



APPROACH

Open Science and Open Data
presentation
October 29, 2025

Jabub Navářík
CATRIN Data Steward

The Approach project receives funding from the European Commission's
Horizon Europe Research Programme under Grant Agreement Number 101120397

The Accelerator project receives funding from the European Commission's
Horizon Europe Research Programme under Grant Agreement Number 101087318



OUTLINE



OPEN SCIENCE

OPEN DATA

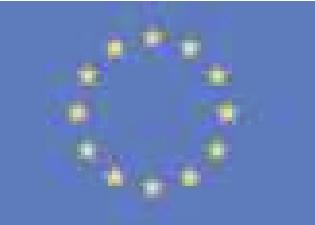
(META)DATA

LABORATORY LEVEL

PUBLICATION LEVEL

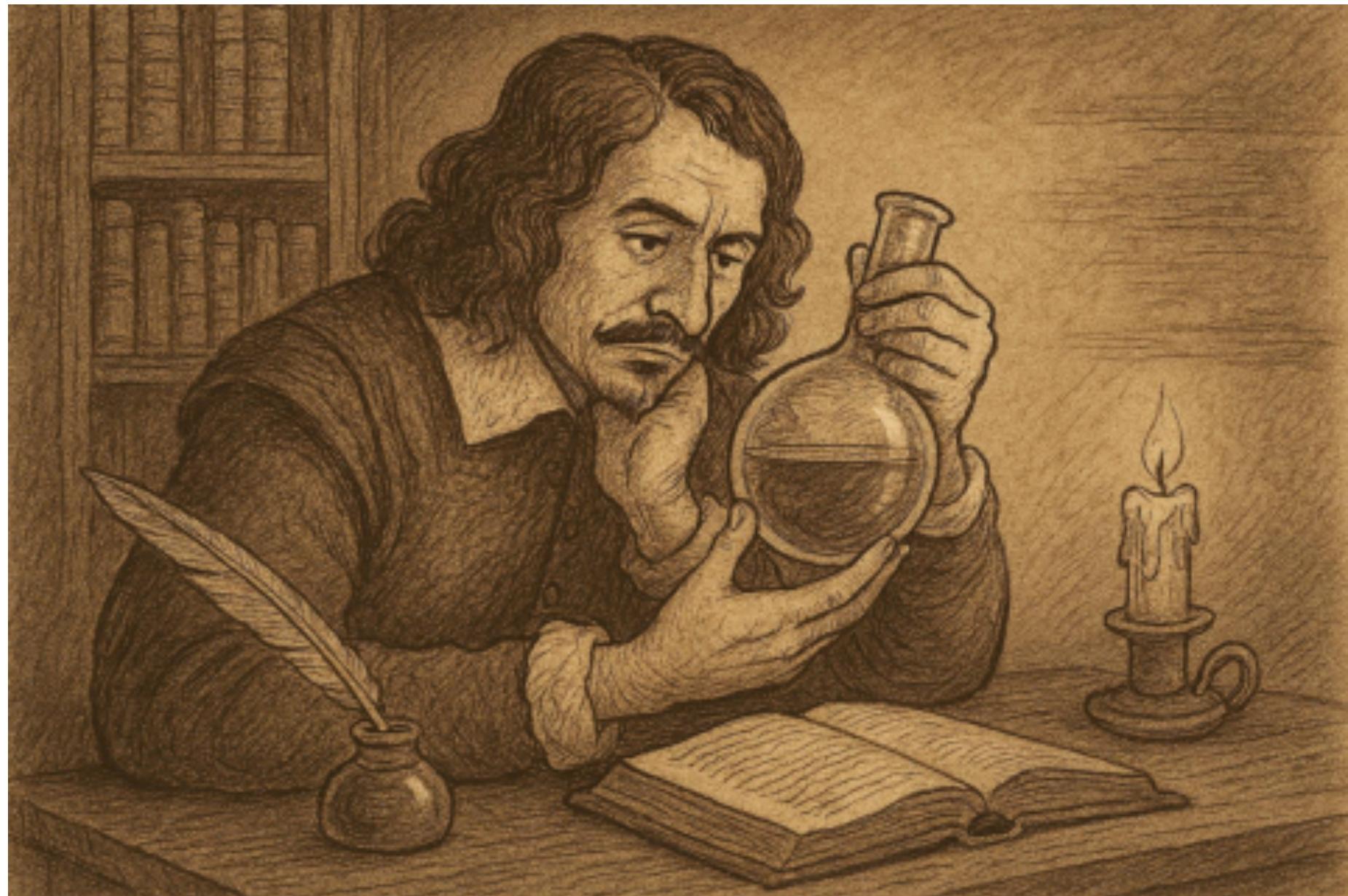
PUBLIC DATA REPOSITORY

QUESTIONS & ANSWERS



History

- 1) Sharing of research results (17th century)
- 2) Paywall (19th – 20th century)
- 3) Secret research (cold war)
- 4) Digital revolution (end of 20th century)
- 5) Open Access (21st century)



Motivation

- 1) Replication crisis
- 2) Fast sharing of data
- 3) Research output increase
- 4) Research quality increase

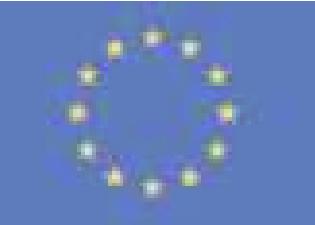


OPEN SCIENCE



RISKS

- 1) Data mining (open data)
- 2) High costs (open access)



OPEN DATA



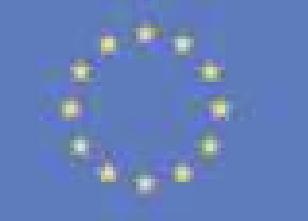
F.A.I.R. Principles

Findable

Accessible

Interoperable

Reusable



Findable data

Laboratory level

- Proper sample identification (unique code, central laboratory evidence, ...)
- All raw data must be accompanied with some kind of metadata (assignability to the sample)
- All measurements must be registered in a laboratory book (hardcopy or electronic)
- A responsible person must be designated

Publication level

- All published data must be safely stored to ensure their full availability during the publication lifetime
- All data must be assignable to individual images, graphs, etc
- A responsible person must be designated (typically corresponding author)

OPEN DATA



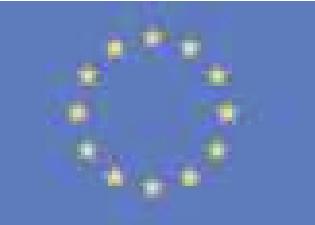
Accessible data

Laboratory level

- Raw data and metadata are available for scientists who are working with them

Publication level

- All published data must be available for broad public
 - Previously on the request upon the containing author's team (typically corresponding author responsibility)
 - New trend is the unlimited public access through online repository



Interoperable data

Laboratory level

- Typically no interoperability request
- Raw data are often processed by specialized software

