A free toolchain from molecular vibrations to detailed combustion

how (some) physical chemists and chemical engineers have escaped proprietary software

Mark E. Fuller, Ph.D. and Kfir Kaplan

LibrePlanet 2022

March 9, 2022

About Us

Mark and Kfir are both active in developing and contributing to FOSS projects within the scientific community including CANTERA, the REACTION MECHANISM GENERATOR, and the AUTOMATIC RATE CALCULATOR.

Mark:

- **Ph.D. in chemical engineering**
- **Linux/FOSS user and contributor over 15+ years**
- **Research: combustion, batteries, fuel cells, and heat transfer**
- **Currently employed in DevOps**
- **Personal page at `stossrohr.net`**

Kfir:

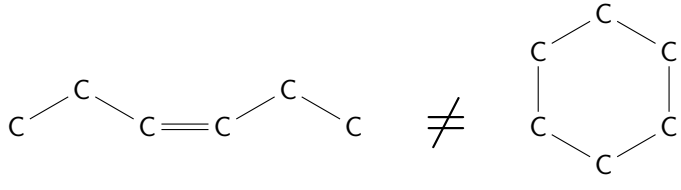
- **Fourth-year undergraduate student at the Technion in Biochemical Engineering**
- **Contributes to FOSS for over a year**
- **Conducts research at `dana.net.technion.ac.il`**

Introduction

Chemistry 101

A few points require some chemistry terminology:

- **Atoms:**
  - ➔ Basic units of matter, composed of protons, neutrons, and electrons
  - ➔ Categorized as elements as a function of number of protons
  - ➔ Protons and neutrons form a nucleus about which electrons orbit
  - ➔ Electrically charged and form bonds through arrangements of electrons
- **Molecules:**
  - ➔ Also called “(chemical) species”
  - ➔ Comprised of connected atoms
  - ➔ Has internal movement (rotation, vibration)
  - ➔ Structure matters: hexene (C_6H_{12}) $\neq$ cyclo-hexene (C_6H_{12})

