Gene regulation: Databases and Integration

Ralf Hofestädt

AG Bioinformatics / Medical Informatics

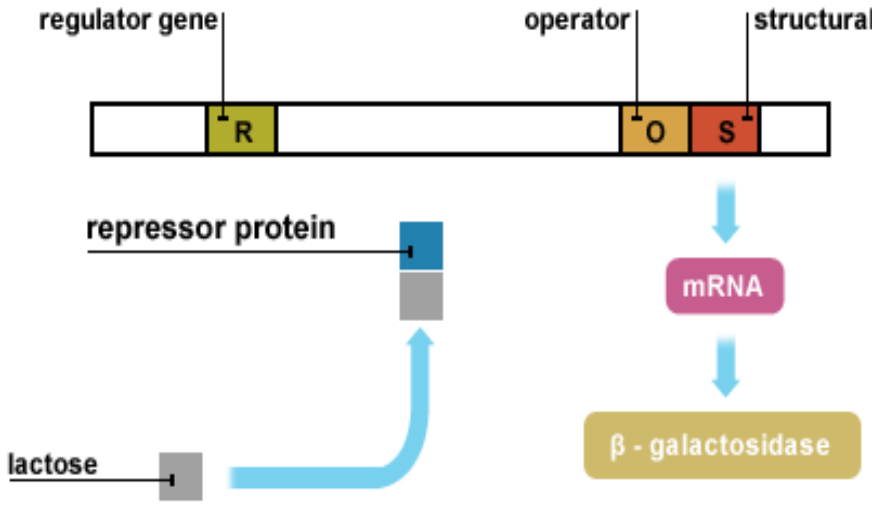
http://cweb.uni-bielefeld.de/agbi/home/index.cw

Göttingen, 7.3.2018

1. Motivation
2. Databases
3. Integration
4. Application
5. Discussion

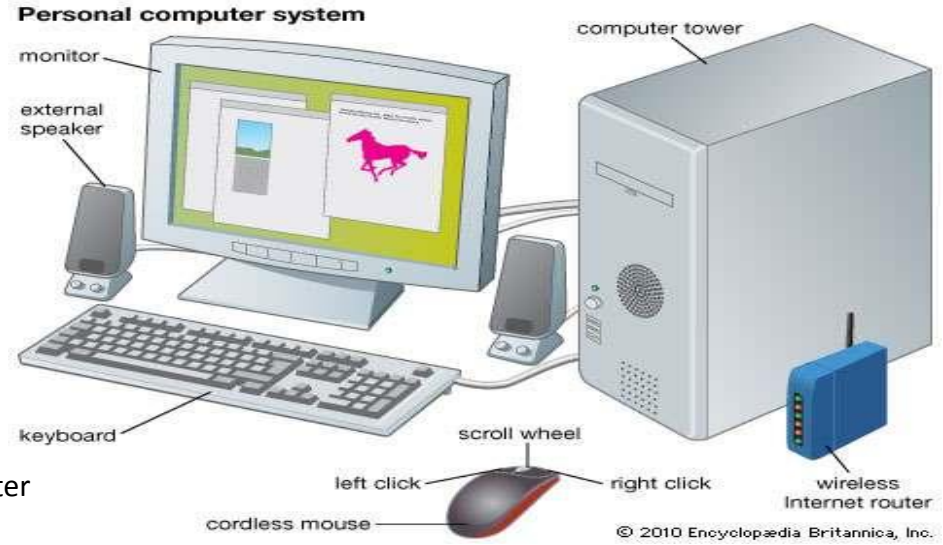
My own startpoint … school (*long time ago*) …

a) **Gene regulation** … Jacob & Monod … easy to understand ! …



http://www.bbc.co.uk/bitesize/higher/biology/control_regulation/genetic_control/revision/1/

b) **Computer** … how to use ?

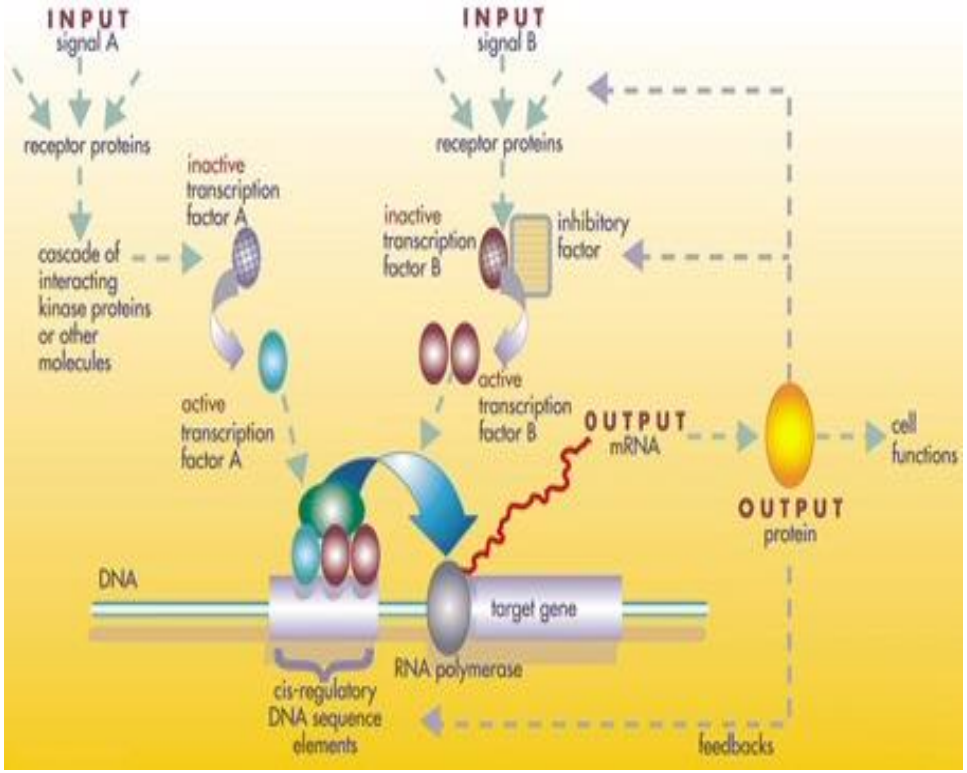


https://www.britannica.com/technology/computer

My vison - during that time:

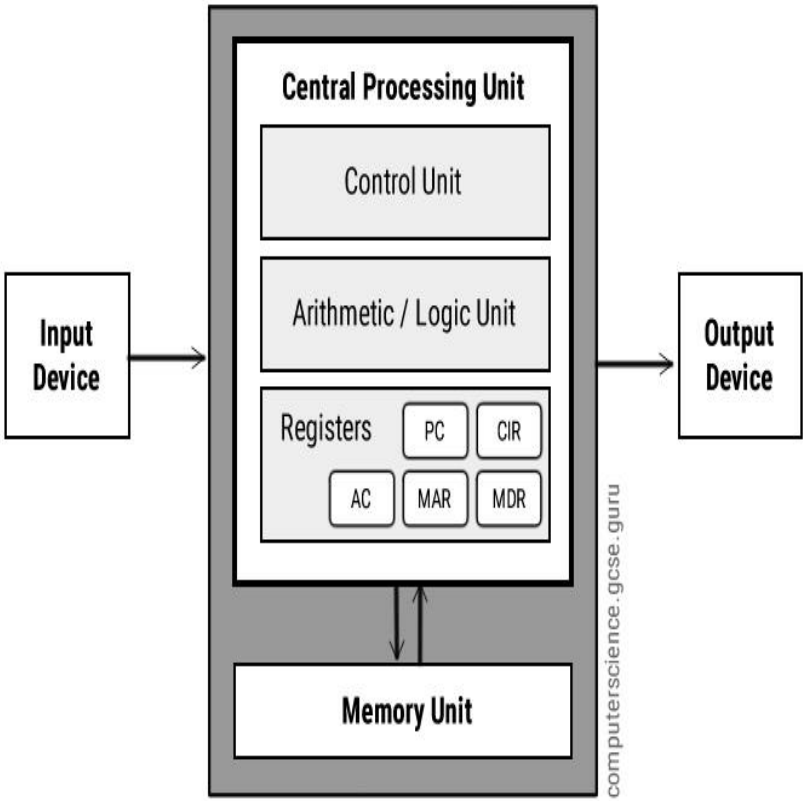
try to understand

a) gene controlled processes.



https://en.wikipedia.org/wiki/Gene_regulatory_network

b) function of the computer

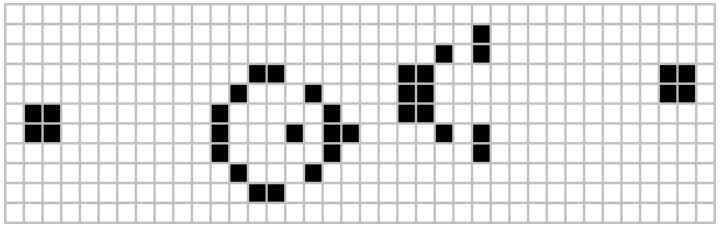


https://www.computerscience.gcse.guru/theory/von-neumann-architecture

Good luck

I could study „bioinformatics“ as a member of the department of theoretical computer science at the University Bonn (Prof. Dr. Böhling)

Why was it possible - during that time ??



https://en.wikipedia.org/wiki/Conway%27s_Game_of_Life

Motivation:

John von Neuman
Cellular automata to model neuronal processes

= *first parallel model of computation* !

