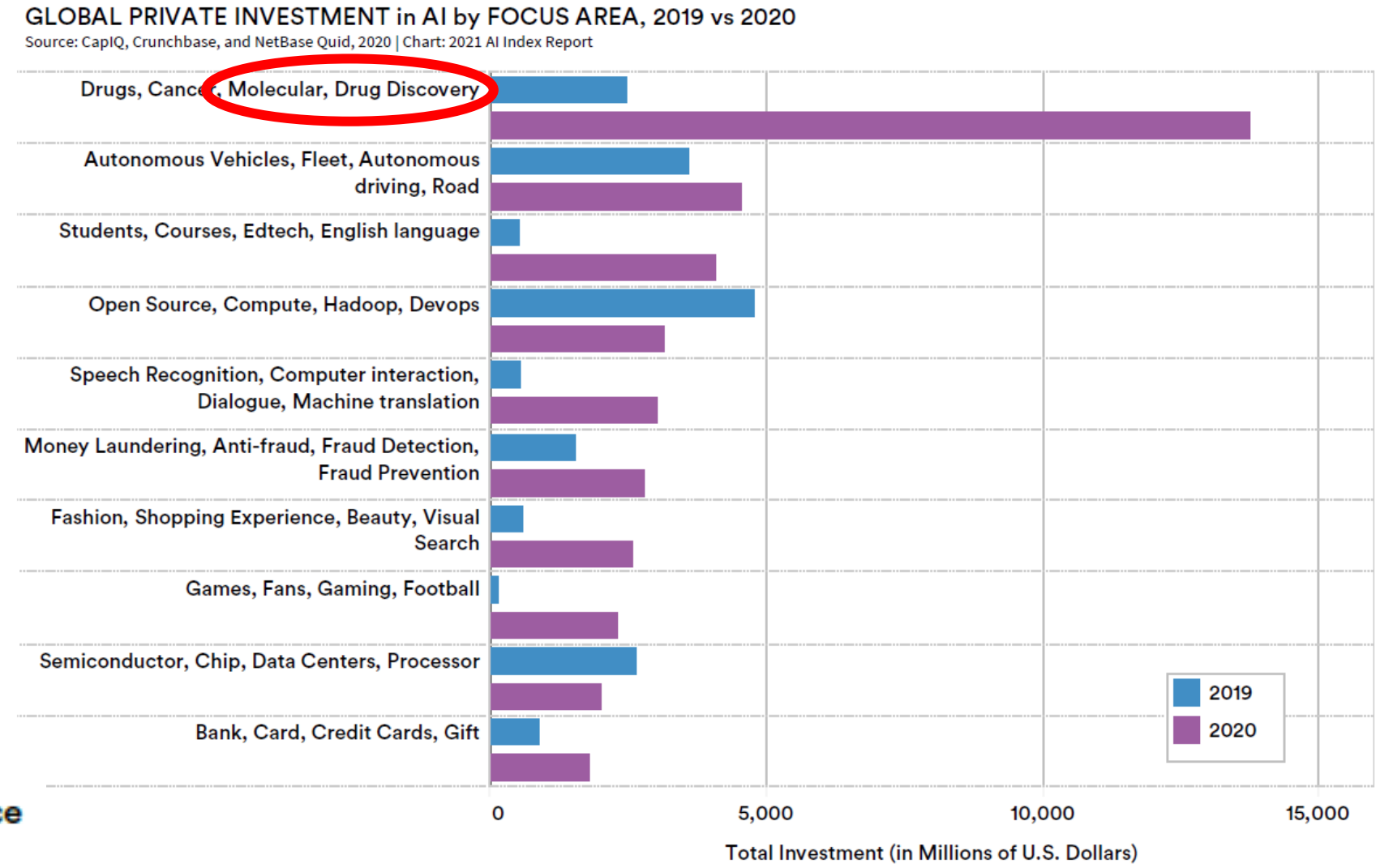
Initiation à l'apprentissage automatique en science des matériaux

6. Datasciences

J.-C. Crivello, LINK : jean-claude.crivello@cnrs.fr
C. Barreteau, ICMPE : celine.barreteau@cnrs.fr

https://link.cnrs.fr/ML/

IA and chemestry



OECD

Artificial Intelligence Index Report 2021

Machine learning notions

1. Supervised / nonsupervised / renforcement
2. Classification / Regression
3. Learning models, penalties, optimization
4. Training / Testing / CV
5. Biais / variance / overfitting → learning curves
6. Metrics / Loss functions
7. Neural networks

Machine learning resources

1. CNRS FIDLE: https://www.youtube.com/c/CNRSFormationFIDLE
2. Coursera ML by Andrew NG: https://www.coursera.org/learn/machine-learning
3. Google AI: https://ai.google/education/
4. Open Classrooms: https://openclassrooms.com/fr/courses/4011851-initiez-vous-au-machine-learning
5. Scikit-learn: https://scikit-learn.org/stable/user_guide.html

…

Conclusions 1/2

Review of machine Learning applied to chemistry and material sciences :

- Large field of applications
- Golden age of Discriminative approaches …
- Welcome to the Generative approaches age !

