Cambridge Cell Networks

Providing an Insight into your Biological Systems

Mechanism-based Applications of Toxicology Ontology

Gordana Apic, PhD
(presented by Dragana Mitic Potkrajac, PhD)

OpenTox 2011 InterAction Meeting Program,
9-12 August 2011, Munich

I. Mechanism-based application requires complex and structured information

II. Ambigious how detailed information must be:

- in some cases inferred protein-chemical interaction networks are sufficient
- in other cases detailed models of *how* a particular pathway brings about a particular outcome is required

CCNet's ToxOntology

1. Organizing and managing information about toxic effects

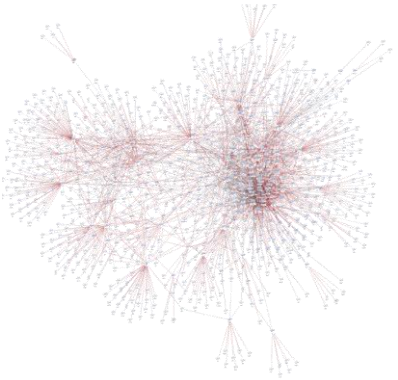
2. Providing as much insight as possible about the mechanisms underlying a toxic effect

3. Helping elucidate underlying biological pathways

Aim is to interpret and explain mechanism of action of drugs/compounds by precisely categorizing terms describing observation/toxicities

CCNet's ToxWiz™ Ontologies

ToxWiz
toxicology ontology



What are they?

What purpose to these ontologies serve?

What are the requirements to create this ontology?

How do we create, validate and optimize the ontology?

