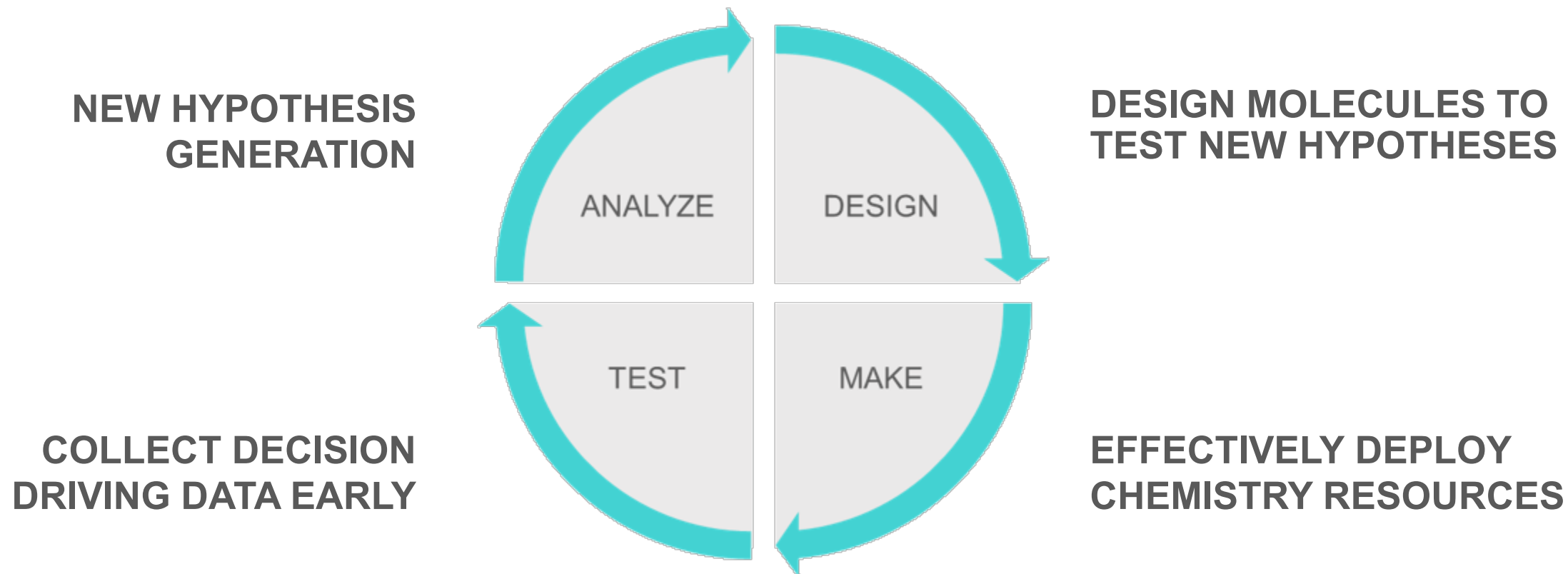


congruence^{TX}

Accelerating Drug Discovery

KNIME Spring Summit – April 15-17, 2024, Austin

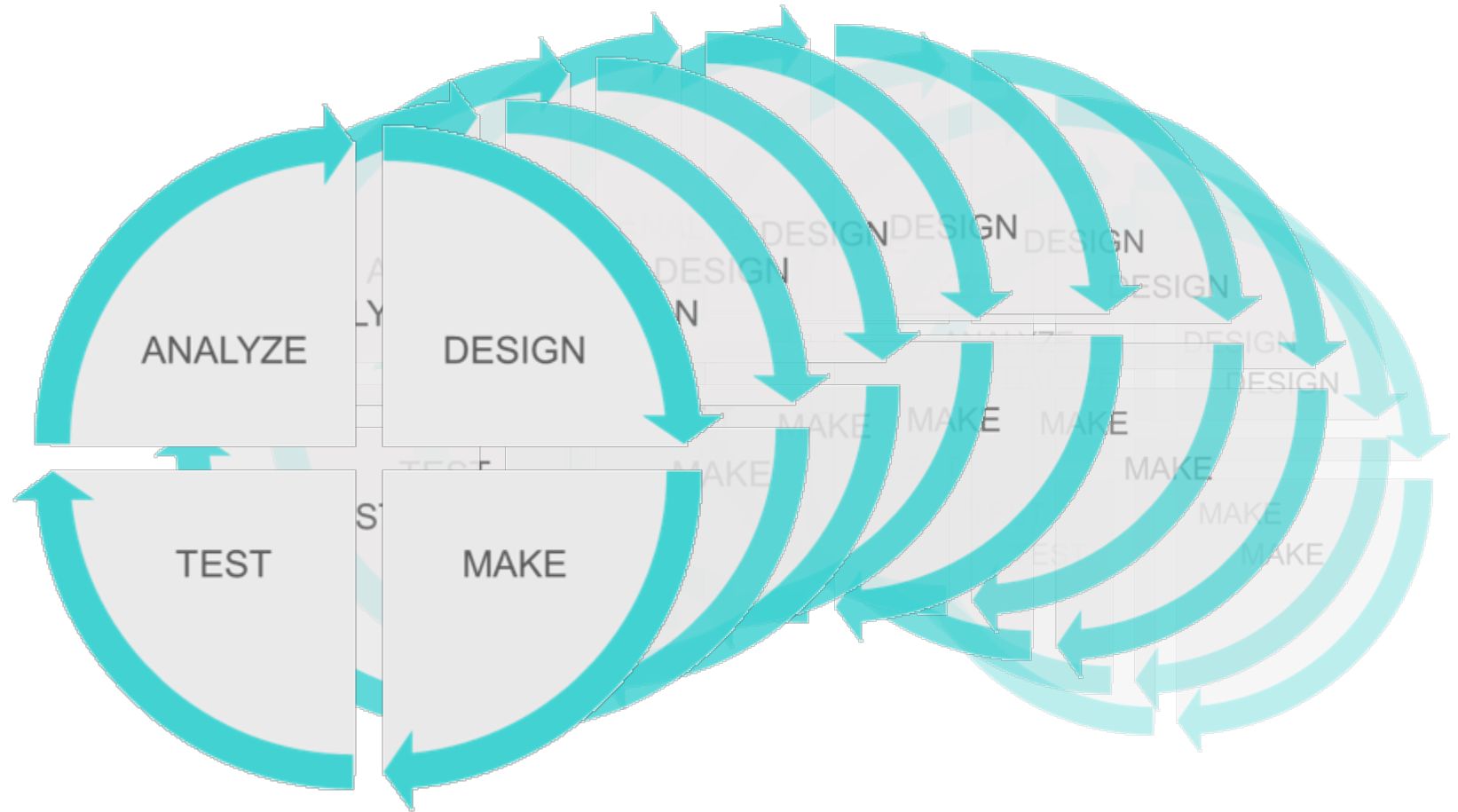
The central paradigm of lead optimization in drug discovery



Iteration through design cycles with efficiency is critical

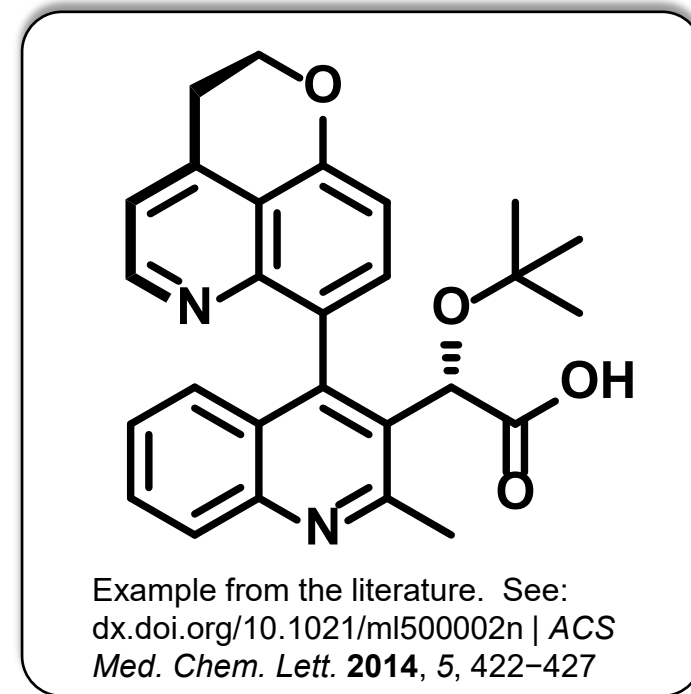
Optimization of design cycle time for efficient hypothesis testing

Minimize the number of iterations for rapid project progression



A key step in cheminformatics is the storage of chemical structures.

