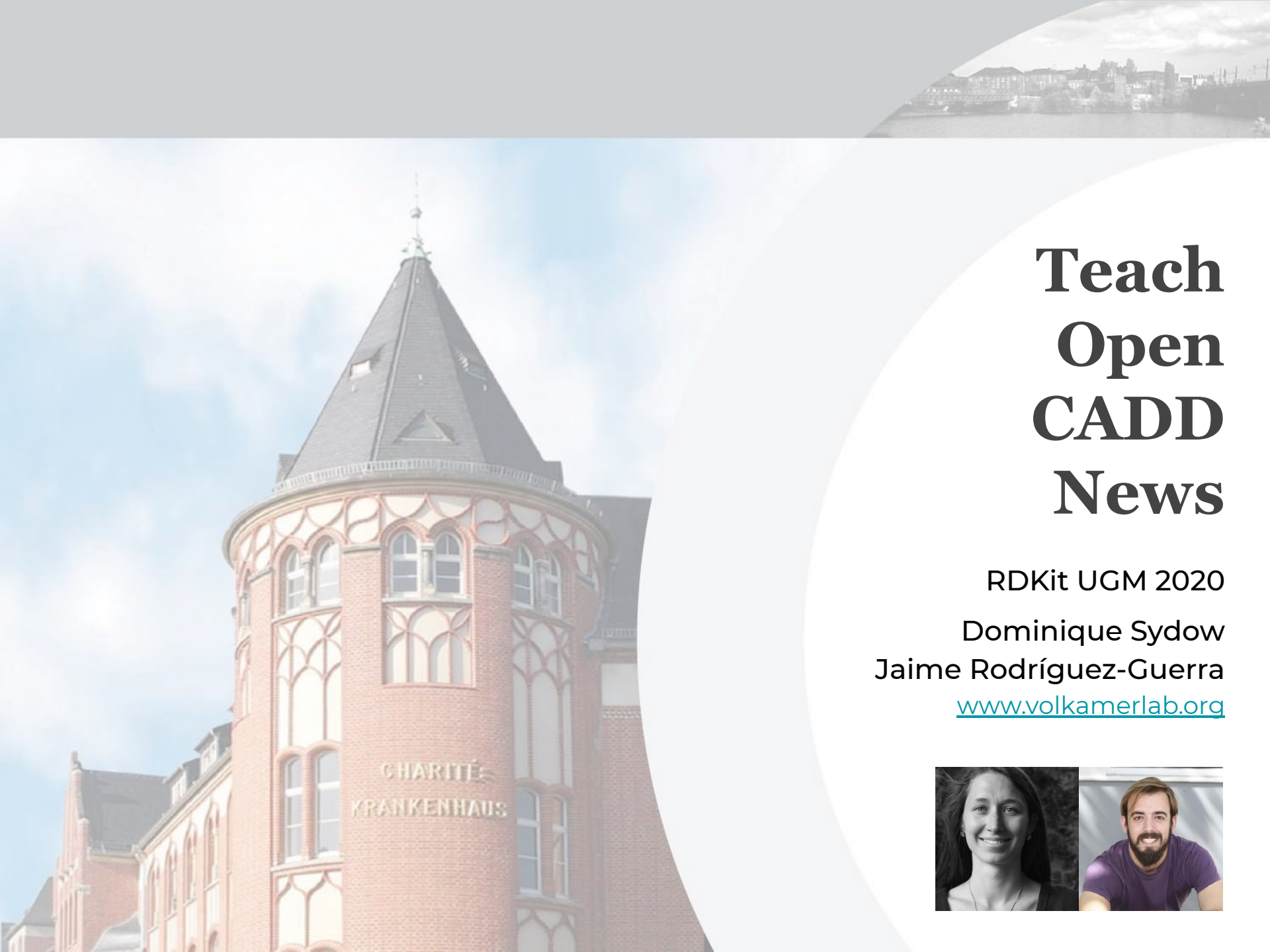
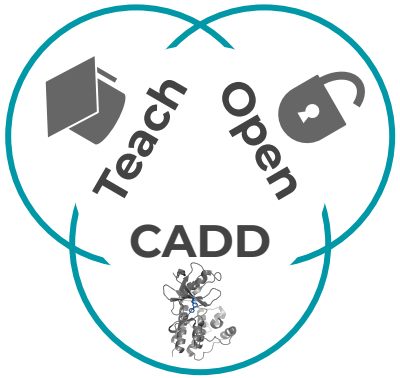


Teach Open CADD News

RDKit UGM 2020

Dominique Sydow
Jaime Rodríguez-Guerra
www.volkamerlab.org

TeachOpenCADD - what is this?



