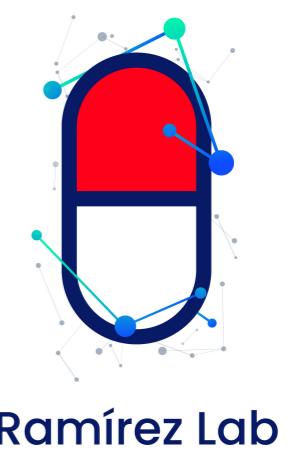


KNIME: Pharmacoinformatics and Drug Design

David Ramirez: dramirezs@udec.cl



*Departamento de Farmacología
Facultad de Ciencias Biológicas
Universidad de Concepción
Web Lab: ramirezlab.github.io*



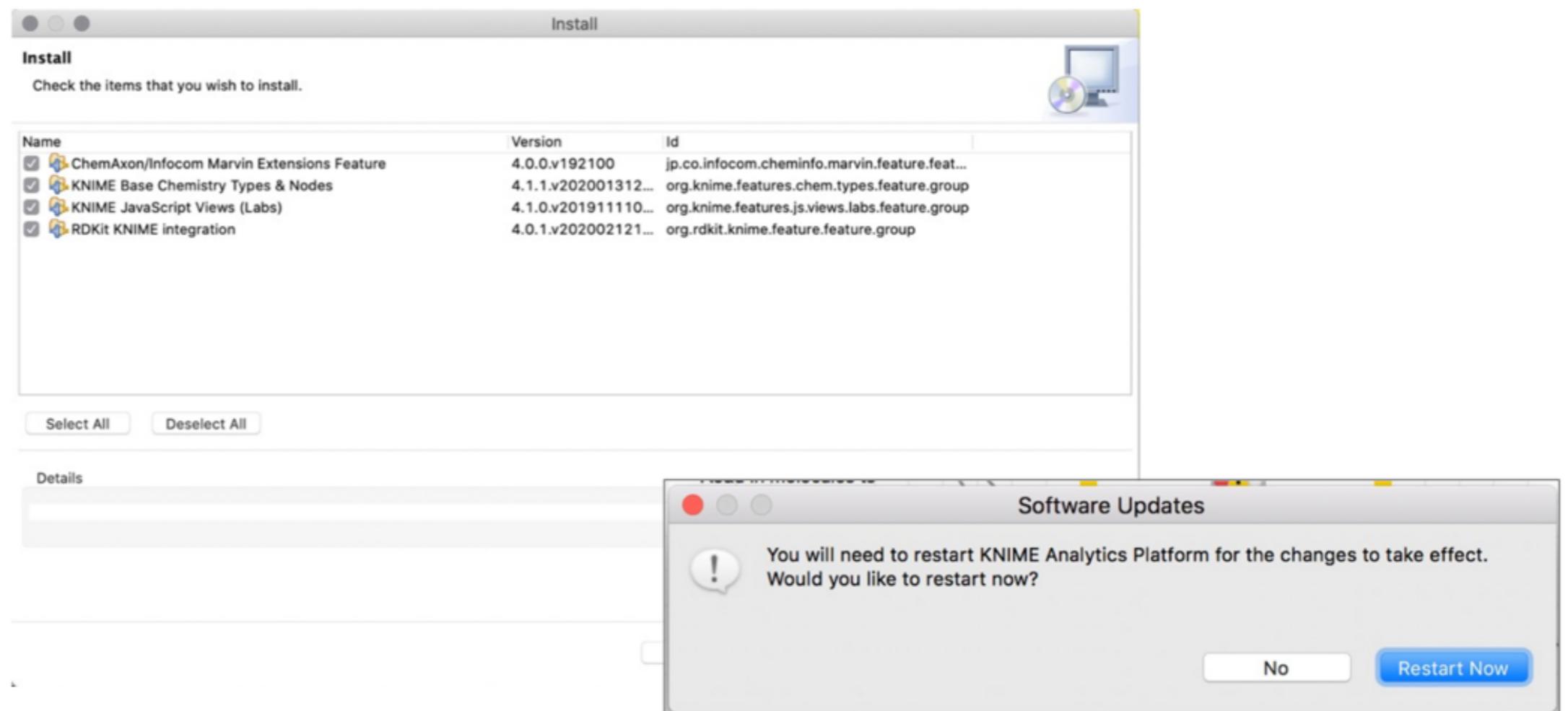
Before we start...

