Open PHACTS
Open Pharmacological Space

Drug discovery FAQs: workflows for answering multidomain drug discovery questions

Daniela Digles
University of Vienna
support@openphacts.org

4th KNIME Cheminformatics Workshop

650 JAHRE universität wien

Open PHACTS
Open Pharmacological Space

Data collection

NCBI Resources How To
PubMed.gov
US National Library of Medicine
National Institutes of Health
PubMed

SciFinder®
The choice for chemistry research.™

EMBL-EBI
Service
ChEMBL
ChEMBL
Downloads
UniChem

DrugBank
Browse Search Downloads About Help Contact Us
DrugBank
DrugBank Version 4.3
unique
matics
iled

EMBL-EBI
ChEBI
Open PHACTS Explorer
Enter search terms
Search
Search for stru
Home Advanced Search Bro

Open PHACTS

Open Pharmacological Space

DisGeNET

WikiPathways
Pathways for the People

GeneOntology
Unifying Biology

UniProt

DrugBank
Open Data Drug & Drug Target Database

ConceptWiki

ChEBI

neXtprot BETA

ChEMBL

ChemSpider
Search and share chemistry

Open PHACTS

Physicochemical data	Identifiers	Pharmacological data
Molecular weight & formula	Synonyms	Activity type, value, concentration
H-Bond acceptors / donors	SMILES	Assay description
Polar surface area, AlogP	InChI / InChIkey	Target organism
Melting point	ChemSpider ID	Target name
...and more	...and more	...and more

Image provided by the Open PHACTS Foundation

Useful links

- Latest version of the API: https://dev.openphacts.org/docs/1.5
- Support portal: http://support.openphacts.org/
- Open PHACTS Explorer: http://explorer.openphacts.org/

- Example Workflows: http://www.myexperiment.org/groups/1125.html

Answering „scientific competency questions“

- 20 questions defined at the beginning of the project.
- Azzaoui K, Jacoby E, Senger S, *et al.* (2013) Drug Discov. Today 18: 843 – 852.
- Two main clusters:
  1.) Compound-target relationships
  2.) includes additional complexity of diseases, pathways, text-mining and patents.

- Example: Give me all oxidoreductase inhibitors active <100 nM in human and mouse.
- Many questions need a combination of queries to the Open PHACTS Platform.
- Workflows published: Chichester C, Digles D, Siebes R, *et al.* (2015) Drug Discov. Today 20: 399 – 405.

Cluster 1 use cases (Q1-Q11): Answers require mainly compound-target pharmacology data		
ID	Use case question	Sequence of API calls with filters used and link to the workflow
Q1	Give me all oxidoreductase inhibitors active <100 nM in human and mouse.	**Target Class Pharmacology** (*target_organism=Homo sapiens|Mus musculus; minEx-pChembl=7*) http://www.myexperiment.org/workflows/4504.html
Q2	For a given compound what is its predicted secondary pharmacology?	**Compound Information> Chemical Structure Search: Similarity> *Compound adverse events***
Q3	Given a target find me all actives against that target. Find/predict polypharmacology of actives.	**Target Pharmacology** (*minEx-pChembl=5*) > **Compound Pharmacology** (*minEx-pChembl=0*) http://www.myexperiment.org/workflows/4505.html
Q4	For a given interaction profile, give me similar compounds.	**Compound Information > Compound Information (Batch) > Chemical Structure Search: Similarity** (*searchOptions.Threshold=0.85*)> **Compound Information** http://www.myexperiment.org/workflows/4516.html
Q5	For molecules that contain substructure X, retrieve all bioactivity data in serine protease assays.	**Chemical Structure Search: Substructure> Compound Pharmacology**, **Target Class Members** http://www.myexperiment.org/workflows/4478.html
Q6	For a specific target family, retrieve all compounds in specific assays	**Target Class Pharmacology** http://www.myexperiment.org/workflows/4506.html
Q7	For a target, give me all active compounds	**Target Pharmacology** (*minEx-pChembl=5*)