A. Lindenmayer
L-Systems to model cell differentiation processes

= *first parallel grammar*

Genetic algorithms
Rechenberg, Schwefel and Holland

= Approximative method to solve hard problems …

and so on…

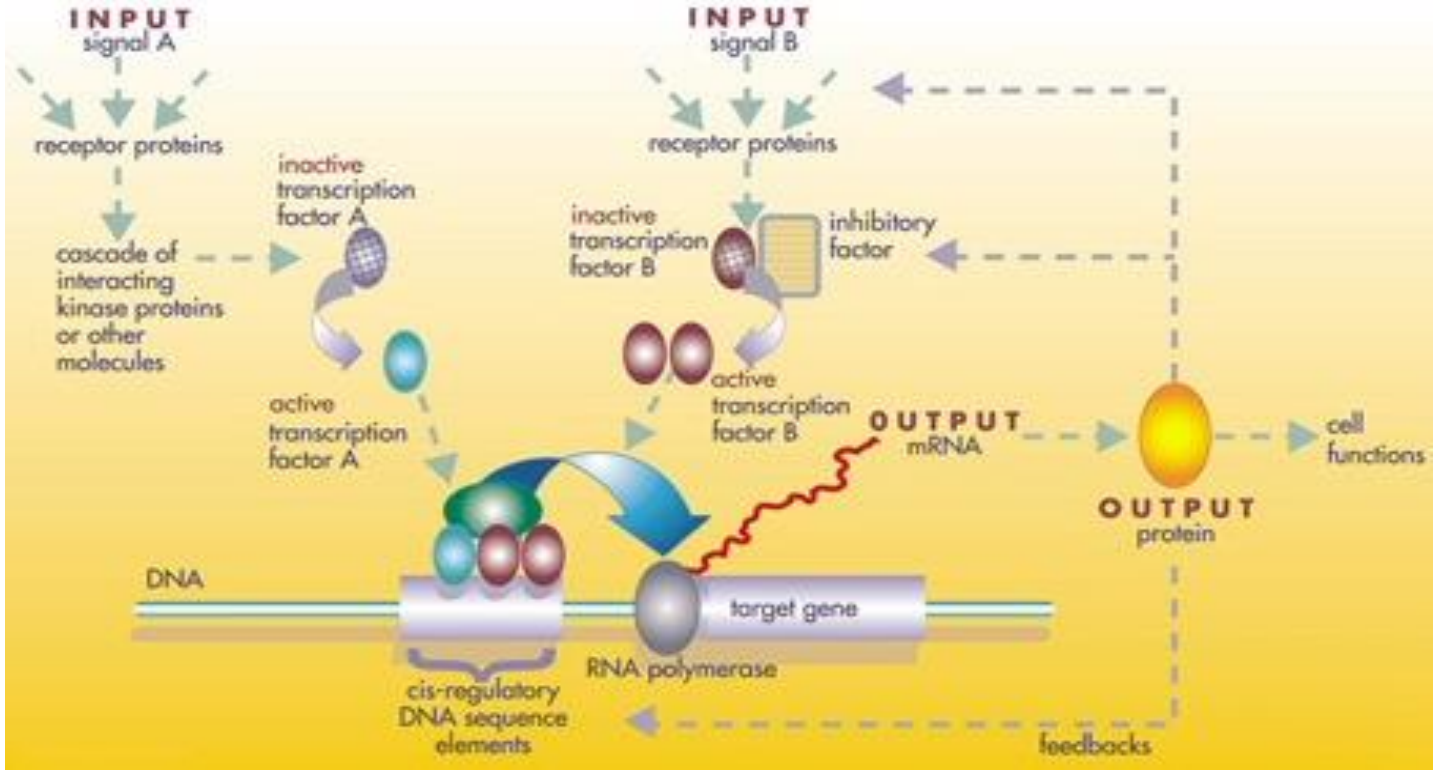
https://www.vexlio.com/blog/drawing-simple-organics-with-l-systems/

Innovations direct and indirect from Bioligy !

Diploma and Dissertation (University Bonn)

To model cell processes using formal languages … biochemical computer….

Cell a chemical machine ?



https://en.wikipedia.org/wiki/Gene_regulatory_network

Rule system = 5-tuple (B, A, E, I, p)

Rule based system to model the cell as a chemical machine …

One important motivation was …

V. Ratner
Molekulargenetische Steuerungssysteme.
Gustav Fischer Verlag, Stuttgart 1977.

DNA-interpretation as a program …

Features of the biochemical machine:

Rule based modeling showed:

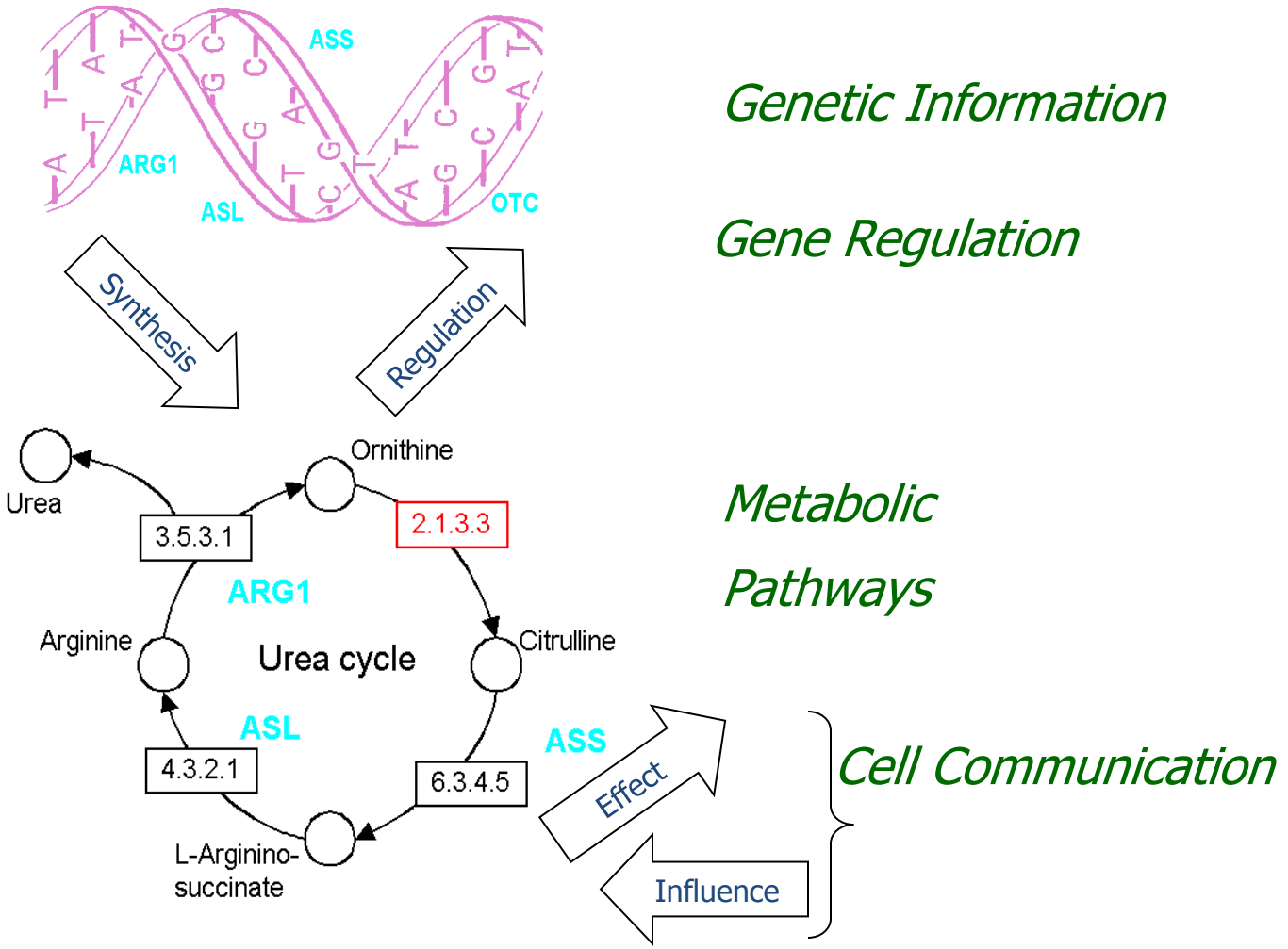
Features:

Dataflow

Parallelism

Probability

Modularity



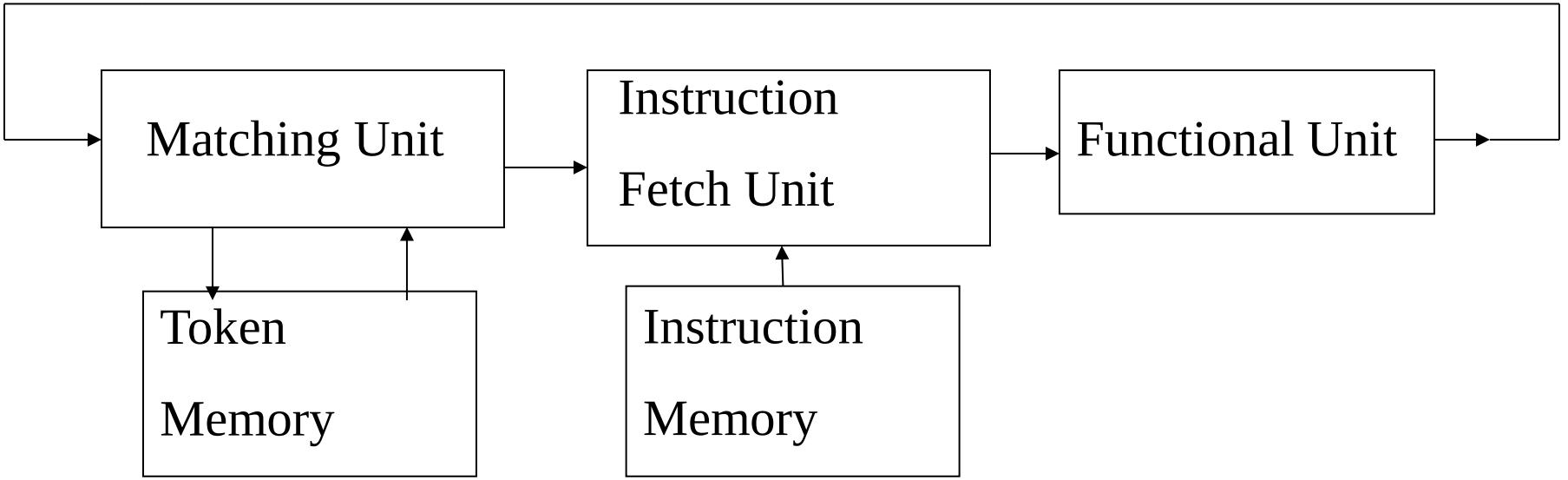
Genetic Information

Gene Regulation

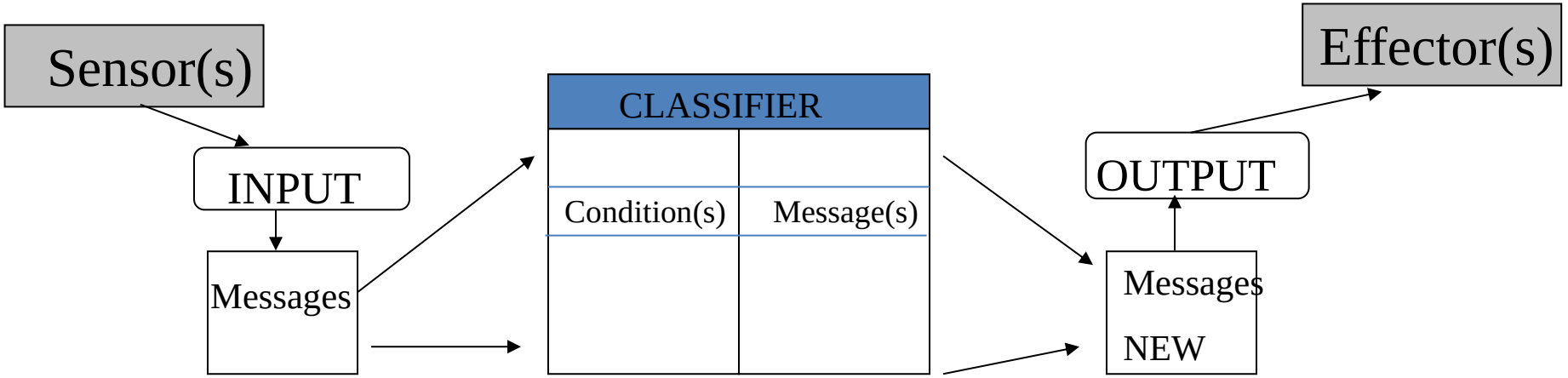
Metabolic Pathways

Cell Communication

Data flow computing (basic concept)