- Register on the website: <https://www.knime.com/>
- Download and install the latest version of the KNIME Analytics Platform (v4.5) from: <https://www.knime.com/downloads>
- Explore the KNIME Hub: <https://hub.knime.com/>
- Explore the Community (<https://www.knime.com/knime-community>) and Learning (<https://www.knime.com/learning>) sections

The image shows the official website for KNIME. At the top left is the KNIME logo with the tagline "Open for Innovation". The top navigation bar includes links for Hub, Blog, Forum, Academia, Events, Solutions, Careers, Contact, a search icon, and a "Download" button. Below the navigation, there are links for SOFTWARE, PRICING, COMMUNITY, LEARNING, PARTNERS, and ABOUT. A "Sign in" button is also present. The main content area features a large banner with the text "End to End Data Science". Below this, a paragraph explains that KNIME builds software to create and productionize data science using one easy and intuitive environment. Two buttons at the bottom are "Explore KNIME Software" and "Read our Open Source Approach". To the right of the text, there is a screenshot of the KNIME interface, which includes a circular sunburst chart, a bar chart, a flowchart, and several line graphs.

Before we start...

- Go to the RamirezLab GitHub repository and download the workflows and tutorial files we will use in this workshop:
 1. KNIME Intro: https://github.com/ramirezlab/WIKI/tree/master/KNIME/Knime_tutorial
 2. KNIME Chemoinformatics: <https://github.com/ramirezlab/WIKI/tree/master/KNIME/Chemoinformatics-Analysis-Tutorial-with-Knime-master>
- Import KNIME workflow and make sure that all extensions are working properly (accept and install the suggests extensions / complements / nodes)
 - Go to “File” and then use the “Import KNIME workflows” option



What is KNIME Analytics Platform?



- A tool for data analysis, manipulation, visualization, and reporting
- Based on the graphical programming paradigm
- Provides a diverse array of extensions:

Text Mining,

Network Mining,

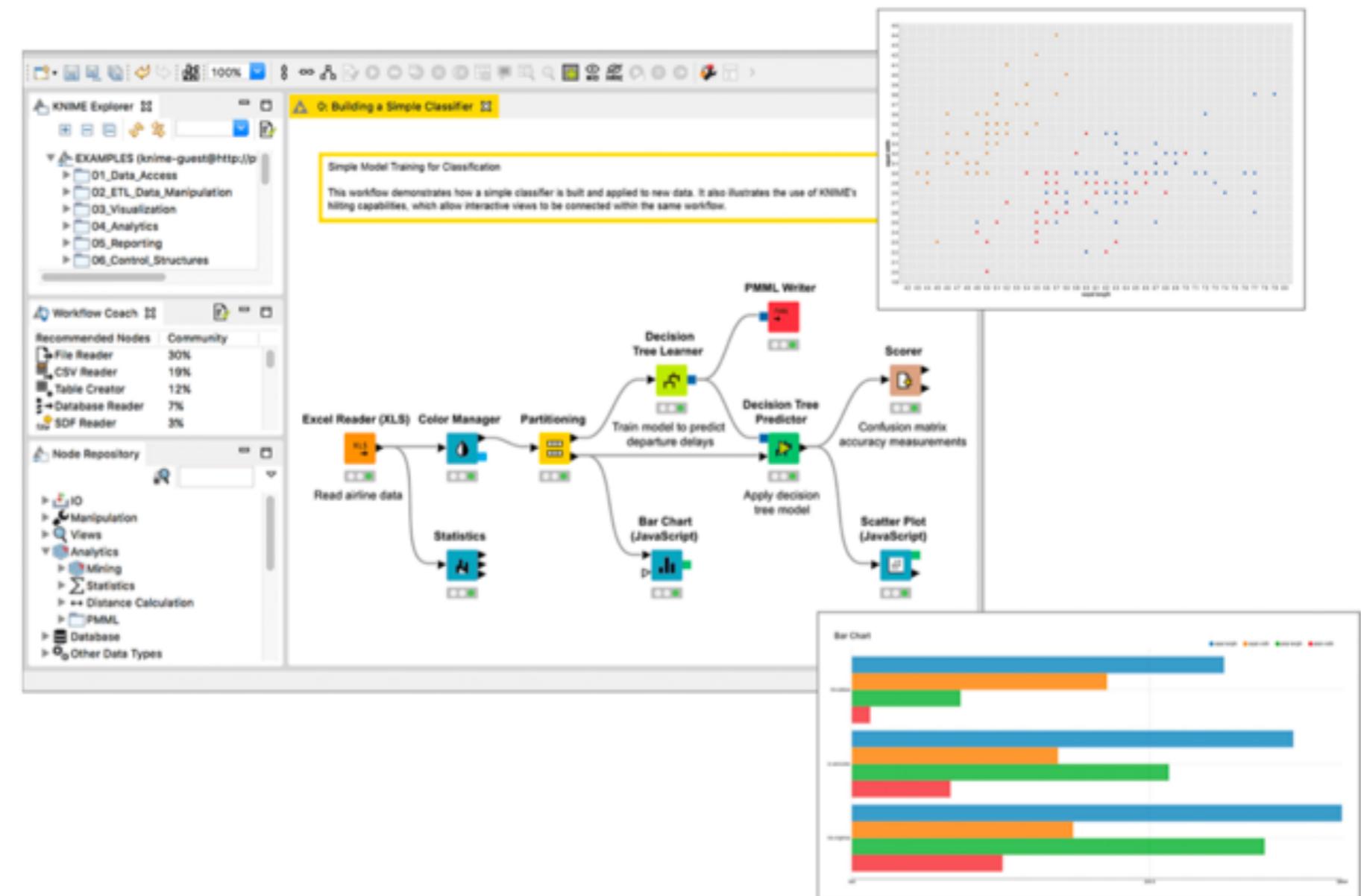
Cheminformatics,

Machine learning, etc.

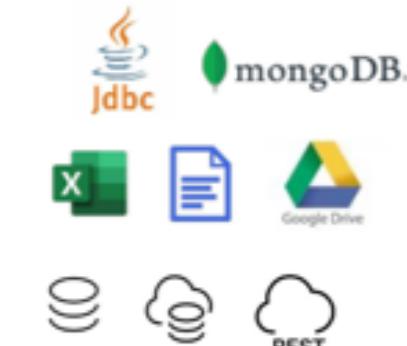
Many integrations,

such as Java, R, Python,

Weka, Keras, Plotly, etc.



