

Discovering and assessing the quality of emergent Adverse Outcome Pathway knowledge

AOP-Wiki 3.0 Idea Generation and Discussion Workshop

August 16th, 2023

Nate Pollesch, PhD

USEPA Office of Research and Development - Duluth, MN

Disclaimer: The views expressed in this presentation are those of the author(s) and do not necessarily represent the views or policies of the Agency. Any mention of trade names or commercial products does not constitute EPA endorsement or recommendation for use.



Discovering and assessing the quality of emergent Adverse Outcome Pathway knowledge



Discovering and assessing the quality of emergent Adverse Outcome Pathway knowledge

Extracting and benchmarking emerging adverse
outcome pathway knowledge *ToxSci* 2019
- Pollesch, NL; Villeneuve, DL; O'Brien, JM



Discovering and assessing the quality of emergent Adverse Outcome Pathway knowledge

Extracting and benchmarking emerging adverse
outcome pathway knowledge *ToxSci* 2019

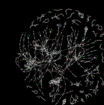
- Pollesch, NL; Villeneuve, DL; O'Brien, JM

Assessing quality of emergent adverse outcome
pathways *Work in Progress*

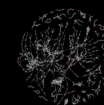
- Pollesch, NL; Olker, JH; Wang, R-L



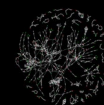
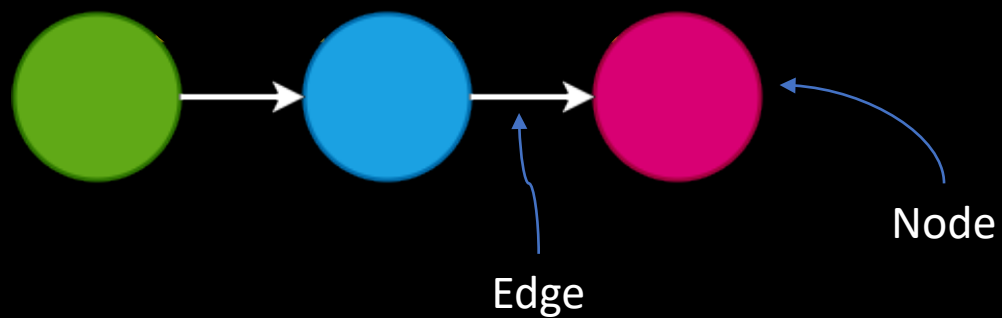
Discovering emergent A O P



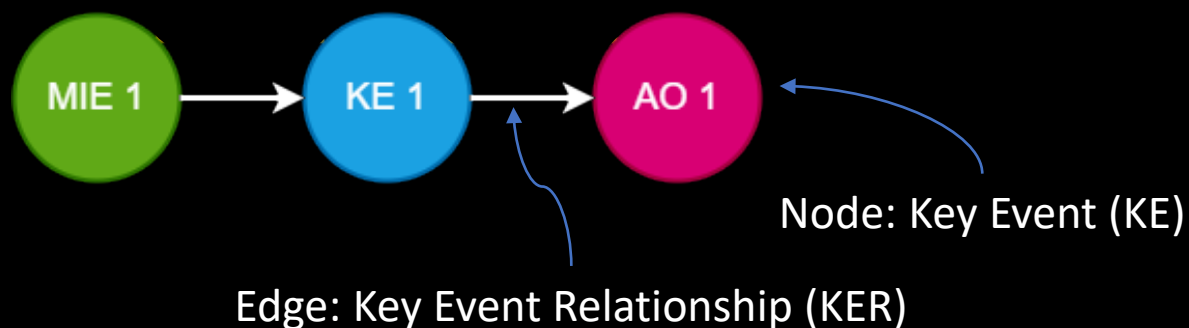
Discovering emergent A O P



Discovering emergent A O P



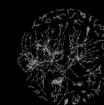
Discovering emergent A O P



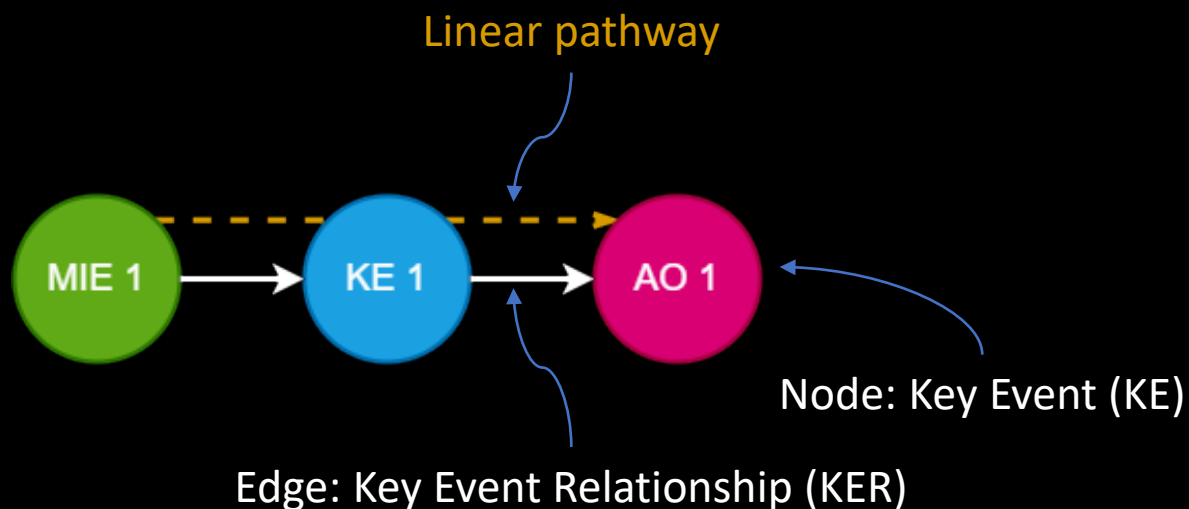
MIE: Molecular Initiating Event

KE: Key Event

AO: Adverse Outcome



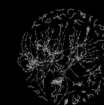
Discovering emergent A O P



MIE: Molecular Initiating Event

KE: Key Event

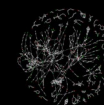
AO: Adverse Outcome



Discovering emergent A O P

A little about the methods:

- Building the Network
 - Key event relationship info was downloaded from AOPwiki.org and was used to create an edge list
 - The edge list was converted to a graph using *igraph* package in R Studio



Discovering emergent AOP

A little about the methods:

- Building the Network
 - Key event relationship info was downloaded from AOPwiki.org and was used to create an edge list
 - The edge list was converted to a graph using *igraph* package in R Studio

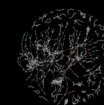
AOP-Wiki AOPs Key Events KE Relationships Stressors Login Register

Search relationships Search Find by ID Find by ID API

Navigational tip: click on "leads to" or the ID column to navigate to the KE Relationship.

KE Relationships

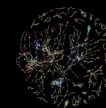
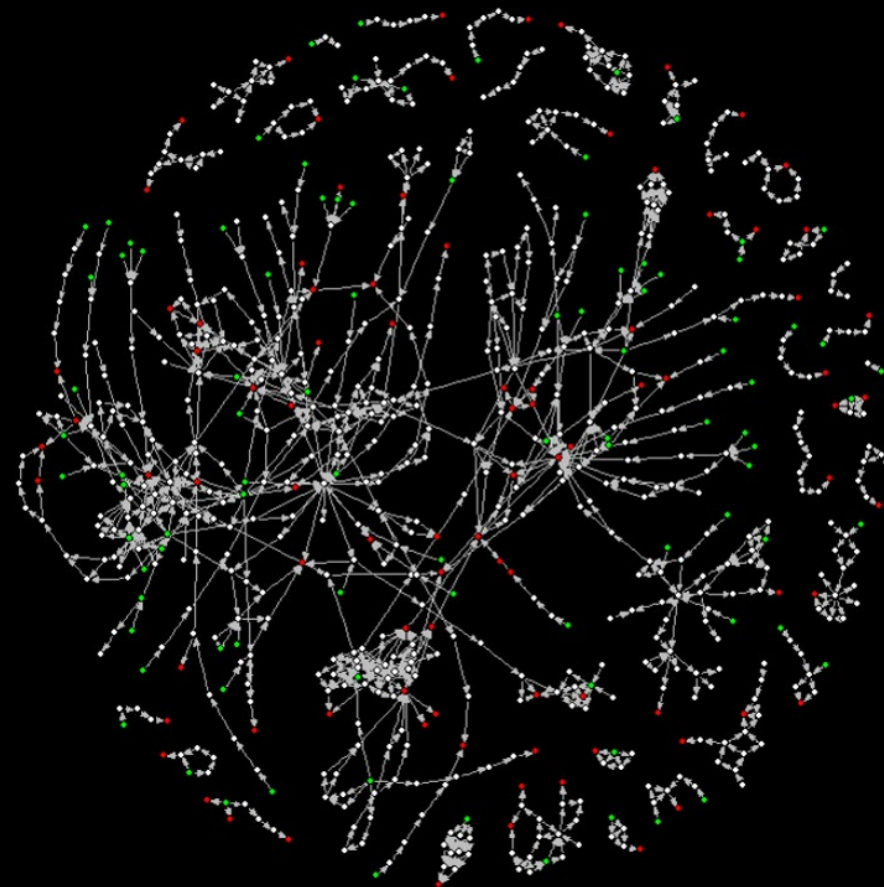
Id	Upstream Event ▲	Relationship Type	Downstream Event
5	Reduction, 17beta-estradiol synthesis by ovarian granulosa cells	leads to	Reduction, Plasma 17beta-estradiol concentrations
9	Decreased, 3-hydroxyacyl-CoA dehydrogenase type-2 activity	leads to	Decreased, Mitochondrial fatty acid beta-oxidation
750	Activation, 5HT2c	leads to	N/A, Unknown
1857	ACh Synaptic Accumulation	leads to	Activation, Muscarinic Acetylcholine Receptors
454	ACh Synaptic Accumulation	leads to	Increased, Atrioventricular block and bradycardia
456	ACh Synaptic Accumulation	leads to	Increased Cholinergic Signaling
459	ACh Synaptic Accumulation	leads to	AChE Inhibition
11	AChE Inhibition	leads to	ACh Synaptic Accumulation
2653	AChE Inhibition	leads to	Activation of Cyp2E1
449	AChE Inhibition	leads to	Increased, Atrioventricular block and bradycardia
450	AChE Inhibition	leads to	Respiratory distress/arrest



Discovering emergent A O P

A little about the methods:

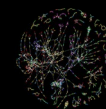
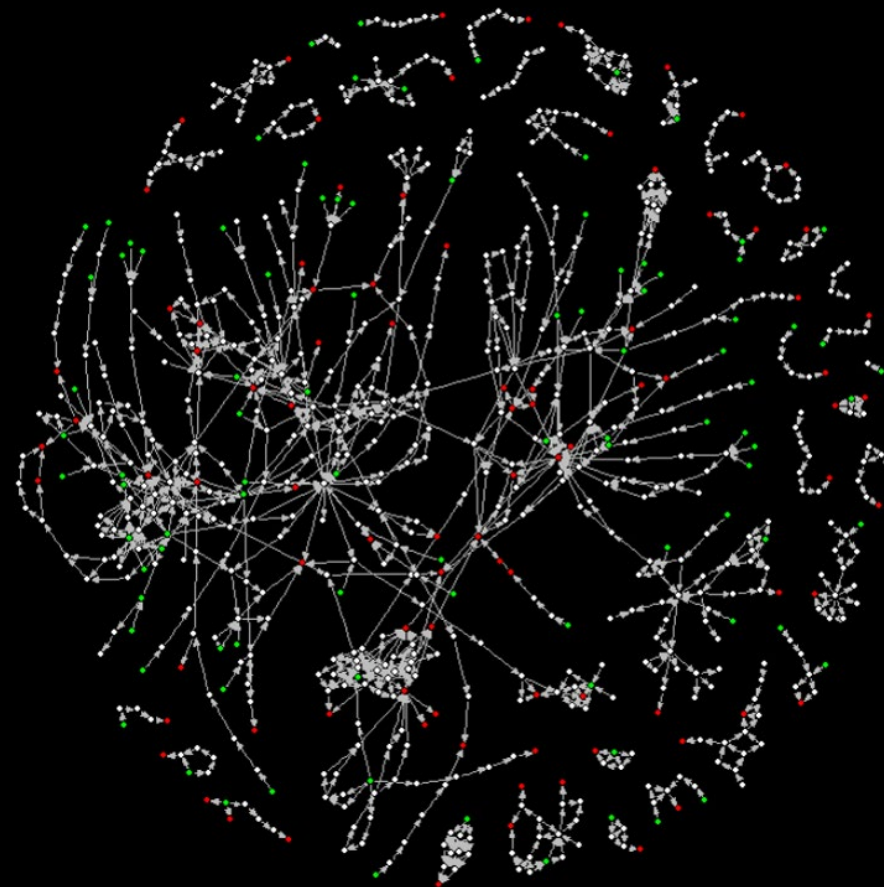
- Building the Network
 - Key event relationship info was downloaded from AOPwiki.org and was used to create an edge list
 - The edge list was converted to a graph using *igraph* package in R Studio