Publication level

- All published data should be usable with standard software tools and standard PC
 - Using of standard open formats, like: ***.TXT (ASCII), *.DOC, *.XLS, *.HTML, *.JPG, ...**
 - Only exception is if there is no such possibility, then the data could be in method specific format

OPEN DATA



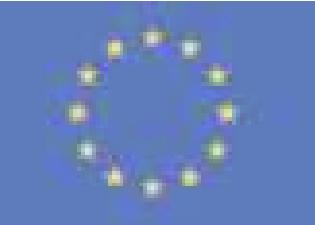
Reusable data

Laboratory level

- Raw data are always reusable from their principle

Publication level

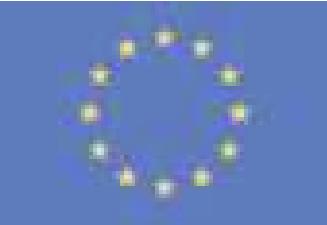
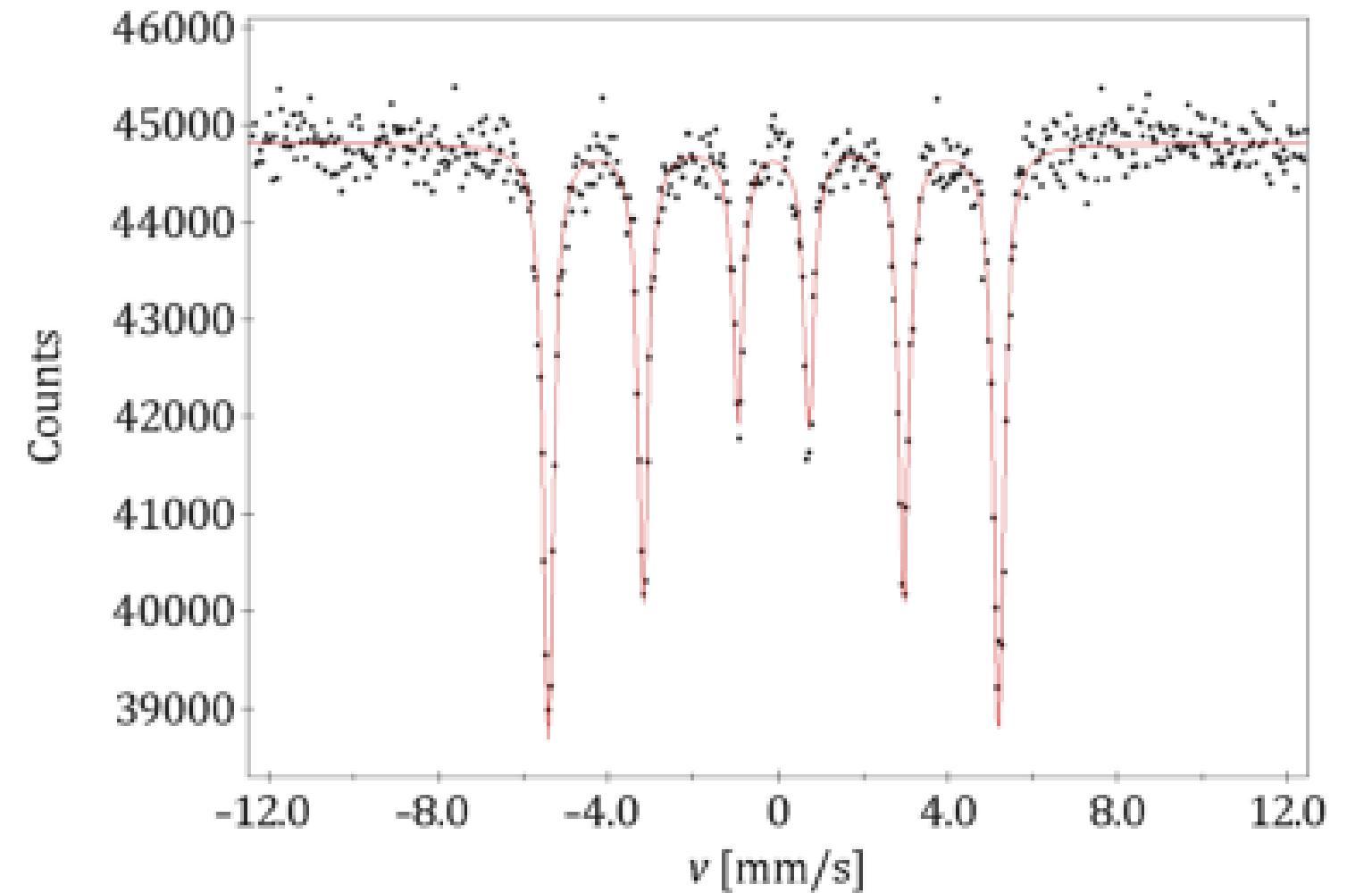
- All published data must be in a format and form which allows their reusage and evaluation by anyone



