CambridgeSoft®
Life Science Enterprise Solutions

KNIME

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

CambridgeSoft
Life Science Enterprise Solutions

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.

CambridgeSoft
Life Science Enterprise Solutions

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.

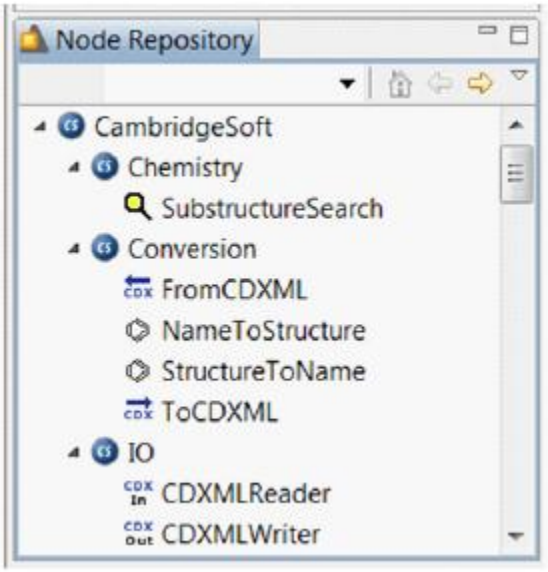
Most (>90%) consultants hold degrees in chemistry or biology in addition to software engineering.

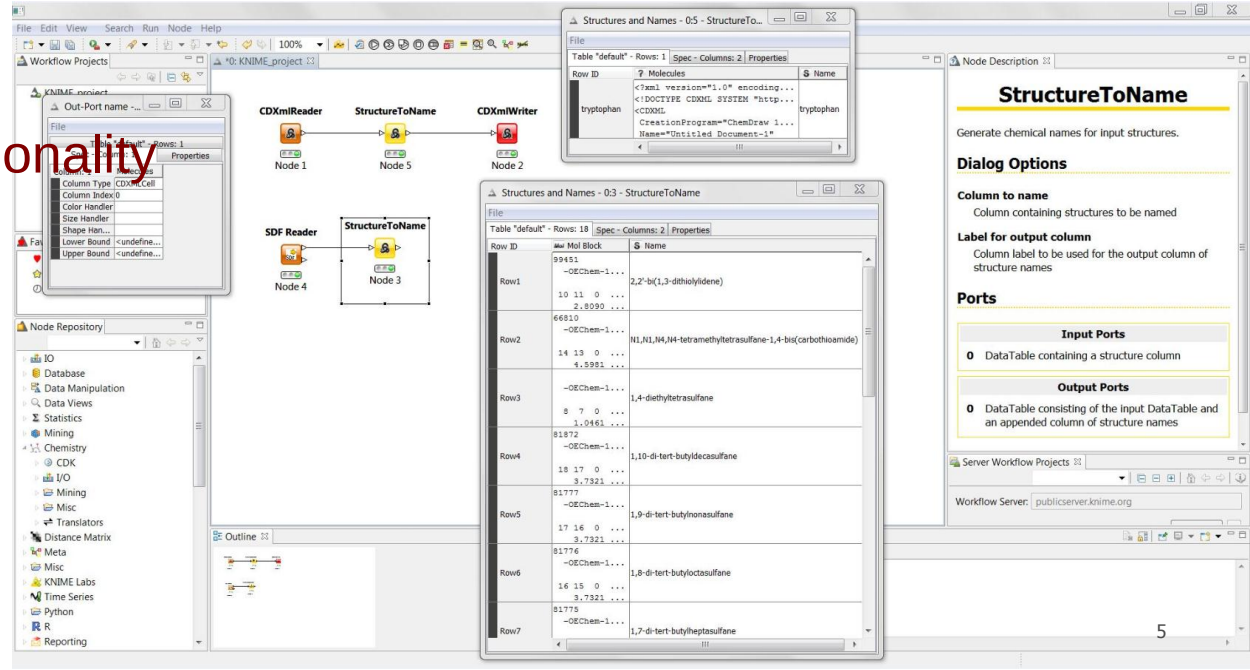
A large percentage of staff hold advanced degrees (Master's, Ph.D.).

