GEO
GROUP ON
EARTH OBSERVATIONS

OPEN DATA
& OPEN KNOWLEDGE
Workshop

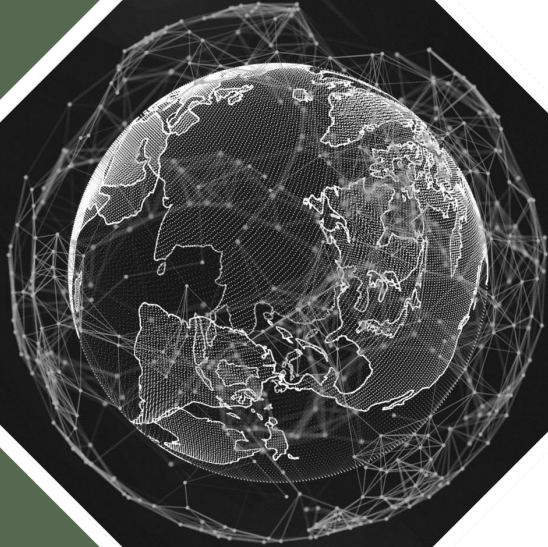**Encoding heterogeneous in-situ data**

2023-06-14

Raphaël Jolivet – MINES Paris PSL University – ARMINES –
raphael.jolivet@minesparis.psl.eu

Raphaël Jolivet

Research Engineer

O.I.E (Observation, Impacts, Energy) @ **Sophia Antipolis**

■ Assessment of **solar resources** from observation of earth

■ Life cycle assessment of **environmental impacts** of energy systems

Encoding heterogeneous in situ data

Common challenges for in-situ data

- **Various providers** of data
- Heterogeneous data : **format**, **protocol**, **timezone**, **units, ...**
- Lack of embedded **meta-data**
- Various policies (licenses)

Encoding heterogeneous in-situ data

Main motivation : what you get ....

Index of /aftp/data/radiation/solrad/abq/2002/

Name	Last modified
Parent Directory	
abq02032.dat	2015-02-05 15:41
abq02033.dat	2015-02-05 15:47
abq02034.dat	2015-02-05 15:47
abq02035.dat	2015-02-05 15:47
abq02036.dat	2015-02-05 15:47
abq02037.dat	2015-02-05 15:41
abq02038.dat	2015-02-05 15:41
abq02039.dat	2015-02-05 15:41
abq02040.dat	2015-02-05 15:41
abq02041.dat	2015-02-05 15:41
abq02042.dat	2015-02-05 15:41
abq02043.dat	2015-02-05 15:41
abq02044.dat	2015-02-05 15:41
abq02045.dat	2015-02-05 15:41
abq02046.dat	2015-02-05 15:41

- Mean of Access: **FTP**, HTTP

- Format: *.dat, csv, ..

Main motivation

	Input : heterogenous
Format	X Mixed (csv, zip, dat, ...)
Granularity	X Mixed (daily, monthly, ...)
Access	X Mixed (http, ftp, …)
Timezone	X Mixed (**local**, **utc**)
Units	X Mixed, implicit
Metadata	X Missing
Standard	X None
Compression	X Optional / external (zip)

Solution

- A **methodology** following **FAIR** principles

- A draft **standard** for **in-situ data,** based on **other standards**

- An open source **library** : *libinsitu,* implementing it

- **Deployed** "live" for our field of research (solar energy)

Solution

Practical data workflow

Raw data from **data logger**
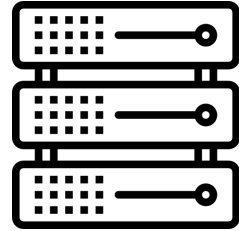
Open standard encoding and **access** protocol

Agnostic access

Dissemination **search** & **discovery**

R
REUSABLE

A
ACCESSIBLE

I
INTER-OPERABLE

F
FINDABLE

"Metadata and data should be well-described"

"User needs to know how data and metadata can be accessed"

"Data need to interoperate with applications"

"Metadata and data should be easy to find"

