# WikiNetworks tutorial

This tutorial demonstrates the use of the WikiNetworks package to download and process pathway representations from WikiPathways database into machine-readable networks that can be used with visualization software or in downstream programming applications

*Author: Mukta Palshikar, mukta\_palshikar@urmc.rochester.edu*

### Install WikiNetworks

**WikiNetworks requires Python 3.6 or newer.**

With pip:

pip install --upgrade wikinetworks

### Import all functions from wikinetworks

from wikinetworks import \*

### Create an instance of the WikiPathways database using the bioservices package

\_\_all\_\_ = ["WikiPathways"]  
s = WikiPathways()

WARNING [bioservices:WikiPathways]: URL of the services contains a double //.Check your URL and remove trailing /

### Specify the pathways from WikiPathways that you wish to convert into networks.

Wikipathways entries are specified by Wikipathways ID - you can find this ID in the URL for the pathway of interest.

For example, the URL for the Apoptosis (*Homo sapiens*) pathway is https://www.wikipathways.org/index.php/Pathway:WP254 and the Wikipathways ID for this pathway is WP254. The ID begins with ‘WP’ and is followed by a string of digits.

Here, we select the Sandbox pathway test https://www.wikipathways.org/index.php/Pathway:WP4 to demonstrate WikiNetworks’ functionality.

Note that the pathway ID is the first element in a list, and is *not* specified as a string. The user can extend this list to include other pathways to download and process.

pathID = "WP4"

### Download and view the curation tags for the pathway of interest

#### Download a dictionary of curation tags for the selected pathway using the WikiPathways API

curationTagsDict = getCurationTags(s, pathID)  
print(curationTagsDict)

{'tags': [{'name': 'Curation:Tutorial', 'displayName': 'Test pathway', 'pathway': {'id': 'WP4', 'url': 'https://www.wikipathways.org/index.php/Pathway:WP4', 'name': 'Sandbox pathway test', 'species': 'Homo sapiens', 'revision': '119193'}, 'revision': None, 'text': '', 'timeModified': '20170121205948', 'userModified': 'AlexanderPico'}, {'name': 'Curation:MissingXRef', 'displayName': 'Annotate nodes', 'pathway': {'id': 'WP4', 'url': 'https://www.wikipathways.org/index.php/Pathway:WP4', 'name': 'Sandbox pathway test', 'species': 'Homo sapiens', 'revision': '119193'}, 'revision': None, 'text': '3 out of 14 DataNodes have a missing external reference. Check DataNode table at the bottom of the page.', 'timeModified': '20200727110051', 'userModified': 'MaintBot'}, {'name': 'Curation:NoInteractions', 'displayName': 'Fix interactions', 'pathway': {'id': 'WP4', 'url': 'https://www.wikipathways.org/index.php/Pathway:WP4', 'name': 'Sandbox pathway test', 'species': 'Homo sapiens', 'revision': '119193'}, 'revision': None, 'text': '2 out of 11 lines are not properly connected.', 'timeModified': '20190924090150', 'userModified': 'MaintBot'}]}

#### Print pathway and curation information

*Caveat*: We advise users to carefully examine the curation tags. Curation tags usually indicate whether a manual curator found issues in the pathway entry on WikiPathways, and that these errors have not yet been fixed by the pathway creator. While WikiNetworks attempts to automatically find and fix these common errors, we encourage the users of WikiNetworks to go over the output network and double-check interactions. Despite our best efforts, it’s possible that the WikiNetworks algorithm does not find or incompletely fixes the pathway issues. This is due to the widely varying drawing styles and quality of the pathways in Wikipathways.

processCurationTags(curationTagsDict)

\*\*\*PATHWAY INFORMATION\*\*\*  
  
Name: Curation:Tutorial   
  
Display Name: Test pathway   
  
id: WP4  
  
url: https://www.wikipathways.org/index.php/Pathway:WP4  
  
name: Sandbox pathway test  
  
species: Homo sapiens  
  
revision: 119193  
  
  
\*\*\*CURATION WARNINGS:\*\*\*  
  
  
\*\*\*CURATION TAG 1\*\*\*  
  
Name: Curation:MissingXRef  
Display Name: Annotate nodes  
Description: 3 out of 14 DataNodes have a missing external reference. Check DataNode table at the bottom of the page.  
  
\*\*\*CURATION TAG 2\*\*\*  
  
Name: Curation:NoInteractions  
Display Name: Fix interactions  
Description: 2 out of 11 lines are not properly connected.  
\*\*\*\*\*\*\*\*\*

### Download and process pathways into networks

This is the driver function for the WikiNetworks processing algorithm. This function downloads GPML files for the requested pathway(s) from WikiPathways, processes the pathway into a network and attempts to correct drawing errors, and finally outputs the network as: 1. a NetworkX digraph object that can be used for downstream programmatic applications, 1. a graphml file that can be imported into other simulation, processing or visualization tools such as CasQ, CellNOpt and Cytoscape, and 1. a simple interaction format (SIF) file that can be used similarly to the graphml file and is also easily human-readable.

graph = runParsePathway(s, pathID)