Node Development

Cheminformatics Nodes

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

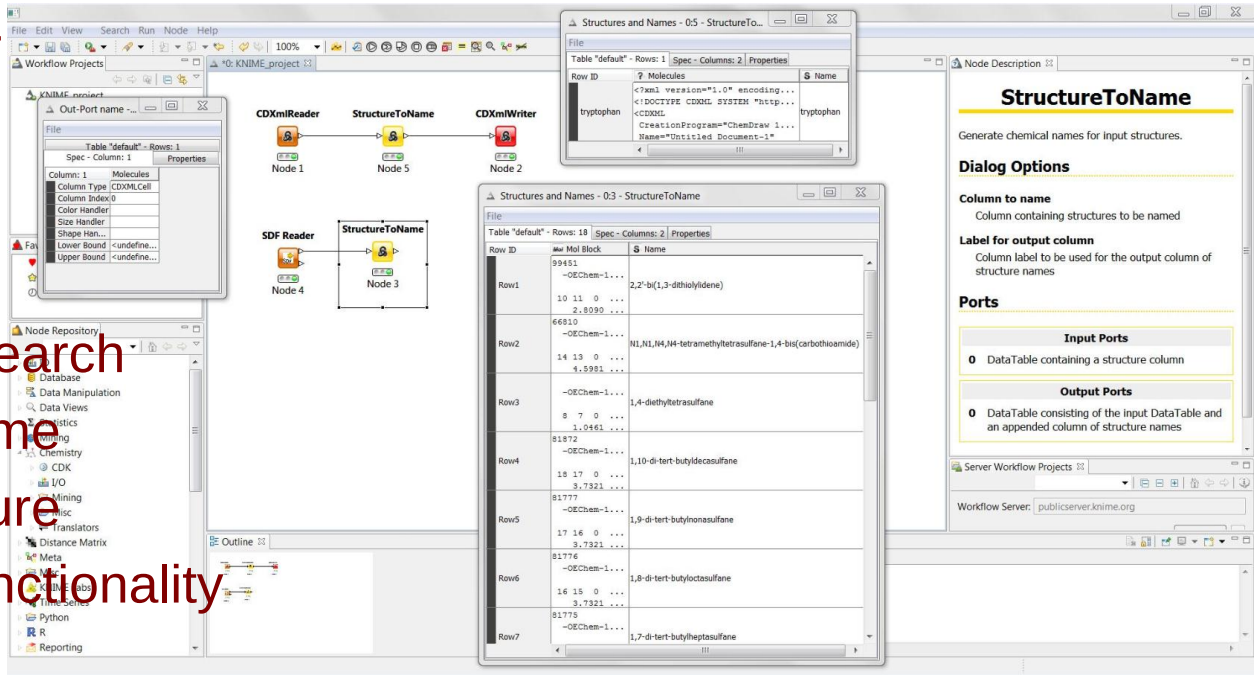




Node Development

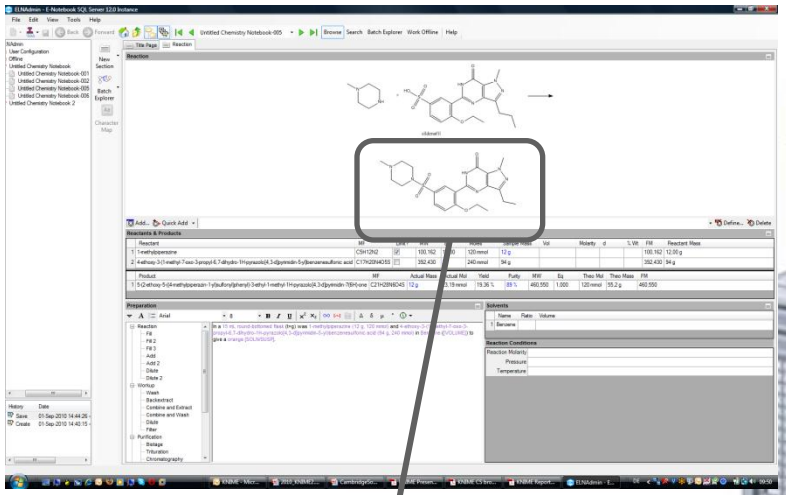
Cheminformatics Nodes

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



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

ChemBioOffice Enterprise & KNIME



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 – Retreve compounds and containers based on parameters
  - GetInventory – Retreive 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

CambridgeSoft
Life Science Enterprise Solutions

Enterprise Market Sample

INSTALLATIONS		RESEARCH				PRE-CLINICAL DEVELOPMENT					KNIME
	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.