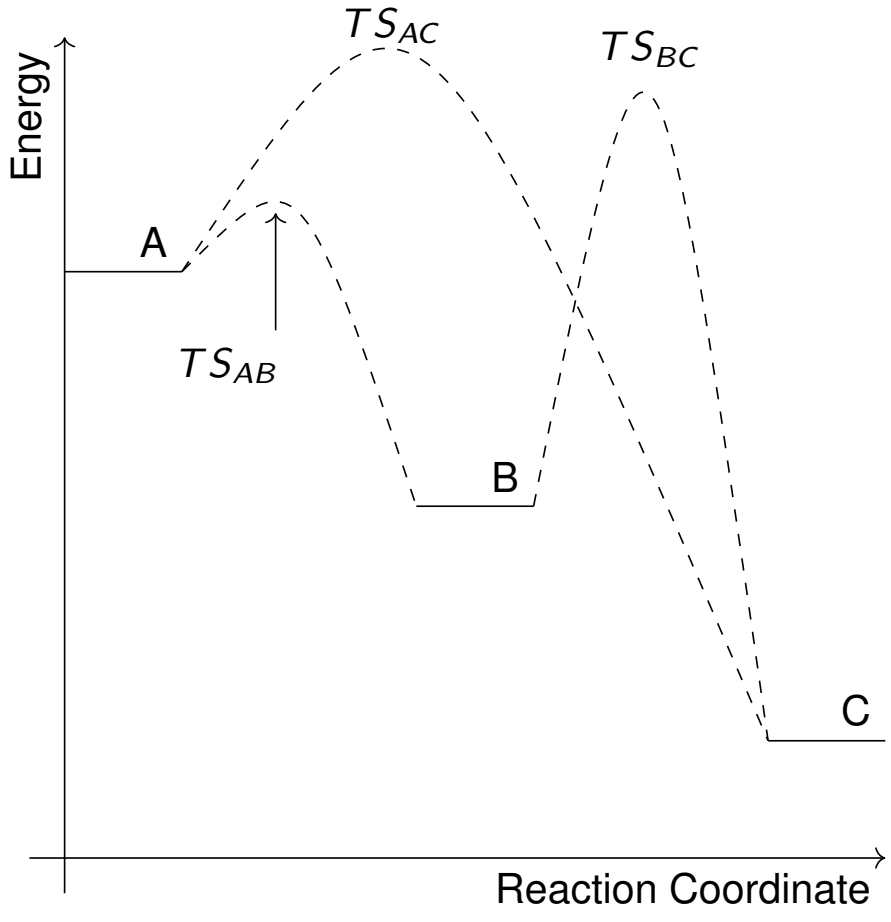


Chemistry 101

Only a few more terms!

- **Reaction:**
  - ➔ A process that changes one or more molecules into others
  - ➔ Has reactants (inputs) and products (outputs), but processes are reversible
  - ➔ Has rate, which determines how fast the process occurs
- **Transition state:**
  - ➔ The highest-energy species that is generated during the reaction
  - ➔ Is a "bottleneck" of the reaction - it controls the rate
  - ➔ Allows for calculations of the reaction rates
- **Mechanism:**
  - ➔ Combinations (sets) of reactions and species
  - ➔ Allow for simulation the time evolution of a chemical system
  - ➔ Depends on external conditions (temperature, pressure, constraints)

Reaction Kinetics and Thermodynamics



- **Thermodynamics favor lowest energy: product C dominates**
- **Kinetics (reaction energy and transition states) show instead that B forms predominantly**

Computational chemistry: an essential science

Just a few examples:

- **Alternative bio and manufactured fuels**
- **Batteries and fuel cells**
- **Drug and pharmaceutical manufacturing and degradation**
- **Proteins and organic molecules**
- **Refining and materials synthesis**

We are moving from postdictive to predictive computational capabilities[1]

[1] Green, W. H. AIChE Journal **2020**, *66*, 1–16.

Toolchain overview: macro to micro-scale

"Complete" simulation of a chemical reactor:

- **Global/macroscopic: temperature, pressure, chemical composition (TPX)**
- **Reactor/condition-specific implementation of laws of thermodynamics, conservation equations**
- **Chemical reactions for all chemical species**
- **Thermodynamic properties for all chemical species**

Toolchain overview: computing and sourcing properties

Various computational approaches:

- **“Guess” new species and thermodynamic properties from tabulated rules**
- **Estimate reaction rates by analogy to other reactions**

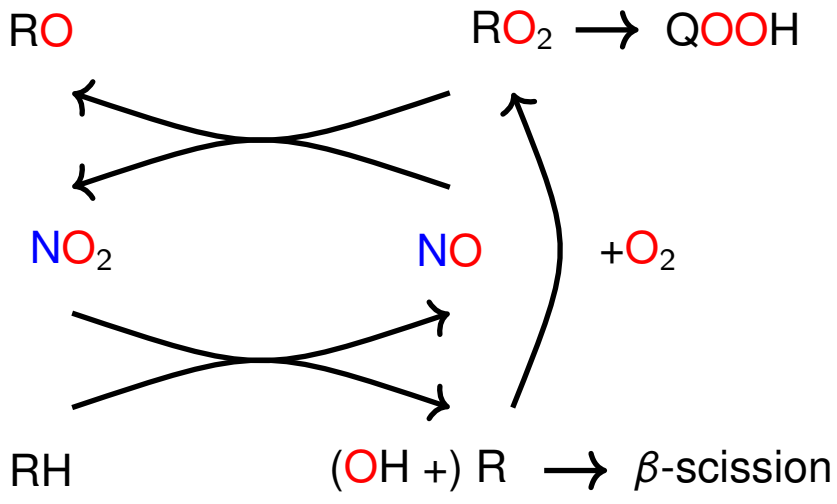
or

- **Compute molecular structures (many different methods)**
- **Calculate thermodynamic properties and reaction rates**

{cheap/fast, low-accuracy} versus {expensive/slow, high-accuracy}
$\rightarrow$we need both approaches to solve real problems