- A key challenge is representing chemical compounds in a way that **both humans and machines** can comprehend.
- Chemists use **2D and 3D visualizations** to represent chemical structures.
- Chemical structure representations may contain **explicit and implicit** information
- **Line notations** benefit both computational processes and human chemists in cheminformatics tasks.



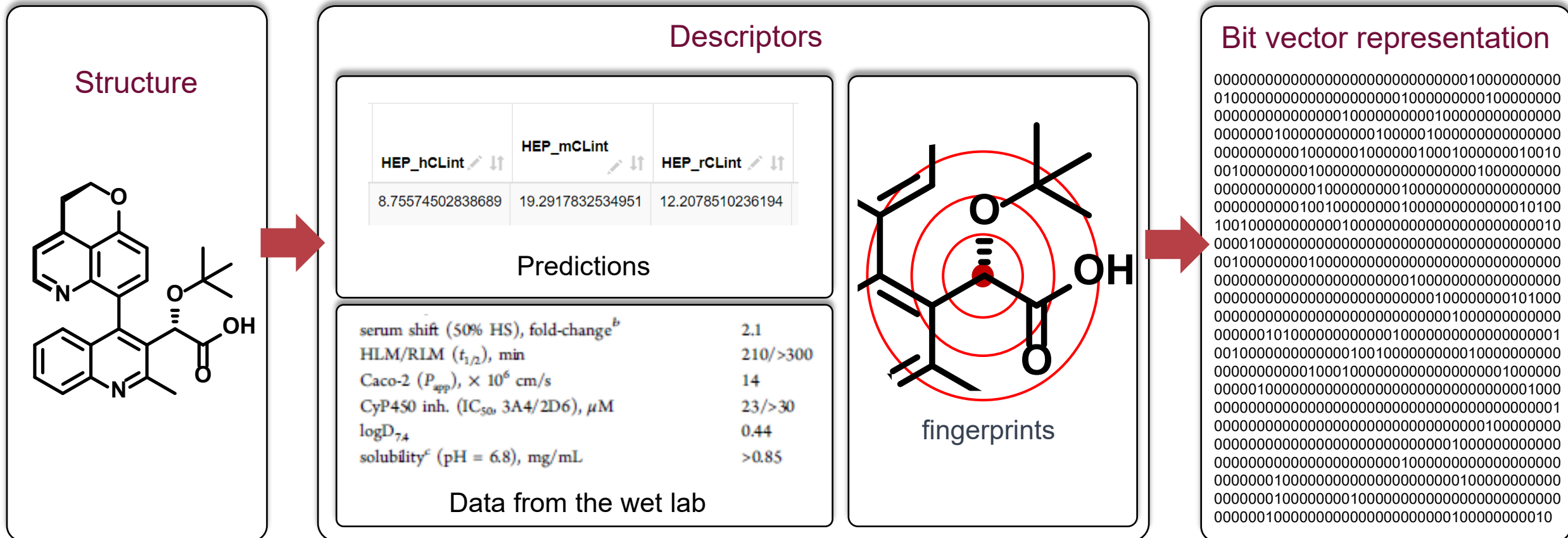
SMILES string: CC1=NC2=CC=CC=C2C(=C1[C@@H](C(=O)O)OC(C)(C)C)C3=C4C5=C(C=C3)OCCC5=CC=N4

InChI: 1S/C27H26N2O4/c1-15-21(25(26(30)31)33-27(2,3)4)23(17-7-5-6-8-19(17)29-15)18-9-10-20-22-16(12-14-32-20)11-13-28-24(18)22/h5-11,13,25H,12,14H2,1-4H3,(H,30,31)/t25-/m0/s1