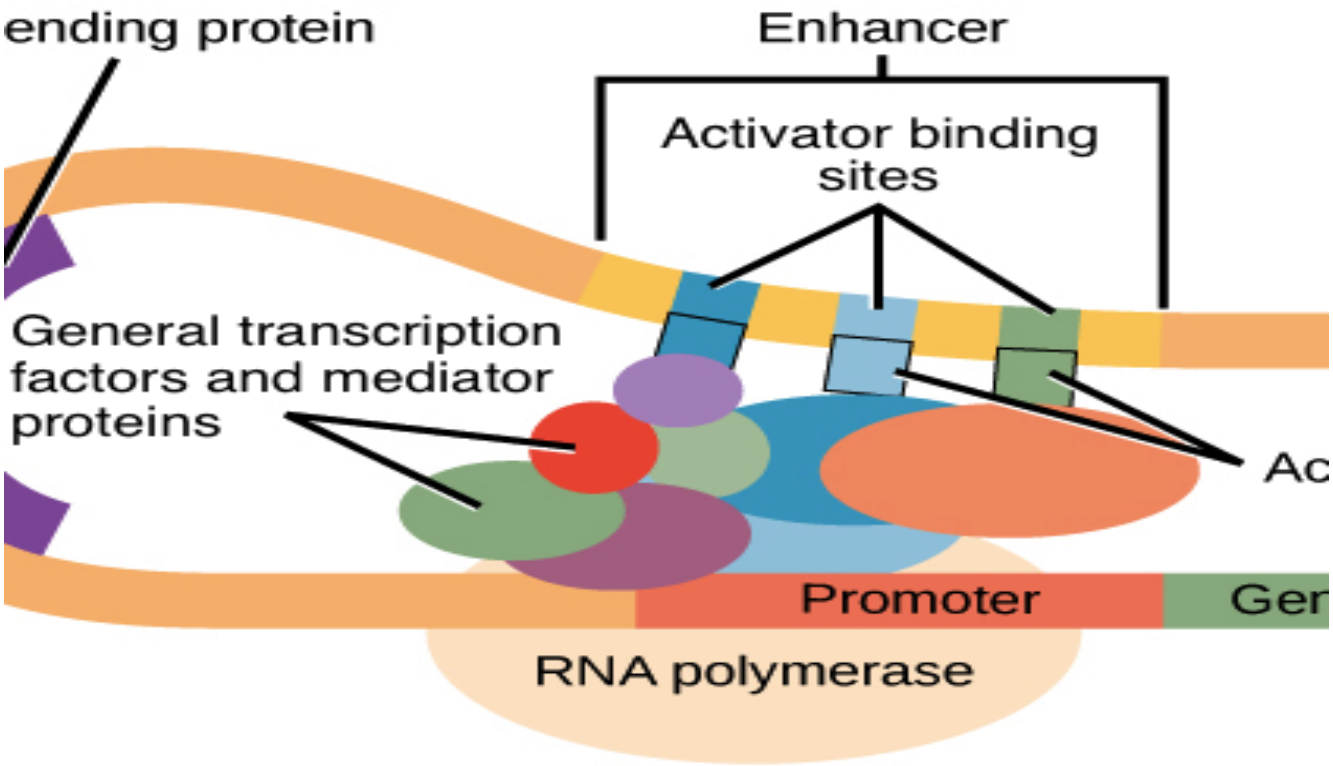
Artificial Intelligence: *Classifier System*



BMBF meeting in Braunschweig (GBF 1989 ?)

Edgar Wingender – **transcrition factors**

I had to learn: Gene regulation is very complicated

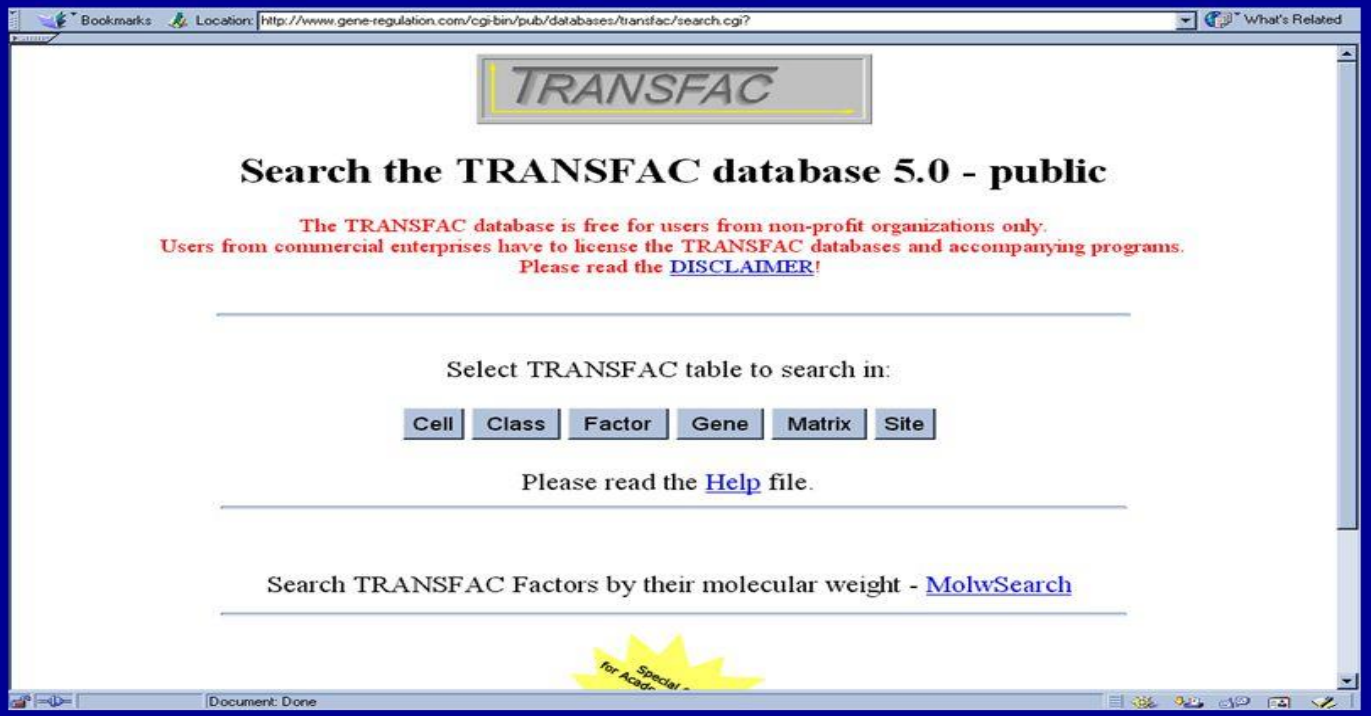


https://scistart.co/articles/482-method-for-predicting-gene-expression-by-modeling-transcription-factor-activity.html

… good news: we have Edgar and other fellows … taking care and make the information available via ***databases*** etc. …

TRANSFAC

http://transfac.gbf.de/TRANSFAC/index.html



http://slideplayer.it/slide/5378126/

Hundreds of data- and information systems (heterogenous !)

Genes

EMBL	http://www.ebi.ac.uk/

.....

Proteins and Enzymes

PDB	http://www.rcsb.org/pdb/
SWISSPROT	http://www.expasy.org/sprot/

.....

Pathways

KEGG	http://www.genome.ad.jp/

.....

Gene Regulation

TRANSFAC	http://www.biobase.de

.....

Metabolic Diseases

OMIM	http://www3.ncbi.nlm.nih.gov/

.....

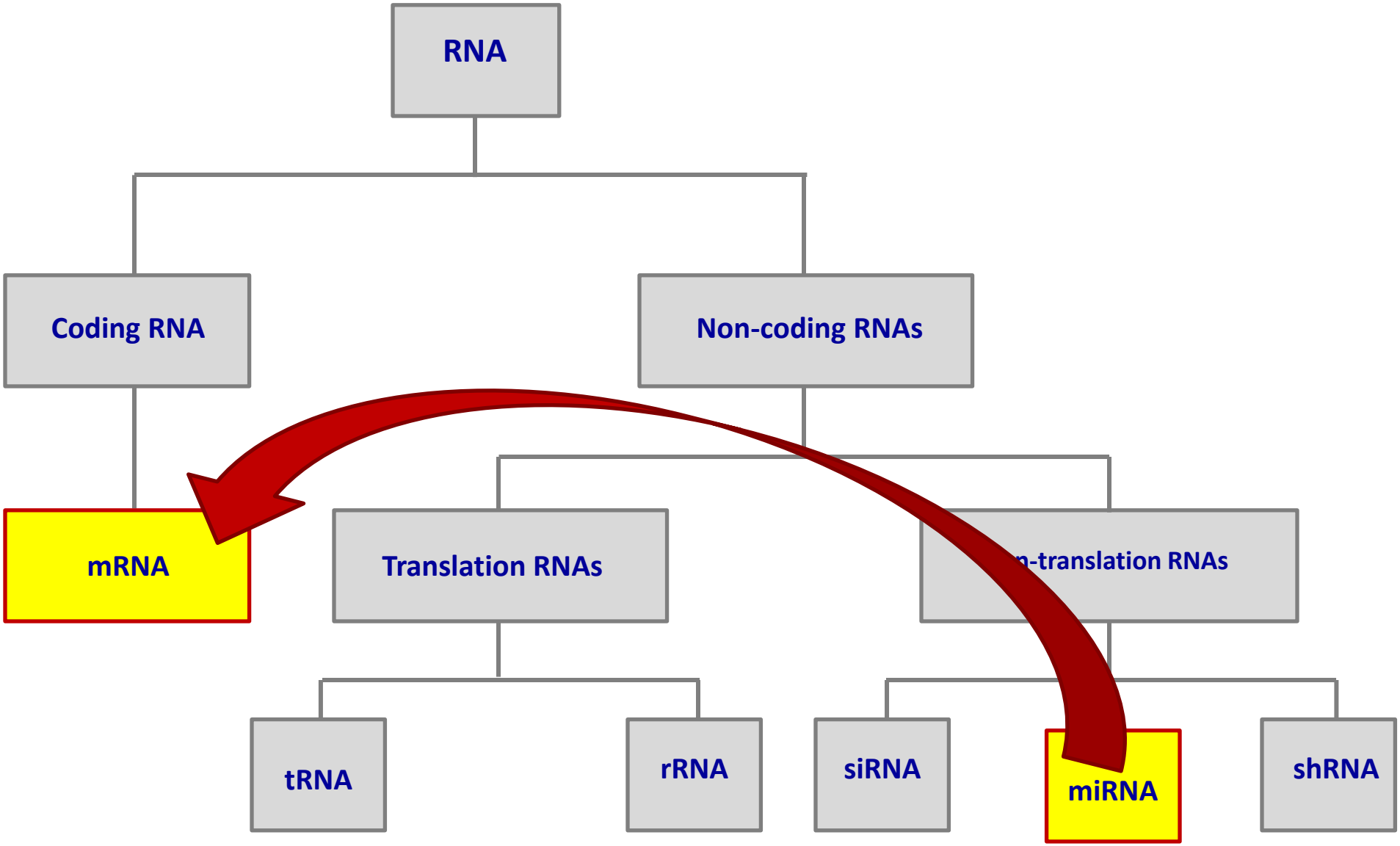
Drugs

DrugBank	http://www.drugbank.ca/

.....