Mix & Match Technologies As Needed



External Data

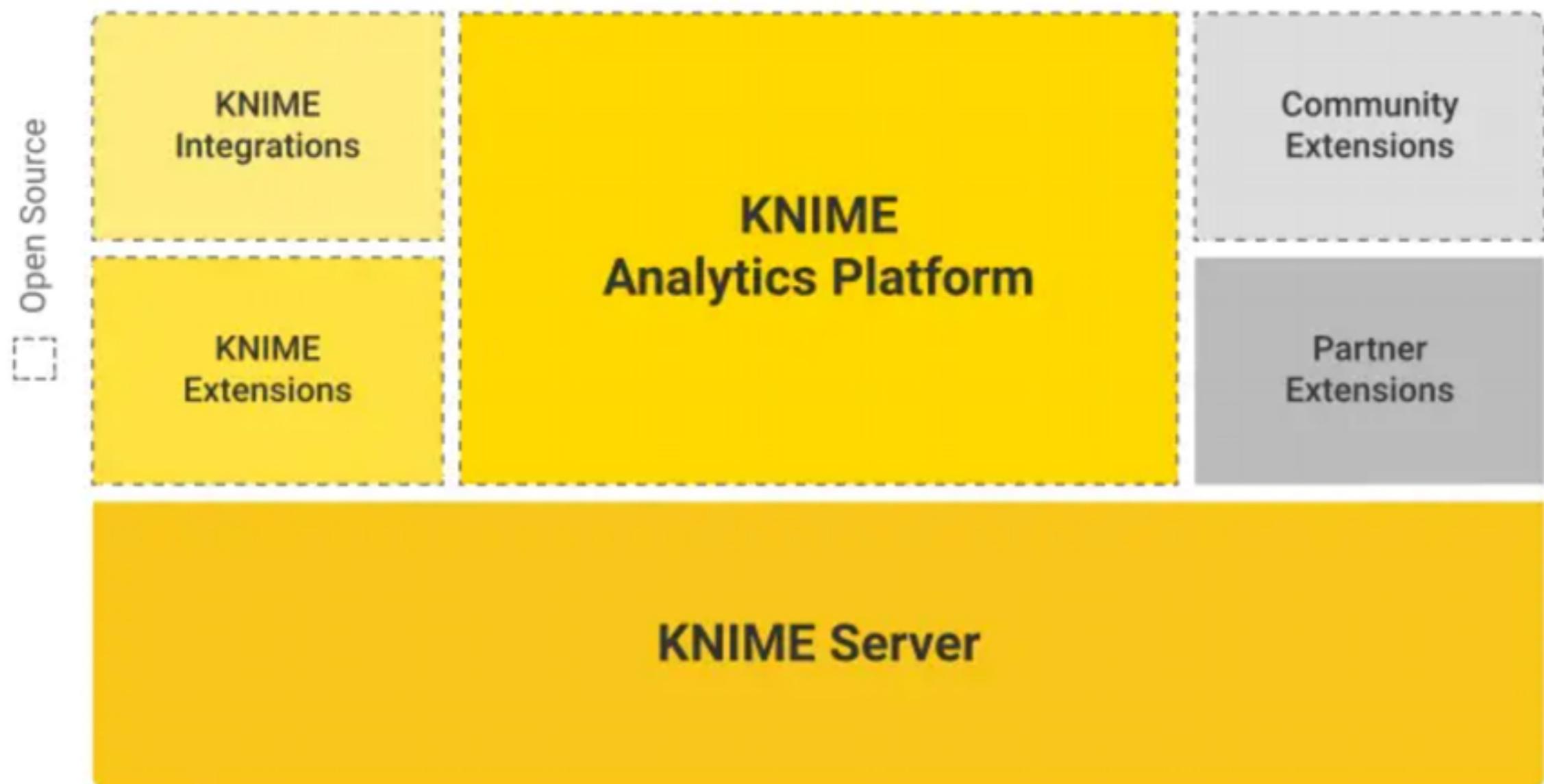


Reporting & Visualization



Cheminformatics Extensions & Molecular Sketchers

Load > Integrate > Transform > Analyze > Visualize