OPEN DATA



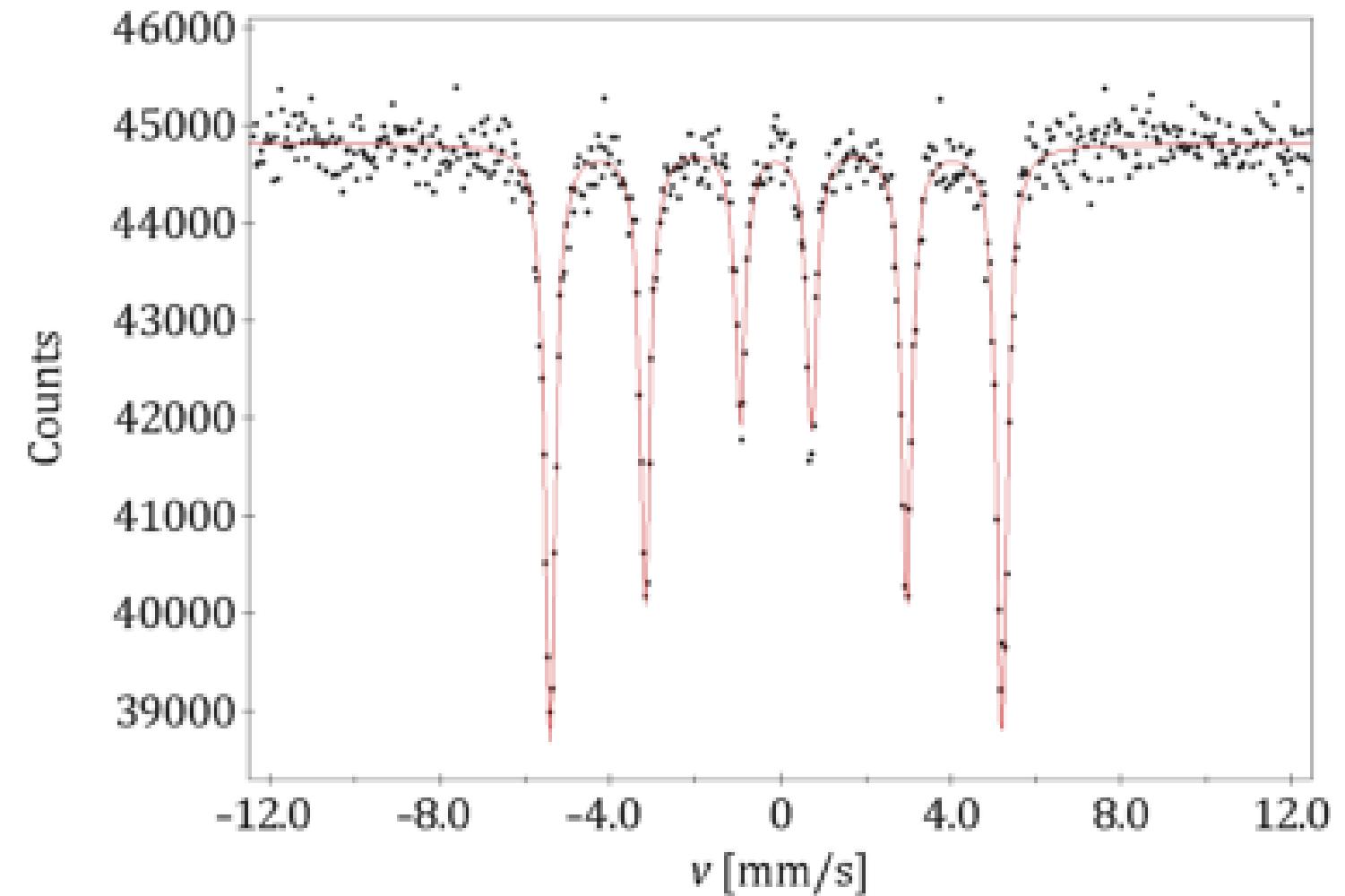
Reusable data example



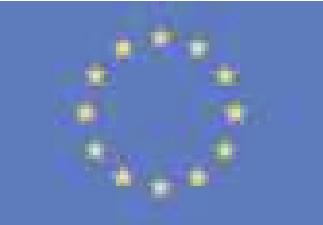
OPEN DATA



Reusable data example



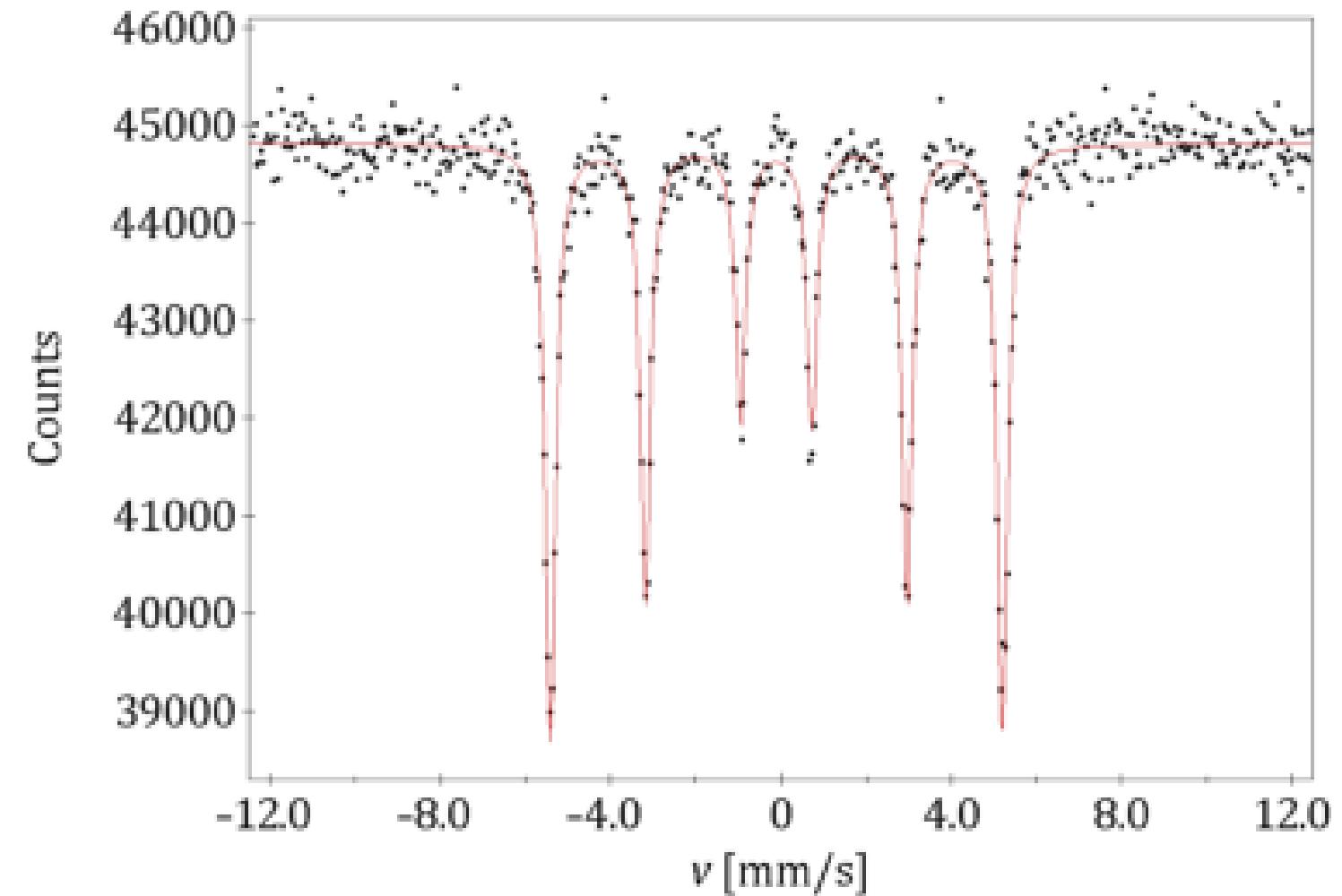
Findable
Accessible
Interoperable
Reusable