CCNet's ToxWiz™ Ontologies

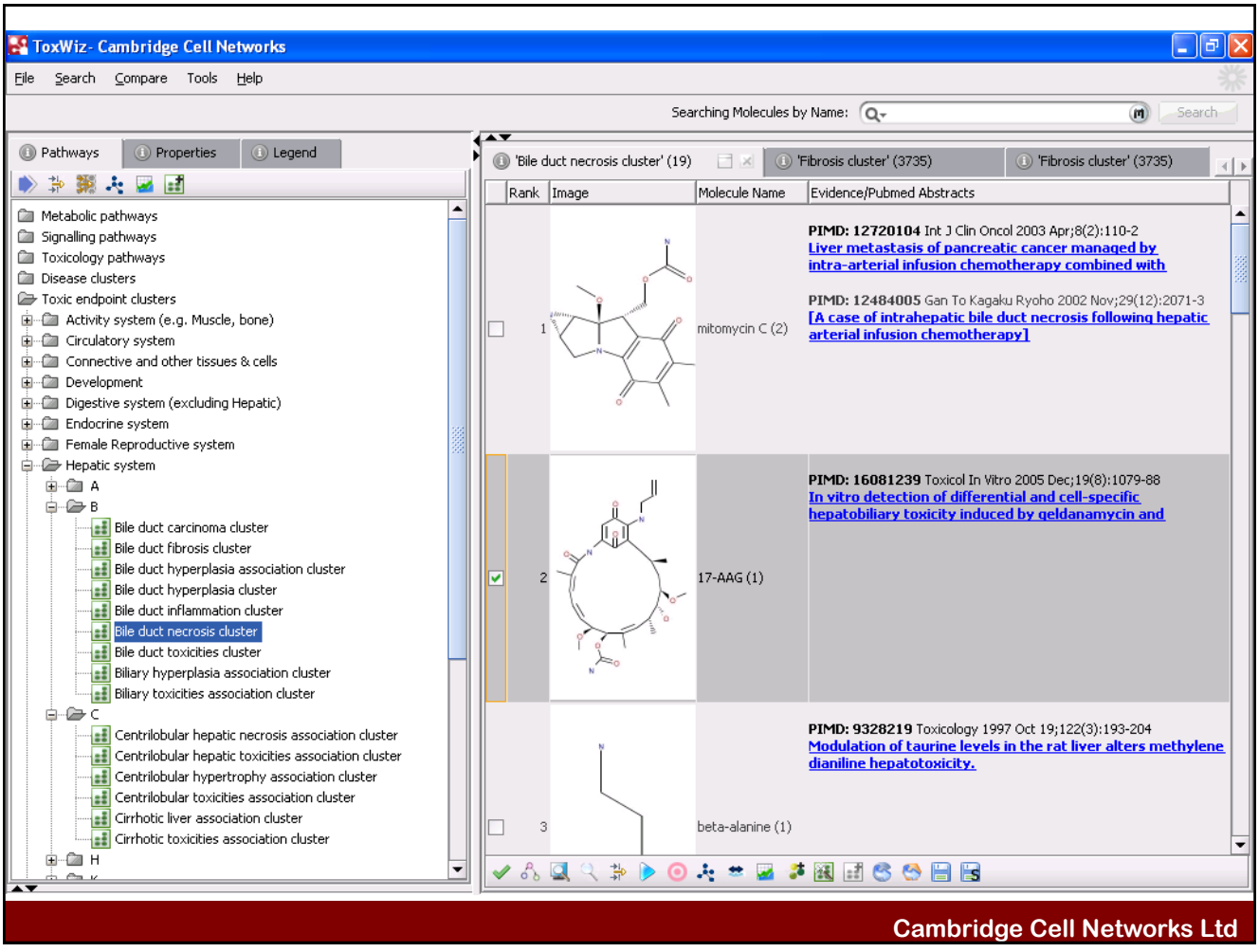
ToxWiz
toxicology ontology

What are they?

- **Controlled vocabularies/terms**
- **Classified into groups & hierarchies**
- **Defined relationships between terms and between groups of terms**

- **A product of over 150 man years of expert curation and toxicology information extraction from safety reports and literature**
- **6 iterations over 8 years for harmonization between projects lead to increased interoperability**
- **912 histopathology terms and 500000 synonyms, over 2 million relationships**

What purpose to these ontologies serve?
What are the requirements to create this ontology?
How do we create, validate and optimize the ontology?



Cambridge Cell Networks Ltd

CCNet's ToxWiz™ Ontologies

ToxWiz
toxicology ontology

What are they?

What purpose do the ontologies serve?

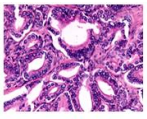
- **To facilitate prediction of toxic effects –** *prospective analysis*
- **To help explain causes of toxic effects –** *retrospective analysis*
- **To elucidate modes of action and create hypothesis for MOA**

- **To support extraction process of knowledge relevant to toxicology from toxicology reports and literature**
- **To enable integration and transfer of findings to clinical observations**

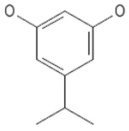
What are the requirements to create this ontology?
How do we create, validate and optimize the ontology?

Ontology supported discovery

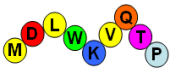
Pathology



Structure

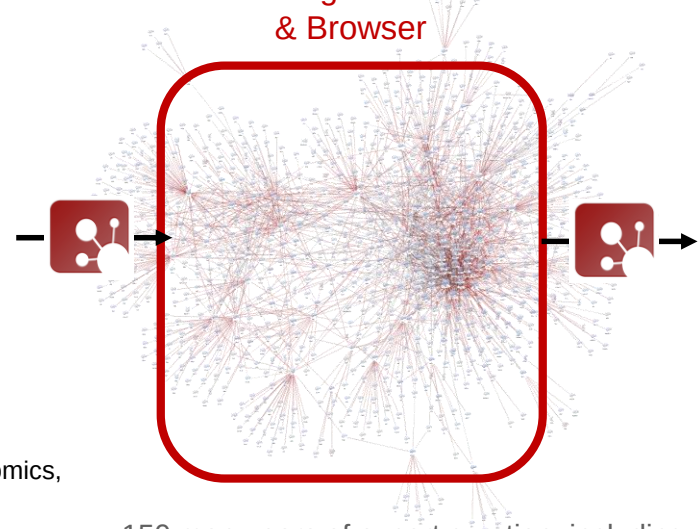


Target/receptor



Genomics, proteomics, metabolomics

Toxwiz
Knowledge Database
& Browser



150 man-years of expert curation, including over 1000 catalogued toxic endpoints & associated mechanistic hypotheses