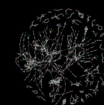
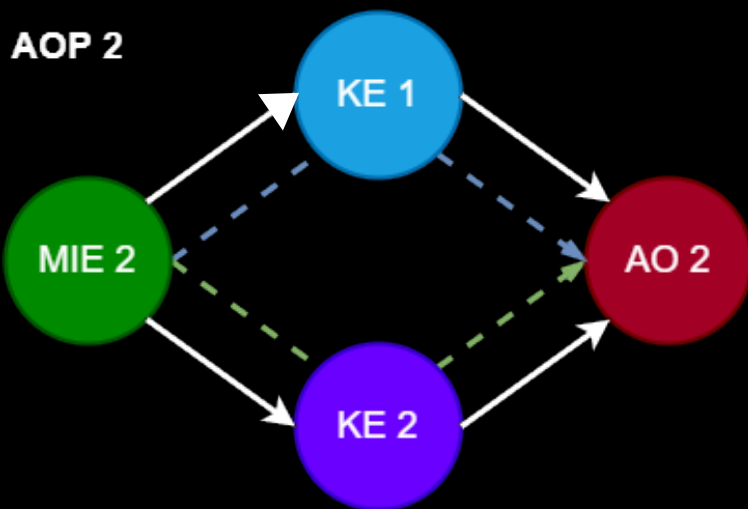
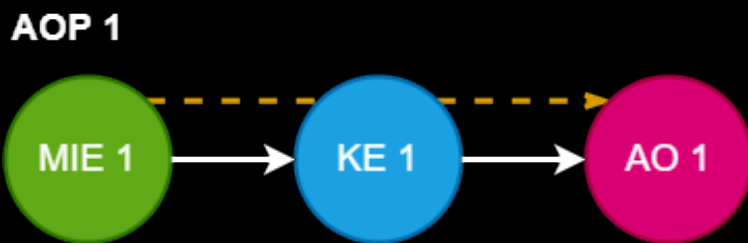
Discovering emergent A O P

A little about the methods:

- Graph attributes assigned to each KE and KER
 - KE: KE ID, AOP ID, designator (MIE, KE, AO), and Level of biological organization (LOBO)
 - KER: KER ID, AOP ID, weight of evidence (WOE), and quantitative understanding (QE)



Discovering emergent AOP

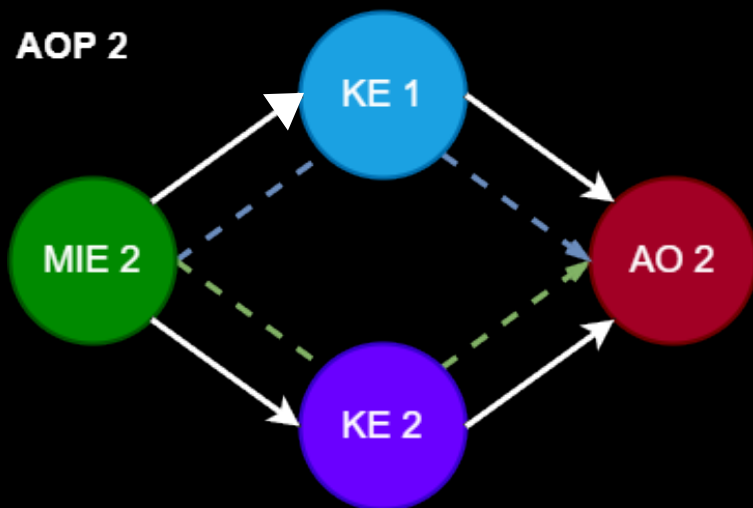


Discovering emergent AOP

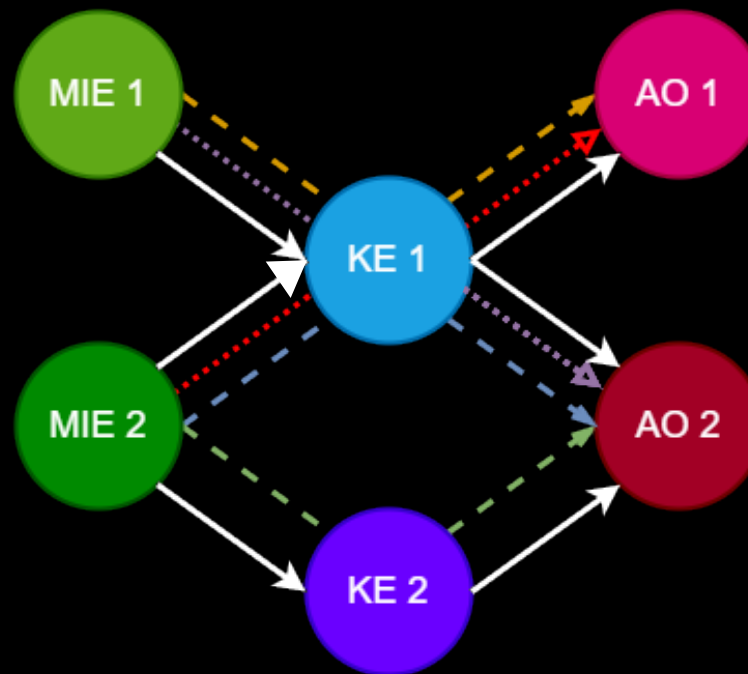
AOP 1



AOP 2

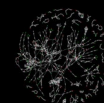


Networked AOP via KE 1 Sharing



LEGEND

-  Key Event Relationship
-  Expert-Specified LAOP
-  Emergent LAOP



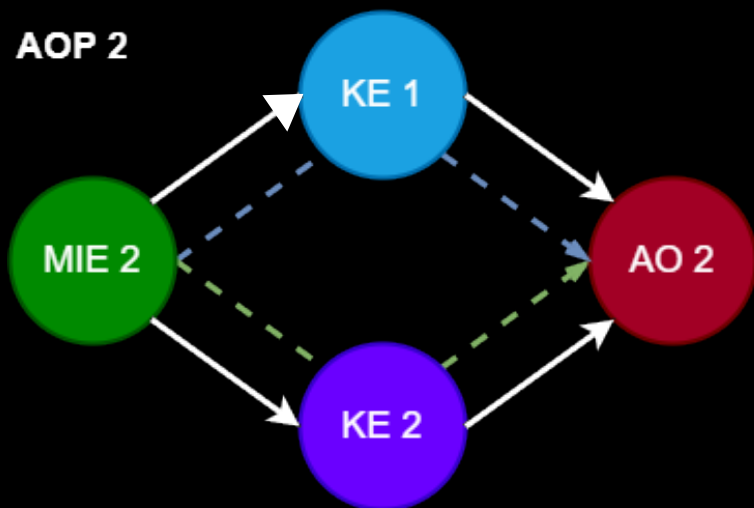
Discovering emergent AOP



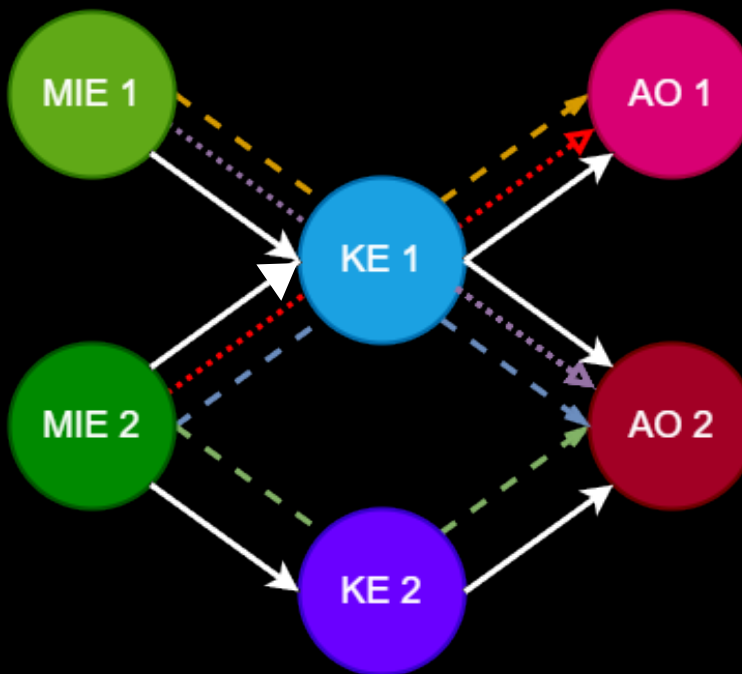
AOP 1



AOP 2



Networked AOP via KE 1 Sharing

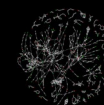


LEGEND

- Key Event Relationship
- Expert-Specified LAOP
- Emergent LAOP

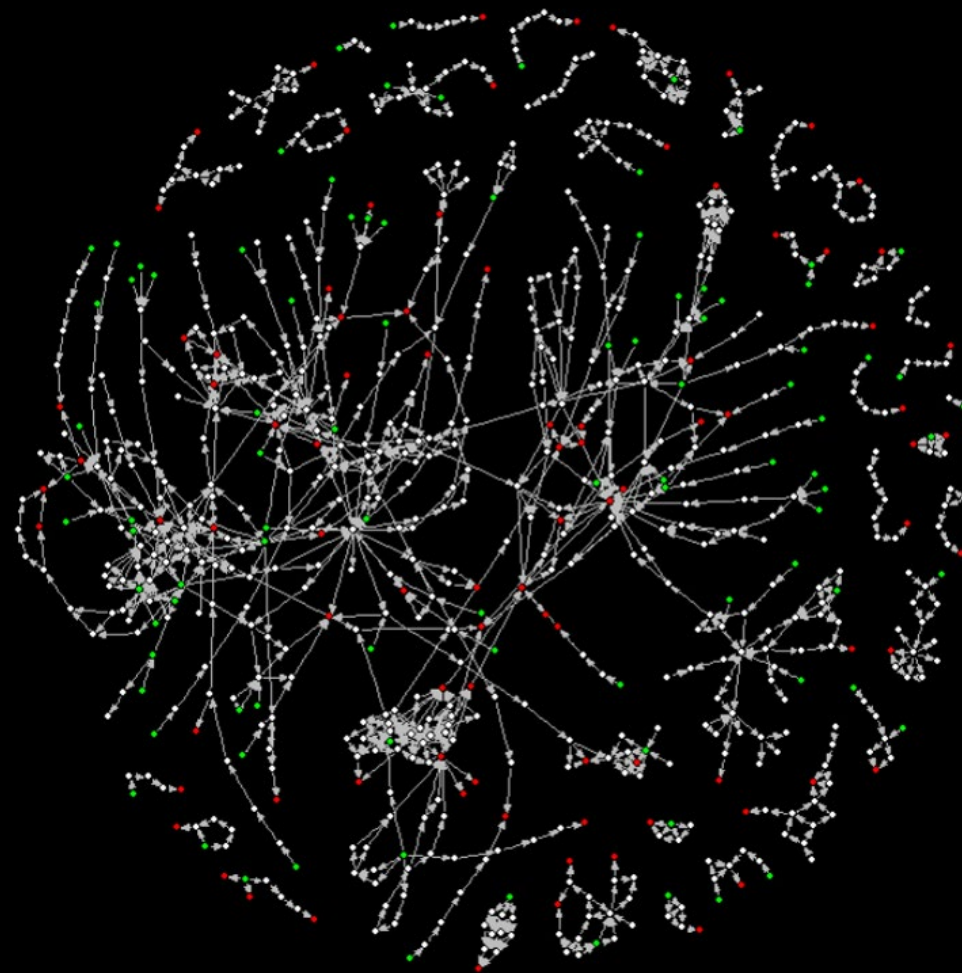
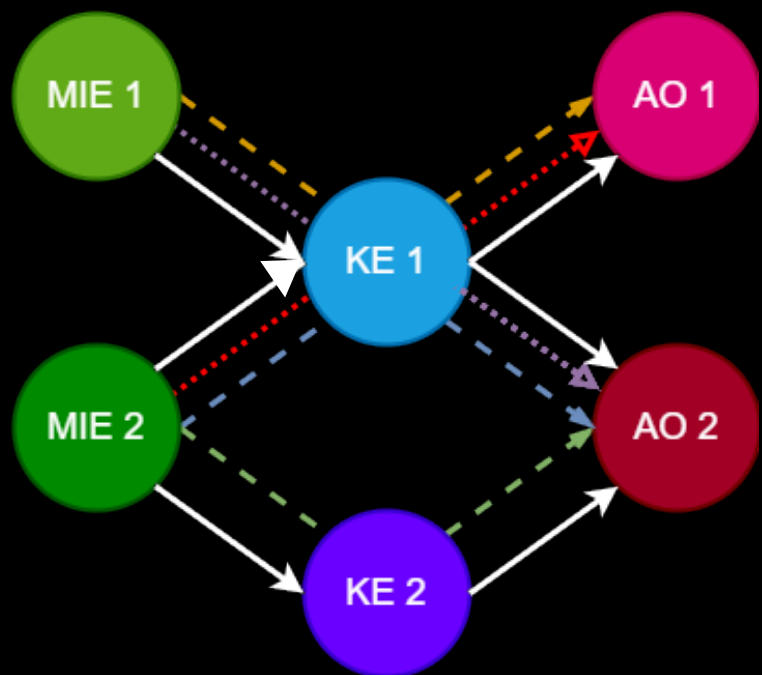
This KE sharing example had

