



CambridgeSoft®
Life Science Enterprise Solutions



Louis J Culot
Vice President, Enterprise Applications & New Business
CambridgeSoft Corporation
lculot@cambridgesoft.com

CambridgeSoft Background

CambridgeSoft

Large life-science enterprise customer base

Market leader for Electronic Notebook implementations

Strong Informatics and Desktop presence

Global Consulting service

Including on-shore and off-shore in major continents

Global Sales and Support

World-wide sales/service/support agreement with KNIME.Com

World-Wide Customer Support

CambridgeSoft Enterprise Customer Support Centers:

US East Coast, US West Coast, UK, Germany, Japan, China

Native languages: English, French, German, Spanish,
Japanese, Chinese, Russian.

Global Enterprise support

8AM Japan Monday -> 5PM US West Friday continuous.

World-Wide Consulting Service

CambridgeSoft Enterprise Consulting:

130 Enterprise Consultants co-located with customers, and offshore development centers.

On-shore co-location U.S., Canada, U.K., France, Germany, Switzerland, Sweden, Spain, Italy, Japan, China, India.

Off-shore development centers: India, China, Lebanon, Argentina, Belarus.

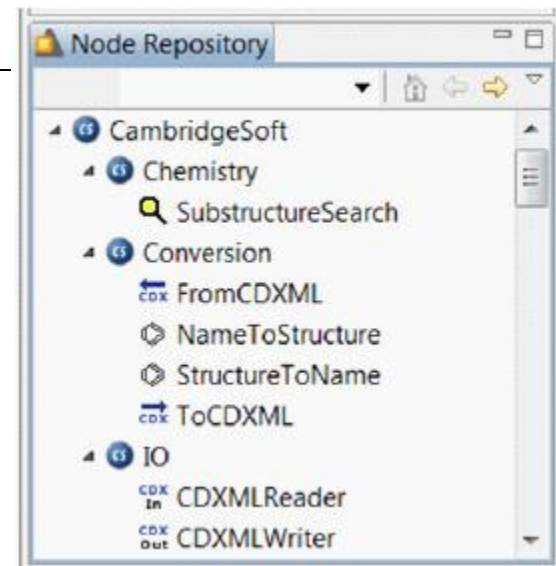
Most (>90%) consultants hold degrees in chemistry or biology in addition to software engineering.

A large percentage of staff hold advanced degrees (Master's, Ph.D.).

Node Development

Cheminformatics Nodes

- CDXMLReader
- CDXML Writer
- ToCDXML
- FromCDXML
- Substructure Search
- StructureToName
- NameToStructure
- Chemsript Functionality



Workflow Editor: Structures and Names - 05 - StructureToName

Row ID	Molecules	Name
1	tryptophan	tryptophan

Workflow Editor: Structures and Names - 03 - StructureToName

Row ID	Mol Block	Name
Row1	99451 10 11 0 ... 2,8090 ...	2,2'-b(1,3-dihydrolydine)
Row2	66810 14 13 0 ... 4,5981 ...	N1,N1,N4,N4-tetramethyltetrasulfane-1,4-bis(carbothioamide)
Row3	81872 18 17 0 ... 3,7321 ...	1,4-diethyltetrasulfane
Row4	81777 17 16 0 ... 3,7321 ...	1,10-di-tert-butyldecaulfane
Row5	81776 16 15 0 ... 3,7321 ...	1,9-di-tert-butylnonasulfane
Row6	81775 17 16 0 ...	1,8-di-tert-butyloctasulfane
Row7		1,7-di-tert-butylheptasulfane

Dialog Box: StructureToName

Generate chemical names for input structures.