CambridgeSoft
Life Science Enterprise Solutions

Enterprise Market Sample

INSTALLATIONS		RESEARCH				PRE-CLINICAL DEVELOPMENT					KNIME
	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’….

CambridgeSoft
Life Science Enterprise Solutions

Mountains of data…

Mountains of data…

+

KNIME

CambridgeSoft
Life Science Enterprise Solutions

Mountains of data…

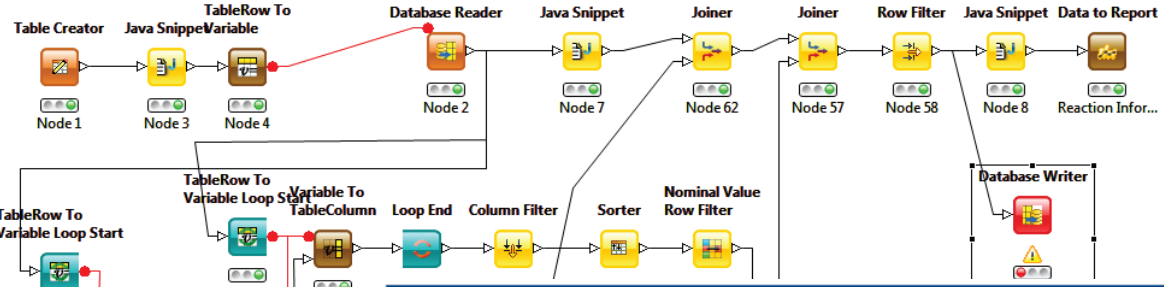
+

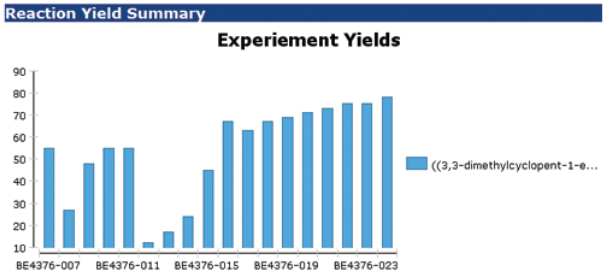
KNIME

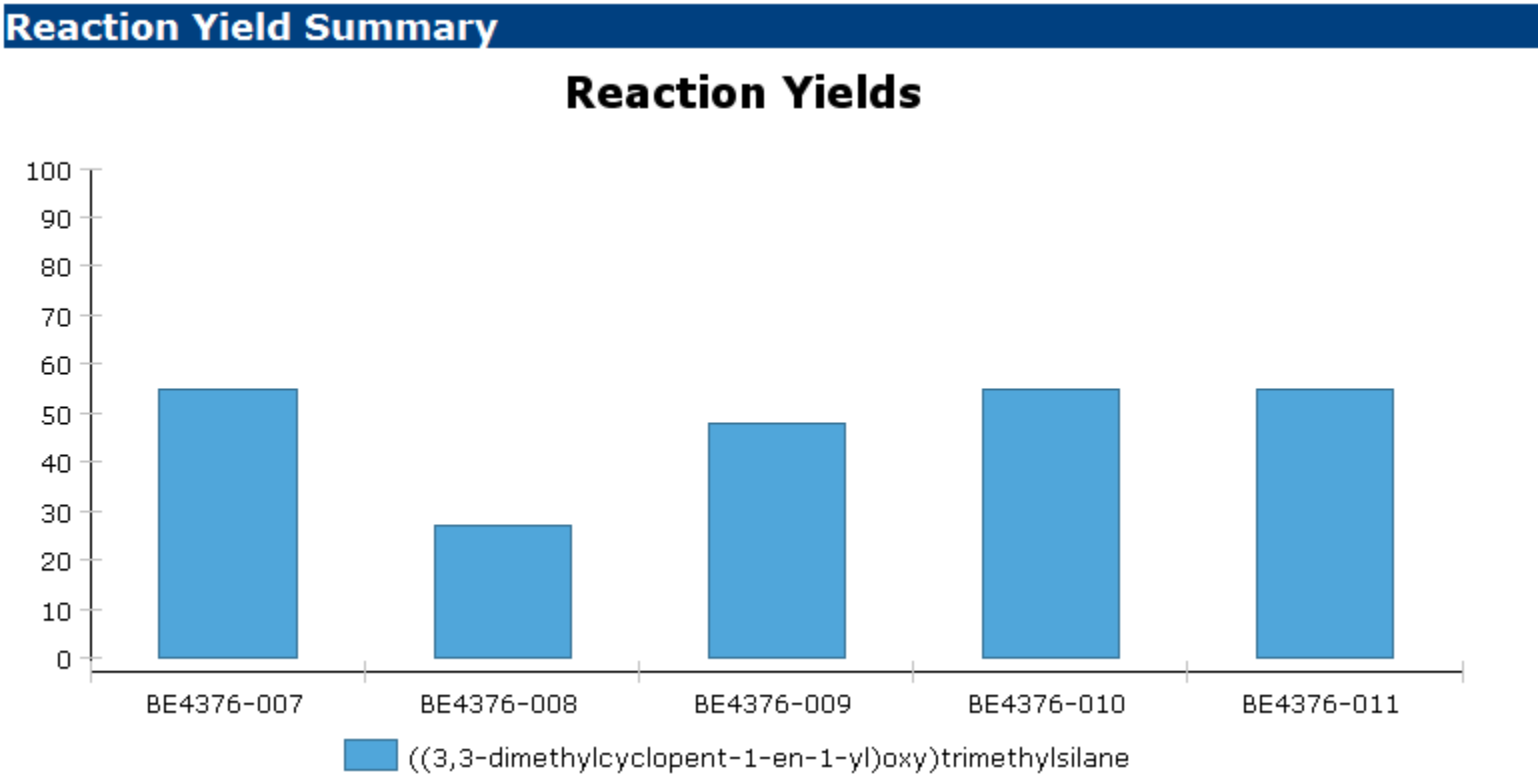
=

CambridgeSoft
Life Science Enterprise Solutions

E-Notebook Data Mining Workflow

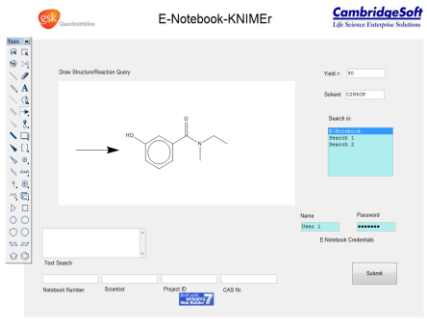






E-Notebook Mining System based on KNIME

Search Page



Search Manager

