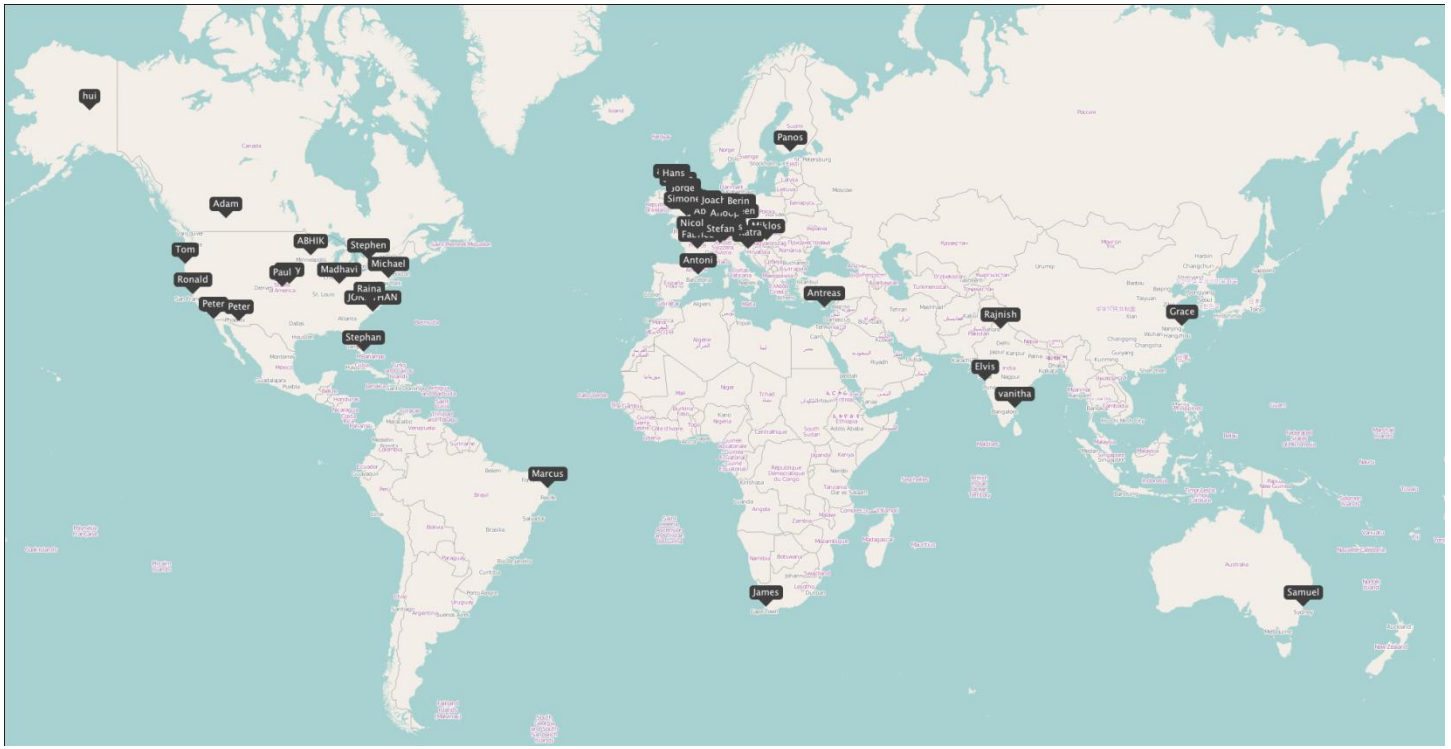
Chemistry in KNIME

For questions and suggestions please contact education@knime.com

Copyright © 2013 by KNIME.com AG. All rights reserved.

Welcome!



- Aaron Hart will be presenting Rosaria Silipo will be moderating.
- Please mute your mic and ask questions by chat.
- This meeting is being recorded, and the video will be posted soon.

Copyright © 2013 by KNIME.com AG. All rights reserved.

Goal of the Course

- Survey KNIME's basic concepts
- Sketch a chemical structure into a KNIME table
- Convert between chemistry types
- Read/write and manipulate in SDF
- Calculate molecular descriptors with RDKit
- Access ChEMBL via webservices (Optional)

Copyright © 2013 by KNIME.com AG. All rights reserved.

Course Material

Prerequisite -> www.knime.org

- KNIME 2.7 or later
- KNIME Extensions (Chemistry, including Marvin + Reporting)
- KNIME Community extensions (including EBI and RDKit nodes)

Supplemetary -> **www.knime.org/files/Intro_to_KNIME_Chem.zip**

- Data
- Example workflows
- These slides

Copyright © 2013 by KNIME.com AG. All rights reserved.

Additional Resources

- **KNIME** pages (www.knime.org)
  - **APPLICATIONS** for example workflows
  - **RESOURCES** with links to a number of resource pages, like downloads, updates, documentation, ...

- **KNIME Tech** pages (http://tech.knime.org/)
  - **FORUM** for questions and answers
  - **DOCUMENTATION** for documentation, FAQ, changelogs, ...
  - **LABS** where to find new experimental nodes
  - **COMMUNITY CONTRIBUTIONS** for development instructions and third party nodes

Copyright © 2013 by KNIME.com AG. All rights reserved.

The KNIME Workbench

Overview of the Platform

Copyright © 2013 by KNIME.com AG. All rights reserved.

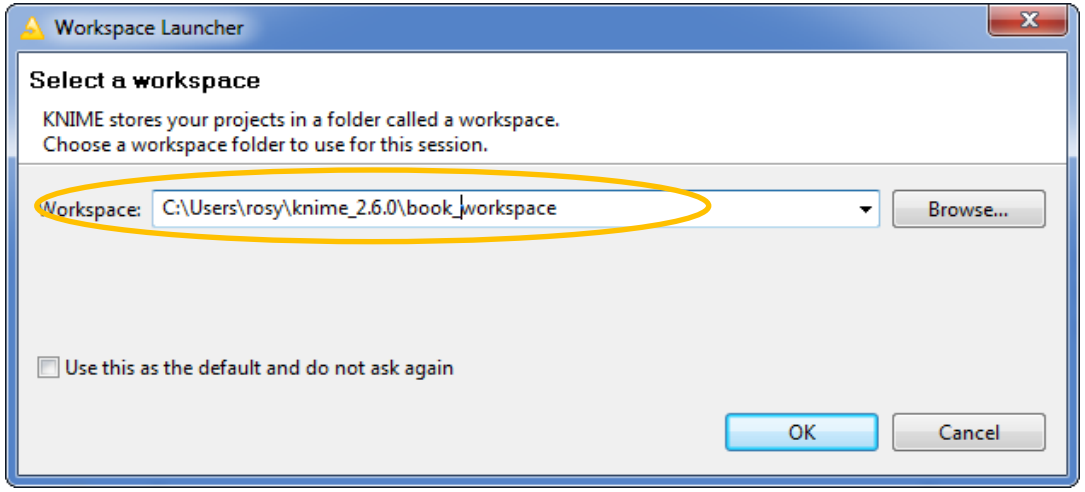
Overview

- The KNIME Workspace
- Installation of KNIME Extensions
- The KNIME Workbench
  - Help
  - Node Description

Copyright © 2013 by KNIME.com AG. All rights reserved.