Dialog Options

Column to name
Column containing structures to be named

Label for output column
Column label to be used for the output column of structure names

Ports

Input Ports

- DataTable containing a structure column

Output Ports

- DataTable consisting of the input DataTable and an appended column of structure names

Node Development

Cheminformatics Nodes

- CDXMLReader
- CDXML Writer
- ToCDXML
- FromCDXML
- Substructure Search
- StructureToName
- NameToStructure
- Chemsript Functionality

The screenshot displays the KNIME software interface with several windows open. The main workspace shows a workflow with nodes: CDXMLReader (Node 1), StructureToName (Node 5), CDXMLWriter (Node 2), SDF Reader (Node 4), and StructureToName (Node 3). A 'Node Repository' pane on the left lists various nodes like Database, Data Manipulation, and CDK. Two data tables are visible:

Table "default" - Rows: 1 | Spec - Columns: 2 | Properties

Row ID	Molecules	Name
1	tryptophan	tryptophan

Table "default" - Rows: 18 | Spec - Columns: 2 | Properties

Row ID	Mol Block	Name
99451	-OEChem-1...	2,2'-bi(1,3-dithiolydene)
10 11 0 ...	2,8090 ...	
66810	-OEChem-1...	N1,N1,N4,N4-tetramethyltetrasulfane-1,4-bis(carbothioamide)
14 13 0 ...	1,5951 ...	
-OEChem-1...		1,4-diethyltetrasulfane
8 7 0 ...	1,0461 ...	
81872	-OEChem-1...	1,10-di-tert-butyldecasulfane
18 17 0 ...	3,7321 ...	
81777	-OEChem-1...	1,9-di-tert-butylnonasulfane
17 16 0 ...	3,7321 ...	
81776	-OEChem-1...	1,8-di-tert-butyloctasulfane
16 15 0 ...	3,7321 ...	
81775	-OEChem-1...	1,7-di-tert-butylheptasulfane

A 'Node Description' window for 'StructureToName' is also open, showing dialog options for column naming and output structure names.

- see Joe for a Demo (the guy with the white hair)

ChemBioOffice Enterprise & KNIME

The screenshot displays the ChemBioOffice Enterprise software interface. At the top, there are navigation tabs for 'Reaction' and 'Structure'. The main workspace shows a chemical reaction scheme with a reactant structure on the left and a product structure on the right, connected by a reaction arrow. Below the reaction, there is a table with columns for 'Reaction', 'Name', 'Rate', and 'Value'. The table contains several rows of data, including chemical names and numerical values. On the left side, there is a sidebar with a tree view showing a hierarchical structure of the project.

The screenshot shows the KNIME Reporting Server web interface. The header includes the KNIME logo and the text 'KNIME Reporting Server'. Below the header, there is a section titled 'Available Reports:' which lists several report types: 'Reporting Workflow', 'Reports', 'Customer Data', 'Data Table', and 'Report with Structure'. The interface is clean and professional, with a navigation bar at the top.



The screenshot displays a complex reaction network within the ChemBioOffice Enterprise software. The network consists of numerous interconnected nodes, each representing a chemical structure or a reaction step. The nodes are arranged in a hierarchical and interconnected manner, showing the flow of the chemical process. The interface includes a sidebar with a tree view and a main workspace with a detailed view of the reaction network.


The screenshot shows a 'DMR Report Example' for the month of October 2009. The report is divided into several sections: 'Month to Date (to 10/09)', 'Year to Date (to 10/09)', and 'In P/L'. Each section contains bar charts, pie charts, and tables. The bar charts compare 'Actual' and 'Budget' values for different categories. The pie charts show the distribution of values. The tables provide detailed numerical data for each category. The report is visually appealing and easy to read, with clear labels and legends.



CambridgeSoft Enterprise Integration Nodes

- **Registration Nodes**
 - SearchRegistration, SendToRegistration
 - GetRegistration – Retrieve information about registered compound from KNIME
- **Inventory Specific**
 - SendToInventory, UpdateInventory (Create/Update Inventory compounds and containers)
 - SearchInventory – Retrieve compounds and containers based on parameters
 - GetInventory – Retrieve specified data from Inventory
- **E-Notebook supporting Nodes**
 - Authentication
 - Substructure/Reaction Search –
 - SectionFormatter – Format data from sections and tables in the E-Notebook to be analyzed by the KNIME workflow