Toxicity & pharmacology prediction

Mechanistic hypotheses

Biomarker identification

Case Study: Phthalic Acid and Testis Degeneration
How Ontologies permit faster interpretation

Parameters and Statistics

Toxic endpoint clusters

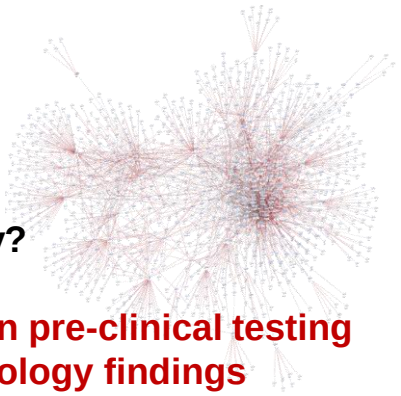
Rank	Molecular Mechanism	Type	Molecules Inside	Molecules Next To
1	Testis atrophy induction cluster	Male Reproductive system	3: Dipentyl phthalate, dibutyl phthalate, Di...	
2	Testis atrophy association cluster	Male Reproductive system	4: Dipentyl phthalate, Mehp, dibutyl phtha...	1: Dioctyl phthalate
3	Testis toxicities association cluster	Male Reproductive system	4: Dipentyl phthalate, Mehp, dibutyl phtha...	1: Dioctyl phthalate
4	Leydig cell hyperplasia cluster	Male Reproductive system	7: Palatinol IC, Benzyl butyl phthalate, Mo...	4: phthalic acid, Dioctyl phthalate, Palatinol N, Didp
5	Leydig cell toxicities cluster	Male Reproductive system	7: Palatinol IC, Benzyl butyl phthalate, Mo...	4: phthalic acid, Dioctyl phthalate, Palatinol N, Didp
6	Testis atrophy cluster	Male Reproductive system	9: Monooctyl phthalate, Ergoplast FDC, Di...	4: Palatinol N, Benzyl butyl phthalate, BUTYL ISODECYL PH...
7	Heart fibrosis inhibition cluster	Circulatory system		4: Mehp, dibutyl phthalate, Dioctyl phthalate, phthalic acid
8	Hepatocyte hyperplasia induction cluster	Hepatic system	1: Dioctyl phthalate	3: Mehp, dibutyl phthalate, phthalic acid
9	Leydig cell hyperplasia association cluster	Male Reproductive system	1: phthalic acid	1: Dioctyl phthalate
10	Spinal cord atrophy induction cluster	Nervous system		10: Palatinol IC, DIETHYL PHTHALATE, Palatinol N, Benzyl ...
11	Hepatocyte hyperplasia association cluster	Hepatic system	1: Dioctyl phthalate	3: Mehp, dibutyl phthalate, phthalic acid
12	Leydig cell toxicities association cluster	Male Reproductive system	1: phthalic acid	1: Dioctyl phthalate
13	Leydig cell hyperplasia induction cluster	Male Reproductive system	1: phthalic acid	1: Dioctyl phthalate
14	Liver steatosis inhibition cluster	Hepatic system		4: Mehp, dibutyl phthalate, Dioctyl phthalate, phthalic acid
15	Hepatocyte hypertrophy association cluster	Hepatic system		4: Mehp, dibutyl phthalate, Dioctyl phthalate, phthalic acid
16	Bile duct inflammation cluster	Hepatic system		7: Dioctyl phthalate, dibutyl phthalate, Mehp, phthalic acid...

Significantly over-represented molecular mechanism groups are shown in Red/Orange

Most significantly over-represented Organ Systems are: Reproductive System & Male Reproductive System

CCNet's ToxWiz™ Ontologies

ToxWiz
toxicology ontology



What are they?