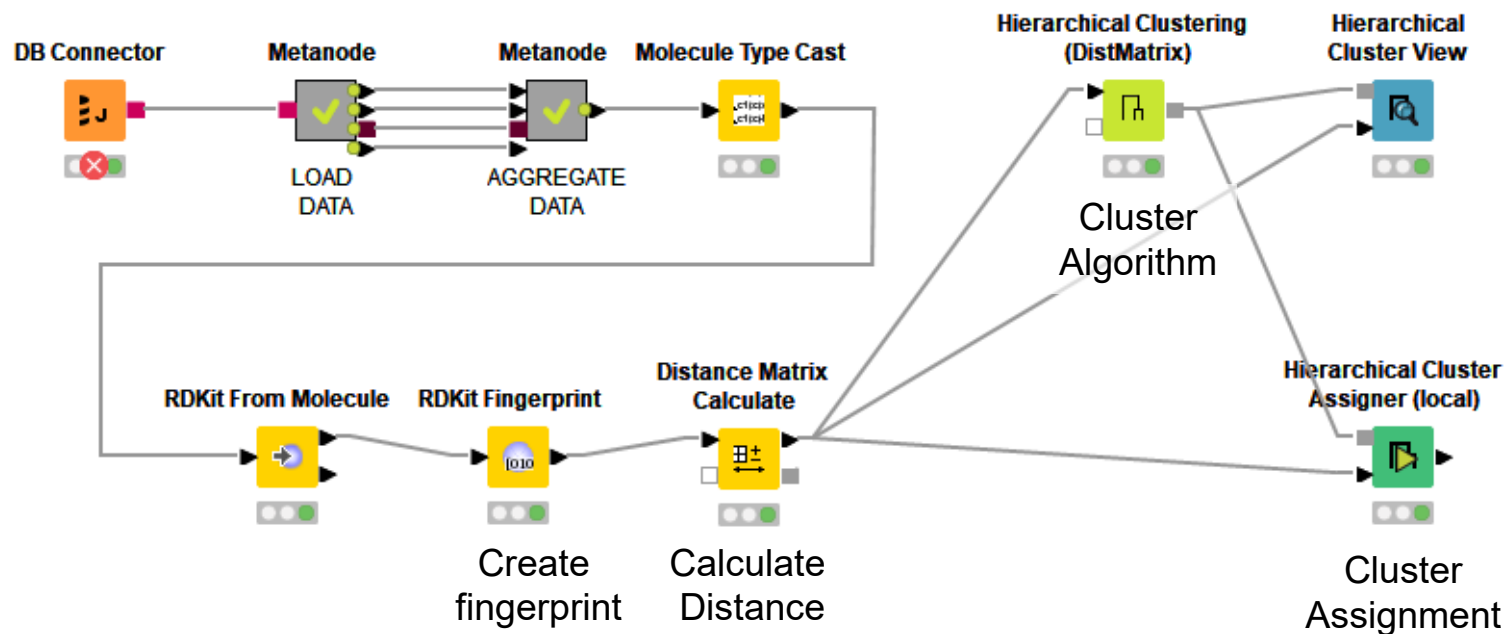
InChIKey: MIXIIJCBELCMCZ-VWLOTQADSA-N

Creation of chemical fingerprints is a crucial step in cheminformatics

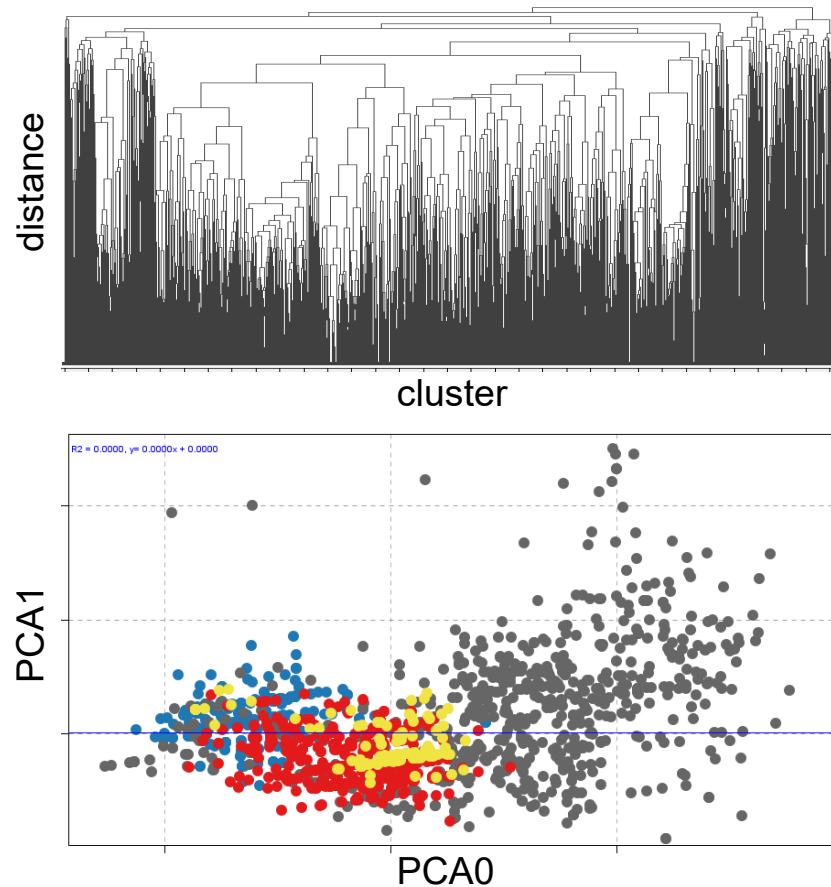
- Chemical fingerprints encode structure-based descriptors



Clustering of compounds through similarity scoring is a common operation involving chemical fingerprints

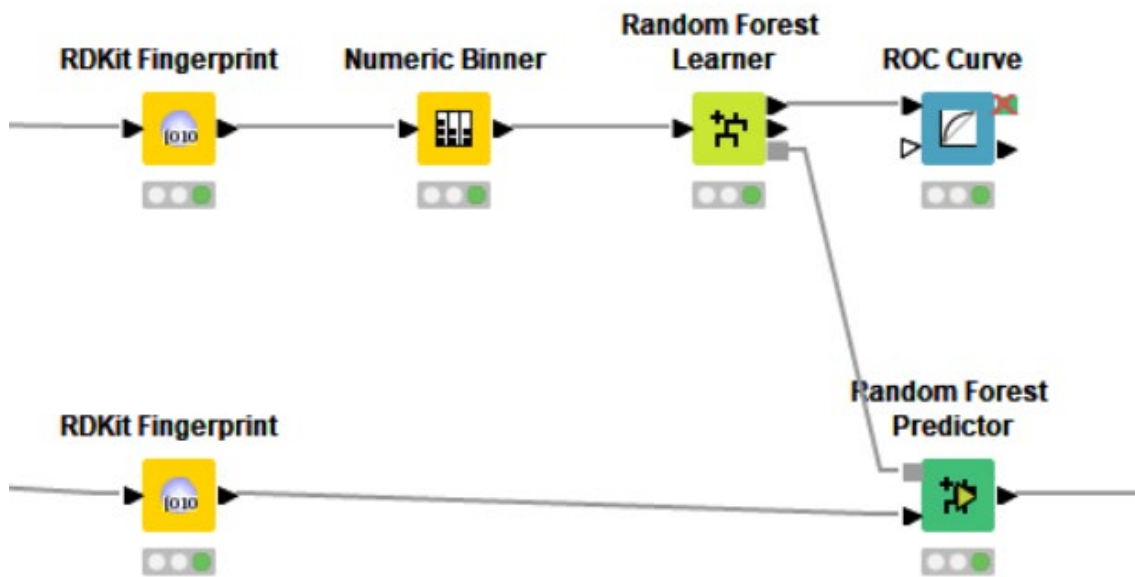


Clustering of Project Compounds



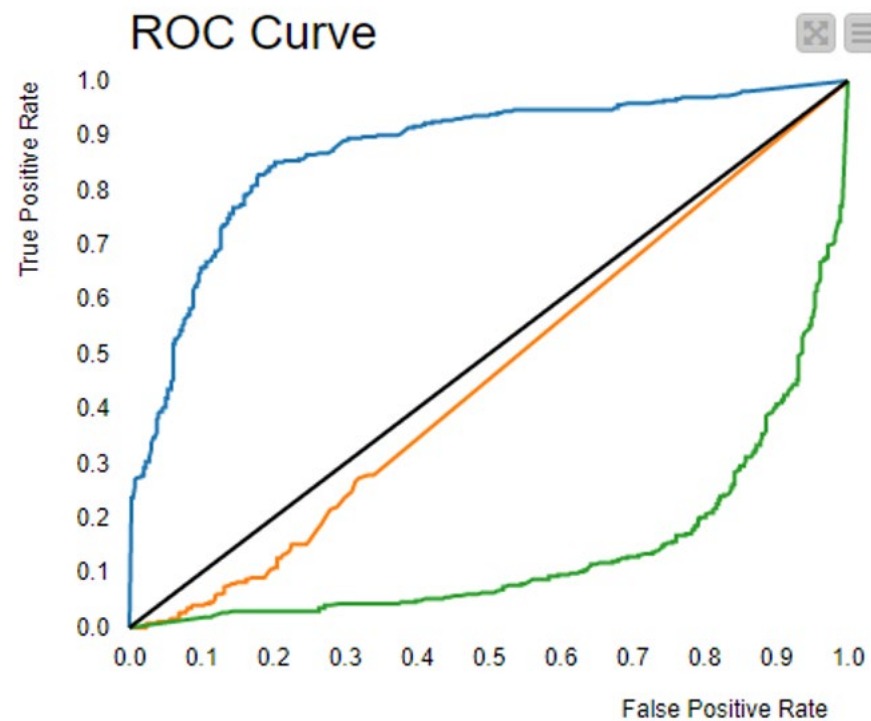
Quantitative Structure Activity Relationships (QSAR) is a contemporary approach to property prediction in Med Chem