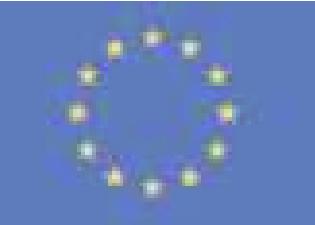
OPEN DATA



Reusable data example



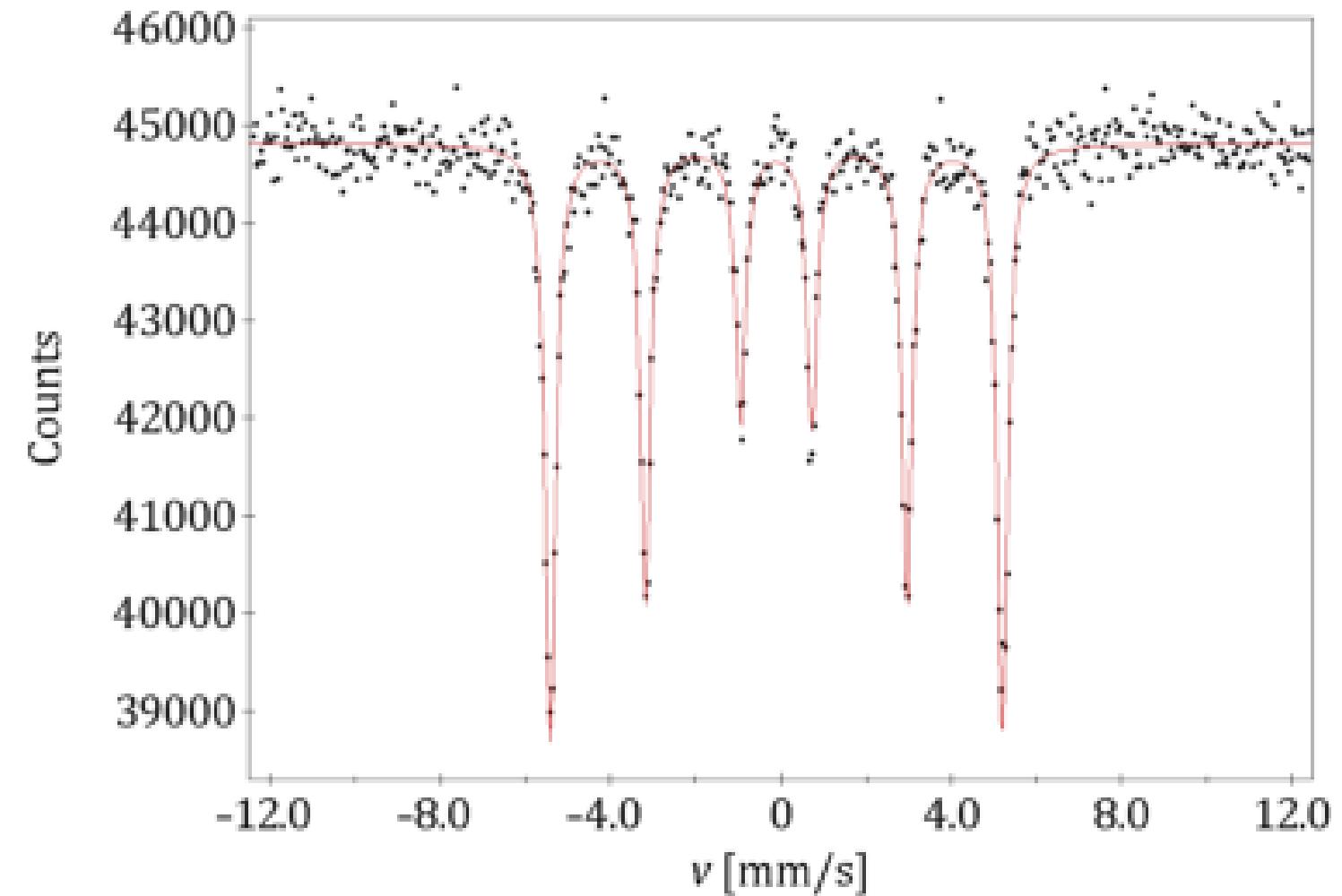
Findable
Accessible
Interoperable
Reusable



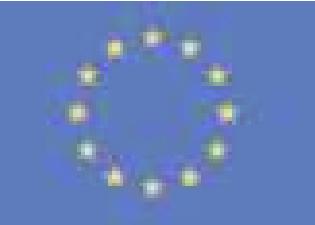
OPEN DATA



Reusable data example



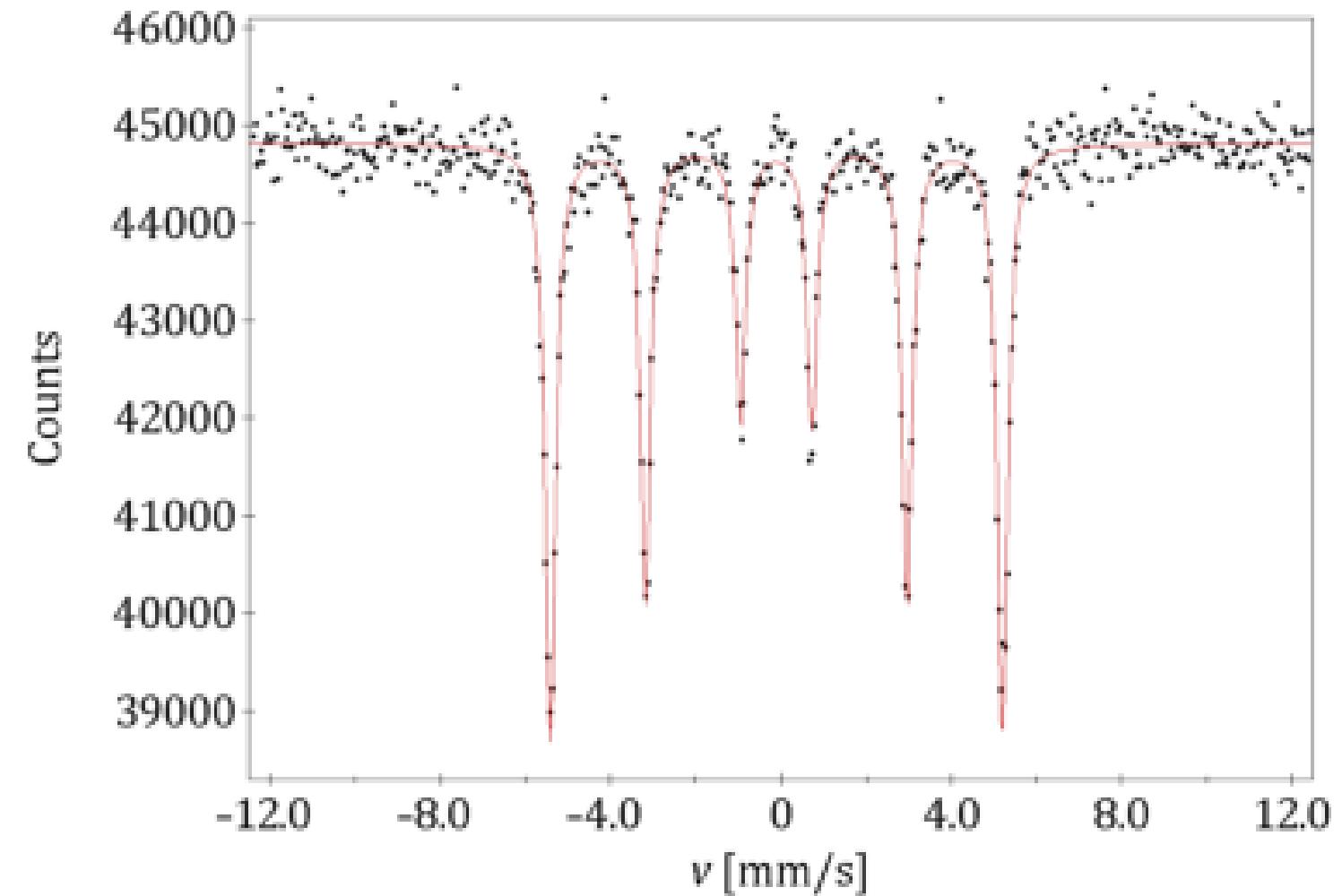
Findable
 Accessible
 Interoperable
 Reusable