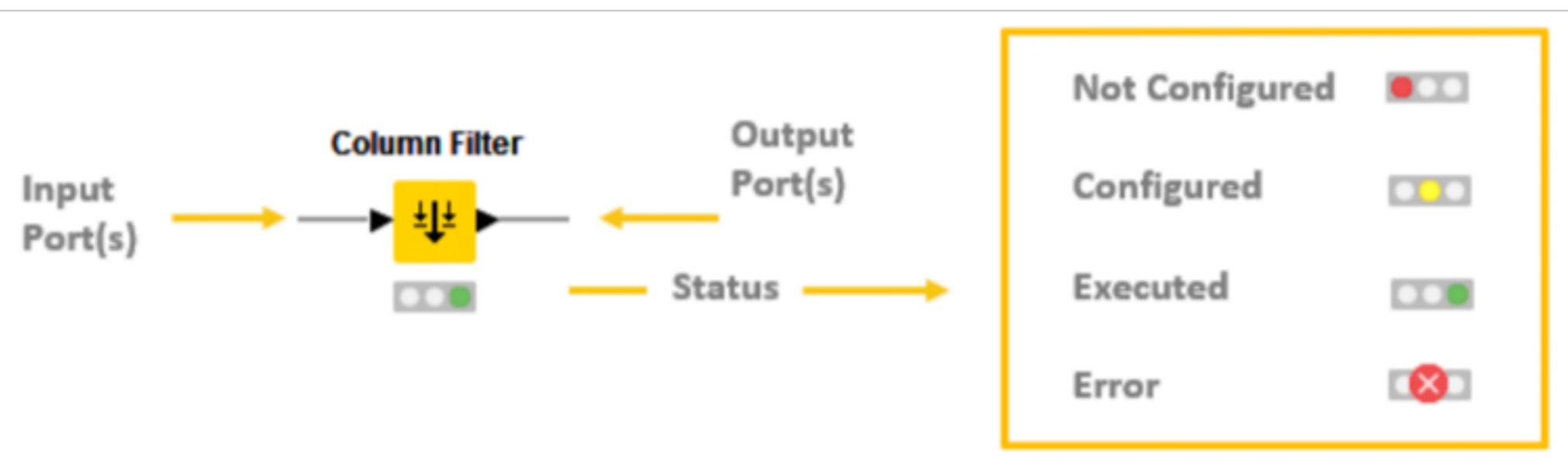


KNIME workbench

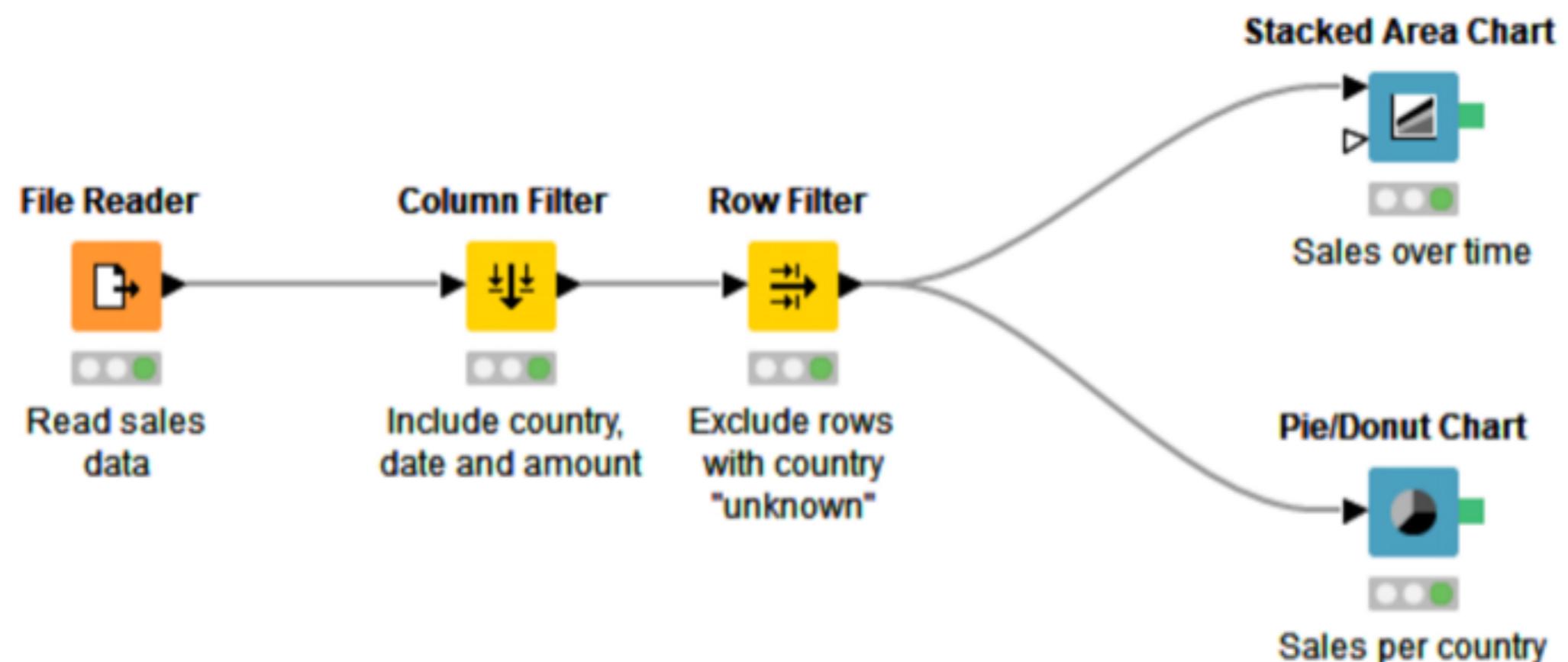
The screenshot shows the KNIME Analytics Platform interface with several annotated sections:

- Servers and Workflows**: Points to the KNIME Explorer panel on the left, which lists projects and their contents.
- Node Recommendations**: Points to the Workflow Coach panel, which displays recommended nodes from the community.
- Node Repository**: Points to the Node Repository panel on the left, listing categories like IO, Manipulation, Views, Analytics, DB, and Chemistry.
- Workflow Editor**: The central workspace where workflows are built. It shows two steps:
 - Step 1.** Read data from different sources. A "File Reader" node is connected to an "XLS Reader" node, which then connects to a "Concatenate" node. A "SDF Reader" node also connects to the "Concatenate" node.
 - Step 2.** Remove duplicates. The output of the "Concatenate" node flows through a "Molecule Type Cast" node, an "RDKit Canon SMILES" node, and finally a "Duplicate Row Filter" node.
- Node Description**: Points to the "Description" panel on the right, which provides details about the current workflow (01_Chemistry_basics).
- KNIME Hub**: Points to the KNIME Hub Search panel, which allows searching for workflows, nodes, and more.
- Console**: Points to the Console panel at the bottom, displaying log messages and system status.
- Outline**: Points to the Outline panel at the bottom, showing a hierarchical tree structure of the workflow.

Nodes and workflows



Nodes and workflows



Selected Open Source Extensions for Cheminformatics

- RDKit**
 - > Converters
 - > Modifiers
 - Calculators**
 - RDKit Descriptor Calculation
 - RDKit Calculate Charges
 - Geometry**
 - RDKit Generate Coords
 - RDKit Optimize Geometry
 - RDKit Add Conformers
 - RDKit Open 3D Alignment
 - RDKit RMSD Filter
 - Fingerprints**
 - RDKit Fingerprint
 - RDKit Count-Based Fingerprint
 - RDKit Fingerprint Reader
 - RDKit Fingerprint Writer
 - RDKit Diversity Picker
 - > Fragments
 - > Searching
 - > Reactions
 - Viewing**
 - RDKit Interactive Table
 - RDKit SMILES Headers
 - RDKit Molecule Highlighting
 - > Experimental
- CDK**
 - > 3D
 - 3D Coordinates
 - 3D D-Moments
 - 3D D-Similarity
 - 3D RMSD
 - 3D Viewer
 - S₃ 3D WHIM
 - > AMBIT
 - > I/O
 - 2D Coordinates
 - Sgn Atom Signatures
 - ChemSpider
 - Connectivity
 - Depiction
 - Element Filter
 - Fingerprint Similarity
 - Fingerprints
 - Hydrogen Manipulator
 - Lipinski's Rule-of-Five
 - Mass Calculator
 - Molecular Properties
 - OPSIN
 - SMARTS Query
 - Structure Sketcher
 - Substructure Search
 - Sugar Remover
 - Sum Formula
 - Symmetry
 - XLogP
- Erlwood Nodes**
 - > IO
 - > Structure Data Format Converters
 - Structure Similarity**
 - Fingerprint Similarity
 - Structure Properties**
 - Plane of Best Fit Calculator
 - Virtual Screening**
 - Virtual Screening Metrics
 - Evaluation and Ranking**