(7)

(1)

(6)

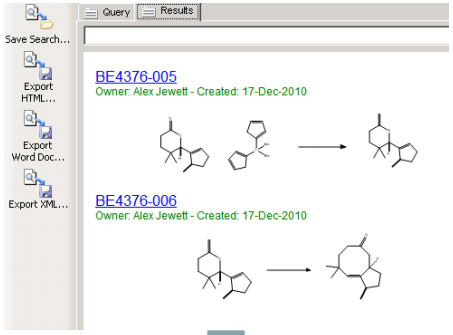
(5)

(2)

(4)

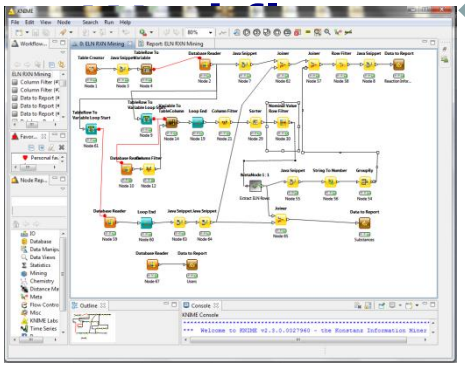
(3)

Results Page

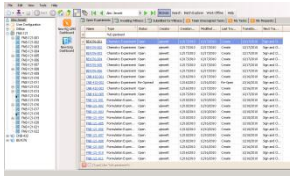


RD Export

KNIME



E-Notebook



(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

CambridgeSoft
Life Science Enterprise Solutions

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

KNIME Reporting Server - Report Listing - Windows Internet Explorer

http://192.168.80.129:10000/com.knime.reporting.server_2.2.2.0027174/listing.jsp

File Edit View Favorites Tools Help

Favorites | Beri... | CRM | Ente... | IThe... | Expe... | Intr... | Who'... | Goog... | Mehr... | Tele... | Inde... | Währ... | KNIM...

