OpenRiskNet

RISK ASSESSMENT E-INFRASTRUCTURE

AOPLink

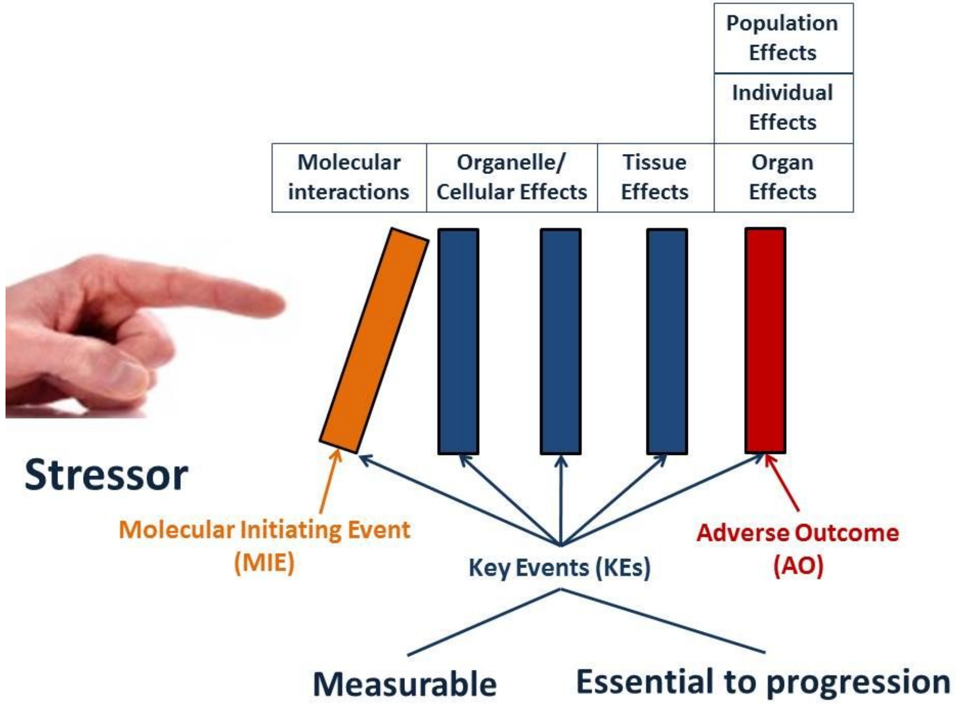
Identification and Linking of Data related to AOP-Wiki

Marvin Martens, Egon Willighagen, Chris Evelo

OpenRiskNet: Open e-Infrastructure to Support Data Sharing, Knowledge Integration and *in silico* Analysis and Modelling in Risk Assessment
Project Number 731075

Central concept: Adverse Outcome Pathways

Framework that captures mechanistic knowledge of toxicological processes to support decision making in risk assessments



Case study objectives

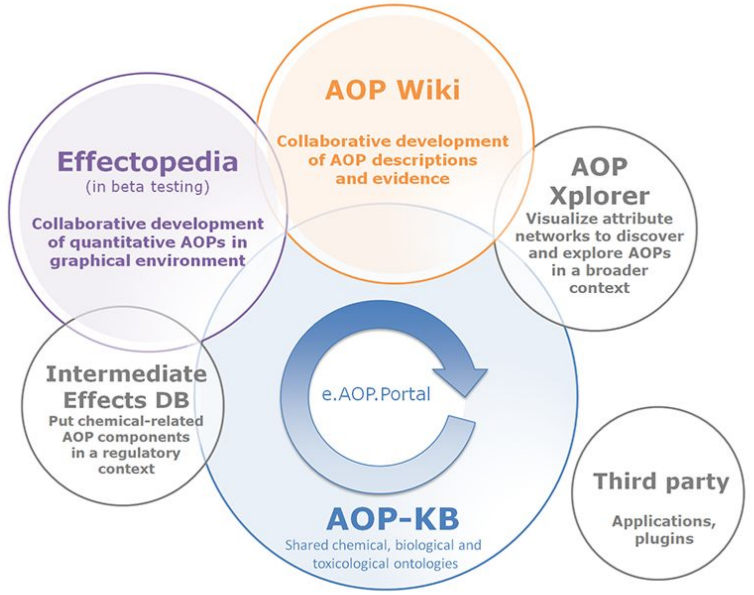
Q1: For an AOP, which experimental data is available to support the AOP?

Q2: Can this experimental data support an existing AOP?

$$\begin{tikzcd} \textbf{Experimental data} \arrow[r, leftrightarrow, "\textbf{AOPLink}"'] & \textbf{Adverse Outcome Pathways} \end{tikzcd}$$

Central repository: AOP-Wiki

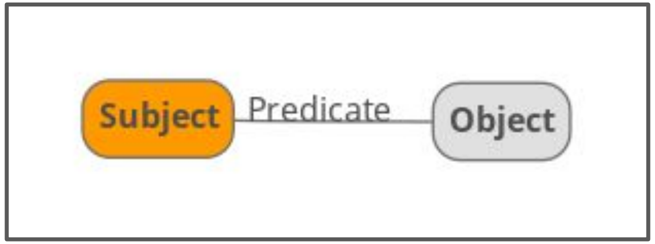
- The main qualitative AOP repository of the AOP-KB
- Joint effort between EC-JRC and US EPA


