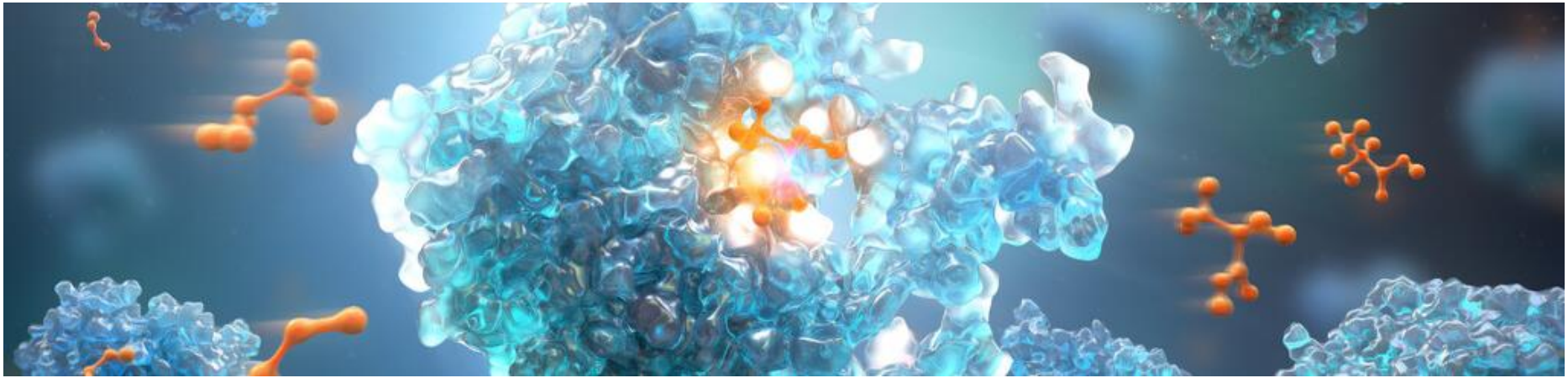
A modern tool for integrating cheminformatics data science in Spotfire

Lars Brive, Justin Morley, Maxime Guitet
PerkinElmer EMEA Nexus 2019



Acknowledgements

Wolfgang Klute

Johan Ulander

Nick Tomkinson

Eric Le Roux

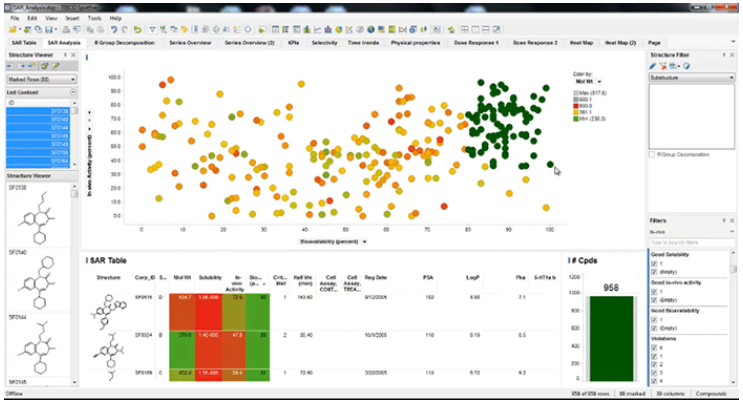
Claire Wallon

Benjamin Pannetier

Michael Bodnarchuk
Peter Brandt
Sameer Kawatkar
Susanne Winiwarter
Kun Song
Magnus Polla
Ekaterina Ratkova
SpotOn team

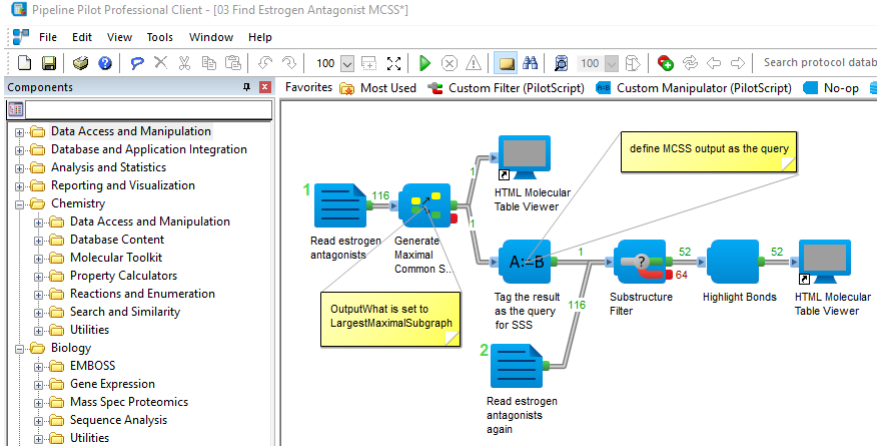
Spotfire ⇌ Biovia Pipeline Pilot ⇌ Databases, Calculation services, AI tools, ….

