

Software library to serve for faster chemical reaction processing

May 14 2019



Credit: CC0 Public Domain

Big Data has become ubiquitous in recent years, and especially so in disciplines with heterogeneous and complex data patterns. This is particularly true for chemistry. In some ways, chemical compounds may

be compared with synonyms in linguistics because one particular compound can be represented in various ways. To further complicate things, some of them don't even have a specific structure and only exist as an amalgamation of forms turning into each other. That's why it's important for researchers to know whether they are dealing with different compounds or with different representations of the same one.

Sometimes, databases also have errors arising from user inattentiveness. Special software is needed to detect and correct such errors. In the case of organic chemistry, reactions are notoriously difficult to analyze. That's why [reaction](#) data in chemoinformatics is much less developed than information about single molecules.

Laboratory of Chemoinformatics and Molecular Modeling (Kazan Federal University) has been working on this problem since 2013. Researchers there have learned to predict reaction characteristics, find optimal reaction conditions, and detect and correct data errors. As a result, a unique database of reaction characteristics has arisen. Currently, it includes 3.5 million entries. KFU is the only Russian member of Reaxys R&D Collaboration, a collective working on chemical databases.

In this new project, titled CGRtools, KFU researchers solved a number of problems to better handle reaction information. The software library is significantly richer in functionality than all the existing tools. CGRtools supports molecules and reaction as objects. CGRtools treats chemical objects similarly to standard Python data types like integers, strings, etc. Every chemical [object](#) is hashable due to atom numbering canonicalization. The objects support transparent class inheritance, which augments existing methods and attributes without breaking up existing ones.

Importantly, the library is freely available at <https://github.com/cimmkzn/CGRtools>.

More information: Ramil I. Nugmanov et al, CGRtools: Python Library for Molecule, Reaction and Condensed Graph of Reaction Processing, *Journal of Chemical Information and Modeling* (2019). [DOI: 10.1021/acs.jcim.9b00102](https://doi.org/10.1021/acs.jcim.9b00102)

Provided by Kazan Federal University

Citation: Software library to serve for faster chemical reaction processing (2019, May 14) retrieved 24 June 2024 from <https://phys.org/news/2019-05-software-library-faster-chemical-reaction.html>

This document is subject to copyright. Apart from any fair dealing for the purpose of private study or research, no part may be reproduced without the written permission. The content is provided for information purposes only.