Welcome, Bjoern! | LinkedIn | Horde :: My Portal | R: Mahalanobis Distance | KNIME Reporting Serve...

KNIME.com

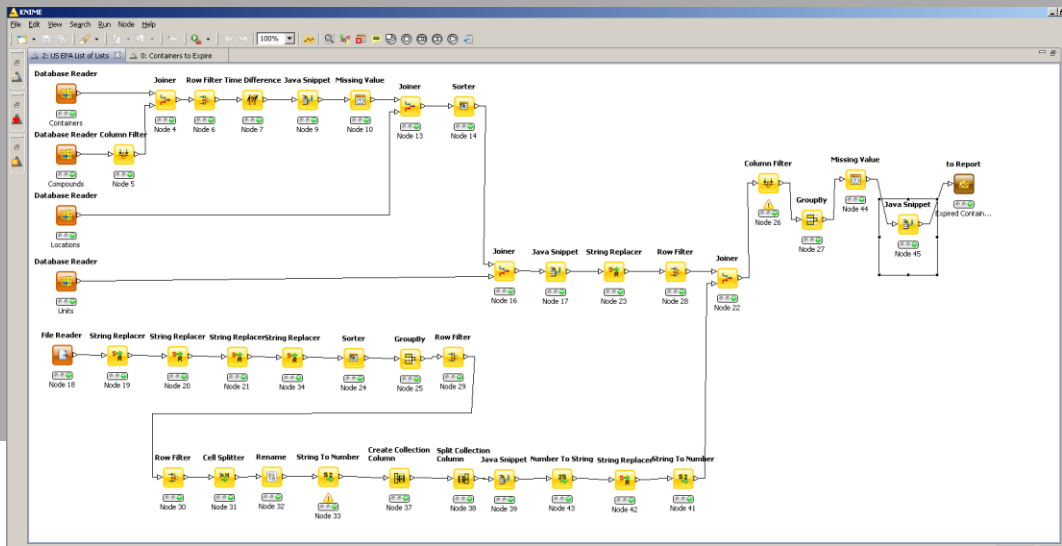
KNIME Reporting Server

Available Reports:

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

Logout

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



KNIME Report Viewer - Windows Internet Explorer

http://192.168.80.129:10000/com.knime.reporting.server_2.2.2.0027174/frameset?__report=US+EPA+List+of+Lists%2Fdefault_report.rptdesign