RNA World

Who can understand gene regulation ?



INTERNET showed up … data available 24 hours - worldwide

http://dheoalfarisi87.blogspot.de/2013/09/definition-internet-and-intranet.html

User requirements:

user friendly access … (integration of heteronegous databases) and analysis tools.

1993 ? – Bioinformatics conference, Talahassee, USA

1995 First Dagstuhl seminar - topics:

Molecular databases
BRENDA, TRANSFAC,
KEGG, …

Integration

Modeling and simulation

Metabolic Engineering
Stephanopoulus

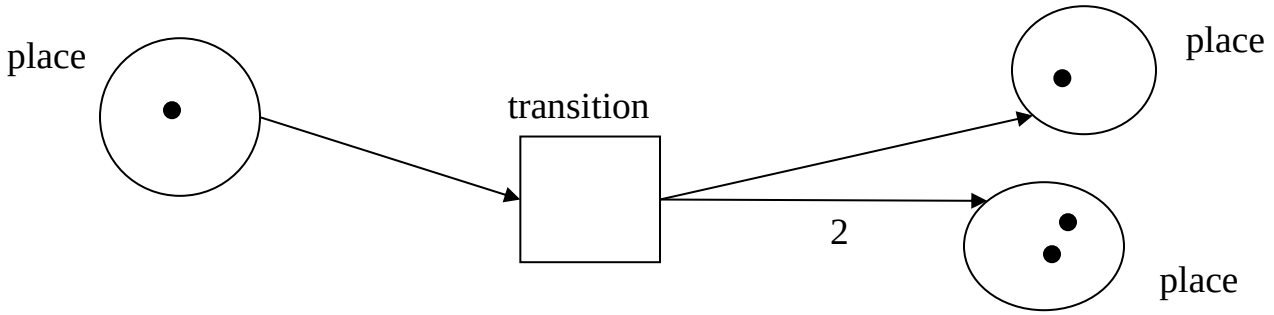
Edgars connection - first contact - Prof. Kolchanov (Novosibirsk).

I started to use Petri-nets … advantages:

a strong theory and

simulation tools are available.

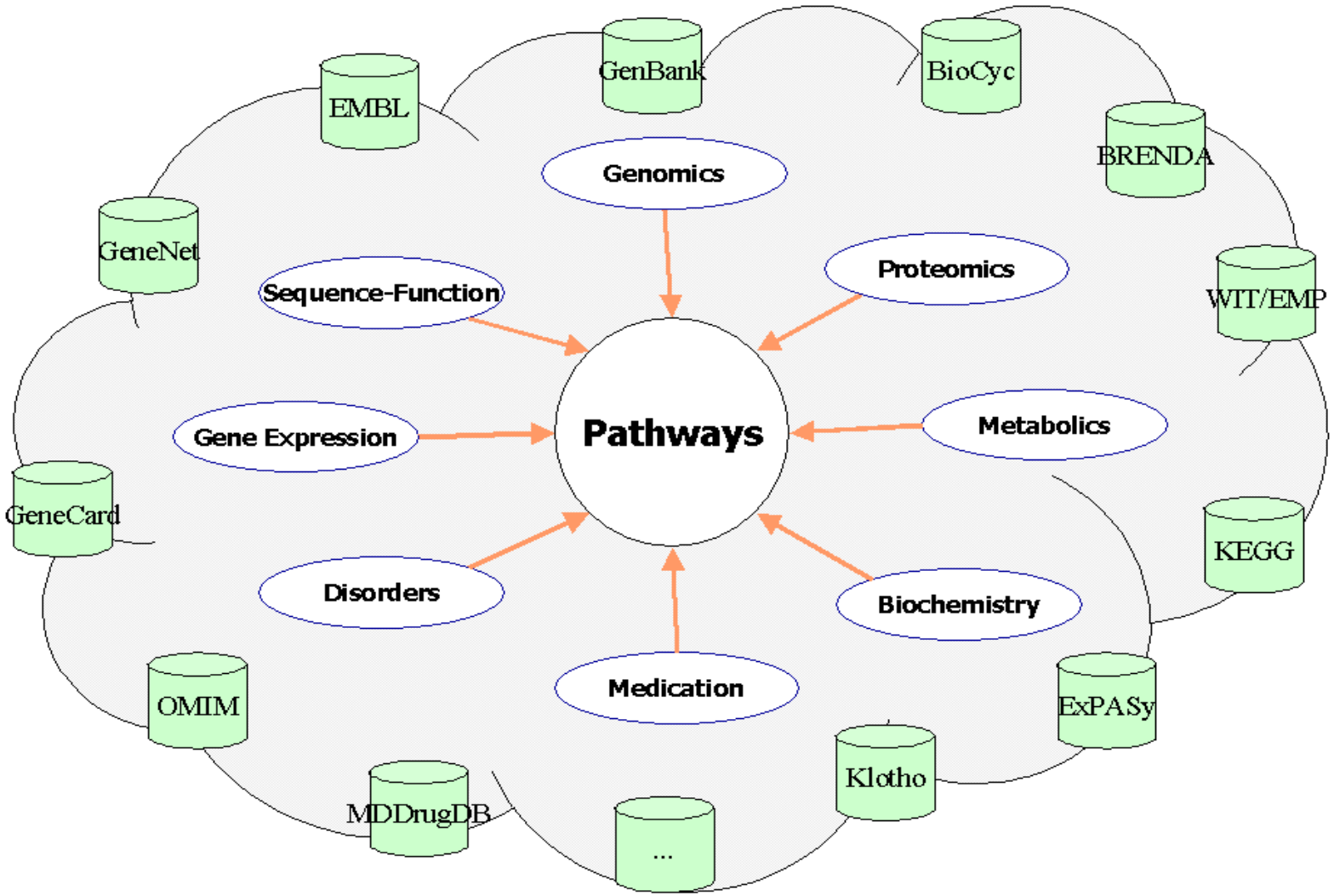
Petri-Net - introduced by C.A. Petri in 1962.



Applications:

Office automation, work-flows, flexible manufacturing, programming languages, protocols and networks, hardware structures, real-time systems, performance evaluation, operations research, embedded systems, defence systems, telecommunications, Internet, e-commerce and trading, railway networks, biological systems.

User specific integration



Database Integration

SRS (Lion Bioscience)

PEDANT (BioMax)

MOBY DICK

HUSAR

BioKleisli

What Is There (WIT)

Biology Workbench

Integrated Genomic Database (IGD)

...

