NTNU – Trondheim
Norwegian University of
Science and Technology

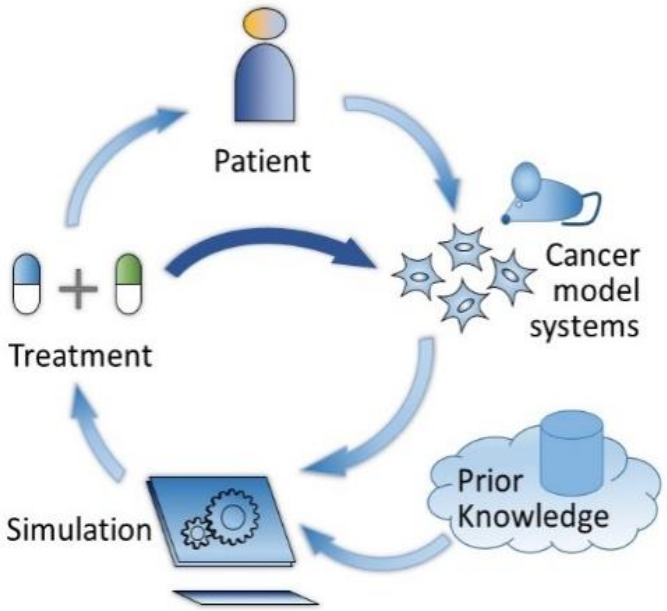
Multidisciplinary Qualities of Systems Medicine

- the pleasure - *as molecular biologist* - of working with bioinformaticians, medical doctors and philosophers

Astrid Lægreid
Department of Clinical and Molecular Medicine
Faculty of Medicine and Health Sciences
NTNU

NTNU – Trondheim
Norwegian University of
Science and Technology

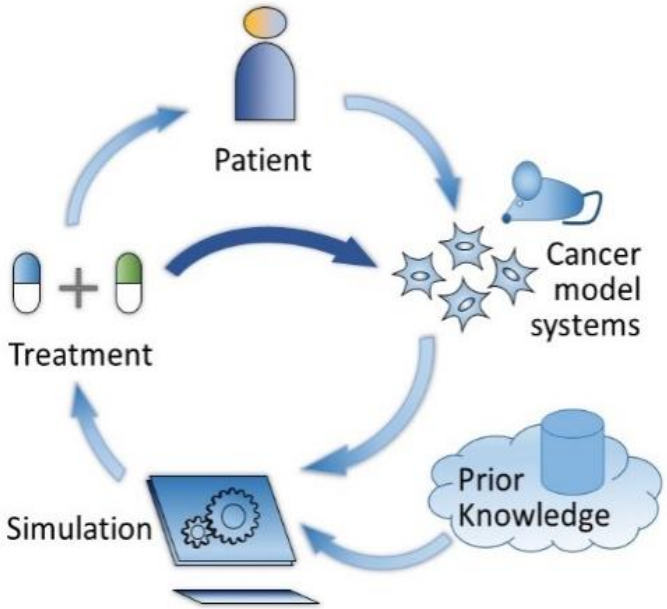
DrugLogics Initiative at NTNU



- drug combinations
- knowledge-based modeling
- experimental testing
- responsible research and innovation, RRI

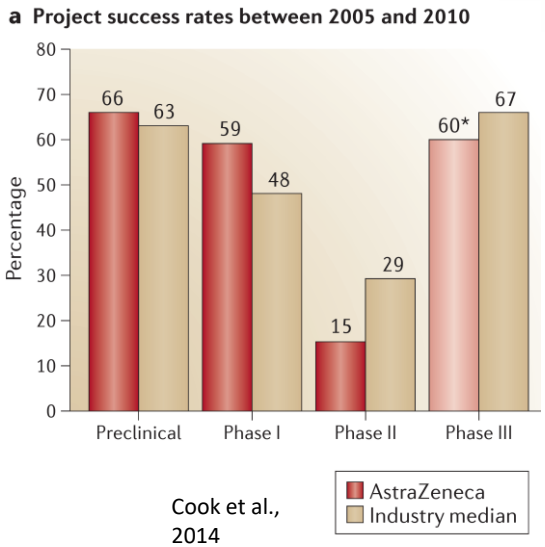
outline

- precision oncology
- predictive modelling
- knowledge management
- extending knowledge commons

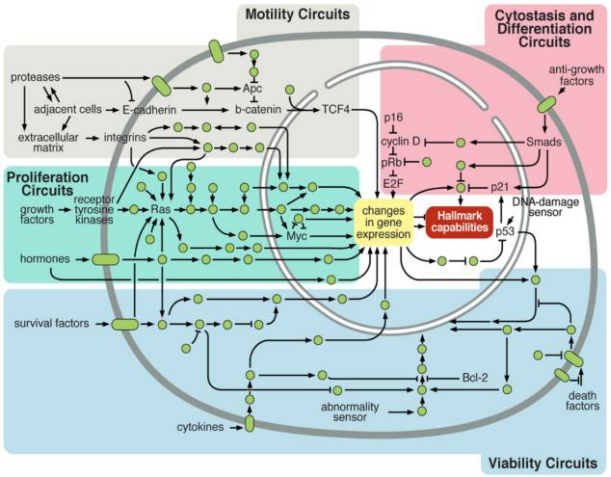


challenges in cancer treatment

- multiple robustness features of cancer
- targeted therapies – relapse
- large number of failed phase II trails



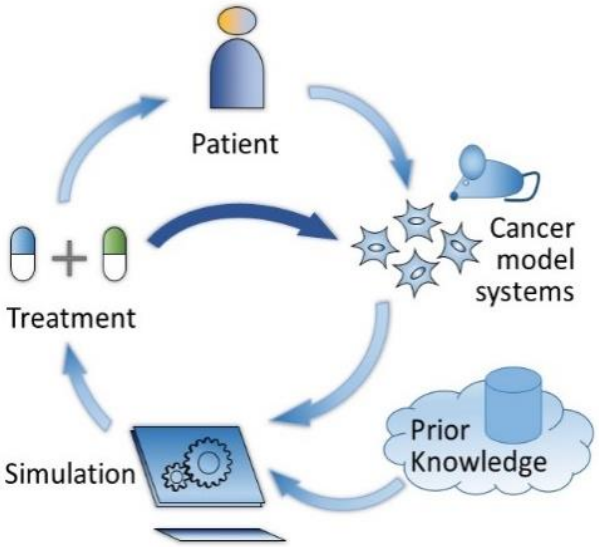
Cook et al., 2014

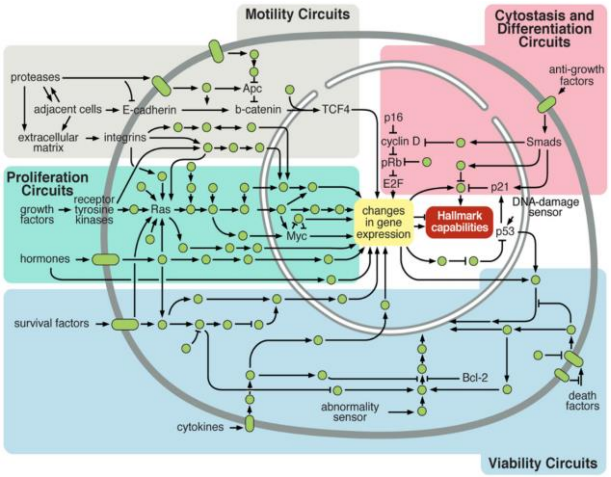


Hanahan & Weinberg, 2011

challenges in cancer treatment

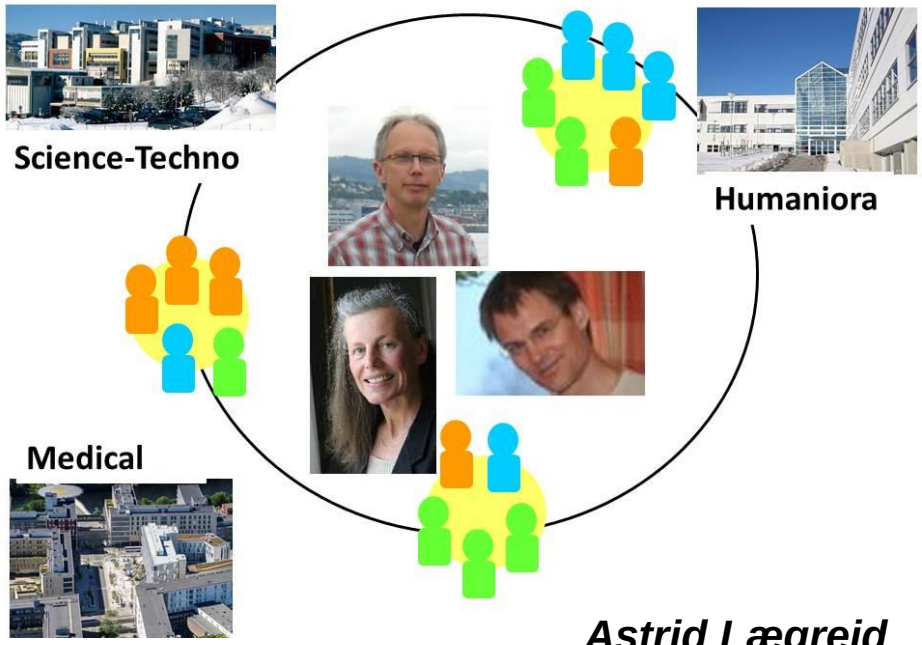
- multiple robustness features of cancer
- targeted therapies – relapse
- large number of failed phase II trails





Hanahan & Weinberg, 2011

multidisciplinary

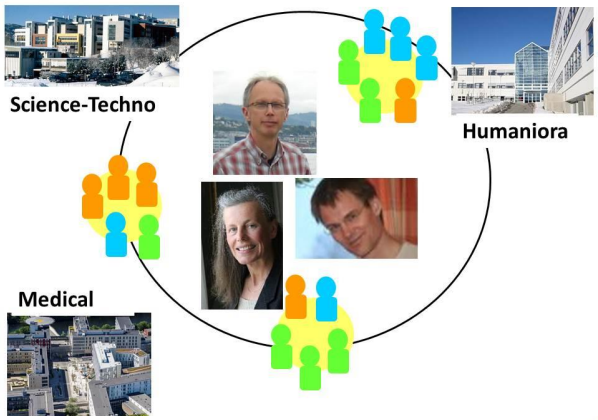


Astrid Lægreid

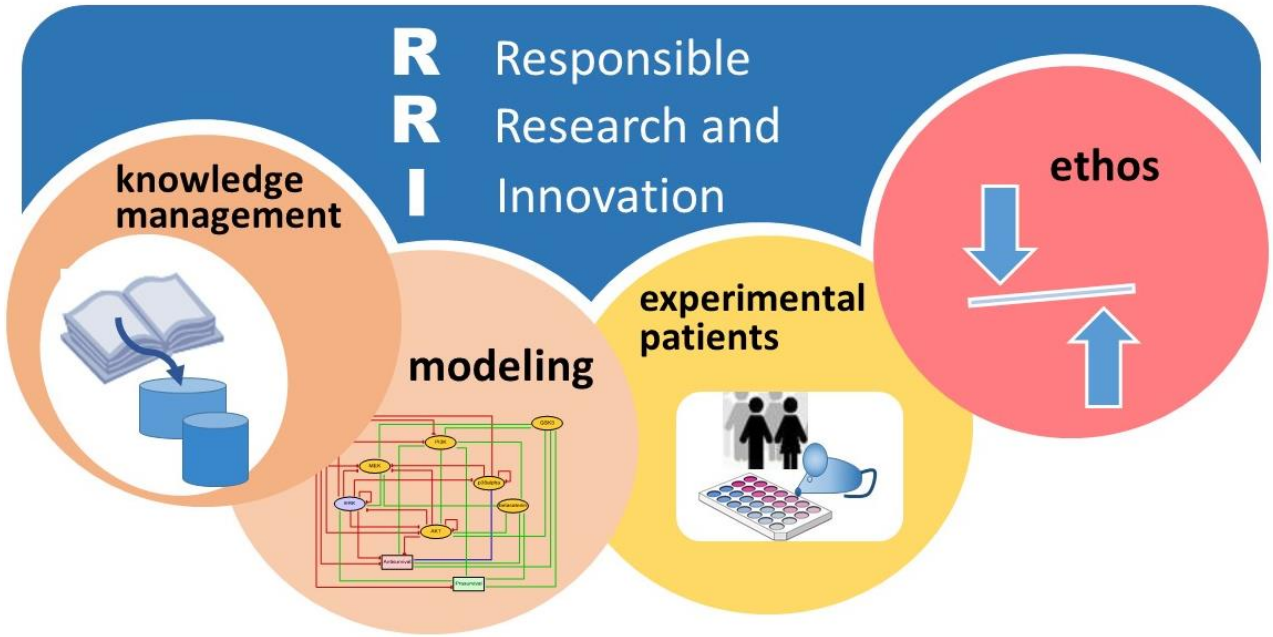
Martin Kuiper

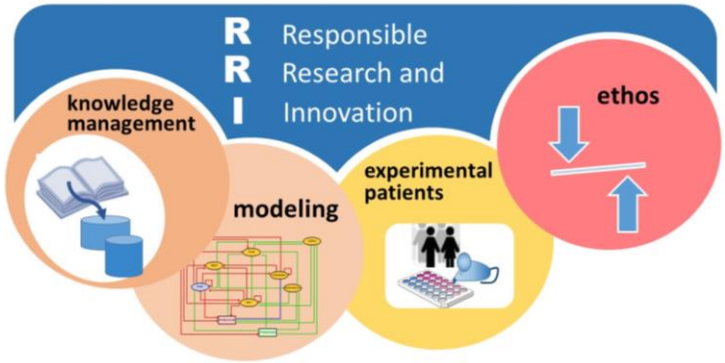
Rune Nydal

organization, funding



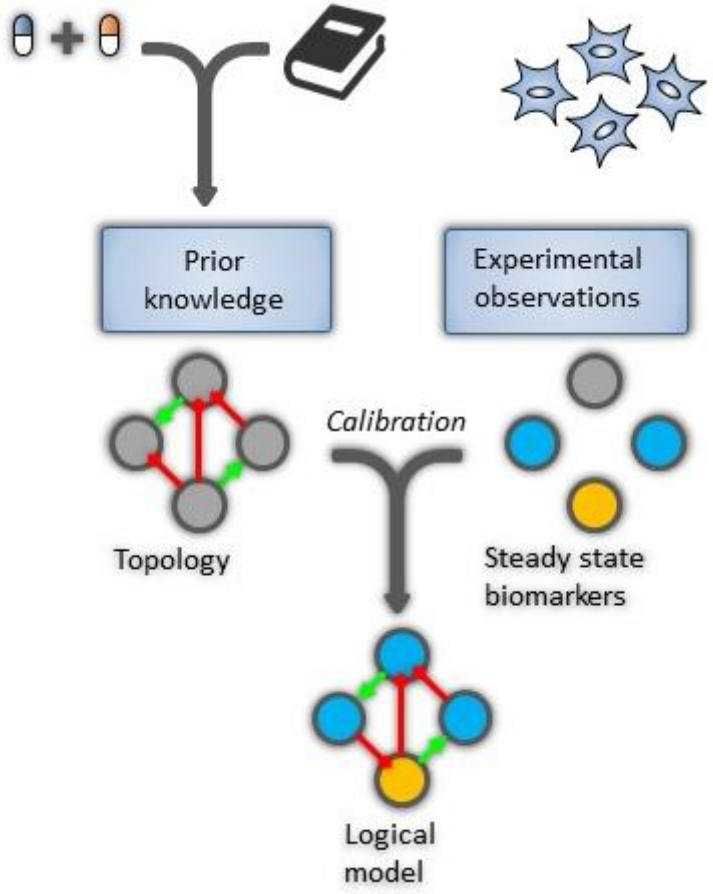
multidisciplinary

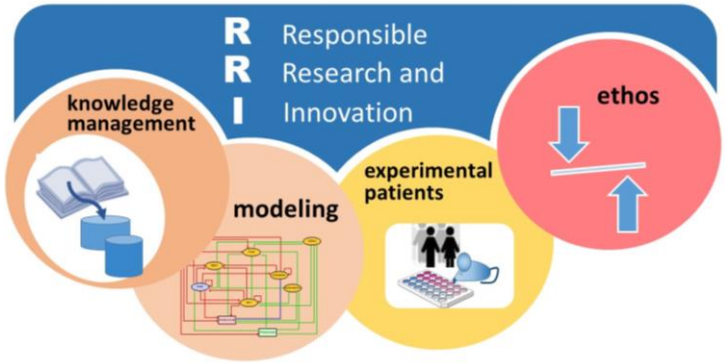




modelling

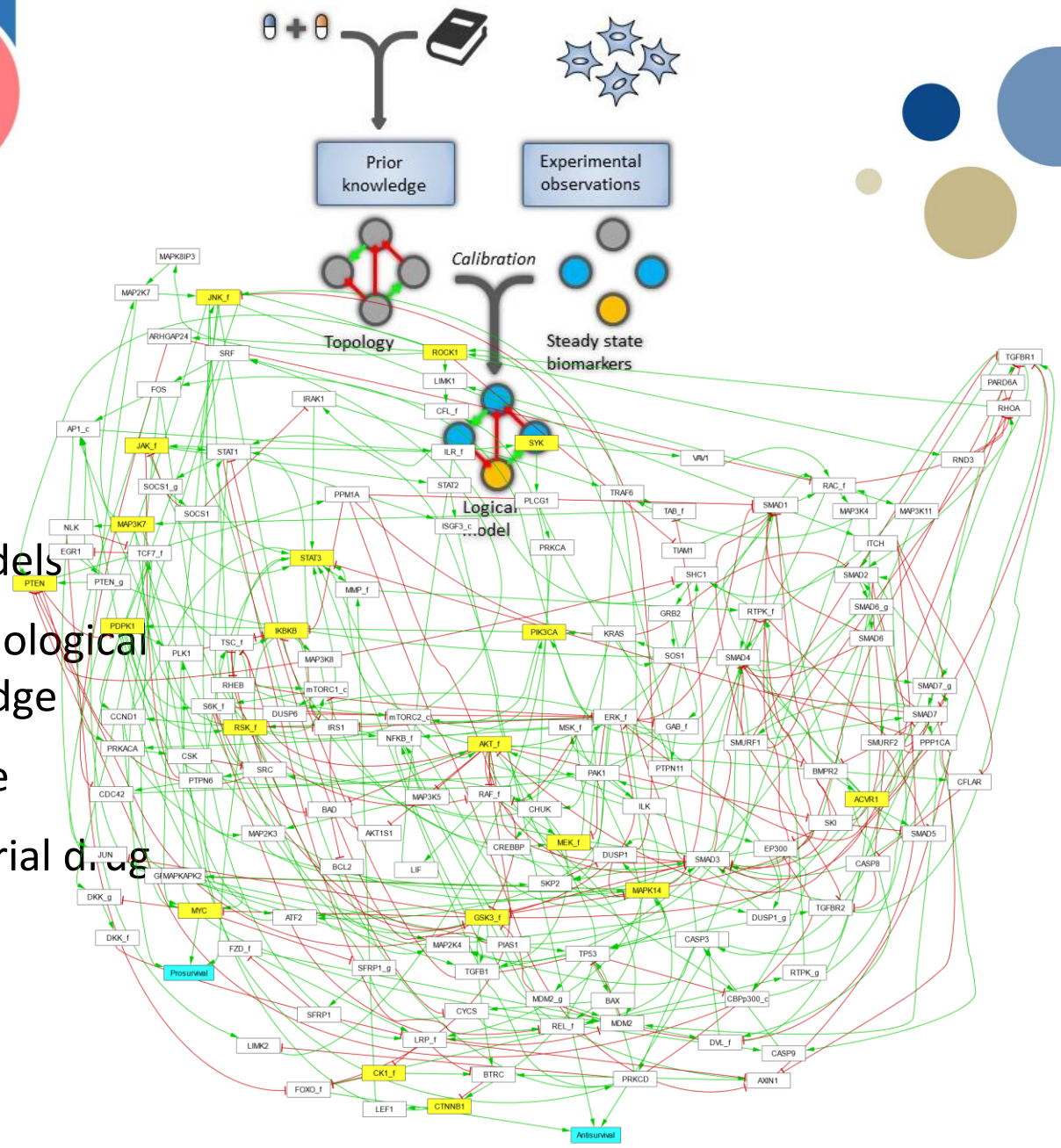
- logical, boolean models
- mechanistic, from biological background knowledge
- calibrated to cell line
- simulate combinatorial drug responses