Conclusions 2/2

- Artificial intelligence is not able to think instead of human
- IA is just efficient for a dedicate learning
- We need data!

Human:

100 .10^9 neurons
10^3 synapses/neurons
200.10^3 electric bonds

7 rules for data scientist workflow

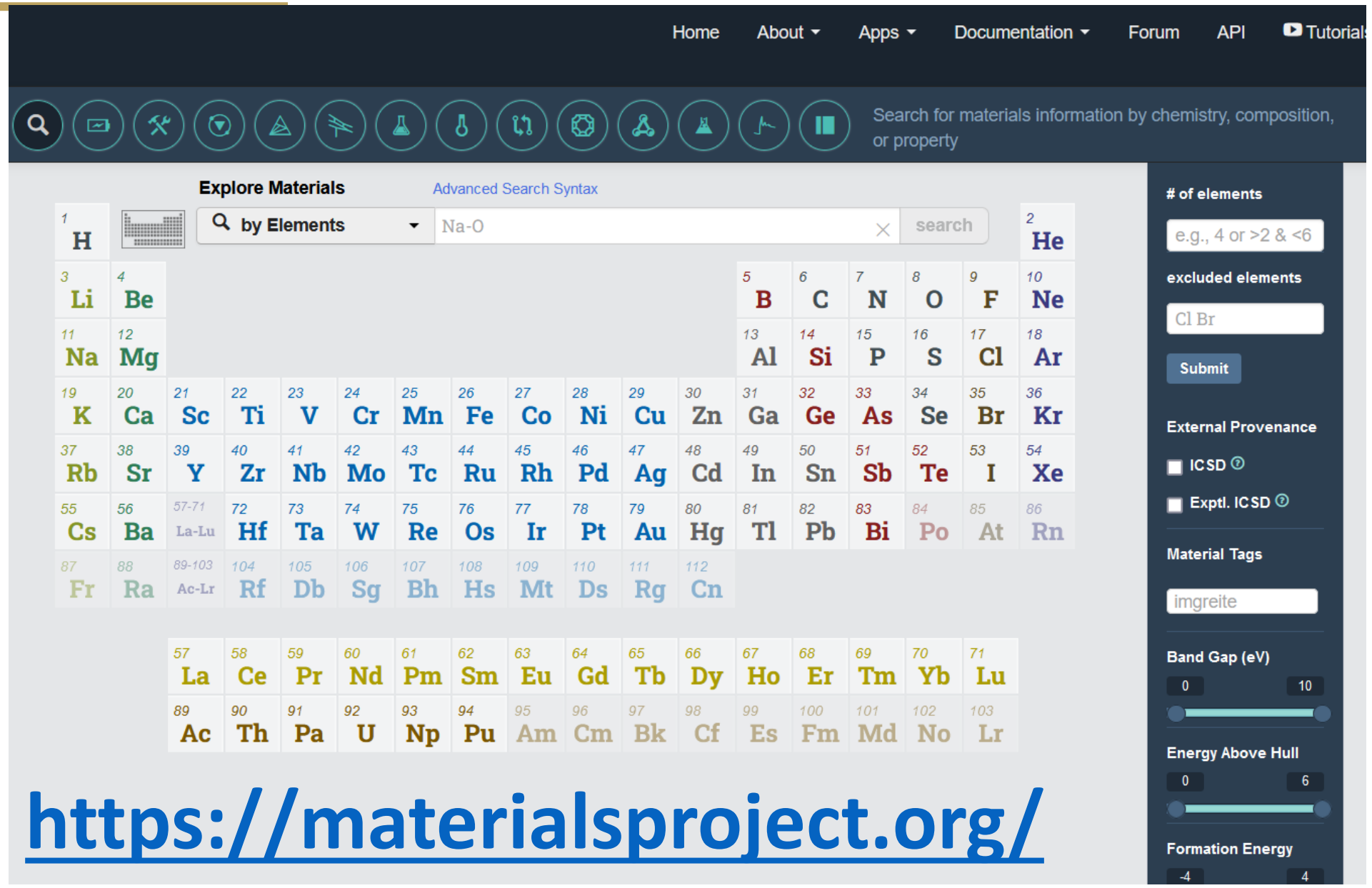
1. Main goal? Definition of the ML problem
2. Collecting data (open data, rights,…)
3. Analyzing data, statistics, visualization
4. Cleaning data: drop missing, normalization, standardization, overlap info?
5. Preparation training/test sets a first training
6. Modelisation, for each model:

- optimization of parameters, grid search

- evaluation of CV score

7. Deployment

Project: based on Material Project



https://materialsproject.org/