NetCDF format with CF Conventions

Thredds Data Server

Open Source Web Clients

GEO Portal and GEO Knowledge Hub

Standards

Encoding heterogeneous in situ data
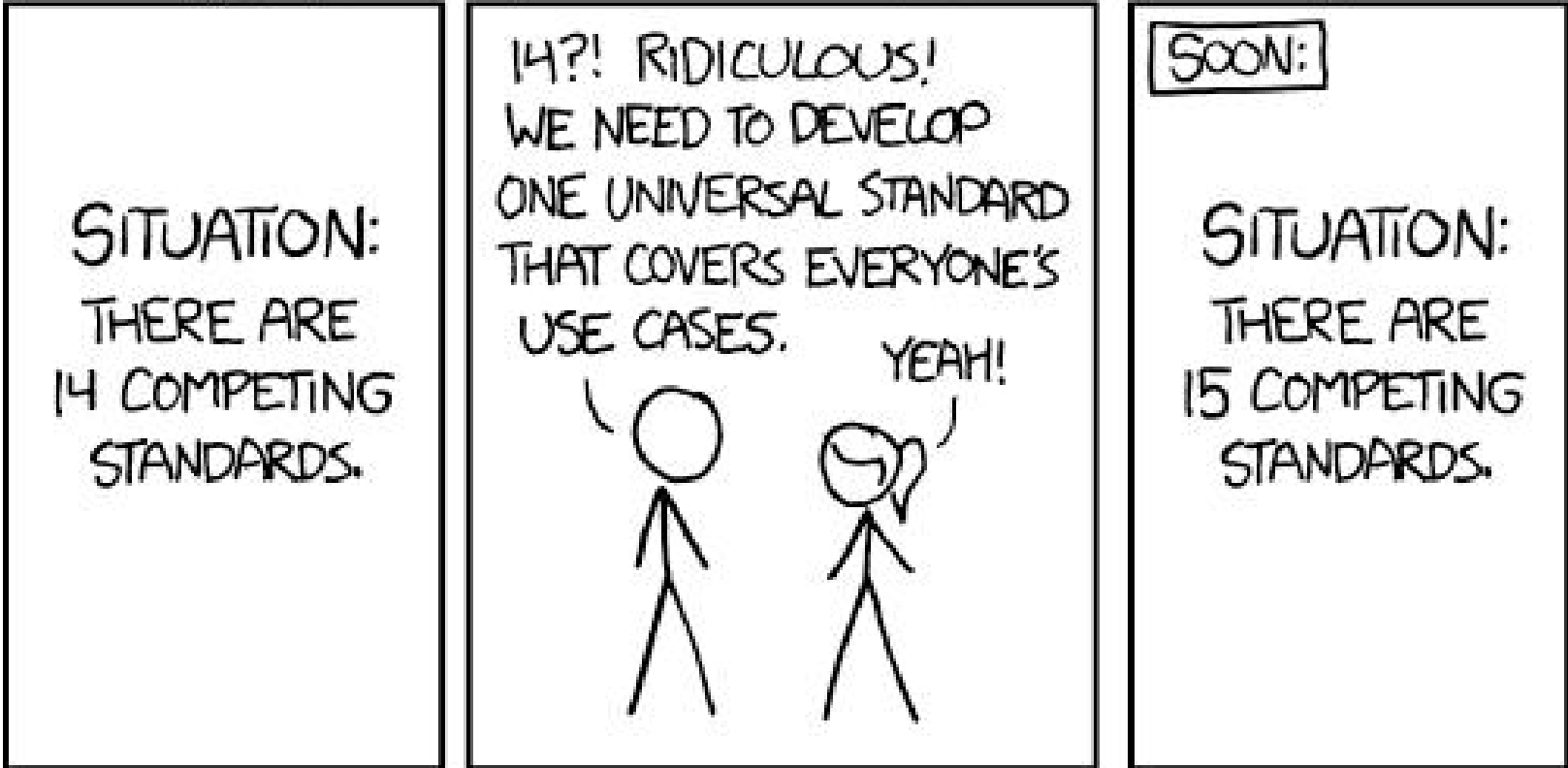
Standards

Proposed standard for in situ data
DRAFT

CF conventions
Semantic of metadata
Standard names for **physical** variables

NetCDF
Efficient binary storage of matrices
Embedded metadata
Remote access (OpenDAP)
Wide support

ACDD (ESIP)
Attribute **C**onvention For **D**ata **D**iscovery
Global Attributes
[Findable]

