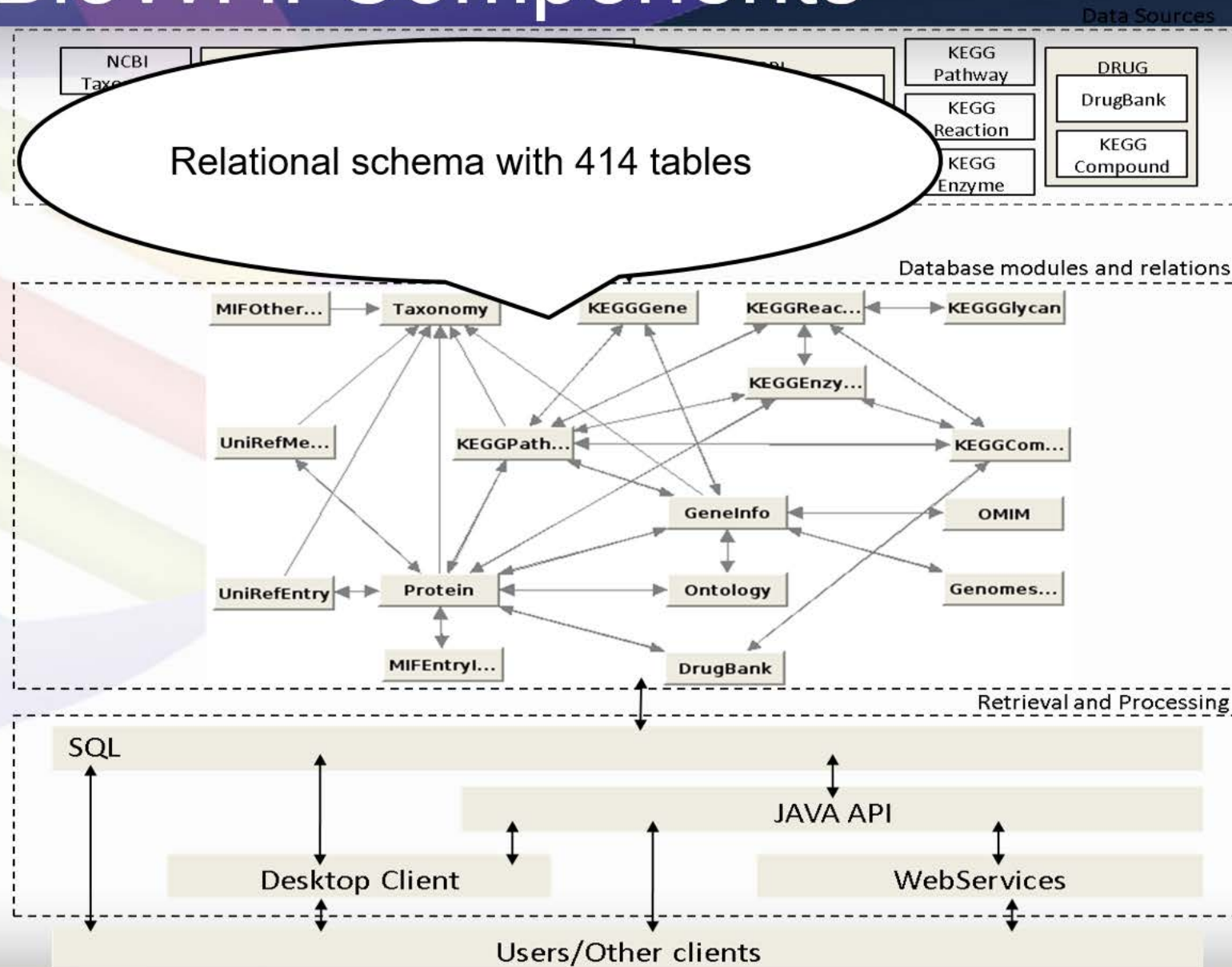
JBioWH: Components



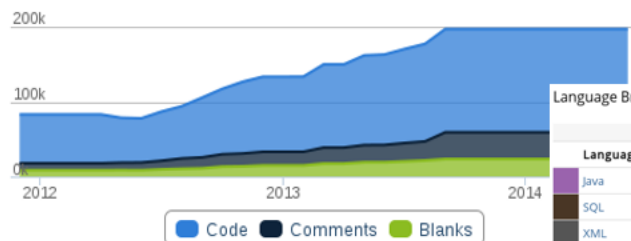
JBioWH: Components



JBioWH: Components



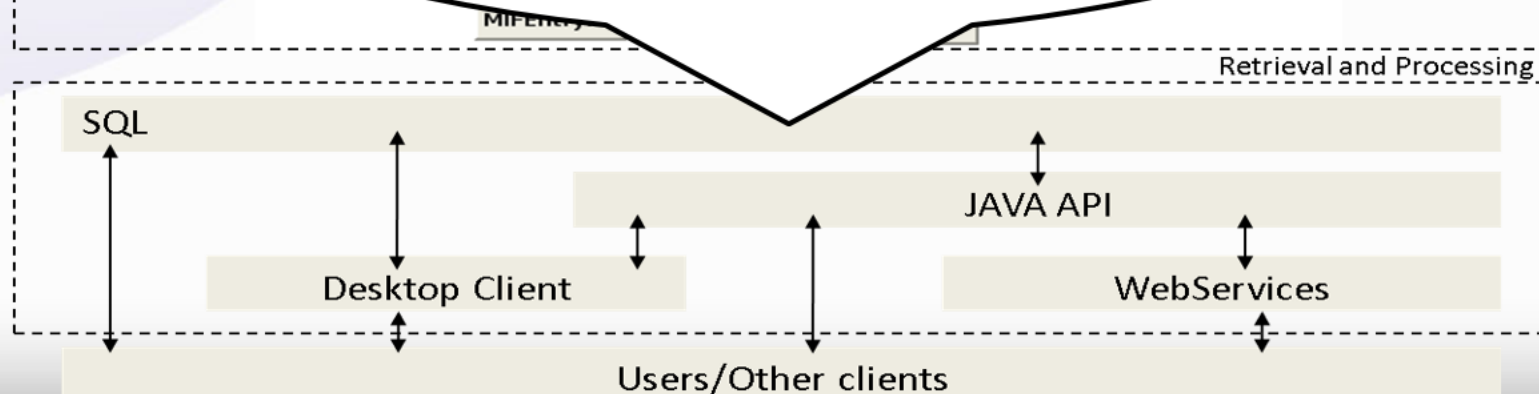
Lines of Code

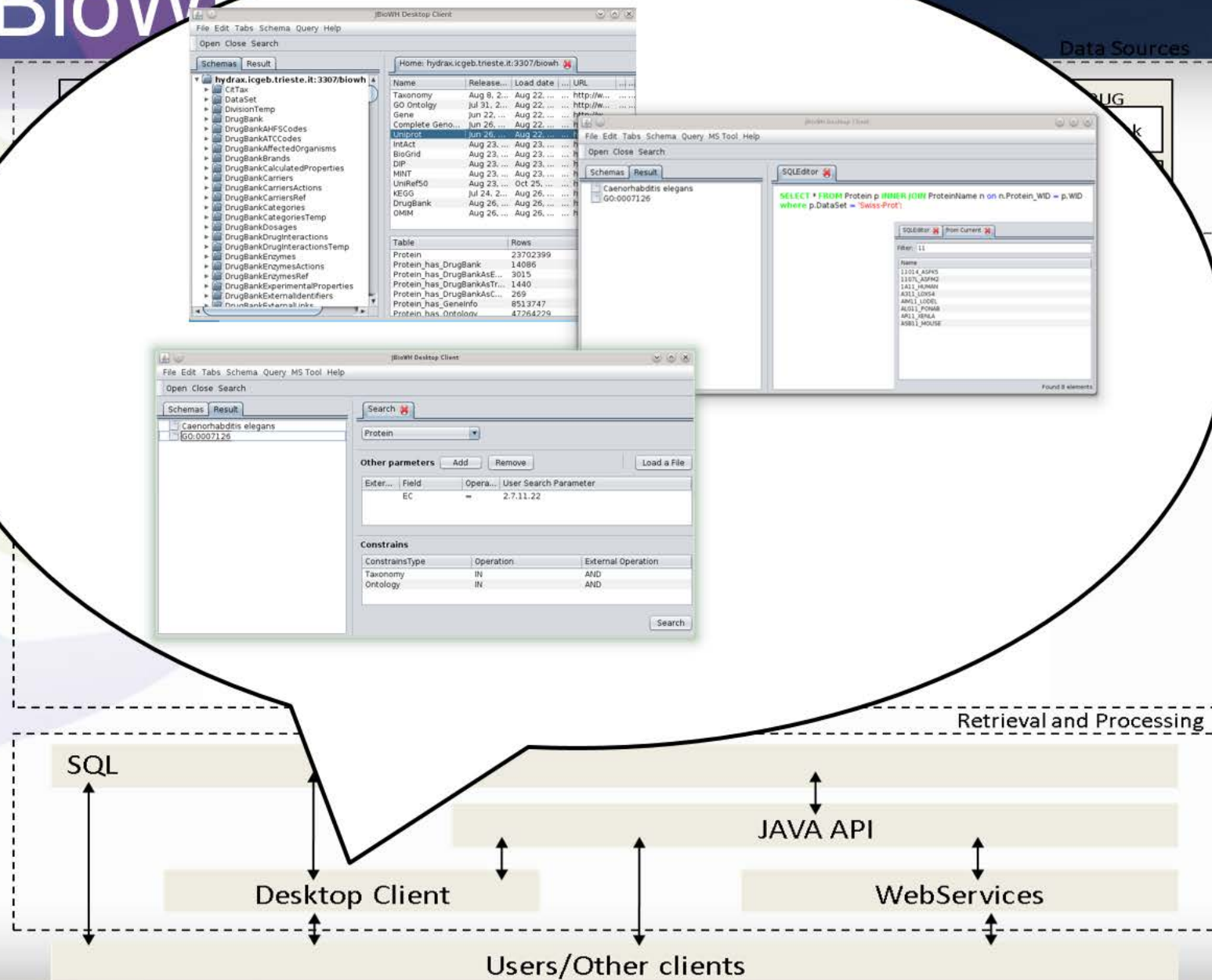


Language Breakdown

Language	Code Lines	Comment Lines	Comment Ratio	Blank Lines	Total Lines	Total Percentage
Java	96,124	26,121	21.4%	14,803	137,048	69.8%
SQL	35,954	8,954	19.9%	10,027	54,935	28.0%
XML	3,972	146	3.5%	138	4,256	2.2%
Totals	136,050	35,221		24,968	196,239	

Task	SQL solution	Java solution
Retrieve the protein sequence for the protein Q8DR59 from UniProt.	