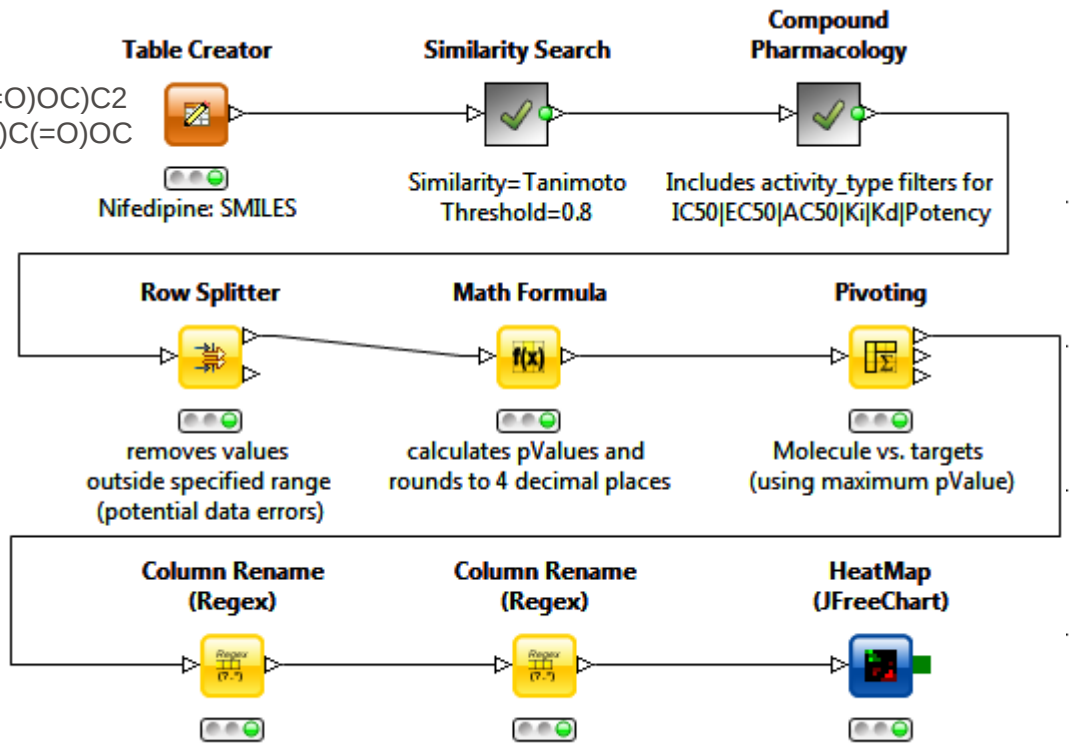
ID	Use case question	Sequence of API calls with filters used and link to the workflow
Q6	For a specific target family, retrieve all compounds in specific assays	**Target Class Pharmacology** http://www.myexperiment.org/workflows/4506.html
Q7	For a target, give me all active compounds with the relevant assay data.	**Target Pharmacology** (*minEx-pChembl=5*) http://www.myexperiment.org/workflows/4507.html
Q8	Identify all known protein-protein interaction inhibitors	**Target Class Pharmacology** (*target_type=ppi, minEx-pChembl=5*) http://www.myexperiment.org/workflows/4508.html
Q9	For a given compound, give me the interaction profile with targets.	**Compound Pharmacology** (*activity_type=IC50|EC50|AC50|Ki|Kd|Potency)* http://www.myexperiment.org/workflows/4509.html
Q10	For a given compound, summarize all similar compounds and their activities.	**Chemical Structure Search: Similarity (***searchOptions.SimilarityType=0; searchOptions.Threshold=0.80)***> Compound Pharmacology** (*activity_type=IC50|EC50|AC50|Ki|Kd|Potency)* http://www.myexperiment.org/workflows/4510.html
Q11	Retrieve all data for a given list of compounds depicted by their chemical structure (SMILES) with options to match stereochemistry.	**Chemical Structure Search: Exact** (*searchOptions.MatchType=2*)**> Compound Pharmacology, Compound Information**, **Compound Classifications** *(tree=chebi)* http://www.myexperiment.org/workflows/4511.html