Easily Trained and deployed in KNIME

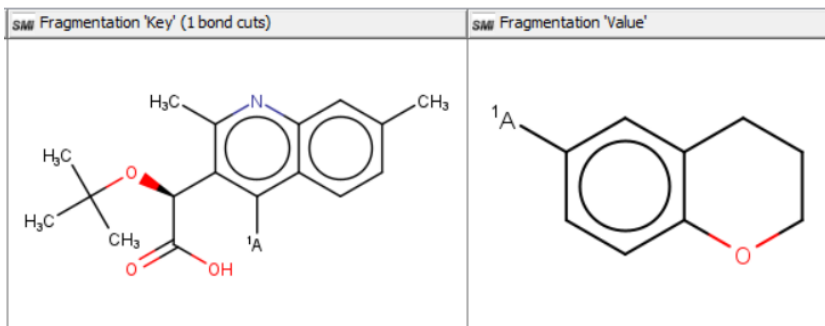
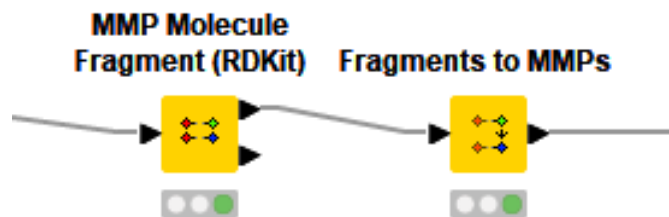


Consensus Approach with complimentary fingerprints often gives best results

Performance is suitable for guiding design

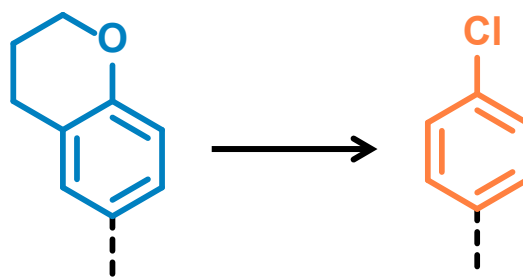
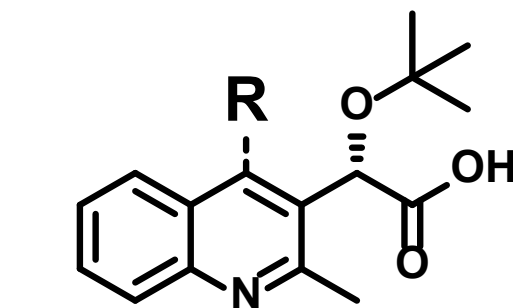


Matched Molecular Pair (MMP) analysis is an effective approach to SAR interpretation

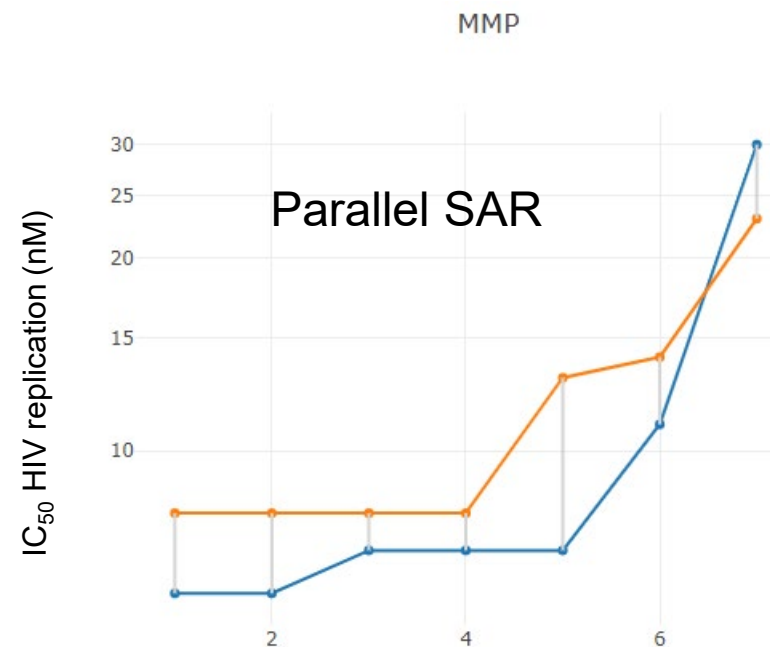


Fragment Key
(constant in a matched pair)

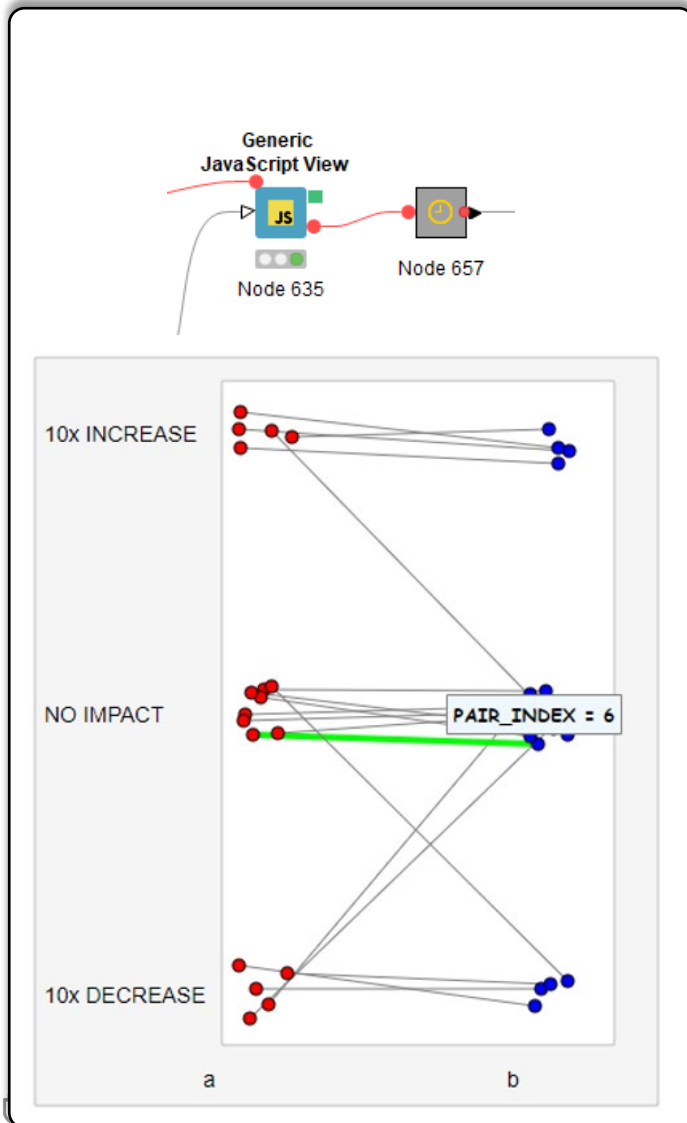
Fragment Value
(variable in a matched pair)



MMP Transformation



The Generic JavaScript View: A workhorse for data visualization-driven decision making



The screenshot shows the 'Generic JavaScript View' dialog box. The 'Flow Variables' section includes 'JSON_axis', 'RowID', 'kname_workspace', and 'JSON_lines'. The 'Dependencies' section lists 'D3 - Version 4.2.6', 'Plotly.js - Version 1.47.4', 'jQuery - Version 3.1.1', and 'jQuery UI - Version 1.12.1'. The 'JavaScript' section contains the following code:

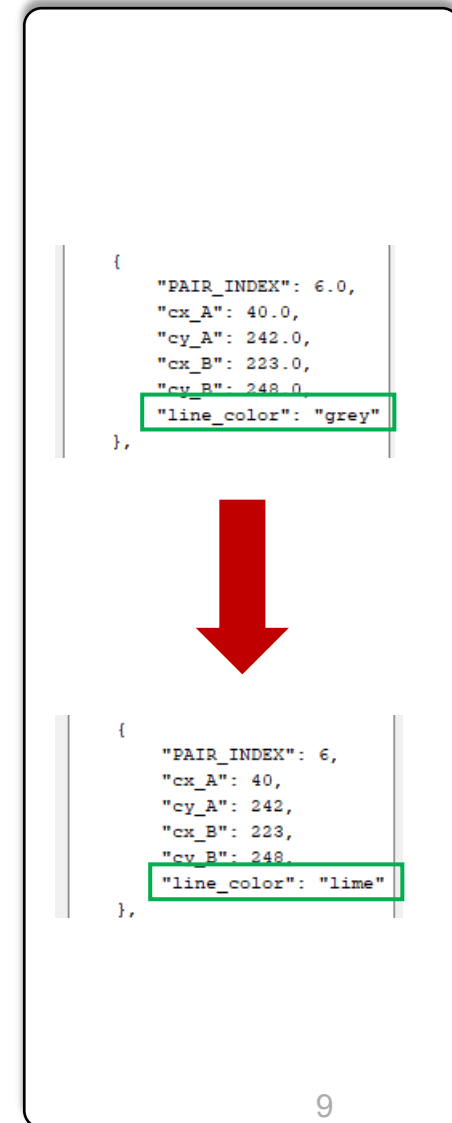
```

49 var lines = svgContainer.selectAll("line")
50 .data(lineData)
51 .enter()
52 .append("line")
53 .attr("x1", function (d) { return d.cx_A+120; })
54 .attr("y1", function (d) { return d.cy_A+20; })
55 .attr("x2", function (d) { return d.cx_B+120; })
56 .attr("y2", function (d) { return d.cy_B+20; })
57 .style("stroke-width", 1)
58 .style("stroke", function (d) { return d.line_color; });
59
60 .on("click", function (d) {
61   d.line_color = "lime";
62   d3.select(this).style("stroke", "lime");
63   FLOW_VARIABLES["JSON2"] = [];
64   svgContainer.selectAll("line").each(function(d,i) {
65     FLOW_VARIABLES["JSON2"].push(d);
66   });
67   FLOW_VARIABLES["JSON2"] = JSON.stringify(FLOW_VARIABLES["JSON2"]);
68 });
69
70 .on("dblclick", function (d) {
71   d.line_color = "grey";
72   d3.select(this).style("stroke", "grey");
73   FLOW_VARIABLES["JSON2"] = [];
74   svgContainer.selectAll("line").each(function(d,i) {
75     FLOW_VARIABLES["JSON2"].push(d);
76   });
77   FLOW_VARIABLES["JSON2"] = JSON.stringify(FLOW_VARIABLES["JSON2"]);
78 });
79

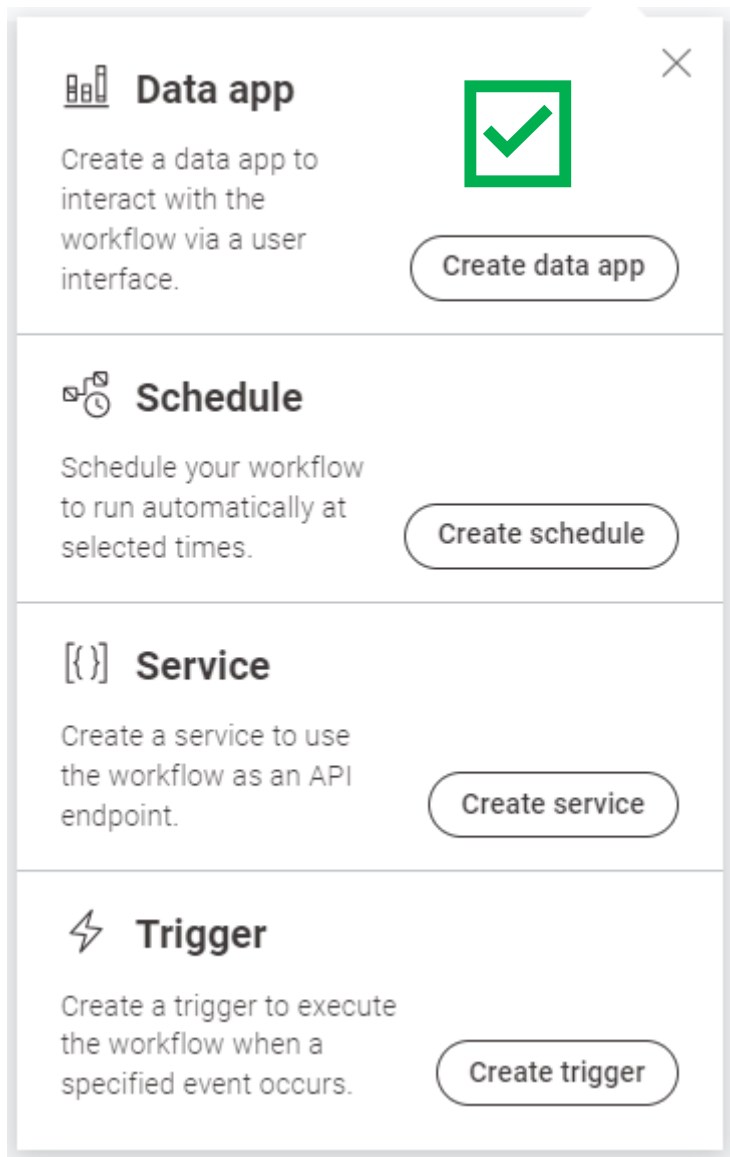
```

The 'Output Flow Variables' table is also visible:

Replace	Flow Variable	Output Type	Script Field	Default Value
<input type="checkbox"/>	JSON2	string	FLOW_VARIABLES["JSON2"]	



Workflows archived and deployed with KNIME Business Hub





























The screenshot displays a vertical menu with four options, each with a description and a 'Create' button. The 'Data app' option is highlighted with a green checkmark icon in a box. A red arrow points from this option towards the text box on the right.

- Data app**
Create a data app to interact with the workflow via a user interface.
Create data app
- Schedule**
Schedule your workflow to run automatically at selected times.
Create schedule
- Service**
Create a service to use the workflow as an API endpoint.
Create service
- Trigger**
Create a trigger to execute the workflow when a specified event occurs.
Create trigger

At congruence we deploy workflows as **Data apps** to:

- Put compound design tools into the hands of as many people as possible
- Centralize decision-driving data to shorten time it takes to respond to data
- Facilitate project management and CRO interactions to maintain a high level of efficiency

The KNIME Hub allows for rapid deployment of cheminformatic tools to guide design and decision making

 ADME+_PREDICT ⓘ Provided by:  CongruenceTX	 ADME_PREDICT ⓘ Provided by:  CongruenceTX	 ADME_REQUESTS_M ⓘ Provided by:  CongruenceTX	 ADME_TRACKER_M ⓘ Provided by:  CongruenceTX
 MMP_ANALYSIS ⓘ Provided by:  CongruenceTX	 PGP_PREDICT ⓘ Provided by:  CongruenceTX	 REGISTRATION_TEMPLATE ⓘ Provided by:  CongruenceTX	 SERIES_ANNOTATION ⓘ Provided by:  CongruenceTX
 SIMILARITY_SCORING_M ⓘ Provided by:  CongruenceTX	 SPOTFIRE_G ⓘ Provided by:  CongruenceTX	 SPOTFIRE_M ⓘ Provided by:  CongruenceTX	 TASK_PRIORITIZATION ⓘ Provided by:  CongruenceTX
 VENN_PLOTTER ⓘ Provided by:  CongruenceTX			