Guessing Species and Reactions

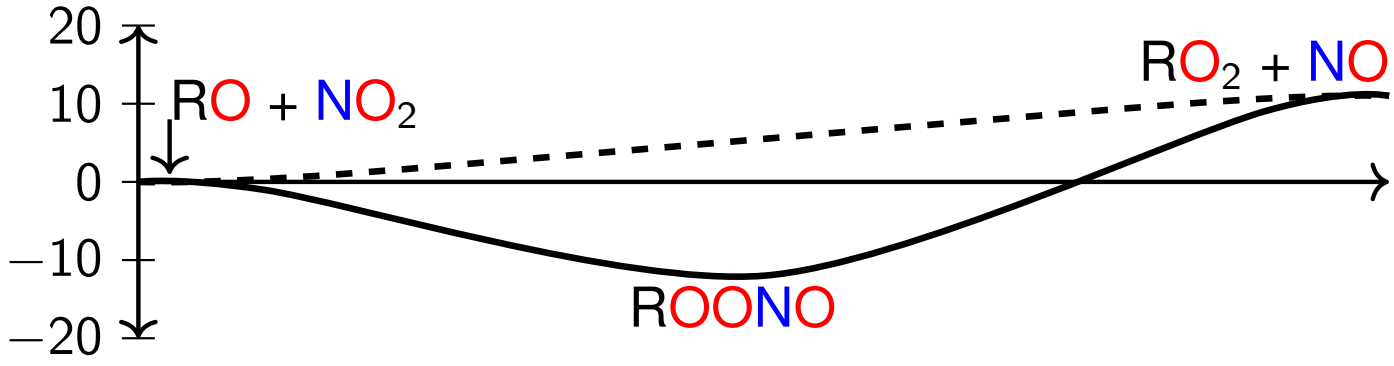
- **Hydrogen abstractions**
  - → $RH + NO_2 \rightleftharpoons R + HONO$
  - → $RH + NO_2 \rightleftharpoons R + HNO_2$
  - → $RH + NO \rightleftharpoons R + HNO$
- **Nitrite/Nitrate/Nitro-/Nitroso-Compounds**
  - → $RONO \rightleftharpoons RO + NO$
  - → $RONO_2 \rightleftharpoons RO + NO_2$
  - → $RNO_2 \rightleftharpoons R + NO_2$
  - → $RNO \rightleftharpoons R + NO$
- **Isomerizations**
  - → $RONO \rightleftharpoons RNO_2$
- **HONO elimination**
  - → $RONO \rightleftharpoons$ alkene + HONO
- **NO_x cycling**
  - → $RO_2 + NO \rightleftharpoons RO + NO_2$
  - → $R + NO_2 \rightleftharpoons RO + NO$



Adding NO_x to a combustion process[2]

[2] Fuller, M. E. et al. Reaction Chemistry & Engineering **2021**, *6*, 2191–2203.

Calculating Species and Reactions



Generalized potential energy surface for alkoxy radical (RO) + NO_2 system. Energies in kcal/mol. Well-skipping occurs at virtually all combustion-relevant temperatures and pressures.

Reaction	A	n	E_a
$CH_3O_2 + NO \rightleftharpoons CH_3O + NO_2$	4.62E+15	-0.38	97.8
$C_2H_5O_2 + NO \rightleftharpoons C_2H_5O + NO_2$	2.11E+14	-0.12	-470.6
n-$C_3H_7O_2 + NO \rightleftharpoons n$-$C_3H_7O + NO_2$	1.07E+14	-0.25	-1302.0

Units: centimeters, kelvin, calories, moles

Software

Our toolchain

- Graphical drawing of structures, basic geometry, and input file generation with AVOGADRO2
- Electronic structure calculations of individual molecules with PSI4
- Conversion of individual molecule results to thermodynamic properties and reaction rates with ARC (also other options)
- Automated model construction including estimating properties with RMG
- Automating decisions to refine estimates with computations using T3
- Reactor simulations with CANTERA
- Comparing to experimental data with standardized formatting (CHEMKED) and tools for validation and manipulation (PYKED)

Reference toolchain

Without the preceding free software projects, our life would be very different!