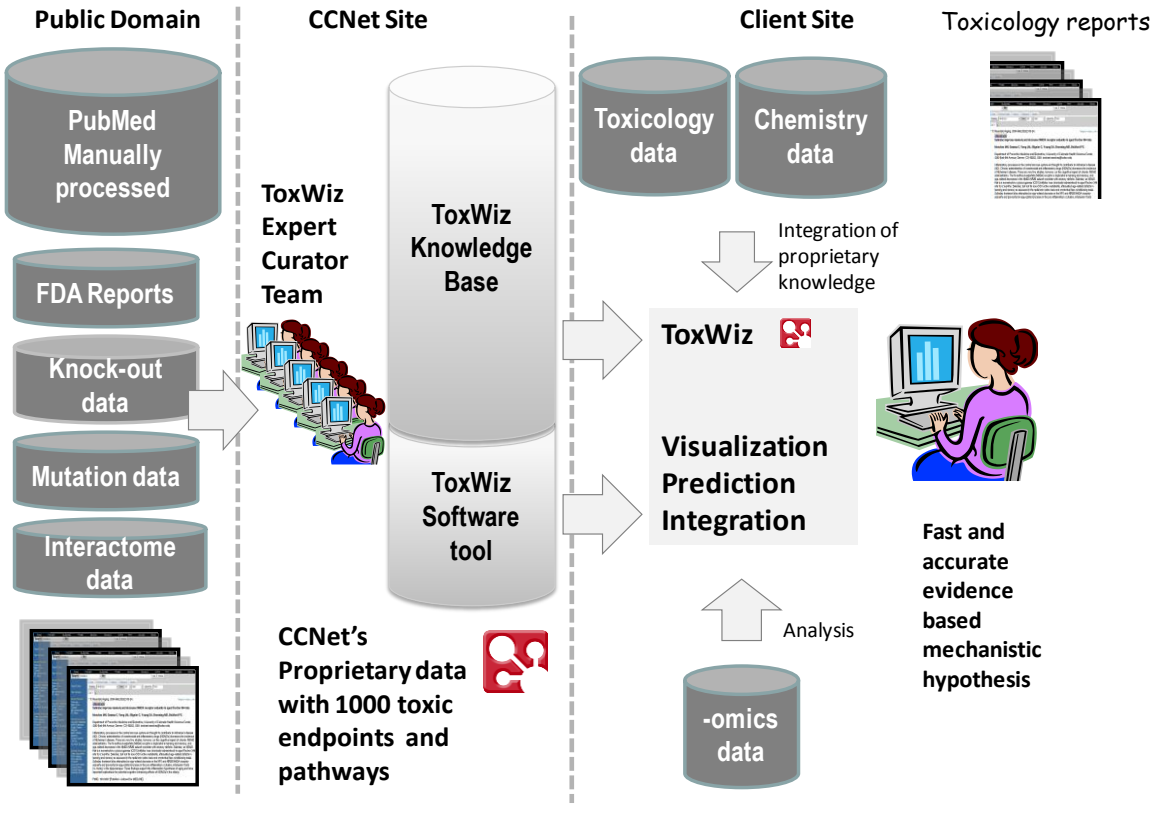
What purpose do they serve?

What are the requirements to create this ontology?

- **To capture adequately toxicology test results in pre-clinical testing**
- **To classify and define a spectrum of histopathology findings**
- **To capture long term toxicity**
- **To capture all available knowledge from the literature and toxicology reports and deal with human way of interpreting and recording the findings**
- **To be interoperable with other ontology efforts**

How do we create, validate and optimize the ontology?

Toxicity information gathering and use

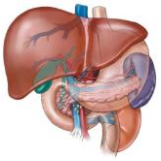


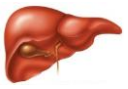
Ontology of toxic endpoints

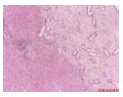
ToxWiz
toxicology ontology

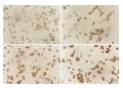
Endpoints are organized into categories: Example liver

Category	Example Toxic Endpoint Cluster	Includes	Description
System	Hepatotoxicity	Gall blader, liver, hepatic system	General observations (e.g. clinical, etc.)
Organ/tissue	Liver toxicities	Liver	Pathology report for any toxicity in liver
Organ/tissue	Liver hypertrophy	Liver	Pathology report of specific toxicity
Cells	Hepatocyte neoplasia	Cells or cell-lines	Result from cell-based assays