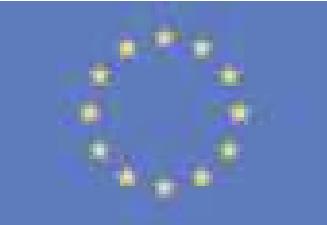
OPEN DATA



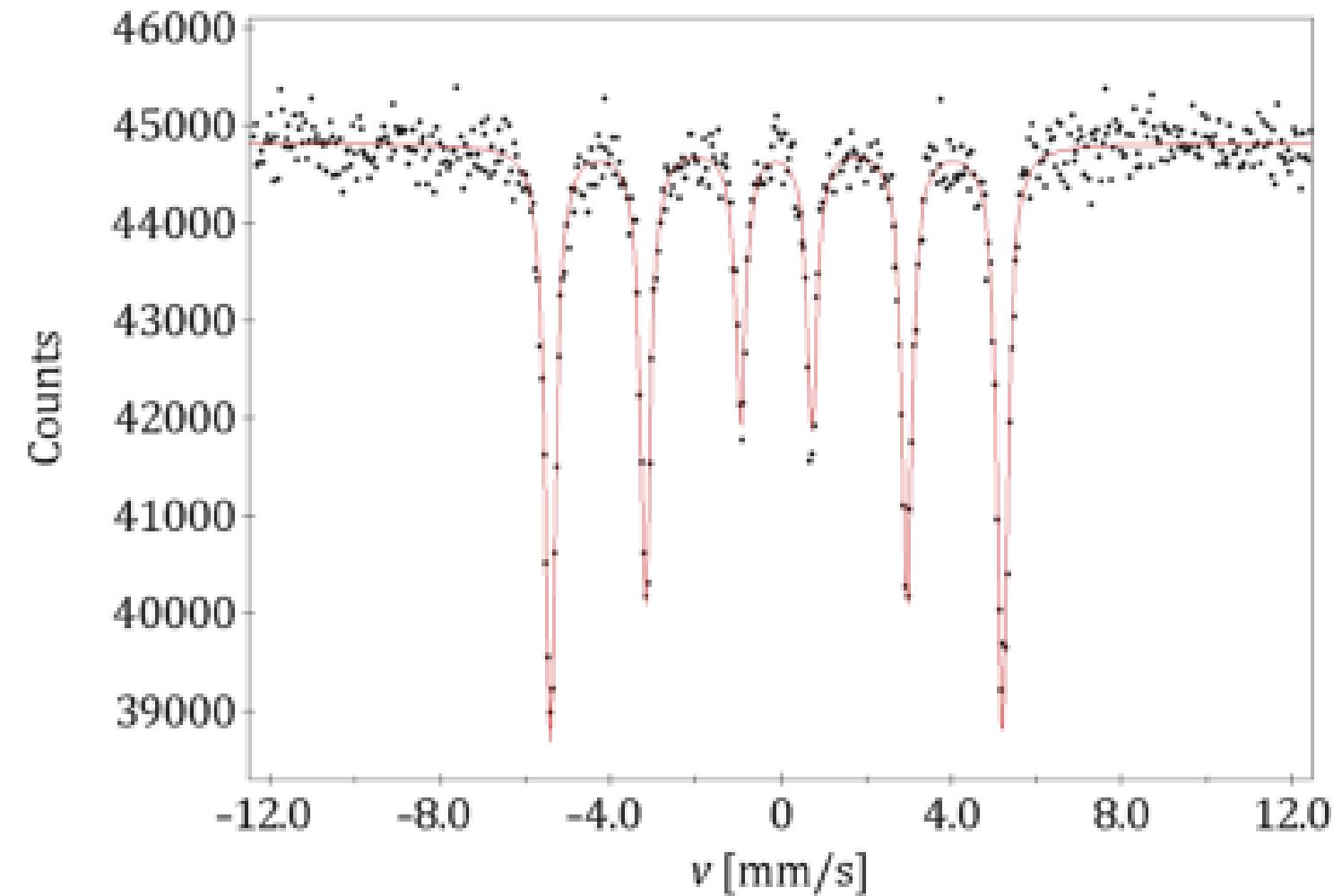
Reusable data example



✓ Findable
✓ Accessible
✓ Interoperable
Reusable

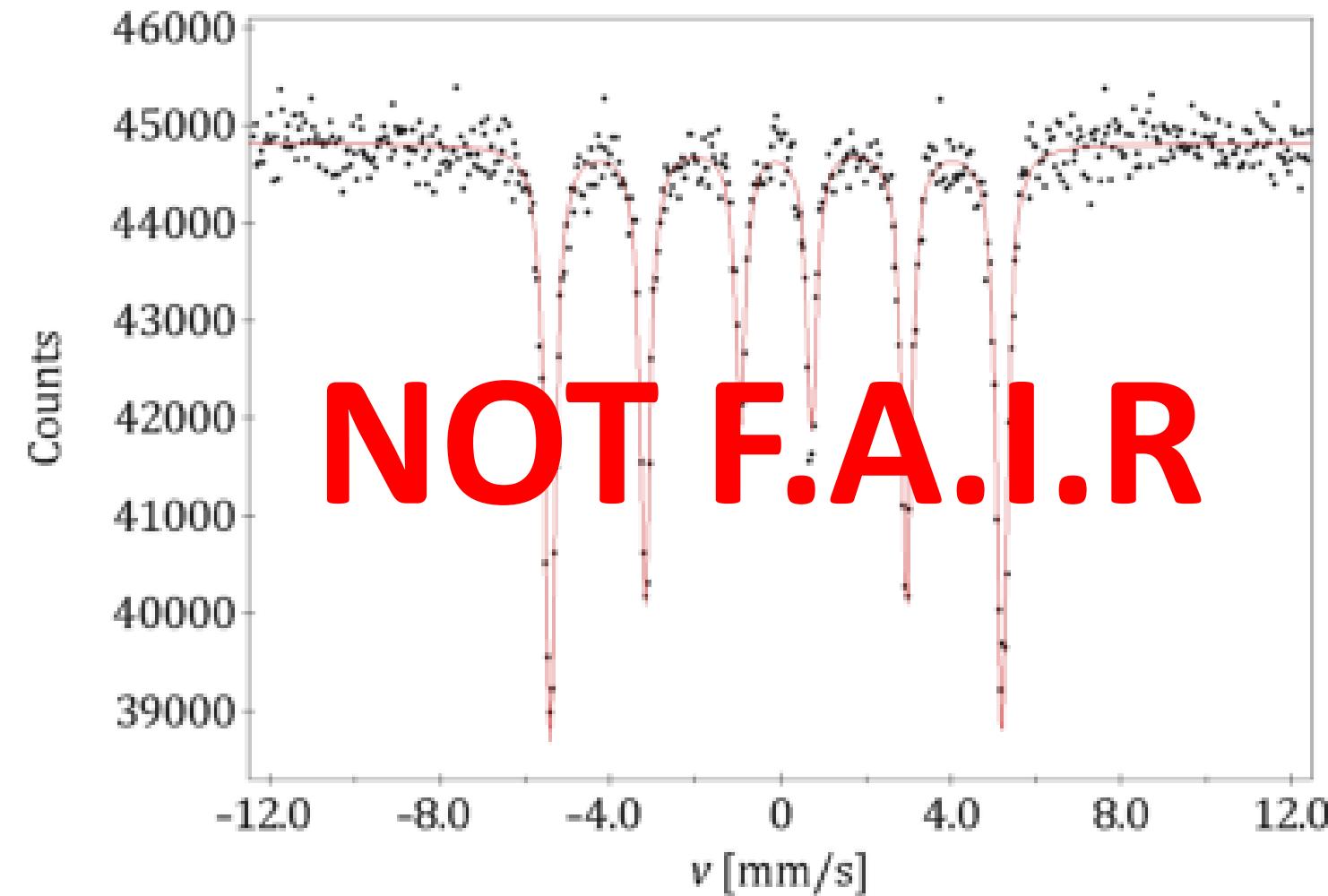