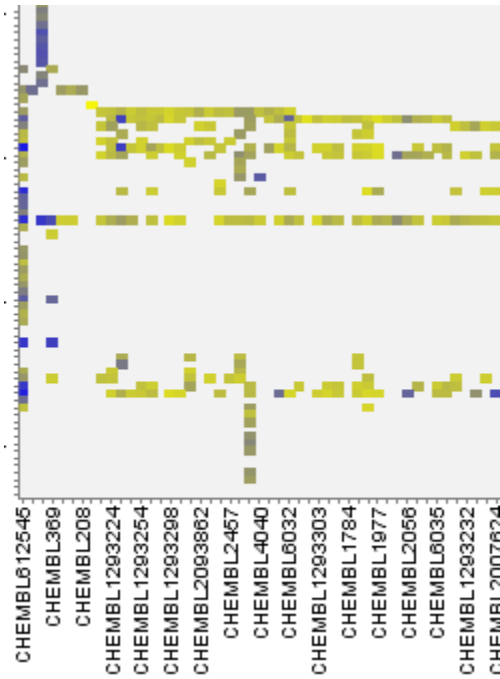
Cluster 2 use cases (Q12-Q20): Answers requiring pharmacology plus disease, pathway, and text mining data		
Q12	For a given compound, which of its targets have been patented in the context of a disease?	**Compound Pharmacology> *PatentsCalls* >Disease for Target**
Q13	For disease X, which targets have ligands in different stages of the development process with publications/patents describing these compounds?	**Targets for Disease> Target Pharmacology** (*minEx-pChembl=5*), **Target Information**>***Patents calls***
Q14	Target druggability: compounds directed against target X have what indications? Which new targets have appeared recently in the patent literature for a disease?	**Target pharmacology** (*minEx-pChembl=5*) >***Indications for Compounds*> *Patent calls*** **>Disease for Targets**
Q15	a) Which chemical series have been shown to be active against target X? b) Which new targets have been associated with disease Y? c) Which companies are working on target X or disease Y?	**a) Classification of Compounds for Target** (*minEx-pChembl=5*) **b) Associations for Disease** c) Competitive Intelligence data not available http://www.myexperiment.org/workflows/4512.html
Q16	Targets in Parkinson's disease or Alzheimer's disease are activated by which compounds?	**Target for Disease> Target Pharmacology** (*minEx-pChembl=5*) http://www.myexperiment.org/workflows/4513.html
Q17	For my specific target, which active compounds have been reported in the literature?	**Target Pharmacology** (*minEx-pChembl=5*) http://www.myexperiment.org/workflows/4507.html
Q18	For pathway X, find compounds that agonize targets assayed in only functional assays with potency <1 µM.	**Pathway Information: Get Targets>Target Pharmacology** (*activity_type=Potency, max-activity_value=1000, activity_unit=nanomolar*) http://www.myexperiment.org/workflows/4514.html
Q19	For the targets in a given pathway, retrieve the compounds that are active with more than one target.	**Pathway Information: Get Targets>Target Pharmacology** (*minEx-pChembl=5*) http://www.myexperiment.org/workflows/4515.html
Q20	For a given disease, retrieve all targets in the pathway and all active compounds hitting them.	**Targets for Disease> Target Pharmacology** (*minEx-pChembl=5*) http://www.myexperiment.org/workflows/4513.html

Example workflow

- Q10: For a given compound, summarize all similar compounds and their activities

CC1=C(C(C(=C(N1)C)C(=O)OC)C2=CC=CC=C2[N+](=O)[O-])C(=O)OC





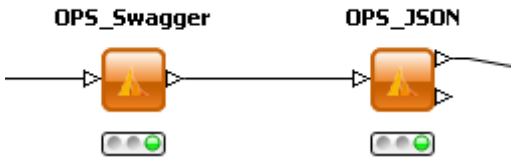
OPS-Knime nodes

- Created by Ronald Siebes, VU Amsterdam.
- No predefined set of nodes for each API call.

- OPS_Swagger:
  - creates the API call
  - Swagger file is used to automatically provide available API calls and parameters

- OPS_JSON (deprecated):
  - executes the API call
  - transforms the output into a flattened spreadsheet format

- available from https://github.com/openphacts/OPS-Knime



Swagger

- Structured format for the generation of API documentation. (https://helloreverb.com/developers/swagger)
- https://raw.githubusercontent.com/openphacts/OPS_LinkedDataApi/1.5.0/api-config-files/swagger.json


```
{
  "basePath": "https://beta.openphacts.org/1.3",
  "apiVersion": "v1.3",
  "apis": [
    {
"path": "/compound" ,
      "operations": [
        {
          "httpMethod": "GET",
          "summary":  "Compound Information" ,
          "description":  "Information about a single compound.
   ...

          "group": "Compound" ,
          "parameters": [
            {
              "name":  "uri" ,
              "description":  "A compound URI. e.g.: http://www.conceptwiki.org/concept/38932552-111f-4a4e-a46a-4ed1c
              "dataType": "string",
              "required": true,
              "paramType": "query"
            },
```

OPS_Swagger details

- Knime node where the user provides a url to a Swagger file (default: Open PHACTS API, v1.4)
- File is parsed and provides a list of the available API calls.
- Parameters tab is updated to the available parameters.
- Parameters can be set in the parameters tab or in the input table.
- Output of the node is an executable API call.