AOP Wiki
Collaborative development of AOP descriptions and evidence
Effectopedia
(in beta testing)
Collaborative development of quantitative AOPs in graphical environment
AOP Xplorer
Visualize attribute networks to discover and explore AOPs in a broader context
Intermediate Effects DB
Put chemical-related AOP components in a regulatory context
e.AOP.Portal
AOP-KB
Shared chemical, biological and toxicological ontologies
Third party
Applications, plugins

https://aopwiki.org

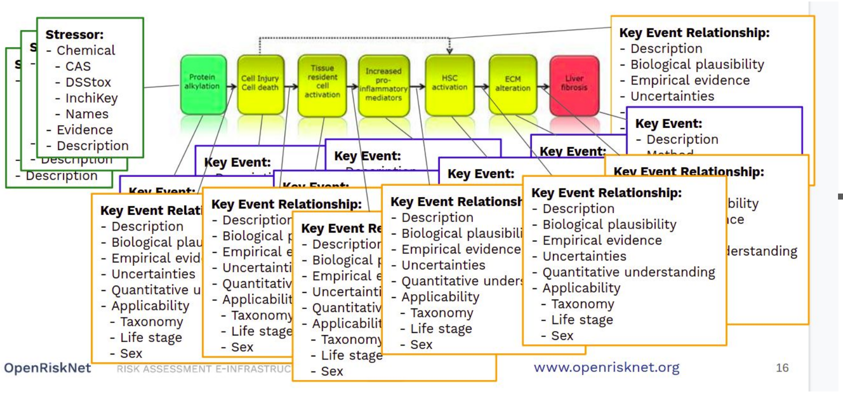
Resource Description Framework (RDF)

- Standard to describe information in web resources
- Information stored in triples



- Generally used in databases
- Large number of libraries and tools

Semantic modelling of AOP-Wiki

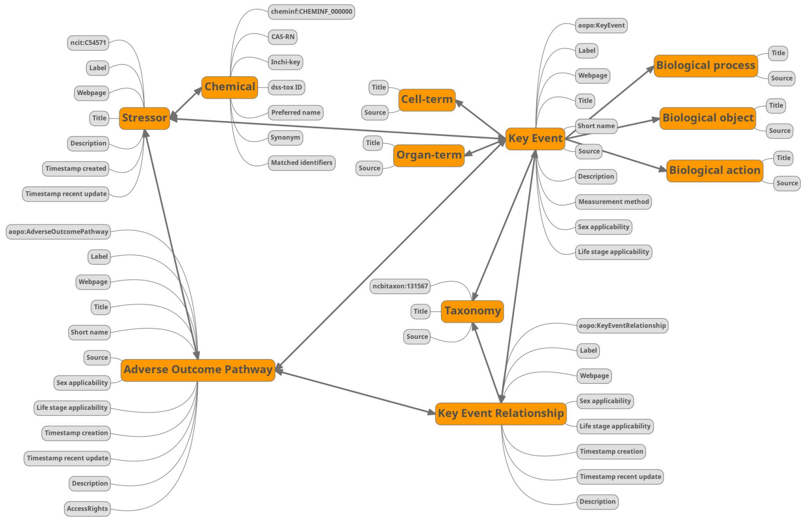


Stressor:
- Chemical
  - CAS
  - DSStox
  - InchiKey
  - Names
- Evidence
- Description

Key Event Relationship:
- Description
- Biological plausibility
- Empirical evidence
- Uncertainties

Key Event:
- Description

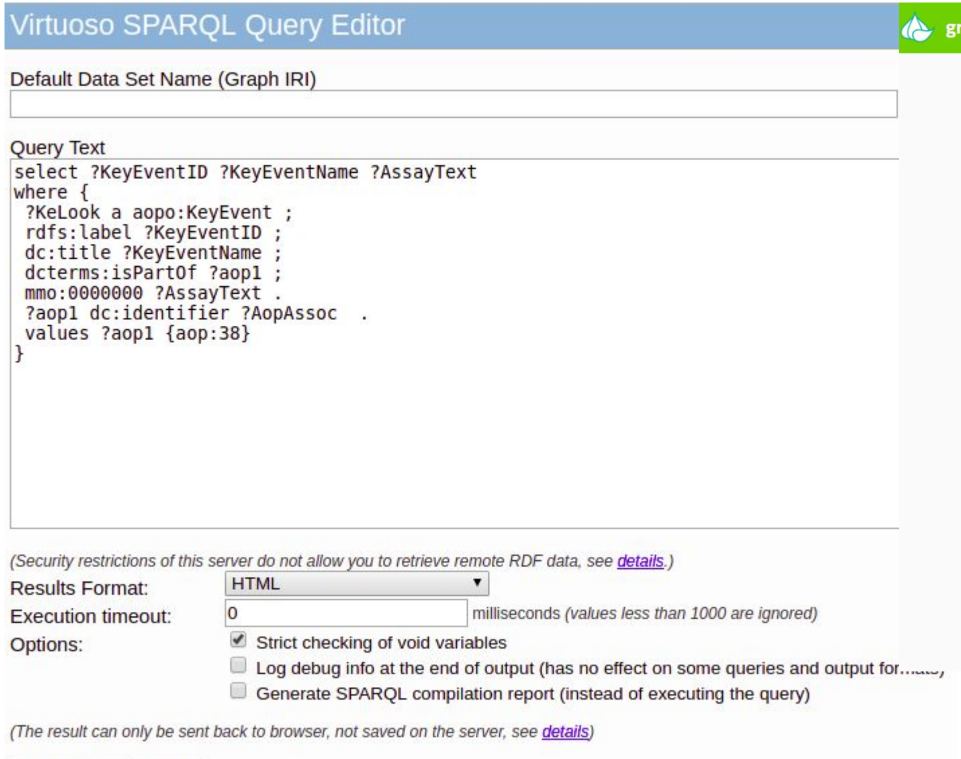
Key Event Relationship:
- Description
- Biological plausibility
- Empirical evidence
- Uncertainties
- Quantitative understanding
- Applicability
  - Taxonomy
  - Life stage
  - Sex