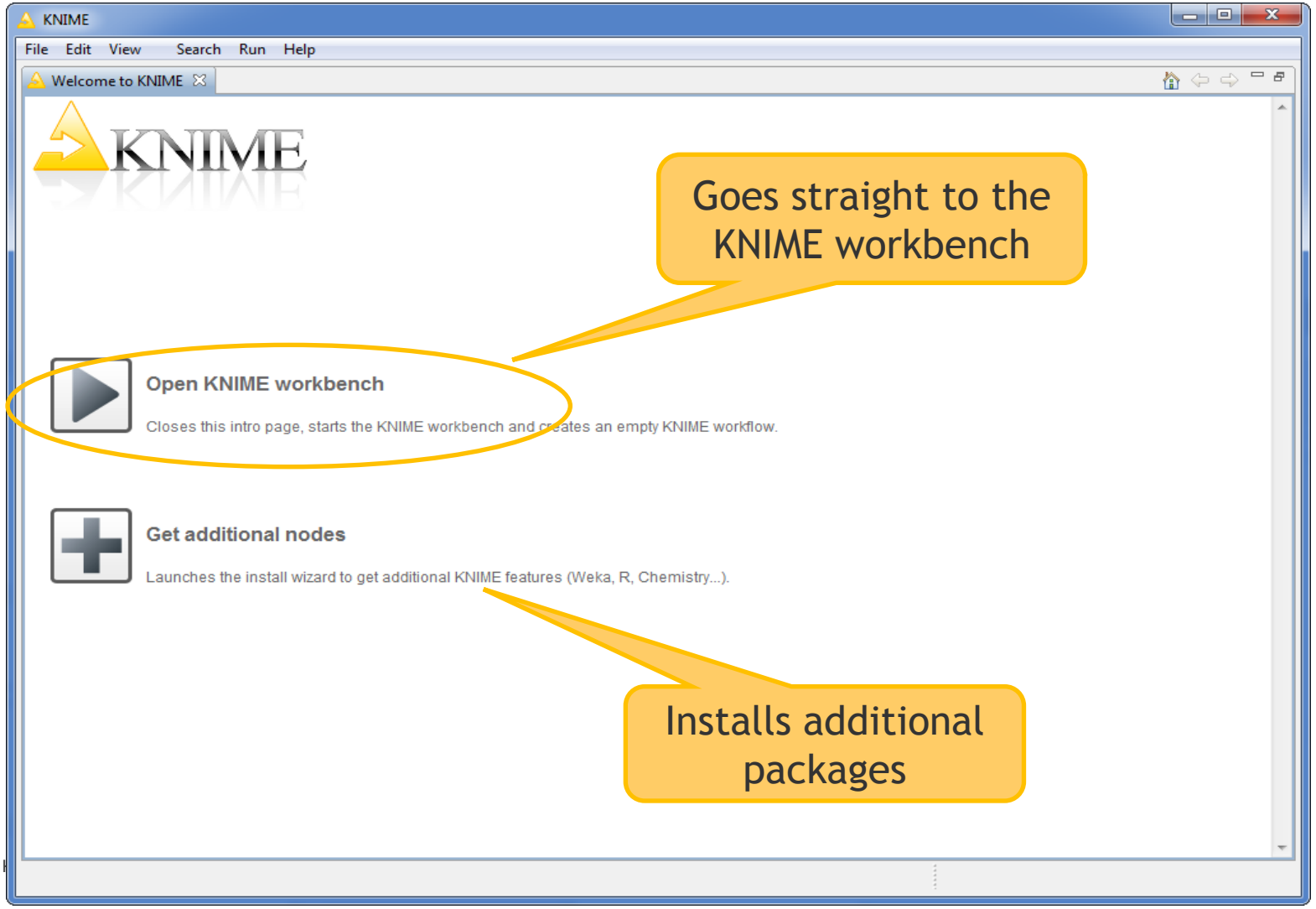
The Workspace

- The workspace is the **folder** in which work for the current KNIME session is stored.
- Using different workspaces for different projects keeps your KNIME work neat and clean.



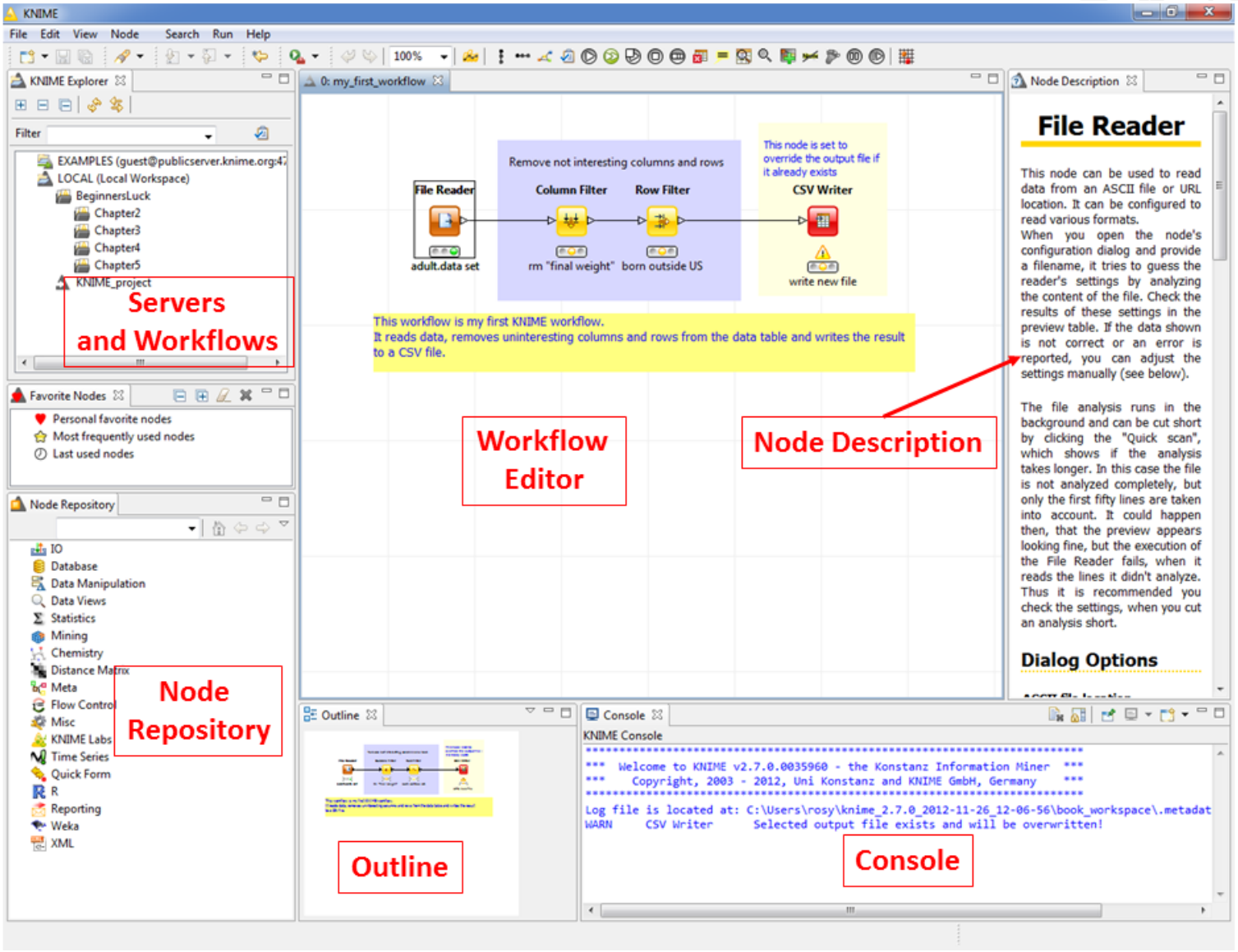
Copyright © 2013 by KNIME.com AG. All rights reserved.