- **Buy an expensive license for a single package to draw structures and then perform calculations**
- **Possibly buy another expensive license to perform finer calculations**
- **Manually copy-paste output values into an in-house code to calculate thermodynamics and reaction rates (?)**
- **Manually assemble a mechanism file from literature and our new results**
- **Purchase yet another license to perform the reactor simulations**
- **Publish the data in a non-machine-readable format (or maybe only as points on a figure)**

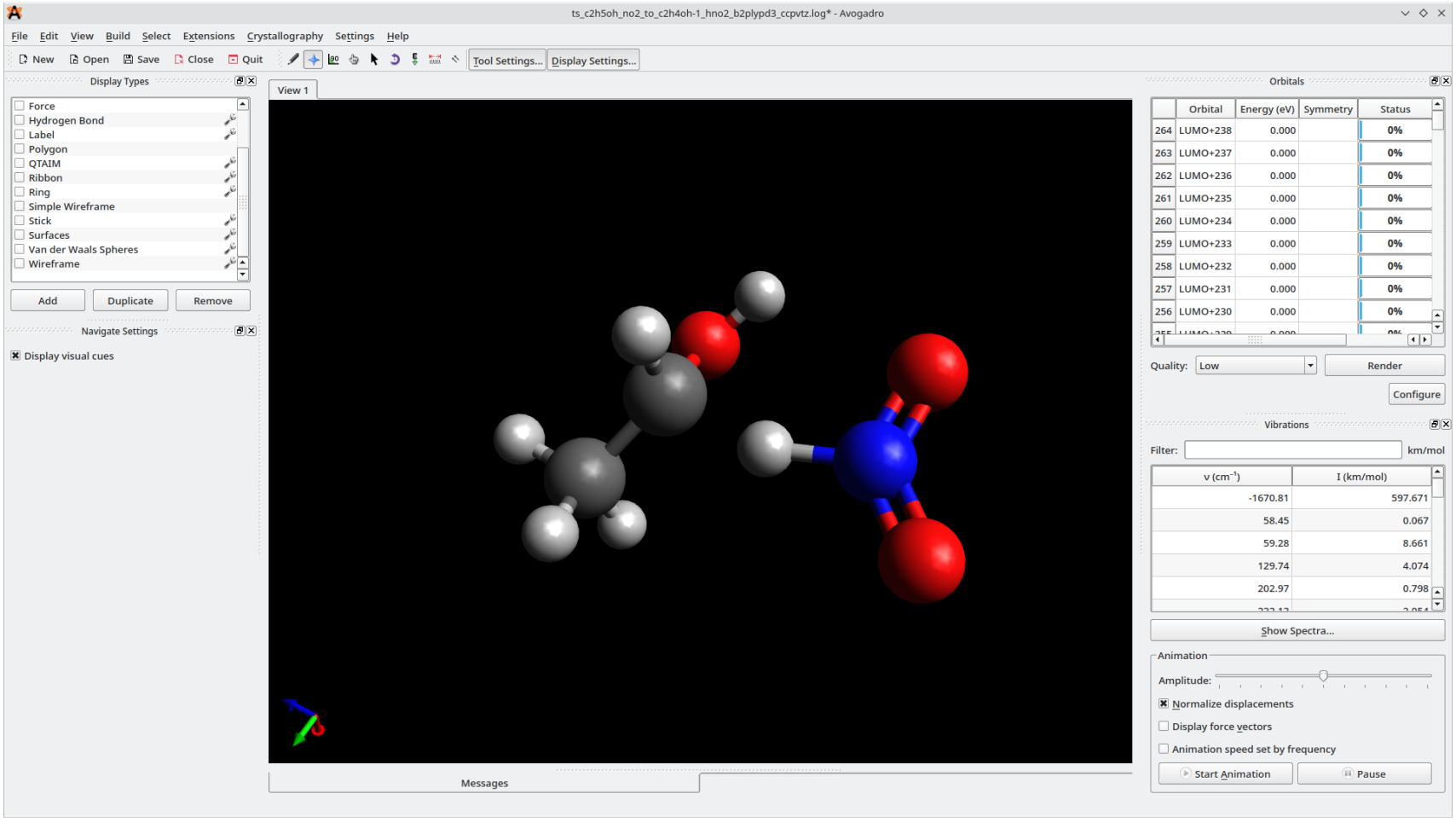
Avogadro2

`two.avogadro.cc`

OpenChemistry/avogadro[app,libs]