File Edit View Favorites Tools Help

KNIME Report Viewer

KNIME.com

KNIME Report Viewer

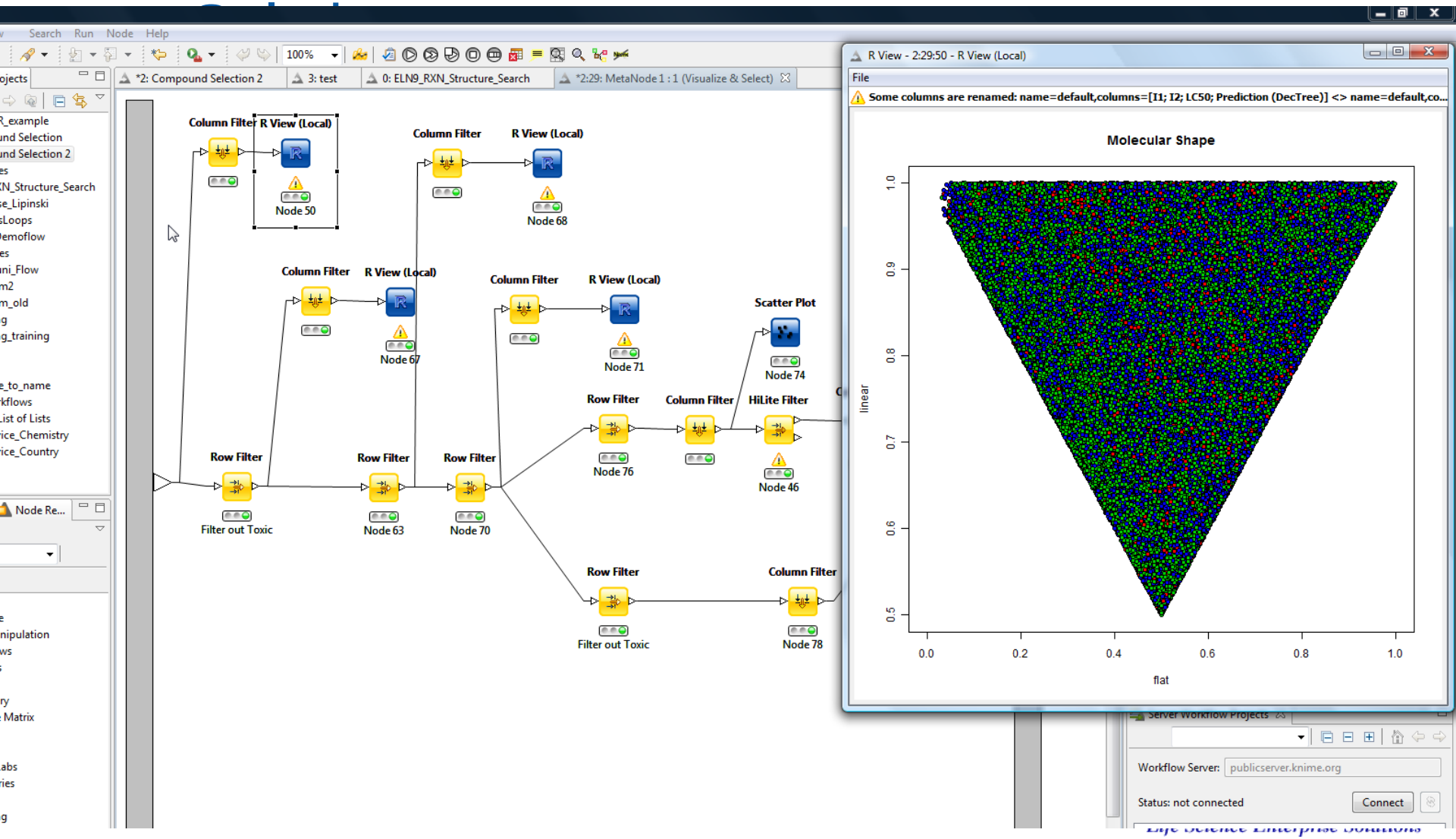
Showing page 1 of 2 Go to page:

US EPA List of Lists Report

CAS	SUBSTANCE_NAME	Sum(AMT_LBS)		
10026-13-8	PHOSPHORUS(V) CHLORIDE	800,00	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	

CambridgeSoft
Life Science Enterprise Solutions

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



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

CambridgeSoft Registration Enterprise

Query Input Form

Current Login: CSSADMIN

File History Queries Hit Lists Marked Hits (1) Help Log Off Home

Search | Clear Form | Retrieve All | Main Menu

Total Searchable Records: 5073

Registration Number (i.e. AB-000001)

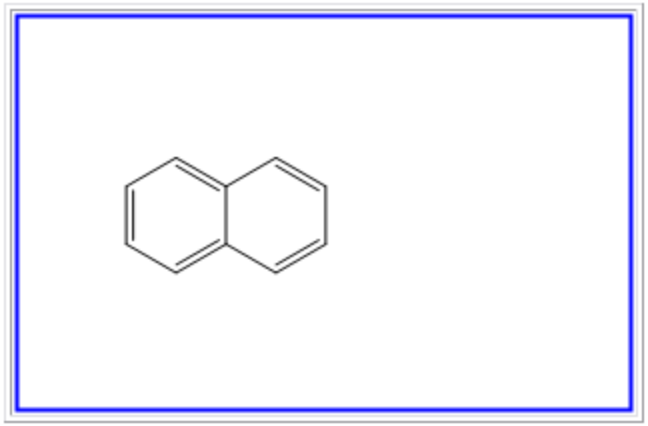
Sequence Number (i.e. 1 or >1)

Reg. Date (dd/mm/yyyy)