How to access the AOP-Wiki RDF

Through the SPARQL endpoint

http://aopwiki-rdf.prod.openrisknet.org/sparql/



Virtuoso SPARQL Query Editor

Default Data Set Name (Graph IRI)

Query Text

```
select ?KeyEventID ?KeyEventName ?AssayText
where {
 ?KeLook a aopo:KeyEvent ;
 rdfs:label ?KeyEventID ;
 dc:title ?KeyEventName ;
 dcterms:isPartOf ?aop1 ;
 mmo:0000000 ?AssayText .
 ?aop1 dc:identifier ?AopAssoc  .
 values ?aop1 {aop:38}
}
```

(Security restrictions of this server do not allow you to retrieve remote RDF data, see details.)

Results Format: HTML

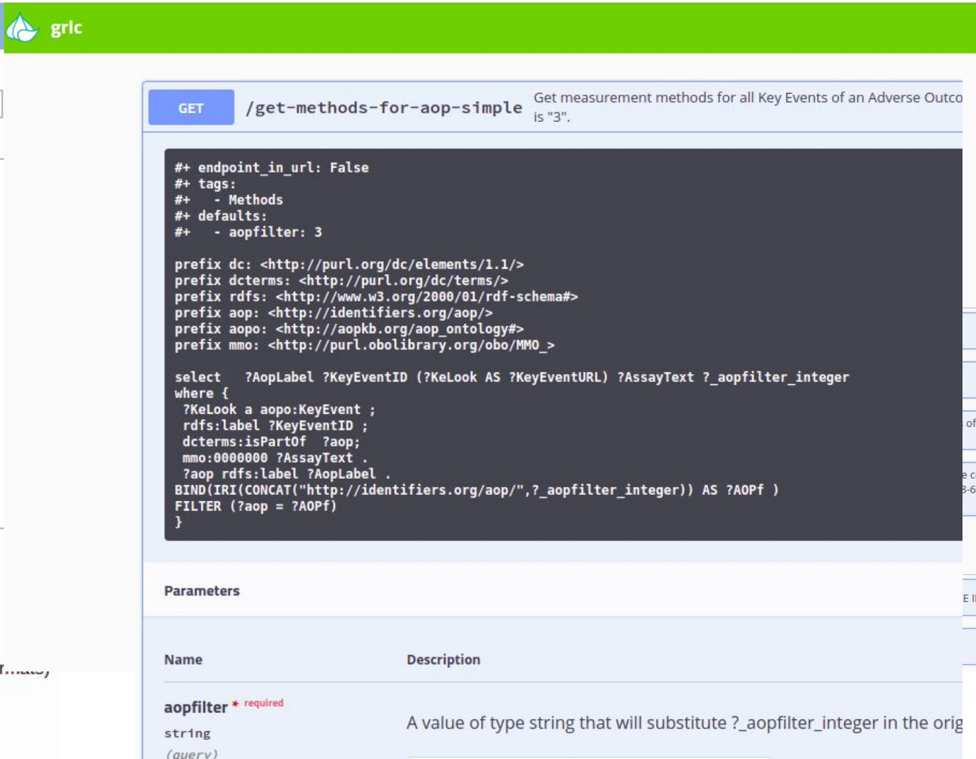
Execution timeout: 0 milliseconds (values less than 1000 are ignored)

Options:
- [✓] Strict checking of void variables
- [ ] Log debug info at the end of output (has no effect on some queries and output for...)
- [ ] Generate SPARQL compilation report (instead of executing the query)

(The result can only be sent back to browser, not saved on the server, see details)

Through the REST API

http://grlc.io/api/marvinm2/AOPWikiQueries



grlc

GET /get-methods-for-aop-simple Get measurement methods for all Key Events of an Adverse Outco
is "3".

```
#+ endpoint_in_url: False
#+ tags:
#+   - Methods
#+ defaults:
#+   - aopfilter: 3

prefix dc: <http://purl.org/dc/elements/1.1/>
prefix dcterms: <http://purl.org/dc/terms/>
prefix rdfs: <http://www.w3.org/2000/01/rdf-schema#>
prefix aop: <http://identifiers.org/aop/>
prefix aopo: <http://aopkb.org/aop_ontology#>
prefix mmo: <http://purl.obolibrary.org/obo/MMO_>

select   ?AopLabel ?KeyEventID (?KeLook AS ?KeyEventURL) ?AssayText ?_aopfilter_integer
where {
 ?KeLook a aopo:KeyEvent ;
 rdfs:label ?KeyEventID ;
 dcterms:isPartOf  ?aop;
 mmo:0000000 ?AssayText .
 ?aop rdfs:label ?AopLabel .
BIND(IRI(CONCAT("http://identifiers.org/aop/",?_aopfilter_integer)) AS ?AOPf )
FILTER (?aop = ?AOPf)
}
```