- 2 AOPs
- 3 unique expert-specified linear AOPs
- 2 unique emergent AOPs



Discovering emergent A O P

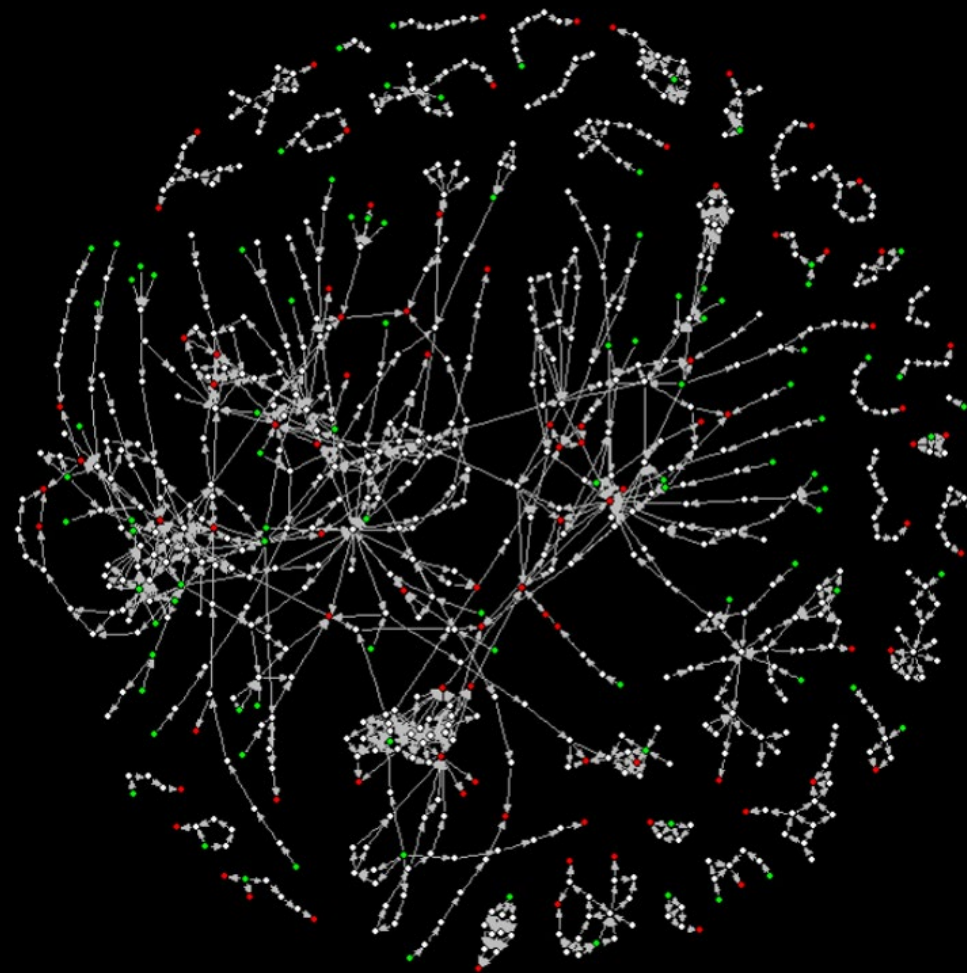
Networked AOP via KE 1 Sharing



Discovering emergent A O P

Our Analysis showed
that the AOP-Wiki had

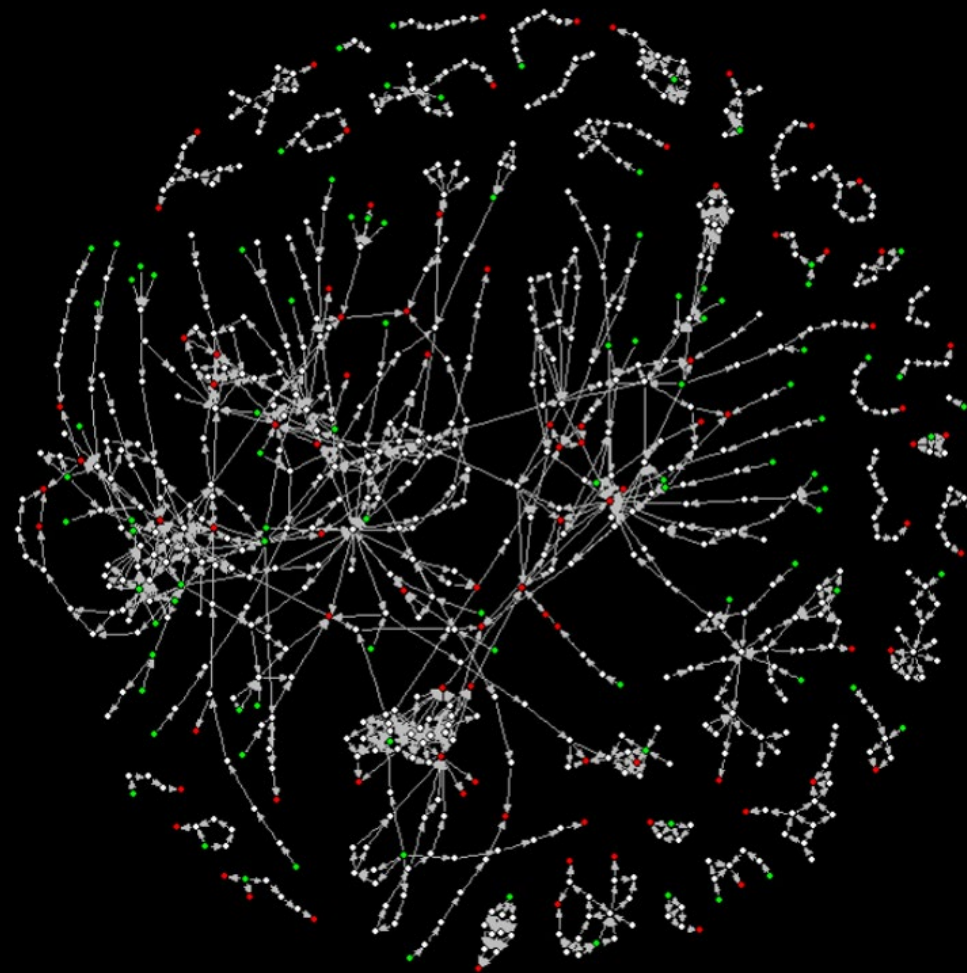
- 187 AOPs
- 471 unique expert-specified linear AOPs
- 9405 unique emergent linear AOPs



Discovering emergent A O P

Our Analysis showed
that the AOP-Wiki had

- 187 AOPs
- 471 unique expert-specified linear AOPs
- 9405 unique
emergent linear AOPs

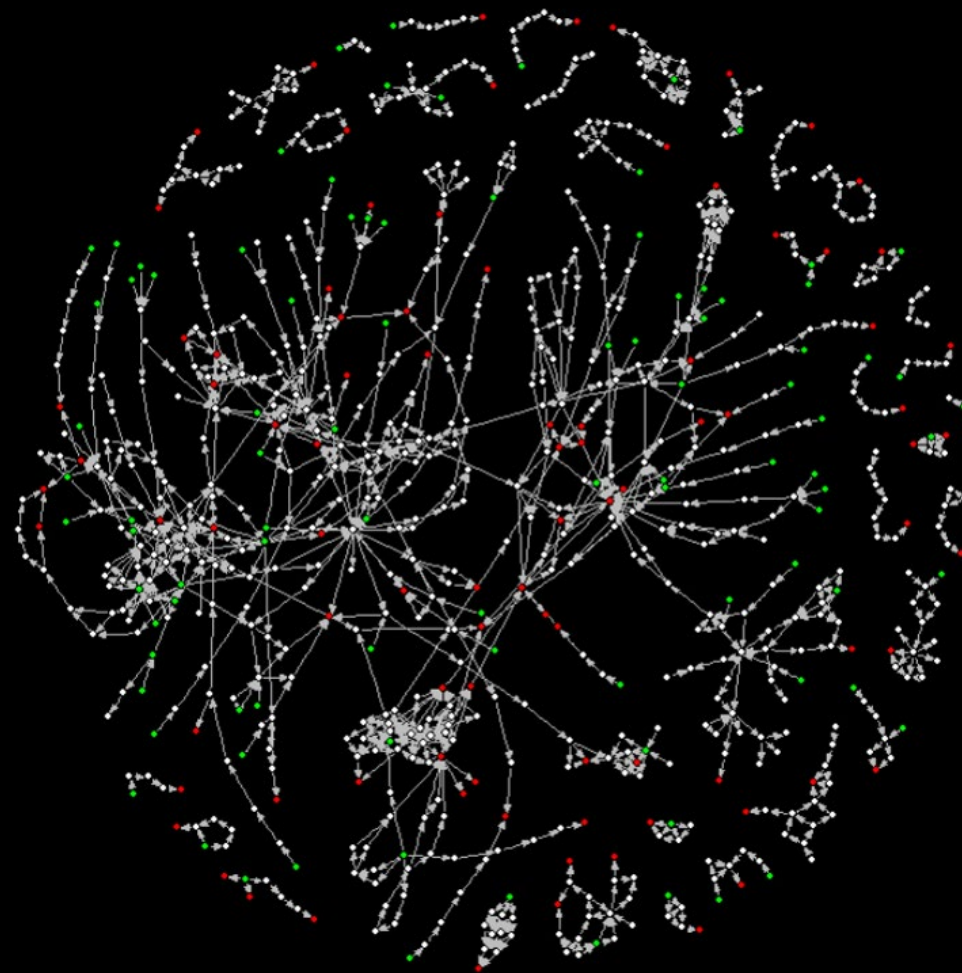


Discovering emergent A O P

Our Analysis showed
that the AOP-Wiki had

- 187 AOPs
- 471 unique expert-specified linear AOPs
- 9405 unique
emergent linear AOPs

We found them, but are
they useful?