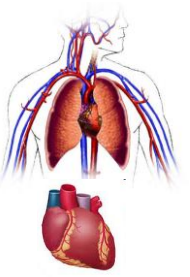


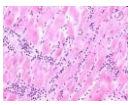
Ontology of toxic endpoints

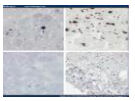
ToxWiz
toxicology ontology

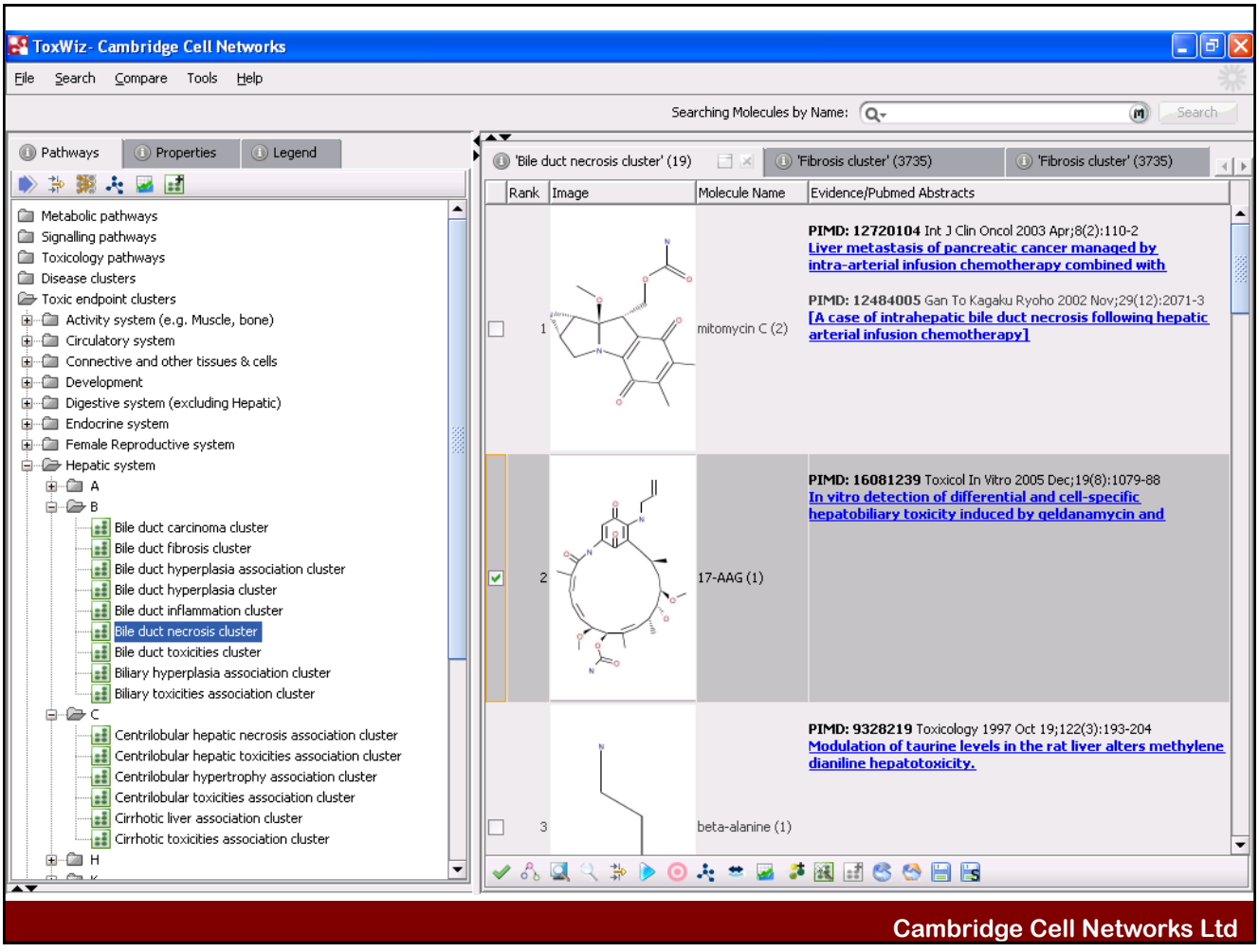
Endpoints are organized into categories: Example heart