Integrative Bioinformatics

Views to (complex) data

- *Visualization*
- *Statistic*
- *Analysis Algorithms …*

Data Mining / Information Fusion

- *Integration of database systems and analysis tools*

Metabolic Network Control

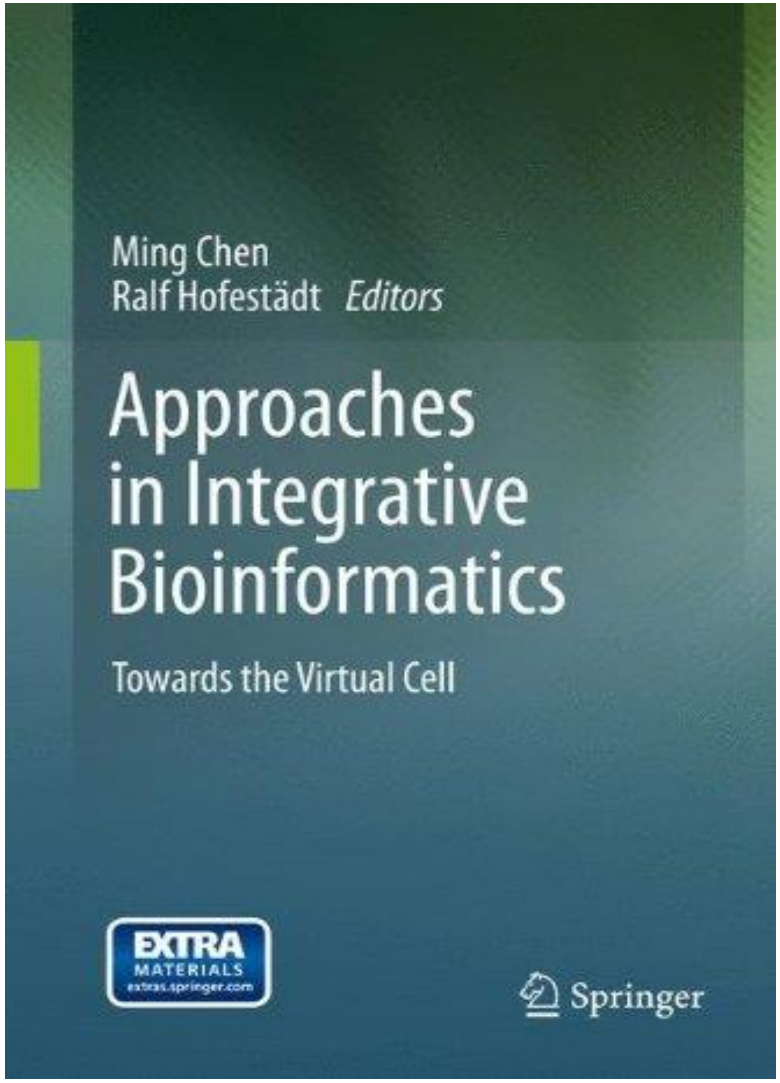
Integration of

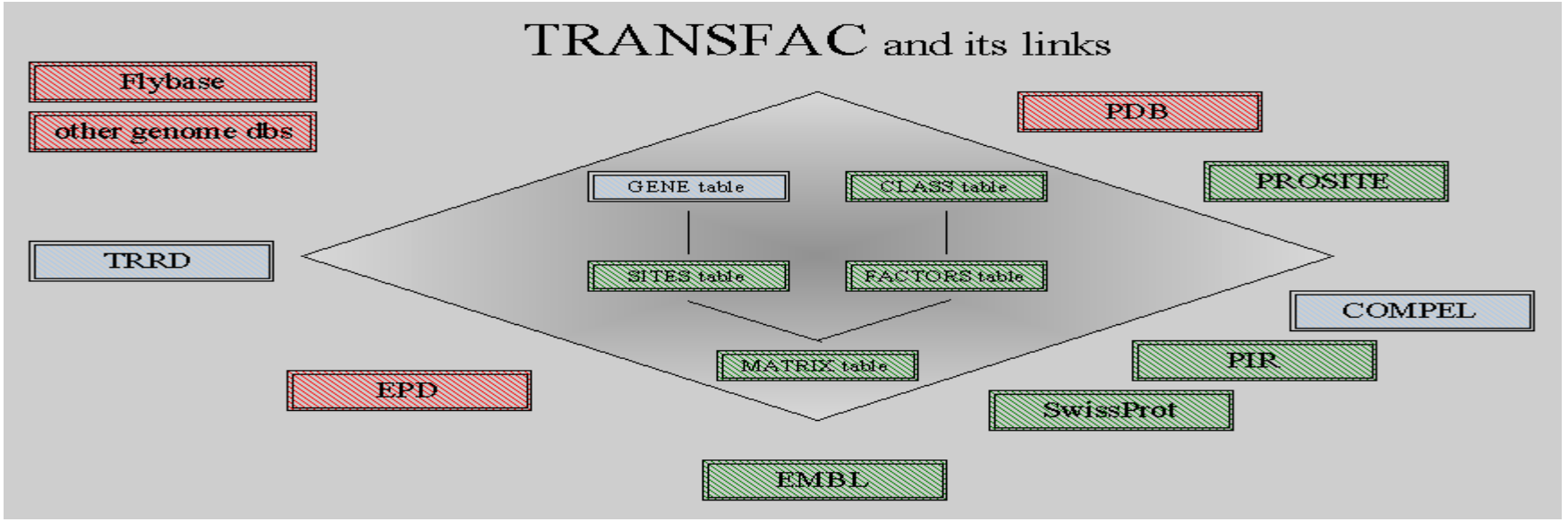
Database systems and

Analysis tools

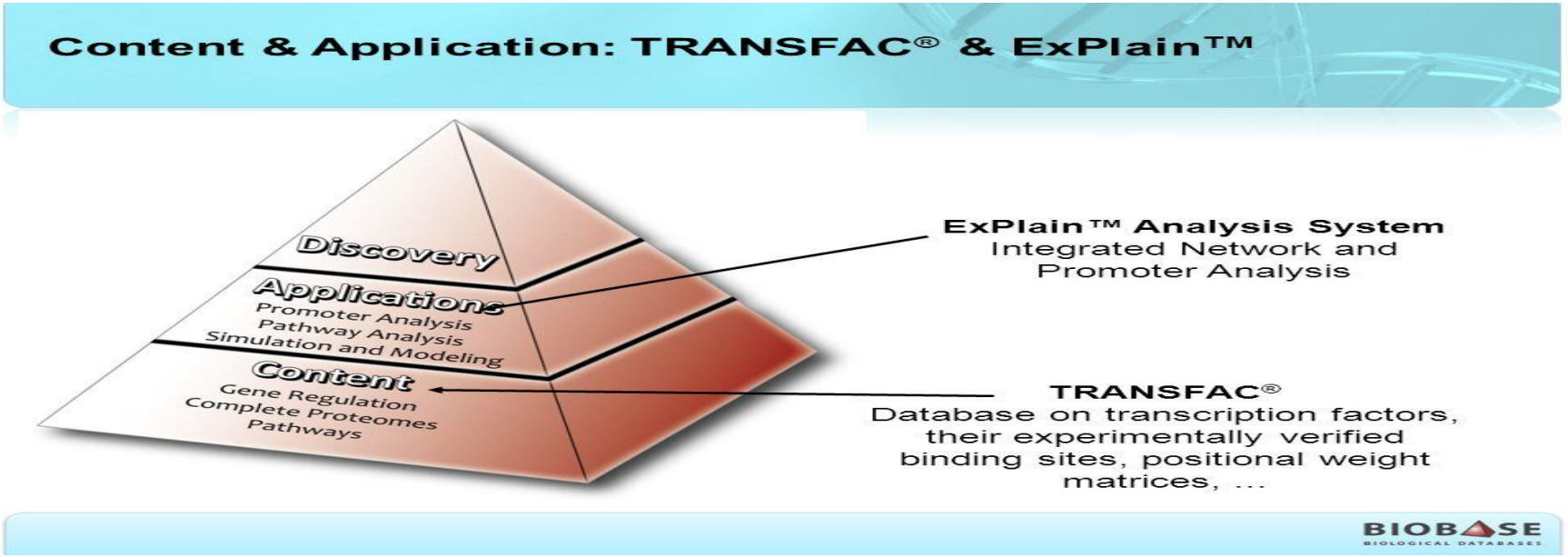
https://www.degruyter.com/view/j/jib

Open access – founded 2004



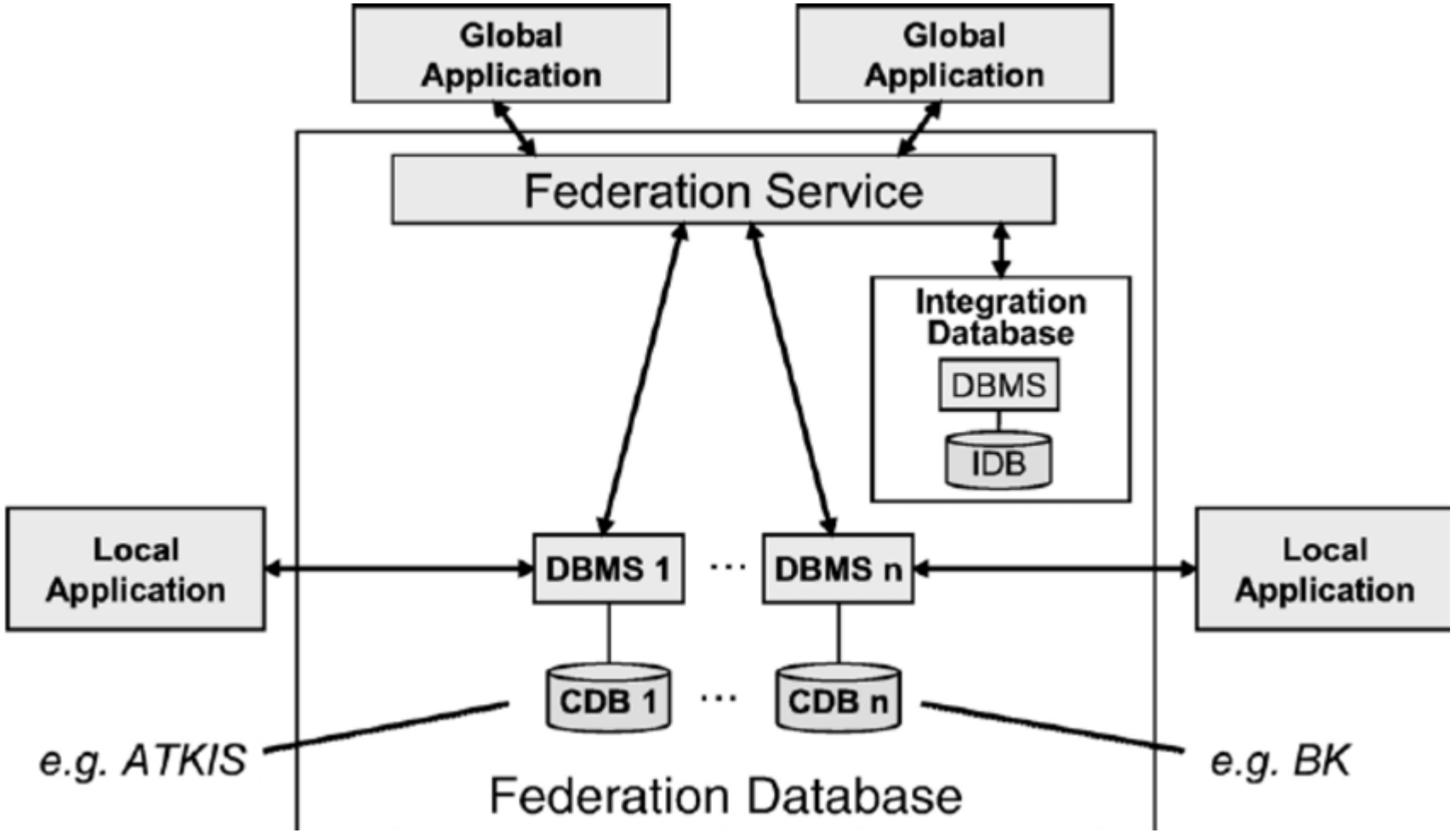


http://www.ai.sri.com/pkarp/mimbd/95/abstracts/karas.html



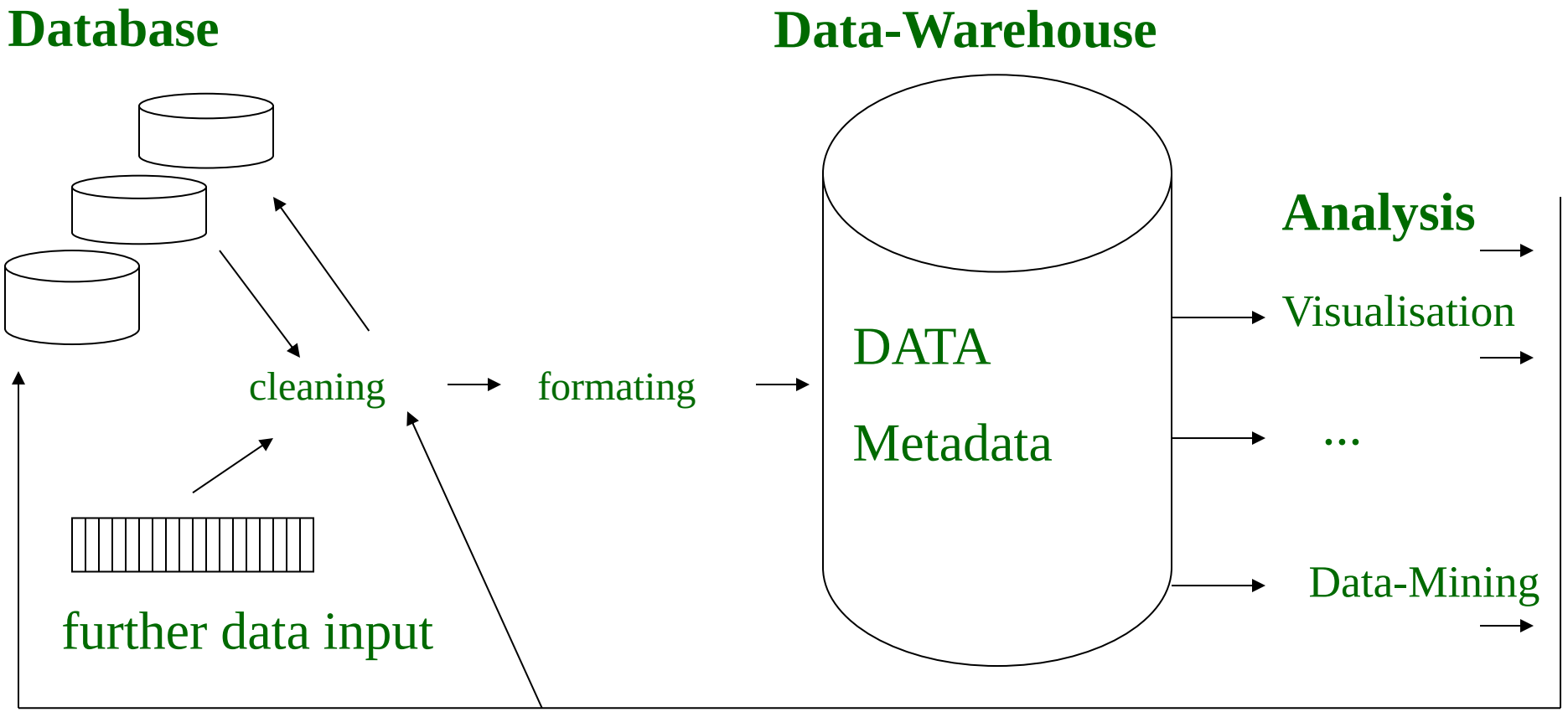
http://slideplayer.com/slide/10838589/

Federated Integration - MARK



https://www.researchgate.net/figure/222418660_fig2_Fig-2-Architecture-of-a-federated-database