First Time Start Screen



Copyright © 2013 by K

The Workbench



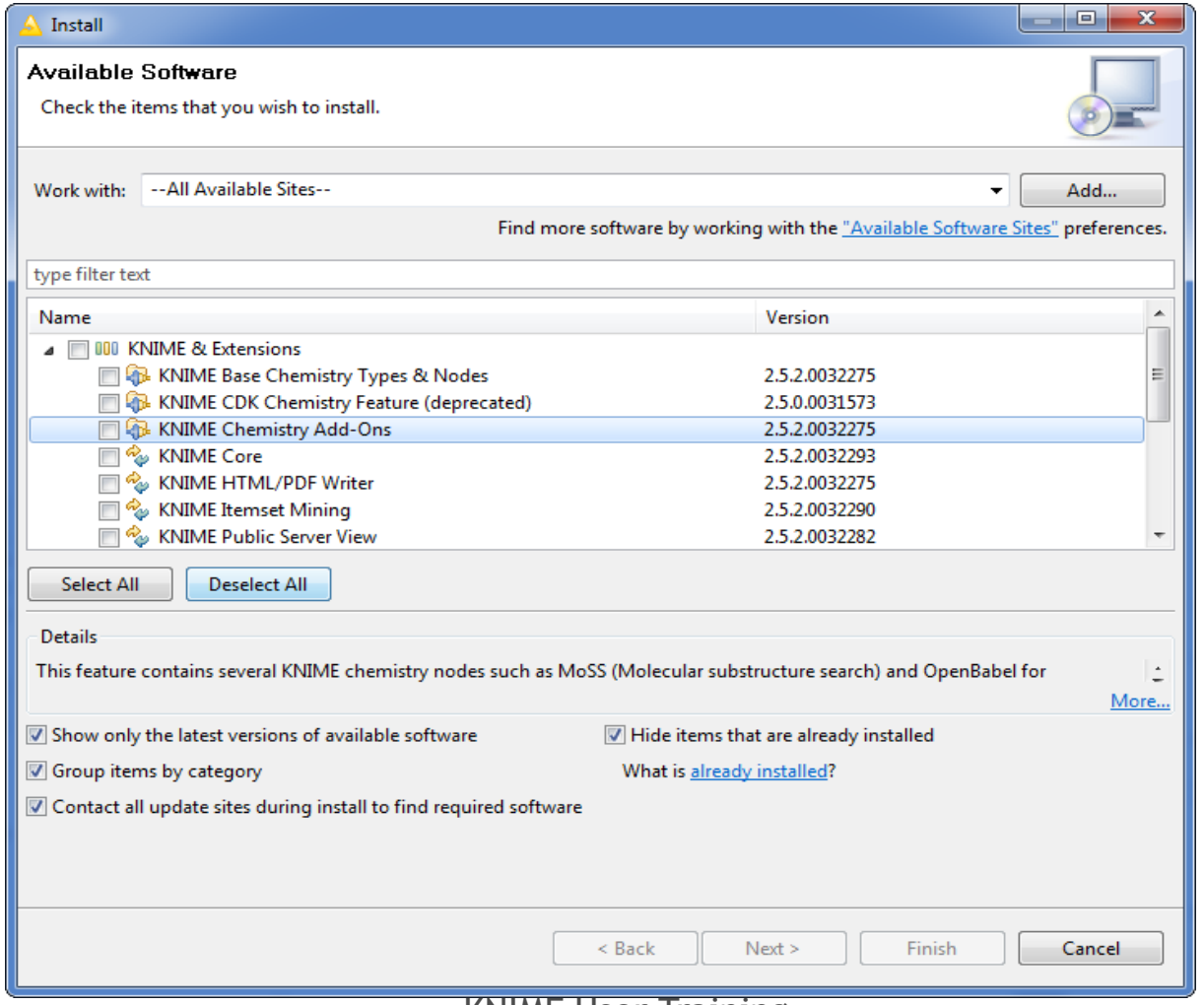
Copyright © 2013 by KNIME.com AG. All rights reserved.

Install KNIME Extensions

- Click Help -> Install New Software
- Work with --All Available Sites--
- To add new sites, click Add
- Expand node **KNIME & Extensions**
- Select your extensions.
- Click Next and follow the Wizard instructions

Copyright © 2013 by KNIME.com AG. All rights reserved.

Install KNIME Extensions



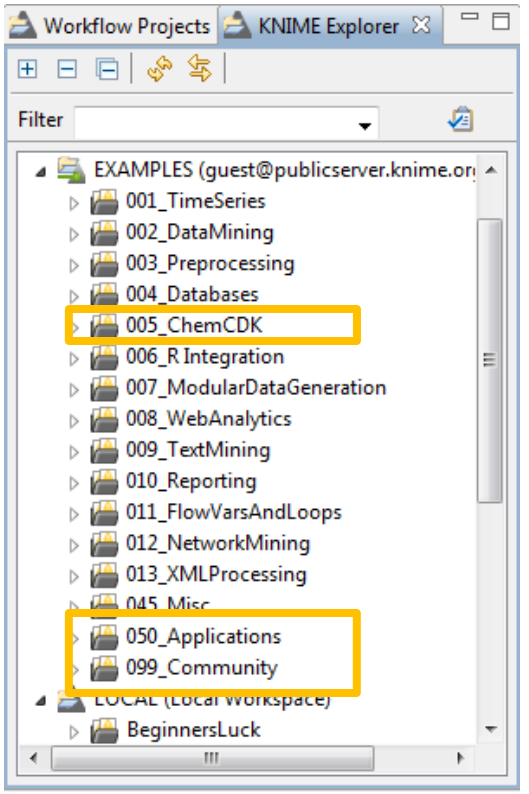
Copyright © 2013 by KNIME.com AG. All rights reserved.

The KNIME Public Server

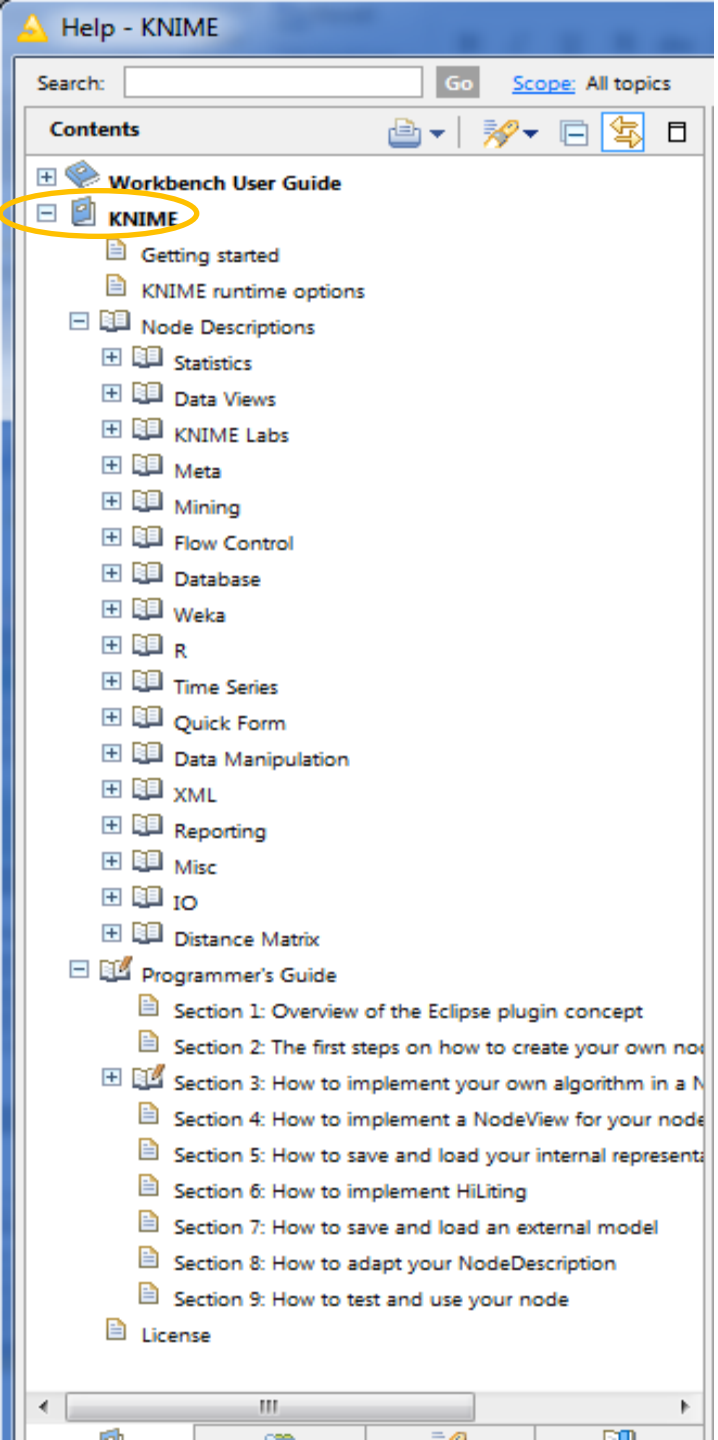
On the KNIME Public Server you can find a number of useful examples on how to implement specific tasks with KNIME.

In the KNIME Explorer panel:

- right click the public server
- select “Login”
- No login credentials required



Copyright © 2013 by KNIME.com AG. All rights reserved.



Help

Go to the Help menu, select Help Contents.

The KNIME section contains:

- A quick guide on getting started;
- The Node Descriptions chapter. This lists all the node descriptions that appear in KNIME when you select a node or category. You can use the search bar for a specific node description.
- The Programmer's Guide chapter gives details on how to create a new node for KNIME.

What is a Node?

A node is the **single processing unit** of a workflow.

A node takes a data set as input, processes it, and makes it available on its output port.

The “processing” action of a node can be:

- Reading a file
- Data manipulation
- Generating a plot
- Calculating properties or statistics
- .. and more

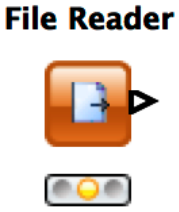
Nodes containing other nodes are called **meta nodes.**

Copyright © 2013 by KNIME.com AG. All rights reserved.