 - Desirability Ranking
 - Pareto Ranking
 - SAR Analysis**
 - Automated Matched Pairs
 - Free-Wilson Matched Pairs
 - Viewers**
 - 2D/3D Scatterplot
 - Activity Cliffs Viewer
 - Similarity Viewer
 - > Testing
- Vernalis**
 - > SQL Databases
 - > European PubMed Central
 - European PubMed Central Advanced Search
 - > Fingerprints
 - > Flow Control
 - > IO
 - > Matched Molecular Pairs (MMPs)
 - Filtering
 - Fragmentation
 - Pair Generation
 - Rendering
 - Transforms
 - MMP Calculate Maximum Cuts (RDKit)
 - MMP Fragmentation Type Loop Start
 - Uniquify IDs
 - > Principal Moments of Inertia (PMIs)
 - > RCSB PDB Tools
 - PDB Connector
 - PDB Connector (XML Query)
 - PDB Connector Combine XML Queries
 - PDB Connector Custom Report
 - PDB Connector Query Only
 - PDB Connector Query Only (XML Query)
 - PDB Connector XML Query Builder
 - PDB Describe Heterogens
 - PDB Downloader (Source)
 - PDB SMILES Query
 - PDB Downloader
 - > Local PDB Tools
 - > Sequence Tools
 - > Speedy SMILES
 - > Testing
 - > Miscellaneous

Commercial Extensions for Cheminformatics

- ▼  BioSolveIT Nodes
 -  CoLibri (Chemistry Spaces)
 -  IO
 -  Assess Affinity with Hyde in SeeSAR
 -  Compute FTrees Similarity
 -  Compute FlexS Alignments
 -  Compute LeadIT Docking
 -  Convert Molecules with Naomi
 -  FTrees Query Generator
 -  Filter Molecules with Naomi
 -  FlexX Docking
 -  Generate 3D Coordinates
 -  Generate Protomers / Tautomers with Naomi
 -  Interactive BioSolveIT Table
 -  Interactive SeeSAR Viewer
 -  Prepare Receptor with LeadIT
 -  Run ReCore Interactively
 -  Search FTrees Fragment Space
 -  SeeSAR Project Generator
- ▼  ChemAxon / Infocom
 - ▼  JChem
 -  IO
 -  Converter
 -  Marvin
 -  Calculator Plugins
 -  JChem Base
 -  JChem Cartridge
 -  Standardizer
 -  Structure Checker
 -  Name to Structure
 -  Screen
 -  JKlustor
 -  Reactor
 -  Markush Viewer
 -  Metabolizer
 -  Fragmenter
 -  Marvin
 - ▼  Cresset
 -  Forge
 -  Models
 -  Project
 -  Forge Align
 -  Activity Miner
 -  FieldTemplater
 -  Spark
 -  Spark Fragment Selector
 -  Generate Spark Database
 -  Spark Database Search
 -  XedTools
 -  XedMin
 -  XedeX
 -  Torch/Forge Molecule Viewer
 -  MOE
 -  Input
 -  Output
 -  Convert
 -  Transform
 -  Process
 -  Calculate
 -  QuaSAR
 -  Fingerprints
 -  Simulations
 -  Bioinformatics
 -  Fragment Based Design
 -  CombiChem
 -  Miscellaneous
 -  Pharmacophore
 -  Materials
 -  Schrödinger
 -  Readers/Writers
 -  Converters
 -  Ligand Preparation
 -  Property Generation
 -  Cheminformatics
 -  Pharmacophore Modeling
 -  Protein Structure Prediction
 -  Docking and Scoring
 -  Molecular Mechanics
 -  Molecular Dynamics
 -  Quantum Mechanics
 -  Workflows
 -  Filtering
 -  Reporting
 -  Scripting
 -  Tools