Category	Example	Includes	Description
System	Cardiotoxicity	Cardiovascular system	General CV health
Organ/tissue	Heart toxicities	Heart	Pathology report for any toxicity in heart
Organ/tissue	Heart ischaemia	Heart	Pathology report of specific toxicity
Cells	Cardiac myocyte necrosis	Cells or cell-lines	Result from cells or cell-assays

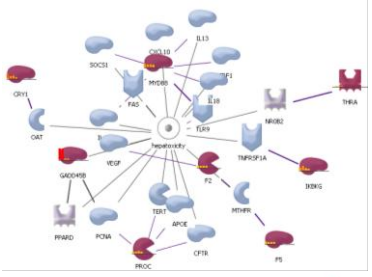




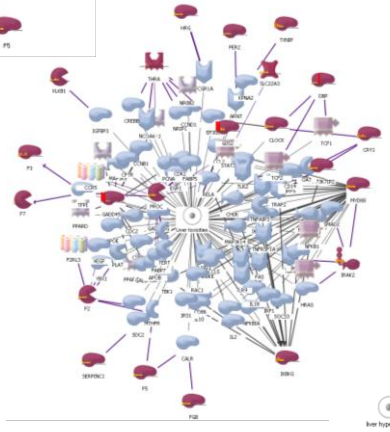



ToxWiz- Cambridge Cell Networks
File
Search
Compare
Tools
Help
Searching Molecules by Name:
Search
Pathways
Properties
Legend
Metabolic pathways
Signalling pathways
Toxicology pathways
Disease clusters
Toxic endpoint clusters
Activity system (e.g. Muscle, bone)
Circulatory system
Connective and other tissues & cells
Development
Digestive system (excluding Hepatic)
Endocrine system
Female Reproductive system
Hepatic system
A
B
Bile duct carcinoma cluster
Bile duct fibrosis cluster
Bile duct hyperplasia association cluster
Bile duct hyperplasia cluster
Bile duct inflammation cluster
Bile duct necrosis cluster
Bile duct toxicities cluster
Biliary hyperplasia association cluster
Biliary toxicities association cluster
C
Centrilobular hepatic necrosis association cluster
Centrilobular hepatic toxicities association cluster
Centrilobular hypertrophy association cluster
Centrilobular toxicities association cluster
Cirrhotic liver association cluster
Cirrhotic toxicities association cluster
H
'Bile duct necrosis cluster' (19)
'Fibrosis cluster' (3735)
'Fibrosis cluster' (3735)
Rank
Image
Molecule Name
Evidence/Pubmed Abstracts
1
mitomycin C (2)
PIMD: 12720104 Int J Clin Oncol 2003 Apr;8(2):110-2
Liver metastasis of pancreatic cancer managed by intra-arterial infusion chemotherapy combined with
PIMD: 12484005 Gan To Kagaku Ryoho 2002 Nov;29(12):2071-3
[A case of intrahepatic bile duct necrosis following hepatic arterial infusion chemotherapy]
2
17-AAG (1)
PIMD: 16081239 Toxicol In Vitro 2005 Dec;19(8):1079-88
In vitro detection of differential and cell-specific hepatobiliary toxicity induced by geldanamycin and
3
beta-alanine (1)
PIMD: 9328219 Toxicology 1997 Oct 19;122(3):193-204
Modulation of taurine levels in the rat liver alters methylene dianiline hepatotoxicity.
Cambridge Cell Networks Ltd

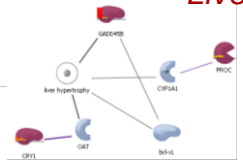
Linking gene expression data to histopathological observations at different levels of granularity of the observation – capturing underlying mechanism of developing pathology



Liver hyperplasia



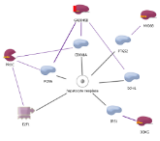
Liver neoplasia



Liver carcinoma

Time

Temporal development of toxicity from in-vivo testing



Hepatocyte neoplasia
In vitro

Cambridge Cell Networks Ltd

Interoperability with other ontology efforts:
Existing ontologies but not for toxicology, although could be implemented.
Important for all: if we know what for to use it, it make sense.