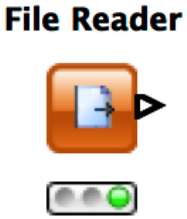
More on Nodes

A node has a one of three states:

Idle:
The node is not yet configured and can not be executed with it's current settings.



Configured:
The node has been set up correctly, and may be executed at any time



Executed:
The node has been successfully executed. Results may be viewed and used in downstream nodes.

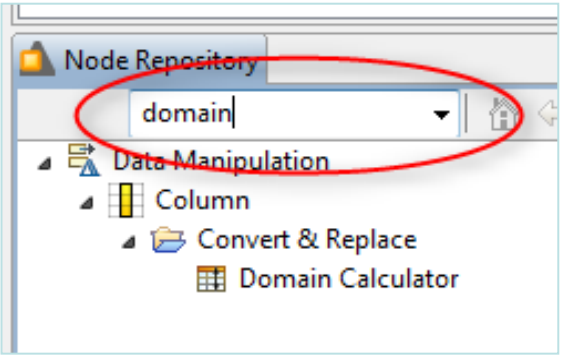
Copyright © 2013 by KNIME.com AG. All rights reserved.

Node Repository Search Box

The Node Repository contains a **search box**.

If you type a keyword in the search box and press Enter, you obtain the list of nodes with that keyword in the name.

Press the Esc key to see all nodes again.



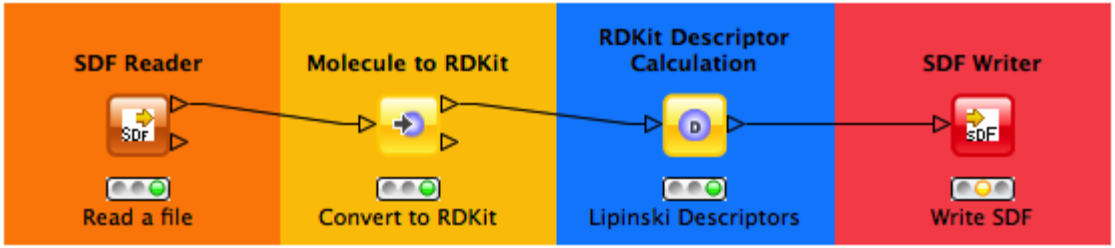
Copyright © 2013 by KNIME.com AG. All rights reserved.

What is a Workflow?

A workflow is an **analysis flow**,
i.e. a **sequence of analysis steps** necessary to reach a given result.

Example:

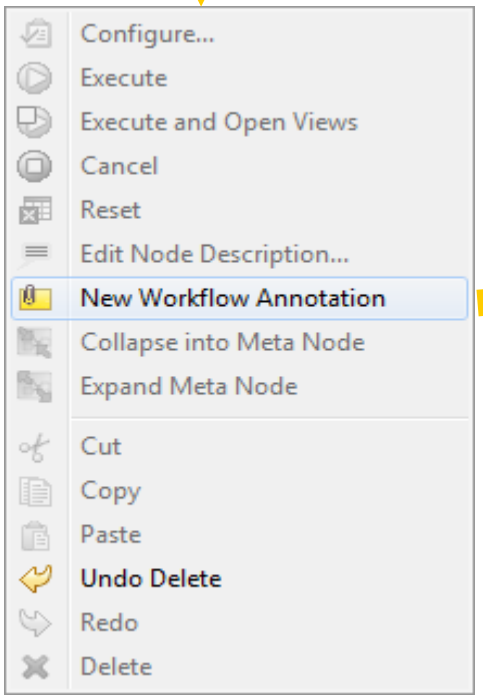
Step 1: Read data file
Step 2: Manipulate types
Step 3: Analyze
Step 4: Export Results



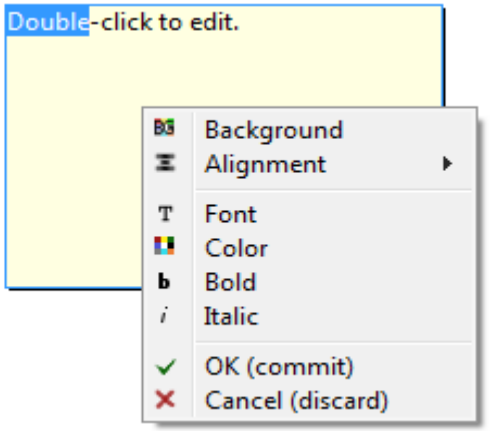
Copyright © 2013 by KNIME.com AG. All rights reserved.

Annotations

Right-click anywhere in the Workflow Editor panel.



Double-click the annotation to write text in it.



Context-menu of the annotation for the annotation editor

Copyright © 2013 by KNIME.com AG. All rights reserved.

Demo

- Survey KNIME Workbench
- Import the demo workflows

Copyright © 2013 by KNIME.com AG. All rights reserved.

Copyright © 2013 by KNIME.com AG. All rights reserved.

The KNIME Chemistry Extensions

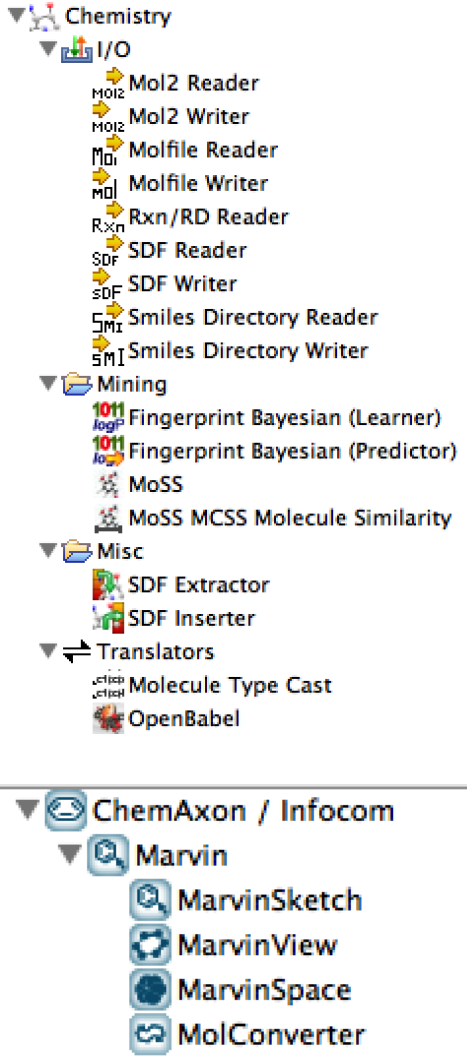
Copyright © 2013 by KNIME.com AG. All rights reserved.

Overview

- KNIME Core chemistry
    - Types, IO, translation, similarity, modeling
- Community Contributions
    - RDKit, CDK, Indigo, Erl Wood etc.
- Commercial Extensions
    - CCG MOE, Chemaxon, Schrodinger etc.

Copyright © 2013 by KNIME.com AG. All rights reserved.