Enterprise Market Sample

INSTALLATIONS		RESEARCH				PRE-CLINICAL DEVELOPMENT					
	Total	Chemistry	Biology	DMPK	Safety	Process Development	Analytical & Bioanalytical	Pharma R&D	Biologics	QC	Customer or eval
Merck	5,000	1300	1600	200	200	400	500	400	250	150	Knime.Com
GSK	3,600	1,100	1500	600	300		50	50			Desktop
Astra Zeneca	1,000	1,000									Knime.Com
Novartis	1,235	850		25		300		60			Knime.Com
BMS	1,540	400	900	100			140				Desktop
Abbott	530	300				200		30			
Sanofi	1,300	1000	100	100			100				Knime.Com
Roche	1,100	800				300					Desktop
Arena	245	50	75	30		30	30	30			
Cubist	225	50	80	25	25	25		10		10	Desktop
VIB	1,000		1000								
Array	230	100	80	20		20		10			
PM USA	400	50	150			150	50				
Firmenich	300	200	100								
IFF	400	200	100				100				Desktop
NCI	50		50								
Dow	300	100	100			25	75				
Infinium	350	100				250					Desktop
Eastman	1,000	300	200			300	200				
Nektar	350	75	75	50	25	25	25	25		50	
Shionogi	175	100				75					
Amgen	150	150									Desktop
Takeda	550	500	50								Desktop
	21,030	8725	6160	1150	550	2100	1270	615	250	210	
In Production		16585				4445					
In Progress											
Pilot											


- 40,000 users in production world-wide at > 100 accounts.

- Global deployments spanning thousands of users and several continents.

- Domains from chemistry, Biology, DMPK, Safety, Process Development, Analytical, and QC.

- First customer in production in 1999.

Enterprise Market Sample

INSTALLATIONS		RESEARCH				PRE-CLINICAL DEVELOPMENT					
	Total	Chemistry	Biology	DMPK	Safety	Process Development	Analytical & Bioanalytical	Pharma R&D	Biologics	QC	Customer or eval
Merck	5,000	1300	1600	200	200	400	500	400	250	150	Knime.Com
GSK	3,600	1,100	1500	600	300		50	50			Desktop
Astra Zeneca	1,000	1,000									Knime.Com
Novartis	1,235	850		25		300		60			Knime.Com
BMS	1,540	400	900	100			140				Desktop
Abbott	530	300				200		30			
Sanofi	1,300	1000	100	100			100				Knime.Com
Roche	1,100	800				300					Desktop
Arena	245	50	75	30		30	30	30			
Cubist	225	50	80	25	25	25		10		10	Desktop
VIB	1,000		1000								
Array	230	100	80	20		20		10			
PM USA	400	50	150			150	50				
Firmenich	300	200	100								
IFF	400	200	100				100				Desktop
NCI	50		50								
Dow	300	100	100			25	75				
Infinium	350	100				250					Desktop
Eastman	1,000	300	200			300	200				
Nektar	350	75	75	50	25	25	25	25		50	
Shionogi	175	100				75					
Amgen	150	150									Desktop
Takeda	550	500	50								Desktop
	21,030	8725	6160	1150	550	2100	1270	615	250	210	
In Production		16585				4445					
In Progress											
Pilot											

- 40,000 users in production world-wide at > 100 accounts.

- Global deployments spanning thousands of users and several continents.

- Domains from chemistry, Biology, DMPK, Safety, Process Development, Analytical, and QC.

- First customer in production in 1999.

Thousands of scientists putting data 'in'....

Mountains of data...



Mountains of data...



+



Mountains of data...