BioPAX.org

Biological Pathway Exchange

the Gene Ontology

The Open Biological and Biomedical Ontologies

Pathway ontology

OBI
Ontology for Biomedical Investigations

THE NATIONAL CENTER FOR BIOMEDICAL ONTOLOGY

BioPortal

ToxWiz
toxicology ontology

ChEBI

The Protein Ontology

SBO

Systems Biology Ontology

OpenTox

Creation, validation and optimization of CCNet's Ontology

CCNet Ontology of toxic endpoints: starting point

1. Hierarchy of organs & tissues for model organisms derived by consultation of various sources (e.g. NCBI, ENSEMBL, UniProt)

System – Organ – Components / Tissue / Cell-type
e.g. Hepatic system – Liver – Bile duct / Kupffer Cell

N.B. includes synonyms (Biliary = Bile duct, etc.)

2. Set of histopathological observational keywords & principles (by consultation with toxicologists, e.g. Frank Bonner)

e.g. Hypertrophy, Hyperplasia, Neoplasia, Carcinoma, Necrosis, Fibrosis, Inflammation, etc.

3. Several principles of progression in time/pathology

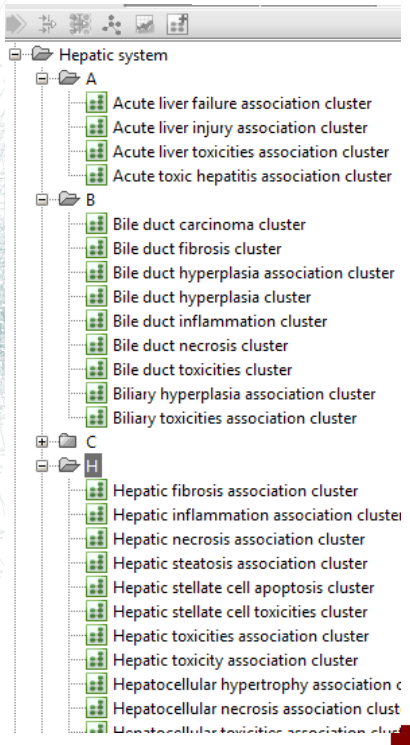
e.g. Carcinoma is a subset of Neoplasia, Hyperplasia precedes Neoplasia

Connections and classes in CCNet ToxWiz™ Ontologies

ToxWiz
toxicology ontology

An integrated set

- **92 high-level histopathological endpoint descriptors**
- **912 tissue/organ terms**
- **500000 systematic English phrases tested in the literature 408K publications) of which 2786 are in use (e.g. “bile duct hyperplasia” = “hyperplasia of the bile duct”)**
- **Basis set for annotation**
- **Thousands of histopathology ontologies in different levels of granularity**



Developing the ontology

ToxWiz
toxicology ontology

Using ontologies always leads to changes and improvements.
Over two years several changes were necessary according to customer feedback

1. Finer definitions of certain pathologies:
   e.g. Carcinoma
   – Adenocarcinoma
   – Squamous cell carcinoma
   – Metastatic carcinoma
   – Etc.

2. Some other broad terms to cover pathological observations, eg. primary & secondary as qualifiers

Generally: terms are added in order to capture better what experimentalists are saying & capturing.

… but this does not mean that ontology standard is a moving target. The changes and improvements only increase the granularity and are captured by higher categories

ToxWiz
toxicology ontology

3. Changes in order to synchronize the key objectives with other running projects and initiatives

INHAND
International Harmonization of Nomenclature and Diagnostic Criteria for Lesions in Rats and Mice

