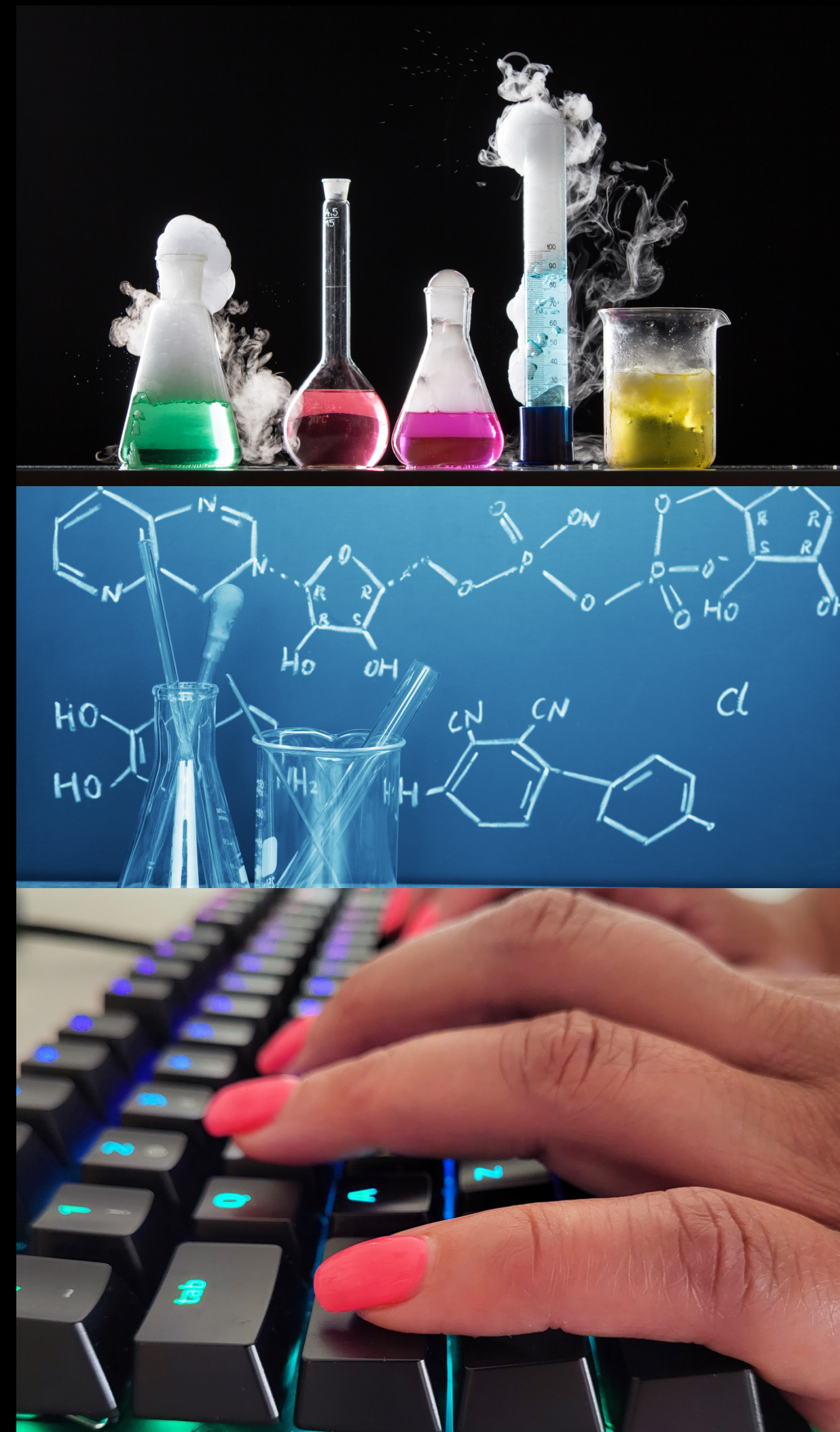


Cheminformatics

with Kimberly Deas, MS, PhD Candidate



Converting SMILES to SDF Using KNIME

Prerequisites

- 1) Basic Chemistry and data knowledge.
- 2) Familiarity with KNIME is helpful.
- 3) Willingness to learn!

Resources

- 1) KNIME Website, [knime.com](https://www.knime.com)
- 2) Introduction to Cheminformatics, David Wild, PhD,
<http://david-wild-knzi.squarespace.com/david-wilds-blog/2018/5/1/introducing-cheminformatics>
- 3) Chemistry Library / LibreTexts Project website,
chem.libretextx.org

Important Terms

- **CDK:** Chemical Development Kit (<https://cdk.github.io/>) - allows 2D rendering of structures, also allows SMILES generation and parsing.
- **RDKit:** Open source cheminformatics and machine learning tools (<https://www.rdkit.org/>) - contains python functions for modifying molecules.
- **SDF:** Structured Data Files consist of atom and bond tables that are used to convey chemical structural information. Molfiles are examples.
- **SMILES:** Defined as The Simplified Molecular-Input Line-Entry System (SMILES), are short ASCII strings for describing chemical structures.

SMILES to SDF Workflow in KNIME

The screenshot displays the KNIME Analytics Platform interface with a workflow titled "SMILEStoSDF" (tab 5). The workflow consists of the following nodes:

- Node 1: CSV Reader** (Orange icon)
- Node 2: Column Rename** (Yellow icon with "A7 LB")
- Node 3: Missing Value** (Yellow icon with "?")
- Node 4: Molecule to CDK** (Yellow icon with "Mol CDK")
- Node 5: RDKit To Molecule** (Yellow icon with a circular arrow)
- Node 6: SDF Writer** (Red icon with "Variable Input")

The workflow is connected as follows: CSV Reader (Node 1) → Column Rename (Node 2) → Missing Value (Node 3) → Molecule to CDK (Node 4) → RDKit To Molecule (Node 5) → SDF Writer (Node 6).

A blue box overlay on the workflow diagram contains the text: **Convert SMILES in a CSV file to an SDF mol file.**

The left sidebar shows the **KNIME Explorer** (Local Workspace) and **Workflow Coach** (Recommended Nodes). The bottom panel shows the **Node Repository** (IO > Read > CSV Reader) and the **Console** output:

```
KNIME Console
*****
***      Welcome to KNIME Analytics Platform v4.6.1.v202207182243      ***
***      Copyright by KNIME AG, Zurich, Switzerland                    ***
*****
Log file is located at: D:\knime-workspace1\.metadata\knime\knime.log
```

Transition to KNIME