<pre>SELECT p.seq FROM Protein p INNER JOIN ProteinAccessionNumber a ON a.Protein_WID = p.WID WHERE a.AccessionNumber = 'Q8DR59';</pre>	<pre>JBioWSearch sProt = new SearchProtein();List prots = sProt.search("Q8DR59", null);for(Protein p : (List<Protein>) prots) System.out.println(p.getSeq());</pre>
Retrieve the protein sequence of all human proteins	<pre>SELECT p.seq FROM Protein p INNER JOIN Protein_has_Taxonomy pt ON pt.Protein_WID = p.WID INNER JOIN TaxonomySynonym ts ON ts.Taxonomy_WID = pt.Taxonomy_WID WHERE ts.Synonym like 'human';</pre>	<pre>JBioWSearch sTax = new SearchTaxonomy();JBioWSearch sProt = new SearchProtein();List taxes = sTax.search('human',null);List c = new ArrayList();List o = new ArrayList();c.add(taxes);o.add("IN");JPLCons trains constrain = new JPLConstraints(c,o,null);List prots = sProt.search("", constrain);for(Protein p : (List<Protein>) prots) System.out.println(p.getSeq());</pre>







Databases by processing stage

- Raw data
- Assembled data
- Machine-annotated data
- Validated data



Biological objects

- Microbiomes
- Cancer mutations



Baktérium-közösségek

a valós világban

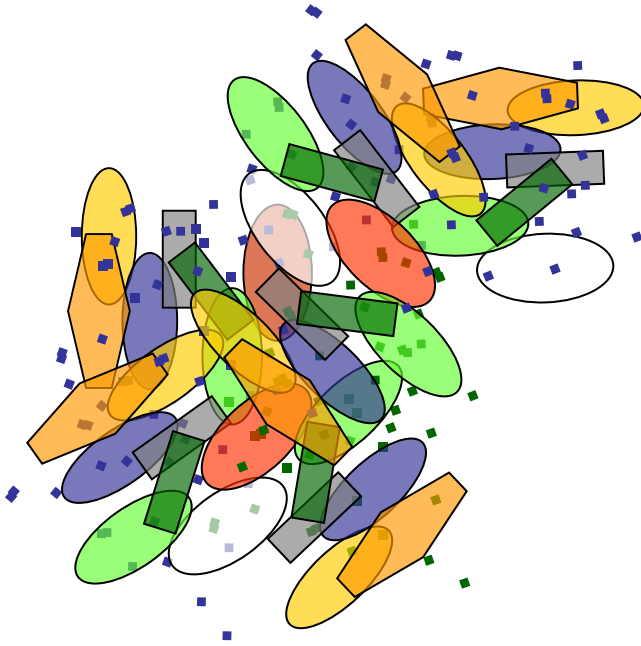
- Baktériumok, gombák, mindenféle egysejtűek
- A Föld ökoszféra legfontosabb elemei
- Sokféle sejt, melyek kommunikálnak és kooperálnak.
- Érdekes analógiák az emberi társadalmakkal: keletkezés, növekedés, összeomlás
- Analógia a mérnöki rendszerekkel: kommunikáció, kooperáció, összeomlás