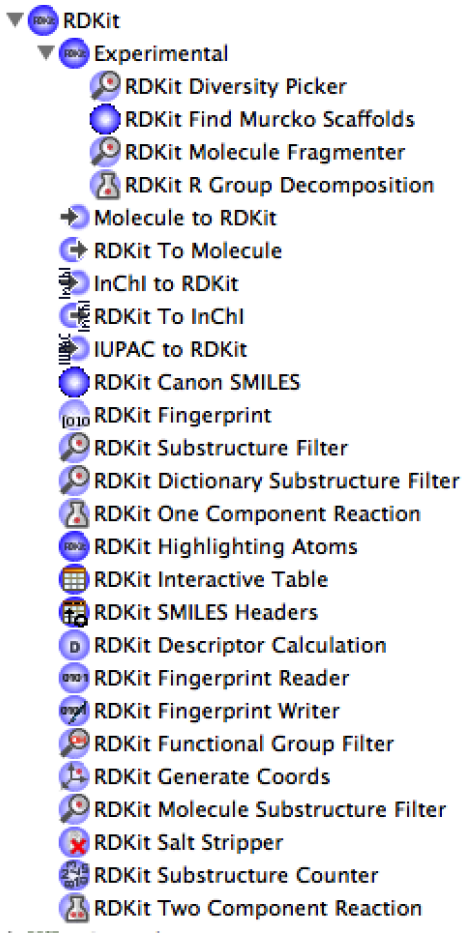
KNIME Core Chemistry

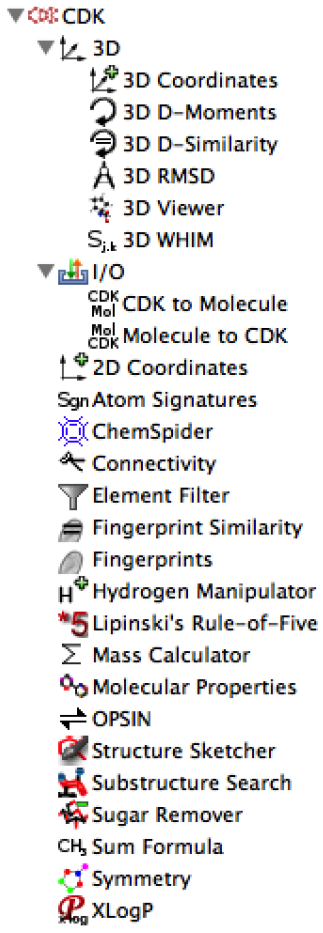


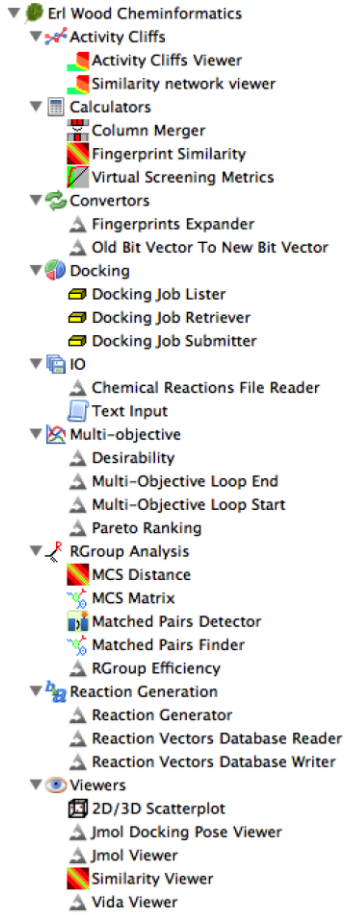
- IO
  - Readers and writers for Mol2, Mol, Rxn, SDF and Smiles.
- Mining
  - Fingerprint learner/predictor
  - MoSS & Moss MCSS similarity
- Misc
  - SDF Property extraction and Insertion
- Translators
  - Molecule Type Cast converts strings to native KNIME chemistry types
  - OpenBabel translates between chemistry types
- Marvin
  - Sketch, Convert and View (2d and 3d)

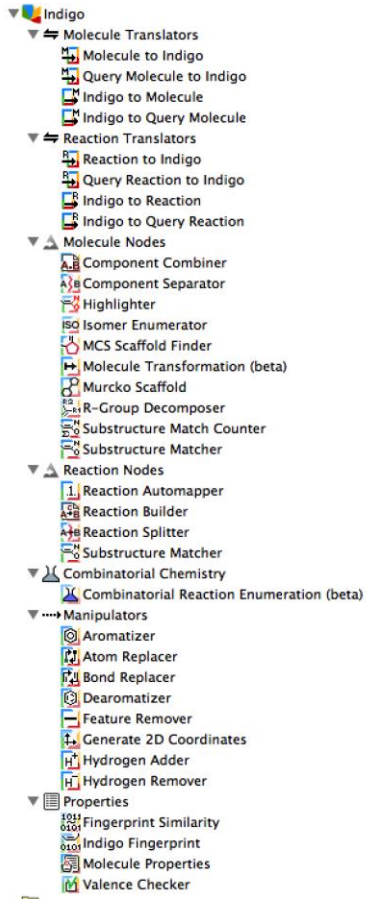
Copyright © 2013 by KNIME.com AG. All rights reserved.

Chemistry Extensions from the Community
(a few tools…)



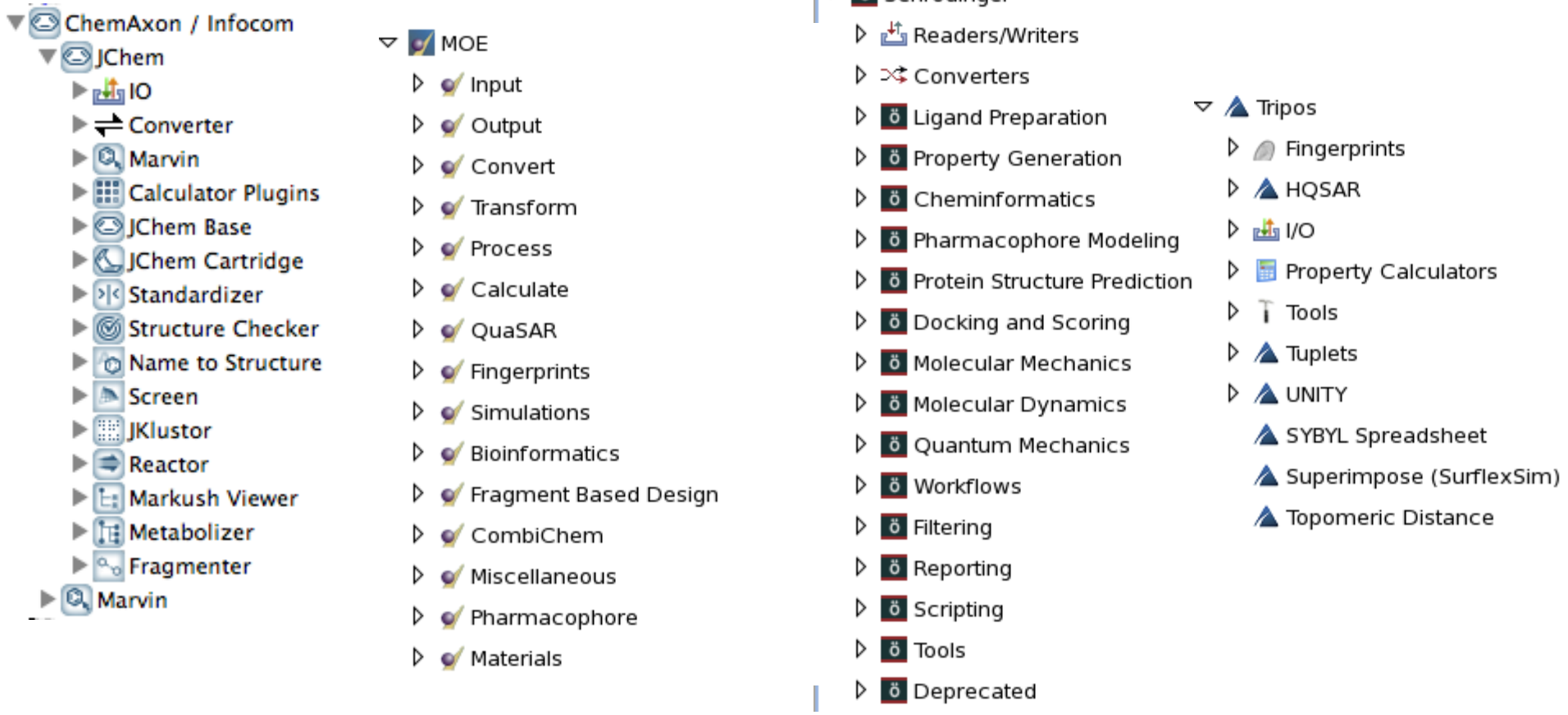






Copyright © 2013 by KNIME.com AG. All rights reserved.

Commercial Chemistry Extensions
(and even more…)



Copyright © 2013 by KNIME.com AG. All rights reserved.

Copyright © 2013 by KNIME.com AG. All rights reserved.

