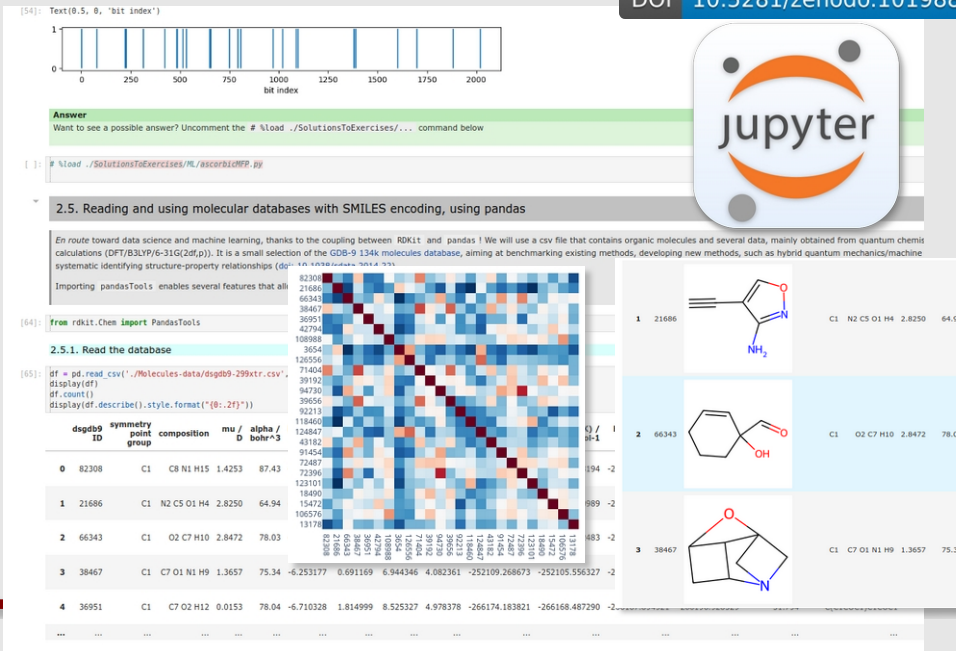


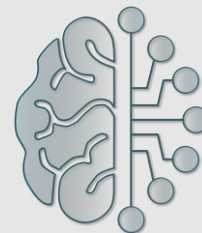
"Talktorials" in physical chemistry and data science / machine learning

S. Christodoulou, Iann C. Gerber, F. Jolibois, R. Poteau
Python in the Physical Chemistry lab (pyPhysChem) github repository, release v. 1.8.0 (2023)

DOI [10.5281/zenodo.10198844](https://doi.org/10.5281/zenodo.10198844)



- interactive python
- images / videos
- mathematical equations
- enriched text (markdown)



Links >

- pdf documents
- data

Motivation

integration of verbal explanations with numerical demonstrations or computer algebra system-based demonstrations proves to be an influential pedagogical tool

let's call them “talktorials”



Python in the
Physical Chemistry Lab



[pyPhysChem]



talktorials specifically tailored for computational
chemistry and data science/machine learning



such mixing is not new, but until recently it was restricted to rather simple applications that required a great deal of development effort

What is new is:

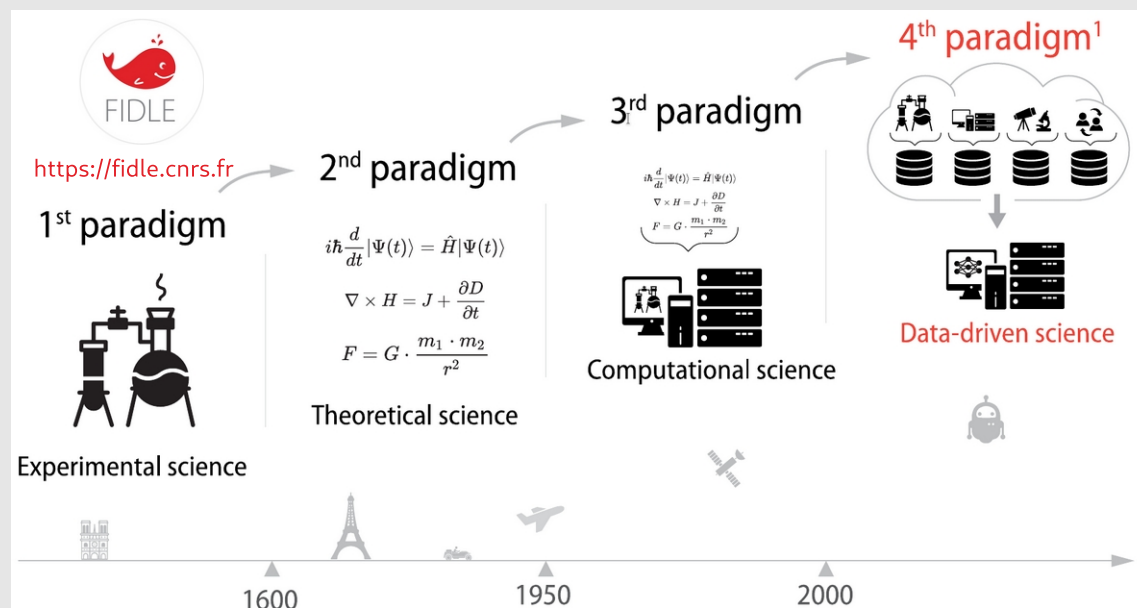
- the combination of Python's popularity and libraries
- the interactive nature of Jupyter Notebooks
- personal computers performance; the prevalence of real-world applications that can quite easily be adapted for students thanks to Python libraries available in a lot of domains
- the strong community support
- the ease of reproducibility that makes tutorials more effective, as learners can directly use the code provided to experiment and build upon it

Motivation

The emergence of such innovative approaches in the realm of computational chemistry is truly encouraging

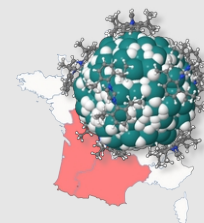
It not only enables learners to grasp theoretical concepts but also offers a practical perspective on their application

For students specialising in computational chemistry who develop their own scripts, they acquire a **dual skill set that could be sought after in various areas of research and industry**



Which students (so far...)?

masters' (graduate) degrees



Réseau Français de
Chimie Théorique

bachelor degree

all chemistry students, 2nd year



Traitement statistique de données
(data science pour débutants)

*Statistical treatment of data
(data science for beginners)*

Lecture et analyse de la base de données "iris" par la bibliothèque pandas

Reading and analysis of the "iris" database with the pandas library

Ce sujet exploite une base de données souvent utilisée pour l'apprentissage de méthodes statistiques, la base **IRIS** :

- elle regroupe les caractéristiques de trois espèces de fleurs d'Iris : Setosa, Versicolor et Virginica
- la base regroupe 50 observations par espèce (soit 150 **individus**)
- chaque observation repose sur 4 caractéristiques (c'est-à-dire 4 **variables**) : longueur et largeur de sépales ainsi que longueur et largeur de pétales

Un [article wikipedia](#) porte sur ce dataset, qui contient à la fois des données numériques (largeur & longueur de pétales et sépales) et descriptives (types d'iris).



This subject uses a database often used for the training of statistical methods, the **IRIS** database:





Python in the
Physical Chemistry Lab



[pyPhysChem]



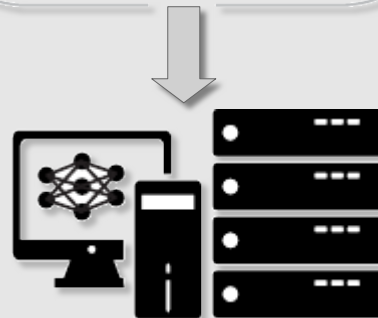
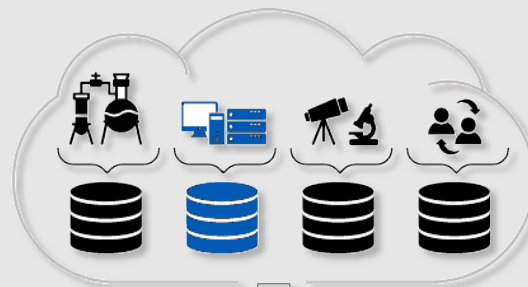
<https://github.com/rpoteau/pyPhysChem>

DOI 10.5281/zenodo.10198844

- python for physicists and chemists in a nutshell
- Computer Algebra System
- Physical chemistry (incl. quantum chemistry)
- coding and use of representations of molecular structures and related data
- Data science and ML (supervised and unsupervised learning, mainly NN)

**strong self-learning
&
project-based learning
dimension**

can we expect a strong convergence between
quantum and computational chemistry, data
science and machine learning?



data-driven science

black boxes?