Parameters

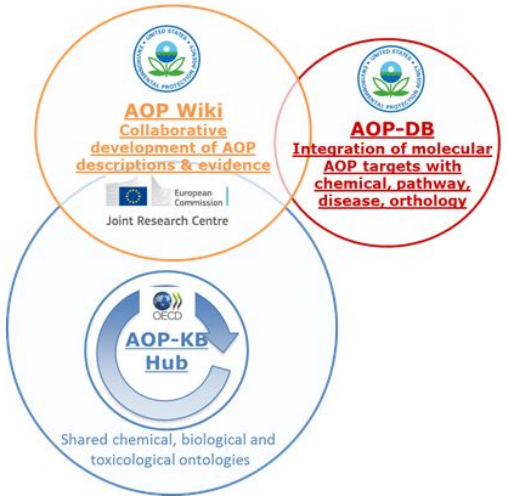
Name	Description
aopfilter * required string (query)	A value of type string that will substitute ?_aopfilter_integer in the orig

Integration of the AOP-DB

Winners of the OpenRiskNet implementation challenge

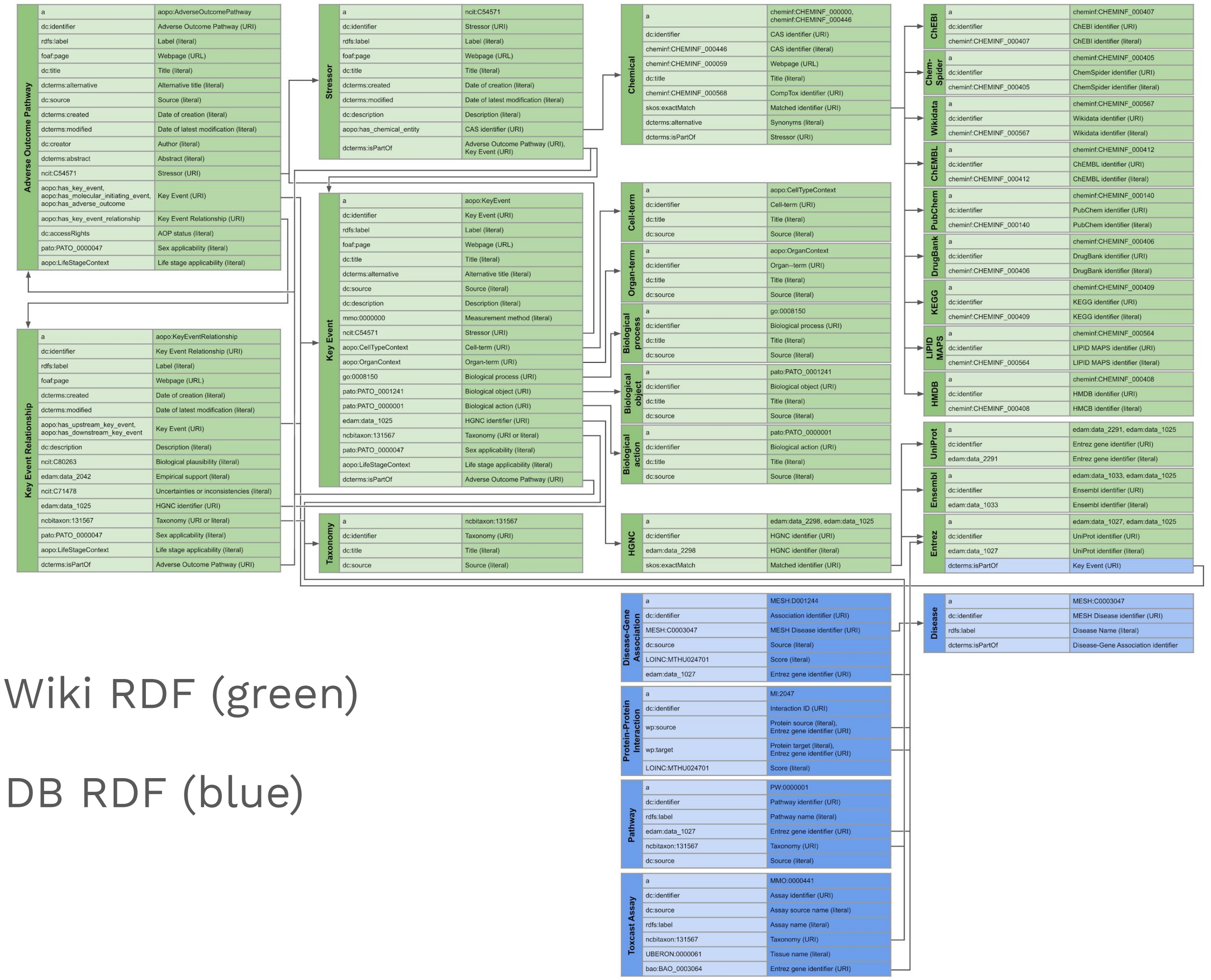
Database of AOPs linked to external data sources for (among others:

- Genes
- Pathways
- ToxCast assays
- Diseases


AOP Wiki
Collaborative development of AOP descriptions & evidence
European Commission
Joint Research Centre
AOP-DB
Integration of molecular AOP targets with chemical, pathway, disease, orthology
OECD
AOP-KB Hub
Shared chemical, biological and toxicological ontologies