Miért stabilisak?
(Miért léteznek?)

Baktérium-közösségek

– logikai modell

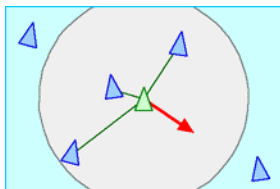


**Többféle sejt a
kibocsátott
kismolekulák
felhőjében**

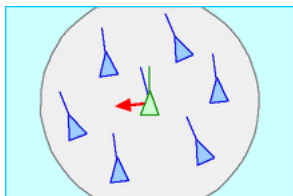
- A sejtek (és a fajok) kommunikálnak és kooperálnak
- Sok kismolekulát bocsátanak ki (antibiotikumok, jelanyagok, tápanyagok)
- Együttes feladatmegoldás
- Kollektív tulajdonságok



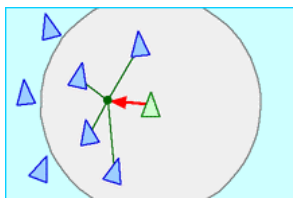
Virtuális világ: 1) “ágensek”



Szeparáció: ne ütközz!



Igazodás: kövesd szomszédodat

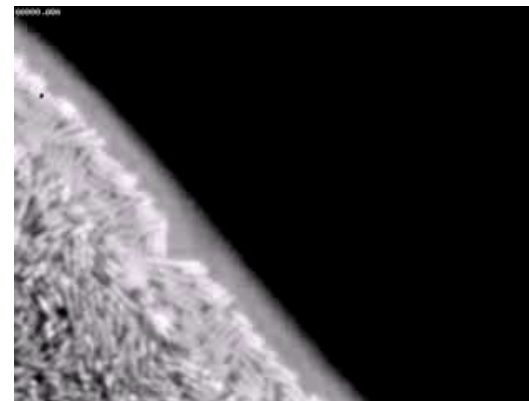


Összetartás: Kövesd a falkát (vagy egy külső parancsot)