Open PHACTS
Open Pharmacological Space

https://dev.openphacts.org/docs/1.5

OpenPHACTS API v1.5

Compound Information /compound GET

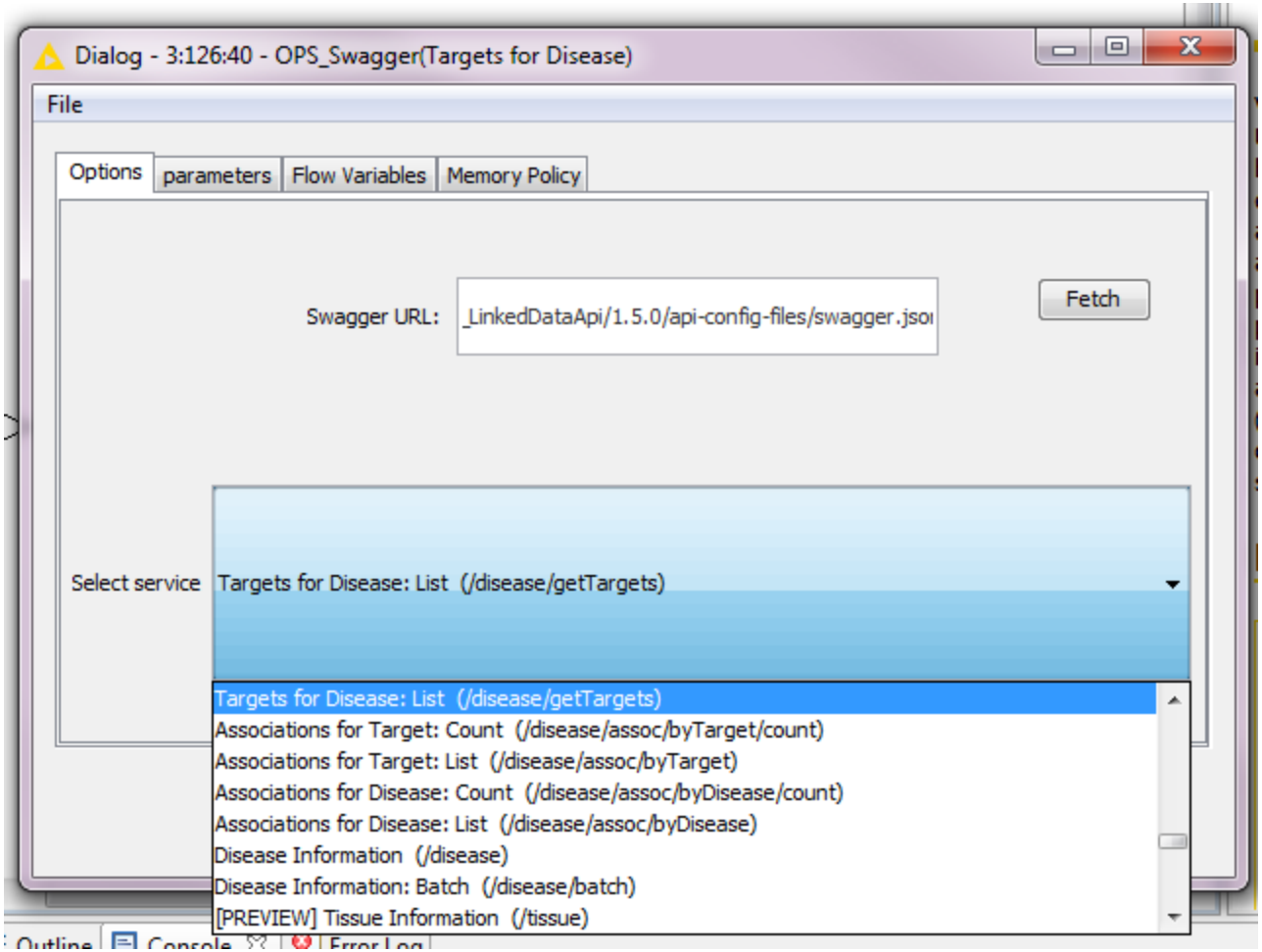
Description

Returns information about a single compound including (but not limited to): molecular weight, biotransformation, protein binding and toxicity.

