Maastricht University

Faculty of Health, Medicine and Life Sciences

Expanding Adverse Outcome Pathway knowledge by creating AOP-Wiki RDF with semantic annotations to facilitate risk assessment of chemicals

Marvin Martens[A], Ryan Miller[A], Chris T Evelo[A,B], Egon Willighagen[A]

A Department of Bioinformatics - BiGCaT, NUTRIM, Maastricht University, The Netherlands
B Maastricht Centre for Systems Biology (MaCSBio), Maastricht University, The Netherlands

Introduction

Toxicology needs faster, more efficient use of existing data and knowledge for risk assessment for the ever-growing number of chemicals to be testing [1]. The Adverse Outcome Pathway (AOP) concept emerged in 2010 to capture all mechanistic information on toxicological processes, and thereby assist in setting up a risk assessment approach [2]. Currently, the major AOP database is the AOP-Wiki, but its data cannot be easily queried for specific content, and lacks semantic annotations that allow the linking to other relevant data sources and tools.

We aim to improve the interoperability of the AOP-Wiki by creating a Resource Description Framework (RDF) version and allowing data retrieval through SPARQL queries to support toxicological risk assessment workflows.

Methods

The AOP-Wiki provides quarterly permanent downloads for the full database XML (https://aopwiki.org/downloads/). This XML was parsed with Python 3.5 and the ElementTree XML API. The data was stored in the Turtle format.

To integrate the RDF, we wrote a variety of federated SPARQL queries executed in Blazegraph (build version 2.1.4).

Results

The created AOP-Wiki RDF has the RDF scheme is represented in Figure 1.

SPARQL queries were created for the following questions:

- What is all measurement/method information for an AOP? (Figure 2)
- For all stressor chemicals what can be linked to WikiPathways [3,4] using Wikidata [5] for identifier mapping (Figure 3).