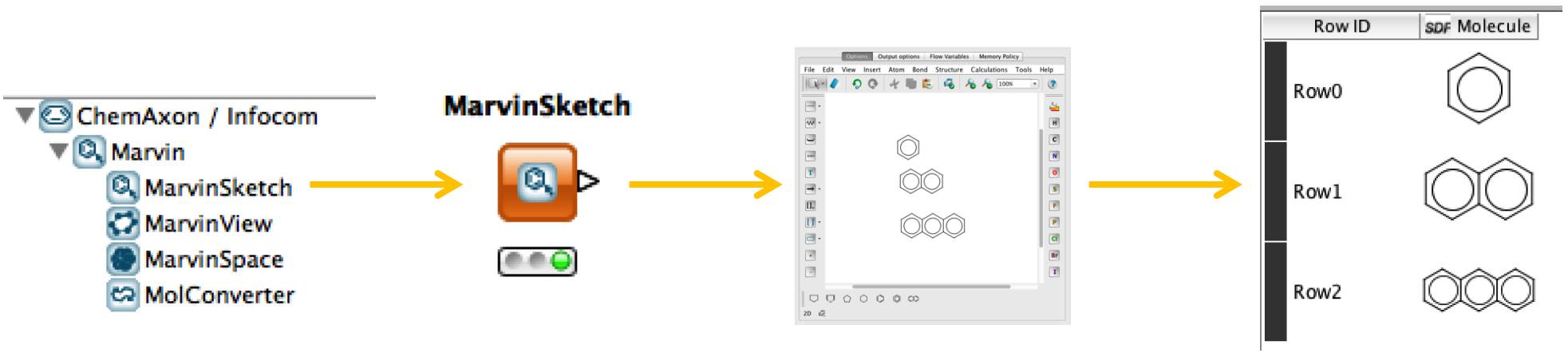
My First Molecule

(Sketching in KNIME)

Copyright © 2013 by KNIME.com AG. All rights reserved.

The Marvin Sketcher

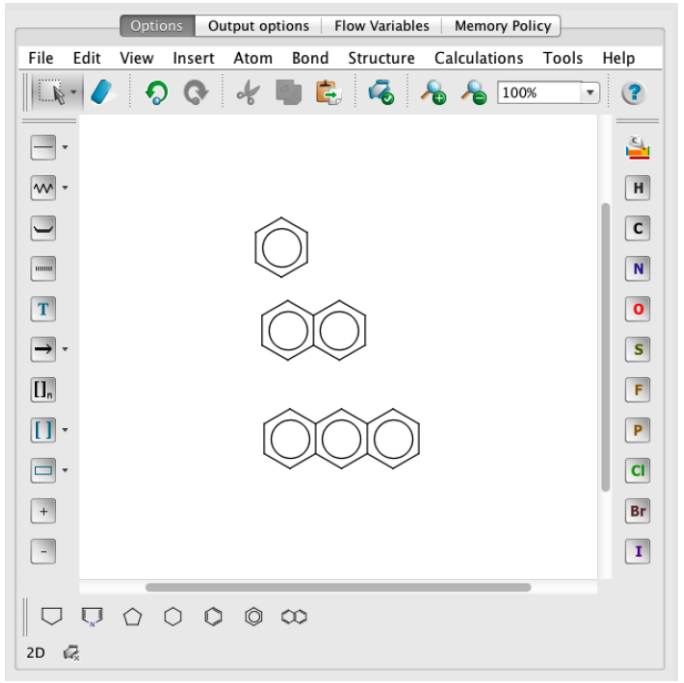
- Available from the KNIME Update Site
- Donated by Chemaxon/Infocom
- Integrates with Webportal to sketch in a browser

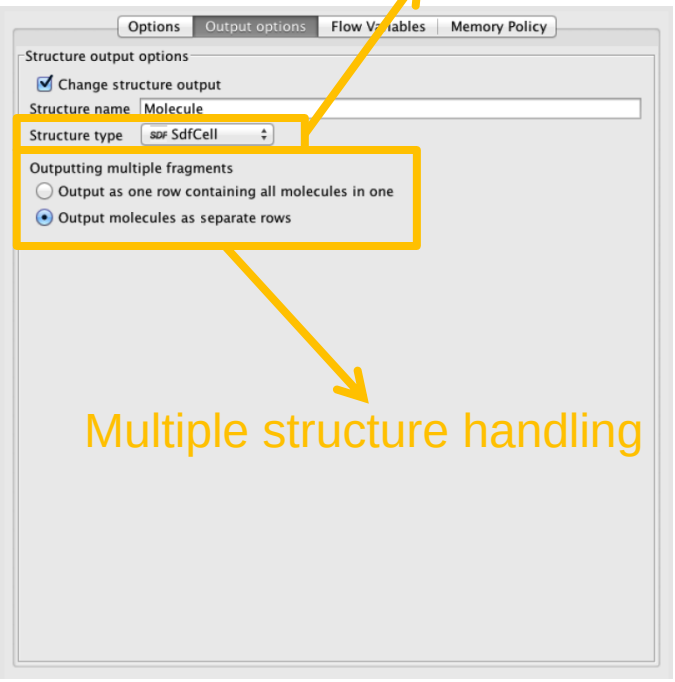


Copyright © 2013 by KNIME.com AG. All rights reserved.

The Marvin Sketcher II

- Full featured Sketcher
- Advanced options shown on in additional tab





Copyright © 2013 by KNIME.com AG. All rights reserved.

DEMO

- Sketch some molecules, export as an SDF cell, generate a new data row for each molecule

Copyright © 2013 by KNIME.com AG. All rights reserved.

Copyright © 2013 by KNIME.com AG. All rights reserved.

Chemistry Data Basics

Types and IO

Copyright © 2013 by KNIME.com AG. All rights reserved.

Chemistry Types Overview

- KNIME Basic types
  - String, Int, Double

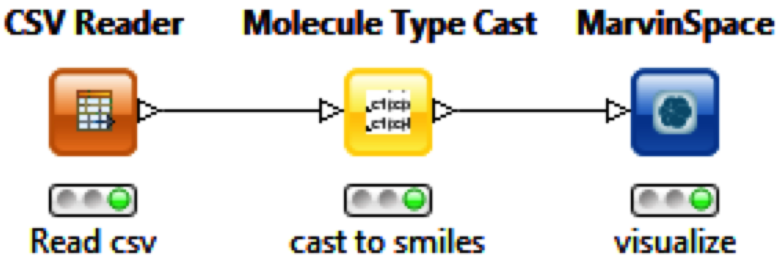
- KNIME Core Chemistry Types
  - SDF, MOL, MOL2, Smiles
  - These types can be rendered in KNIME Tables

- Special nodes convert between types.

Copyright © 2013 by KNIME.com AG. All rights reserved.

Node: Molecule Type Cast

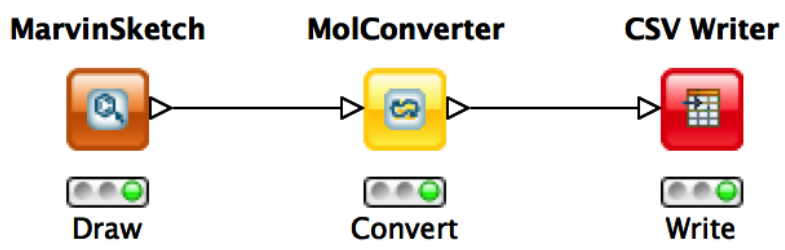
- KNIME Core Chemistry
- Convert string cells to a renderable smiles type
- Use to convert structures from string text files or databases to smiles cells.

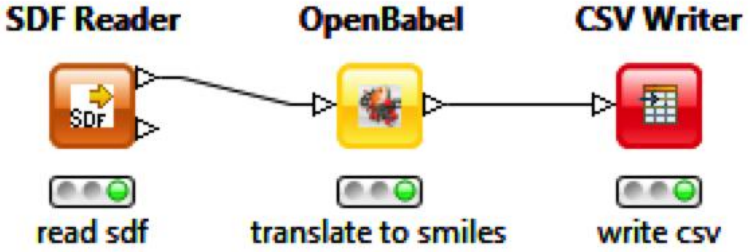


Copyright © 2013 by KNIME.com AG. All rights reserved.

Nodes: MolConverter & OpenBabel

- KNIME Core Chemistry
- Convert to/from virtually any format
- MolConverter provided by Chemaxon/Infocom
- www.openbabel.org





Copyright © 2013 by KNIME.com AG. All rights reserved.