Bio Data-Warehouse



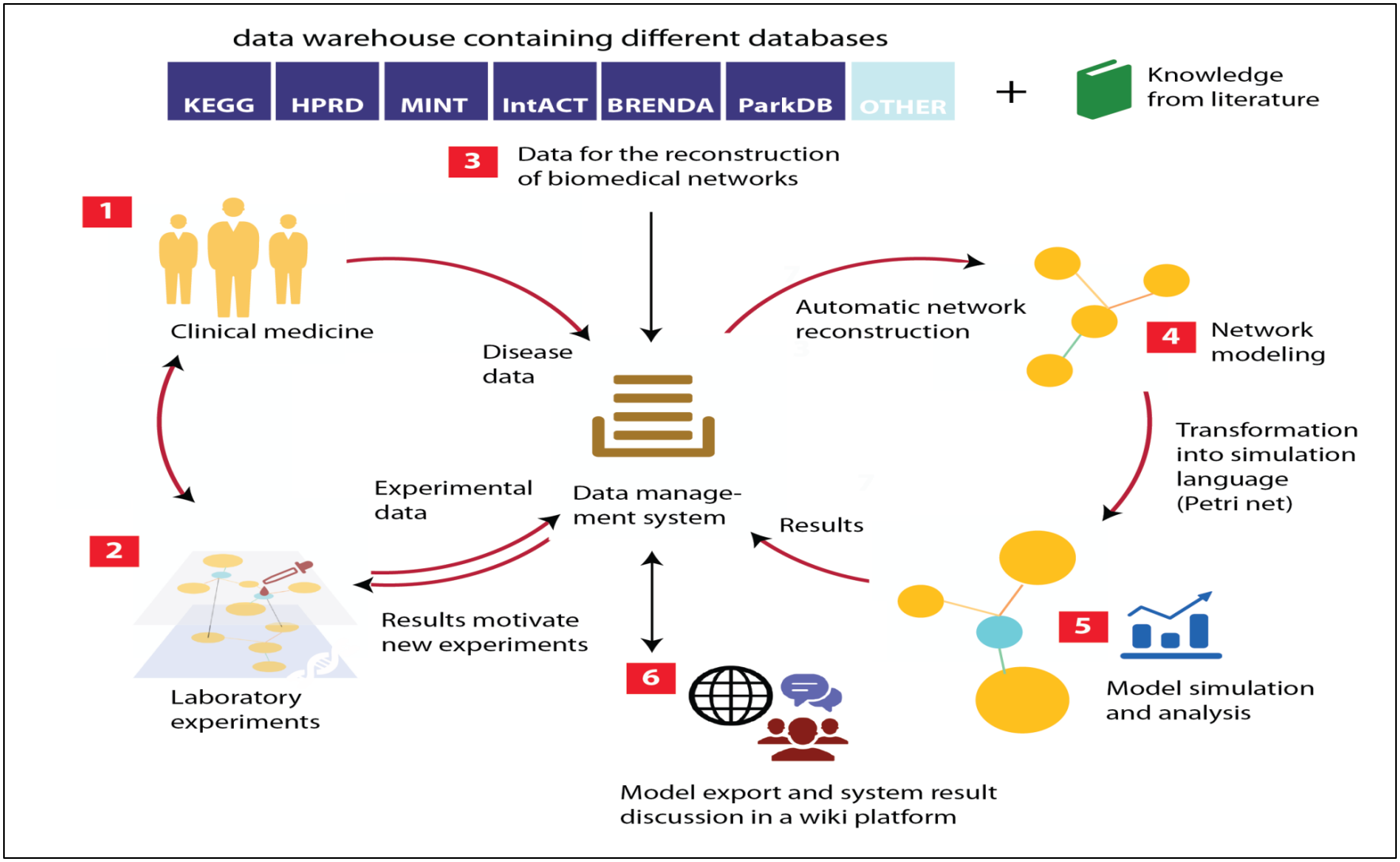
BioDWH: Bioinformatics Data Warehouse

http://sourceforge.net/projects/biodwh/

Kormeier, Hippe, Toepel (Bielefeld University)

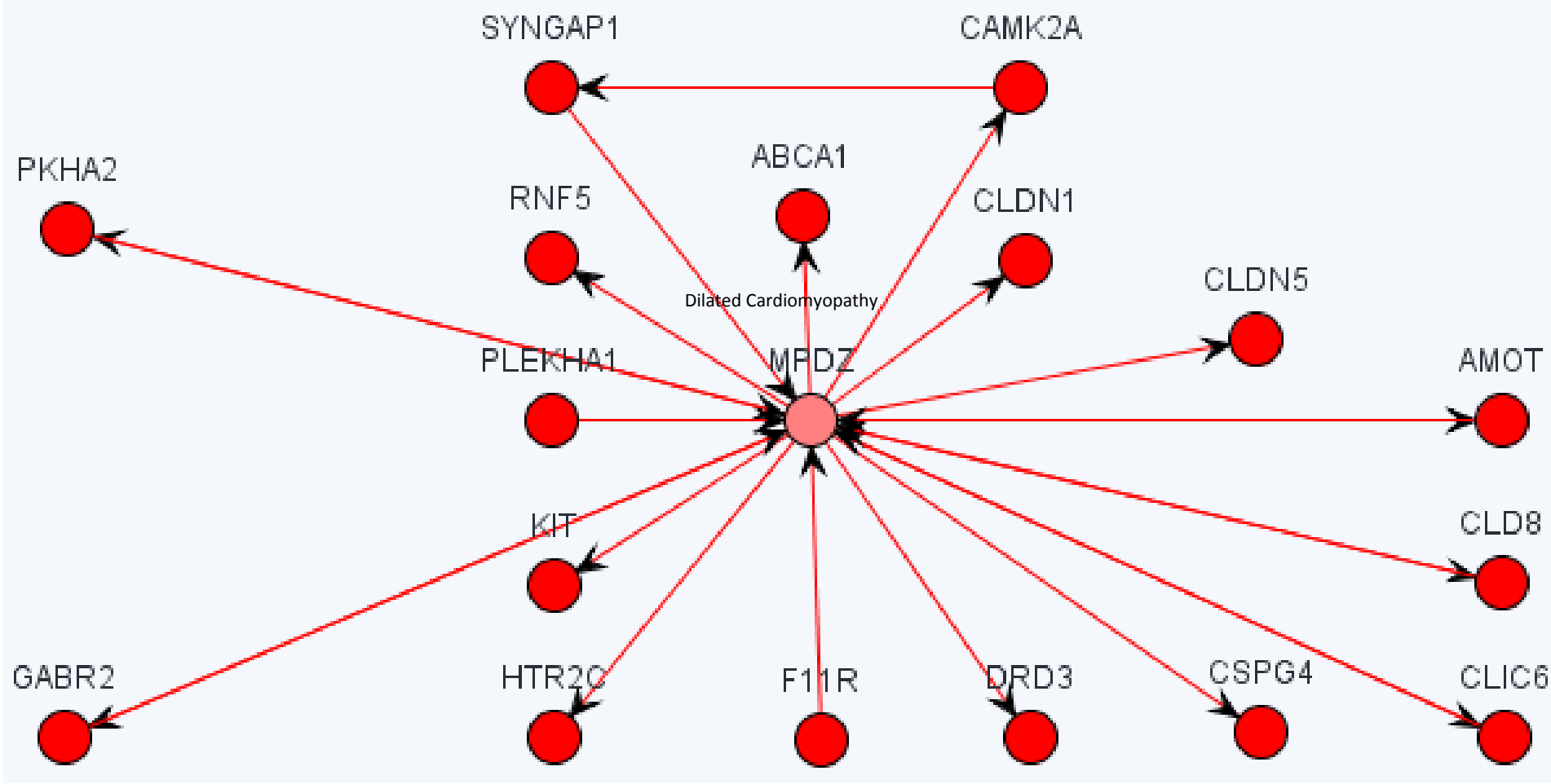
Modeling, analysis and simulation - Networks

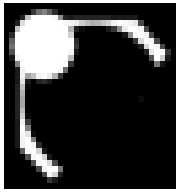
http://agbi.techfak.uni-bielefeld.de/vanesa/



Input: Protein-Interaction Network

Cardiovascular-related network





Cell microcosmos project – 3D virtual cell

Björn Sommer (University Konstanz)