Teach common CADD tasks using open resources (packages, databases) for beginners and advanced users!

www.github.com/volkamerlab/TeachOpenCADD

jupyter
python

Open for Innovation
KNIME

How?

- Coding-based Jupyter notebooks[1]
- GUI-based KNIME workflows[2]
- Highly interactive
- All in one place: Theory + code/nodes + output

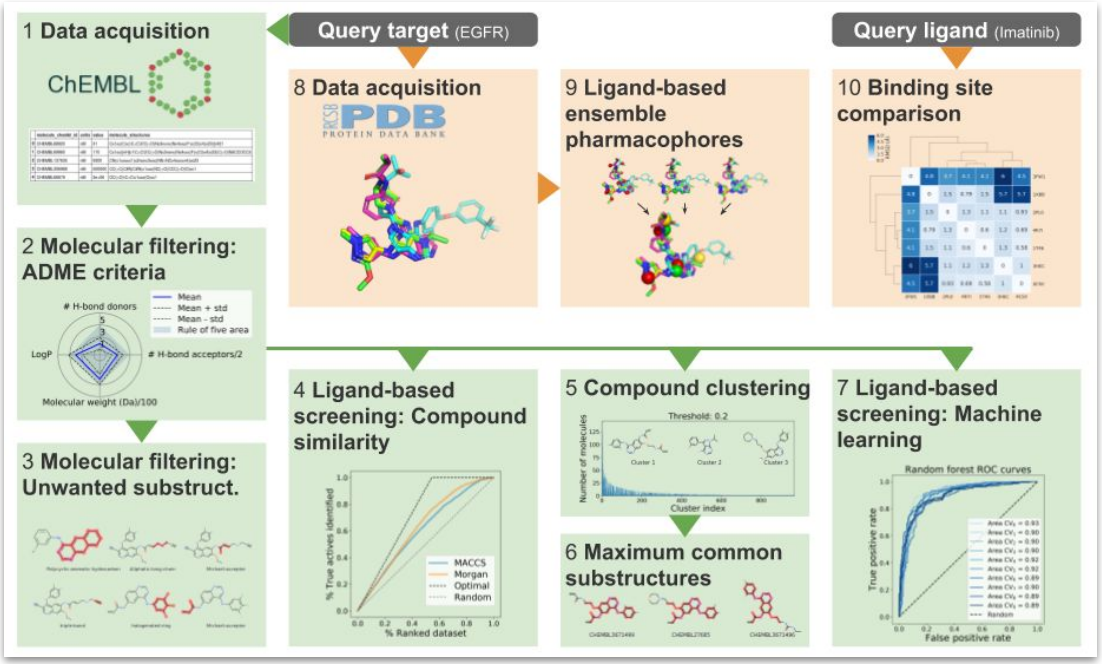
[1]Sydow et al. *J. Chem.* (2019), [2]Sydow and Wichmann et al. *JCIM* (2019)

TeachOpenCADD - what is offered so far?

Open-Source Cheminformatics and Machine Learning

Cheminformatics and structural bioinformatics

Published content



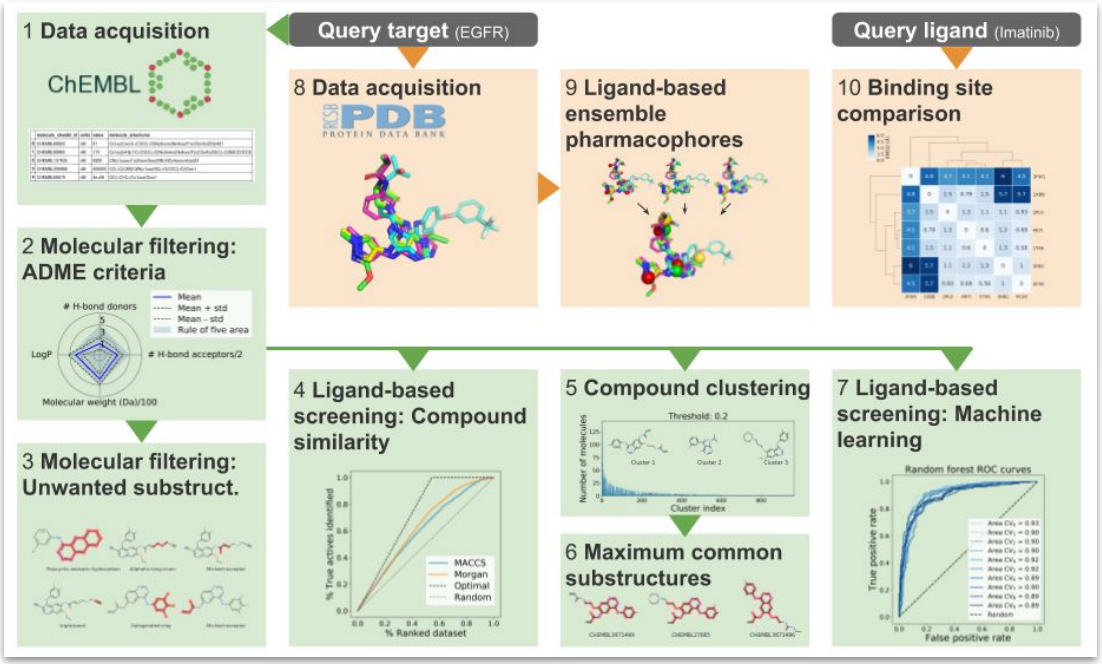
WIP: Towards more pythonic code - refactoring 2020
https://github.com/volkamerlab/TeachOpenCADD/pull/42