**Lokális kölcsönhatás
globális reakciók**



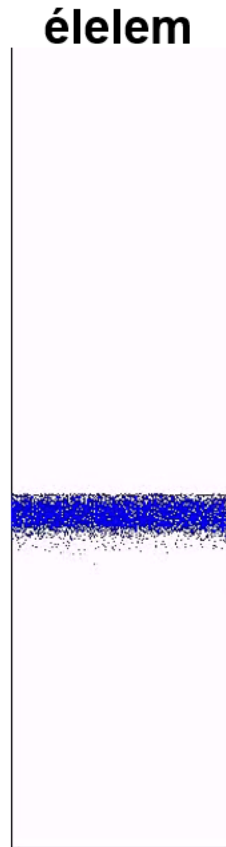
Craig Reynolds, 1982



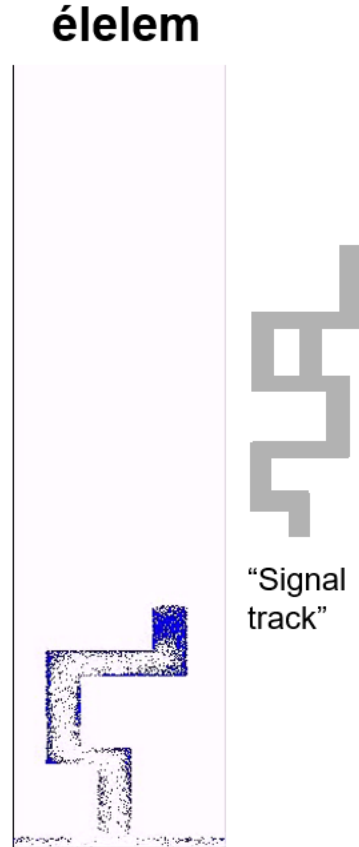
Howard Berg, 1993



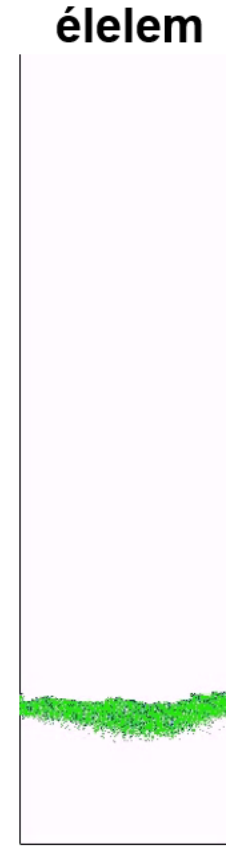
A kommunikáció új közösségi tulajdonságokat hoz létre!



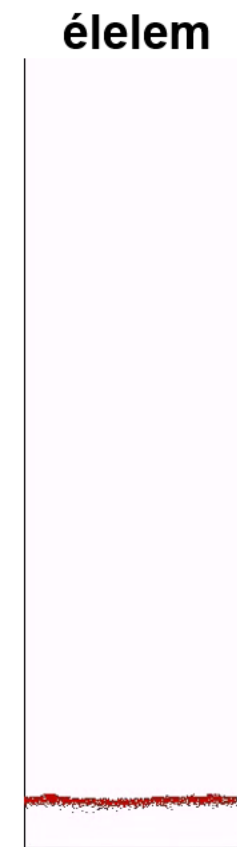
**Közösség
formálás**



**Jelkövetés=
migráció**



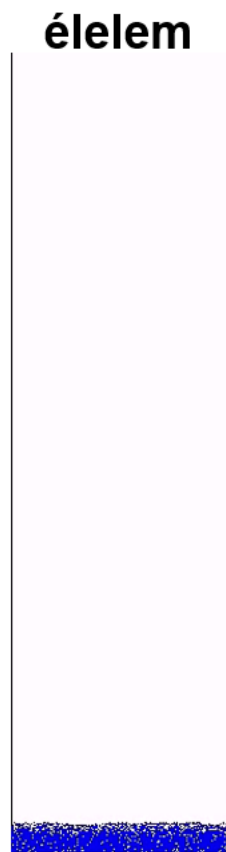
**Normál + nem-
kommunikáló
sejtek =
társadalmi
rétegződés**



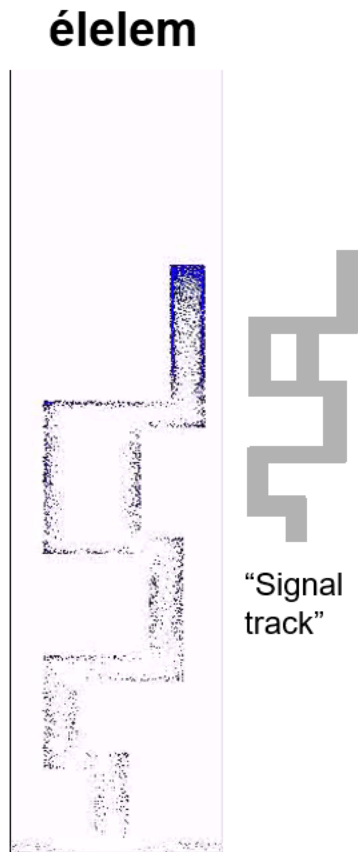
**Normál + nem-
kooperáló sejtek=
összeomlás**



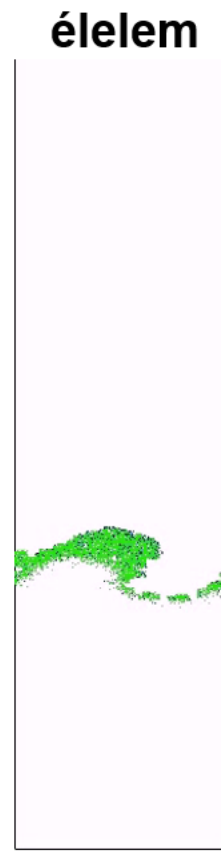
A kommunikáció új közösségi tulajdonságokat hoz létre!