Spotfire:
Analysis, visualisations, PerkinElmer Lead Discovery

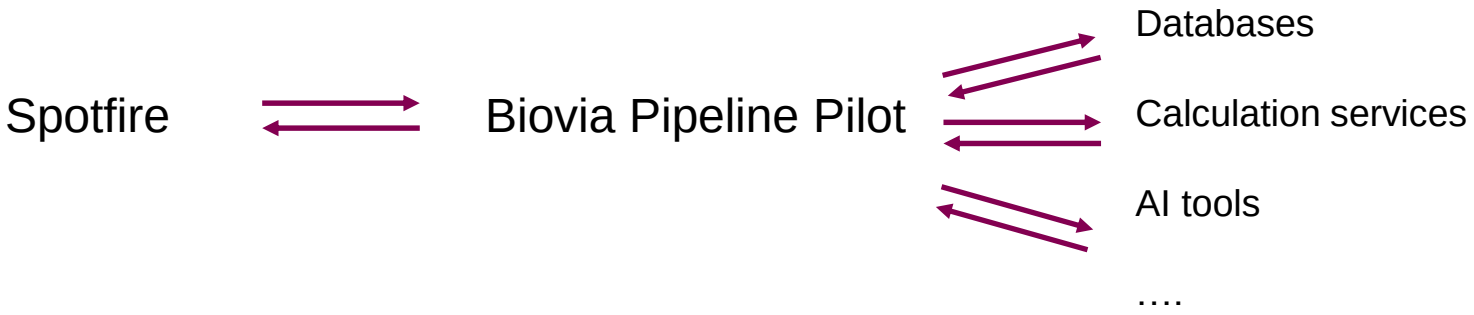


PerkinElmer.com

Biovia Pipeline Pilot: Data processing



Pipeline Pilot example protocol



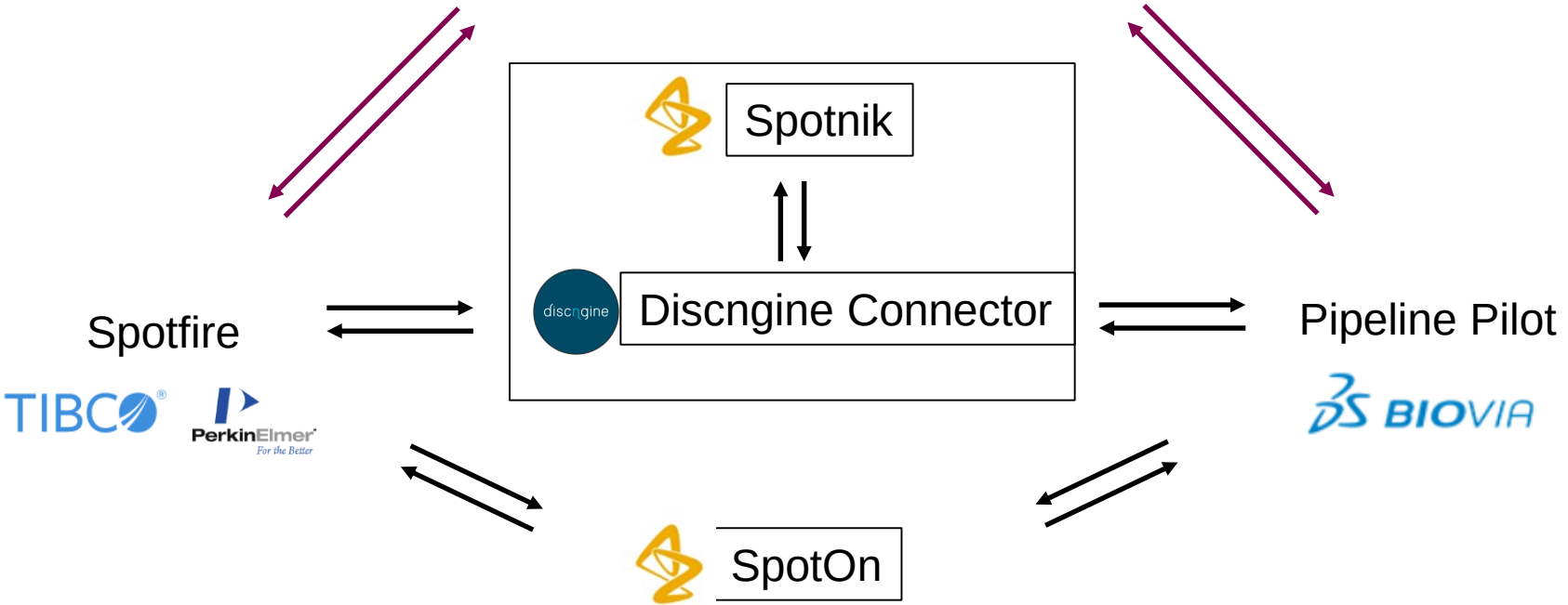
Requirements

- Interactive and flexible analysis of data
- Complex queries made easy.
- Default visualisations.
- Range of user expectations: One green button or full control over parameters.
- Allow advanced users to implement their own protocols.
- Modular design
- Rapid implementation. Global usage.

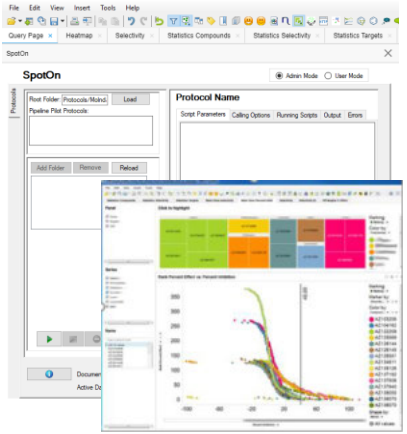
”Are these compounds available in our compound collection”

”Given a list of compound names, extract the bioactivity data in test, comparison with internal and external data. Group by target type.”

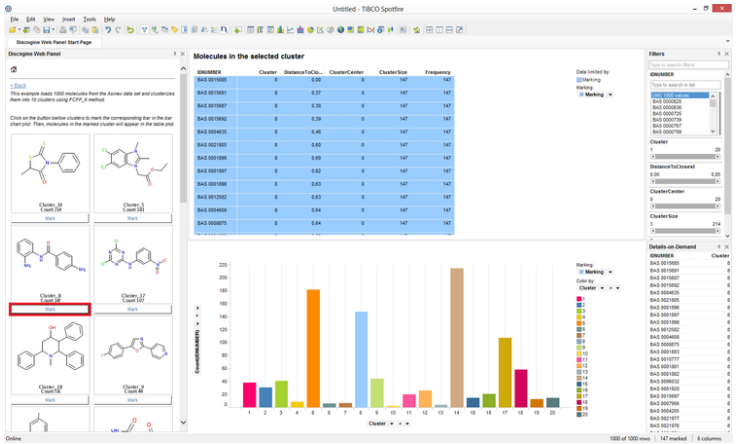
The Future


Spotnik
discngine
Discngine Connector
Spotfire
TIBCO
PerkinElmer
For the Better
Pipeline Pilot
BIOVIA
SpotOn

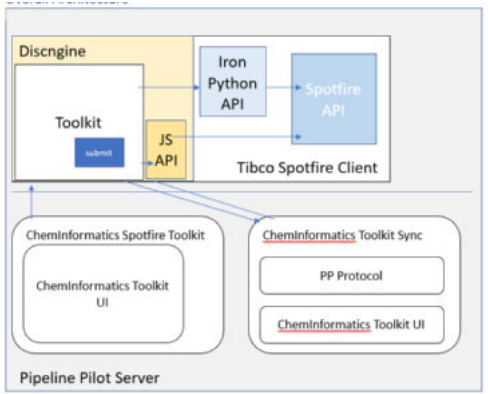
AZ SpotOn



Discngine Connector



AZ Spotnik



Justin Morley

Discngine: Run Pipeline Pilot protocols from Tibco Spotfire

Properties

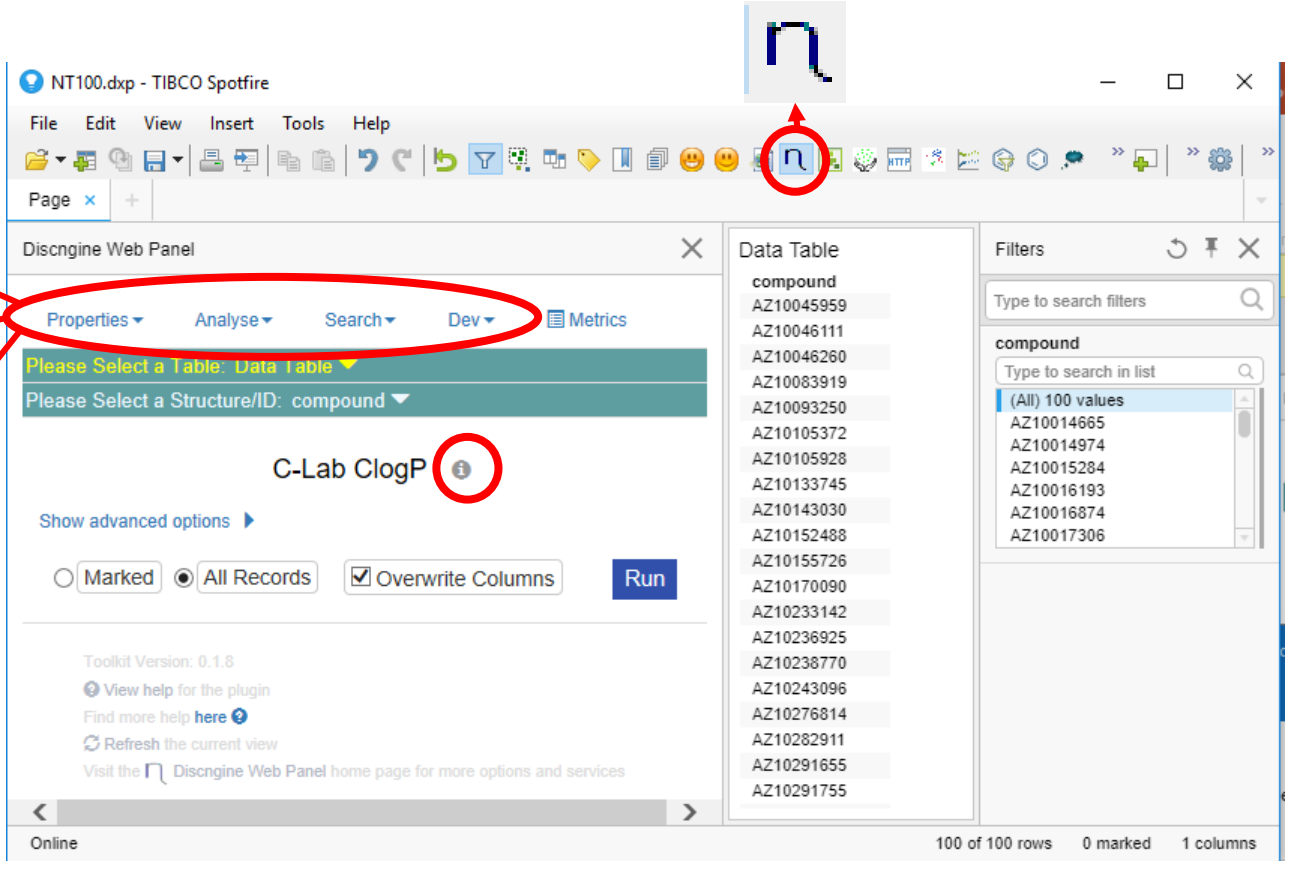
- AZlogD74
- ClogP
- structural alert
- Top 20
- Count bonds and rings
- LogBSF
- PAINS filter