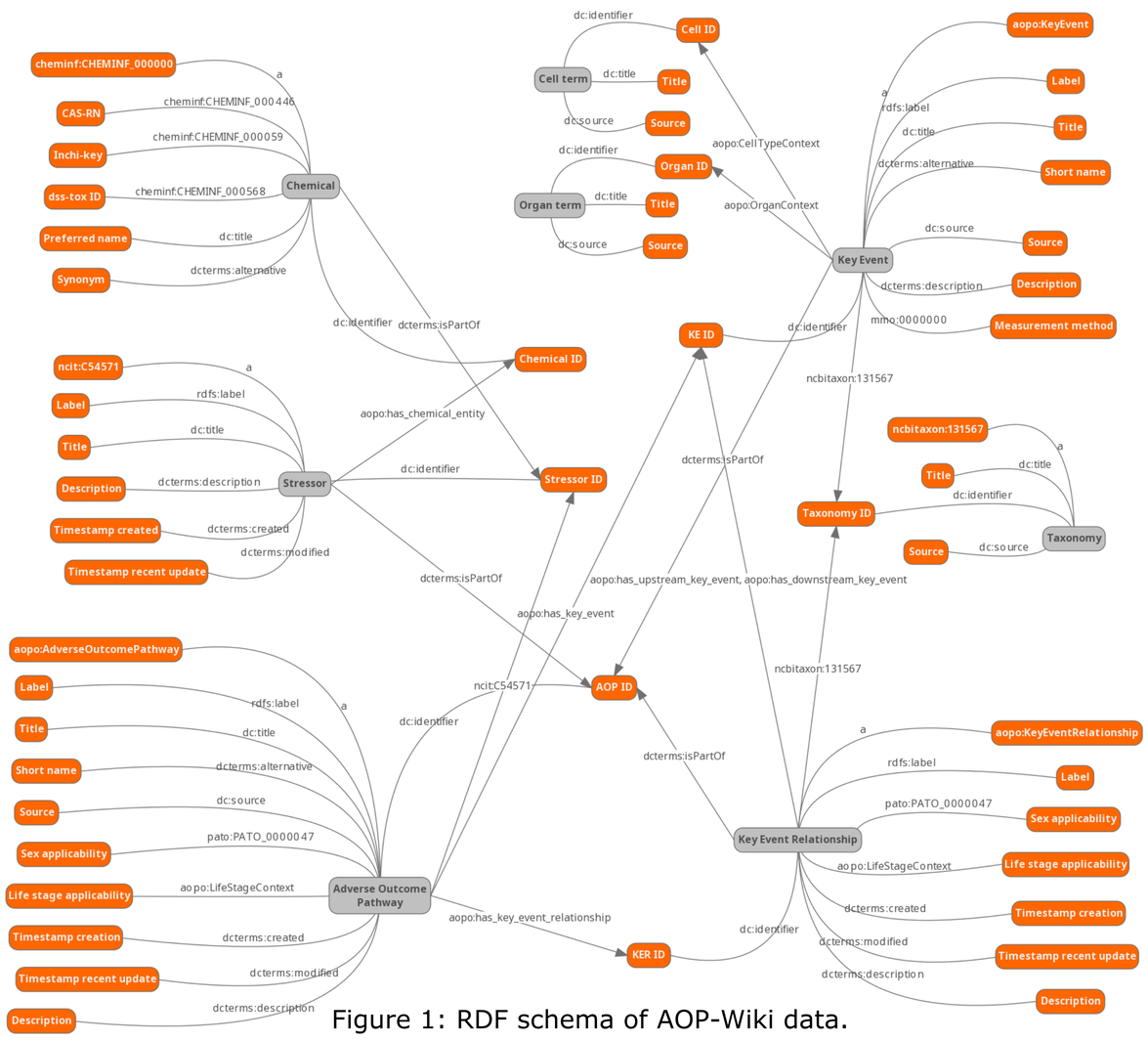

Figure 1: RDF schema of AOP-Wiki data.

```
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?KeLookUp ?AssayText
where { ?KeLook a aopo:KeyEvent ; rdfs:label ?KeLookUp ; dcterms:isPartOf ?aop1 ;  mmo:0000000 ?AssayText .
 ?aop1 dc:identifier ?AopAssoc ; rdfs:label ?AopLabel .
 filter (?aop1 = aop:12)
}
```

AopLabel	KeLookUp	
AOP 12	KE 188	Neuroinflammation, i.e. the activation of glial cells can be measured by quantification of cellular markers (most commonly), or of released Microglial activation can be detected based on the increased numbers of labeled microglia per volume element of brain tissue (due to The most frequently used astrocyte marker is GFAP (99% of all studies) (Eng et al., 2000). This protein is highly specific for astro All immunocytochemical methods can also be applied to cell culture models. In patients, microglial accumulation can be monitored by PET imaging, using [11C]-PK 11195 as a microglial marker (Banati et al., 20 Activation of glial cells can be assessed in tissue or cell culture models also by quantification of sets of activation markers. Thi Pro- and anti-inflammatory cytokine expression (IL-1β; TNF-α, Il-6, IL-4); or expression of immunostimmulatory proteins (e.g. MHC-II Itgam, CD86 expression as markers of M1 microglial phenotype Arg1, MRC1, as markers of M2 microglial phenotype (for descriptions of techniques, see also Falsig 2004; Lund 2006 ; Kuegler 2010; Monnet-Tschudi et al., 2011; Sandström et al., 2014; von T Regulatory example using the KE:Measurement of glial fibrillary acidic protein (GFAP) in brain tissue, whose increase is a marker of astrocy
AOP 12	KE 195	Methods that have been previously reviewed and approved by a recognized authority should be included in the Overview section above. All other methods, including those well established in the published literature, should be described here.