Parts of AOP-DB converted into RDF and exposed in a SPARQL endpoint

AOP-DB webinar: https://openrisknet.org/events/60/



Adverse Outcome Pathway

a	aopo:AdverseOutcomePathway
dc:identifier	Adverse Outcome Pathway (URI)
rdfs:label	Label (literal)
foaf:page	Webpage (URL)
dc:title	Title (literal)
dcterms:alternative	Alternative title (literal)
dc:source	Source (literal)
dcterms:created	Date of creation (literal)
dcterms:modified	Date of latest modification (literal)
dc:creator	Author (literal)
dcterms:abstract	Abstract (literal)
ncit:C54571	Stressor (URI)
aopo:has_key_event, aopo:has_molecular_initiating_event, aopo:has_adverse_outcome	Key Event (URI)
aopo:has_key_event_relationship	Key Event Relationship (URI)
dc:accessRights	AOP status (literal)
pato:PATO_0000047	Sex applicability (literal)
aopo:LifeStageContext	Life stage applicability (literal)

Key Event Relationship

a	aopo:KeyEventRelationship
dc:identifier	Key Event Relationship (URI)
rdfs:label	Label (literal)
foaf:page	Webpage (URL)
dcterms:created	Date of creation (literal)
dcterms:modified	Date of latest modification (literal)
aopo:has_upstream_key_event, aopo:has_downstream_key_event	Key Event (URI)
dc:description	Description (literal)
ncit:C80263	Biological plausibility (literal)
edam:data_2042	Empirical support (literal)
ncit:C71478	Uncertainties or inconsistencies (literal)
edam:data_1025	HGNC identifier (URI)
ncbitaxon:131567	Taxonomy (URI or literal)
pato:PATO_0000047	Sex applicability (literal)
aopo:LifeStageContext	Life stage applicability (literal)
dcterms:isPartOf	Adverse Outcome Pathway (URI)

Stressor

a	ncit:C54571
dc:identifier	Stressor (URI)
rdfs:label	Label (literal)
foaf:page	Webpage (URL)
dc:title	Title (literal)
dcterms:created	Date of creation (literal)
dcterms:modified	Date of latest modification (literal)
dc:description	Description (literal)
aopo:has_chemical_entity	CAS identifier (URI)
dcterms:isPartOf	Adverse Outcome Pathway (URI), Key Event (URI)

Key Event

a	aopo:KeyEvent
dc:identifier	Key Event (URI)
rdfs:label	Label (literal)
foaf:page	Webpage (URL)
dc:title	Title (literal)
dcterms:alternative	Alternative title (literal)
dc:source	Source (literal)
dc:description	Description (literal)
mmo:0000000	Measurement method (literal)
ncit:C54571	Stressor (URI)
aopo:CellTypeContext	Cell-term (URI)
aopo:OrganContext	Organ-term (URI)
go:0008150	Biological process (URI)
pato:PATO_0001241	Biological object (URI)
pato:PATO_0000001	Biological action (URI)
edam:data_1025	HGNC identifier (URI)
ncbitaxon:131567	Taxonomy (URI or literal)
pato:PATO_0000047	Sex applicability (literal)
aopo:LifeStageContext	Life stage applicability (literal)
dcterms:isPartOf	Adverse Outcome Pathway (URI)

Taxonomy

a	ncbitaxon:131567
dc:identifier	Taxonomy (URI)
dc:title	Title (literal)
dc:source	Source (literal)

Chemical

a	cheminf:CHEMINF_000000, cheminf:CHEMINF_000446
dc:identifier	CAS identifier (URI)
cheminf:CHEMINF_000446	CAS identifier (literal)
cheminf:CHEMINF_000059	Webpage (URL)
dc:title	Title (literal)
cheminf:CHEMINF_000568	CompTox identifier (URI)
skos:exactMatch	Matched identifier (URI)
dcterms:alternative	Synonyms (literal)
dcterms:isPartOf	Stressor (URI)

Cell-term

a	aopo:CellTypeContext
dc:identifier	Cell-term (URI)
dc:title	Title (literal)
dc:source	Source (literal)

Organ-term

a	aopo:OrganContext
dc:identifier	Organ--term (URI)
dc:title	Title (literal)
dc:source	Source (literal)

Biological process

a	go:0008150
dc:identifier	Biological process (URI)
dc:title	Title (literal)
dc:source	Source (literal)

Biological object

a	pato:PATO_0001241
dc:identifier	Biological object (URI)
dc:title	Title (literal)
dc:source	Source (literal)

Biological action

a	pato:PATO_0000001
dc:identifier	Biological action (URI)
dc:title	Title (literal)
dc:source	Source (literal)

HGNC

a	edam:data_2298, edam:data_1025
dc:identifier	HGNC identifier (URI)
edam:data_2298	HGNC identifier (literal)
skos:exactMatch	Matched identifier (URI)

ChEBI

a	cheminf:CHEMINF_000407
dc:identifier	ChEBI identifier (URI)
cheminf:CHEMINF_000407	ChEBI identifier (literal)

Chem-Spider

a	cheminf:CHEMINF_000405
dc:identifier	ChemSpider identifier (URI)
cheminf:CHEMINF_000405	ChemSpider identifier (literal)

Wikidata

a	cheminf:CHEMINF_000567
dc:identifier	Wikidata identifier (URI)
cheminf:CHEMINF_000567	Wikidata identifier (literal)

ChEMBL

a	cheminf:CHEMINF_000412
dc:identifier	ChEMBL identifier (URI)
cheminf:CHEMINF_000412	ChEMBL identifier (literal)

PubChem

a	cheminf:CHEMINF_000140
dc:identifier	PubChem identifier (URI)
cheminf:CHEMINF_000140	PubChem identifier (literal)

DrugBank

a	cheminf:CHEMINF_000406
dc:identifier	DrugBank identifier (URI)
cheminf:CHEMINF_000406	DrugBank identifier (literal)