+

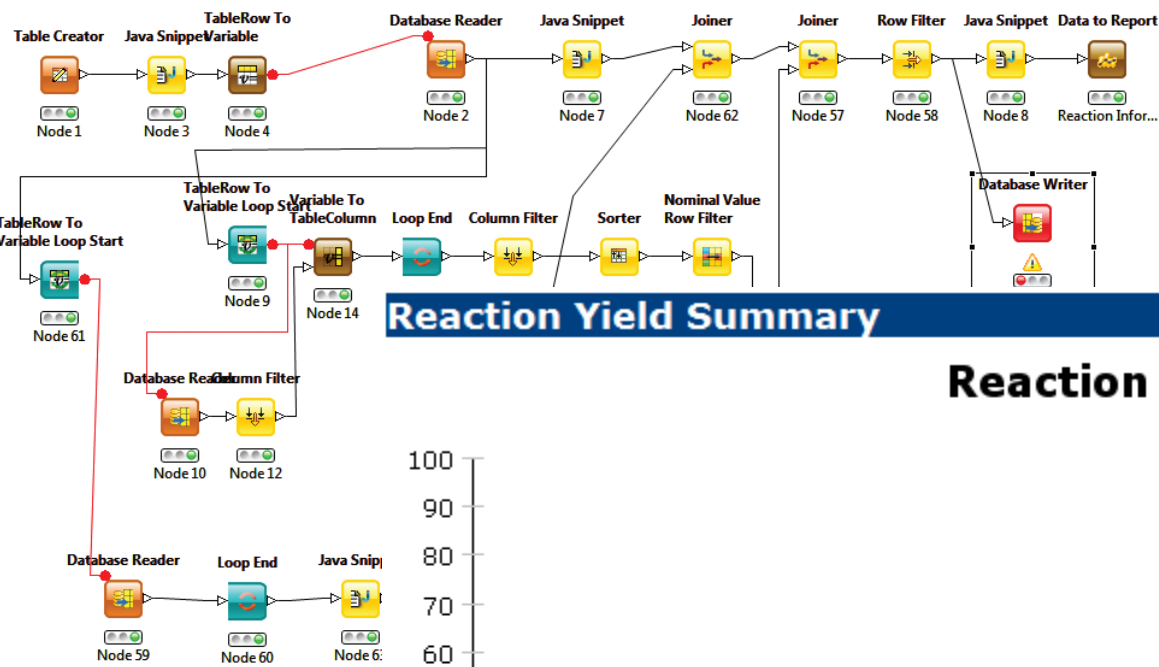


KNIME

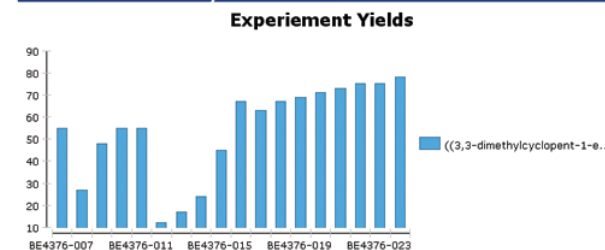
=



E-Notebook Data Mining Workflow

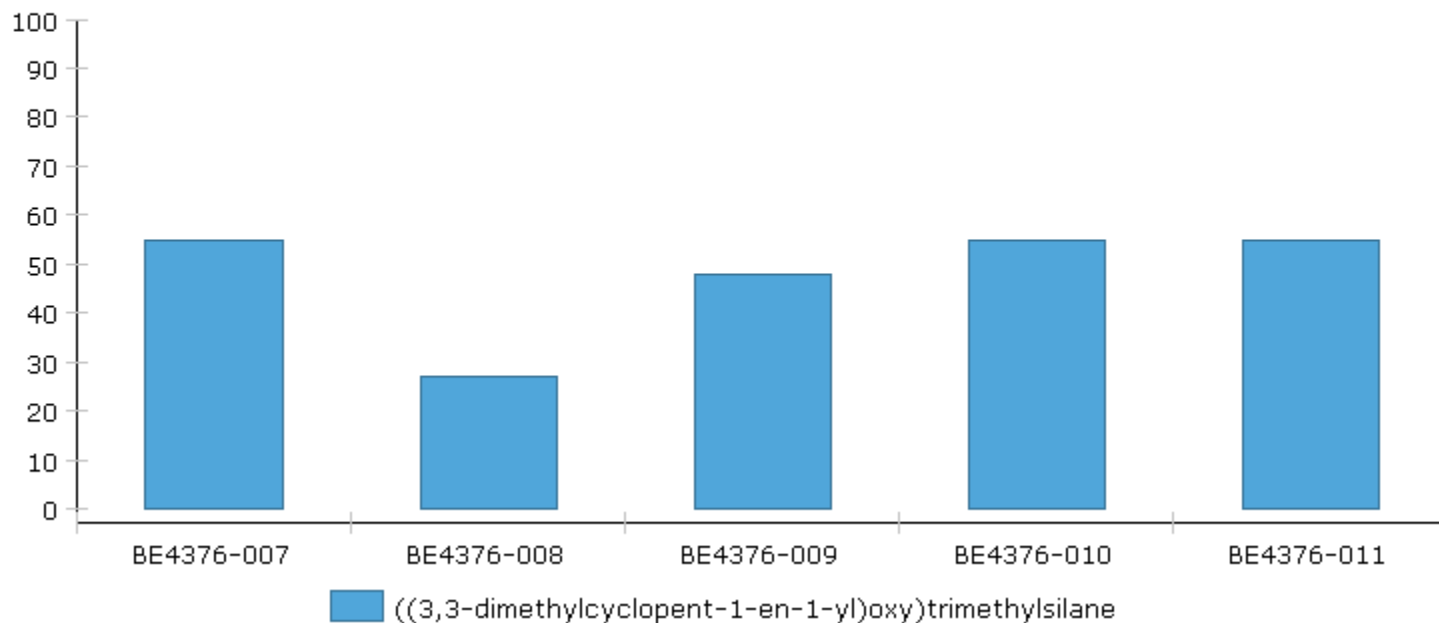


Reaction Yield Summary



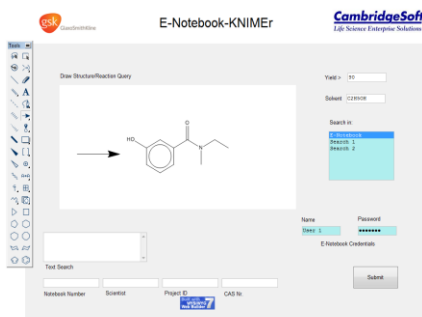
Reaction Yield Summary

Reaction Yields

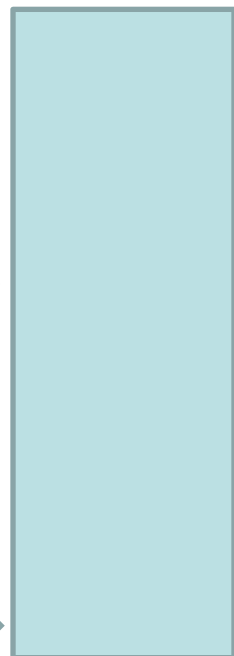


E-Notebook Mining System based on KNIME

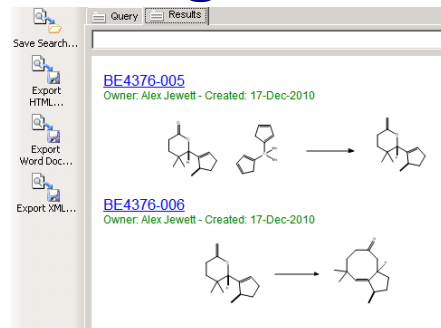
Search Page



Search Manager



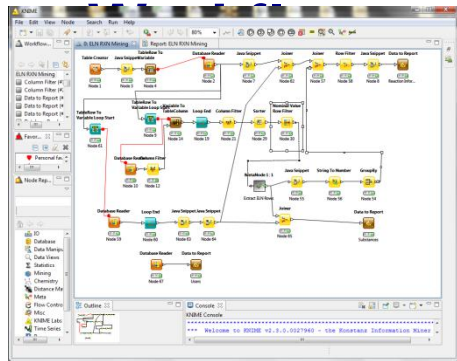
Results Page



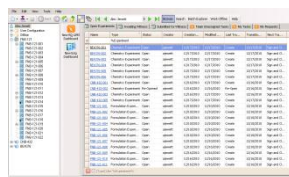
RD Export

- (1) User search entry
- (2) Middle tier saves search info and initiates KNIME via Web Service
- (3) ELN is queried and results are returned to KNIME
- (4) KNIME uploads results into Search Manager
- (5) Results Page for result visualization
- (6) Results can be selected and saved
- (7) Saved lists can be refined based on search criteria

KNIME



E-Notebook



More use cases

- Project 1: Enterprise Chemical Inventory System: Is current Compound handling in compliance to US EPA Standards?
- Project 2: Compound Selection, Screening database enrichment
- Project 3: Add calculated Structure Data to Enterprise Registration System