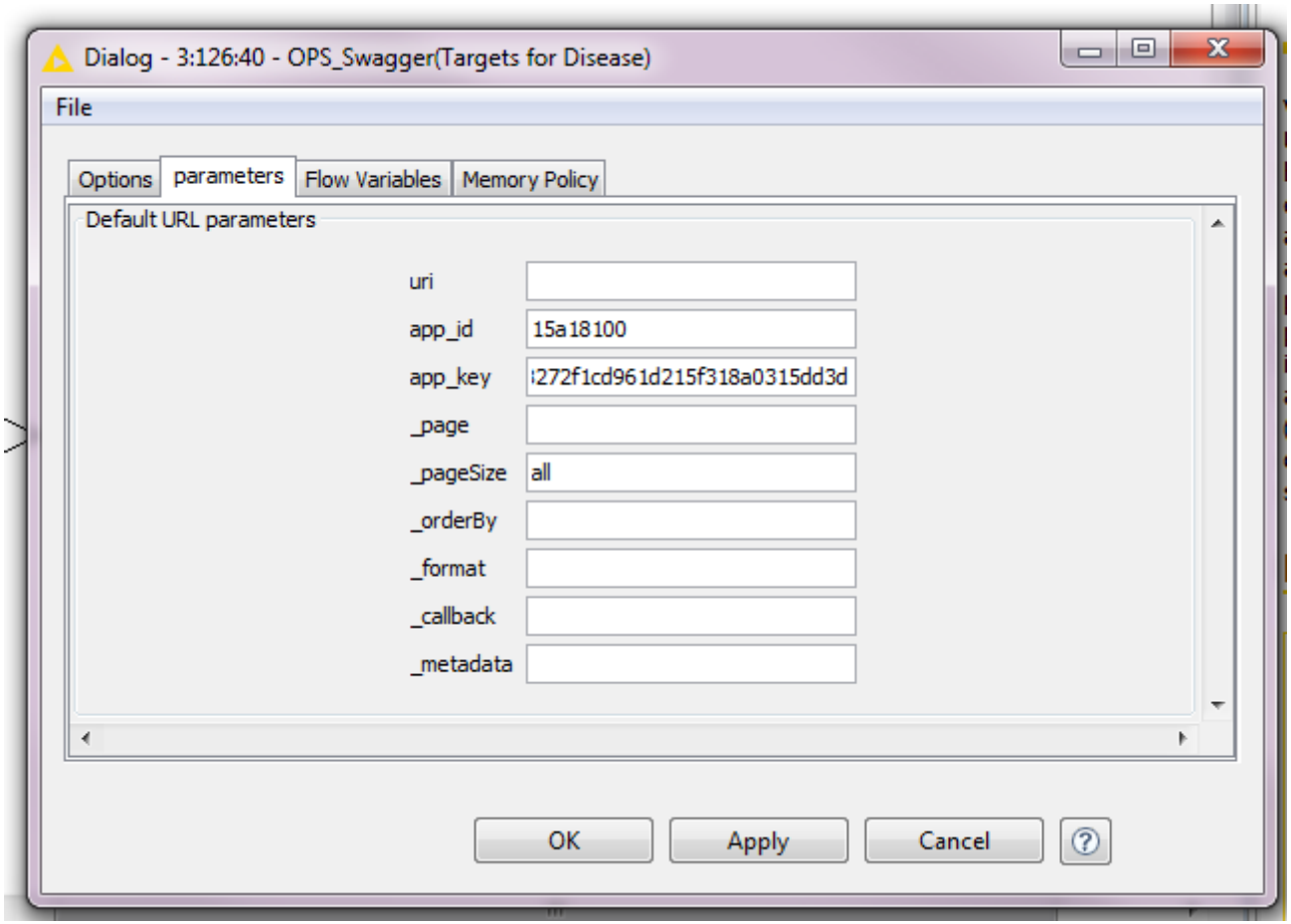
...