Analyse

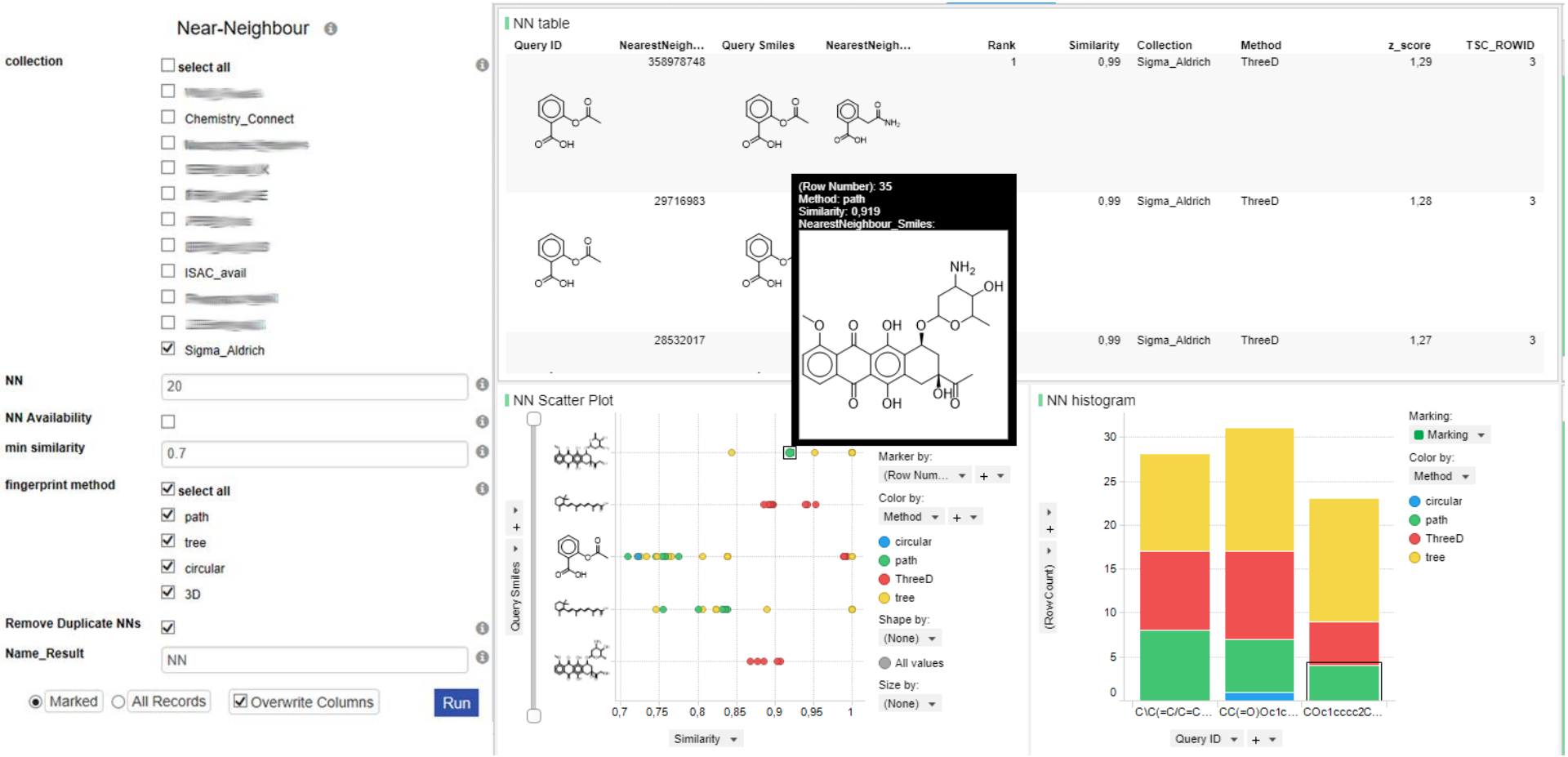
- Cluster
- CNS_MPO
- Diverse selection
- Similarity