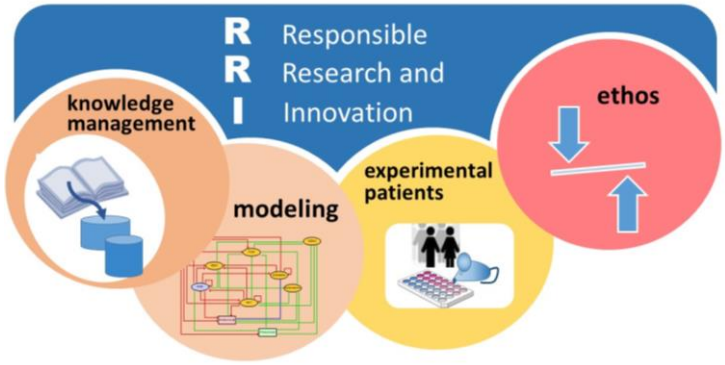




modelling

- logical, boolean models
- mechanistic, from biological background knowledge
- calibrated to cell line
- simulate combinatorial drug responses

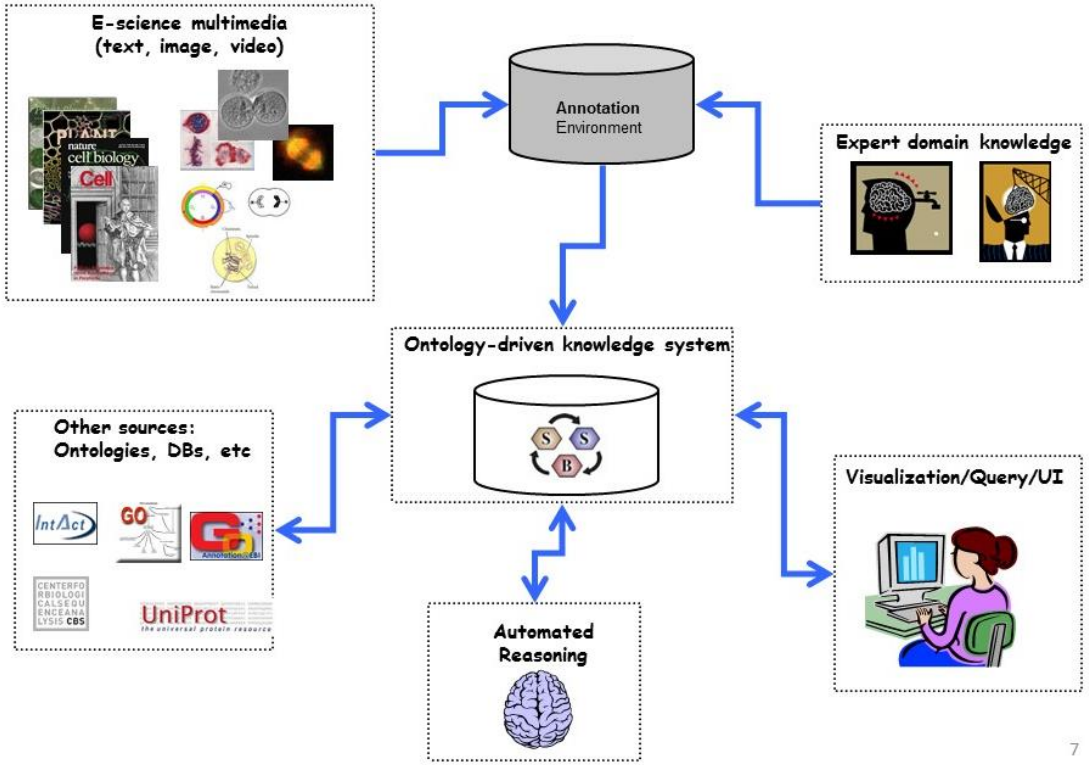


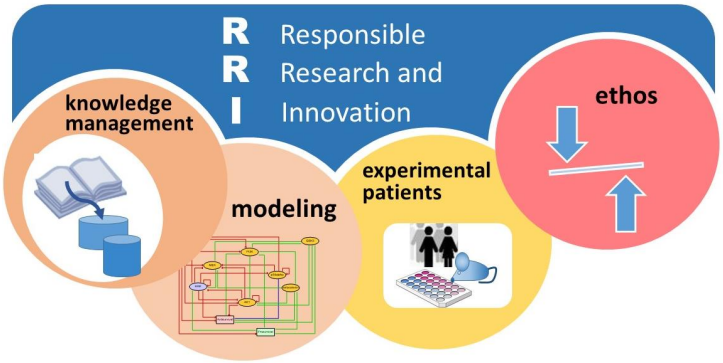


knowledge management

for systems medicine including model building

- data
- knowledge
- causal statements
- standards, interoperability

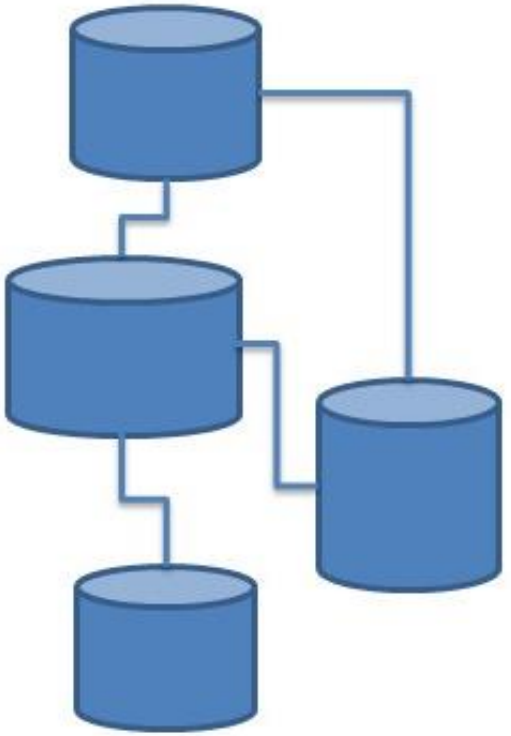


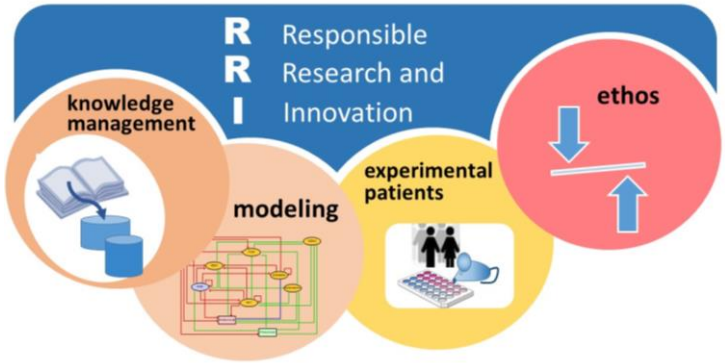


knowledge commons

ecosystem knowledge- and data bases

- structure
- function
- processes
- causality

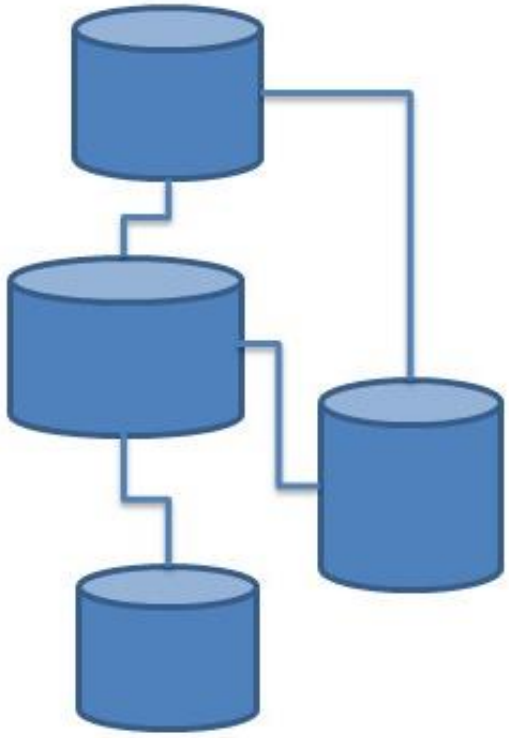




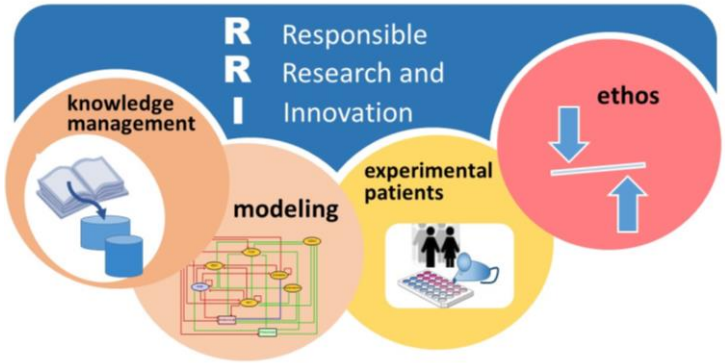
ethos

knowledge commons

- novel *e-infrastructures* enabling precision medicine
- digital knowledge representation
- stakeholders concerns
- trust, confidence



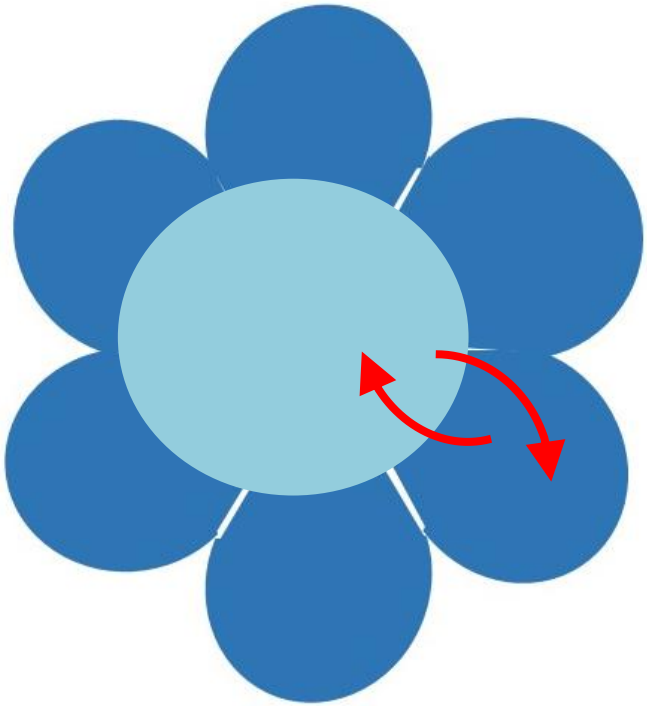
ethos – moral character, guiding beliefs and ideals



multidisciplinary

how to best work together?

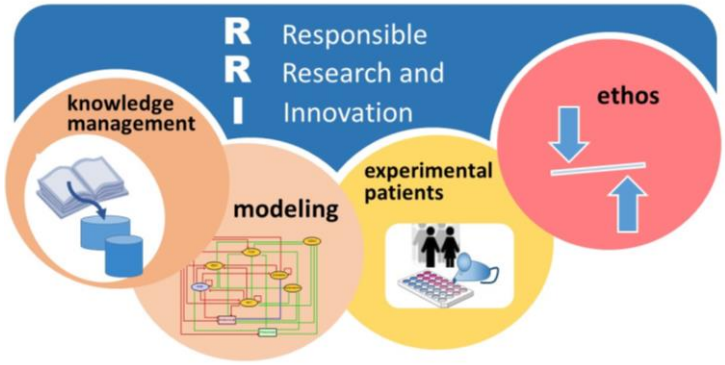
integrating disciplines



- pursue own research while achieving common objective
- mobilise different disciplines for mutual contributions to common objective

integrating disciplines

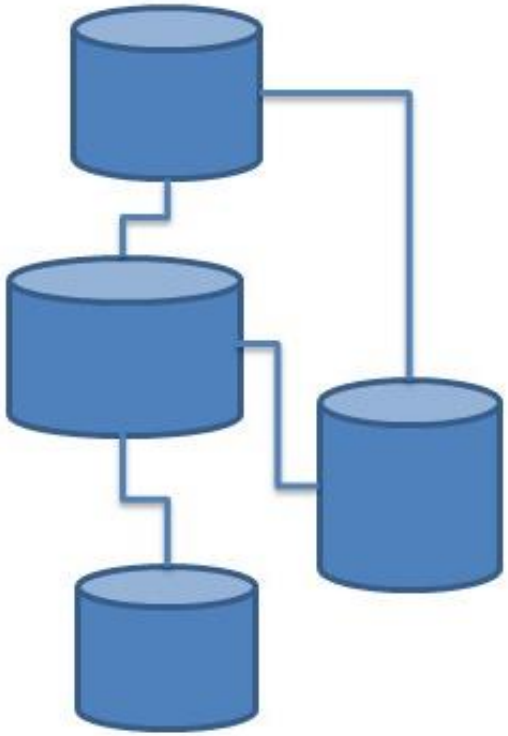
- pursue own research while achieving common objective
- mobilise different disciplines for mutual contributions to common objective
  - **major joint focus:**
    Knowledge Commons for the Life Sciences as foundation for precision oncology research and clinical decision support

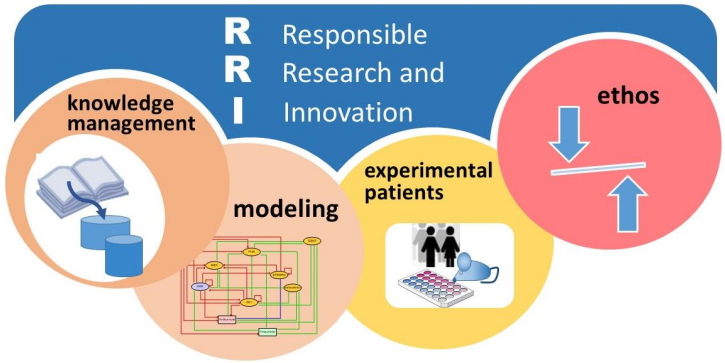


ethos

knowledge commons

- novel *e-infrastructures* enabling precision medicine
- digital knowledge representation
- stakeholders concerns
- trust, confidence