Discovering and **assessing the quality** of emergent Adverse Outcome Pathway knowledge

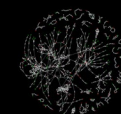


Extracting and benchmarking emerging adverse outcome pathway knowledge *ToxSci* 2019

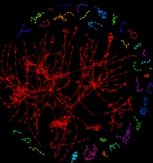
- Pollesch, NL; Villeneuve, DL; O'Brien, JM

Assessing quality of emergent adverse outcome pathways *Work in Progress*

- Pollesch, NL; Olker, JH; Wang, R-L



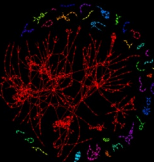
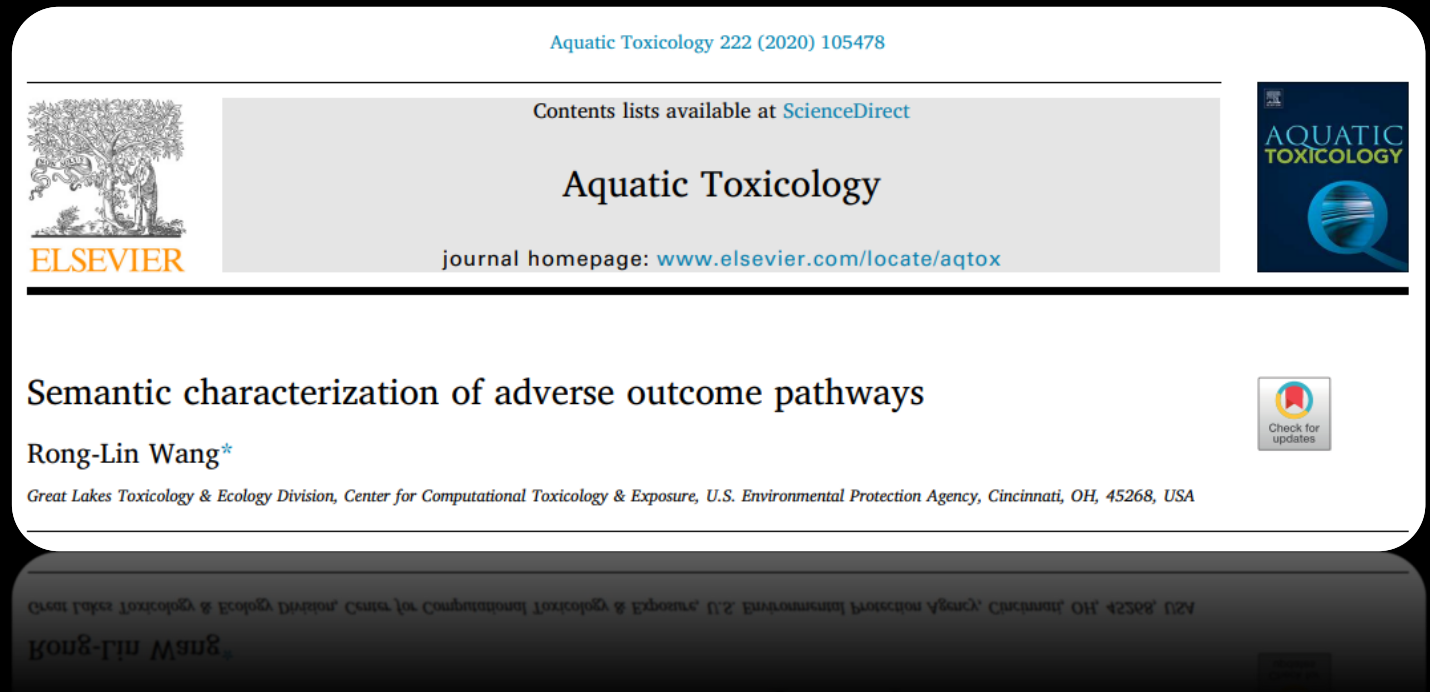
Assessing the quality of emergent AOP knowledge



Assessing the quality of emergent AOP knowledge



In 2020, Rong-Lin Wang (CCTE/GLTED) showed that semantic analysis provides a method for quantifying the quality of AOPs based on semantic coherence



Assessing the quality of emergent AOP knowledge



In 2020, Rong-Lin Wang (CCTE/GLTED) showed that semantic analysis provides a method for quantifying the quality of AOPs based on **semantic coherence**

This is a method of converting expressions into logical definitions and looking at similarity/closeness of those definitions to one another based on their structured content

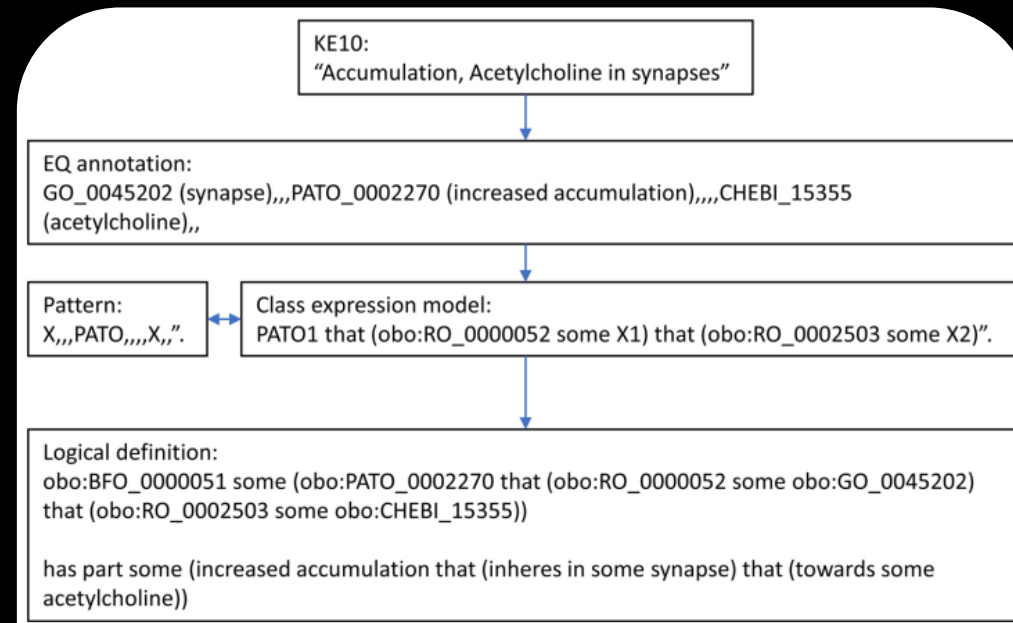
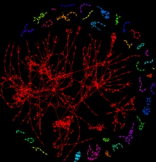


Figure 1b: Creation of logical definitions from KEs (Wang, 2020)



Assessing the quality of emergent AOP knowledge



From Wang, (2020)

- *A logical definition here is defined as several phenotypic expressions composed of more atomic terms from multiple reference ontologies and linked together by appropriate object properties from the Relations Ontology.*

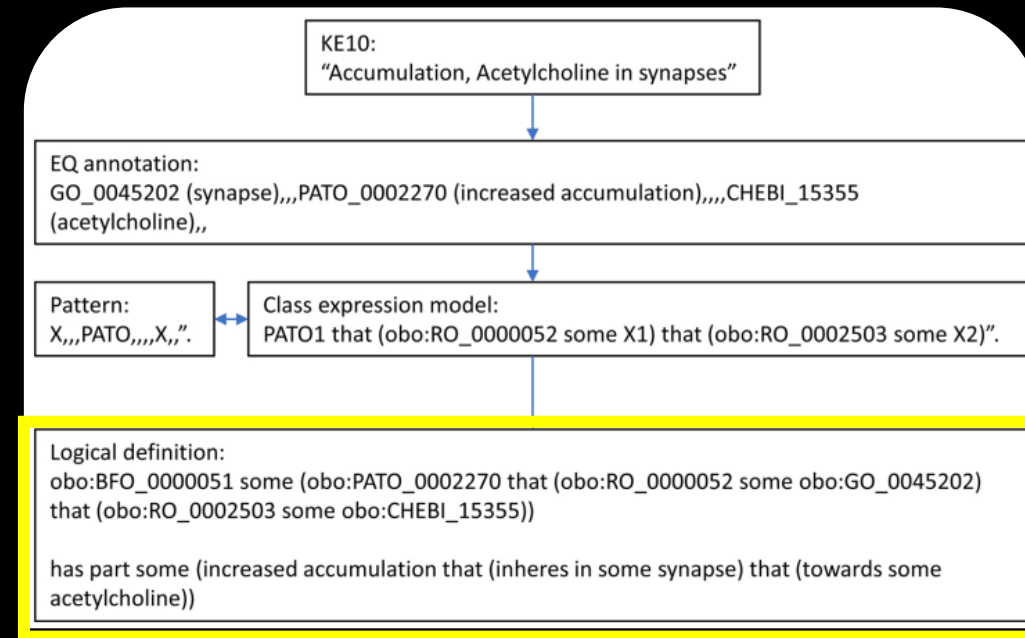
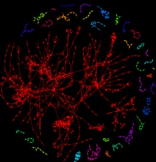


Figure 1b: Creation of logical definitions from KEs (Wang, 2020)



Assessing the quality of emergent AOP knowledge



From Wang, (2020)

- A logical definition here is defined as several phenotypic expressions composed of more atomic terms from multiple reference ontologies and linked together by appropriate object properties from the Relations Ontology.
- There are many reference domain ontologies developed thus far, such as the Gene Ontology, Chemical Entities of Biological Interest Ontology, Cell Ontology, Phenotype and Trait Ontology, and various anatomy ontologies (<http://obofoundry.org/>).

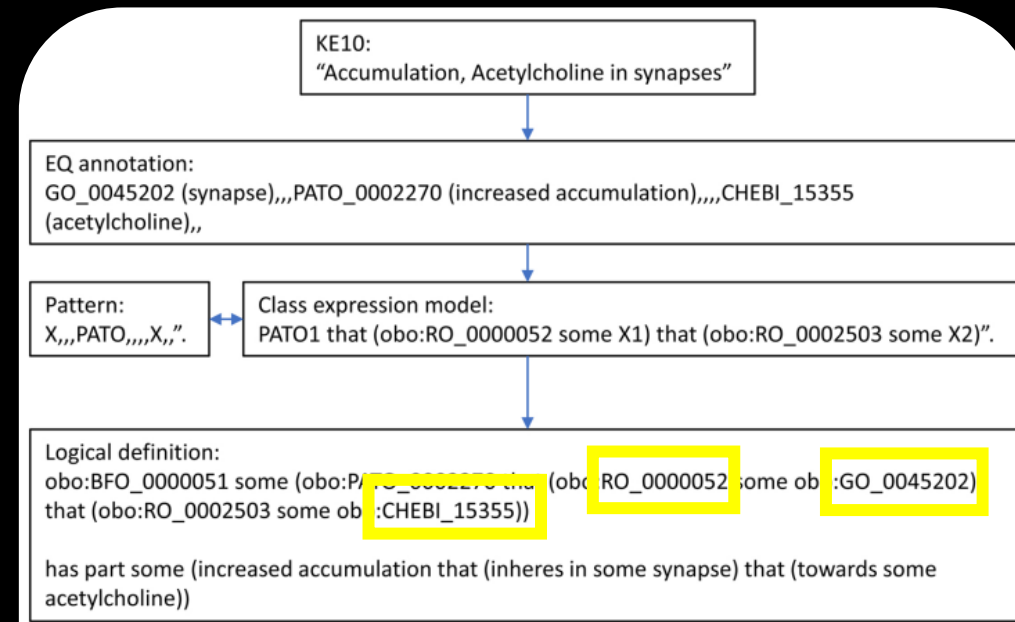
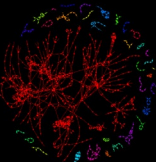


Figure 1b: Creation of logical definitions from KEs (Wang, 2020)



Assessing the quality of emergent AOP knowledge



Rong-Lin did this for **all the KEs in the AOP-wiki**. Once these definitions were created, he could then look at semantic similarity of KEs and AOPs

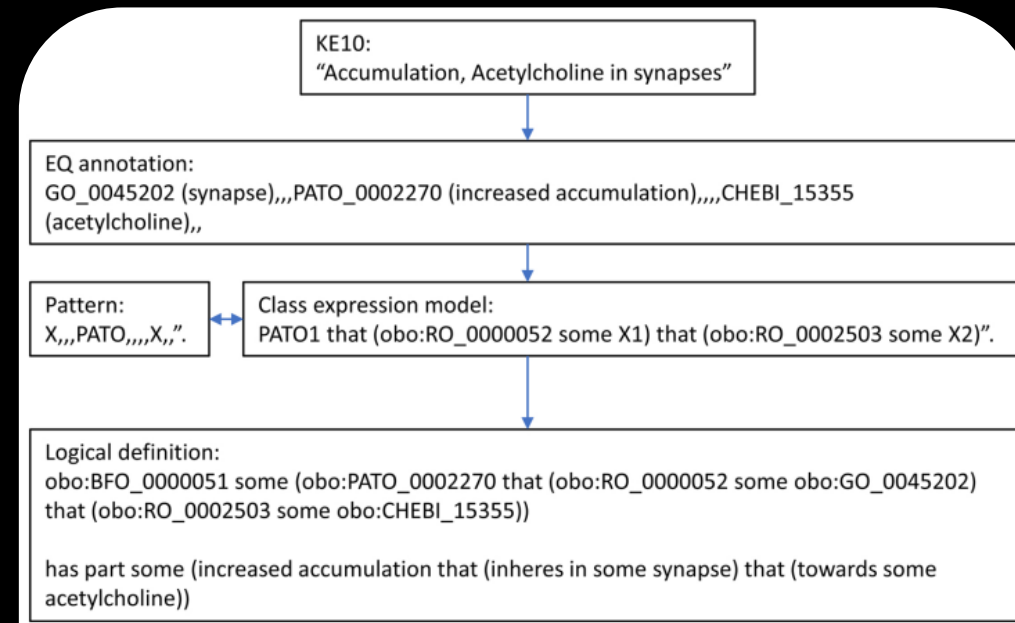
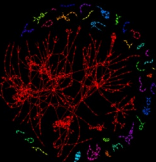