Chemistry IO

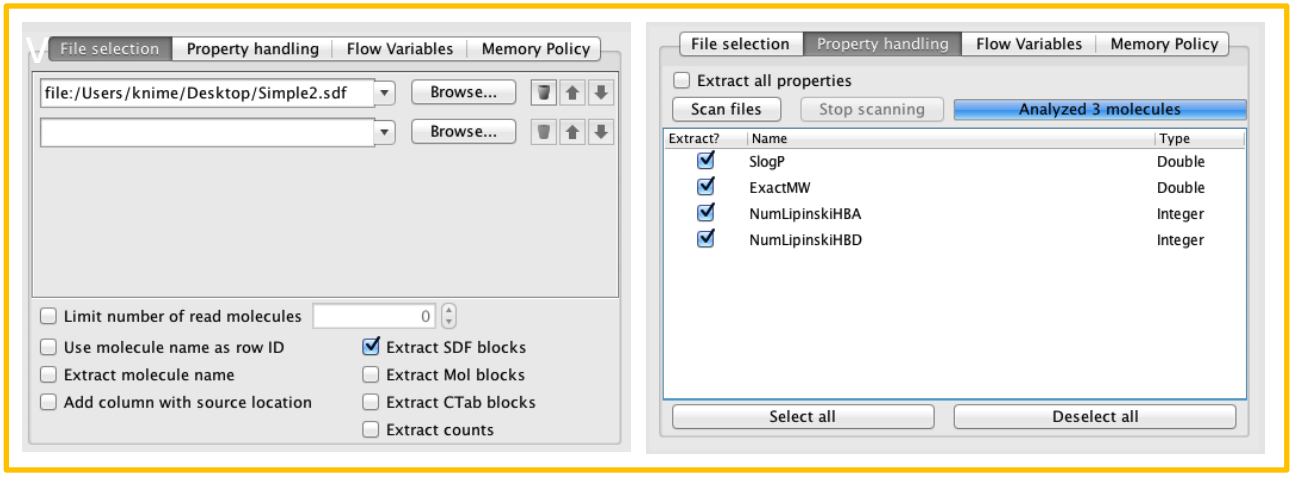
- File Readers and writers available for:
  - SDF, MOL, MOL2, Smiles

Copyright © 2013 by KNIME.com AG. All rights reserved.

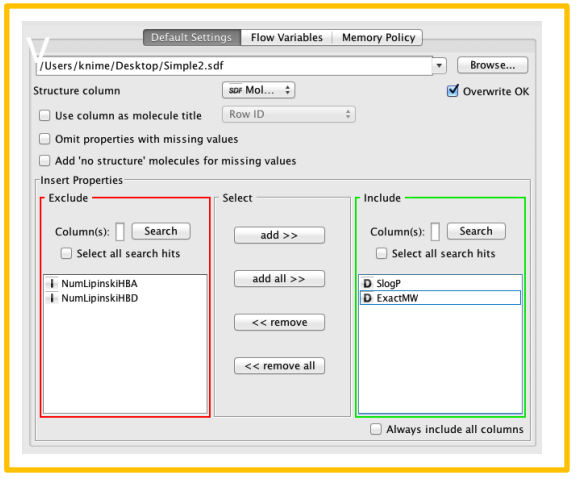
Chemistry IO II: SDF

- SDF Reader can import properties
- SDF Writer can write columns as properties
- Dedicated nodes also exist for this (SDF Inserter and SDF Extractor)

SDF Reader



SDF Writer



Copyright © 2013 by KNIME.com AG. All rights reserved.

Demo

- Create new workflow named Exercise 3
- Read structures from 124mols.sdf
- Write smiles to csv
- Read resulting csv
- Cast molecule strings to smiles and translate to sdf
- Visualize in report with BIRT

Copyright © 2013 by KNIME.com AG. All rights reserved.

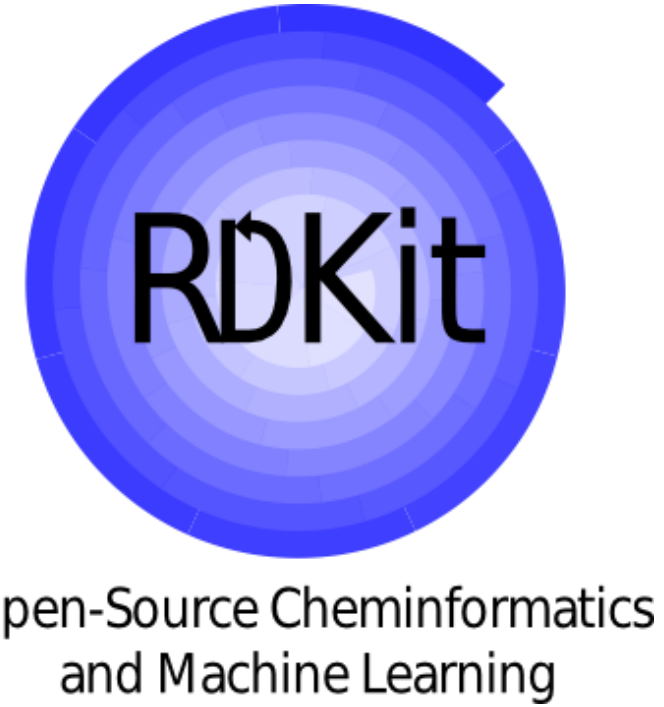
Copyright © 2013 by KNIME.com AG. All rights reserved.

Molecular Descriptors

An RDKit example

Overview

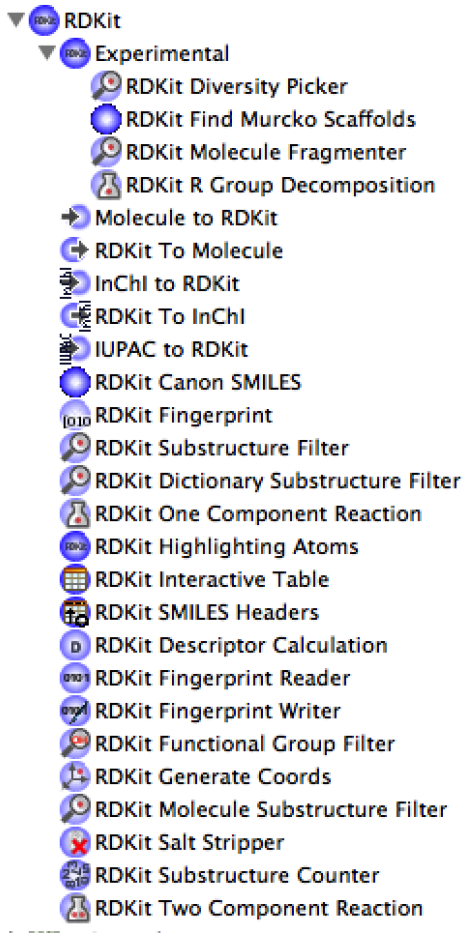
- Intro to RDKit
- Special Type Conversion
- Molecular properties
- InChI keys
- Smiles



Copyright © 2013 by KNIME.com AG. All rights reserved.

RDKit in KNIME

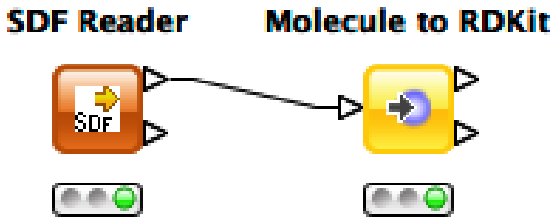
- Cheminformatics and Machine Learning
- Written in C++ and python.
- www.rdkit.org for more information

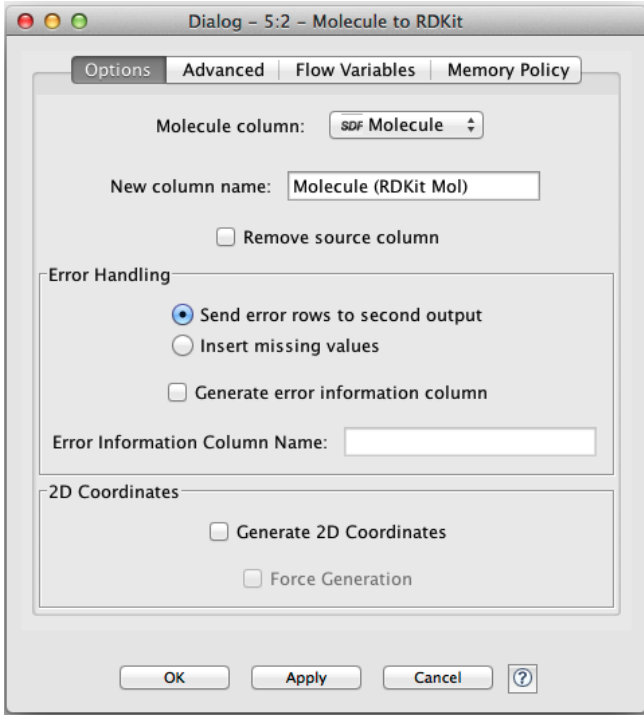


Copyright © 2013 by KNIME.com AG. All rights reserved.