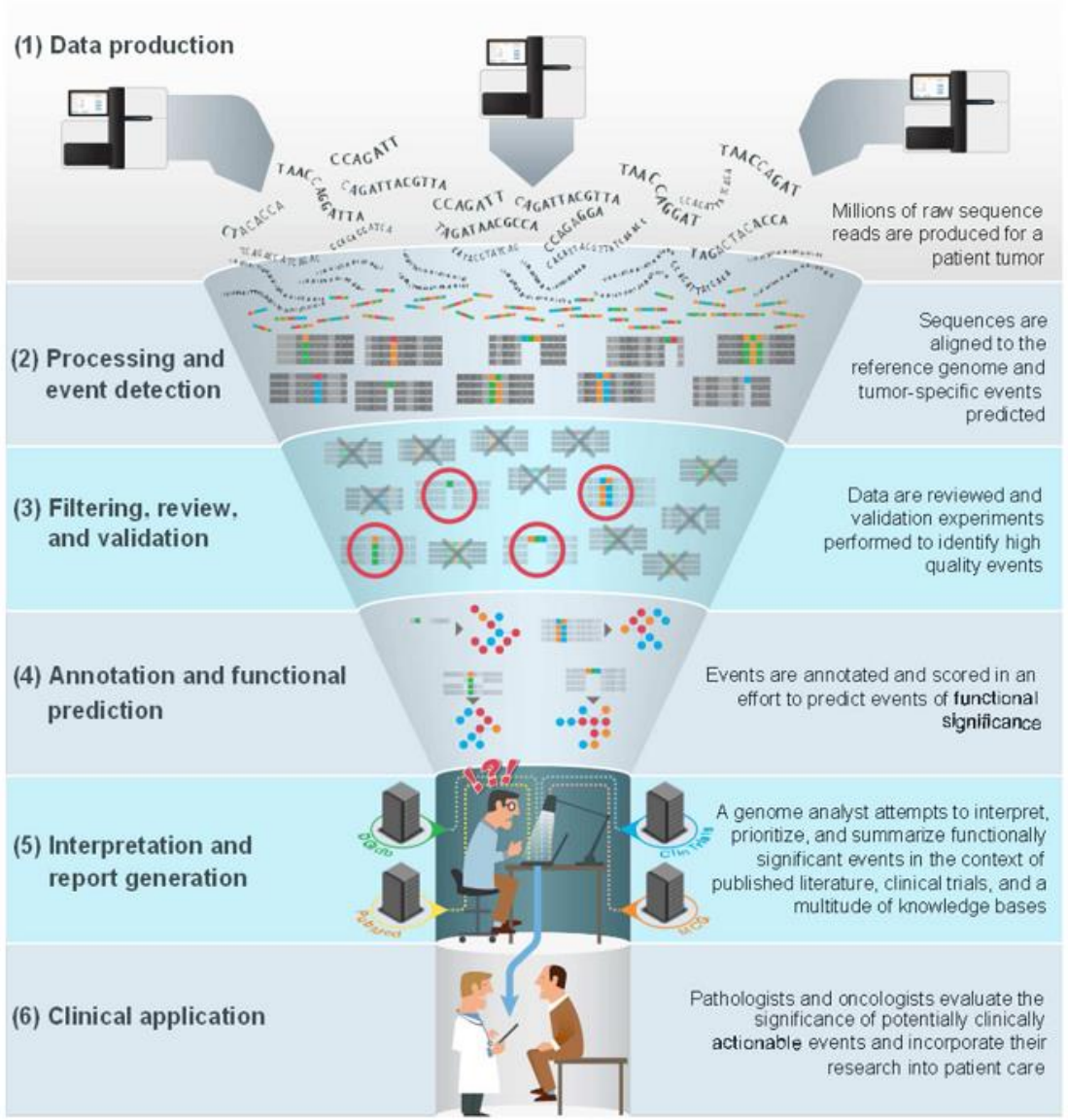




ethos

knowledge commons

- novel *e-infrastructures* enabling precision medicine
- digital knowledge representation
- stakeholders concerns
- trust, confidence



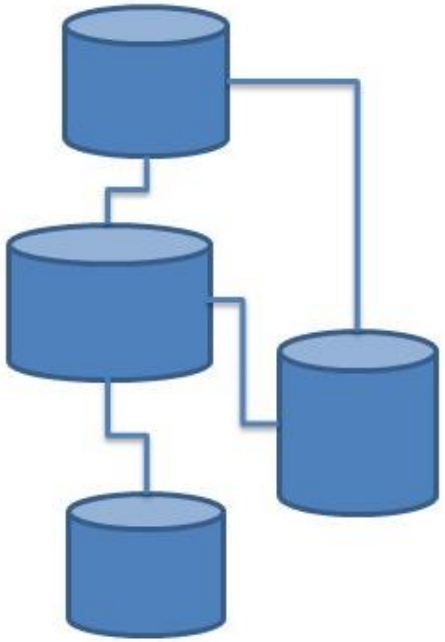
Good BM, et al; *Genome Biol.* 2014 Aug 27;15(8):438

life science - computing

- life science increasingly becoming dependent on computer technology
- these technologies need to be appropriated in life science daily practice.
- is there a scientific obligation to engage in the work of appropriation?
- is trust required?
- is confidence sufficient?

life science - knowledge commons

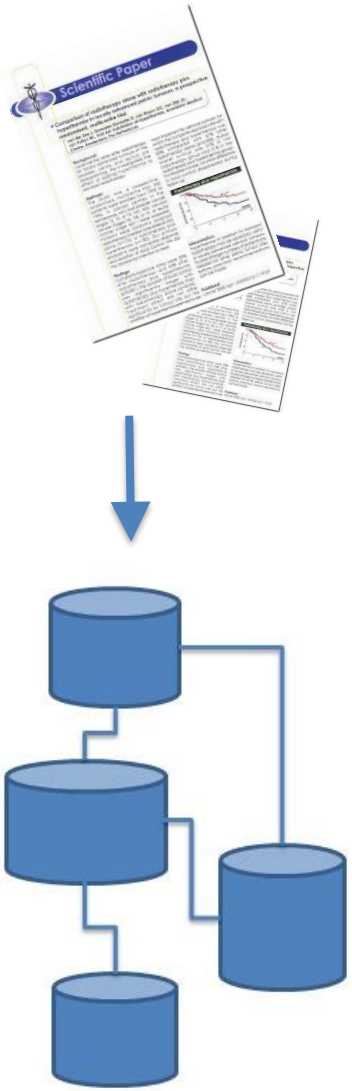
- life science increasingly becoming dependent on the Knowledge Commons
- the Knowledge Commons needs to be appropriated in life science daily practice.
- is there a scientific obligation to engage in the work of appropriation?
- is trust required?
- is confidence sufficient?



knowledge commons

-> destabilising

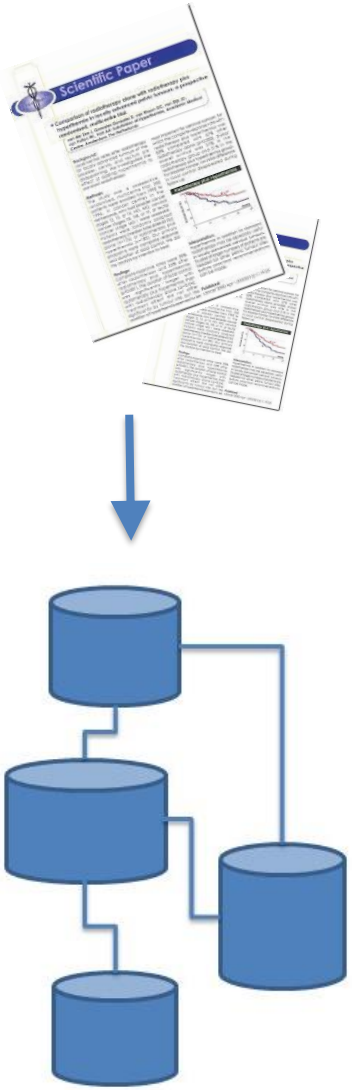
- destabilising effect of accomodating the Knowledge Commons
- changes relationships to past achievements
- biologists and medical doctors used to rely on personal accumulated knowledge and capacity to find, read and evaluate scientific journals
- now also need to rely on institutions, procedures and expertise involved in building and maintaining the Knowledge Commons offered to her



knowledge commons

-> shift in epistemic practices

- the work of building and appropriating the Knowledge Commons is part of a significant shift of epistemic practices
- life scientists' relationships to 'prior knowledge' changes; including the relationships between the scientists and her community

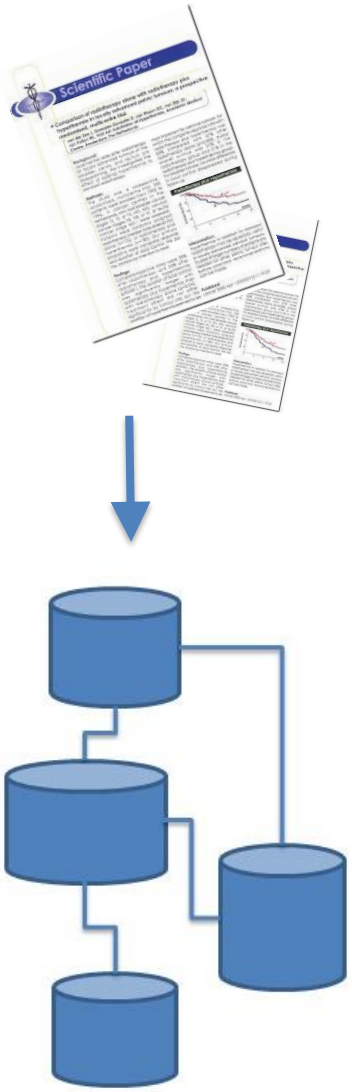


knowledge commons

-> what is at stake?

-> knowledge infrastructures restructured

- knowledge infrastructures: “robust internetworks of people, artefacts, and institutions which generate, share, and maintain specific knowledge about the human and natural worlds”
- past collective achievements are managed and mobilised through these infrastructures

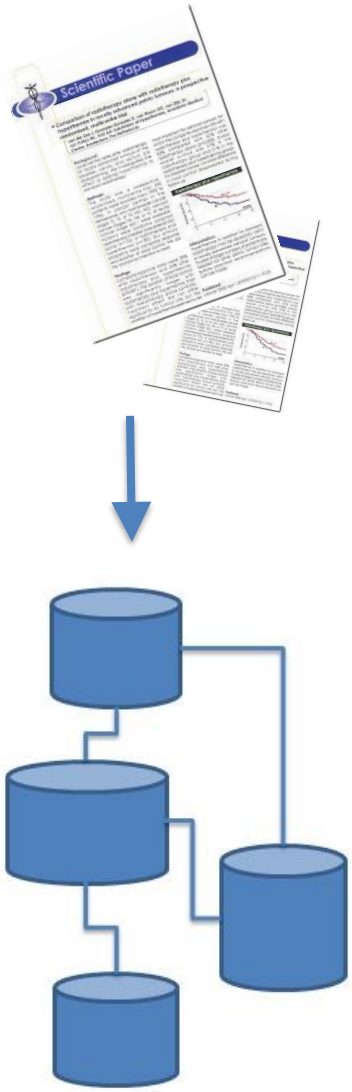


knowledge commons

-> re-engineering knowledge infrastructures

complex new infrastructures of people artefacts and institutions.

- text-centric infrastructure
  - publications, journals, peer review, vancouver declaration, paper genre, style of writing, impact-factor, libraries, network of libraries, editors, publishing houses, librarians, Google Scholar, Research Gate, etc ….
- infrastructure reengineered to include digital objects
  - digital objects, ontologies, controlled vocabularies, interoperability, knowledge bases, curators, stewards, provenance, nano-publications, Gene ontology consortium, EBI, FAIR, data-information-knowledge, etc..

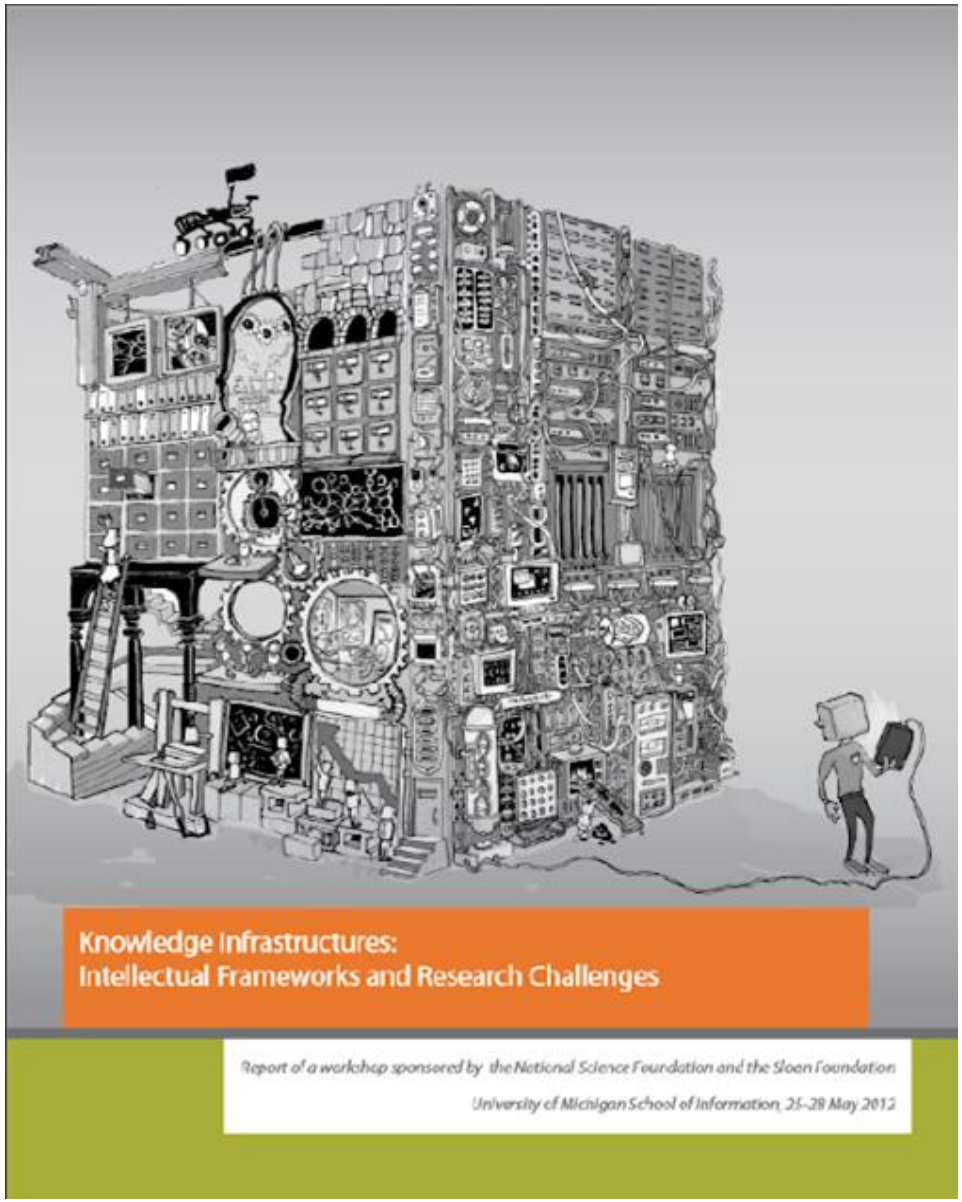


knowledge commons

-> re-engineering knowledge infrastructures

complex new infrastructures of people artefacts and institutions.

- text-centric infrastructure
  - publications, journals, peer review, vancouver declaration, paper genre, style of writing, impact-factor, libraries, network of libraries, editors, publishing houses, librarians, Google Scholar, Research Gate, etc ….
- infrastructure reengineered to include digital objects
  - digital objects, ontologies, controlled vocabularies, interoperability, knowledge bases, curators, stewards, provenance, nano-publications, Gene ontology consortium, EBI, FAIR, data-information-knowledge, etc..



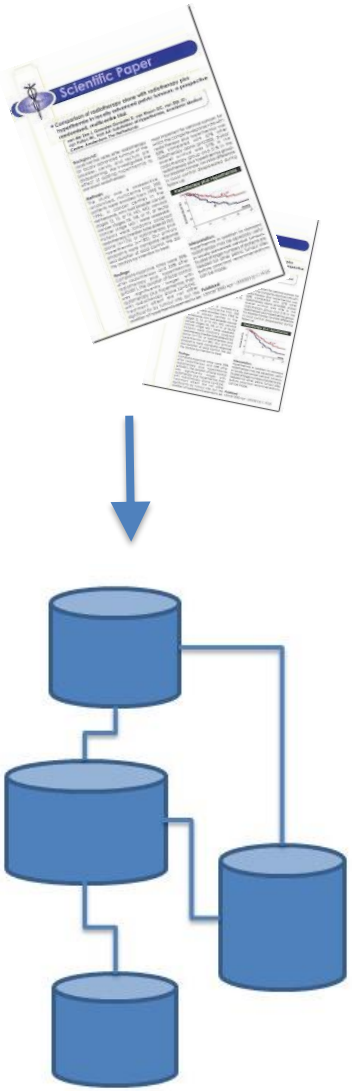
Edwards et al, 2012

knowledge commons

-> restabilising, re-engineering

-> trust-building activities

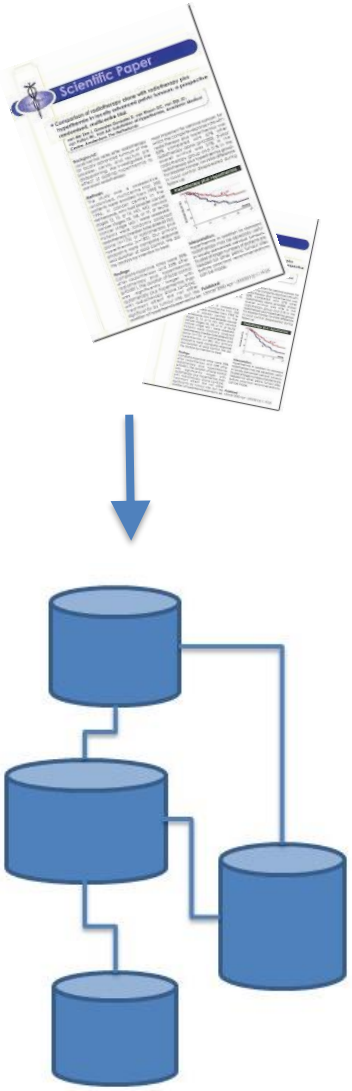
- quality issues translated into issues of trust
- translating reliability-metrics from established (text-centric) to new digital knowledge infrastructure



knowledge commons

-> restabilising, re-engineering

- ***relevance:*** knowledge base adequate for the subfield of the scientist? Can it be integrated in her work flow?
- ***metadata:*** biological context of digital objects provided by adequatemetadata, is it rich or thick enough?
- ***provenance:*** sources and procedures available? – like the journal sources, or curation guidelines?
- ***evidence:*** evidence codes satisfactory for evaluations of the reliability of the claims?
- ***confidence:*** how to translate confidence into a number?
- ***interoperability:*** how well do digital objects to move and connect to other digital objects in other databases?
- ***coverage:*** how well is the knowledge of the sub-field covered?
- ***maintenance:*** reflects the state of the art: updated on empirical findings, adequately structured by new theoretica principles?