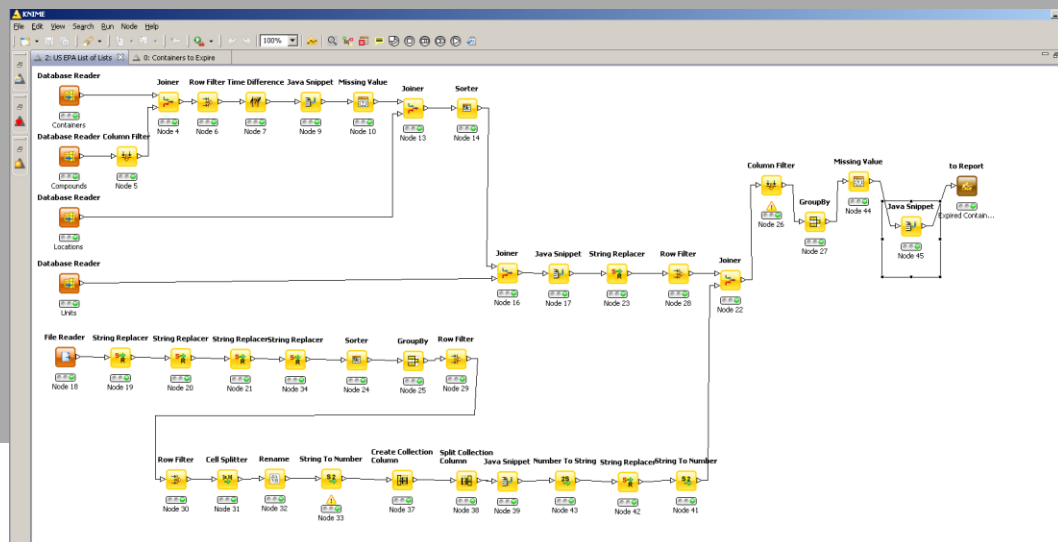
- Solution:
- KNIME server workflow for searching the Enterprise Inventory System for dangerous compounds and summarize results in a report

KNIME Reporting Server

Available Reports:

- ▲ Containers to Expire
- ▲ Molecules
- ▲ US EPA List of Lists

Logout





KNIME Report Viewer

Showing page 1 of 2

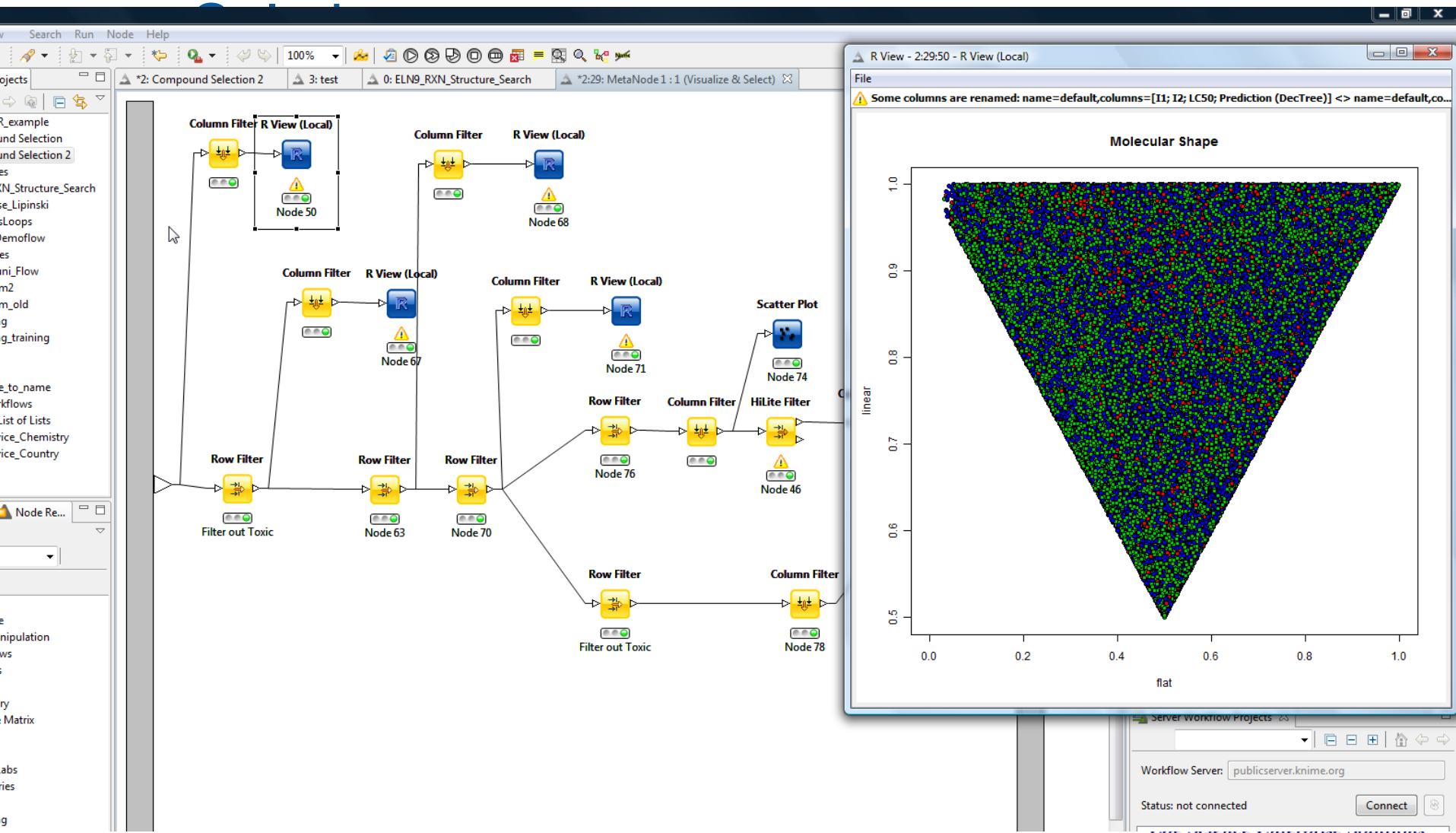
Go to page:

US EPA List of Lists Report

CAS	SUBSTANCE_NAME	Sum(AMT_LBS)
10026-13-8	PHOSPHORUS(V) CHLORIDE	800,00
10034-93-2	HYDRAZINIUM SULFATE	1,00
10099-74-8	LEAD(II) NITRATE	-2,00
106-42-3	P-XYLENE	-2,00
107-06-2	1,2-DICHLOROETHANE	-22,00

CAS	SUBSTANCE_NAME	Sum(AMT_LBS)	
10026-13-8	PHOSPHORUS(V) CHLORIDE	800,00	Section 313 Section 302 (EHS) TPQ 500,00 Section 304 EHS RQ 500,00 CERCLA RQ RCRACODE CAA 112(r) TQ
10034-93-2	HYDRAZINIUM SULFATE	1,00	Section 313 313 Section 302 (EHS) TPQ Section 304 EHS RQ CERCLA RQ RCRACODE CAA 112(r) TQ
10099-74-8	LEAD(II) NITRATE	-2,00	Section 313 313c Section 302 (EHS) TPQ Section 304 EHS RQ CERCLA RQ 10,00 RCRACODE CAA 112(r) TQ
106-42-3	P-XYLENE	-2,00	Section 313 X Section 302 (EHS) TPQ Section 304 EHS RQ CERCLA RQ 100,00 RCRACODE U239 CAA 112(r) TQ
107-06-2	1,2-DICHLOROETHANE	-22,00	Section 313 313 Section 302 (EHS) TPQ

Compound Selection Library Enrichment based on molecular shape descriptors calculated with Chemsript