TeachOpenCADD - what is offered so far?

Open-Source Cheminformatics and Machine Learning

Cheminformatics and structural bioinformatics

Published content



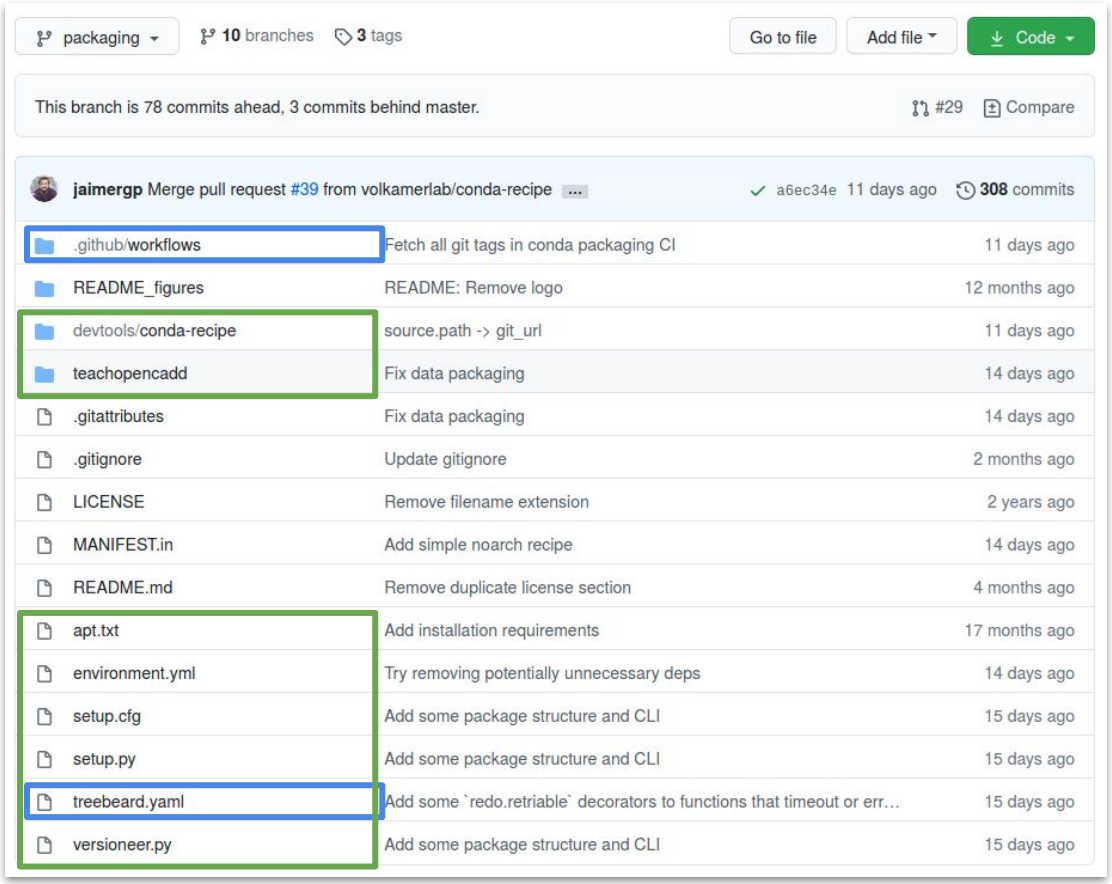
WIP: Towards more pythonic code - refactoring 2020
https://github.com/volkamerlab/TeachOpenCADD/pull/42

New content / WIP

- Programmatic access to online databases
  - PubChem
  - KLIFS
- Protein-ligand docking
- Protein-ligand interactions
- Binding site detection
- Molecular dynamics
- Deep learning and one-hot encoding

TeachOpenCADD - coming up...