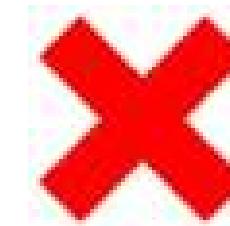
Reusable data example



-  **Findable**
-  **Accessible**
-  **Interoperable**
-  **Reusable**

Reusable data example

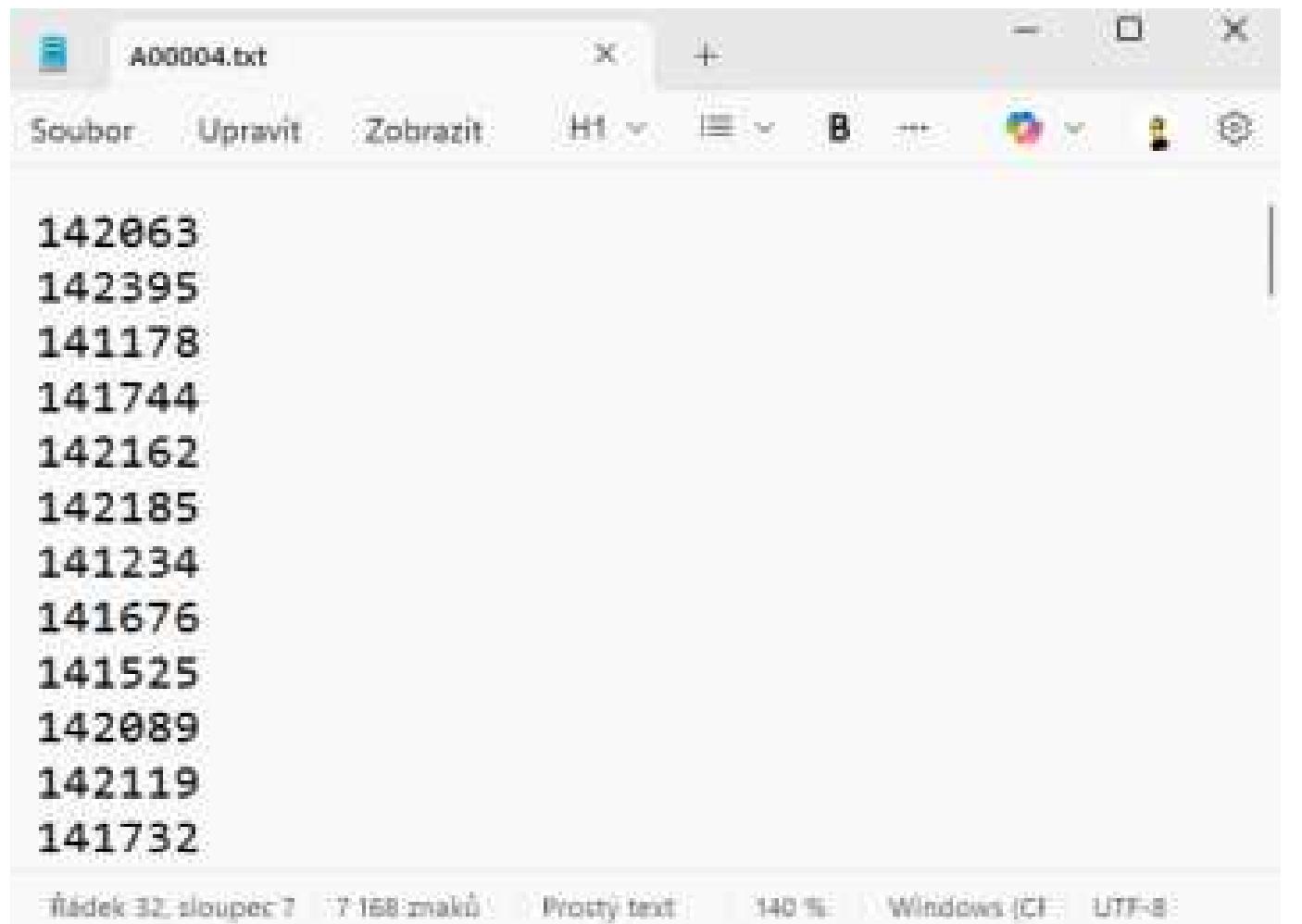
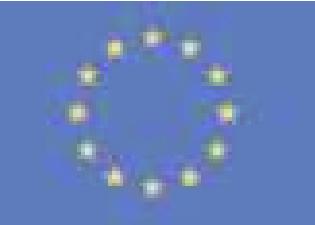