KEGG

a	cheminf:CHEMINF_000409
dc:identifier	KEGG identifier (URI)
cheminf:CHEMINF_000409	KEGG identifier (literal)

LIPID MAPS

a	cheminf:CHEMINF_000564
dc:identifier	LIPID MAPS identifier (URI)
cheminf:CHEMINF_000564	LIPID MAPS identifier (literal)

HMDB

a	cheminf:CHEMINF_000408
dc:identifier	HMDB identifier (URI)
cheminf:CHEMINF_000408	HMCB identifier (literal)

UniProt

a	edam:data_2291, edam:data_1025
dc:identifier	Entrez gene identifier (URI)
edam:data_2291	Entrez gene identifier (literal)

Ensembl

a	edam:data_1033, edam:data_1025
dc:identifier	Ensembl identifier (URI)
edam:data_1033	Ensembl identifier (literal)

Entrez

a	edam:data_1027, edam:data_1025
dc:identifier	UniProt identifier (URI)
edam:data_1027	UniProt identifier (literal)
dcterms:isPartOf	Key Event (URI)

Disease-Gene Association

a	MESH:D001244
dc:identifier	Association identifier (URI)
MESH:C0003047	MESH Disease identifier (URI)
dc:source	Source (literal)
LOINC:MTHU024701	Score (literal)
edam:data_1027	Entrez gene identifier (URI)

Disease

a	MESH:C0003047
dc:identifier	MESH Disease identifier (URI)
rdfs:label	Disease Name (literal)
dcterms:isPartOf	Disease-Gene Association identifier

Protein-Protein Interaction

a	MI:2047
dc:identifier	Interaction ID (URI)
wp:source	Protein source (literal), Entrez gene identifier (URI)
wp:target	Protein target (literal), Entrez gene identifier (URI)
LOINC:MTHU024701	Score (literal)

Pathway

a	PW:0000001
dc:identifier	Pathway identifier (URI)
rdfs:label	Pathway name (literal)
edam:data_1027	Entrez gene identifier (URI)
ncbitaxon:131567	Taxonomy (URI)
dc:source	Source (literal)

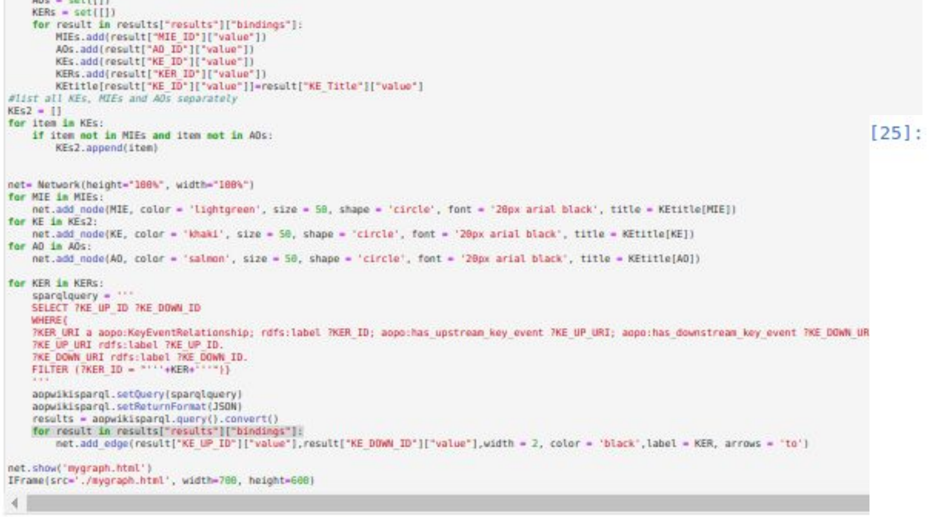
Toxcast Assay

a	MMO:0000441
dc:identifier	Assay identifier (URI)
dc:source	Assay source name (literal)
rdfs:label	Assay name (literal)
ncbitaxon:131567	Taxonomy (URI)
UBERON:0000061	Tissue name (literal)
bao:BAO_0003064	Entrez gene identifier (URI)

AOP-Wiki RDF (green)

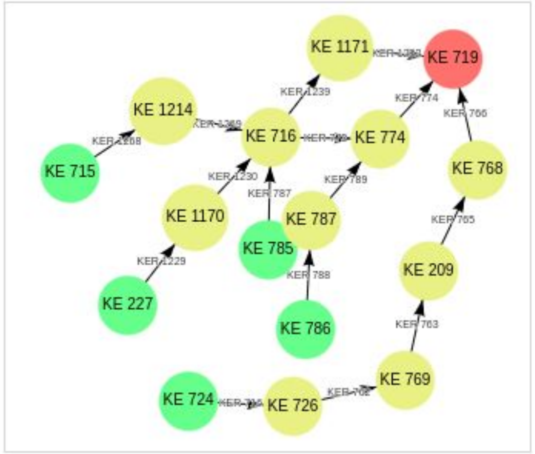
AOP-DB RDF (blue)