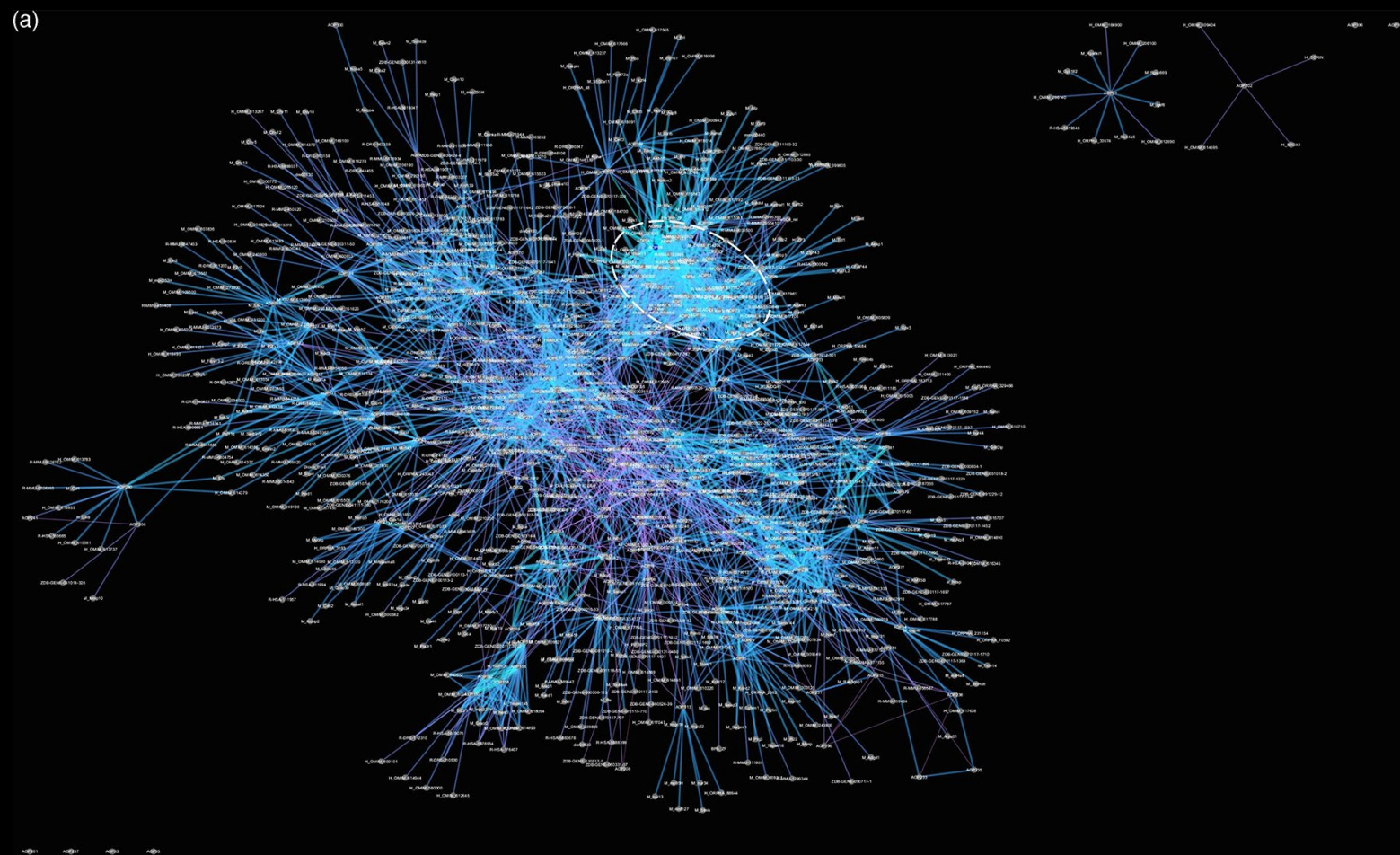
Figure 1b: Creation of logical definitions from KEs (Wang, 2020)



Assessing the quality of emergent AOP knowledge

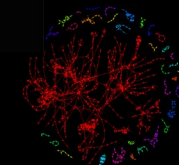


Semantic similarity networks of AOPs, CSPPs, genes, pathways, and diseases (Wang, 2020)



Assessing quality of emergent adverse outcome pathways *Work in Progress*

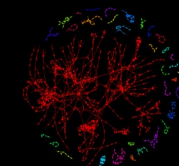
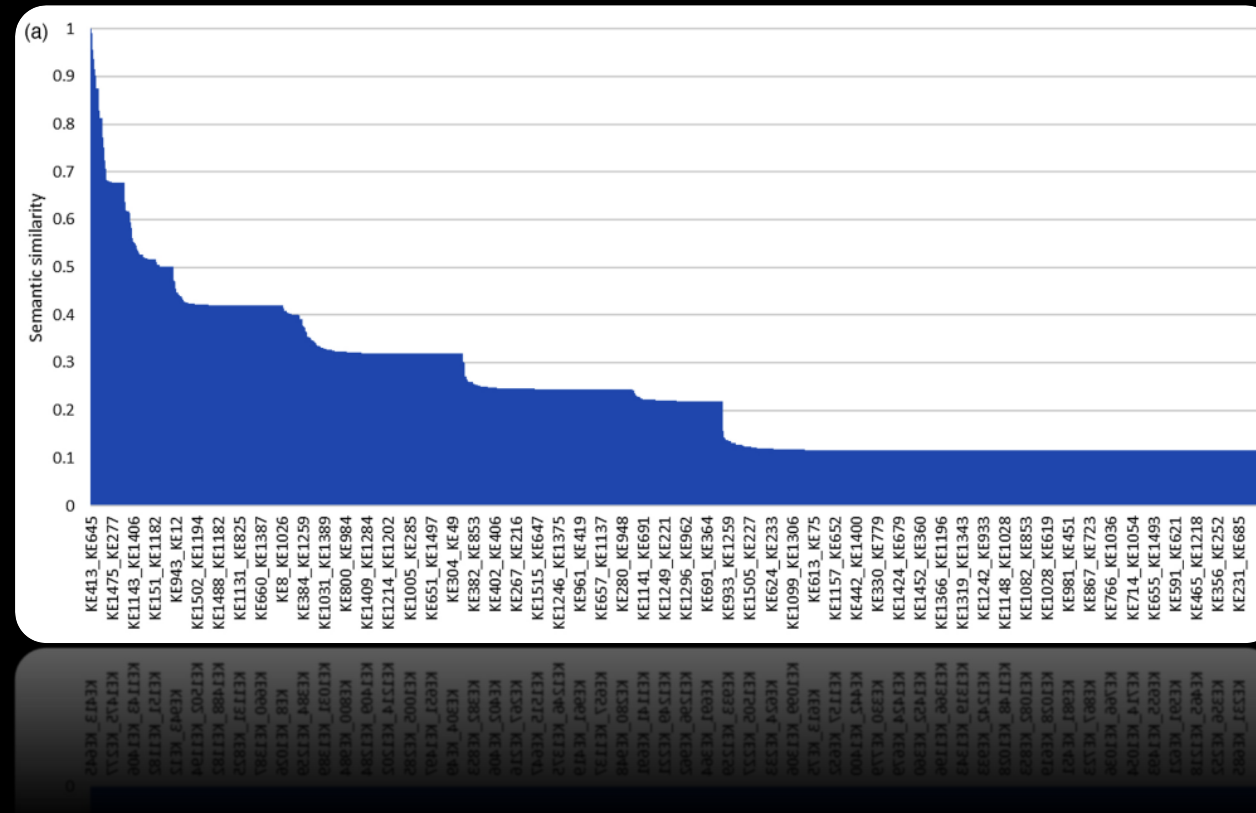
- Pollesch, NL; Olker, JH; Wang, R-L



Assessing the quality of emergent AOP knowledge

Rong-Lin did this for all the KEs in the AOP-wiki. Once these definitions were created, he could then look at semantic similarity of KEs and AOPs

Figure 2a: Pairwise similarities of logical KE definitions (Wang, 2020)



Assessing the quality of emergent AOP knowledge



Rong-Lin did this for all the KEs in the AOP-wiki. Once these definitions were created, he could then look at semantic similarity of KEs and AOPs

Figure 2a: Pairwise similarities of logical KE definitions (Wang, 2020)

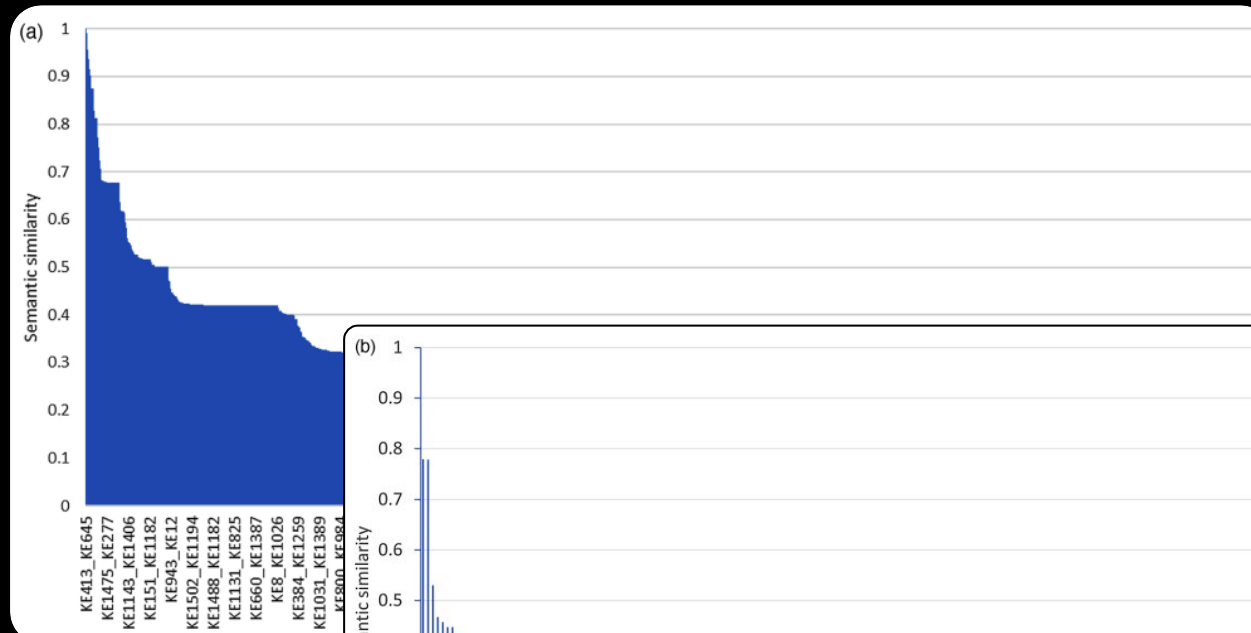
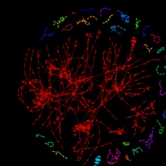
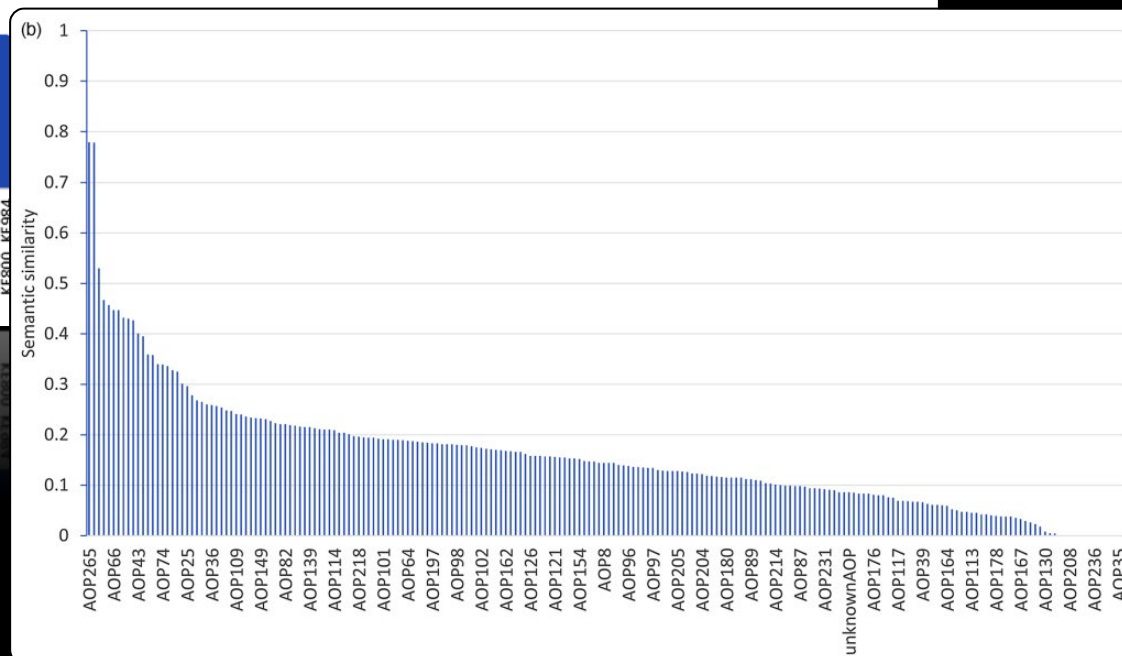
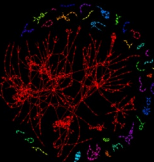