- **Written in C++, released under the BSD 3 Clause License**
- **Rewrite of the original Avogadro under the Open Chemistry projects (not all features ported yet, both version in use)**
- **Cross-platform (Linux/Mac/Windows)**
- **Molecular editor for drawing molecules, creating input files for other programs, viewing output files**
- **Able to provide rough rules-based guess of geometry**
- **Plugins to many calculators including Psi4**

Viewing results with Avogadro



`psicode.org` psi4/psi4

- **Written primarily in C++ with Python interfaces, released under the LGPL-3.0 License**
- **Inputs: molecule(s), unrefined geometry, and other optional parameters**
- **Calculates:**
  - ➔ Optimized geometry
  - ➔ Energy
  - ➔ Harmonic vibrations

The Automatic Rate Calculator

ReactionMechanismGenerator/ARC

- **Written in Python 3, released under the MIT License**
- **The goal is to automatically calculate chemical species thermochemistry and reaction rate coefficients**
- **Uses output files from quantum chemistry solvers (like Psi4)**
- **Calculates reaction coordinates and themochemical properties**

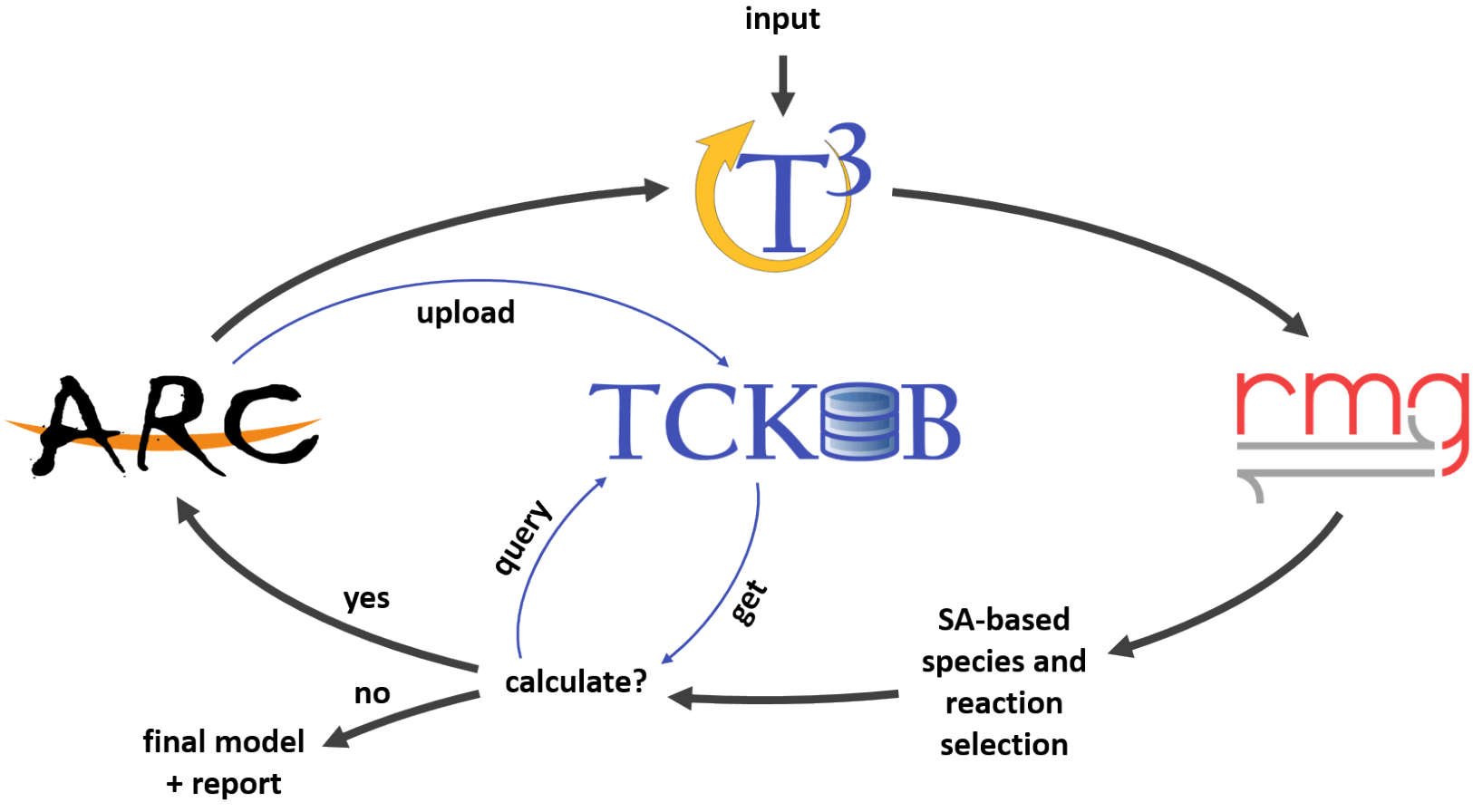
The Reaction Mechanism Generator

`rmg.mit.edu` ReactionMechanismGenerator/RMG-Py

- **Written in Python 3, released under the MIT License**
- **The goal is to automatically build kinetic models from elementary reaction sets**
- **Defines a "core" and an "edge", where the core contains the most important reactions and species**
- **Iterative calls to ARC, adding species and reactions to the core and edge.**

The Tandem Tool

ReactionMechanismGenerator/T3

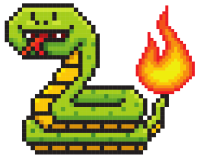


- **Written in Python 3, released under the MIT License**

Cantera

cantera.org Cantera/cantera

- **“Cantera is an open-source suite of tools for problems involving chemical kinetics, thermodynamics, and transport processes.”**
- **BSD 3-Clause license**