cell microcosmos

BioDWH

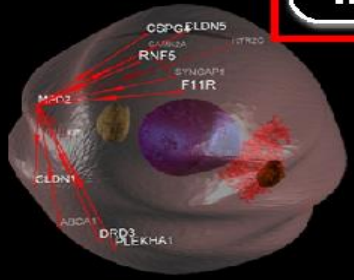
experimental data

datawarehouse data

ANDVisio

VANESA

network construction

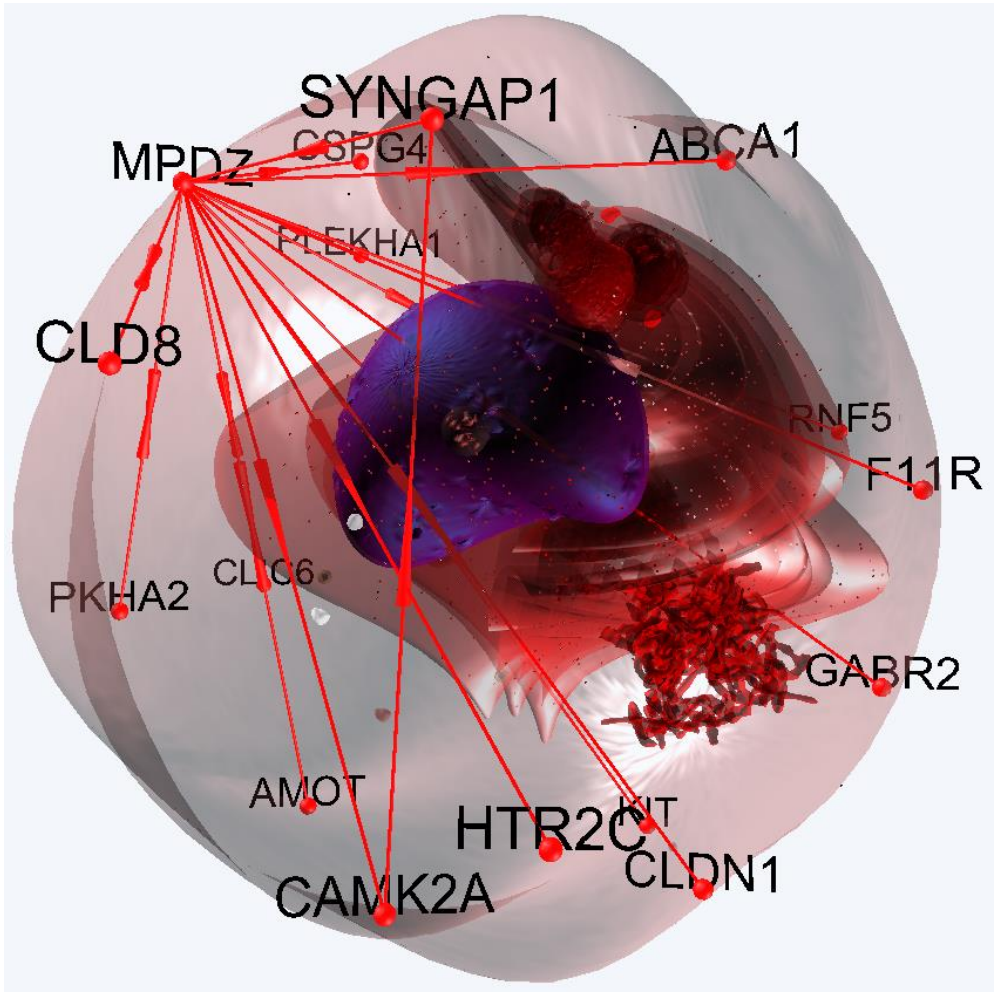


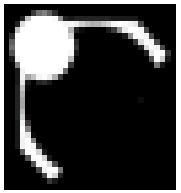
CmPI

enzyme localization

3D Visualization

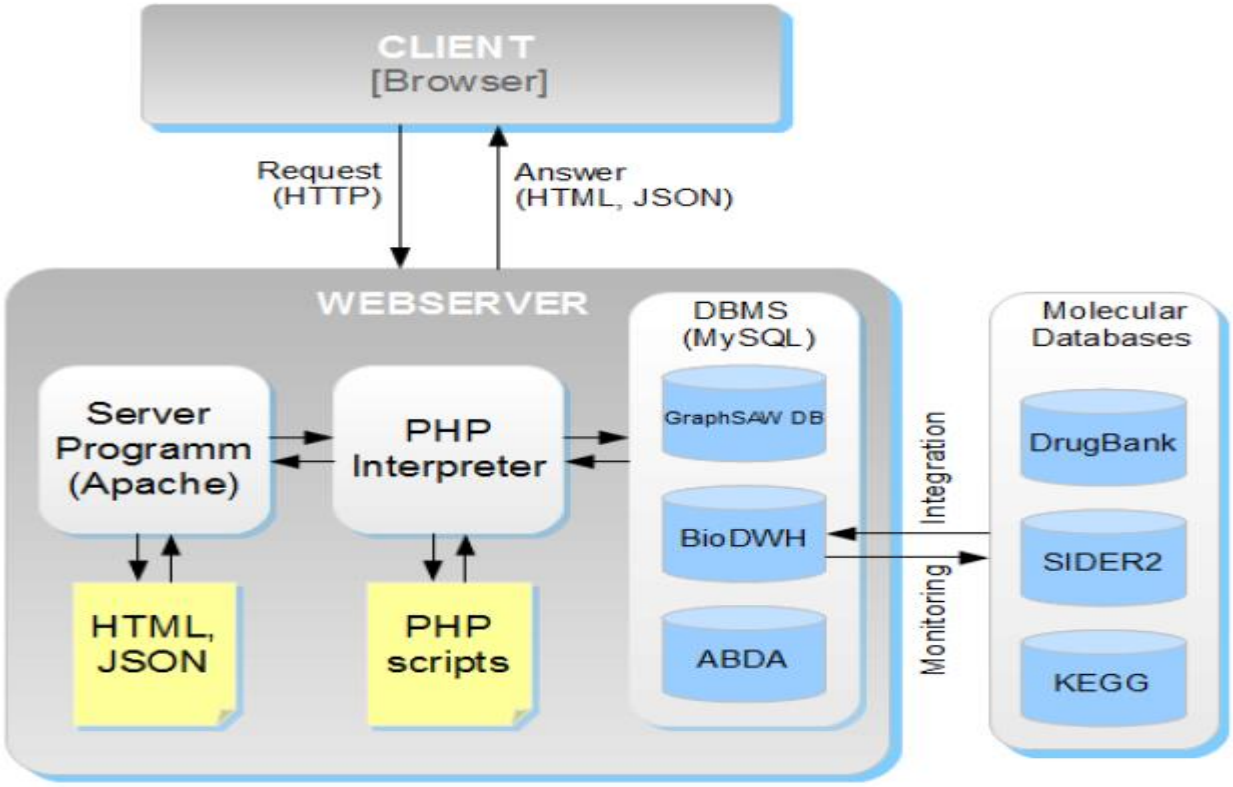
Cardiovasular-related network





GraphSAW

Graphical analysis of molecular and pharmaceutical side effects and interactions



A. Shoshi, T. Hoppe, B. Kormeier, V. Ogultarhan, R. Hofestädt
GraphSAW: A web-based system for graphical analysis of drug interactions and side effects using pharmaceutical and molecular data.
BMC Medical Informatics and Decision Making 15:15, 2015.

Data integration

BioDWH: Bielefelder Bio-Data Warehouse

DrugBank

DrugBank - XML representation: http://www.drugbank.ca/downloads
Drug data information system - Version DrugBank 3.0 represents:

6811 drugs

chemical, pharmaceutical and
biological information of each drug.

SIDER2

http://sideeffects.embl.de/download/

Drug information system represents:

996 drugs with
4192 different side effects.

ABDA (ABDATA Pharma Daten Service)

http://www.dimdi.de/static/de/amg/abda/

Drug information system represents:

Drug informtion (German market)
Information coming direct from the manufacturer.

Drugs
Information about agents and their chemical and physical features.

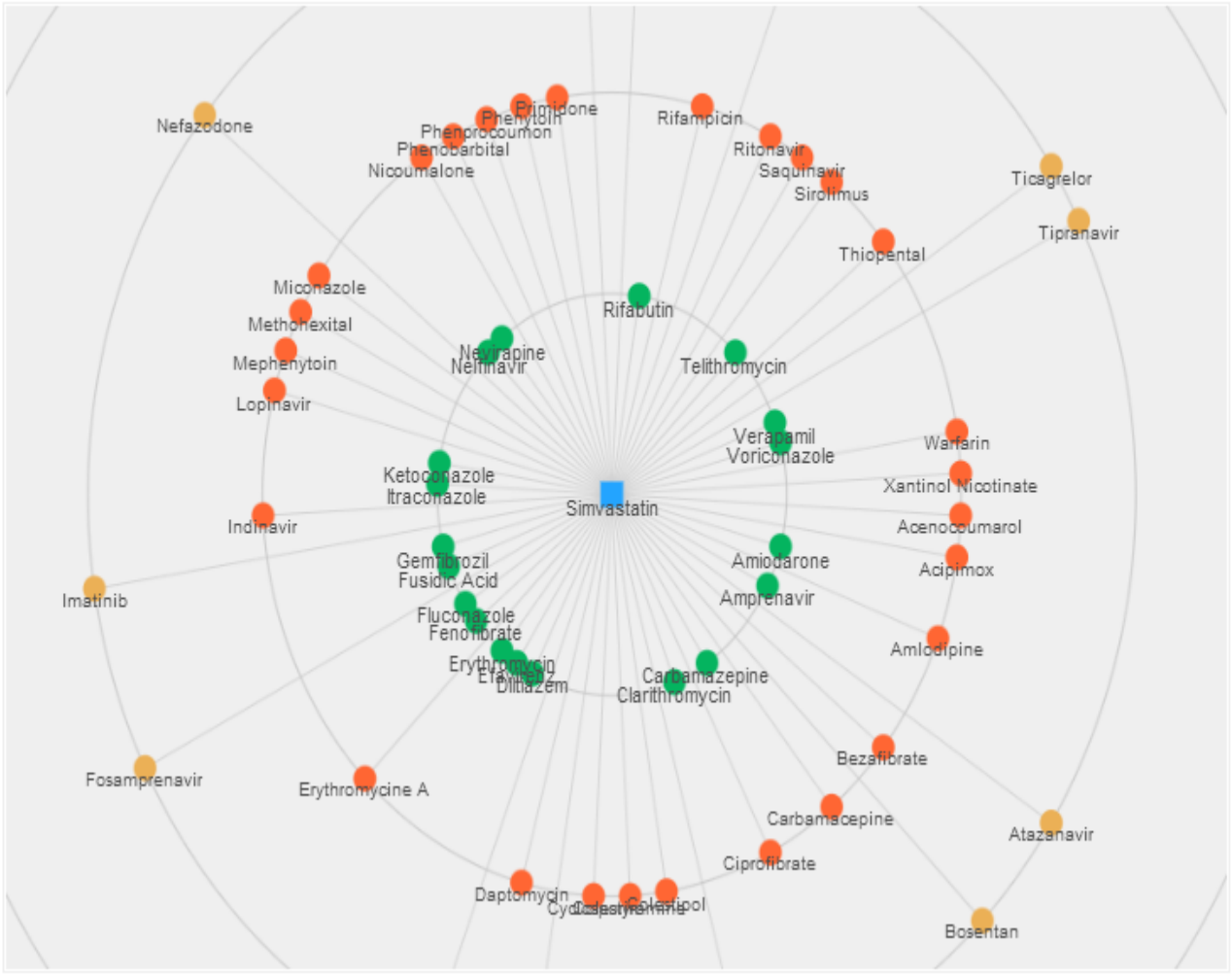
Interactions
Information about drug interactions (from the literature).

Producer/Marketing
Infos about the producer and the market.

Agent information
Producer independent informations about agents and drugs.

Daily news
Infos about new drugs, modifications and call backs.