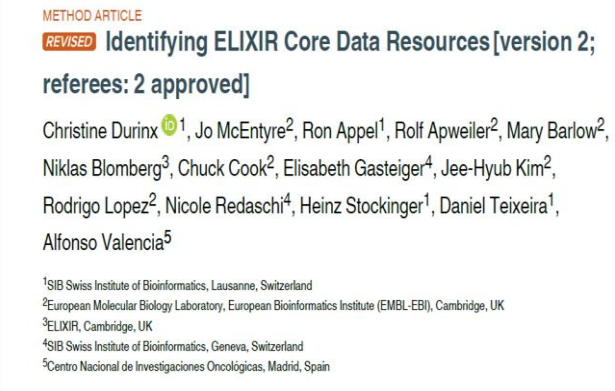


knowledge commons

-> restabilising, re-engineering

-> ELIXIR identifies Core Data Resources

- governmental support: to ensure long time preservation, maintenance, commitment and relevance
- resources should:
  - be known to key stakeholders, journals & funders
  - be well known by scientific community
  - have authority in the field
  - cover the sub-domain
  - have well-understood dependencies to other databases
  - be able to co-exist with databases having other motivations
- governance:
  - short period of immature construction phase (2y)
  - give notice in good time before withdrawal (1y)



METHOD ARTICLE

REVISED **Identifying ELIXIR Core Data Resources [version 2; referees: 2 approved]**

Christine Durinx[1], Jo McEntyre[2], Ron Appel[1], Rolf Apweiler[2], Mary Barlow[2], Niklas Blomberg[3], Chuck Cook[2], Elisabeth Gasteiger[4], Jee-Hyub Kim[2], Rodrigo Lopez[2], Nicole Redaschi[4], Heinz Stockinger[1], Daniel Teixeira[1], Alfonso Valencia[5]

[1]SIB Swiss Institute of Bioinformatics, Lausanne, Switzerland
[2]European Molecular Biology Laboratory, European Bioinformatics Institute (EMBL-EBI), Cambridge, UK
[3]ELIXIR, Cambridge, UK
[4]SIB Swiss Institute of Bioinformatics, Geneva, Switzerland
[5]Centro Nacional de Investigaciones Oncológicas, Madrid, Spain

knowledge commons

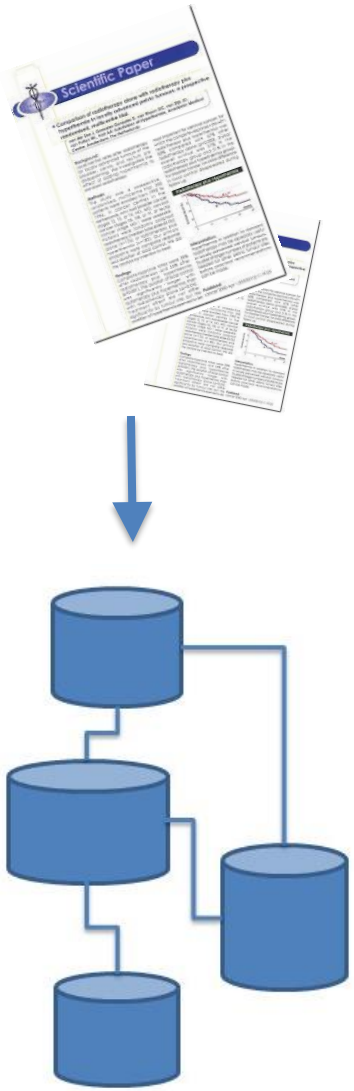
-> restabilising

-> requires trust or confidence?

Luhman's distinction between confidence & trust

- **Trust** is required if a bad outcome would make you regret – in a situation where the possible damage may be grater than the advantage you seek.
- **Confidence** is the normal case – you are confident your expectations will not be disappointed.

- the epistemic shift cast as issue of trust or confidence?



Luhman: "Familiarity, confidence and trust: Problems and alternatives" (2000)

knowledge commons

-> restabilisation

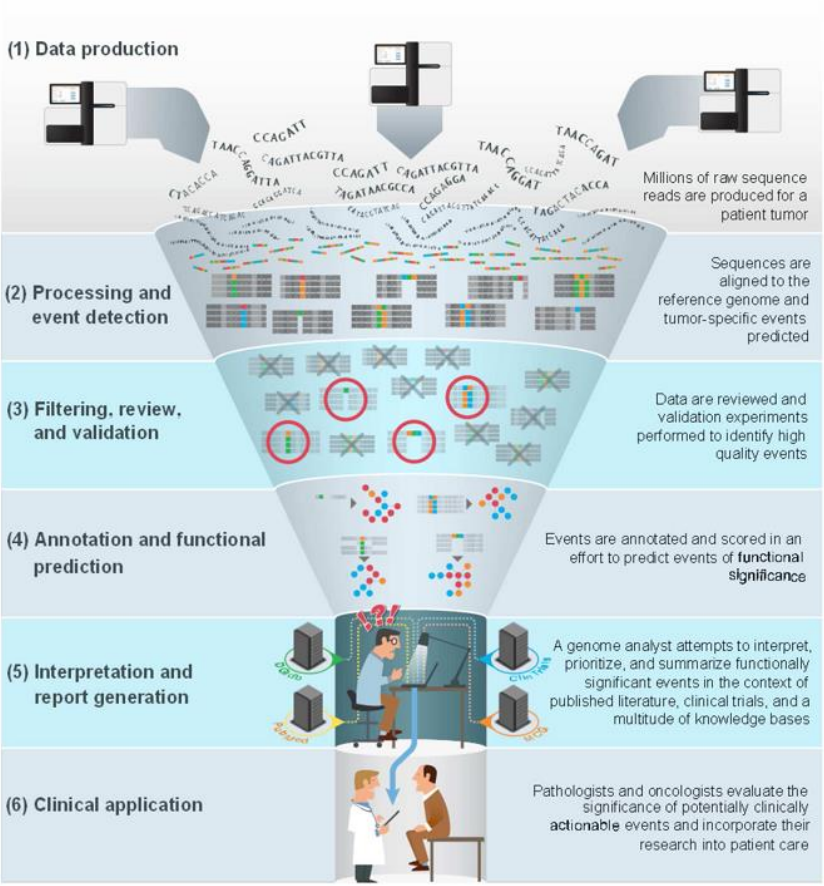
-> confidence vs trust: different approaches?

Confidence

- do not consider alternatives (no choice made).
- make incremental improvements in the digitalized research practices of scientists
- convince the sceptics, recruit and mobilise relevant stakeholders

Trust

- requires engagement – pathos
- emphasis on choices and actions involving a risk one is willing to take to achieve something (or willing to loose in avoiding risk taking)
- mobilise trust is to mobilise engagement and activities.



knowledge commons

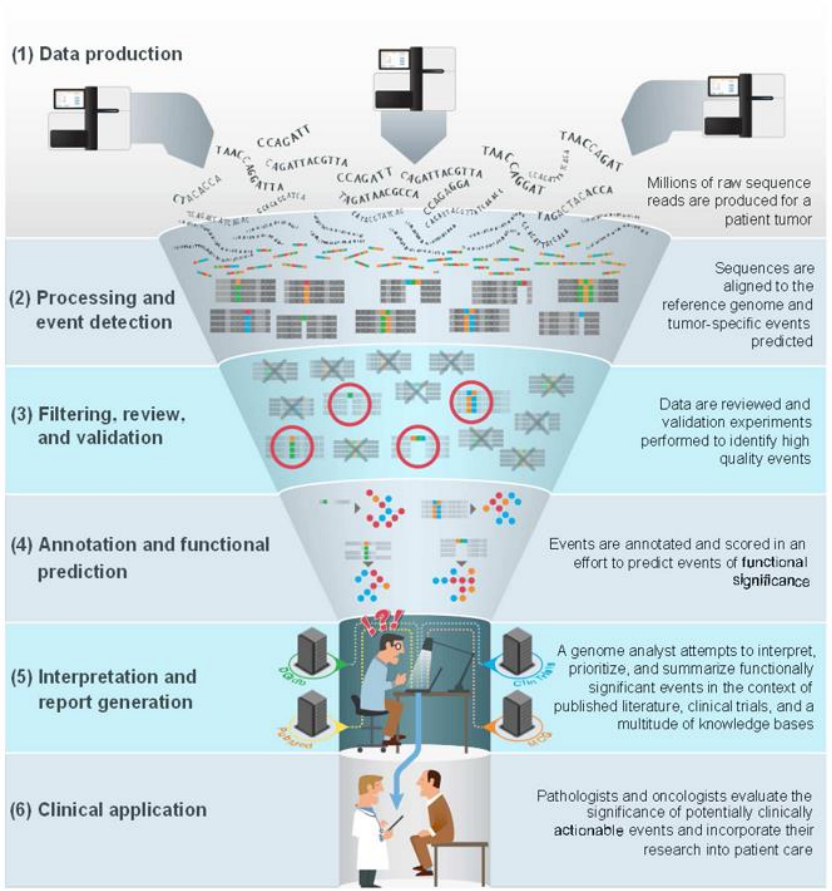
-> restabilisation – building trust

Trust

- requires engagement – pathos
- emphasis on choices and actions involving a risk one is willing to take to achieve something (or willing to loose in avoiding risk taking)
- mobilise trust is to mobilise engagement and activities.

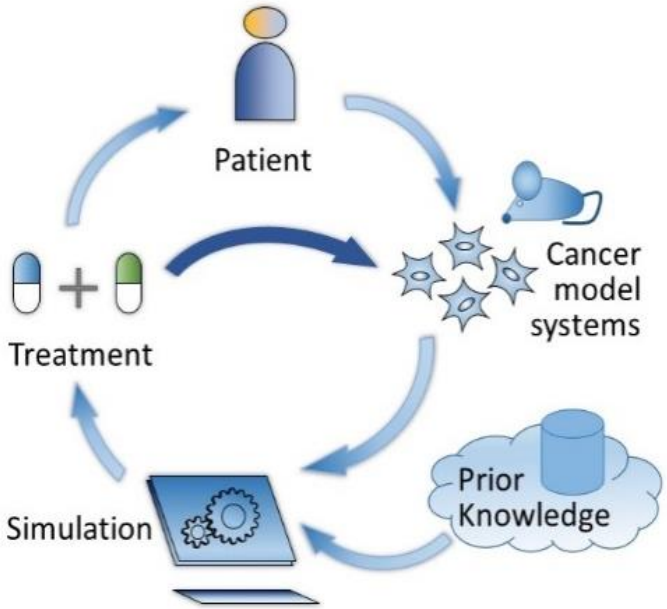
Engaging the pathos and ethos; Explicating

- what is worth doing?
- what risk is worth taking?
- what choices are available and what are the risks involved?
- what one may loose in case of risk aversion



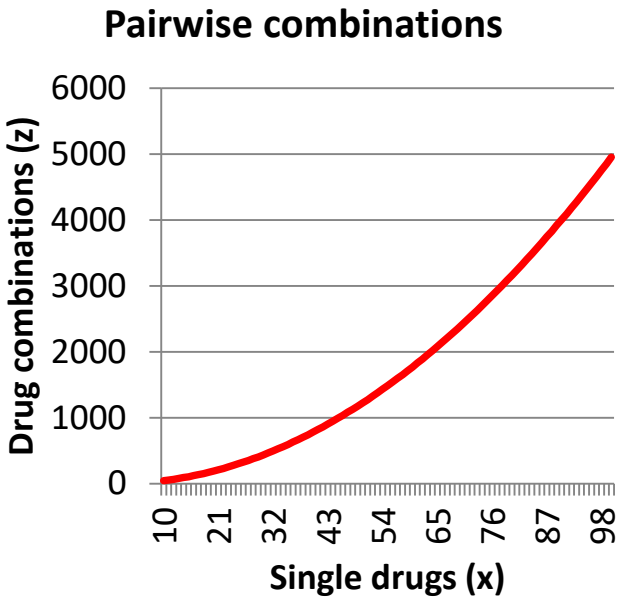
outline

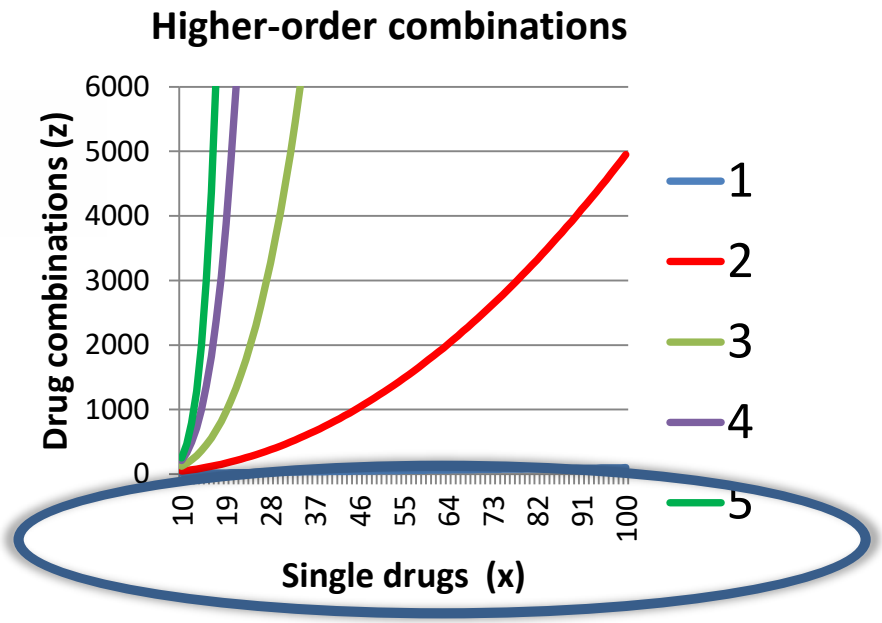
- precision oncology
- **predictive modelling**
- knowledge management
- extending knowledge commons



drug combinations

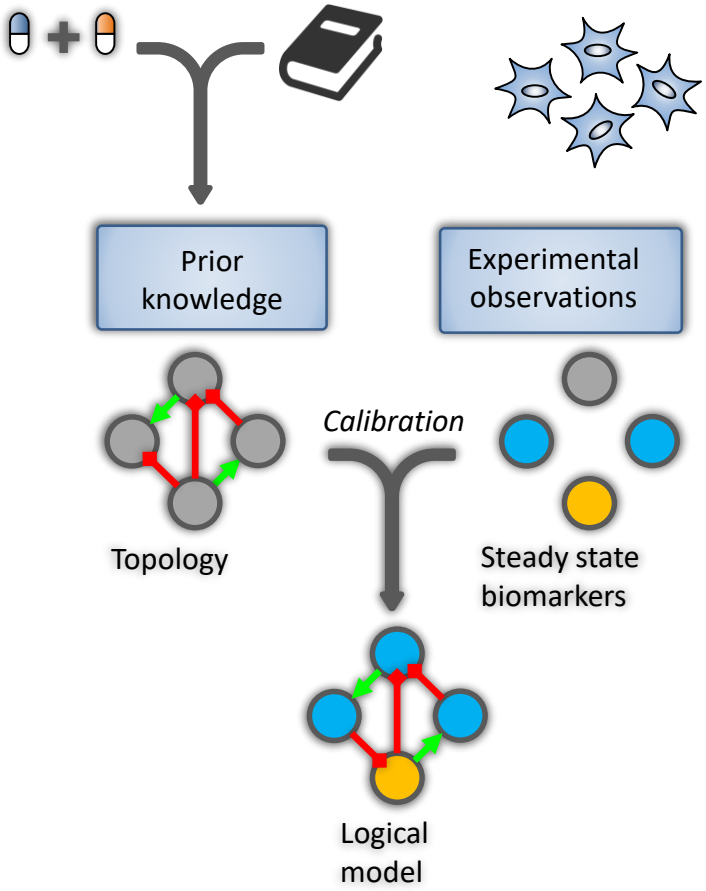
–> combinatorial explosion



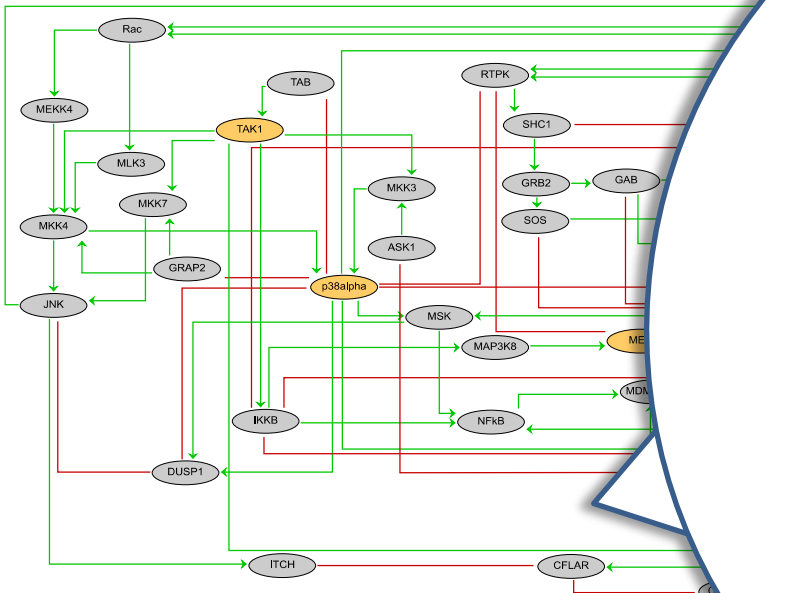


modelling

- logical, boolean models
- mechanistic, from biological background knowledge
- calibrated to cell line
- simulate combinatorial drug responses

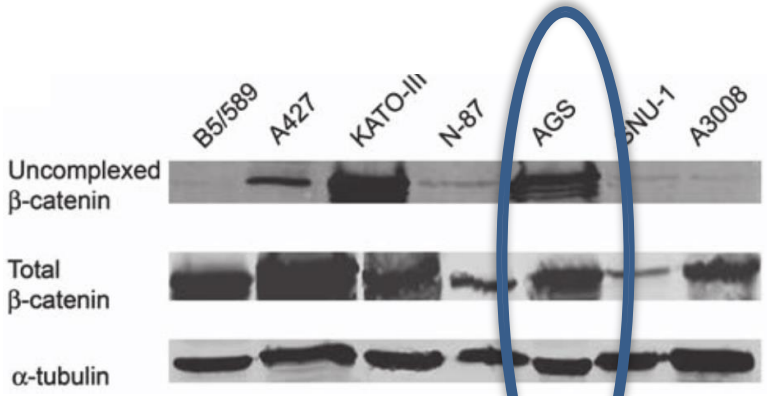


logical model for AGS cell line (GINsim)