Additional Resources

KNIME pages (<https://www.knime.com>)

Life Sciences landing page <https://www.knime.com/why-knime-for-life-science>

RESOURCES **LEARNING HUB** <https://www.knime.com/learning-hub>

RESOURCES **HUB** <https://hub.knime.com/>

BOOK **WILL THEY BLEND** <https://www.knime.com/knimepress/will-they-blend>

KNIME Tech pages

FORUM for questions and answers <https://forum.knime.com>

- **DOCUMENTATION** for docs, FAQ, changelogs, ... <https://docs.knime.com/>

- **COMMUNITY CONTRIBUTIONS** for dev instructions and third party nodes <https://www.knime.com/community>

KNIME TV on **YouTube** <https://www.youtube.com/user/KNIMETV>. https://www.youtube.com/watch?v=mGv0Nle_Nr

KNIME TV on **YouTube**

https://www.youtube.com/watch?v=mGv0Nle_NrQ

More information

TeachOpenCADD-KNIME -> [Volkamer Lab](#)

JCIM JOURNAL OF CHEMICAL INFORMATION AND MODELING Application Note Cite This: *J. Chem. Inf. Model.* 2019, 59, 4083–4086 pubs.acs.org/jcim

TeachOpenCADD-KNIME: A Teaching Platform for Computer-Aided Drug Design Using KNIME Workflows

Dominique Sydow,^{†,||} Michele Wichmann,^{†,||} Jaime Rodríguez-Guerra,[†] Daria Goldmann,[‡] Gregory Landrum,[§] and Andrea Volkamer^{*,†}

[†]In Silico Toxicology, Institute of Physiology, Charité - Universitätsmedizin Berlin, Charitéplatz 1, 10117 Berlin, Germany
[‡]KNIME GmbH, Körtestr. 10, 10967 Berlin, Germany
[§]KNIME AG, Technoparkstr. 1, 8005 Zurich, Switzerland

ABSTRACT: Open-source workflows have become more and more an integral part of computer-aided drug design (CADD) projects since they allow reproducible and shareable research that can be easily transferred to other projects. Setting up, understanding, and applying such workflows involves either coding or using workflow managers that offer a graphical user interface. We previously reported the TeachOpenCADD teaching platform that provides interactive Jupyter Notebooks (talktutorials) on central CADD topics using open-source data and Python packages. Here we present the conversion of these talktutorials to KNIME workflows that allow users to explore our teaching material without any line of code. TeachOpenCADD KNIME workflows are freely available on the KNIME Hub: <https://hub.knime.com/volkamerlab/space/TeachOpenCADD>.

<https://hub.knime.com/volkamerlab/spaces/Public/latest/TeachOpenCADD/TeachOpenCADD~xYhrR1mfFcGNxz7I>

More information

TeachOpenCADD-Jupyter Notebook -> [Volkamer Lab](#)

Sydow et al. *J Cheminform* (2019) 11:29
<https://doi.org/10.1186/s13321-019-0351-x>

Journal of Cheminformatics

SOFTWARE

Open Access



TeachOpenCADD: a teaching platform for computer-aided drug design using open source packages and data

Dominique Sydow , Andrea Morger , Maximilian Driller and Andrea Volkamer ^{*}

<https://github.com/dominiquesydow/TeachOpenCADD>