Research data management and ELNs

ACS Fall 2023 – August 14th 2023

Who am I

Nicolas CARPI

- Worked 15 years in a cell biology research lab
- Creator of **eLabFTW** ELN software
- Founder and President of **Deltablot** company

Deltablot

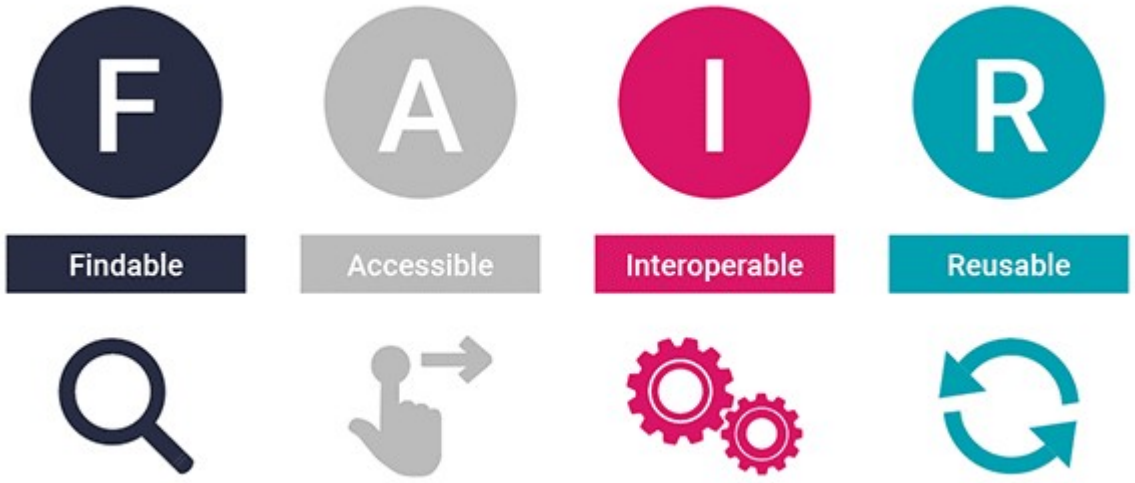
ACS Fall 2023 – August 14th 2023

What’s an ELN?

ELN = Electronic Laboratory Notebook

- Vastly superior to paper notebooks
- First place where research data is referenced (close to the bench)
- Daily tool of any researcher in any branch of Science

What’s FAIR



ACS Fall 2023 – August 14th 2023

100% FAIR is utopia

- Research is messy by design
- Labs are highly dynamics, people come and go all the time
- Improving the FAIRness MUST NOT be a hassle and should be as automatic/transparent as possible

=> Using a progressive enhancement approach is the most likely to be successful

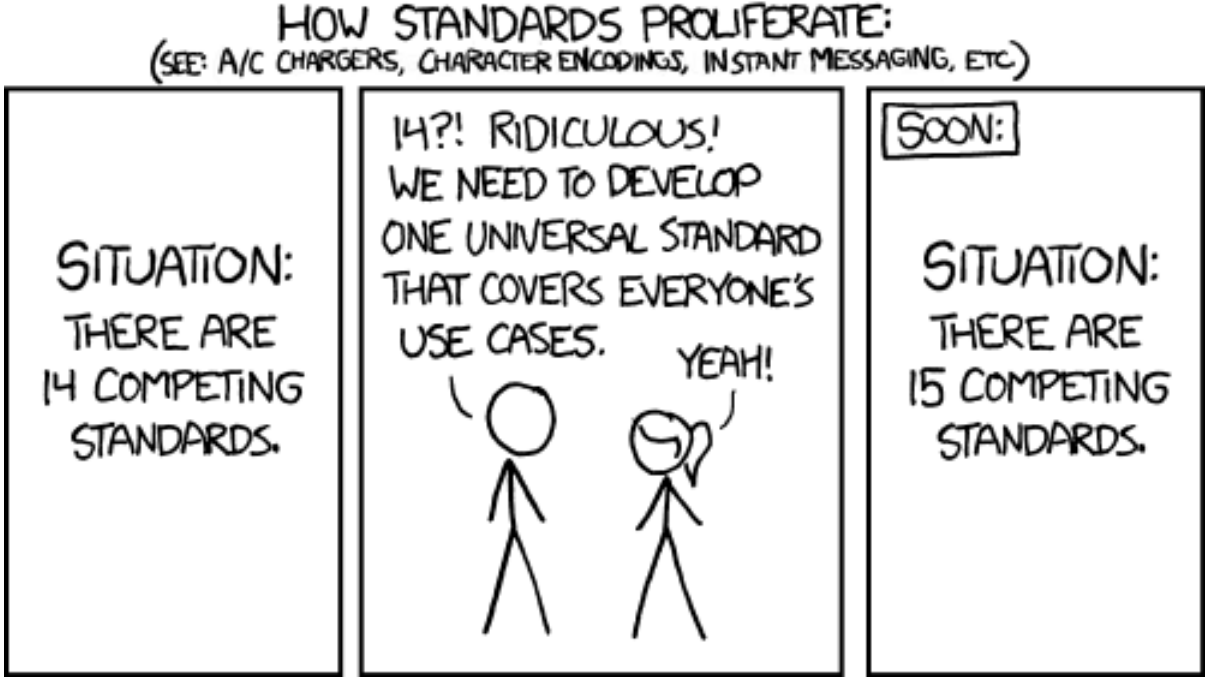
First step is to get the data in the ELN