Search

- IsCompoundAZ
- IsPublic
- Name to structure
- Structure to name



Advanced visualisations

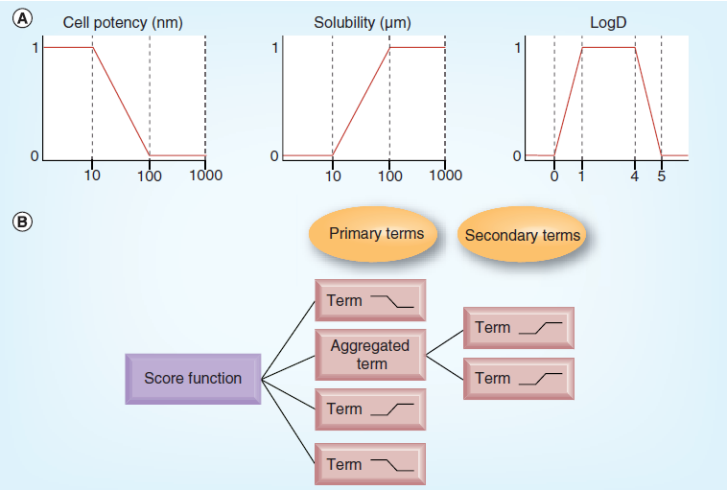


Support of AI in drug discovery

Generate large numbers of entirely novel compounds

Multi-parameter optimization (MPO) of desirability scores used for selection

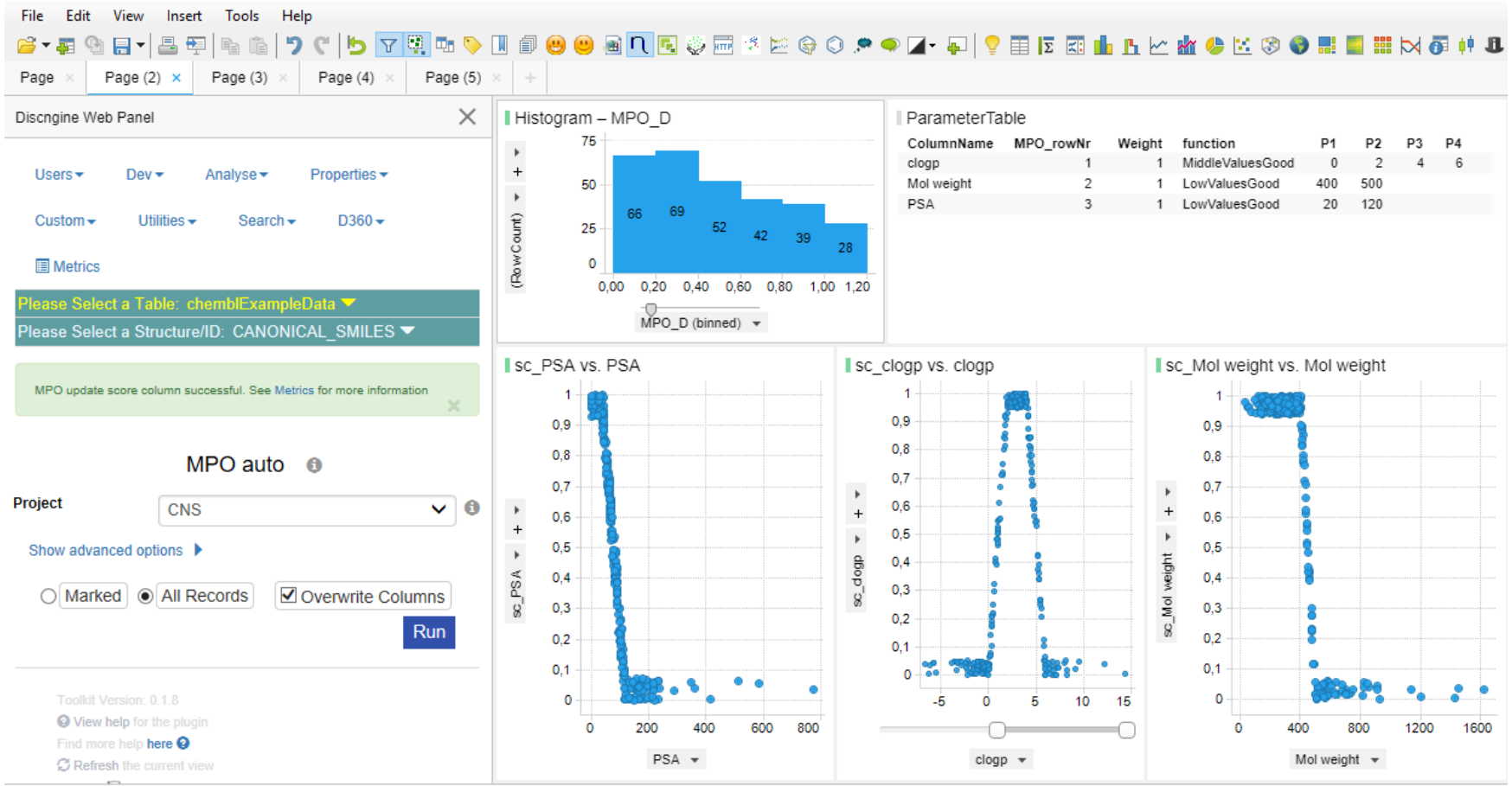
Example: MPO with default properties and parameters for a particular project



Nissink and Degorce (2013) Future Med. Chem. 5:753

ColumnName	Weight	function	P1	P2	P3	P4
clogp	1	MiddleValuesGood	0	2	4	6
Mol weight	1	LowValuesGood	400	500		
PSA	1	LowValuesGood	20	120		

CNS MPO; based on Wager et al (2010) ACS Chem. Neurosci. 1:435


File
Edit
View
Insert
Tools
Help
Page
Page (2)
Page (3)
Page (4)
Page (5)
Discngine Web Panel
Users
Dev
Analyse
Properties
Custom
Utilities
Search
D360
Metrics
Please Select a Table: chemblExampleData
Please Select a Structure/ID: CANONICAL_SMILES
MPO update score column successful. See Metrics for more information
MPO auto
Project
CNS
Show advanced options
Marked
All Records
Overwrite Columns
Run
Toolkit Version: 0.1.8
View help for the plugin
Find more help here
Refresh the current view
Histogram – MPO_D
(Row Count)
MPO_D (binned)
ParameterTable
ColumnName MPO_rowNr Weight function P1 P2 P3 P4
clogp 1 1 MiddleValuesGood 0 2 4 6
Mol weight 2 1 LowValuesGood 400 500
PSA 3 1 LowValuesGood 20 120
sc_PSA vs. PSA
sc_PSA
PSA
sc_clogp vs. clogp
sc_clogp
clogp
sc_Mol weight vs. Mol weight
sc_Mol weight
Mol weight
Online
296 of 1 474 rows
0 marked

File Edit View Insert Tools Help

Page | Page (2) | Page (3) | Page (4) | Page (5)

Discngine Web Panel

Users Dev Analyse Properties

Custom Utilities Search D360

Metrics

Please Select a Table: chemblExampleData

Please Select a Structure/ID: CANONICAL_SMILES

MPO update score column successful. See Metrics for more information

MPO auto

Project: CNS

Show advanced options

Marked | All Records | Overwrite Columns

Run

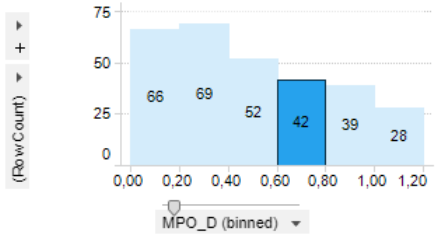
Toolkit Version: 0.1.8

View help for the plugin

Find more help here

Refresh the current view

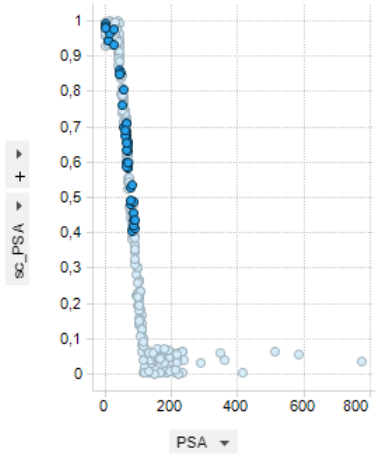
Histogram – MPO_D



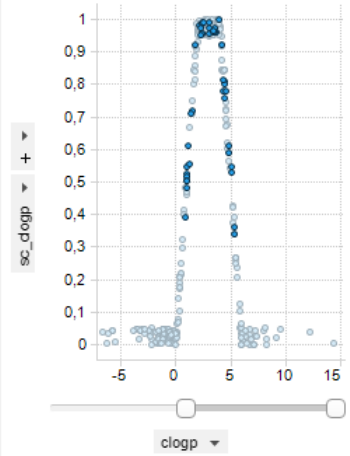
ParameterTable

ColumnName	MPO_rowNr	Weight	function	P1	P2	P3	P4
clogp	1	1	MiddleValuesGood	0	2	4	6
Mol weight	2	1	LowValuesGood	400	500		
PSA	3	1	LowValuesGood	20	120		