PARAMETER	VALUE	DESCRIPTION
uri	(required)	A compound URI. e.g.: http://www.conceptwiki.org/concept/38932552-111f-4a4e-a46a-4ed1d7bdf9d5
app_id		Your access application id
app_key		Your access application key
_format		The desired result format.

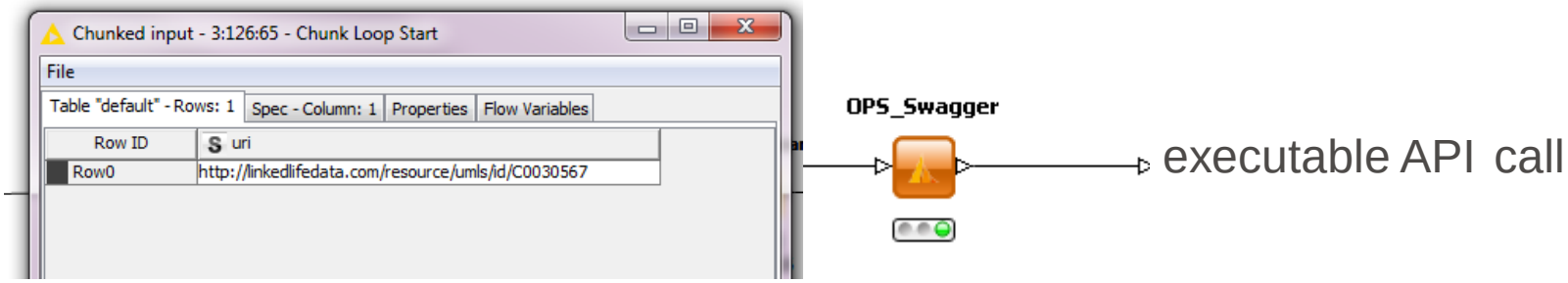
OPS_Swagger details



OPS_Swagger details

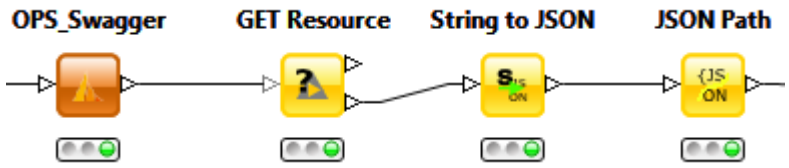


OPS_Swagger details



Retrieving and parsing the results

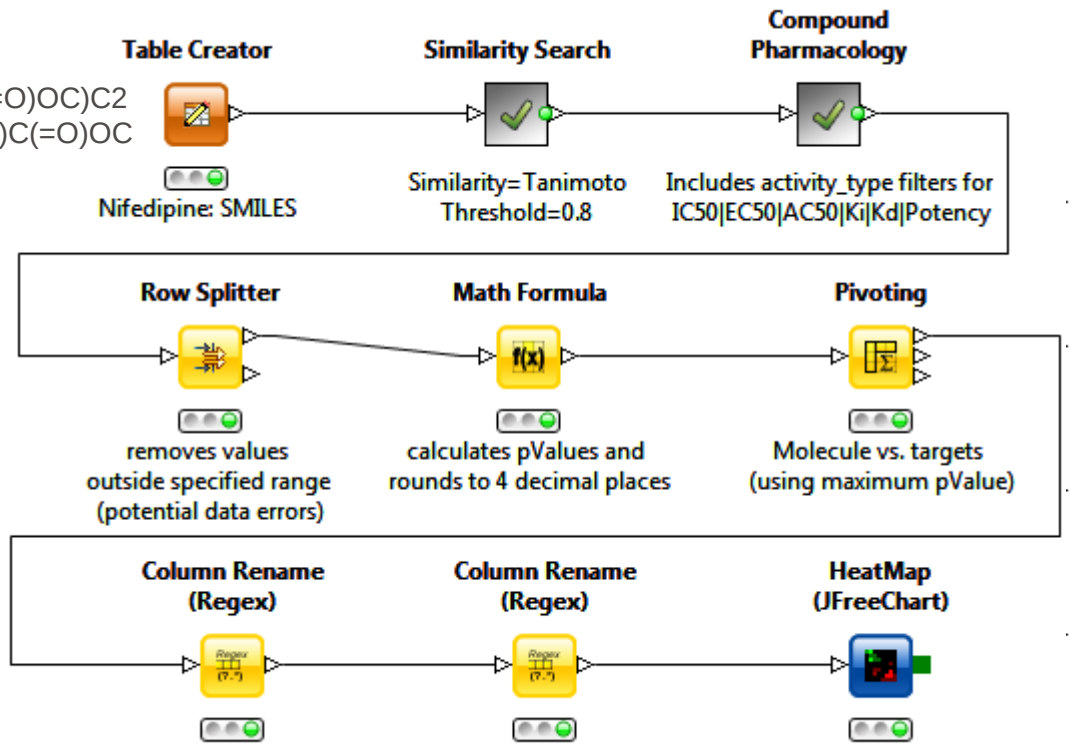
- Either use OPS_Json (deprecated) or the REST and JSON nodes available as addin from KNIME.
- GET Resource: retrieves the actual data from the server. Configure the node to use the column url as input. Response representation cell type: Autodetection.
- String to JSON: transforms the result to a JSON column type.
- JSON Path: allows the individual selection of the data which is transformed into a tabular structure.