Encoding pipeline

Solution

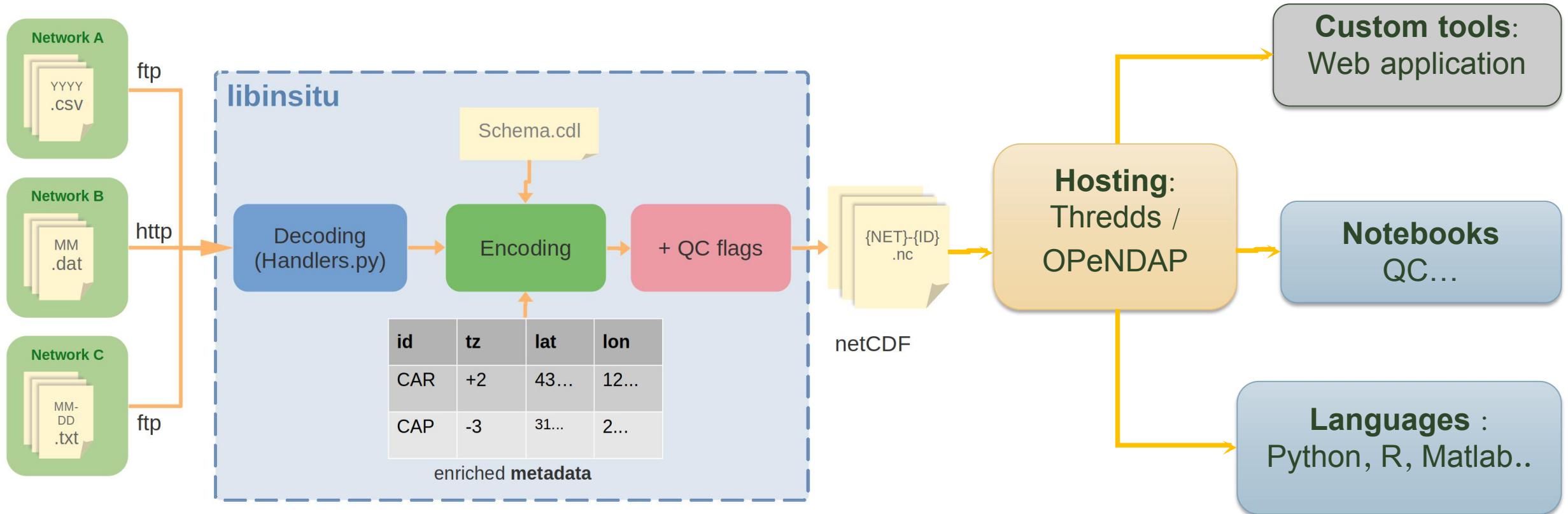|	Input : heterogenous	Output : standardized
Format	X Mixed (csv, zip, dat, ...)	V **NetCDF**
Granularity	X Mixed (daily, monthly, ...)	V Single file per station
Access	X Mixed (http, ftp, …)	V **TDS** & OpenDAP (random subset)
Timezone	X Mixed (**local**, **utc**)	V **UTC**, explicit
Units	X Mixed, implicit	V **SI**, explicit
Metadata	X Missing	V **Embedded**
Standard	X None	V **CF** conventions
Compression	X Optional / external (zip)	V **Native** (~10% ratio)

GROUP ON
EARTH OBSERVATIONS

Demo

Binder : http://bit.ly/3X6e6IG

Take aways

- **Libinsitu** :

  Use it… **It's open, free, documented and exemplified**

- **Replicable** for any domain

- **Tested** and **implemented** by Marine Radioactivity Information System (*MARIS*) community

- **New breath** for in-situ measurement networks towards **standard and interoperable practices**

Resources

Documentation	https://libinsitu.readthedocs.io
Source code	https://git.sophia.mines-paristech.fr/oie/libinsitu/
Methodology	https://doi.org/10.23646/AC2M-8504
Standard	https://libinsitu.readthedocs.io/en/latest/conventions.html
Mailing list	solar-insitu@groupes.mines-paristech.fr
	https://groupes.minesparis.psl.eu/wws/info/solar-insitu

CONTACT DETAILS

✉ EMAIL ADDRESS

raphael.jolivet@minesparis.psl.eu

GEO
GROUP ON
EARTH OBSERVATIONS