WARNING [bioservices:WikiPathways]: URL of the services contains a double //.Check your URL and remove trailing /

#### Print basic network statistics

Print the number of nodes and edges in the network.

print("Nodes:", len(graph.nodes())) #print number of nodes in the pathway  
print("Edges:", len(graph.edges())) #print number of edges in the pathway

Nodes: 13  
Edges: 18

#### Print the names of the output files

print("SIF graph:", '\_'.join([pathID, "edgeList.csv"])) #location of the edgelist file. SIF = simple interaction format  
print("Graphml: ", ''.join([pathID, "\_graph.graphml"])) #location of the graphml file

SIF graph: WP4\_edgeList.csv  
Graphml: WP4\_graph.graphml

### Download and process multiple pathways into networks

To download and process multiple pathways from WikiPathways, it is straightforward to specify a list of pathways of interest and process them in a loop, as shown below.

testPathways = ["WP4", "WP2727","WP2456"] # Specify the test pathways using the WikiPathways IDs  
for pathID in testPathways:  
 curationTagsDict = getCurationTags(s, pathID) # download a dictionary of curation tags for the selected pathway  
 processCurationTags(curationTagsDict) # print pathway and curation information  
 graph = runParsePathway(s, pathID) # download and process pathways, generate a NetworkX digraph object  
 print("Nodes:", len(graph.nodes())) # print number of nodes in the pathway  
 print("Edges:", len(graph.edges())) # print number of edges in the pathway  
 print("SIF graph:", '\_'.join([pathID, "edgeList.csv"])) # location of the edgelist file. SIF = simple interaction format  
 print("Graphml: ", ''.join([pathID, "\_graph.graphml"])) # location of the graphml file  
 print("\*\*\*\*\*\*\*\*\*\*\n")

\*\*\*PATHWAY INFORMATION\*\*\*  
  
Name: Curation:Tutorial   
  
Display Name: Test pathway   
  
id: WP4  
  
url: https://www.wikipathways.org/index.php/Pathway:WP4  
  
name: Sandbox pathway test  
  
species: Homo sapiens  
  
revision: 119193  
  
  
\*\*\*CURATION WARNINGS:\*\*\*  
  
  
\*\*\*CURATION TAG 1\*\*\*  
  
Name: Curation:MissingXRef  
Display Name: Annotate nodes  
Description: 3 out of 14 DataNodes have a missing external reference. Check DataNode table at the bottom of the page.  
  
\*\*\*CURATION TAG 2\*\*\*  
  
Name: Curation:NoInteractions  
Display Name: Fix interactions  
Description: 2 out of 11 lines are not properly connected.  
\*\*\*\*\*\*\*\*\*  
  
  
  
WARNING [bioservices:WikiPathways]: URL of the services contains a double //.Check your URL and remove trailing /  
  
  
Nodes: 13  
Edges: 18  
SIF graph: WP4\_edgeList.csv  
Graphml: WP4\_graph.graphml  
\*\*\*\*\*\*\*\*\*\*  
  
  
\*\*\*PATHWAY INFORMATION\*\*\*  
  
Name: Curation:Reactome\_Approved   
  
Display Name: Community: Reactome   
  
id: WP2727  
  
url: https://www.wikipathways.org/index.php/Pathway:WP2727  
  
name: Cellular response to hypoxia  
  
species: Homo sapiens  
  
revision: 114982  
  
  
\*\*\*CURATION WARNINGS:\*\*\*  
  
  
\*\*\*CURATION TAG 1\*\*\*  
  
Name: Curation:NoInteractions  
Display Name: Fix interactions  
Description: 20 out of 114 lines are not properly connected.  
\*\*\*\*\*\*\*\*\*  
  
  
  
WARNING [bioservices:WikiPathways]: URL of the services contains a double //.Check your URL and remove trailing /  
  
  
Nodes: 124  
Edges: 3121  
SIF graph: WP2727\_edgeList.csv  
Graphml: WP2727\_graph.graphml  
\*\*\*\*\*\*\*\*\*\*  
  
  
\*\*\*PATHWAY INFORMATION\*\*\*  
  
Name: Curation:AnalysisCollection   
  
Display Name: Approved version   
  
id: WP2456  
  
url: https://www.wikipathways.org/index.php/Pathway:WP2456  
  
name: HIF1A and PPARG regulation of glycolysis  
  
species: Homo sapiens  
  
revision: 119185  
  
  
\*\*\*CURATION WARNINGS:\*\*\*  
  
  
\*\*\*CURATION TAG 1\*\*\*  
  
Name: Curation:NeedsReference  
Display Name: Add references  
Description:   
  
\*\*\*CURATION TAG 2\*\*\*  
  
Name: Curation:NeedsWork  
Display Name: Needs work  
Description:   
\*\*\*\*\*\*\*\*\*  
  
  
  
WARNING [bioservices:WikiPathways]: URL of the services contains a double //.Check your URL and remove trailing /  
  
  
Nodes: 19  
Edges: 23  
SIF graph: WP2456\_edgeList.csv  
Graphml: WP2456\_graph.graphml  
\*\*\*\*\*\*\*\*\*\*