Node: Molecule to RDKit

- Input smiles or SDF
- Creates RDKit data column
- Required before most RDKit nodes

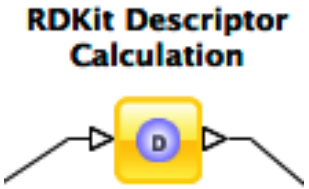




Copyright © 2013 by KNIME.com AG. All rights reserved.

Node: RDKit Descriptor Calculator

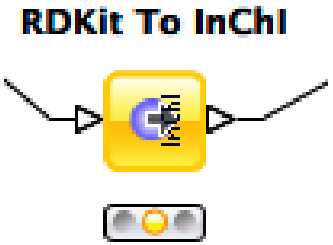
- Input RDKit Molecules
- Calculates many descriptors
  - MW, logp, h-bond donors etc.

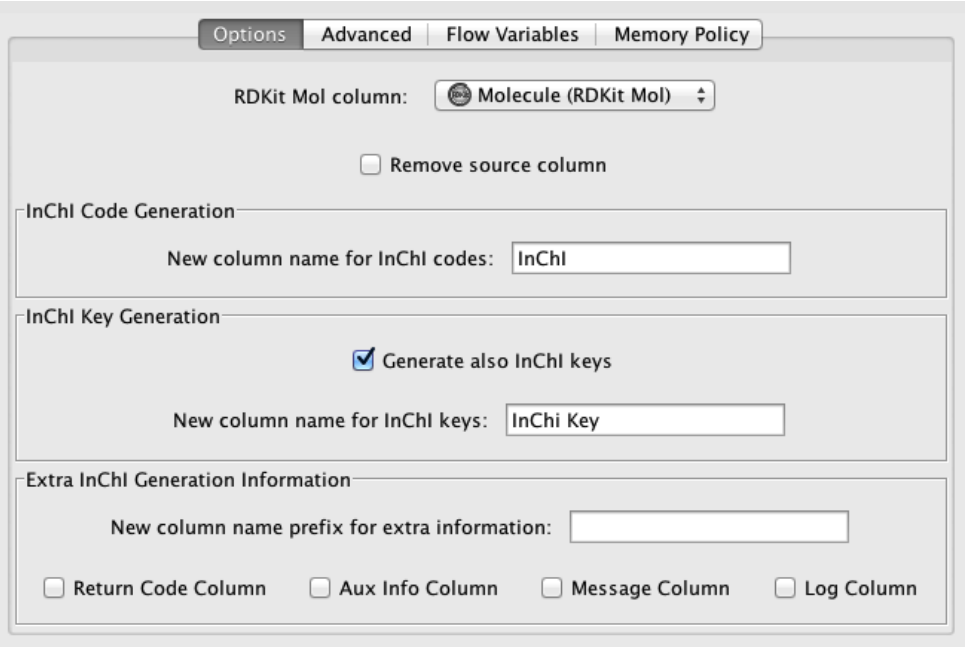


Copyright © 2013 by KNIME.com AG. All rights reserved.

Node: InChI Keys

- Input RDKit Molecules
- Generates InChI Keys

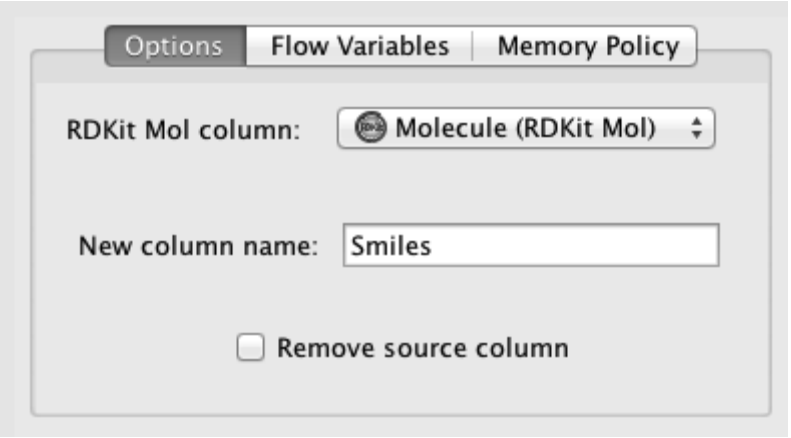




Copyright © 2013 by KNIME.com AG. All rights reserved.

NodeRDKit Cannon Smiles

- Input RDKit Molecule
- Generates cannonical smiles



Copyright © 2013 by KNIME.com AG. All rights reserved.

Demo

- Read structures from 124mols.sdf
- Convert to RDKit
- Calculate Properties
- Generate InChI keys
- Generate smiles
- Write to SDF

Copyright © 2013 by KNIME.com AG. All rights reserved.

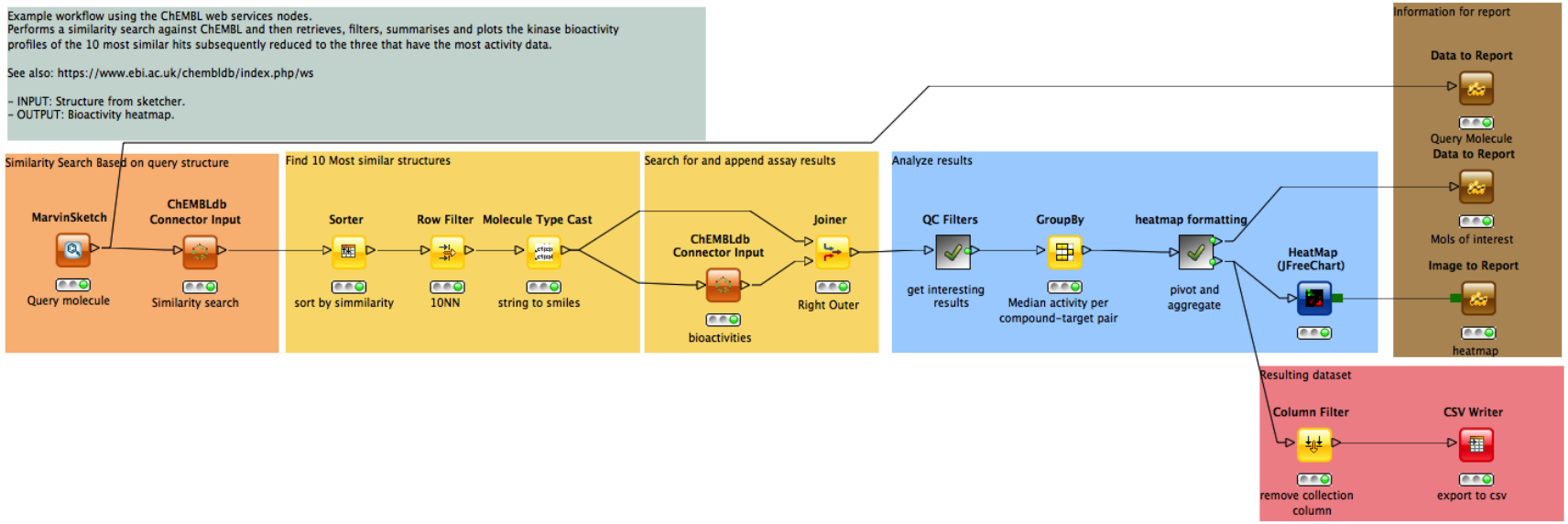
Copyright © 2013 by KNIME.com AG. All rights reserved.

KNIME Chemistry

Similarity and Bioactivity from ChEMBL

Workflow Background

- Donated by EBI (Along with the nodes!)
- Similarity search and bioactivty query with ChEMBL Webservices nodes
- Available on the public examples server.



Copyright © 2013 by KNIME.com AG. All rights reserved.

ChEMBL Webservices Demo

- Modify the workflow to include assay data from only the 3 molecules with the most activity data.

Copyright © 2013 by KNIME.com AG. All rights reserved.

Copyright © 2013 by KNIME.com AG. All rights reserved.

Thank you

education@knime.com

Copyright © 2013 by KNIME.com AG. All rights reserved.