-  Findable
-  Accessible
-  Interoperable
-  Reusable

OPEN DATA

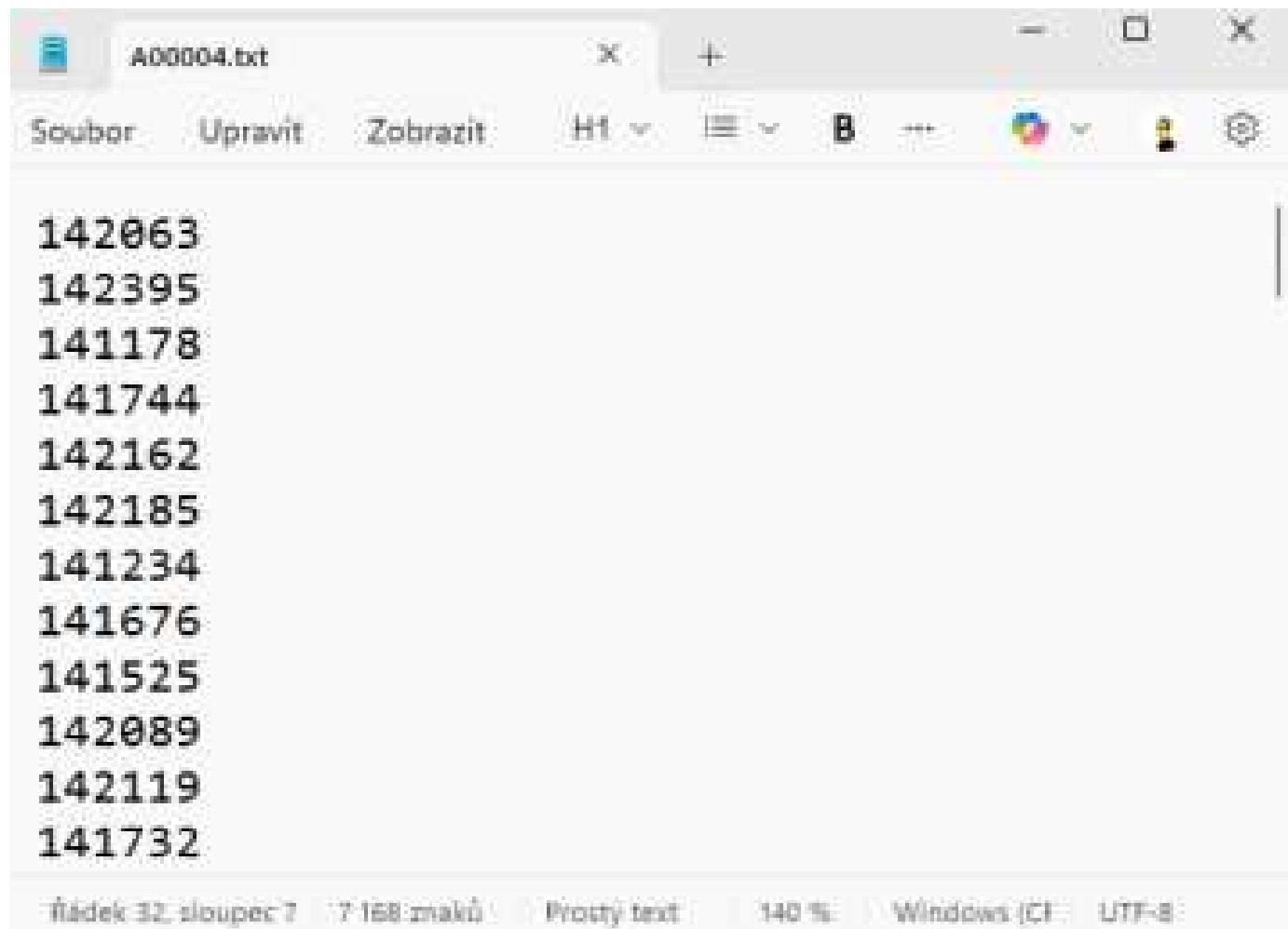


Reusable data example

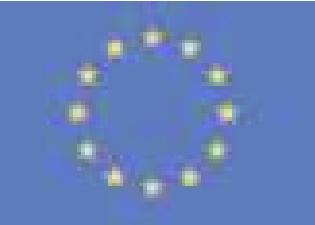
A screenshot of a Windows-style text editor window titled 'A00004.txt'. The window contains a list of 15 numerical IDs, each on a new line. The IDs are: 142063, 142395, 141178, 141744, 142162, 142185, 141234, 141676, 141525, 142089, 142119, and 141732. The text editor interface includes a menu bar with 'Soubor', 'Upravit', 'Zobrazit', and 'Hl', as well as various toolbar icons. The status bar at the bottom shows 'Rádek 32, sloupec 7, 7 168 znaků, Prostý text, 140 %, Windows-1250, UTF-8'.



Reusable data example

A screenshot of a Microsoft Word document titled 'A00004.txt'. The document contains a single column of 14 IDs, each on a new line. The IDs are: 142063, 142395, 141178, 141744, 142162, 142185, 141234, 141676, 141525, 142089, 142119, and 141732. The document has a standard Word header with tabs for 'Soubor', 'Upravit', 'Zobrazit', and 'H1'. The status bar at the bottom shows 'Rámeček 32 sloupec 7 7168 znaků' and 'Příkazy 140 % Windows (C) UTF-8'.

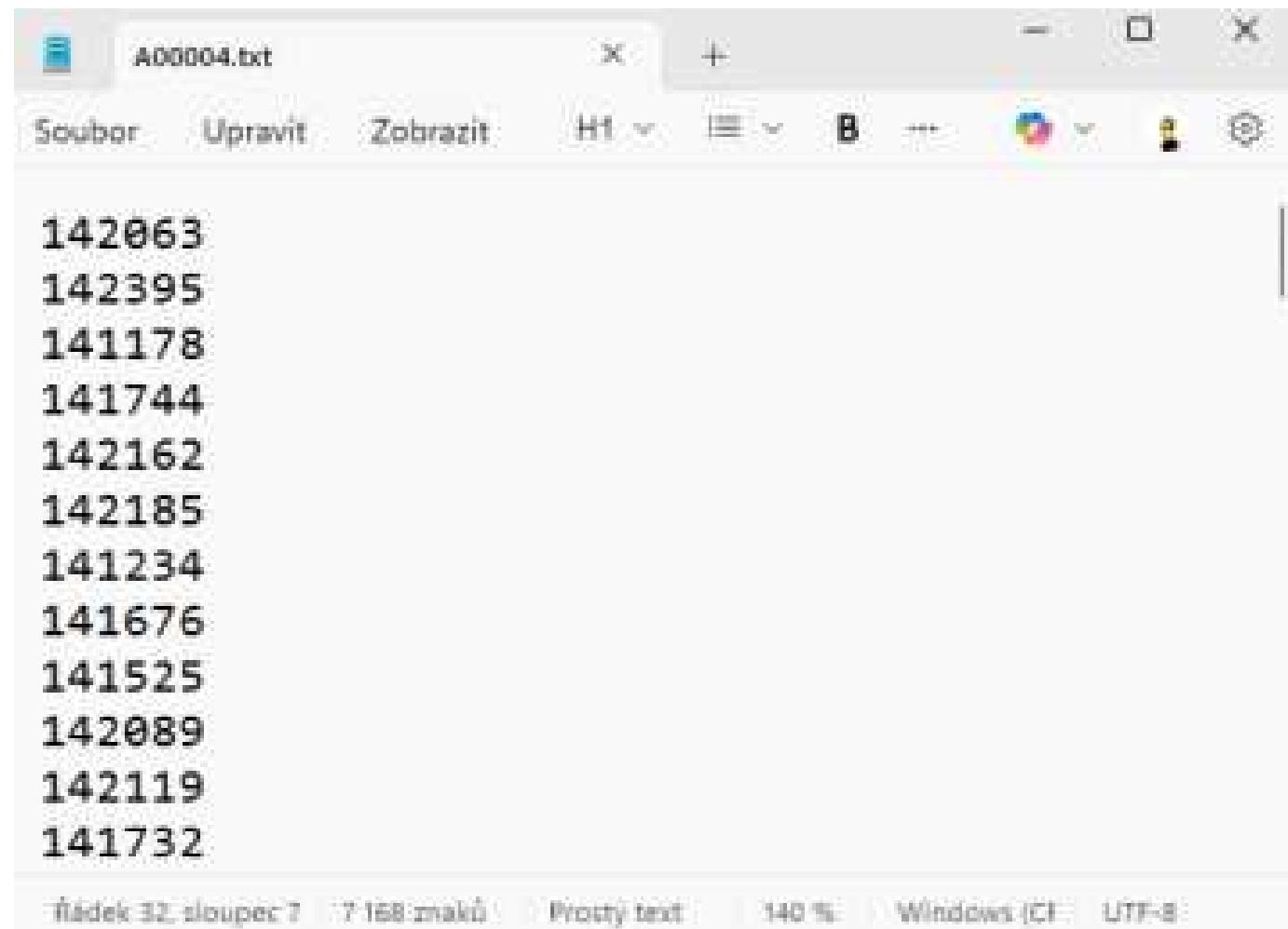
Findable
Accessible
Interoperable
Reusable