Figure 2b: Mean pairwise similarity within AOPs (Wang, 2020)



Assessing the quality of emergent AOP knowledge

I was excited to learn about Rong-Lin's work and asked him if he would be able to use his methods on the emergent AOPs that we found in Pollesch et al., (2019)



Assessing the quality of emergent AOP knowledge

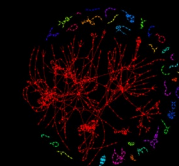


I was excited to learn about Rong-Lin's work and asked him if he would be able to use his methods on the emergent AOPs that we found in Pollesch et al., (2019)

He agreed and ran the analyses on the nearly 10000 LAOPs I sent

20 of 471 expert-specified LAOPs identified in Pollesch et al., (2019)

	A	B	C	D	E	F	G	H	I	J
1	1	2	3	4	5	6	7	8	9	10
2	888	887	177	889	890	896				
3	888	887	177	890	896					
4	103	1264	1265	988	1266	1267	1268	993	990	351
5	998	1000	858	859	860	861	862	863	864	
6	998	1000	858	860	861	862	863	864		
7	998	1000	858	861	862	863	864			
8	408	3	219	405						
9	408	3	219	405	406					
10	667	64	669	682	616	613				
11	201	195	52	381	55	188	352			
12	201	195	52	381	55	188	352	341		
13	201	195	52	381	382	385	386	341		
14	201	195	52	381	383	385	386	341		
15	201	195	52	381	55	385	386	341		
16	97	155	185	336						
17	97	185	336							
18	97	336								
19	12	10	444	351						
20	12	10	39	445	351					
21	12	351								



Assessing the quality of emergent AOP knowledge



I was excited to learn about Rong-Lin's work and asked him if he would be able to use his methods on the emergent AOPs that we found in Pollesch et al., (2019)

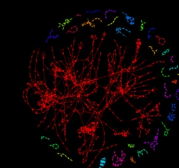
He agreed and ran the analyses on the nearly 10000 LAOPs I sent

20 of 471 expert-specified LAOPs identified in Pollesch et al., (2019)

20 of 9408 emergent LAOPs identified in Pollesch et al., (2019)

	A	B	C	D	E	F	G	H	I	J
1	1	2	3	4	5	6	7	8	9	10
2	888	887	177	889	890	896				
3	888	887	177	890	896					
4	103	1264	1265	988	1266	1267	1268	993	990	351
5	998	1000	858	859	860	861	862	863	864	
6	998	1000	858	860	861	862	863	864		
7	998	1000	858	861	862	863	864			

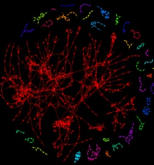
	A	B	C	D	E	F	G	H	I	J
1	1	2	3	4	5	6	7	8	9	10
2	888	887	177	55	188	890	896			
3	888	887	177	55	352	188	890	896		
4	888	887	177	889	890	188	55	352		
5	888	887	177	889	890	188	352			
6	888	887	177	890	188	55	352			
7	888	887	177	890	188	352				
8	888	887	177	55	188	352				
9	888	887	177	55	352					
10	888	887	177	889	890	188	55	352	341	
11	888	887	177	889	890	188	55	352	618	341
12	888	887	177	889	890	188	55	385	386	341
13	888	887	177	889	890	188	55	386	341	
14	888	887	177	889	890	188	352	341		
15	888	887	177	889	890	188	352	618	341	
16	888	887	177	890	188	55	352	341		
17	888	887	177	890	188	55	352	618	341	
18	888	887	177	890	188	55	385	386	341	
19	888	887	177	890	188	55	386	341		
20	888	887	177	890	188	352	341			
21	888	887	177	890	188	352	618	341		



Assessing the quality of emergent AOP knowledge

He agreed and ran the analyses on the nearly 10000 LAOPs I sent

The results from the first round of analyses are exciting



Assessing the quality of emergent AOP knowledge



He agreed and ran the analyses on the nearly 10000 LAOPs I sent

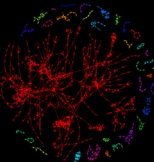
The results from the first round of analyses are exciting

Semantic analysis metrics

For an AOP with N KEs, $AOP_N = KE_1 \rightarrow KE_2 \rightarrow \dots \rightarrow KE_N$, and with $SS(KE_x, KE_y) \in [0,1]$ as the semantic similarity between KEs x and KE y :

- $SSS_{Mean}(AOP_N) = \frac{1}{N-1} \sum_{i=1}^{N-1} SS(KE_i, KE_{i+1})$
- $SSS_{Min}(AOP_N) = \min_{i=1, \dots, N-1} SS(KE_i, KE_{i+1})$
- $PSS_{Mean}(AOP_N) = \frac{1}{\binom{N}{2}} \sum_{i < j} SS(KE_i, KE_j)$ for $i, j \in \{1, \dots, N\}$

Where KE_1 is the *MIE* and KE_N is the *AO*. Each metric captures different AOP semantic qualities

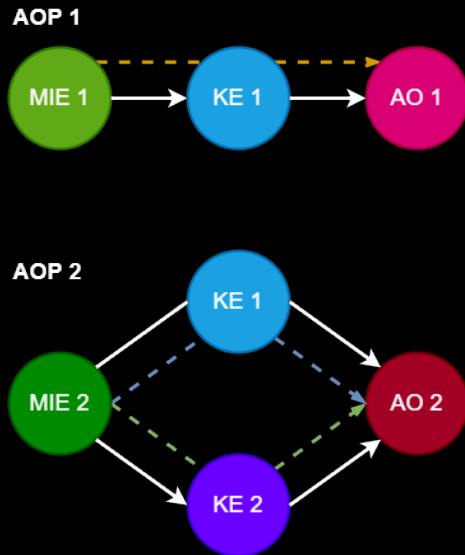


Assessing the quality of emergent AOP knowledge



He agreed and ran the analyses on the nearly 10000 LAOPs I sent

The results from the first round of analyses are exciting



Semantic analysis metrics

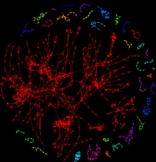
For an AOP with N KEs, $AOP_N = KE_1 \rightarrow KE_2 \rightarrow \dots \rightarrow KE_N$, and with $SS(KE_x, KE_y) \in [0,1]$ as the semantic similarity between KEs x and KE y :

$$- SSS_{Mean}(AOP_N) = \frac{1}{N-1} \sum_{i=1}^{N-1} SS(KE_i, KE_{i+1})$$

$$- SSS_{Min}(AOP_N) = \min_{i=1, \dots, N-1} SS(KE_i, KE_{i+1})$$

$$- PSS_{Mean}(AOP_N) = \frac{1}{\binom{N}{2}} \sum_{i < j} SS(KE_i, KE_j) \text{ for } i, j \in \{1, \dots, N\}$$

Where KE_1 is the *MIE* and KE_N is the *AO*. Each metric captures different AOP semantic qualities



Assessing the quality of emergent AOP knowledge



He agreed and ran the analyses on the nearly 10000 LAOPs I sent

The results from the first round of analyses are exciting

Semantic analysis metrics

For an AOP with N KEs, $AOP_N = KE_1 \rightarrow KE_2 \rightarrow \dots \rightarrow KE_N$, and with $SS(KE_x, KE_y) \in [0,1]$ as the semantic similarity between KEs x and KE y :

$$- SSS_{Mean}(AOP_N) = \frac{1}{N-1} \sum_{i=1}^{N-1} SS(KE_i, KE_{i+1})$$

$$- SSS_{Min}(AOP_N) = \min_{i=1, \dots, N-1} SS(KE_i, KE_{i+1})$$

$$- PSS_{Mean}(AOP_N) = \frac{1}{\binom{N}{2}} \sum_{i < j} SS(KE_i, KE_j) \text{ for } i, j \in \{1, \dots, N\}$$

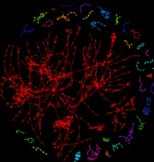
Where KE_1 is the *MIE* and KE_N is the *AO*. Each metric captures different AOP semantic qualities

Results

	SSS_{Mean}	SSS_{Min}	PSS_{Mean}
AOP: Mean (n=236)	0.153 ^a	.025	0.151 ^b
• Max (ID)	• 0.553 (66)	• 0.421 (41)	• 0.391 (66)
EAOP: Mean (n=9408)	0.144 ^a	.007	0.141 ^b
• Max (ID)	• 0.409 (9402)	• 0.319 (^c)	• 0.424 (9403)

^{a,b} Difference in means not statistically significant ($\alpha = 0.05$)

^c 12 EAOPs tied for the maximum SSS_{Min} value: 6761, 6762, 6766, 6767, 6768, 6772, 6778, 6784, 9401, 9402, 9403, 9404. Note: EAOP IDs were assigned for analysis purposes and details of specific EAOPs can be shared upon request