goRENI
Registry Nomenclature Information System

SEURAT-1

COLIPA
The European Cosmetics Association

FDA U.S. Food and Drug Administration
LTKB- Liver Toxicity Knowledge Base

imi
Innovative Medicines Initiative

eTRIKS:
ontology synchronization among all EU projects and Initiatives

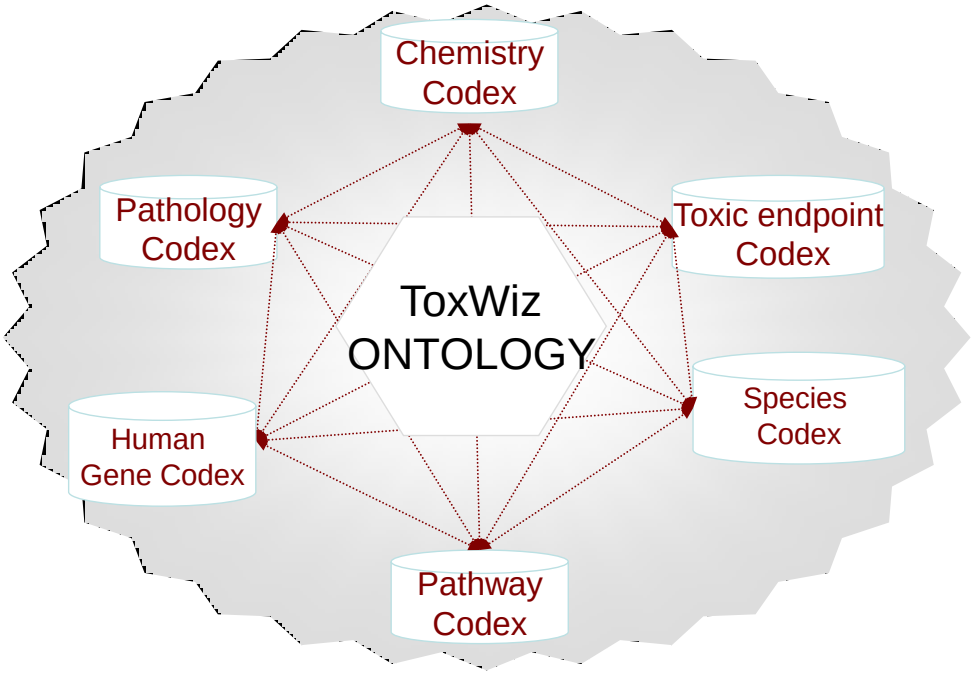
CCNet Expert Databases: 150 man years of expert curation: ToxWiz CODECES

* 17 million processed literature references, 4 million catalogued references supporting each database entry, 10 million relationships

Codex Database	Content
Chemistry Codex	50 000 chemical structures with all reported targets (drugs, industrials, metabolites, toxins)
Pathology Codex	1000 pathologies linked to 100.000 molecules
Toxic Endpoint Codex	1300 toxic-endpoints linked to 100.000molecules
Pathway Codex	500 pathways (signalling, metabolic)
Gene Codex	20 000 human targets/proteins/genes
Species Codex	cross species genes for 17 species

Interoperability achieved via harmonization of ontology

ToxWiz
toxicology ontology



Summary

ToxWiz
toxicology ontology

➢ to our knowledge the only one currently developed with capability to support a mechanistic interpretation

ToxWiz ontologies

- Derived from 150 man years of expert curation of safety information
- Describes histopathological terms with 912 terms
- Classifies further 500 000 terms used by pathologists
- Designed with relevant structure to describe MoA hypothesis and biological pathways
- Modeled on over 8 years of information extraction from safety reports

Benefits

- Extracting and mapping information about chemical structures related to
  - Toxicity
  - Disease
  - Hypothesis
- Allows exchange of data between user groups
- Deals with the different ways pathologists report
- Aids information classification for QSAR prediction
- Supports –omics interpretations
- Supports in-vitro findings

➢ CCNet base ontology framework as open source, with proprietary specialized add on modules' ?

A few words on Ontologies

Ontology is the philosophical discipline which aims to understand how things in the world are divided into categories and how these categories are related together. This is exactly what information scientists aim for in creating structured, automated representations, called 'ontologies,' for managing information in fields such as science, government, industry, and healthcare.

Applied Ontology: An Introduction, Katherine Munn, Barry Smith (Eds.)

Often an ontology of the domain is not a goal in itself. Developing an ontology is akin to defining a set of data and their structure for other programs to use.

Ontology Development 101: A Guide to Creating Your First Ontology, Noy & McGuinness, Stanford

Key point: Ontologies only make sense if they are used for something. Gene-ontology (GO) is perhaps the best example in the life sciences. They can seem mysterious until you use them.