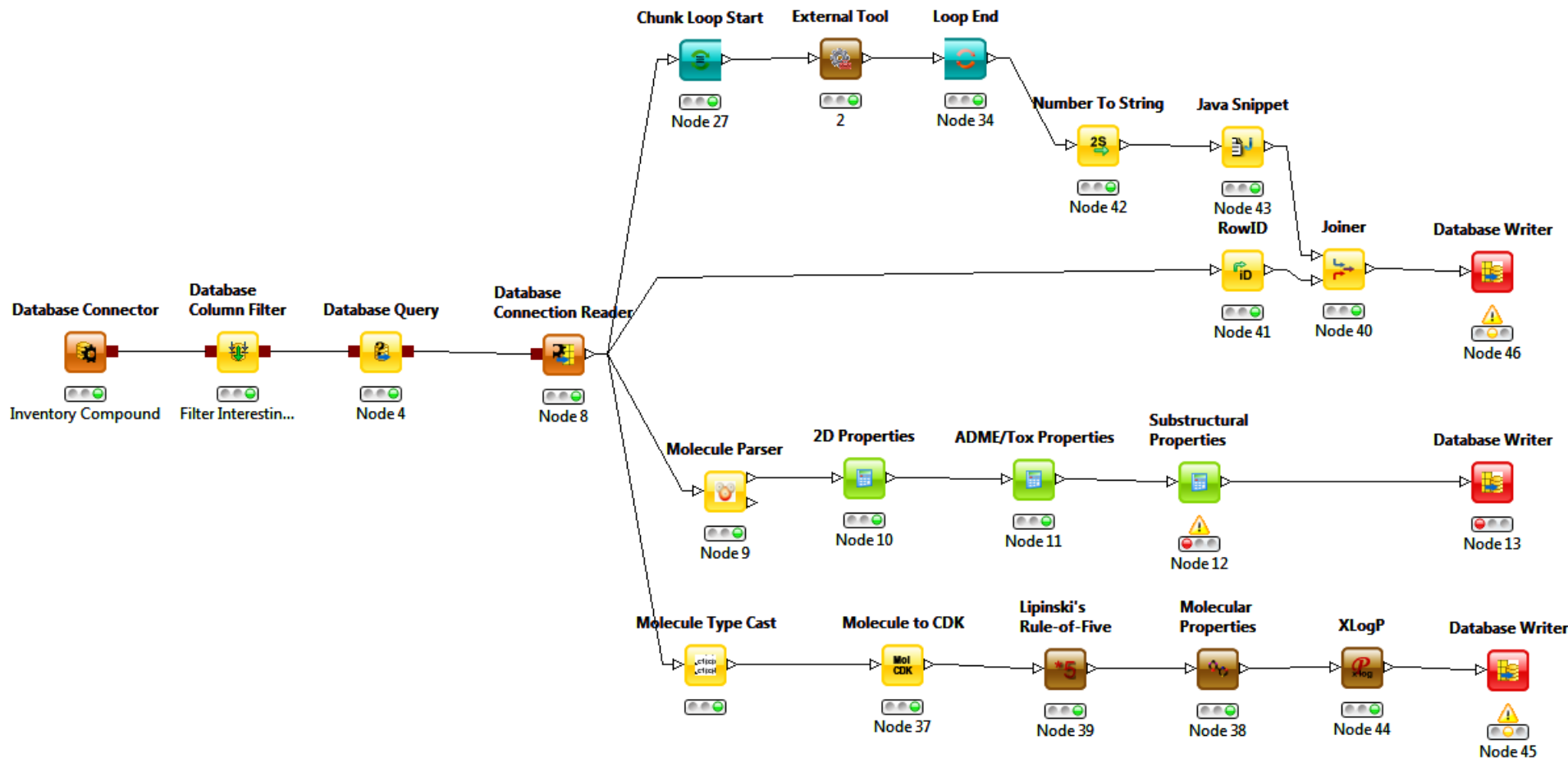


Project 3

- Company uses CambridgeSoft Enterprise registration system
- Now want to add predicted Lipinski Parameters
 - # H-Bond Donors, # H-Bond Acceptors, LogP, MW, # of Lipinsky violations as fields to be visible in side the Chemical registration system

Project 3

- Solution
- KNIME workflow + Oracle trigger



KNIME and CambridgeSoft

- CambridgeSoft core KNIME Team
 - Louis Culot – management-team sponsor
16 years CambridgeSoft
ex-Ciba-Geigy
 - Patrick Donovan
US-based
Materials science and chemistry background
5 years at CambridgeSoft
 - Dr. Bjoern Loeprecht
Former Tripos
Extensive chemistry and chem-informatics
 - Dr. Joe Durant
Former Symyx/MDL “KNIME” core developer
Responsible for CambridgeSoft KNIME R&D