- Make it easy to input data
- It’s okay if the data doesn’t have 184 ontology defined fields
- It’s okay if it’s incomplete, at least we have something
- Now we can search for it : that’s the F

Making it Accessible

- That’s the job of a data repository, and should happen after publication
- eLabFTW’s approach to research data is private by default

Making it interoperable



https://xkcd.com/927

Making it interoperable

https://the.elnconsortium.org

We need to describe research data in a way that any program can understand.

The ELN Consortium: we'll use RO-Crates and ZIP archives. => re-use existing technology!

RO-Crate contains *ro-crate-metadata.json* using **specific terminology** to describe the content of the archive.

The ELN file format is officially recognized by IANA (*application/vnd.eln+zip*).

Making it interoperable

Result:

- export from one ELN to another
- export from an ELN to a data repository
- export from anything to anything that can open a zip and read JSON

https://the.elnconsortium.org

Making it interoperable

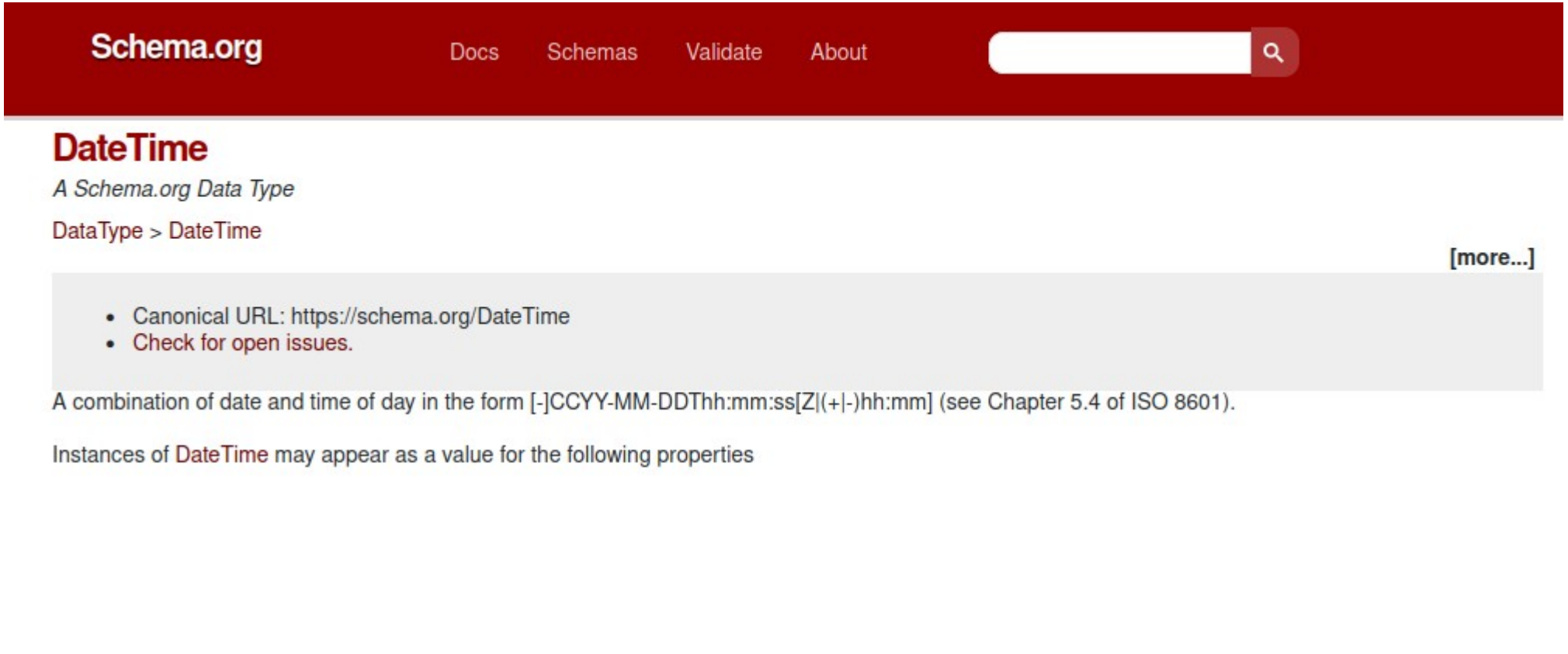
How it works:

- eLabFTW: I use `date`
- Kadi4Mat: I use `start_date`
- PASTA: I use `experiment_date`
- SampleDB: I use `main_datetime`

Problem: we all have our vocabulary for different aspects

Solution: common vocabulary

Making it interoperable



Making it interoperable

Now everyone uses DateTime in ISO8601 format.

We can import/export from different software.

Want to share a sample? Share also the .eln so the history of its creation is not lost!

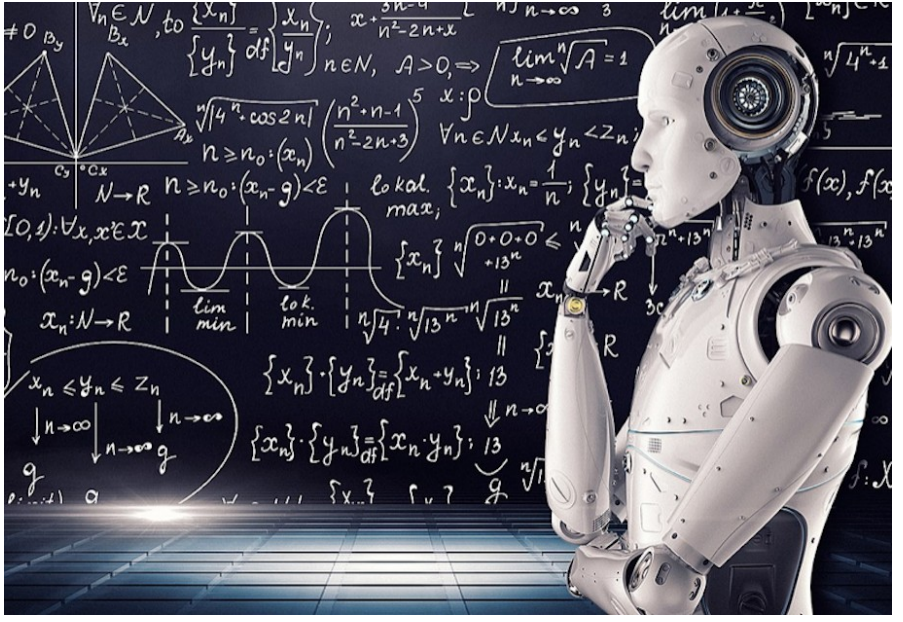
Want to publish a project along with your scientific article? Export all the related entries in a .eln and push that to the data repository which can interpret all the metadata correctly.

Making it reusable

The research data is now convenably annotated

More importantly: the experiment that produced that data is available with all the conditions and information needed to reproduce it

We can do some machine learning AI magic on it!



Next steps

Improve automatic metadata extraction from raw data

Improve ELN comprehension from Dataverse

Increase ELN support across the research software space

Use more precise ontology (bioschemas.org, ...)

Links

- eLabFTW ELN: https://www.elabftw.net
- The ELN Consortium: https://the.elnconsortium.org
- This presentation