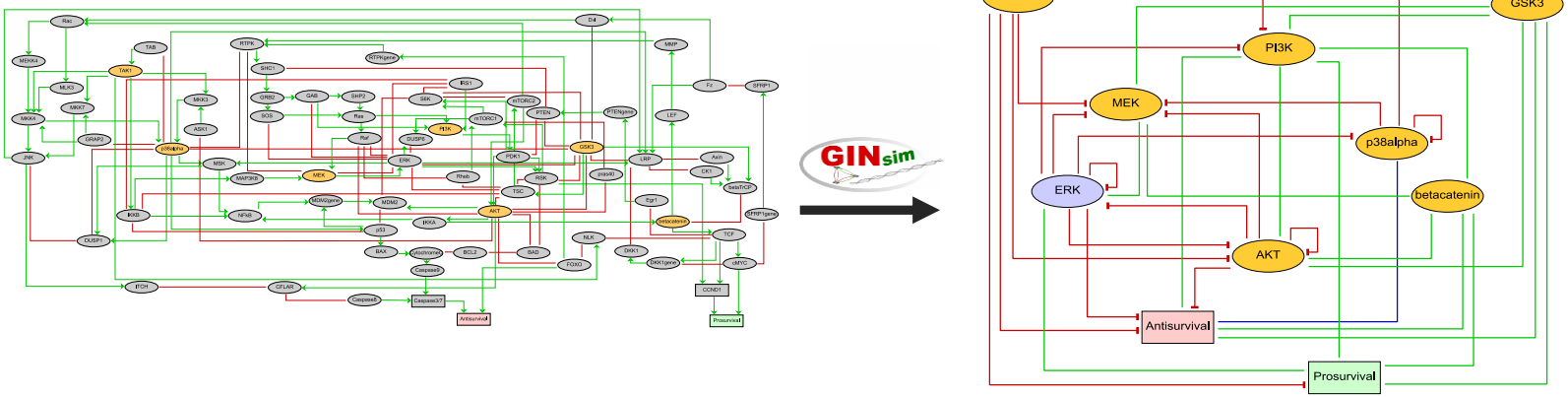
Drug targets in orange

configure with baseline biomarker states from scientific literature



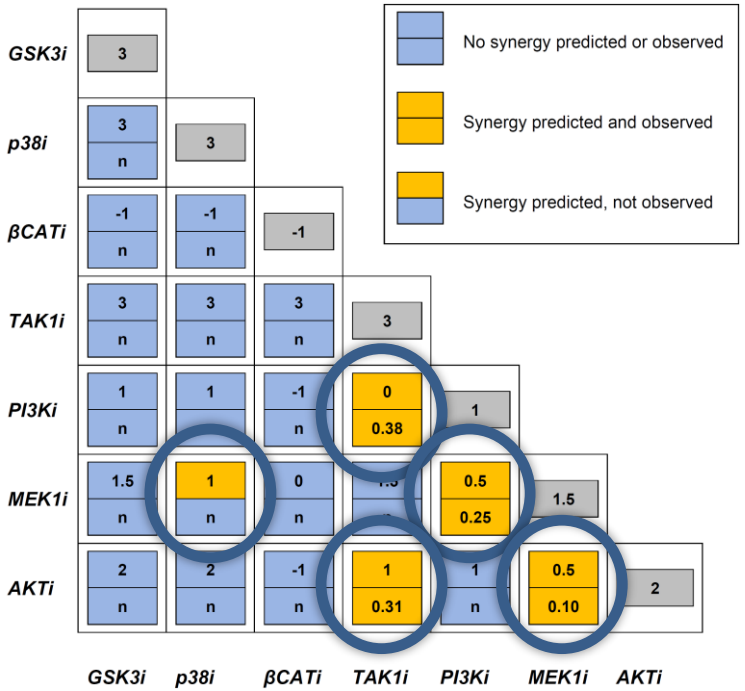
Asciutti et al. 2011

reduction: enable exhaustive simulations



Naldi et al., 2009

testing of predicted synergies



- 7 drugs: 21 pairwise combinations tested
- 20 of 21 combination effects correctly predicted
- 5 synergies predicted, 4 verified in AGS cells
- Assessing <25% of possible combinations would suffice to discover 4 synergies (without *a priori* drug screen)

extending model and testing against multiple cell lines

144 nodes

366 interactions

test predictions

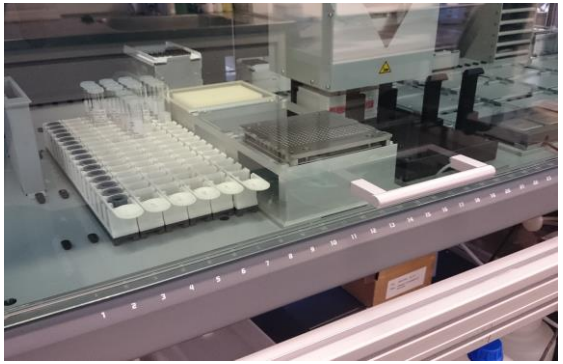
high throughput screening, cancer cell lines

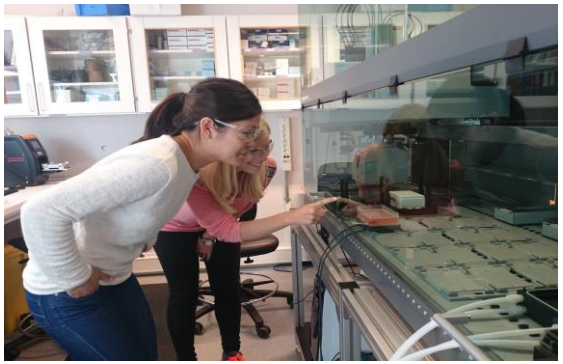
19 drugs

171 combinations

8 cell lines

A498	kidney cancer
AGS	gastric adenocarcinoma
COLO 205	colorectal cancer
DU-145	prostate cancer
MDA-MB-468	breast cancer
SF-295	glioblastoma
SW-620	colorectal cancer
UACC-62	melanoma





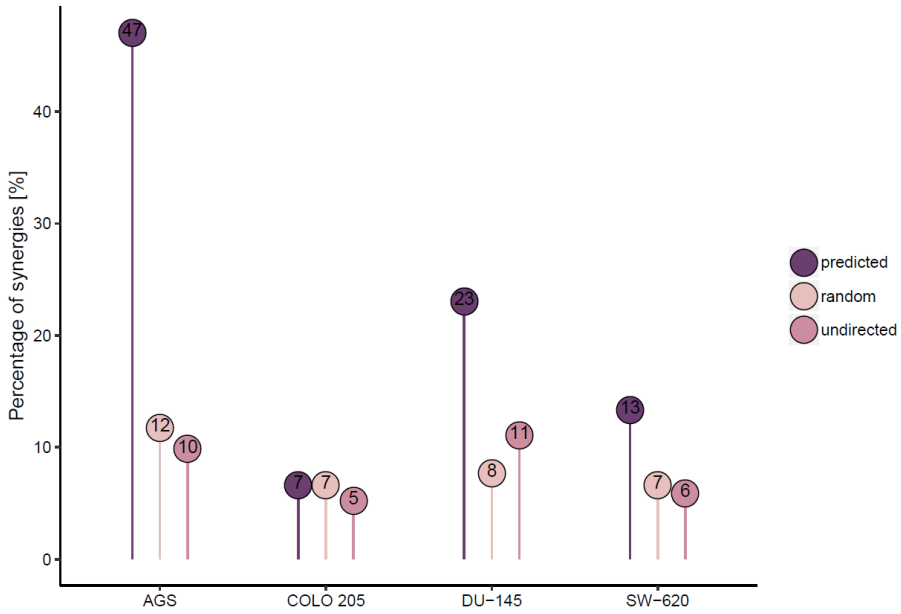
predictions, testing

configure models with baseline biomarker states

- derived from scientific literature
- inferred from large scale data using PARADIGM (Vaske, 2010)

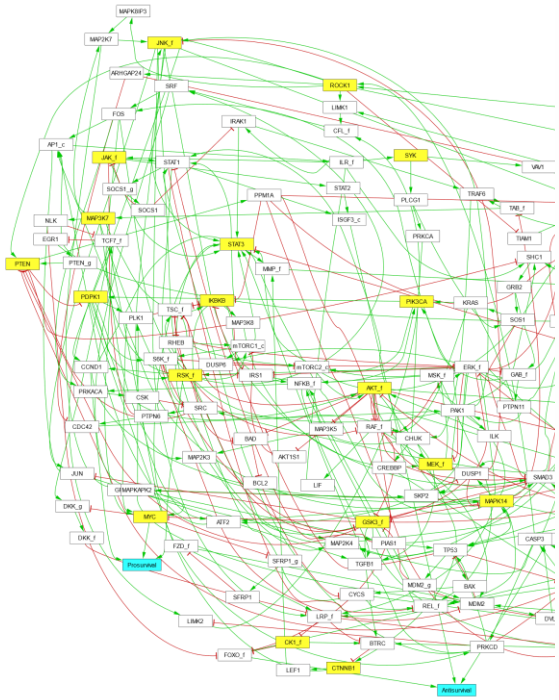
test against cell line screening

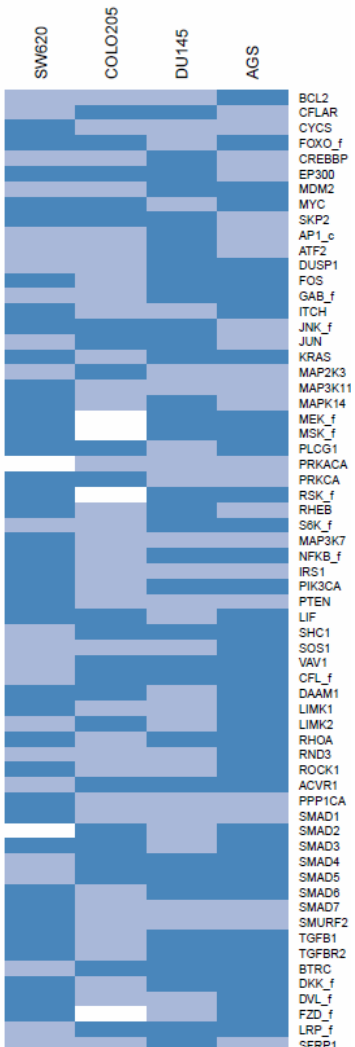
- some cell line specificity observed



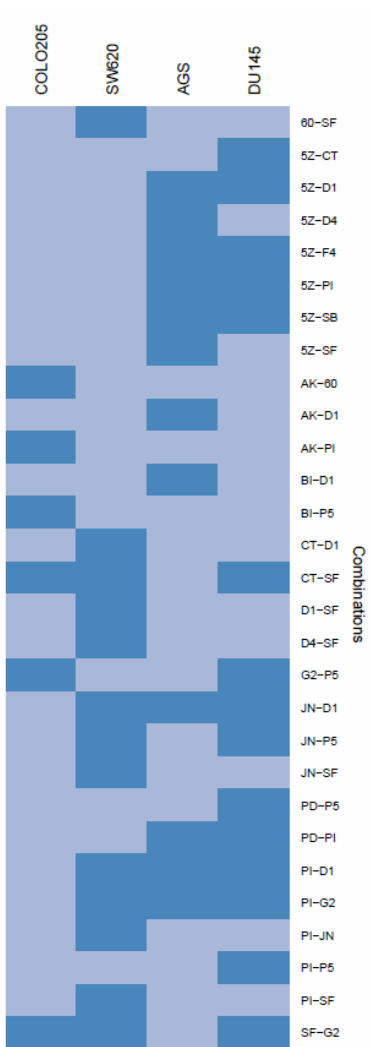
cell specific biomarker states and responses

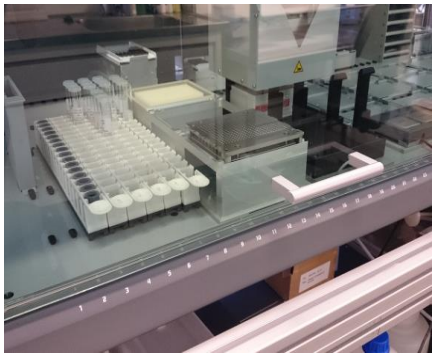
biomarker states



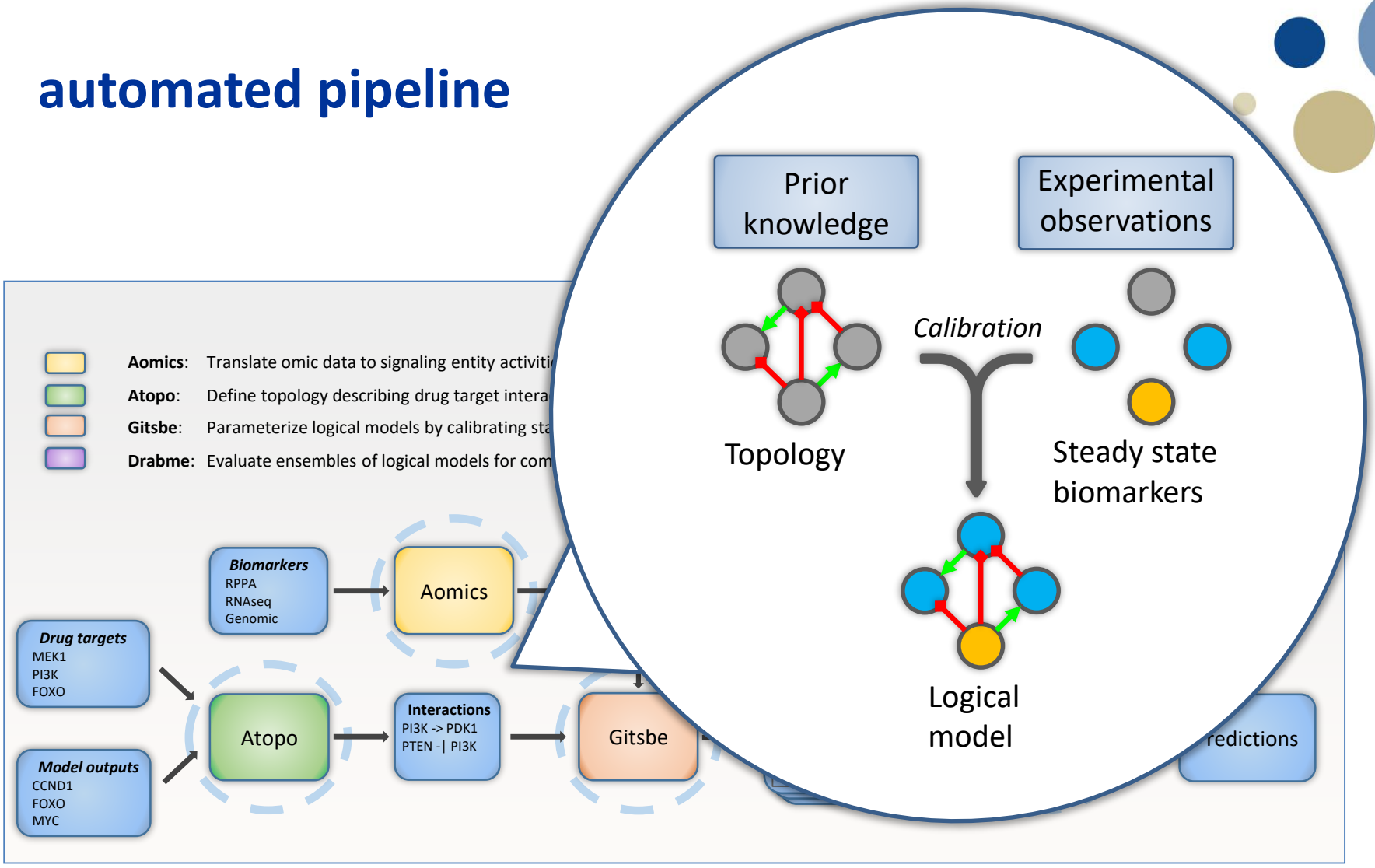


cell drug responses





automated pipeline



genetic algorithms for model configuration

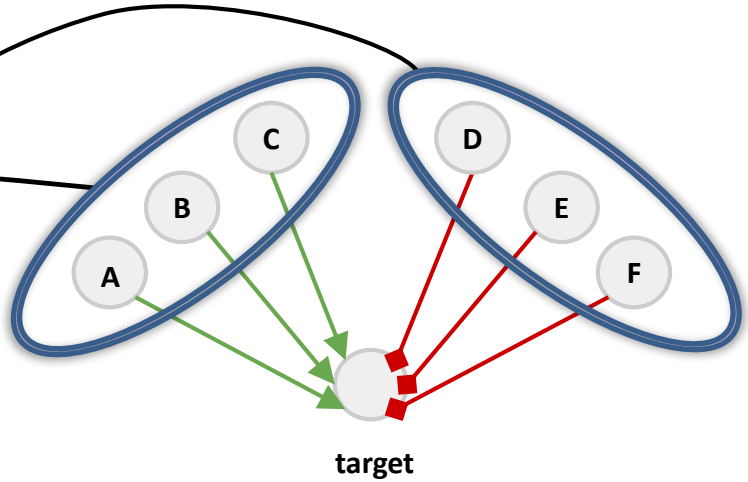
Parameterization modified

General formula:

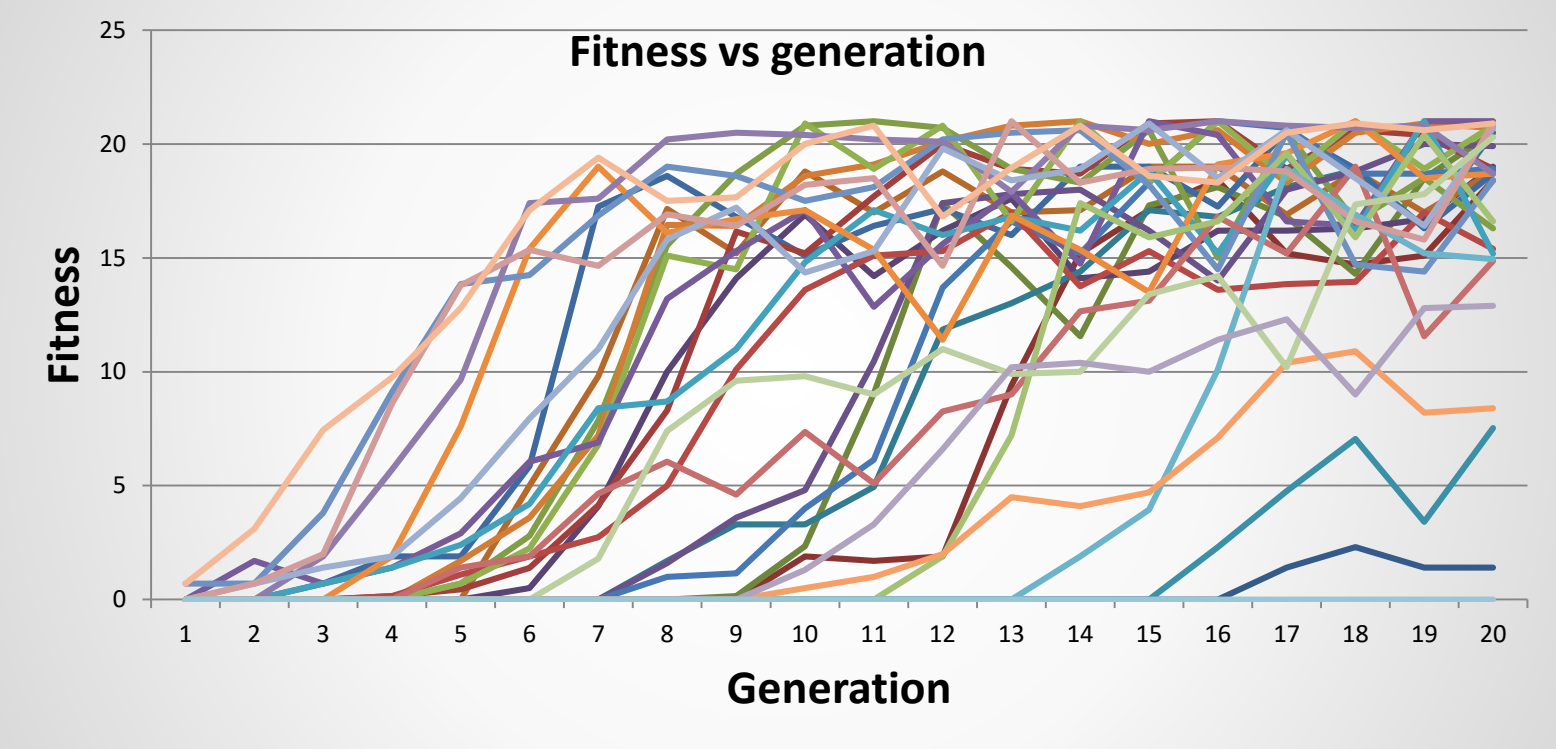
target = (A *or* B *or* C) ***and not*** (*D* or E *or* F)

Alternative formula:

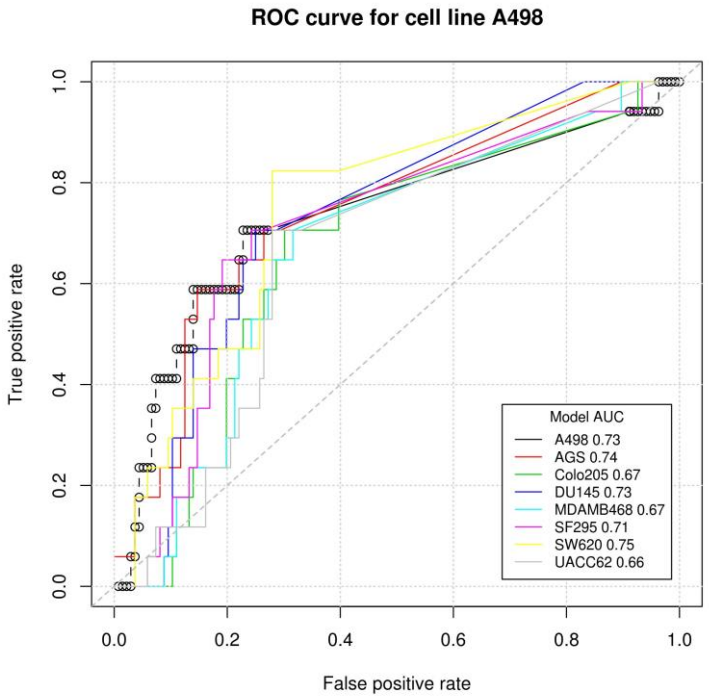
target = (A *or* B *or* C) ***or not*** (D *or* E *or* F)

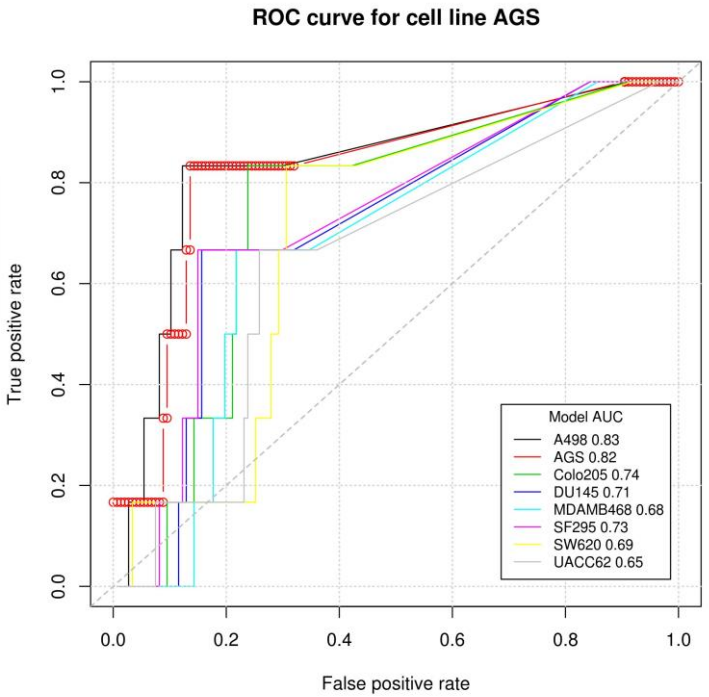


genetic algorithms for model configuration



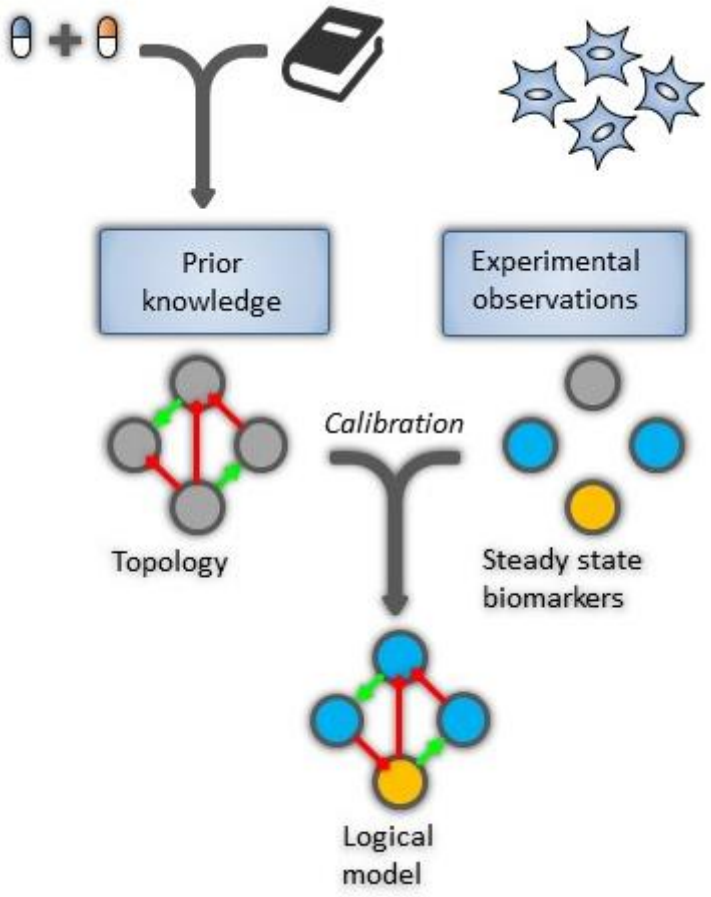
testing against cell line screening

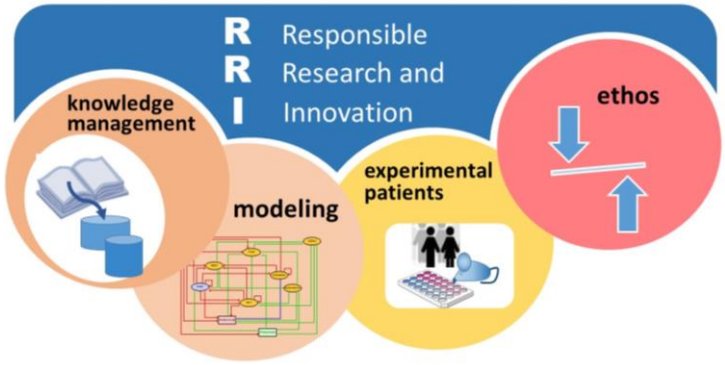




challenges, predictive modelling

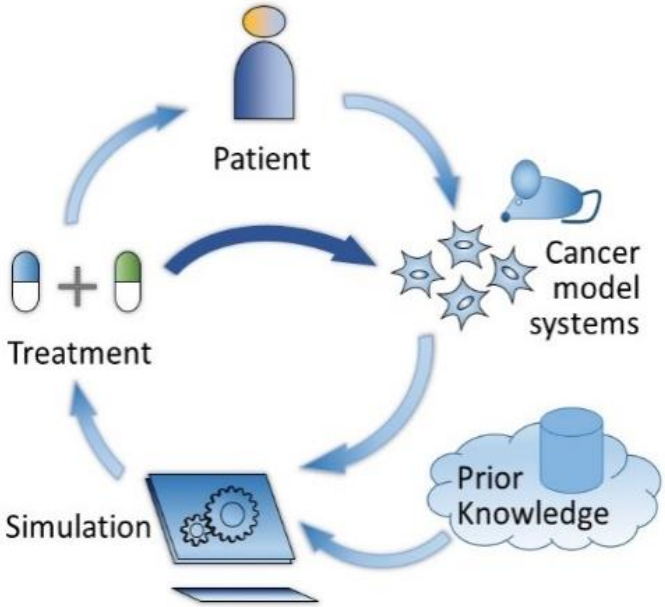
- biomarkers for configuration
  - improved inference from large scale data
  - phosphoproteomics
- knowledge for topologies
  - improved *Knowledge Commons f*or signallng mechanisms





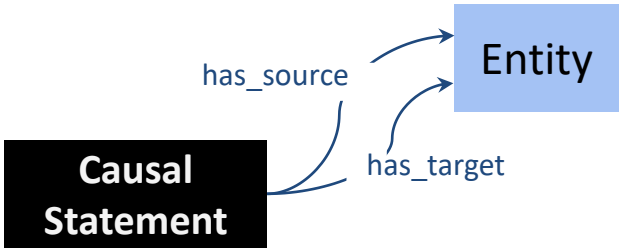
outline

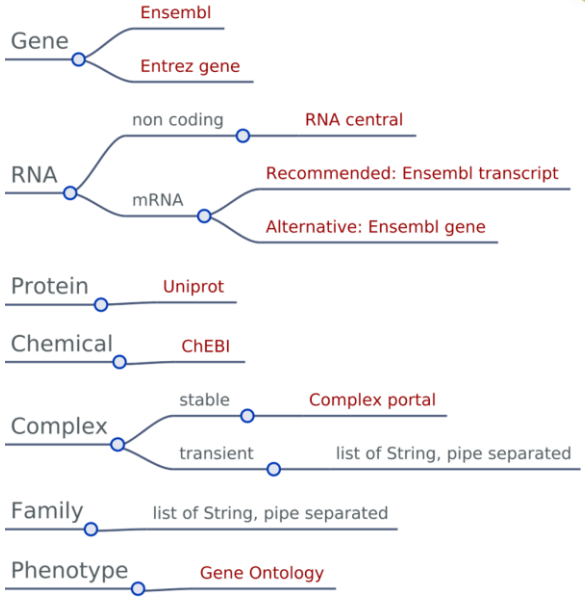
- precision oncology
- predictive modelling
- knowledge management
- **extending knowledge commons**



causal statements

standards, MiCaST

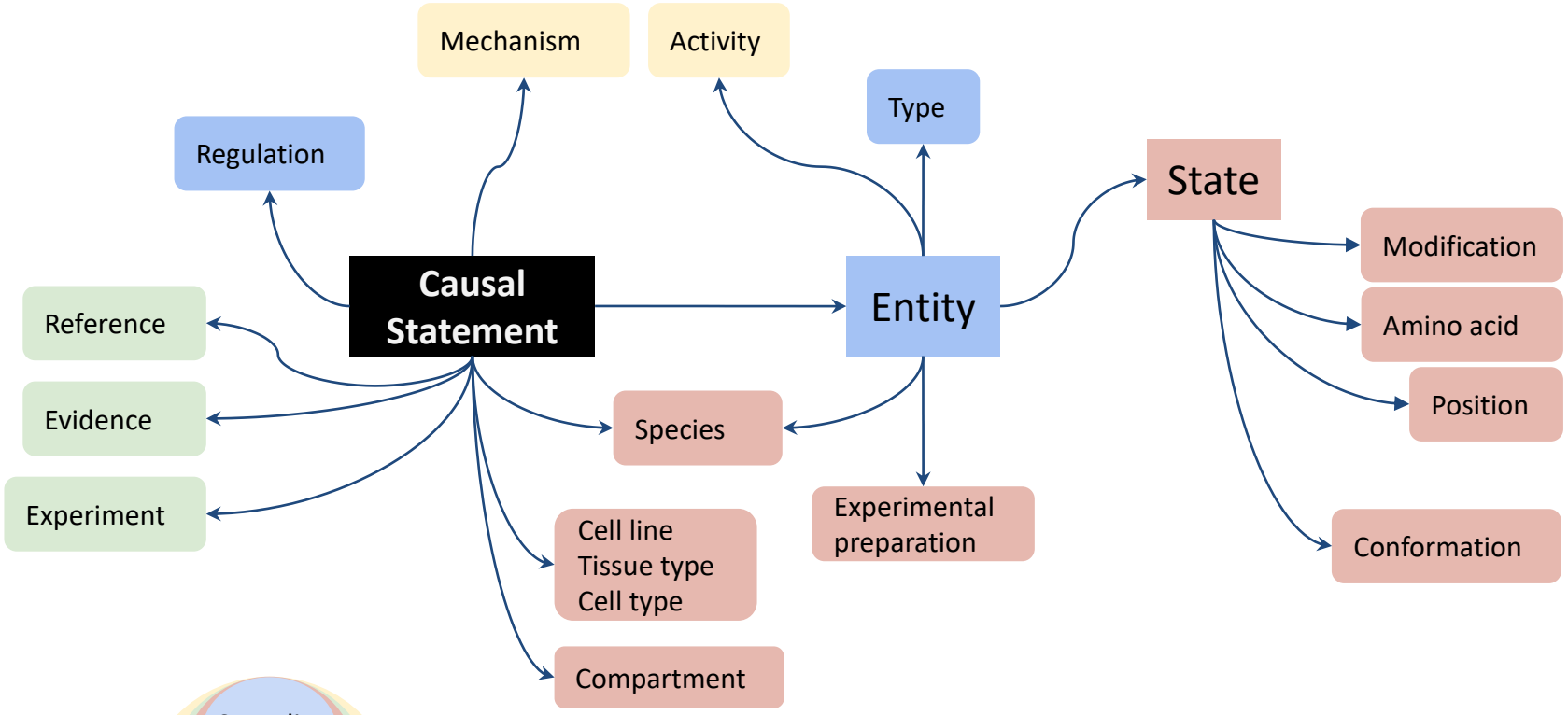


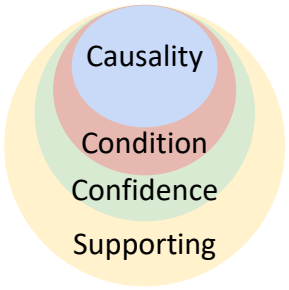


Example

Source Entity	Target Entity
P31749 (AKT1_HUMAN)	**O43524** (FOXO3_HUMAN)

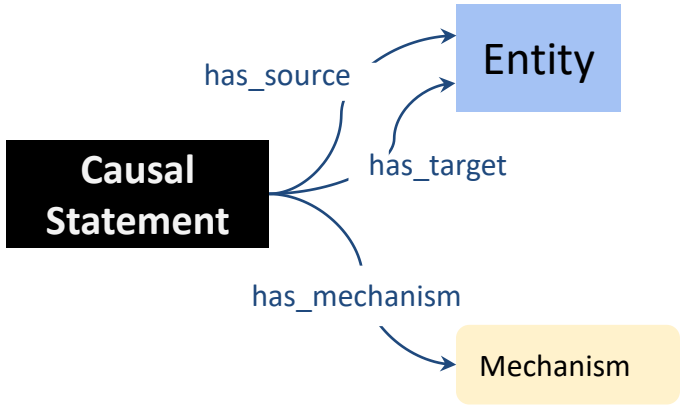
causal statements

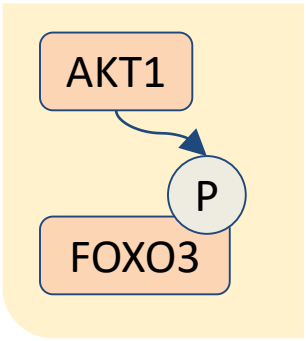




-> apply to knowledge bases in the *Knowledge Commons*

causal statements





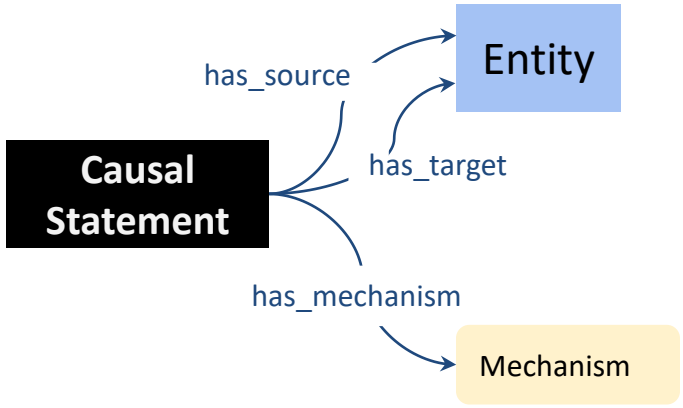
Example

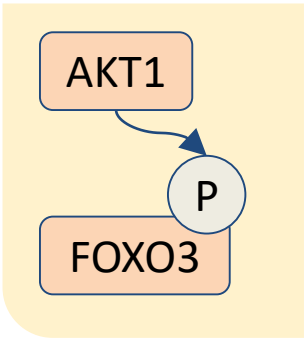
Source Entity	Regulation	Target Entity
AKT1	MI:2241	FOXO3