Bringing it all together: Using KNIME Business Hub to accelerate drug discovery

ADME_REQUESTS_M

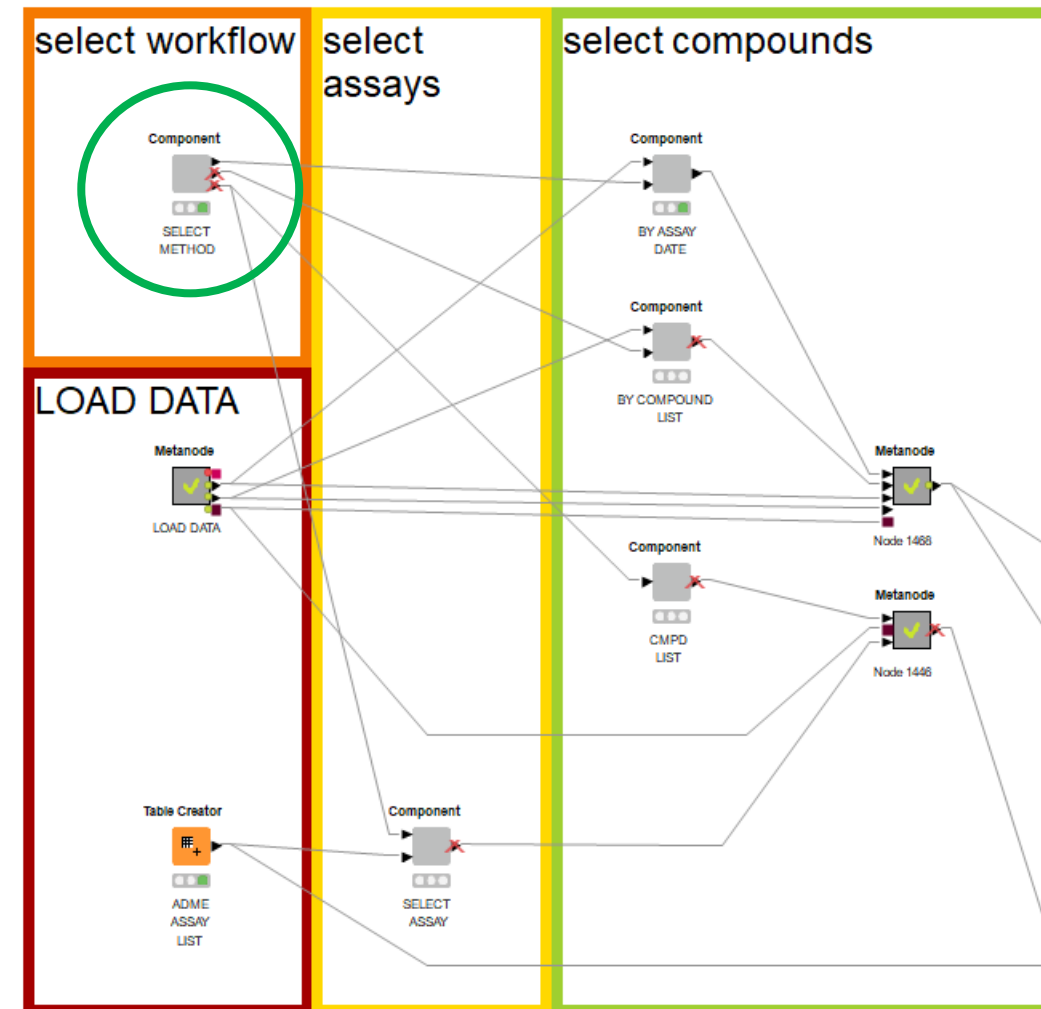
Provided by: CongruenceTX



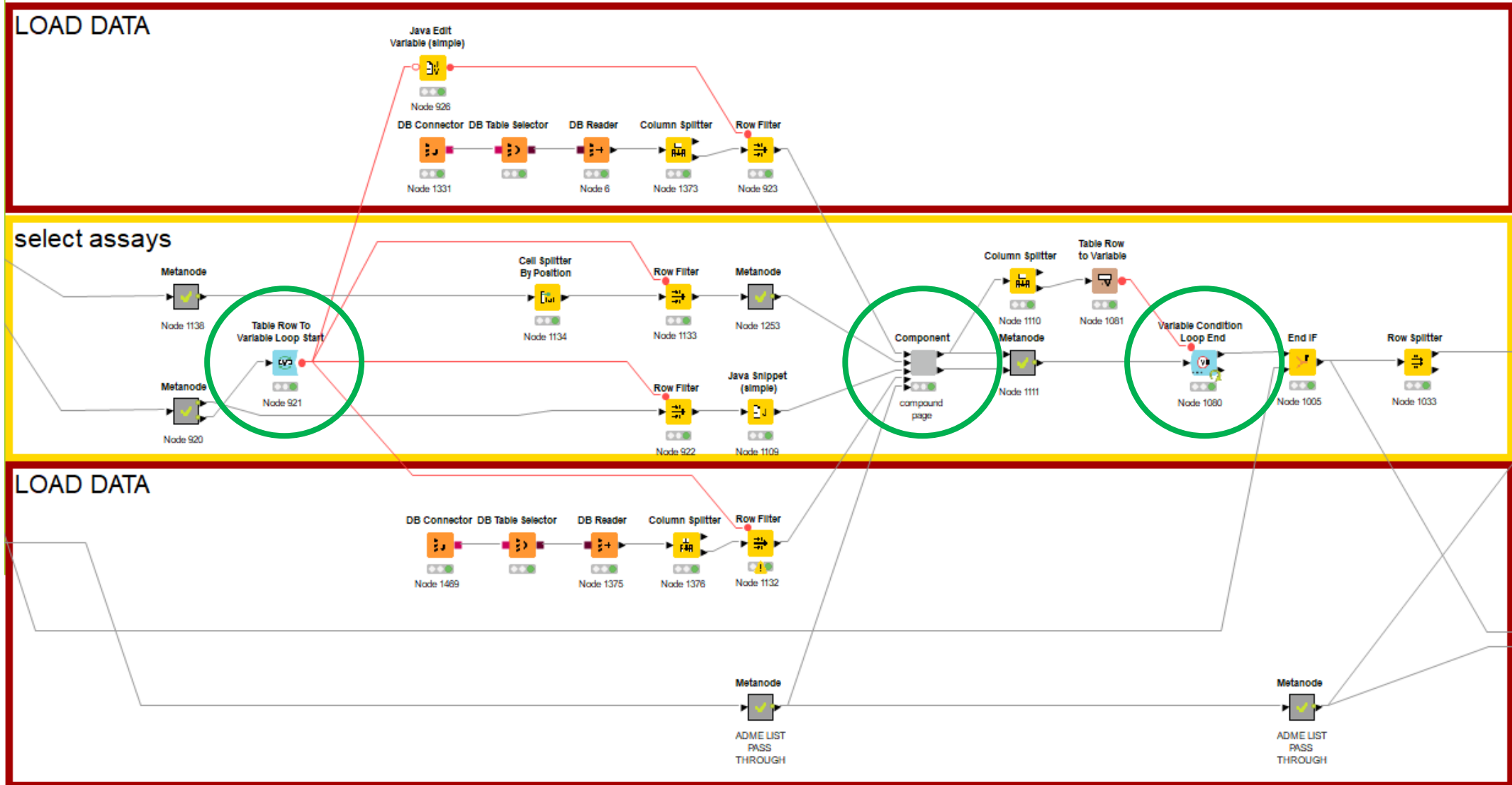
congruence^{TX} **From Topology to Treatments**