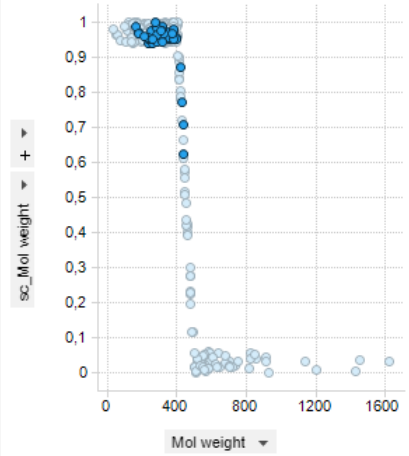
sc_PSA vs. PSA



sc_clogp vs. clogp

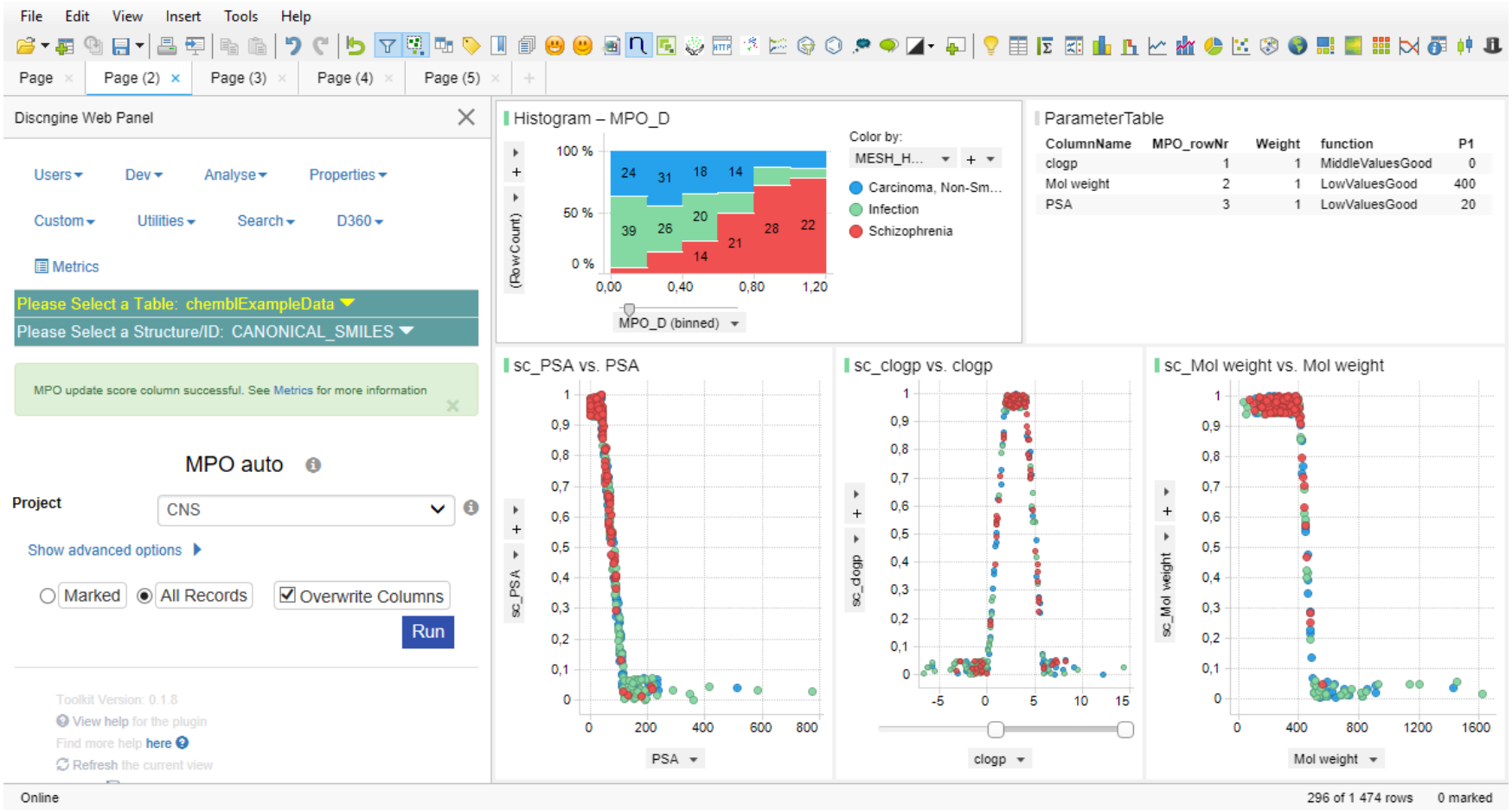


sc_Mol weight vs. Mol weight



Online

296 of 1 474 rows 42 marked


File
Edit
View
Insert
Tools
Help
Page
Page (2)
Page (3)
Page (4)
Page (5)
Discngine Web Panel
Users
Dev
Analyse
Properties
Custom
Utilities
Search
D360
Metrics
Please Select a Table: chemblExampleData
Please Select a Structure/ID: CANONICAL_SMILES
MPO update score column successful. See Metrics for more information
MPO auto
Project
CNS
Show advanced options
Marked
All Records
Overwrite Columns
Run
Toolkit Version: 0.1.8
View help for the plugin
Find more help here
Refresh the current view
Histogram – MPO_D
Color by:
MESH_H...
Carcinoma, Non-Sm...
Infection
Schizophrenia
100 %
50 %
0 %
(Row Count)
MPO_D (binned)
ParameterTable
ColumnName
MPO_rowNr
Weight
function
P1
clogp
1
1
MiddleValuesGood
0
Mol weight
2
1
LowValuesGood
400
PSA
3
1
LowValuesGood
20
sc_PSA vs. PSA
sc_PSA
PSA
sc_clogp vs. clogp
sc_clogp
clogp
sc_Mol weight vs. Mol weight
sc_Mol weight
Mol weight
Online
296 of 1 474 rows
0 marked

Summary

Pros:

Advanced chemistry workflows available from Spotfire

Advanced visualisations that gives users a quick start for analysis

Interface that works well for a range of users

Implementation of complex workflows doesn’t require programming skills

Good debugging options

Rapid implementation of new tools to scientists

Cons:

3D structure viewer

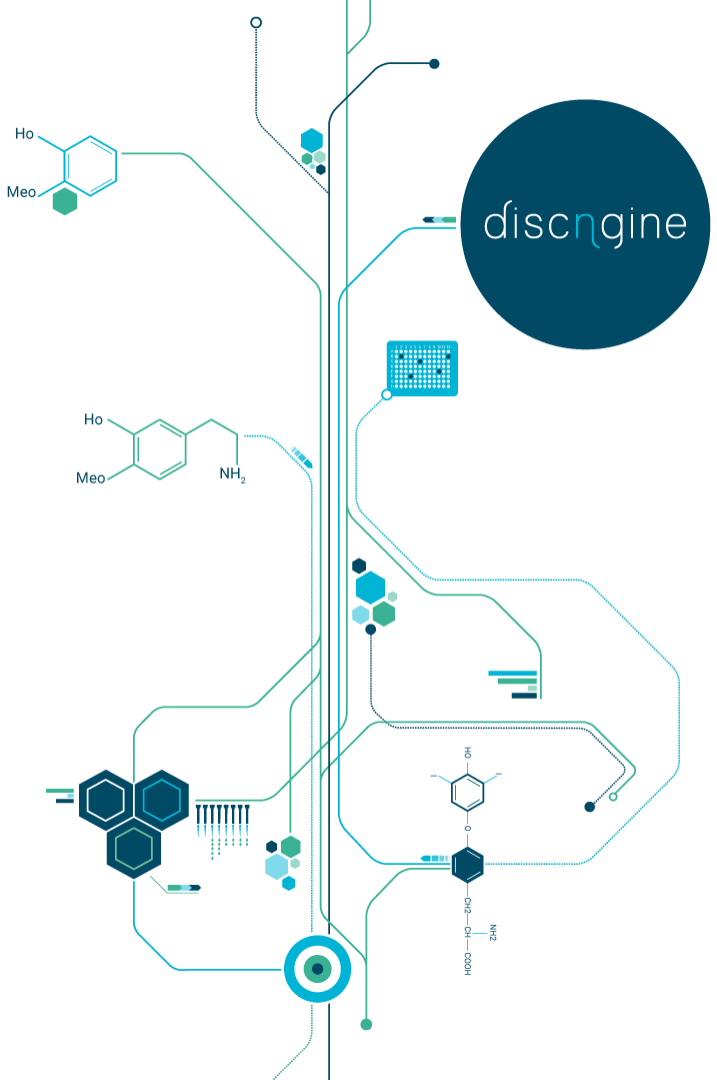
Data transfer to Pipeline Pilot limited to one table