Mechanism

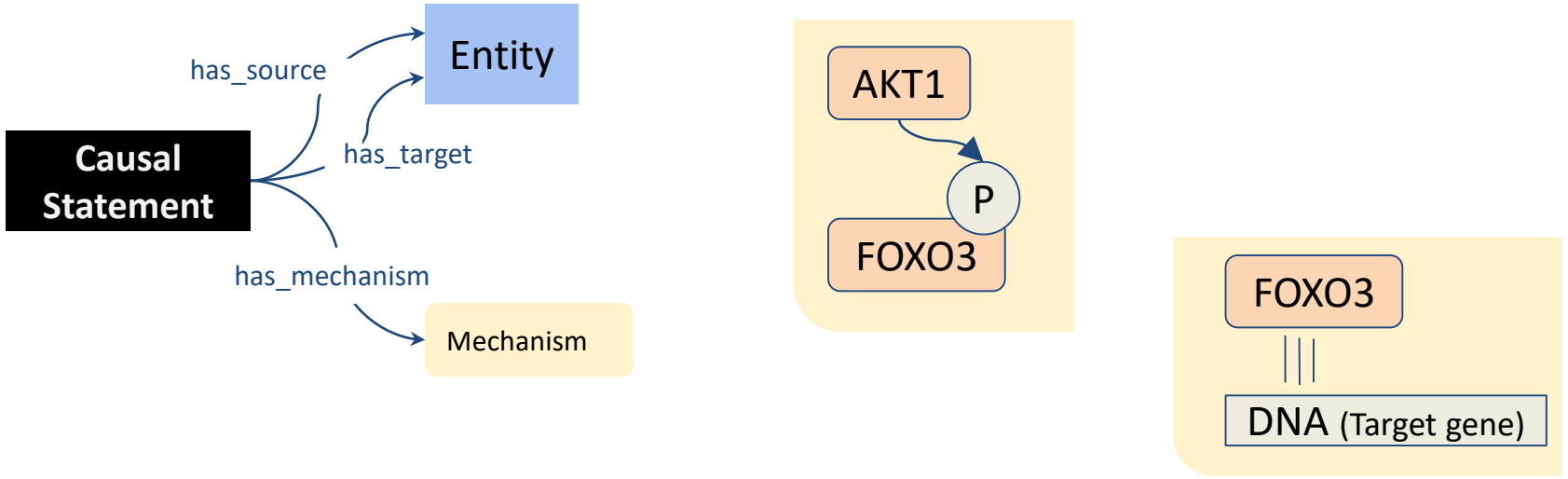
MI:0217 (phosphorylation reaction)

causal statements -> VSM visual synthax method

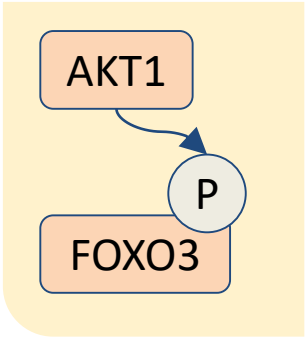


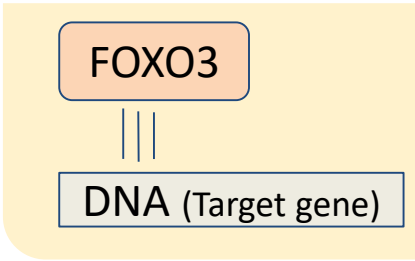


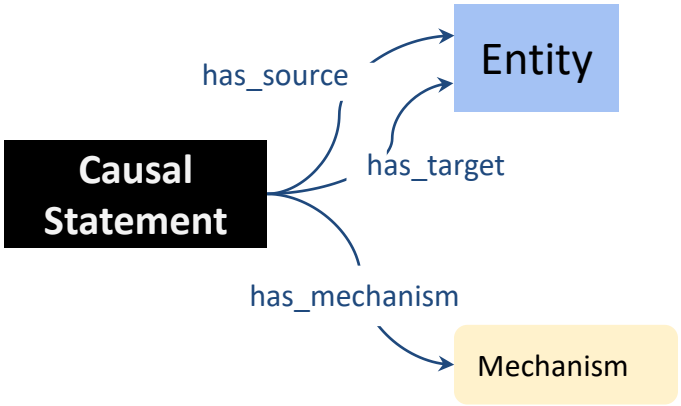
causal statements -> VSM visual synthax method

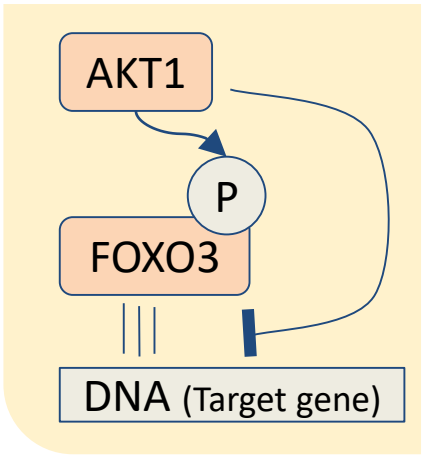


causal statements -> VSM

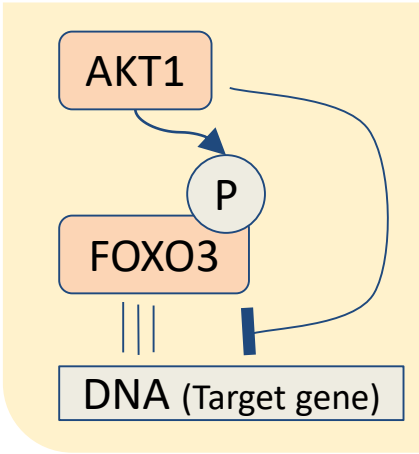


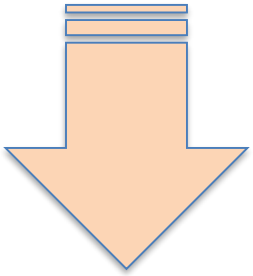




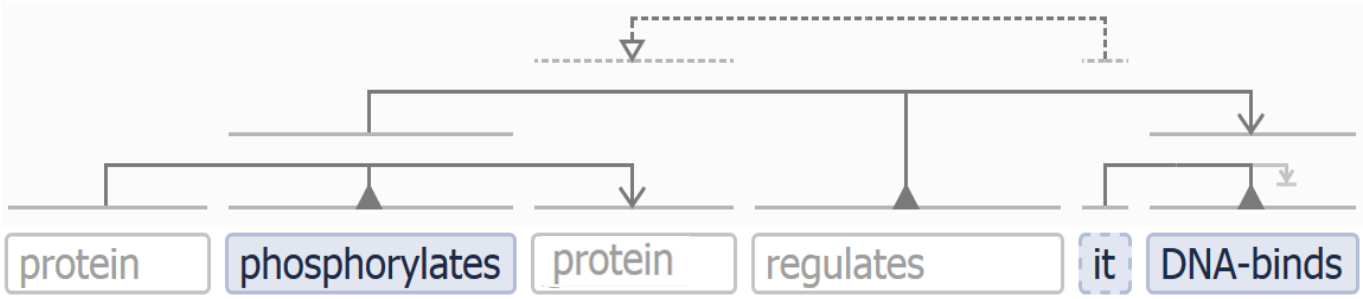


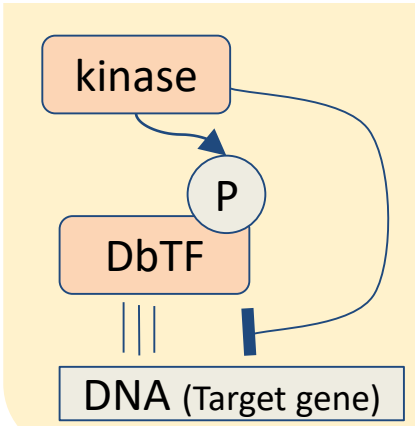
causal statements -> VSM





curation templates

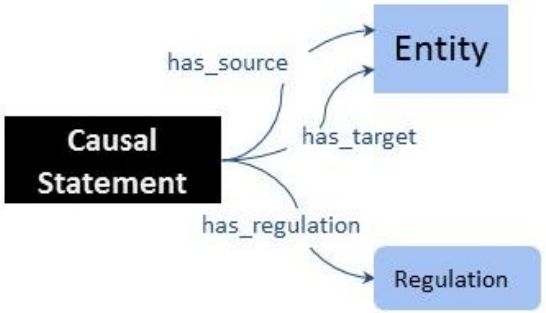




ExTRI

Extraction of Transcription Regulation Interactions

Text mining of scientific literature



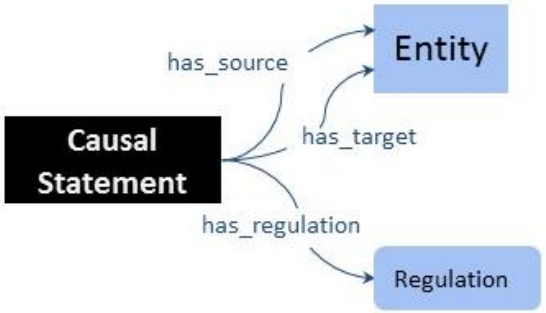
cnio
Centro Nacional de Investigaciones Oncológicas

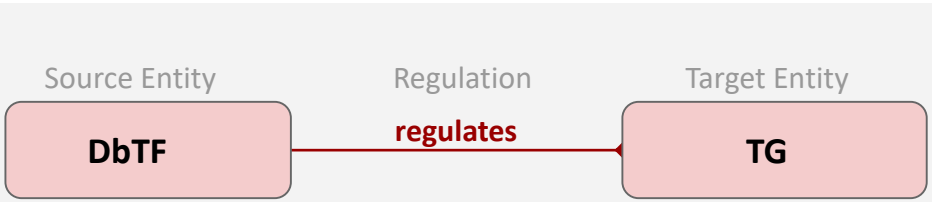
BSC Barcelona Supercomputing Center
Centro Nacional de Supercomputación

ExTRI

Extraction of Transcription Regulation Interactions

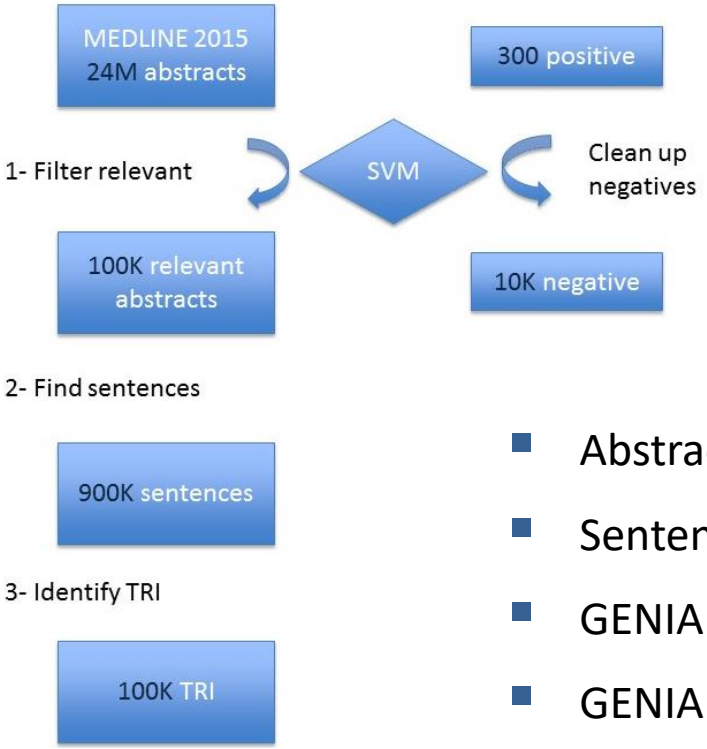
Text mining of scientific literature





ExTRI

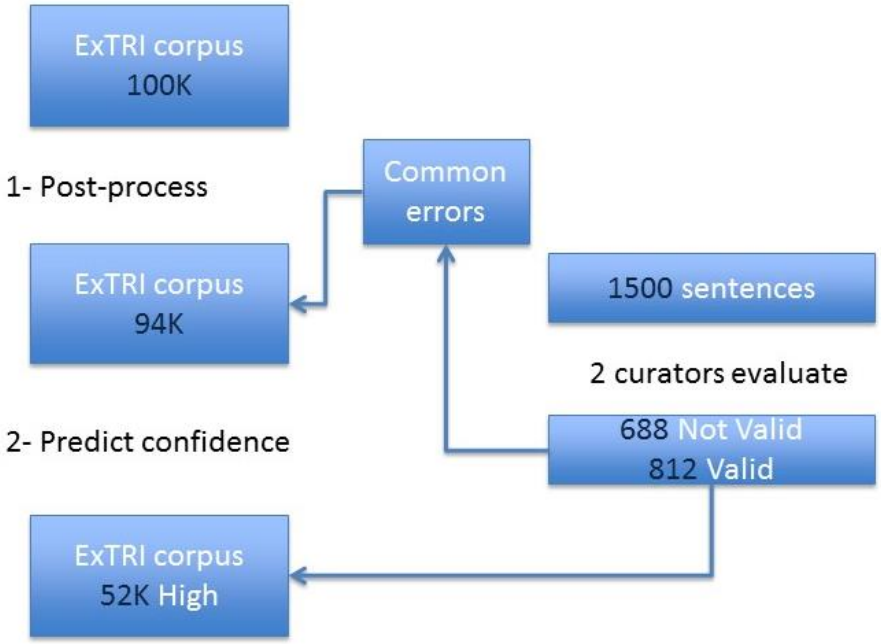
Extraction of Transcription Regulation Interactions



- Abstract classification key
- Sentence splitting by segtok
- GENIA Tagger annotation, Part of speech, Chunking
- GENIA NER
- Gene names normalized into mamalian nomenclature
- Special consideration for DbTFs

ExTRI

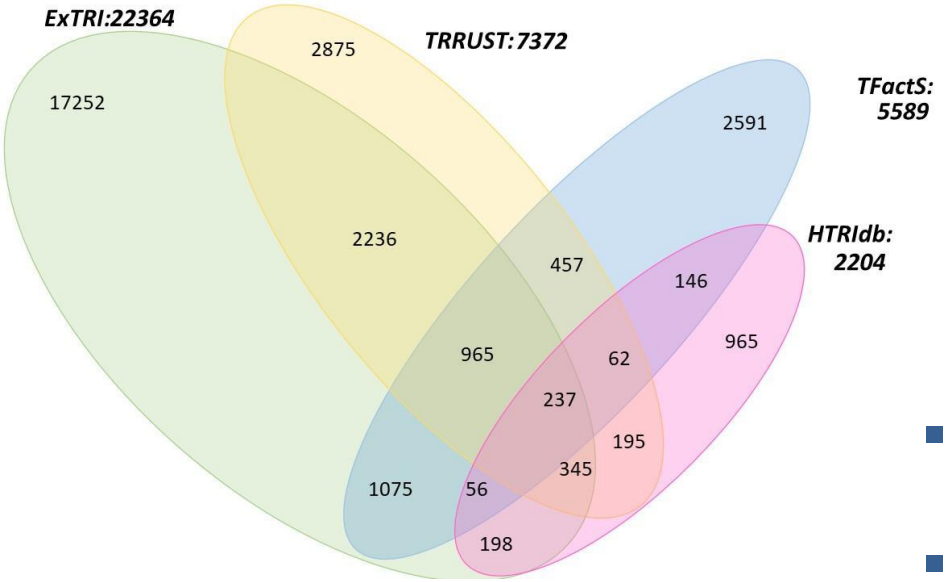
Extraction of Transcription Regulation Interactions



- sentence evaluation, roughly half correct
- recurrent errors identified; expressed as post-processing rules
- Classification for high and low confidence using Random Forest
  - Number of interactions per sentence
  - Number of times interaction appears in abstract
  - Number of abstracts with that interaction
  - Score from the TRI classifier
- F-Score 0.74 (CV) of for high confidence in the evaluation set

ExTRI

Extraction of Transcription Regulation Interactions



- many DbTF – TG overlapping with info from other resources
- 17,5k DbTF-> TG ***only from ExTRI***
- covers 800 of 1600 DbTFs
- roughly half of DbTFs described in literature for TG relations

ExTRI -2 *on-going*

Extraction of Transcription Regulation Interactions

Text mining of scientific literature

NTNU – Trondheim
Norwegian University of
Science and Technology

THE BIOGATEWAY RESOURCE

A Semantic Systems Biology Database

BioGateway consists of a graph-based database built on Semantic Web principles, a SPARQL endpoint allowing users to query it, and a Cytoscape app which integrates the query functionality directly into your network building workflow.

Get the Cytoscape App

THE BIOGATEWAY RESOURCE

A Semantic Systems Biology Database

BioGateway consists of a graph-based database built on Semantic Web principles, a SPARQL endpoint allowing users to query it, and a Cytoscape app which integrates the query functionality directly into your network building workflow.

What is BioGateway?