Assessing the quality of emergent AOP knowledge

He agreed and ran the analyses on the nearly 10000 LAOPs I sent

The results from the first round of analyses are exciting

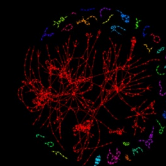
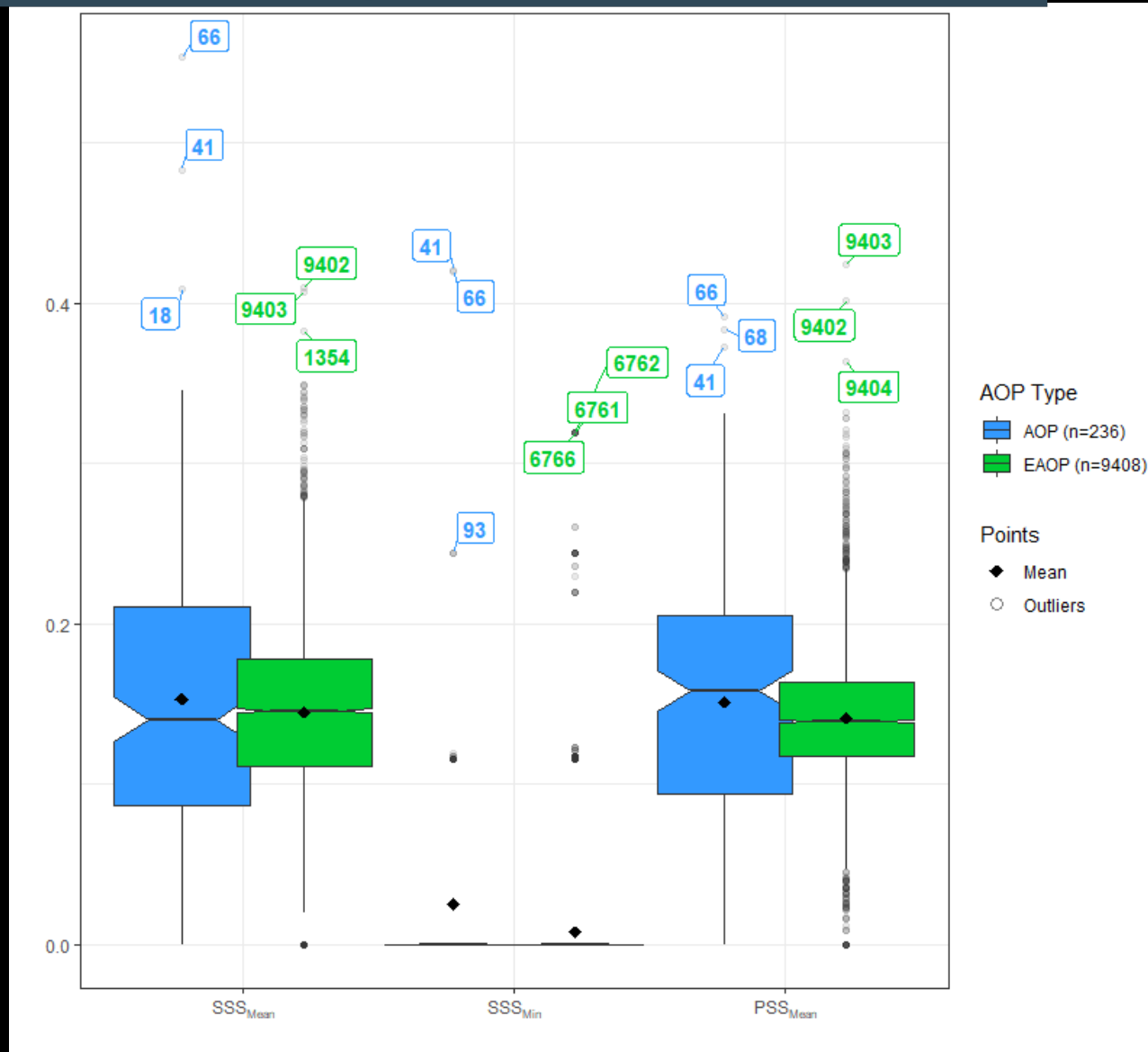
Semantic analysis metrics

For an AOP with N KEs, $AOP_N = KE_1 \rightarrow KE_2 \rightarrow \dots \rightarrow KE_N$, and with $SS(KE_x, KE_y) \in [0,1]$ as the semantic similarity between KEs x and KE y :

- $SSS_{Mean}(AOP_N) = \frac{1}{N-1} \sum_{i=1}^{N-1} SS(KE_i, KE_{i+1})$
- $SSS_{Min}(AOP_N) = \min_{i=1, \dots, N-1} SS(KE_i, KE_{i+1})$
- $PSS_{Mean}(AOP_N) = \frac{1}{\binom{N}{2}} \sum_{i < j} SS(KE_i, KE_j)$ for $i, j \in \{1, \dots, N\}$

Where KE_1 is the MIE and KE_N is the AO. Each metric captures different AOP semantic qualities

Semantic similarity metrics for AOPs and EAOPs

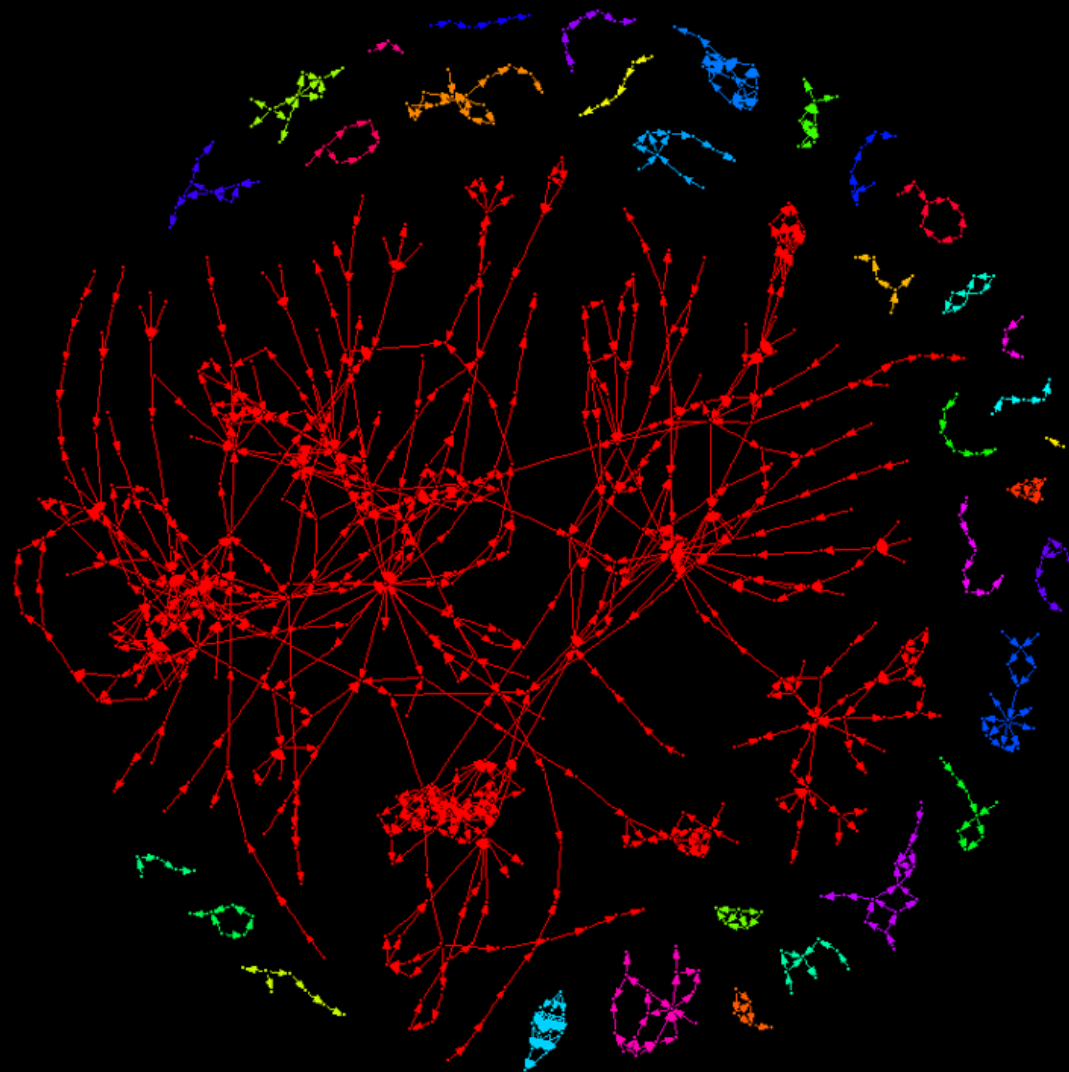


Assessing the quality of emergent AOP knowledge

The results from the first round of analyses are exciting

Why?

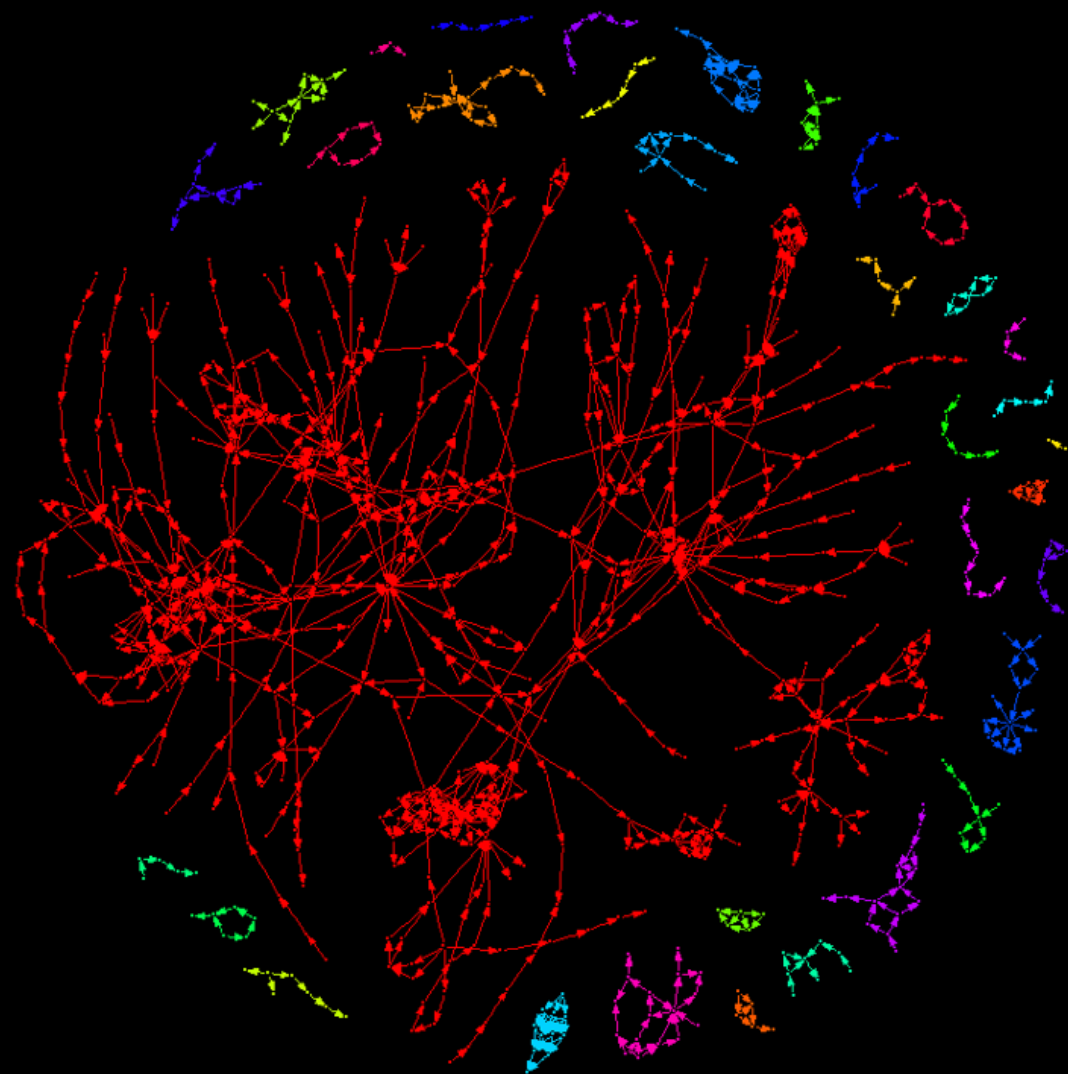
If these results hold up, that means that we have computationally identified (likely to be thousands of) unique emergent AOPs that are of high semantic quality



Assessing the quality of emergent AOP knowledge

If these results hold up, that means that we have computationally identified (likely to be thousands of) unique emergent AOPs that are of high semantic quality