- **Written in C++; interfaces for programming with Python, C++, Fortran, and Matlab**
- **Built-in classes to represent wide range of gas-phase and surface chemical kinetics, multiple transport models, and reactor classes to consolidate determination of governing equations**
- **Implements Eigen and SUNDIALS libraries for solving equations**
- **Binary distribution on Fedora, RHEL, Ubuntu, Gentoo, FreeBSD, Mac and Windows plus Conda installation**



PyKED and ChemKED

pr-omethe-us/PyKED

- **ChemKED is a standard human and machine-readable file format for experimental data typical in combustion (`github.com/pr-omethe-us/ChemKED-database`)**
- **PyKED is a Python interface for validating ChemKED files and implements standard interactions and routines for use with the data (`github.com/pr-omethe-us/PyKED`)**
- **Written in Python, released under BSD 3-Clause license**

Help wanted

There is a lot that can be contributed by non-experts in chemistry (actually our biggest deficit):

- **Cleanup of Conda environments and updating versions of dependencies (e.g. migrating away from NOSETESTS) in RMG and ARC**
- **Developing database for TCKDB with reactions and interfacing to T3**
- **Binary packages and distribution in mainstream repositories on Linux distributions**
- **Extending the ChemKED standard to include additional types of experiment**
- **Overhauling data validating and type-checking in PyKED (old version of Cerberus currently)**


Dana Research Group
Fundamental and Applied Chemical Kinetics
is looking for
MSc / PhD
candidates
TECHNION
Israel Institute
of Technology
alon@technion.ac.il
dana.net.technion.ac.il

Q&A

References

(1) Green, W. H. AIChE Journal **2020**, *66*, 1–16.

(2) Fuller, M. E.; Morsch, P.; Goldsmith, M. P. C. F.; Heufer, K. A. Reaction Chemistry & Engineering **2021**, *6*, 2191–2203.

Mark E. Fuller, Ph.D. and Kfir Kaplan – fuller@fedoraproject.org

mefuller & kfir4444

fuller