On our way to a `teachopencadd` conda package!



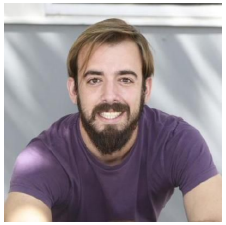
https://github.com/volkamerlab/TeachOpenCADD/pull/29

Packaging!

- ✓ Apply package structure to repo
- ✓ Add CLI
- ▢ Make conda package

Continuous integration!

- ✓ GH Action & Treebeard
- ✓ Automatically spot raised errors in notebooks
- ▢ Test notebook output (selected cells)



Driven by Jaime

TeachOpenCADD - coming up...

On our way to a `sphinx`-based TeachOpenCADD website!

Molecular filtering: ADME and lead-likeness criteria

Search

TeachOpenCADD
145 Stars · 52 Forks

TeachOpenCADD

Getting started

Installing

Talktorials

All talktorials

- Compound data acquisition (ChEMBL)
- Molecular filtering: ADME and lead-likeness criteria
- Molecular filtering: unwanted substructures
- Ligand-based screening: compound similarity
- Ligand-based screening: compound similarity
- Compound clustering
- Maximum common substructure
- Ligand-based screening: machine learning
- Protein data acquisition: Protein Data Bank (PDB)
- Ligand-based pharmacophores
- Binding site similarity and off-target prediction
- Structure-based CADD using online APIs/servers
- Structure-based CADD using online APIs/servers
- Structure-based CADD using online APIs/servers
- Structure-based CADD using online APIs/servers

Developers

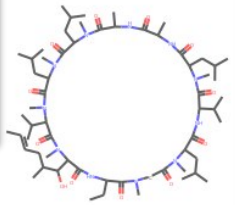
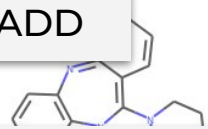
Define and visualize example molecules

Before working with the whole dataset retrieved from ChEMBL, we pick four example compounds to investigate their chemical properties. We import the nece… …aw them.

```
[3]: smiles = [
         "CCC1C(=O)N(CC(=O…                    …(C(=O)N(C(C(=O)N1)C(C(C
         "CN1CCN(CC1)C2=C3…
         "CC1=C(C(CCC1)(C)C)C=CC(=CC=CC(=CC=CC=C(C)C=CC=C(C)C=CC2=C(CCCC2(C)C)C)C)C",
         "CC…CCC1=CC(=C(C(=C1)O)C2C=C(CCC2C(=C)C)C)O"
     ]
     names = ["cyclosporin…
```

```
[4]: molecules = pd.DataFr…({"name" : names, "smiles" : smiles})
     PandasTools.AddMoleculeColumnToFrame(molecules, "smiles")
     molecules
```

[4]:

	name		ROMol
0	cyclosporine	CCC1C(=O)N(…	
1	clozapine	CN1CCN(CC1)C2=C3C=CC=CC3=NC4=C(N2)C=C(C=C4)C	

Find easy installation instructions (conda installation)

Explore notebooks online

Driven by Jaime

Find detailed information on how to contribute to TeachOpenCADD

https://github.com/volkamerlab/TeachOpenCADD/pull/40

Contents

Acknowledgements

Talktorials T1-10

Students from CADD course 2017 & 2018!!
Dominique Sydow
Andrea Morger
Maximilian Driller
Andrea Volkamer

KNIME workflows W1-8

Michele Wichmann
Dominique Sydow
Jaime Rodríguez-Guerra
Greg Landrum
Daria Goldmann
Andrea Volkamer

RDKit UGM 2019 Hackathon

Richard Gowers
Floriane Montanari
Jaime Rodríguez-Guerra
Dominique Sydow
...

Refactoring 2020

Andrea Volkamer
Jaime Rodríguez-Guerra
Dominique Sydow
Talia Kimber
Mareike Leja
Sakshi Misra
Yonghui Chen
David Schaller

Additional contributors

Greg Landrum
Jeff Wagner
Ross Johnstone
Thomas Holder
Andrew Wilkinson

Freie Universität Berlin

EINSTEIN Foundation.de

Bundesministerium für Bildung und Forschung

DFG Deutsche Forschungsgemeinschaft

CHARITÉ UNIVERSITÄTSMEDIZIN BERLIN