Compound Registrar -- Select one --

Last Mod Person -- Select one --

SearchType: Substructure



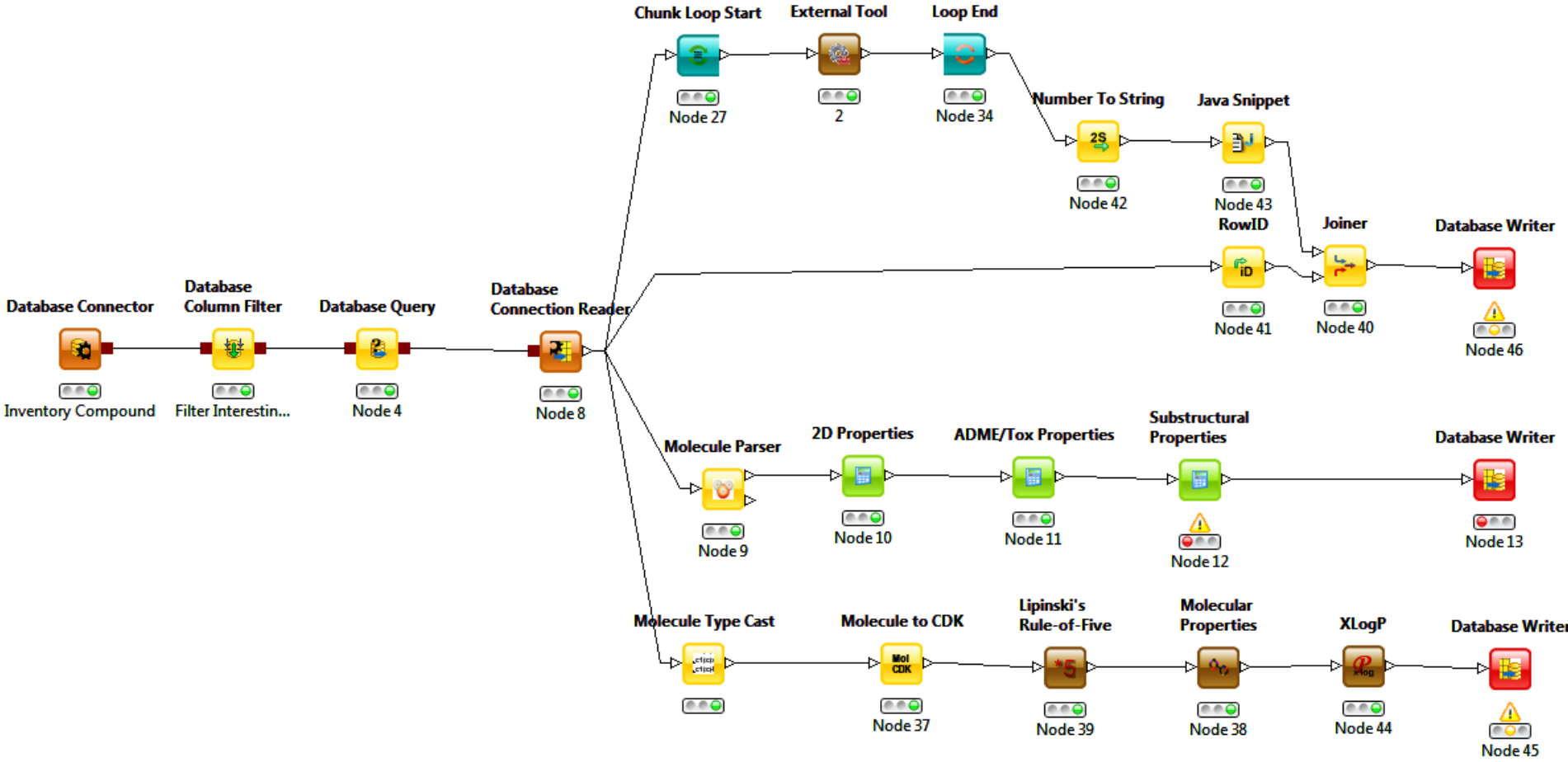
Compound Information

Project	-- Select one --
Prefix	-- Select one --
CAS No	
# Lipinski Violations	
MW (calc)	
MF (calc)	
CLogP	
#H Bond Donors	
#H Bond Acceptors	

CambridgeSoft
Life Science Enterprise Solutions

Project 3

- Solution
- KNIME workflow + Oracle trigger



CambridgeSoft
Life Science Enterprise Solutions

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