BioGateway is an initiative that enables a Semantic Systems Biology approach. It provides an entry point to access a data warehouse where biological data is gathered in the form of triples (using RDF). The systems can be queried using SPARQL. The BioGateway system can also be explored using the SPARQL browser. With this browser, SPARQL results can be visually seen as a network of resources.

The BioGateway Database

Unified Identifiers

Every entity in BioGateway has a unique identifier URI across all datasets – allowing queries across data from different sources.

High-confidence Data

The data in BioGateway is a combination of the most trusted datasets from UniProt, IntAct and other curated sources.

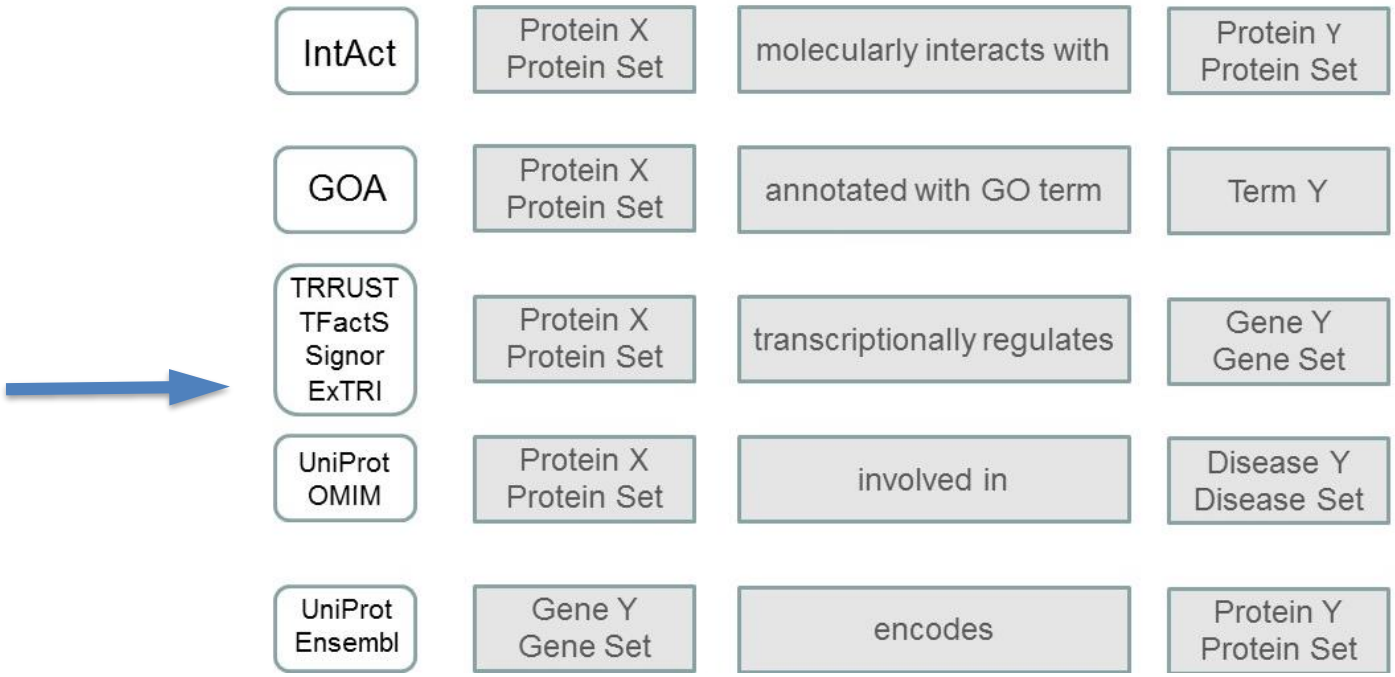
Semantic Web Technologies

BioGateway combines Systems Biology and Semantic Web technologies for more effective modelling of biological systems.

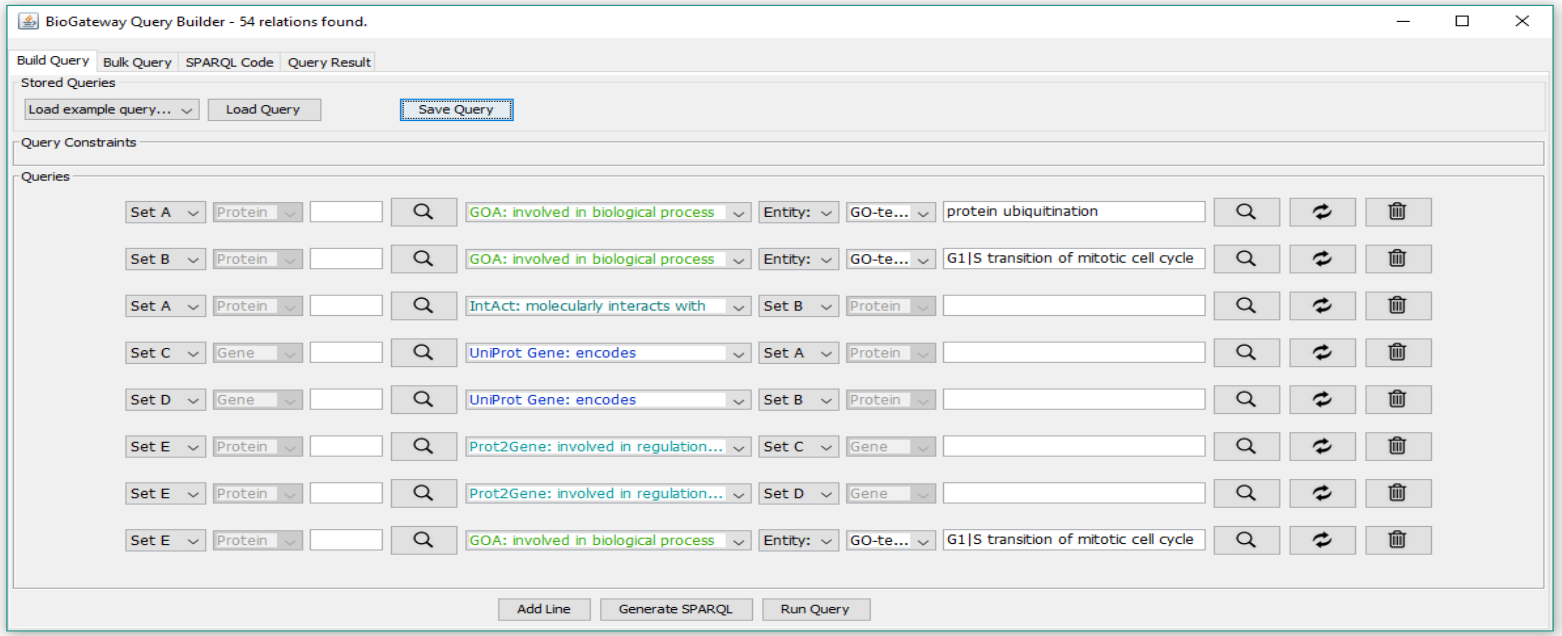
Explorative Data

Network builders can explore the use of candidate Transcription factor – Target gene network connections obtained through text mining and check their validity in the original abstract.

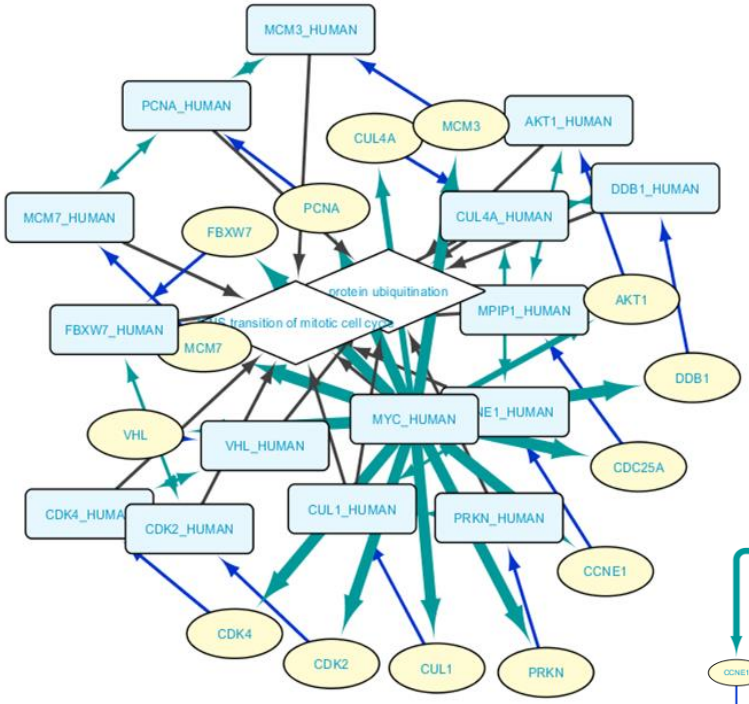
database content



querying BIOGATEWAY

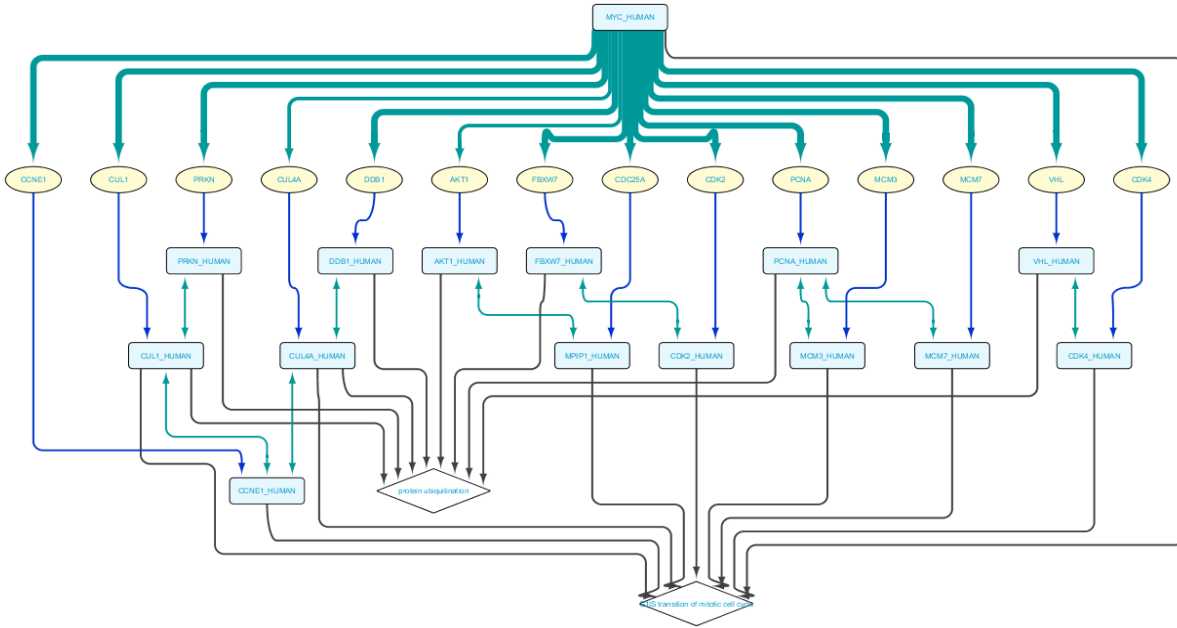


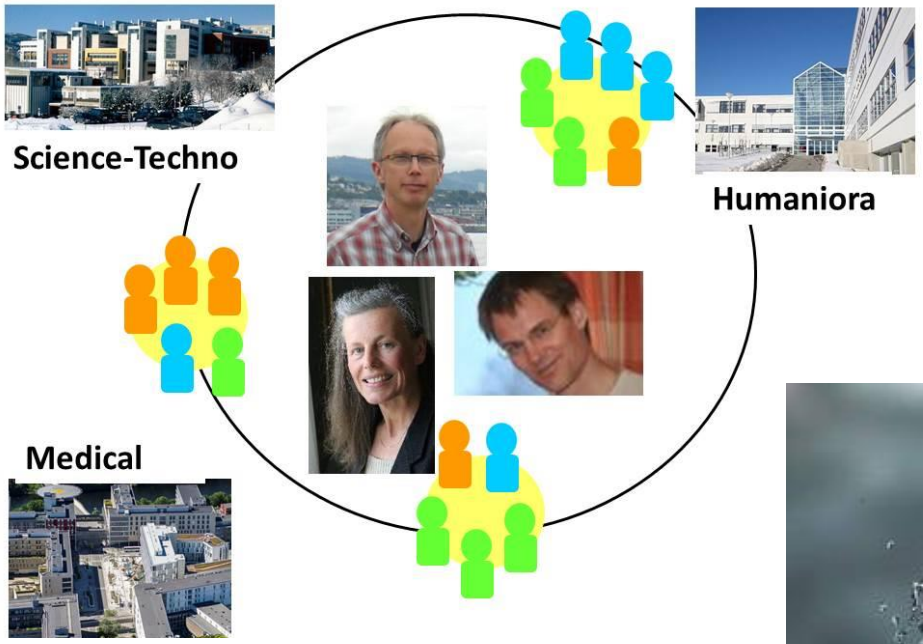
BIOGATEWAY Cytoscape app



visual representation

confidence level filtering




Science-Techno
Humaniora
Medical

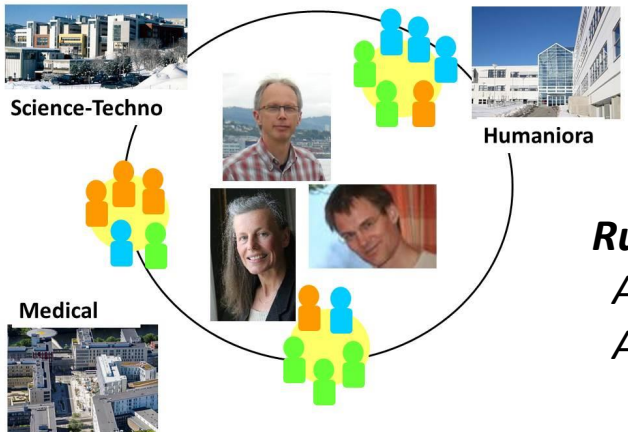
DrugLogics www.druglogics.eu

Martin Kuiper
Steven Vercruysse
Wim de Mulder
Vasundra Touré
John Zobolas
Vladimir Mironov
Rafel Riudavets
Stian Holmås

SINTEF

High throughput screening facility, Trondheim

Geir Klinkenberg
Vu To Nakstad



Rune Nydal
Ane Møller Gabrielsen
Anamika Chaterjee

Astrid Lægreid
Åsmund Flobak
Miguel Vazquez
Marcio Luis Acencio
Liv Thommesen
Eva Hofsli
Barbara Niederdorfer
Evelina Folkesson
Kathleen Heck

Strategic Research Area 2014–2023
NTNU HEALTH

COLOSYS
ERACO SYSMED

CENTRE FOR DIGITAL LIFE NORWAY

DrugLogics

Martin Kuiper

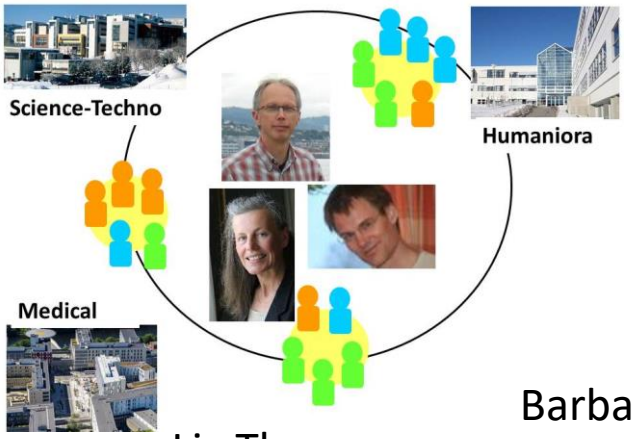
Steven Vercruysse
Wim de Mulder
Vasundra Touré
John Zobolas
Vladimir Mironov
Rafel Riudavets
Stian Holmås

Ane Møller Gabrielsen

Anamika Chaterjee

Rune Nydal

Ane Møller Gabrielsen
Anamika Chaterjee



Vasundra Touré

Astrid Lægreid

Åsmund Flobak
Miguel Vazquez
Marcio Luis Acencio
Liv Thommesen
Eva Hofsli
Barbara Niederdorfer
Evelina Folkesson
Kathleen Heck

Liv Thommesen

Barbara Niederdorfer

Evelina Folkeson

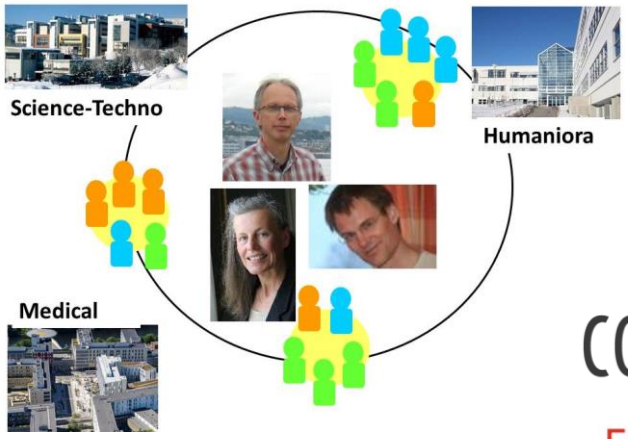
DrugLogics www.druglogics.eu

Martin Kuiper
Steven Vercruysse
Wim de Mulder
Vasundra Touré
John Zobolas
Vladimir Mironov
Rafel Riudavets
Stian Holmås

SINTEF

High throughput screening facility, Trondheim

Geir Klinkenberg
Vu To Nakstad



Rune Nydal
Ane Møller Gabrielsen
Anamika Chaterjee

Astrid Lægreid
Åsmund Flobak
Miguel Vazquez
Marcio Luis Acencio
Liv Thommesen
Eva Hofsli
Barbara Niederdorfer
Evelina Folkesson
Kathleen Heck

COLOSYS

ERACO SYSMED

cnio Centro Nacional de Investigaciones Oncológicas

Strategic Research Area 2014–2023
NTNU HEALTH

CENTRE FOR DIGITAL LIFE NORWAY

BSC Barcelona Supercomputing Center
Centro Nacional de Supercomputación