Asynchronous jobs not straightforward

Framework needs local support, maintenance and upgrades

Confidentiality Notice

Spotfire Web Application for Pipeline Pilot



Discngine

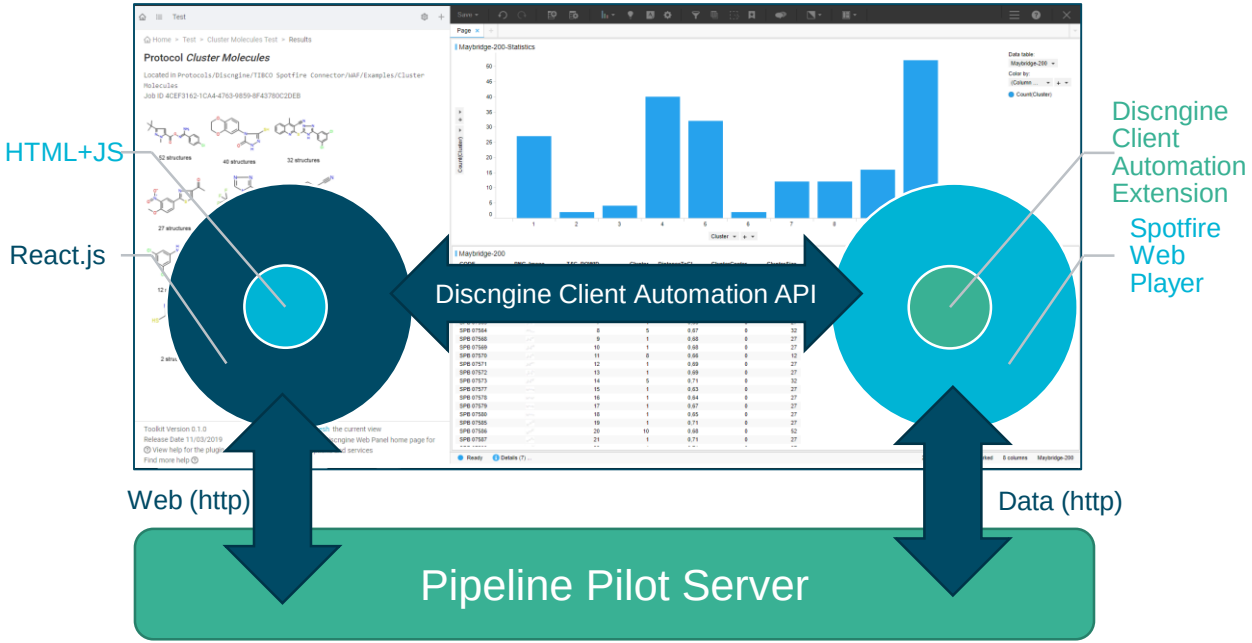
Connector

https://connector.discngine.com

Client Automation


```
spotfireDocumentEditor
   .addDataTable(sbdfFileURL)
   .addPage("Comparison Of Mean Ranges")
   .addBoxPlot(jsonConfig);
```

S.W.A.P.P: Spotfire Web Application for Pipeline Pilot



in Web Player *and* Analyst

SWAPP + Connector Client Automation

A new framework to streamline Pipeline Pilot protocol integration

- Low code

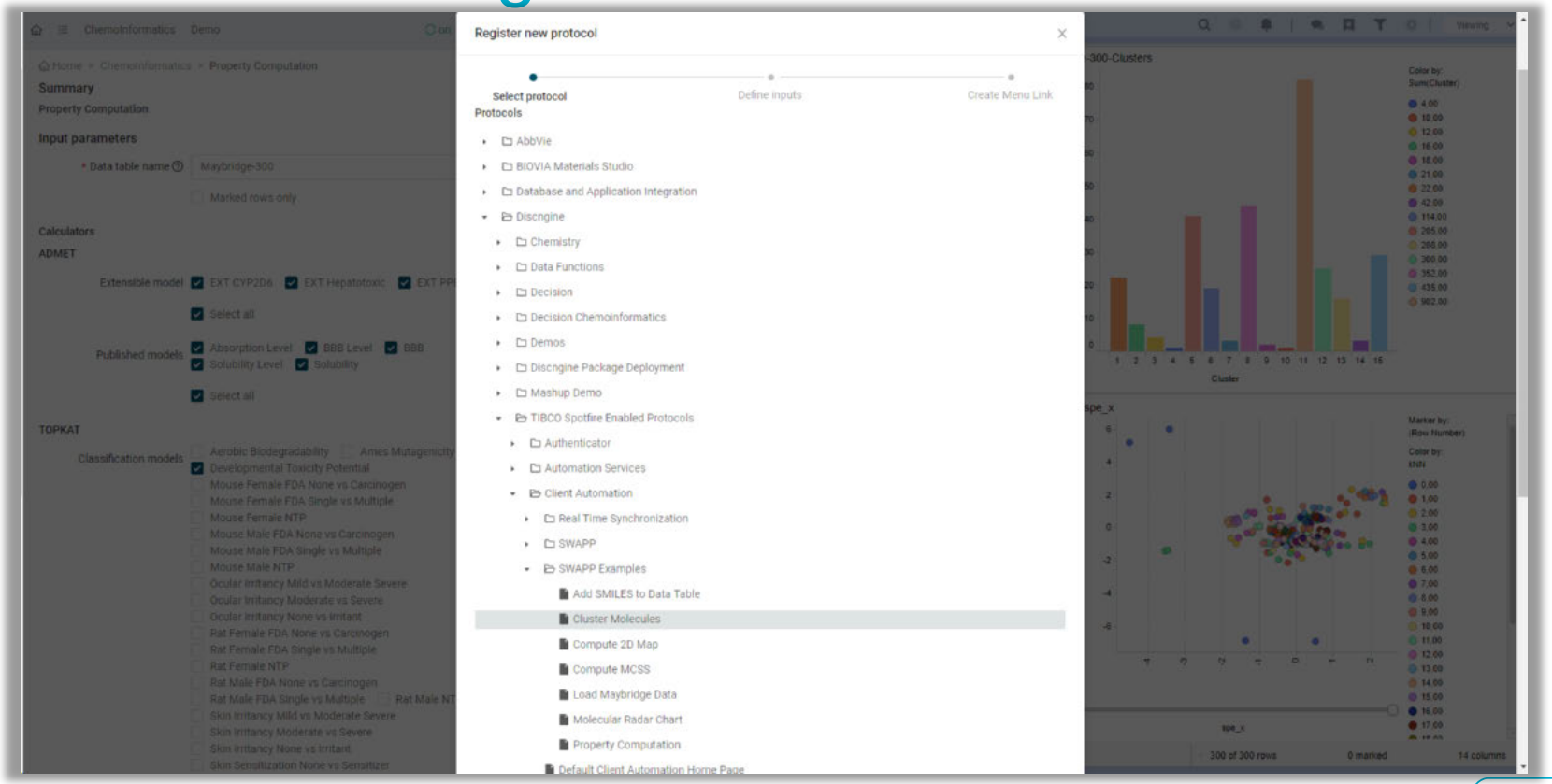
1. Declare the Pipeline Pilot protocol
2. Map the parameters to a web form (Wizard)
3. Define actions on the Spotfire document based on protocol output files