YES and no... better
explainability of models
than usually supposed



increase in the number of students in the master's programmes in theoretical and computational chemistry?

... unless we do not really give them a dual skill set that could be sought after in various areas of research and industry



Summary & Outlook: digital twins

Version : Thursday 23 November 2023



CNRS Presse CNRS Info Agenda Er

ACTUAL

[Accueil](#) / [Appels à projets](#) / Jumeaux numériques : nouvelles frontières et futurs développements

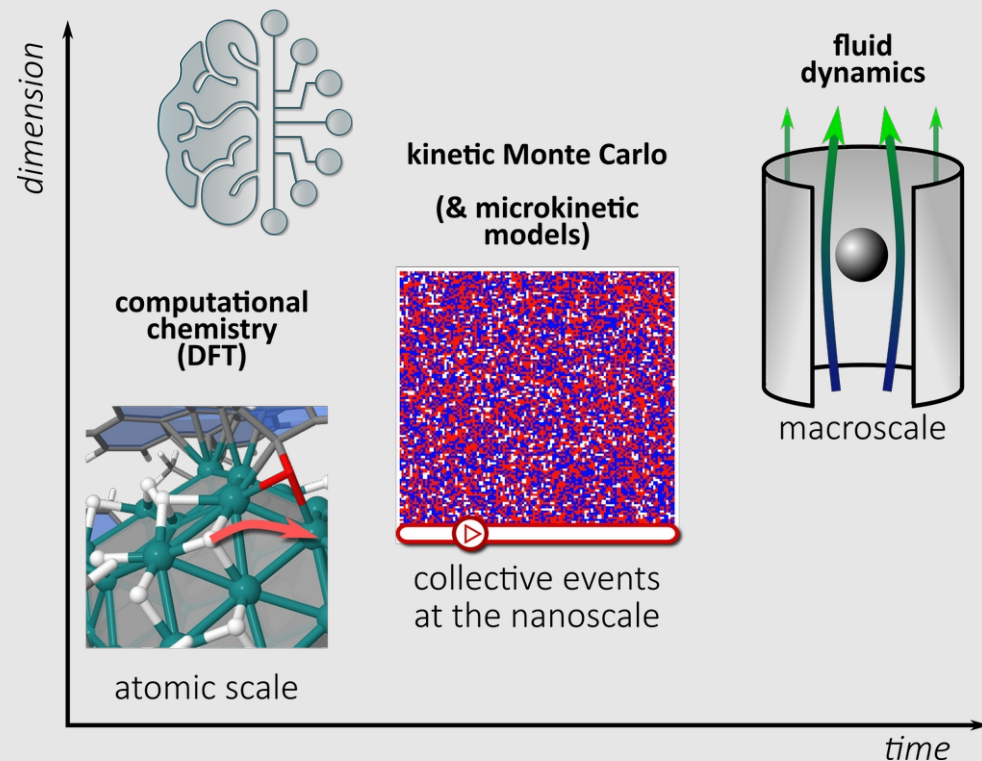
Jumeaux numériques : nouvelles frontières et futurs développements



La MITI a pour objectif de promouvoir, animer et coordonner l'interdisciplinarité au CNRS, et en particulier l'interaction entre ses dix instituts. Dans ce cadre, elle lance en 2024 l'appel à projets « **Jumeaux numériques : nouvelles frontières et futurs développements** ».

Le concept de jumeau numérique est relatif à la représentation virtuelle en temps réel d'un objet, d'un processus, d'un milieu, d'un instrument scientifique, d'un ensemble de mécanismes voire d'un système complexe du monde réel. Ce concept, assez ancien, suscite un regain d'intérêt certain des communautés scientifiques et du monde industriel à l'ère de la production massive de données. La capacité des jumeaux numériques à accompagner la prise de décision des acteurs de terrain sur la base de scénarios variés ouvre de nombreuses applications, dans de multiples domaines : médecine, patrimoine, urbanisme, ingénierie, environnement, climat...

Le but du présent appel à projets est de soutenir des projets relatifs aux **nouvelles frontières méthodologiques d'une recherche interdisciplinaire basée sur le développement de jumeaux numériques**. Ces frontières pourront concerner différents angles détaillés dans l'appel. Les projets devront associer au moins deux équipes issues de disciplines différentes et devront démontrer en quoi leurs travaux se distinguent d'une problématique de modélisation numérique classique, ou de l'utilisation d'un jumeau numérique existant.



création de diplômes communs avec le génie des procédés ?