Function: Single ***drug-interactions***



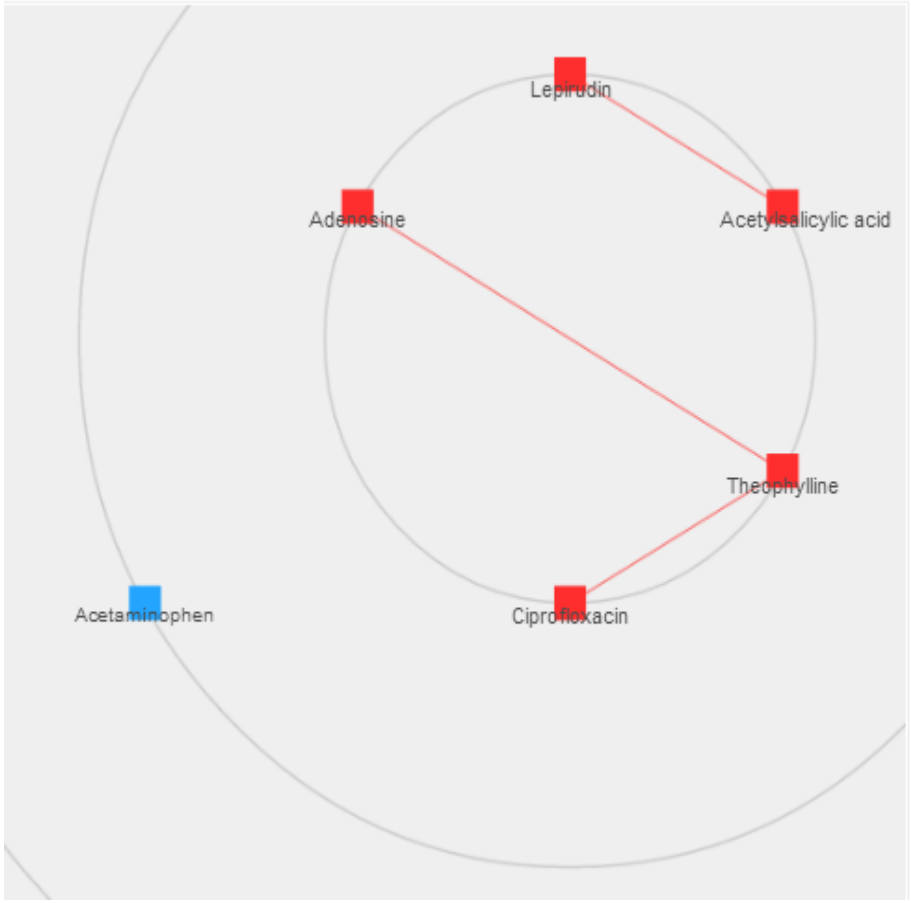
Interaktionspartner für Simvastatin [60]

Interaktionen aus ABDA: 47
Interaktionen aus Drugbank: 32
Korrelierende Interaktionen: 19

Interaktionspartner	Datenquelle
Acenocoumarol	ABDA
Acipimox	ABDA
Amiodarone	Korrelierend
Amlodipine	ABDA
Amprenavir	Korrelierend
Atazanavir	Drugbank
Bezafibrate	ABDA
Bosentan	Drugbank
Carbamacepine	ABDA
Carbamazepine	Korrelierend
Ciprofibrate	ABDA
Clarithromycin	Korrelierend
Colchicine	Drugbank
Colestipol	ABDA
Colestyramine	ABDA
Cyclosporin	ABDA
Cyclosporine	Drugbank
Daptomycin	ABDA
Delavirdine	Drugbank

Graph and table with interaction-partners of drug Simvastatin.

Function: Combined drug-interaction



Kombinierte Arzneimittel-Interaktionen

Interaktionspartner für Theophylline	Datenquelle
Adenosine	ABDA/Drugbank
Ciprofloxacin	ABDA/Drugbank

Interaktionspartner für Ciprofloxacin	Datenquelle
Theophylline	ABDA/Drugbank

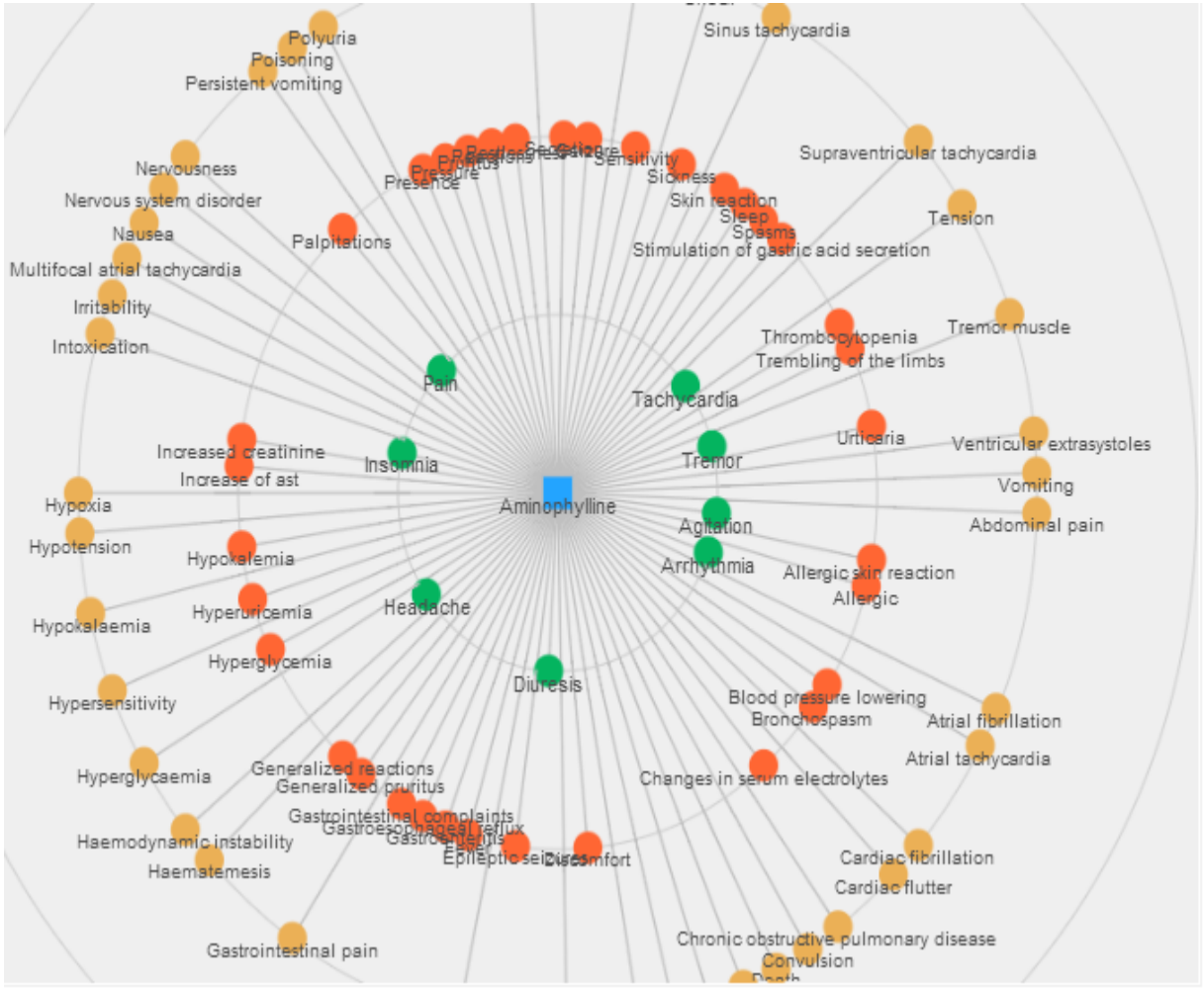
Interaktionspartner für Adenosine	Datenquelle
Theophylline	ABDA/Drugbank

Interaktionspartner für Lepirudin	Datenquelle
Acetylsalicylic acid	ABDA

Interaktionspartner für Acetylsalicylic acid	Datenquelle
Lepirudin	ABDA

Combined drug-interaction for the agents *Theophyllin, Adenosin, Ciprooxacin, Acetaminophen, Lepirudin und Acetylsalicylic acid.*
Left: Visualization of the result.
Right: textual representation of the result.

Function: Single drug-side effects



Nebenwikrungen für Aminophylline [83]

Nebenwirkungen aus ABDA: 43
Nebenwirkungen aus SIDER: 48
Korrelierende Nebenwirkungen : 8

Nebenwirkung	Datenquelle
Abdominal pain	SIDER
Agitation	Korrelierend
Allergic	ADBA
Allergic skin reaction	ADBA
Arrhythmia	Korrelierend
Atrial fibrillation	SIDER
Atrial tachycardia	SIDER
Blood pressure lowering	ADBA
Bronchospasm	ADBA
Cardiac fibrillation	SIDER
Cardiac flutter	SIDER
Changes in serum electrolytes	ADBA
Chronic obstructive pulmonary disease	SIDER
Convulsion	SIDER
Death	SIDER
Dermatitis	SIDER
Dermatitis contact	SIDER
Dermatitis exfoliative	SIDER
Diarrhoea	SIDER

Result for the single drug-side search of Aminophyllin.

Application case*

Drugs

Erythromycin-ratiopharm 500 DB/-1000 DB; -500

-> ***Erythromycin***

Simvastatin - 1 A Pharma 5 mg/-10 mg/-20 mg/-30 mg/-40 mg/-80 mg Filmtabletten

-> ***Simvastatin***

Paracetamol-ratiopharm 500 mg Brausetabletten; -500 Tabletten; -Lösung; -125/-250/-500/-1000 Zäpfchen

-> ***Acetaminophen***

Ibuprofen Sandoz 400 mg/-600 mg Filmtabletten; -800 mg Retardtabletten

- **> Ibuprofen**

Side effects

Vomiting
Body temperature decreased
Dizziness

*based on medical guidlines (09/2013)

Application - result*

Molekulare Medikationsanalyse

Die Tabellen zeigen die Arzneimittel-Interaktionen, Enzym-Interaktionen basierend auf genetischen Enzymdefekten sowie Nebenwirkungen für jedes Arzneimittel. Durch Auswahl eines Eintrages in den Tabellen können weitere Informationen abgerufen werden.

Acetaminophen
Vomiting

Erythromycin
Simvastatin
Vomiting
Dizziness

Ibuprofen
Vomiting
Dizziness

Simvastatin
Erythromycin
Vomiting
Body temperature increased
Dizziness

Für die folgenden Arzneimittel wurden Interaktionen gefunden:
[»] = Alternative Medikation

ARZNEIMITTELNAME	ARZNEIMITTELNAME
Ciprofloxacin STADA®... [»]	Theophyllin STADA® 2... [»]

Die eingegebenen Nebenwirkungen können durch folgende Arzneimittel verursacht werden:
[»] = Alternative Medikation

ARZNEIMITTELNAME	NEBENWIRKUNGEN
Theophyllin STADA® 2... [»]	Kopfschmerzen
Ciprofloxacin STADA®... [»]	Schwindel

Folgende Arzneimittel der gleichen Wirkstoffgruppe stehen Ihnen zur Auswahl anstatt:
Ciprofloxacin STADA® 100 mg/-250 mg/-500 mg/-750 mg Filmtabletten

ARZNEIMITTELNAME
Tarivid® 200/-400 Filmtabletten; -i.v. 200/-400
Uro-Tarivid® 100
Ofloxacin AL 100/-200/-400
Oflox-Sandoz 100 mg/-200 mg Filmtabletten

Discussion

Are we able to understand gene regulation ?

We build up the electronical infrastructure (first steps).

But my feeling – NO

Why ?

1. Always new molecular mechanisms – RNAs

2. In theory we do not understand machines, which represent all the detected features at once:
parallel, probabilistic, dynamic, modular, dataflow …