Features

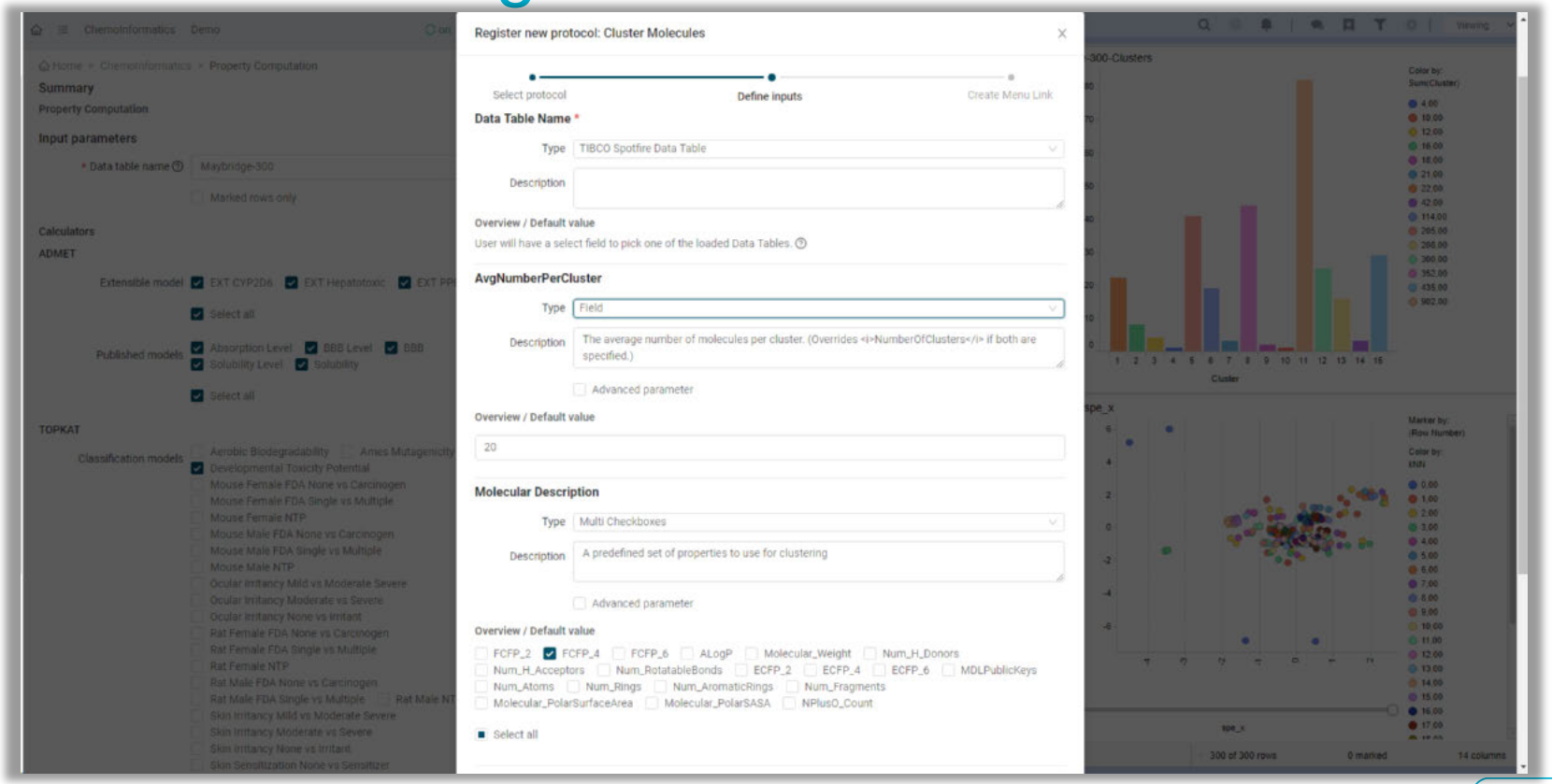
- Authentication
- Job Control (start, poll, stop) including asynchronous run
- Push/Pull Data from Web Service endpoint
- Allow injection of HTML/JS Snippets