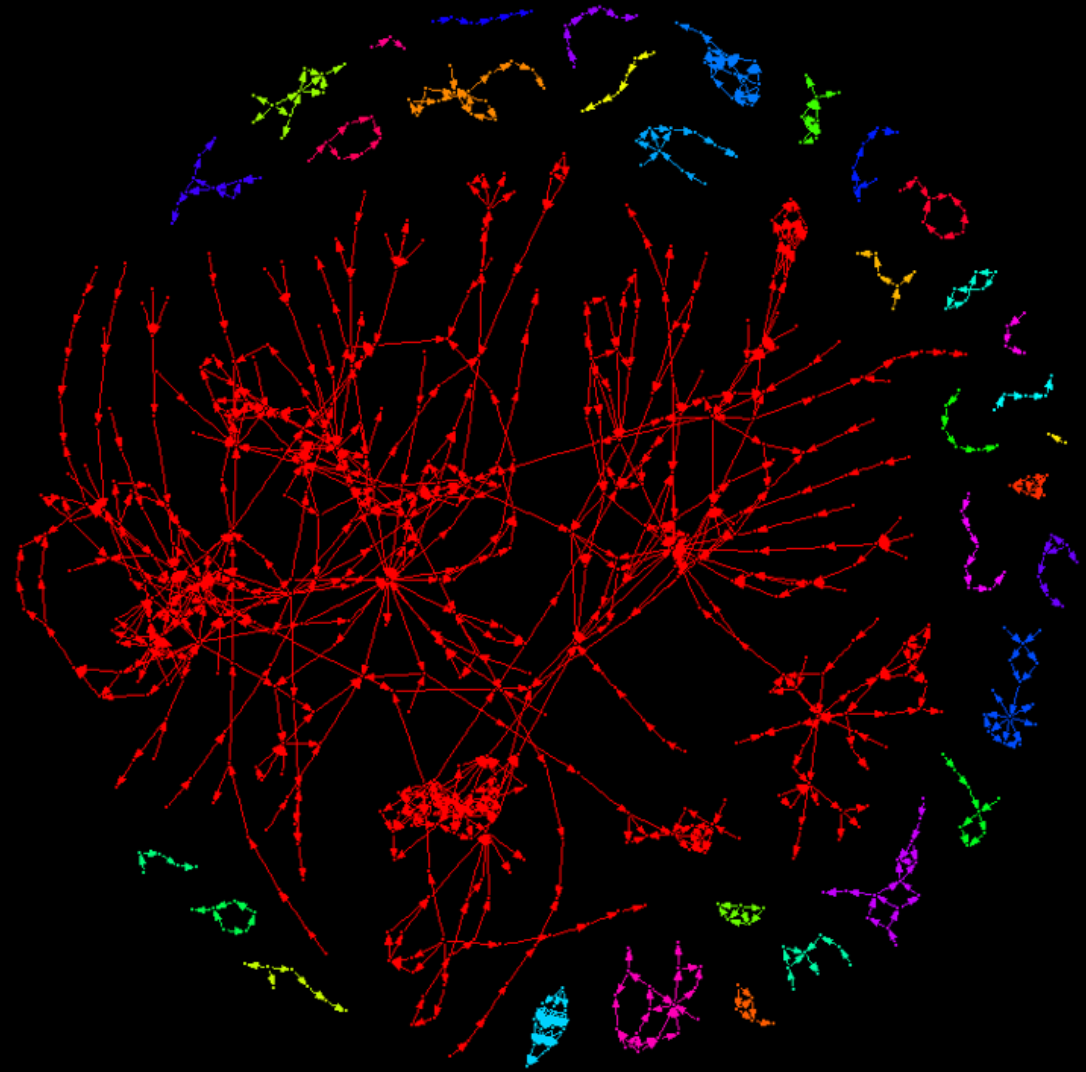
So, what is next?



Assessing the quality of emergent AOP knowledge

So, what is next?

- Manual inspection of computational results
 - Are there patterns in highly coherent LAOPs
 - Do these results pass the sniff test?
 - Identifying a few of the outstanding emergent LAOPs to use as case studies



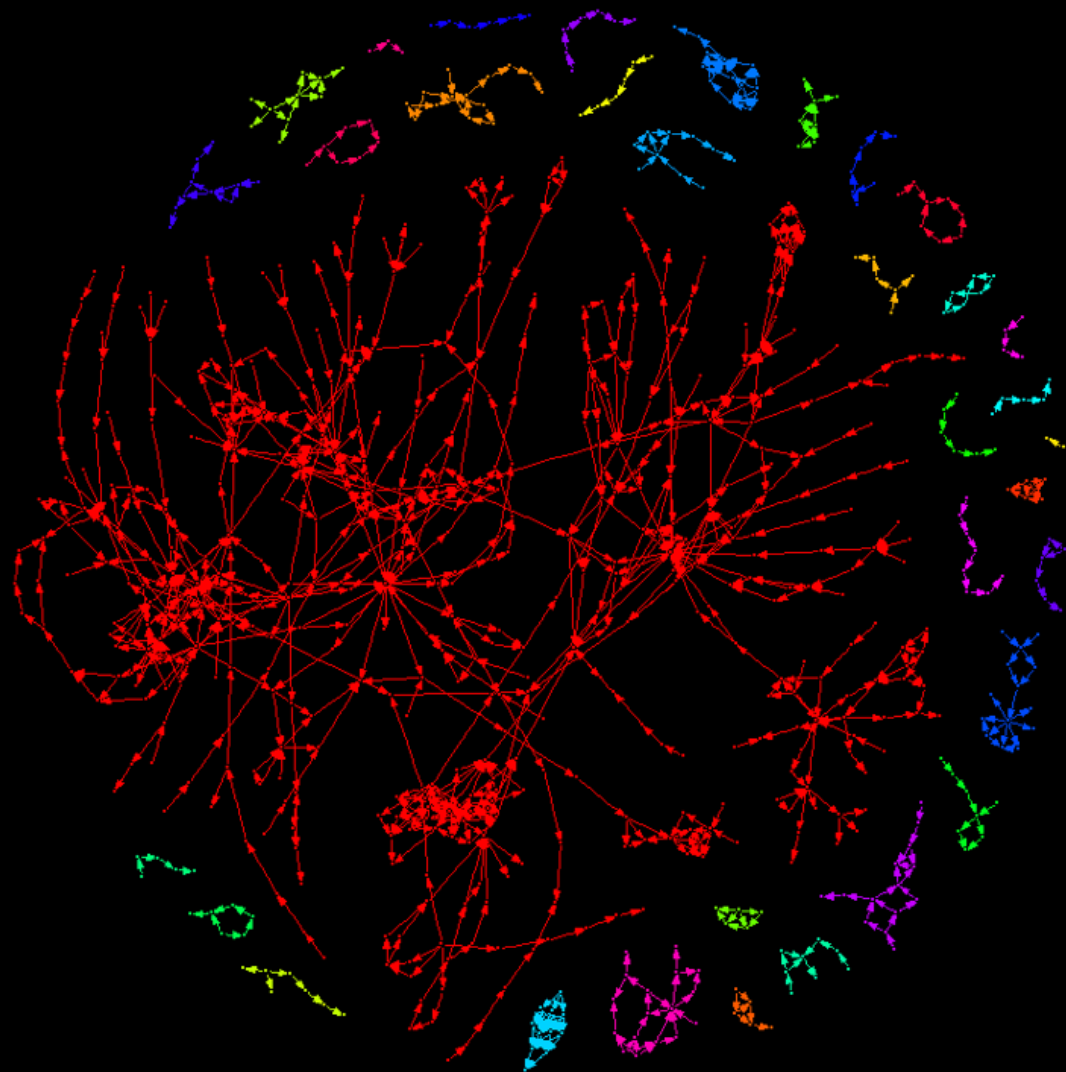
Assessing the quality of emergent AOP knowledge



So, what is next?

- Manual inspection of computational results
 - Are there patterns in highly coherent LAOPs
 - Do these results pass the sniff test?
 - Identifying a few of the outstanding emergent LAOPs to use as case studies

We are also comparing how other metrics of quality compare to semantic quality metrics



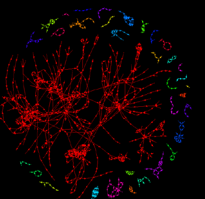
Assessing the quality of emergent AOP knowledge

What aspects of the AOP-Wiki enabled these analyses?

Emergent AOP:

Graph attributes assigned to each KE and KER

- KE: KE ID, AOP ID, and designator (MIE, KE, AO)
- KER: KER ID, AOP ID, weight of evidence (WOE), and quantitative understanding (QE)



Assessing the quality of emergent AOP knowledge



What aspects of the AOP-Wiki enabled these analyses?

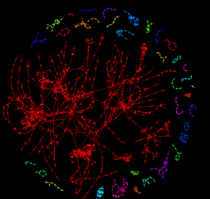
Emergent AOP:

Graph attributes assigned to each KE and KER

- KE: KE ID, AOP ID, and designator (MIE, KE, AO)
- KER: KER ID, AOP ID, weight of evidence (WOE), and quantitative understanding (QE)

Semantic Analysis:

KE descriptions and supplementary information on applicability



Assessing the quality of emergent AOP knowledge



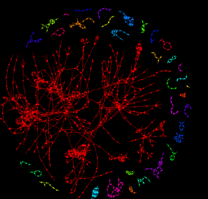
What aspects of these analyses could be useful in the AOP-Wiki?

Emergent AOP:

- How involved are KEs in emergent AOPs? (We calculated this)
- High quality emergent AOPs could be identified, evaluated for quality, and added explicitly to the AOP-Wiki

Semantic Analysis:

- Semantic Similarity Scores for all KERS *and* for any pair of KEs in the WIKI
- Semantic similarity metrics for all AOPs
- Suggested KEs for developers (based on high SS scores)

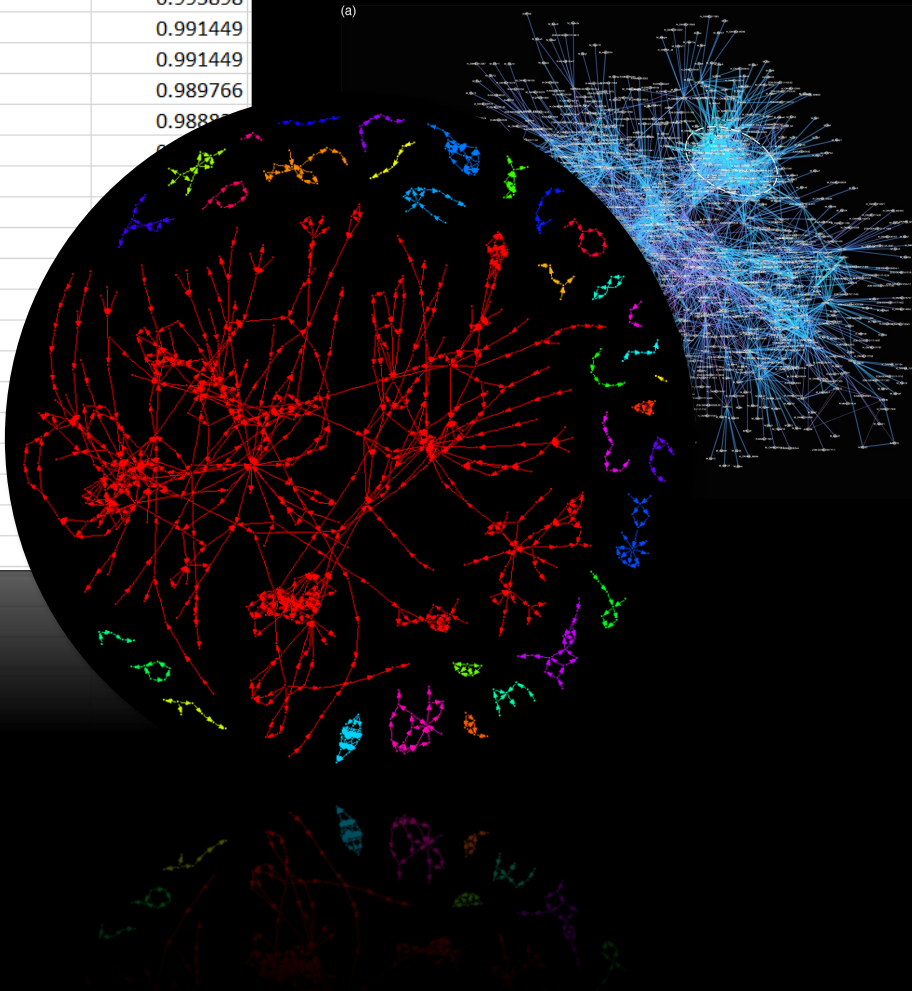


Discovering and Assessing the quality of emergent Adverse Outcome Pathway knowledge



If you have questions

262	AOP158	AOP155	0.995898
263	userAOP94	AOP18	0.991449
264	userAOP95	AOP18	0.991449
265	userAOP23	AOP23	0.989766
266	userAOP112	AOP216	0.9888
267	EmergingAOP1814	AOP216	
268	userAOP33	AOP25	
269	EmergingAOP6816	AOP216	
270	userAOP348	AOP6	
271	userAOP143	AOP42	
272	EmergingAOP6812	AOP216	
273	userAOP113	AOP238	
274	userAOP163	AOP54	
275	EmergingAOP1815	AOP238	
276	userAOP171	AOP134	
277	userAOP160	AOP7	
278	AOP238	AOP268	
279	AOP268	AOP238	
280	EmergingAOP1869	AOP42	



Please reach out to me at Pollesch.Nathan@epa.gov with any questions.

EXTRA:

So, what can we do with AOP networks?

- That depends on the network, in particular:
 - The structure of the network
 - The data available to describe the relationships
- Some network analyses:
 - Shortest path analyses
 - Topological sorting for visualization
- From a quantitative modeling standpoint, it also depends on the network, in particular:
 - The structure of the network
 - The data available to describe the relationships

Perkins et al describe how qAOP can utilize networks of increasing specificity

Bayesian networks are coming under increasing use in ecotoxicology.

IEAM special issue, Burgoon et al, Landis et al

Petri nets (Edhlund et al)