Choose the type of Request

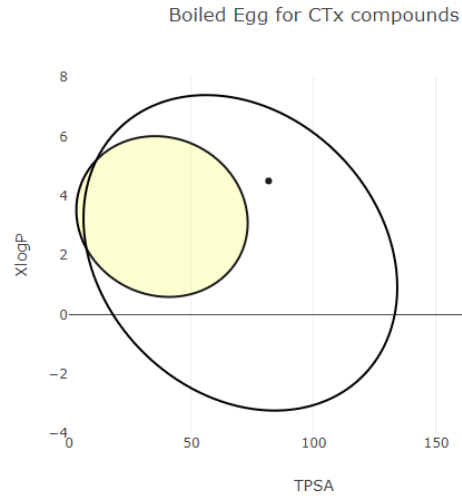
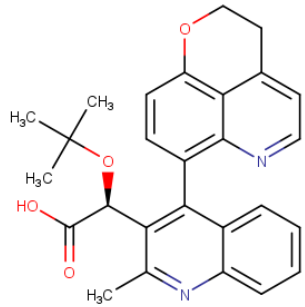
- Select by Compound Codes
- Select by Assay Date
- Select by Assay List



Data Apps with Components within Loops: A versatile approach to decision making



With KNIME and KNIME Hub, decision making tools designed to retrieve and analyze data are easily assembled and deployed



PROGRESS

compound 10 out of 11

Check this box to stop after this entry.

RowID	Row0
Topological Polar Surface Area	81.54
Molecular Weight	442.18925731200073
XLogP	4.499
CYP_HLM_CLint	10.424613703671
HEP_hCLint	8.75574502838692
HEP_mCLint	19.291783253495
HEP_rCLint	12.2078510236195
UGT1A1	Yes (85%)
UGT1A10	No (83%)
UGT1A3	Yes (78%)
UGT1A4	No (64%)
UGT1A6	No (97%)

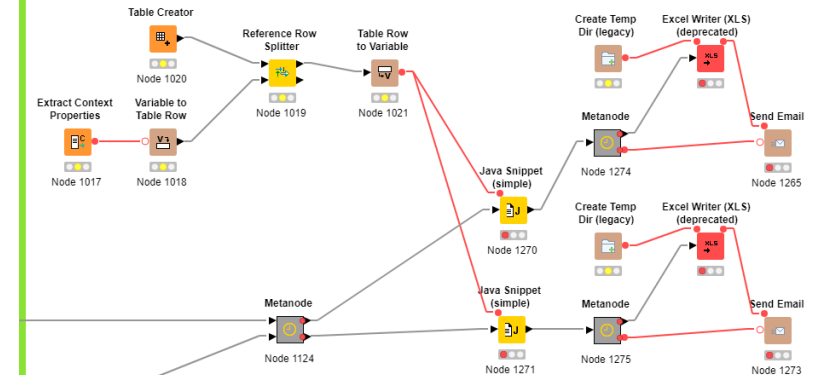
Assay	Days since request
Stability in liver microsomes (5 time points, only NADPH)_mouse	3
Stability in liver microsomes (5 time points, only NADPH)_mouse	3
Stability in liver microsomes (5 time points, only NADPH)_human	3
Stability in liver microsomes (5 time points, only NADPH)_human	3
hERG single point automated patch-clamp system (% inh at 10 uM, minimum 10 compounds)	3

Showing 1 to 5 of 5 entries

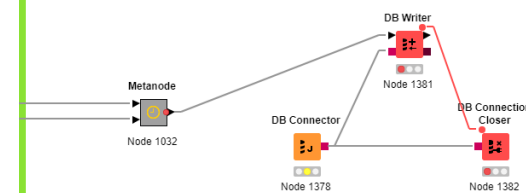
ASSAY REQUEST LIST

ASSAY	Compound Number
<input type="checkbox"/> Caco-2 (A to B)	not requested
<input type="checkbox"/> Stability in liver microsomes (5 time points, only NADPH)_mouse	pending
<input type="checkbox"/> Stability in liver microsomes (5 time points, only NADPH)_human	pending

send assay requests



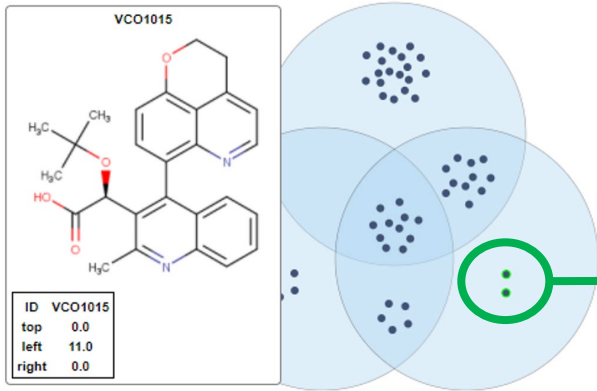
log assay requests



Example: the Venn Diagram-Scatter Plot hybrid

A custom visualization for guided analytics

Refresh



Filter based
on selection

```
.on("click", function (d) {  
  d.line_color = "lime";  
  d3.select(this).style("stroke", "lime");  
  FLOW_VARIABLES["JSON2"] = [];  
  svgContainer.selectAll("line").each(function(d,i) {  
    FLOW_VARIABLES["JSON2"].push(d);  
  });  
  FLOW_VARIABLES["JSON2"] = JSON.stringify(FLOW_VARIABLES["JSON2"]);  
});
```

Number of Compounds with all three data points:

56

left: HLM 2pt (%QH)

max

75

min

0

top: CYP Inhibition HTL: % Inh 3A4-T

max

50

min

0

right: hERG Inhibition APC Single Point: % Inhibition at 10 uM

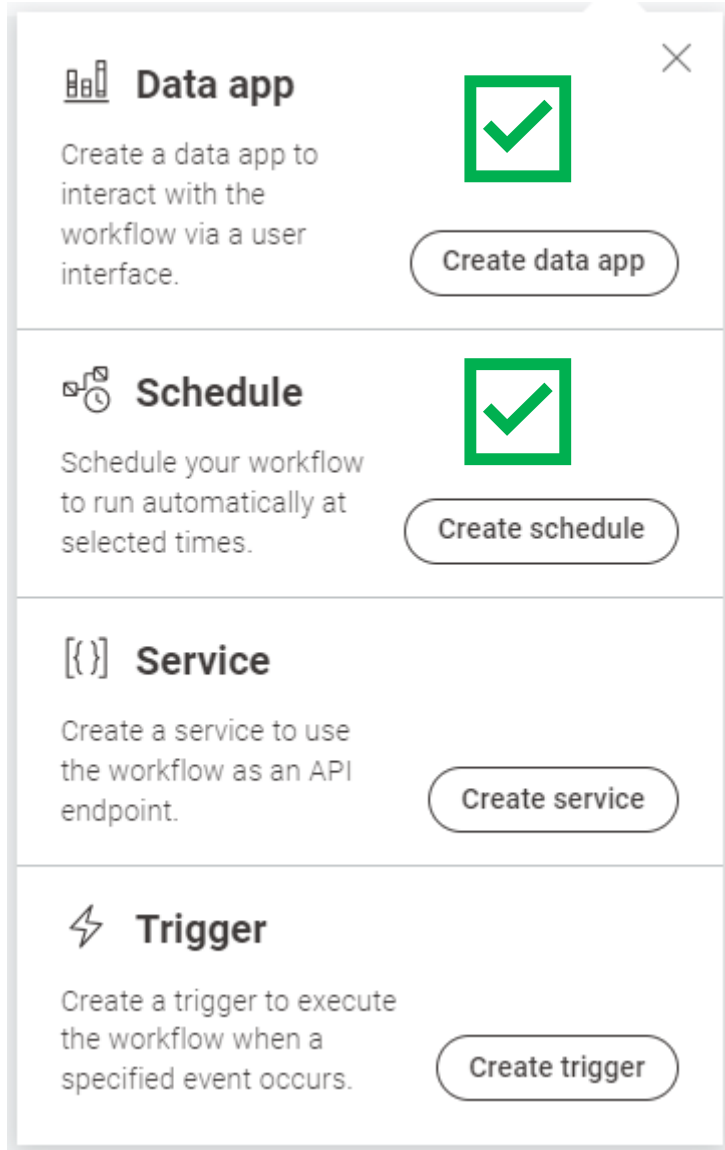
max


50


min

0

Workflows archived and deployed with KNIME Business Hub



Data app
Create a data app to interact with the workflow via a user interface.