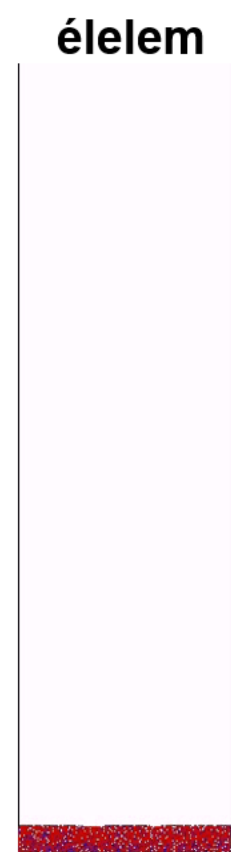
**Közösség
formálás**



**Jelkövetés=
migráció**



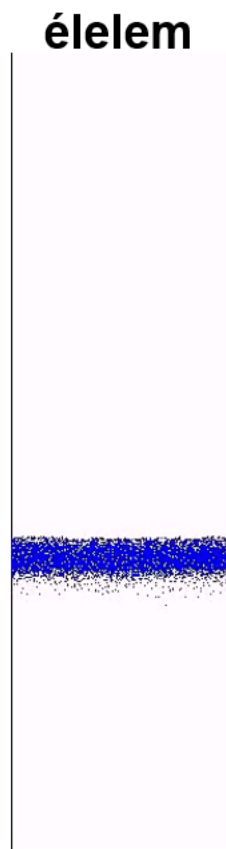
**Normál + nem-
kommunikáló
sejtek =
társadalmi
rétegződés**



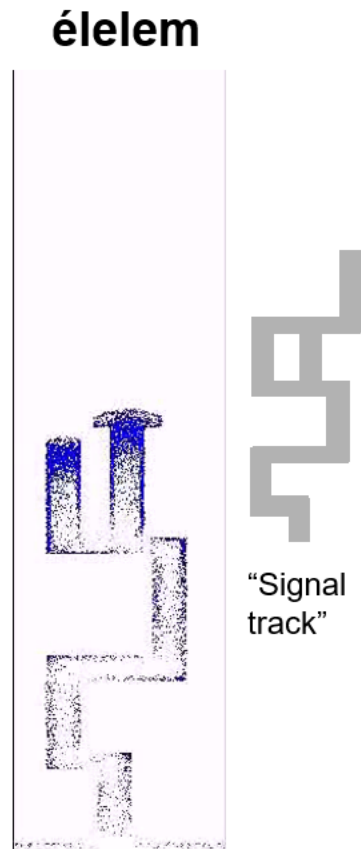
**Normál + nem-
kooperáló sejtek=
összeomlás**



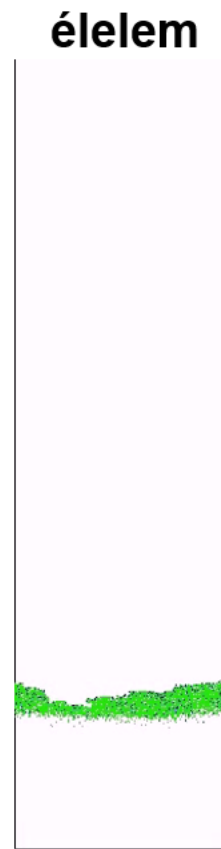
A kommunikáció új közösségi tulajdonságokat hoz létre!



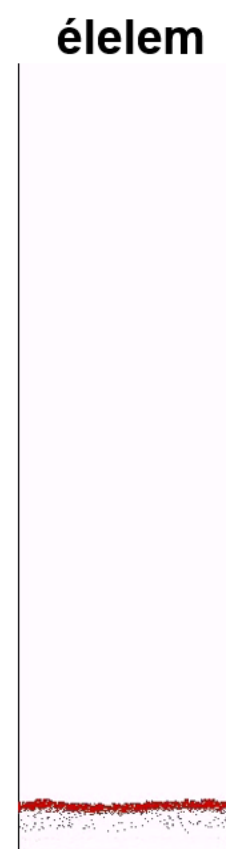
**Közösség
formálás**



**Jelkövetés=
migráció**



**Normál + nem-
kommunikáló
sejtek =
társadalmi
rétegződés**



**Normál + nem-
kooperáló sejtek=
összeomlás**



Köszönöm a figyelmet!

- Ligeti Balázs
- Juhász János
- Bihary Dóra
- Gelencsér Zsolt
- Kerényi Ádám