> Control Spotfire document
> Display content

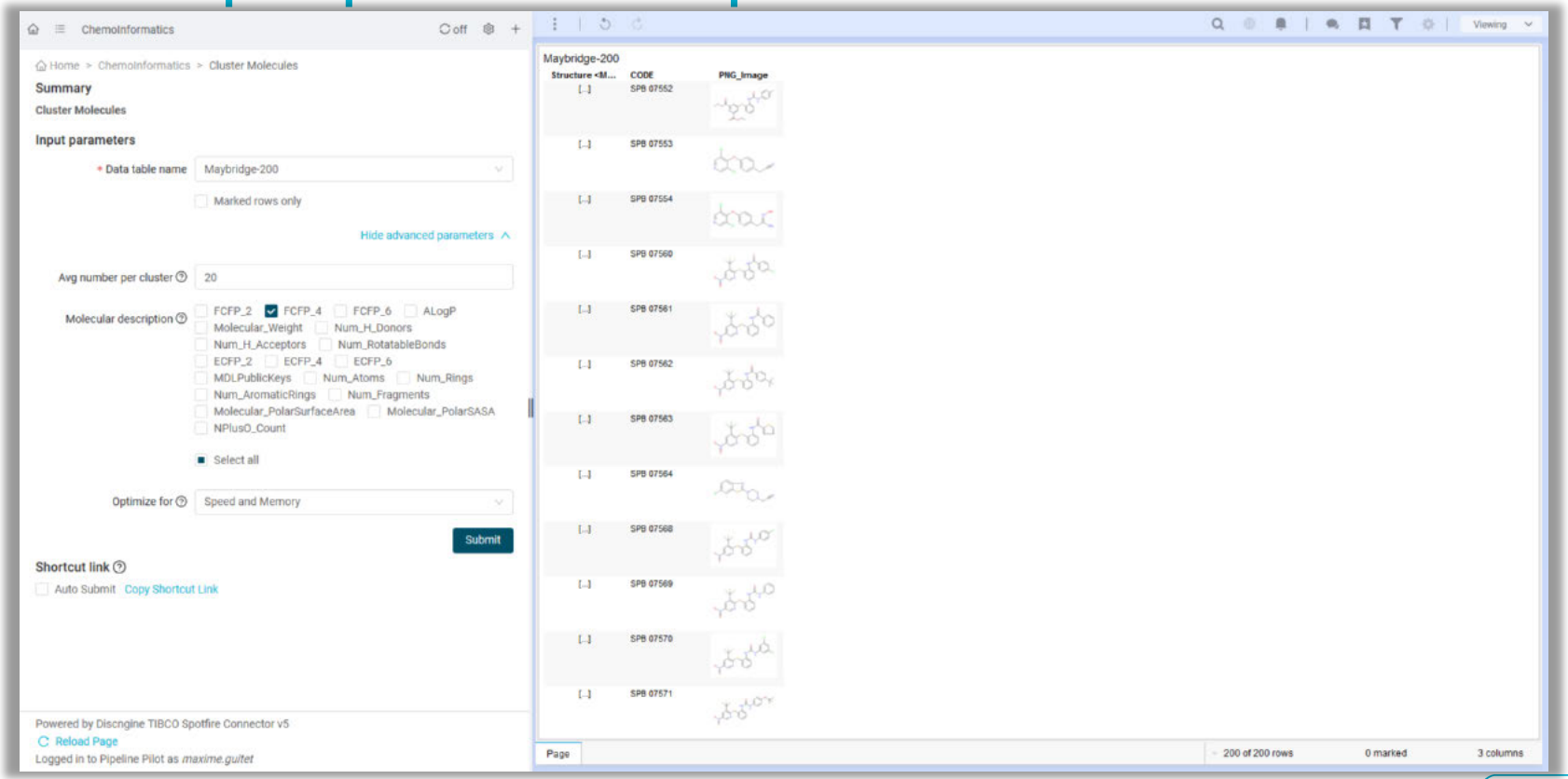
S.W.A.P.P: Register Protocol 1/2



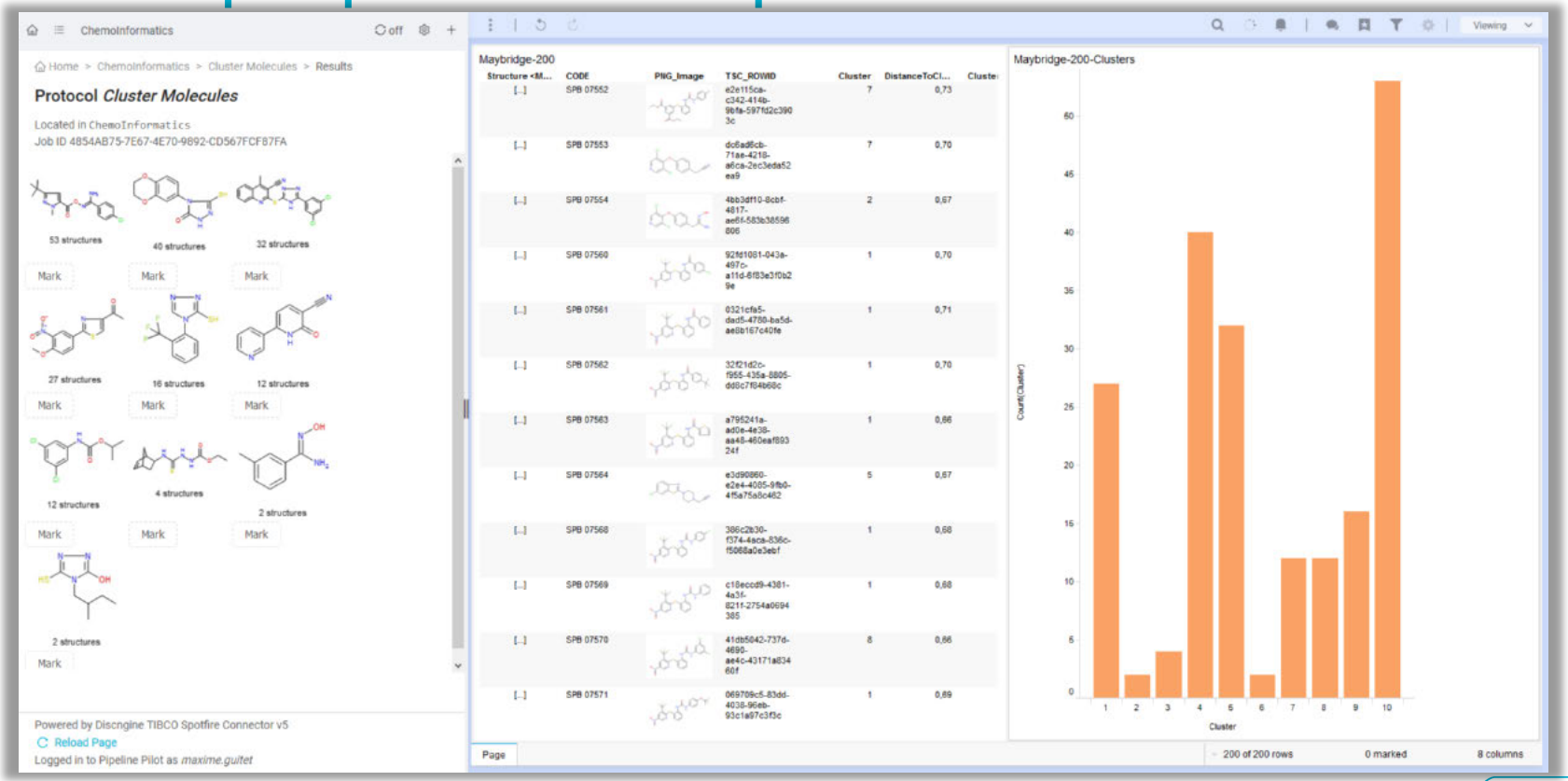
S.W.A.P.P: Register Protocol 2/2



Example protocol: intput



Example protocol: output

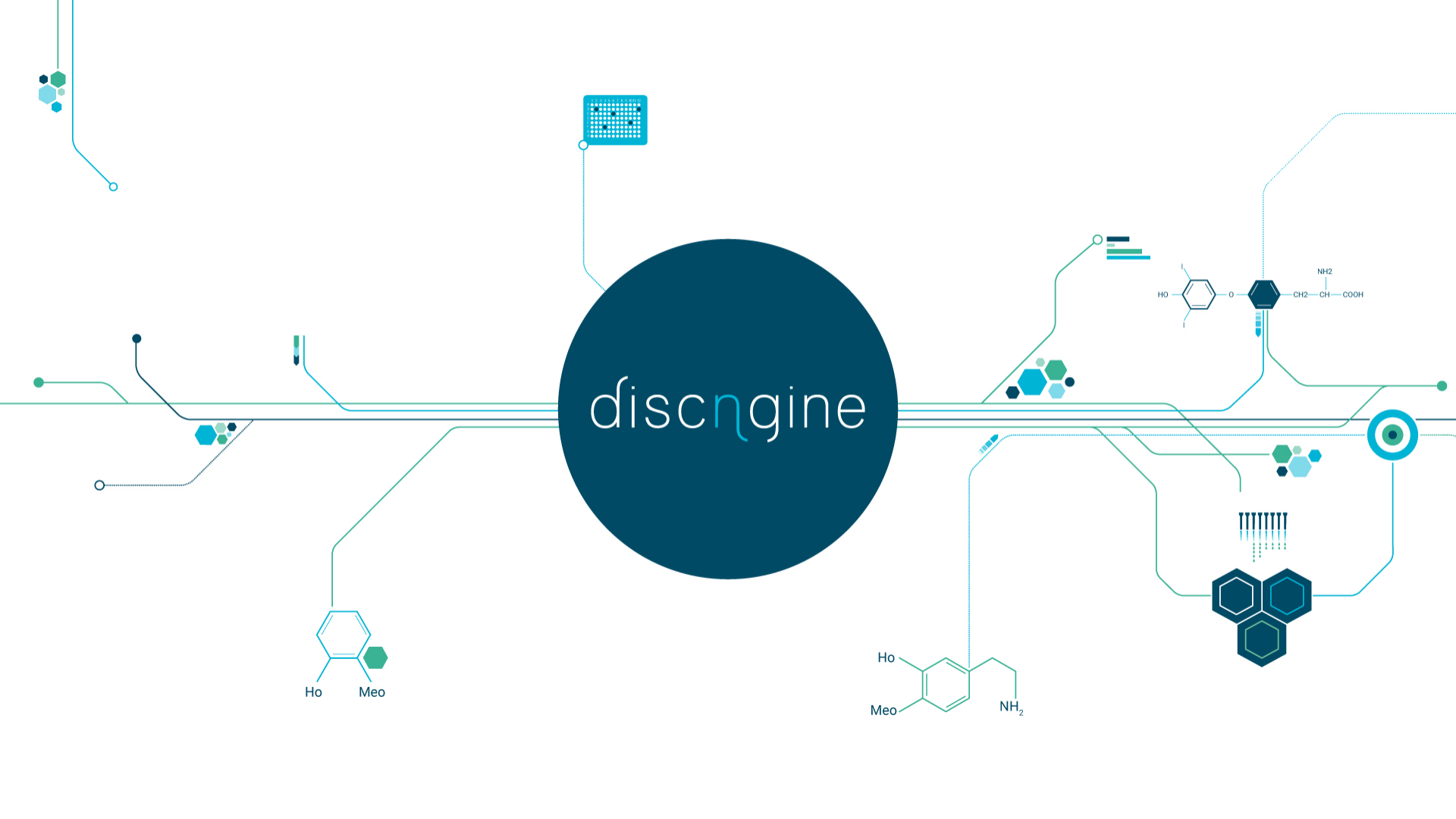


Future work

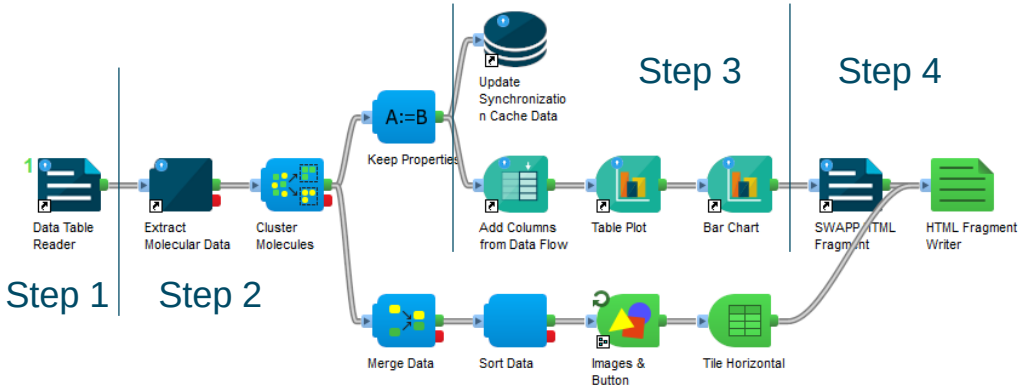
Extend to Knime, Talend

Implement more features

- Chain protocols
- Extend user permissions to control visibility of protocols
- Allow conditional linking of parameters


discngine
Ho
Meo
Ho
Meo
NH2
NH2
HO
O
CH2
CH
COOH

Existing application: SWAPP - Example



- Step 1: reading data from SBDF
- Step 2: data transformation
- Step 3: preparing API calls
- Step 4: combining API calls and creating HTML/JS file