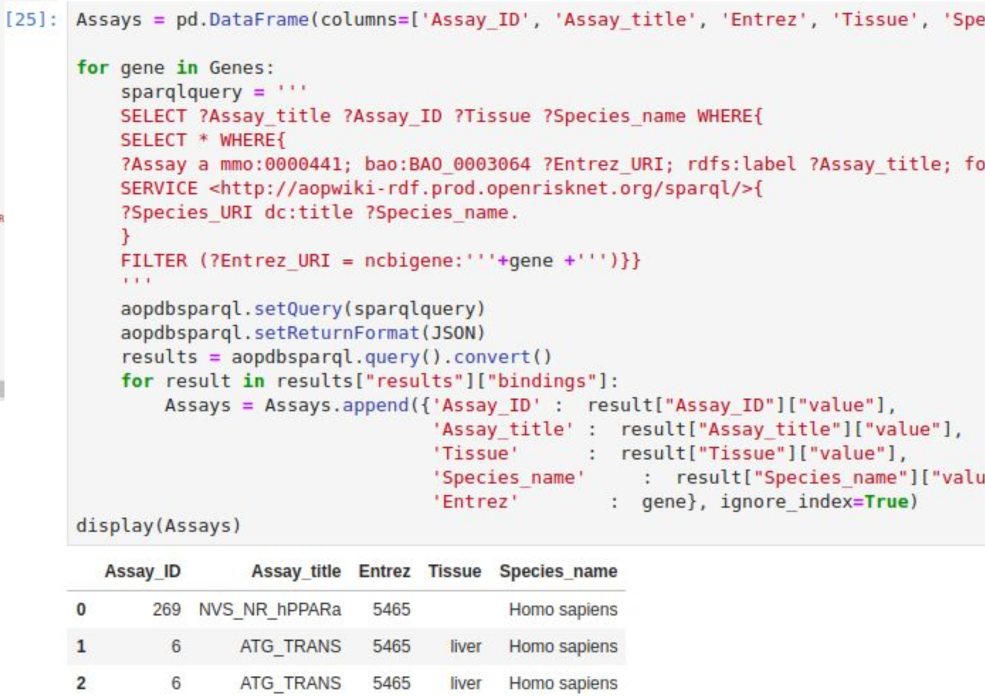
OpenRiskNet



```
    AOs = set([])
    KERs = set([])
    for result in results["results"]["bindings"]:
        MIEs.add(result["MIE_ID"]["value"])
        AOs.add(result["AO_ID"]["value"])
        KEs.add(result["KE_ID"]["value"])
        KERs.add(result["KER_ID"]["value"])
        KEtitle[result["KE_ID"]["value"]]=result["KE_Title"]["value"]
#list all KEs, MIEs and AOs separately
KEs2 = []
for item in KEs:
    if item not in MIEs and item not in AOs:
        KEs2.append(item)


net= Network(height="100%", width="100%")
for MIE in MIEs:
    net.add_node(MIE, color = 'lightgreen', size = 50, shape = 'circle', font = '20px arial black', title = KEtitle[MIE])
for KE in KEs2:
    net.add_node(KE, color = 'khaki', size = 50, shape = 'circle', font = '20px arial black', title = KEtitle[KE])
for AO in AOs:
    net.add_node(AO, color = 'salmon', size = 50, shape = 'circle', font = '20px arial black', title = KEtitle[AO])

for KER in KERs:
    sparqlquery = '''
    SELECT ?KE_UP_ID ?KE_DOWN_ID
    WHERE{
    ?KER_URI a aopo:KeyEventRelationship; rdfs:label ?KER_ID; aopo:has_upstream_key_event ?KE_UP_URI; aopo:has_downstream_key_event ?KE_DOWN_UR
    ?KE_UP_URI rdfs:label ?KE_UP_ID.
    ?KE_DOWN_URI rdfs:label ?KE_DOWN_ID.
    FILTER (?KER_ID = "'''+KER+'''")}
    '''
    aopwikisparql.setQuery(sparqlquery)
    aopwikisparql.setReturnFormat(JSON)
    results = aopwikisparql.query().convert()
    for result in results["results"]["bindings"]:
        net.add_edge(result["KE_UP_ID"]["value"],result["KE_DOWN_ID"]["value"],width = 2, color = 'black',label = KER, arrows = 'to')

net.show('mygraph.html')
IFrame(src='./mygraph.html', width=700, height=600)
```

[20]:

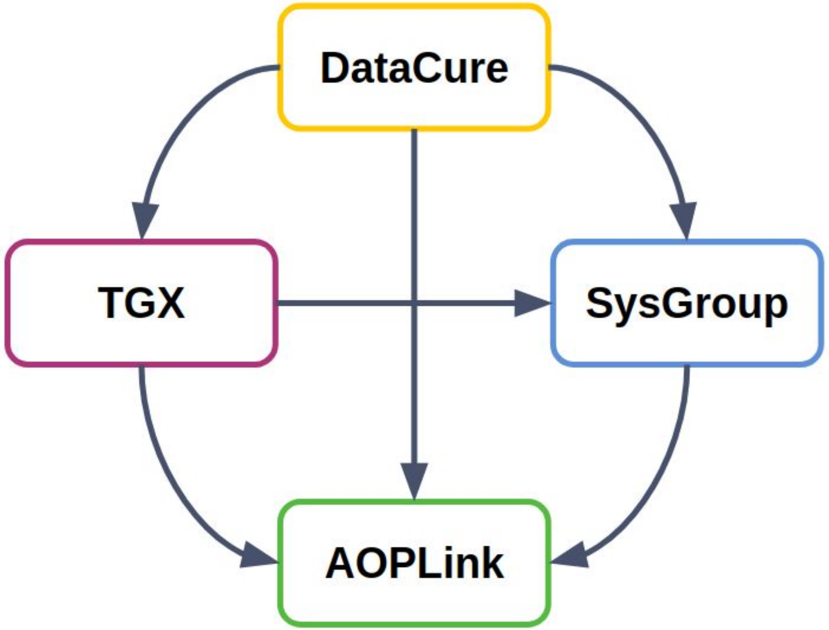


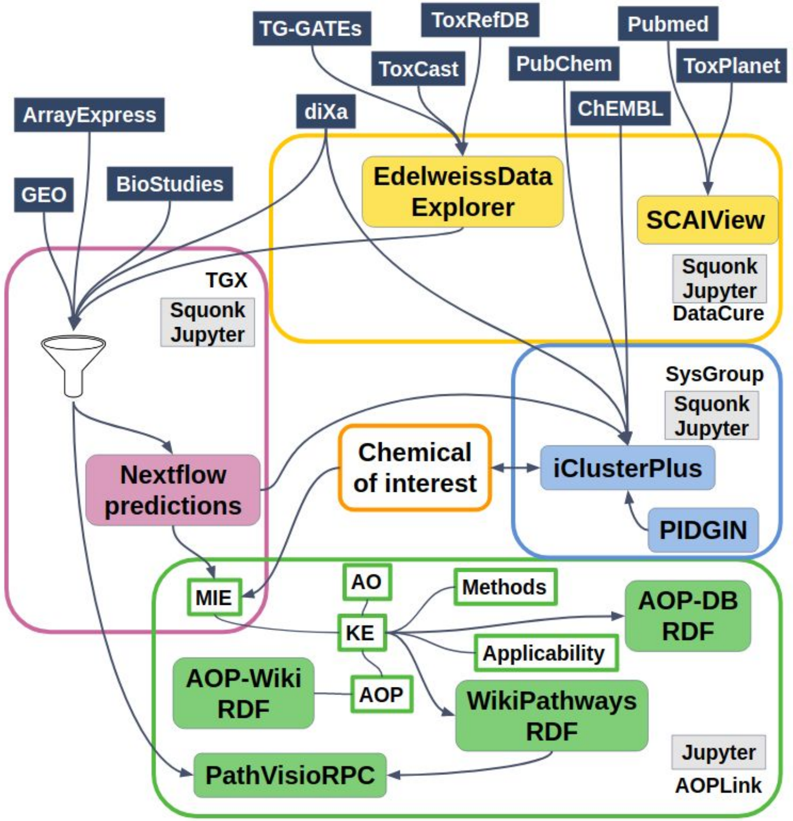


```
[25]: Assays = pd.DataFrame(columns=['Assay_ID', 'Assay_title', 'Entrez', 'Tissue', 'Spe

      for gene in Genes:
          sparqlquery = '''
          SELECT ?Assay_title ?Assay_ID ?Tissue ?Species_name WHERE{
          SELECT * WHERE{
          ?Assay a mmo:0000441; bao:BAO_0003064 ?Entrez_URI; rdfs:label ?Assay_title; fo
          SERVICE <http://aopwiki-rdf.prod.openrisknet.org/sparql/>{
          ?Species_URI dc:title ?Species_name.
          }
          FILTER (?Entrez_URI = ncbigene:'''+gene +''')}}
          '''
          aopdbsparql.setQuery(sparqlquery)
          aopdbsparql.setReturnFormat(JSON)
          results = aopdbsparql.query().convert()
          for result in results["results"]["bindings"]:
              Assays = Assays.append({'Assay_ID' :  result["Assay_ID"]["value"],
                                      'Assay_title' :  result["Assay_title"]["value"],
                                      'Tissue'      :  result["Tissue"]["value"],
                                      'Species_name'     :  result["Species_name"]["valu
                                      'Entrez'        :  gene}, ignore_index=True)
      display(Assays)
```

	Assay_ID	Assay_title	Entrez	Tissue	Species_name
0	269	NVS_NR_hPPARa	5465		Homo sapiens
1	6	ATG_TRANS	5465	liver	Homo sapiens
2	6	ATG_TRANS	5465	liver	Homo sapiens

AOPLink links with other case studies





Identification and Linking of Data related to AOP-Wiki [AOPLink]

CS leader: Marvin Martens, Egon WIllighagen, Chris Evelo (UM)
Involved: EwC, UoB, CRG

Outcome:
- Discoverable annotated BridgeDb API
- Development of the AOPLink RDF (AOP-Wiki + WikiPathways + AOP-DB), and loaded and exposed as Virtuoso SPARQL endpoints
- Implementation challenge service: AOP-DB RDF
- Workflows utilizing the AOPLink RDF, linking knowledge repositories and experimental data to AOPs.

To do:
- Improved linking of AOPs with WikiPathways via KE genes.
- Development of pathway analysis Jupyter notebook with gene expression data related to AOP stressor chemicals.
- Knowledge base linking nanomaterials to MIEs

Risk Assessment Framework
Tier 0.3, 0.4 (Support Data), 1.6 (MOA)

Databases
- AOP-Wiki, AOP-DB: AOP knowledgebase (AOP-KB);
- WikiPathways, Reactome: biological pathway database;
- eNanoMapper, EPA Chemistry Dashboard, NORMAN Network: experimental data.

Tools / APIs
- BridgeDb, ChemIdConverter: identifier mapping;
- PathVisioRPC: pathway analysis;
- eNanoMapper database test instance

Results
- BridgeDb service, AOP-Wiki and AOP-DB SPARQL endpoints, operational in VRE
- Report on AOPWiki<>WikiPathways linking options

Activities
- Continued development identifier mapping databases
- Semantification of AOPWiki
- AOP Portal (http://aop.wikipathways.org)
- Exploration of APIs around semantic web technologies

https://openrisknet.org/development/case-studies/case-study-aoplink/