[Create data app](#)

Schedule
Schedule your workflow to run automatically at selected times.

[Create schedule](#)

Service
Create a service to use the workflow as an API endpoint.
[Create service](#)

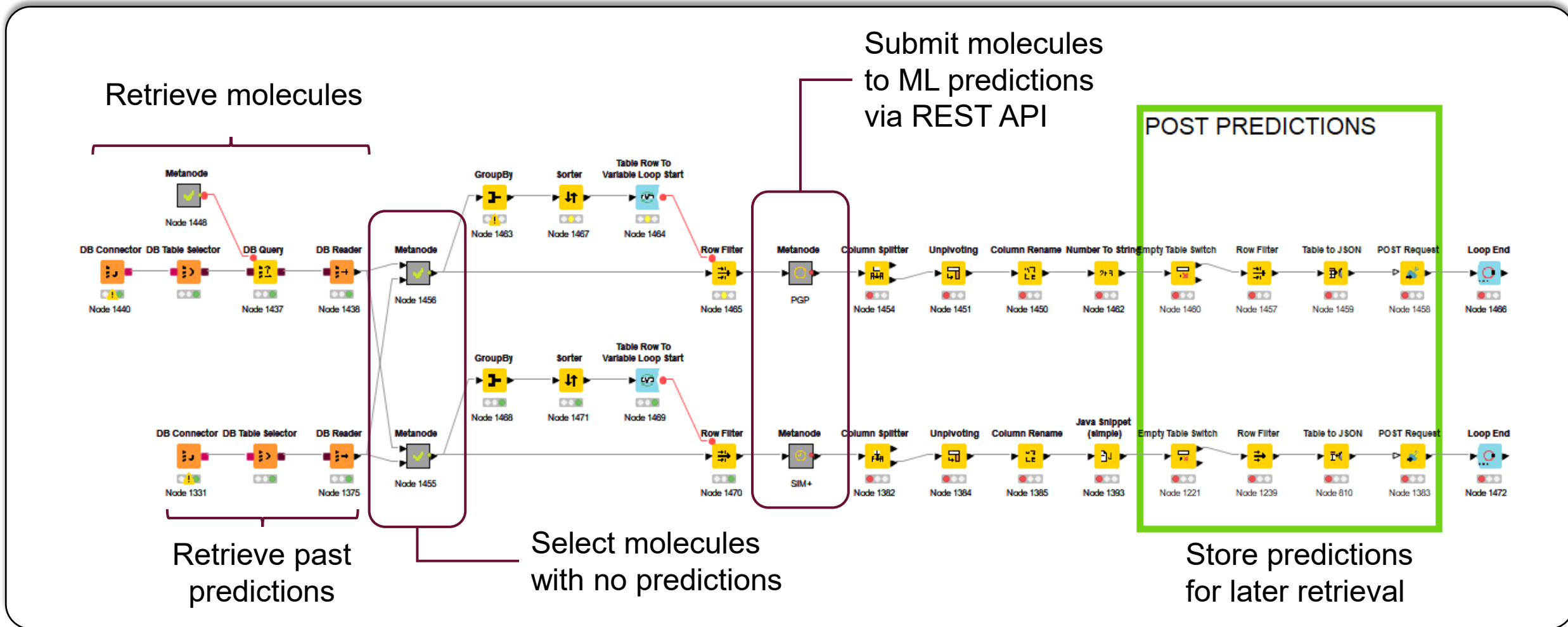
Trigger
Create a trigger to execute the workflow when a specified event occurs.
[Create trigger](#)



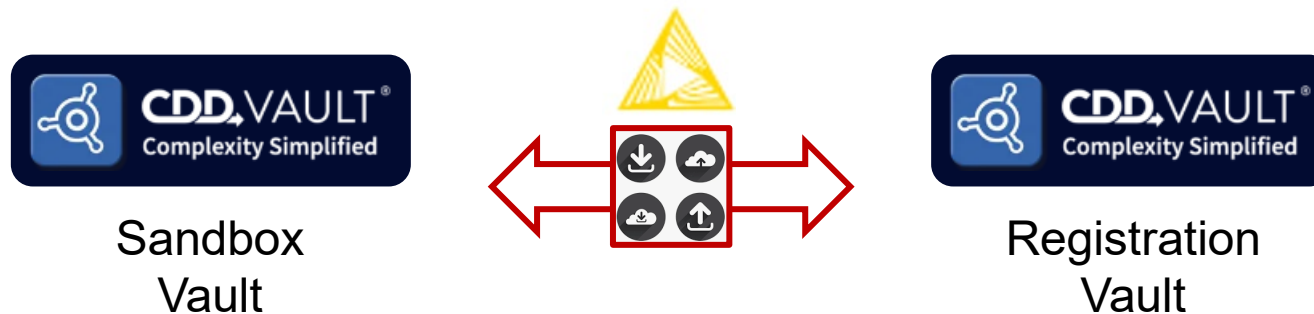
At congruence we deploy workflows as **Schedules to:**

- Automate regular calculations
- Conduct data aggregation tasks
- Preparation and transfer of information needed for more sophisticated tools

Services deployments automate the machine-learning prediction of new properties for molecules that are added to our databases



Services deployments help manage our compound synthesis queues and automate registration of new compounds



825 Selected: [Launch Visualization](#) [Export](#) [Add to collection](#) [Build model](#) [Flag outliers](#) [Customize your report](#) [Save this search](#)

Select... all none	Molecule	Properties		Molecule Fields					Batch Fields						
		Molecular weight (g/mol)	log P	Synonyms	Series	Designer	Hypothesis	Subseries	Batch Name	Created Date	Projects	Molecule-Batch ID	Status	Priority	Initial Amount
<input checked="" type="checkbox"/>	 VIRTUAL-0012128 Congruence Sandbox SMILES	423.940	5.5	LF-4066-2900	BI-SSS	Congruence LF	Pyr better ER	6-Ph	001	2024-03-18	TEST	VIRTUAL-0012128-001	Shipped	High	40.0

Service deployment automates exchange of information between CDD Vaults



Key Takeaways

- In medicinal chemistry, the central paradigm of compound optimization follows a Design-Make-Test-Analyze cycle that is iterated over as efficiently as possible.
- To do this effectively, different representations of chemical structures that are interpretable by machines and can be translated to human interpretable representations are needed.
- Chemical descriptors can be derived from chemical structures and assembled into fingerprints, which are used for myriad analytical tasks, modelling and machine learning.
- Aggregating predictions, biological data and other cheminformatic outputs enables informed decision making in drug discovery.
- The KNIME Business Hub allows for flexible construction of workflows, services and data apps that facilitate decision making and accelerate drug discovery.