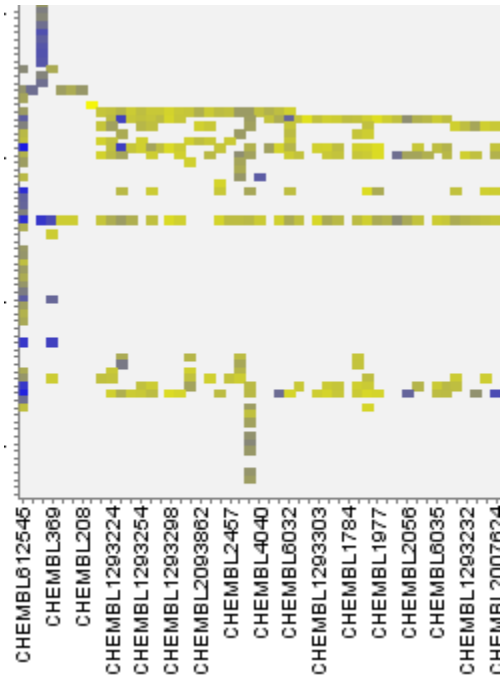


Example workflow

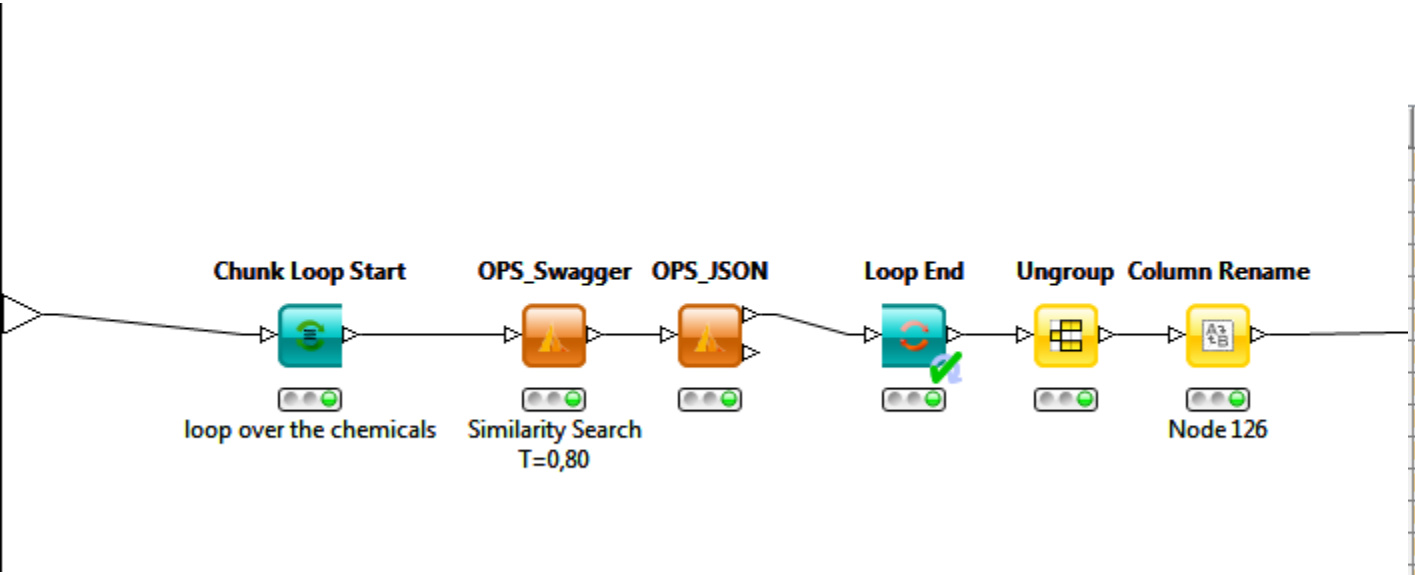
- Q10: For a given compound, summarize all similar compounds and their activities