Figure 2: SPARQL query and result to retrieve all measurement/method information for AOP 12 from the AOP-Wiki data.

```
prefix dc: <http://purl.org/dc/elements/1.1/>
prefix dcterms: <http://purl.org/dc/terms/>
prefix cheminf: <http://semanticscience.org/resource/>
prefix wdt: <http://www.wikidata.org/prop/direct/>
prefix wp: <http://vocabularies.wikipathways.org/wp#>
prefix aopo: <http://aopkb.org/aop_ontology#>

select distinct ?ChemicalName ?LinkedAOP ?LinkedAOPURI ?CASRN ?ChemicalURI ?PathwayID ?PathwayURI where{
select *where { ?cheLook a cheminf:CHEMINF_000000 ; dc:identifier ?ChemicalURI ; dc:title ?ChemicalName ; cheminf:CHEMINF_000446 ?CASRN ; dcterms:isPartOf ?LinkedStressor.
 ?LinkedStressor dcterms:isPartOf ?LinkedAOPURI .
 ?LinkedAOPURI a aopo:AdverseOutcomePathway ; rdfs:label ?LinkedAOP.
 service <https://query.wikidata.org/bigdata/namespace/wdq/sparql>{ ?wikidata wdt:P231 ?CASRN.
   service <http://sparql.wikipathways.org/>{ ?metabolite wp:bdbWikidata ?wikidata ; dcterms:isPartOf ?PathwayURI.
     ?PathwayURI a wp:Pathway ; dcterms:identifier ?PathwayID.
}}}} order by ?ChemicalName
```

ChemicalName	LinkedAOP	LinkedAOPURI	CASRN	ChemicalURI	PathwayID	PathwayURI
Acetaminophen	AOP 260	<http://identifiers.org/aop/260>	103-90-2	<http://identifiers.org/cas/103-90-2>	WP4111	<http://identifiers.org/wikipathways/WP4111_r93592>
Acetaminophen	AOP 260	<http://identifiers.org/aop/260>	103-90-2	<http://identifiers.org/cas/103-90-2>	WP4085	<http://identifiers.org/wikipathways/WP4085_r93465>
Acetaminophen	AOP 260	<http://identifiers.org/aop/260>	103-90-2	<http://identifiers.org/cas/103-90-2>	WP4265	<http://identifiers.org/wikipathways/WP4265_r98060>
Acetaminophen	AOP 260	<http://identifiers.org/aop/260>	103-90-2	<http://identifiers.org/cas/103-90-2>	WP2886	<http://identifiers.org/wikipathways/WP2886_r88729>
Acetaminophen	AOP 260	<http://identifiers.org/aop/260>	103-90-2	<http://identifiers.org/cas/103-90-2>	WP4268	<http://identifiers.org/wikipathways/WP4268_r98052>

Figure 3: SPARQL query and result to integrate chemical knowledge from AOP-Wiki with WikiPathways, by utilizing Wikidata RDF.

Conclusion

Here we showed that the RDF transformation of AOP-Wiki content can assist in the accessibility and expansion of toxicological knowledge by allowing semantic interoperability. This opens many opportunities to increase the usefulness of RDF for toxicological risk assessment, such as the increased annotation of biological processes, genes and proteins, and pathways involved in the AOPs.

We are working making available a SPARQL endpoint Docker image to simplify the use of the data, and include it in the OpenRiskNet e-infrastructure to provide AOP knowledge for automated risk assessment workflows.

References

[1] Schechtman, L. M. (2002). *ILAR J.* 43, S85–S94. doi:10.1093/ilar.43.Suppl_1.S85.

[2] Ankley, G. T., et al. (2010). *Environ. Toxicol. Chem.* 29, 730–741. doi:10.1002/etc.34.

[3] Slenter, D. N., et al. (2018). *Nucleic Acids Res.* 46, D661–D667. doi:10.1093/nar/gkx1064.

[4] Waagmeester, A., et al. (2016). *PLOS Comput Biol.* 12(6) doi:10.1371/journal.pcbi.1004989.

[5] Erxleven, F., et al. (2014). *Springer, Cham.* doi:10.1007/978-3-319-11964-9_4.

Correspondence to:
Marvin Martens
ORCID: 0000-0003-2230-0840
marvin.martens@maastrichtuniversity.nl
@marvinm2

Maastricht University
Department of Bioinformatics – BiGCaT
http://www.bigcat.unimaas.nl
T +3143 388 1343

This project has received funding from the European Union's Horizon 2020 research and innovation programme project EU-ToxRisk under grant agreement No. 681002 and EINFRA-22-2016 programme project OpenRiskNet under grant agreement No. 731075.