CC1=C(C(C(=C(N1)C)C(=O)OC)C2=CC=CC=C2[N+](=O)[O-])C(=O)OC



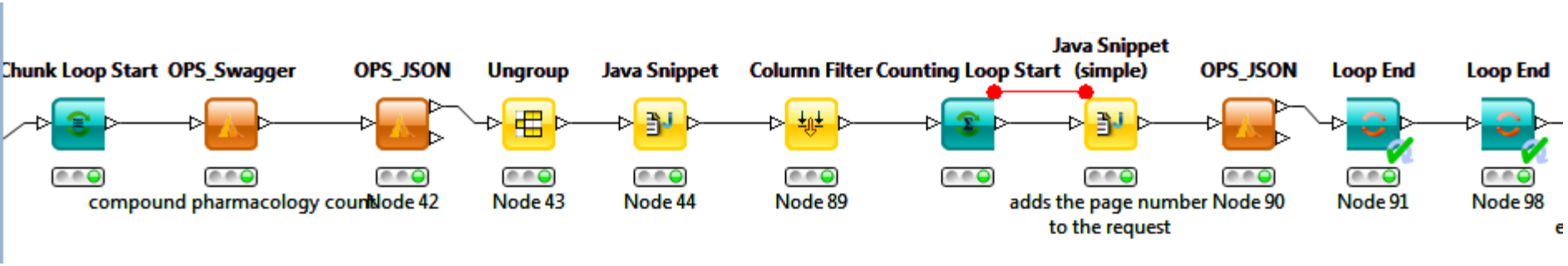


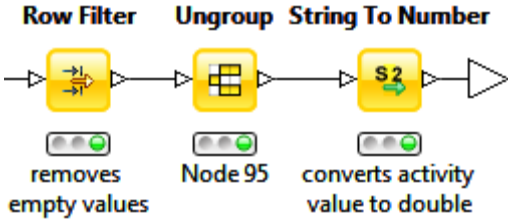
Similarity Search Metanode



S uri
http://ops.rsc.org/OPS980348
http://ops.rsc.org/OPS1864091
http://ops.rsc.org/OPS1299748
http://ops.rsc.org/OPS1306882
http://ops.rsc.org/OPS1188404
http://ops.rsc.org/OPS16172
http://ops.rsc.org/OPS1872614
http://ops.rsc.org/OPS1298823
http://ops.rsc.org/OPS1322374
http://ops.rsc.org/OPS95449
http://ops.rsc.org/OPS1864790
http://ops.rsc.org/OPS1340181
http://ops.rsc.org/OPS1000948
http://ops.rsc.org/OPS89613
http://ops.rsc.org/OPS1296791
http://ops.rsc.org/OPS1261414
http://ops.rsc.org/OPS1345323
http://ops.rsc.org/OPS57223

Compound Pharmacology Metanode





Difficulties

- Definition of „active“ compounds
- Wish of retrieving „all“ results
- Requested data not available from open access data providers (e.g. Which companies are working on target X or disease Y?)
- Data not yet available in the Open PHACTS Discovery Platform (Patents; upstream/downstream information from Pathways)
- Long execution times

Current activities

- Improvement of the available Open PHACTS Knime nodes
- Improvement of the existing workflows:
  - Replacement of OPS_JSON node with existing nodes available from Knime
  - Update of the API to the latest available version
- Integration of new data sources (e.g. patent data)

Open PHACTS
Open Pharmacological Space

Acknowledgements

- Pharmacoinformatics research group, University of Vienna
  - Gerhard F. Ecker
  - Barbara Zdrazil
  - Lars Richter
- Open PHACTS – KNIME
  - Ronald Siebes, VU Amsterdam
  - Christine Chichester, SIB
  - Evan Tzanis, QMUL

imi innovative medicines initiative

efpia

650 JAHRE universität wien