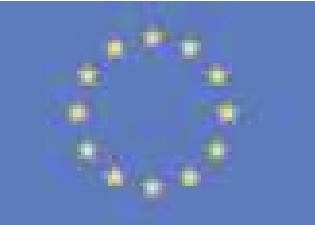
OPEN DATA



Reusable data example

A screenshot of a Microsoft Word document window titled 'A00004.txt'. The document contains a single column of 14 IDs, each on a new line. The IDs are: 142063, 142395, 141178, 141744, 142162, 142185, 141234, 141676, 141525, 142089, 142119, and 141732. The window shows standard Microsoft Word ribbon tabs like 'Soubor', 'Upravit', and 'Zobrazit' at the top.

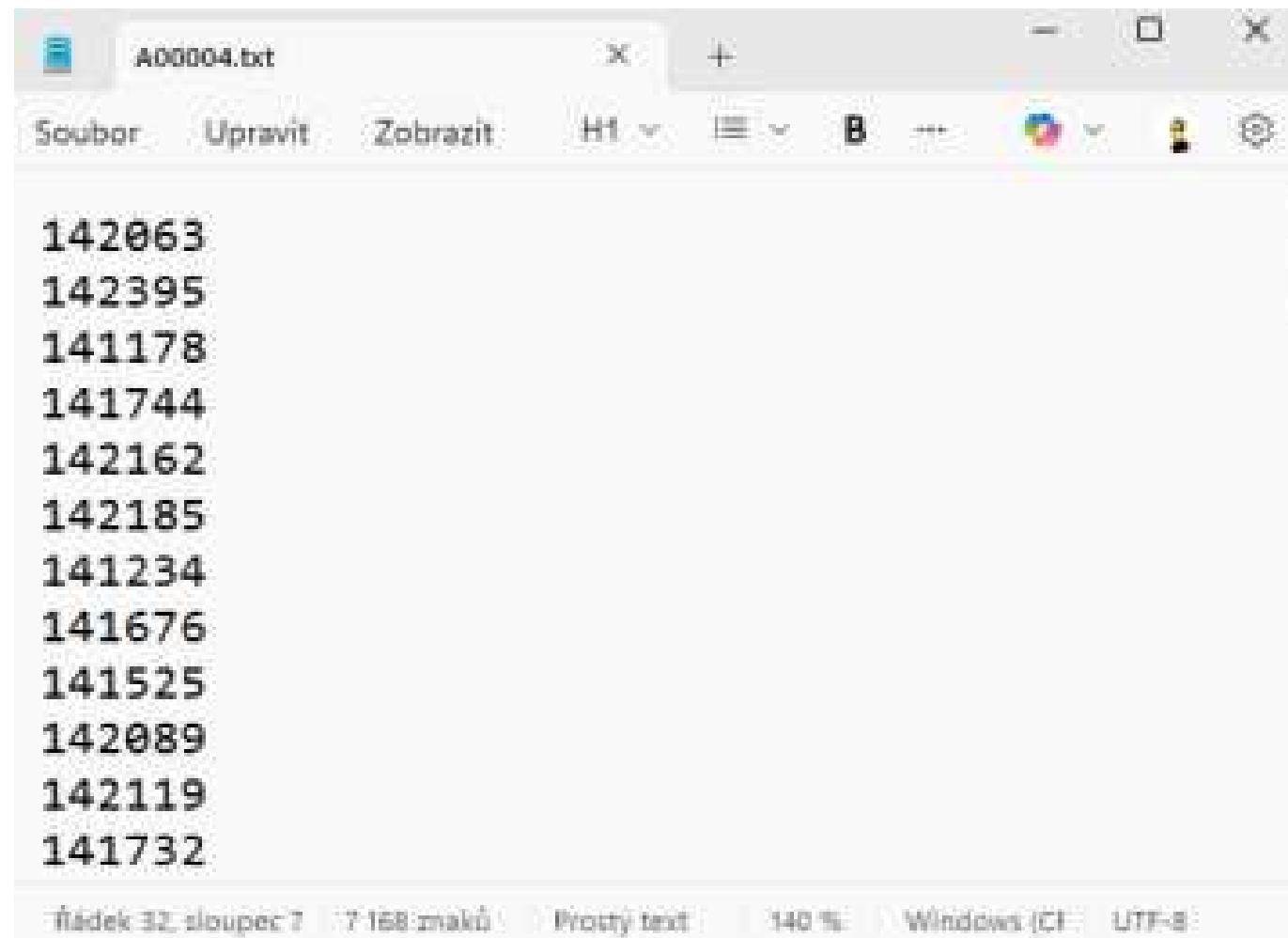
**Findable
Accessible
Interoperable
Reusable**



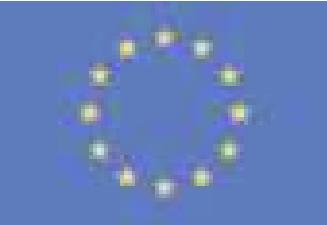
OPEN DATA



Reusable data example

A screenshot of a Microsoft Word document titled 'A00004.txt'. The document contains a list of 15 IDs, each on a new line. The IDs are: 142063, 142395, 141178, 141744, 142162, 142185, 141234, 141676, 141525, 142089, 142119, and 141732. The document has a standard Word header with tabs for 'Soubor', 'Upravit', 'Zobrazit', and 'H1'. The status bar at the bottom shows 'Rámeček 32 sloupec 7 7168 znaků' and 'Příkazy 140 % Windows (C) UTF-8'.

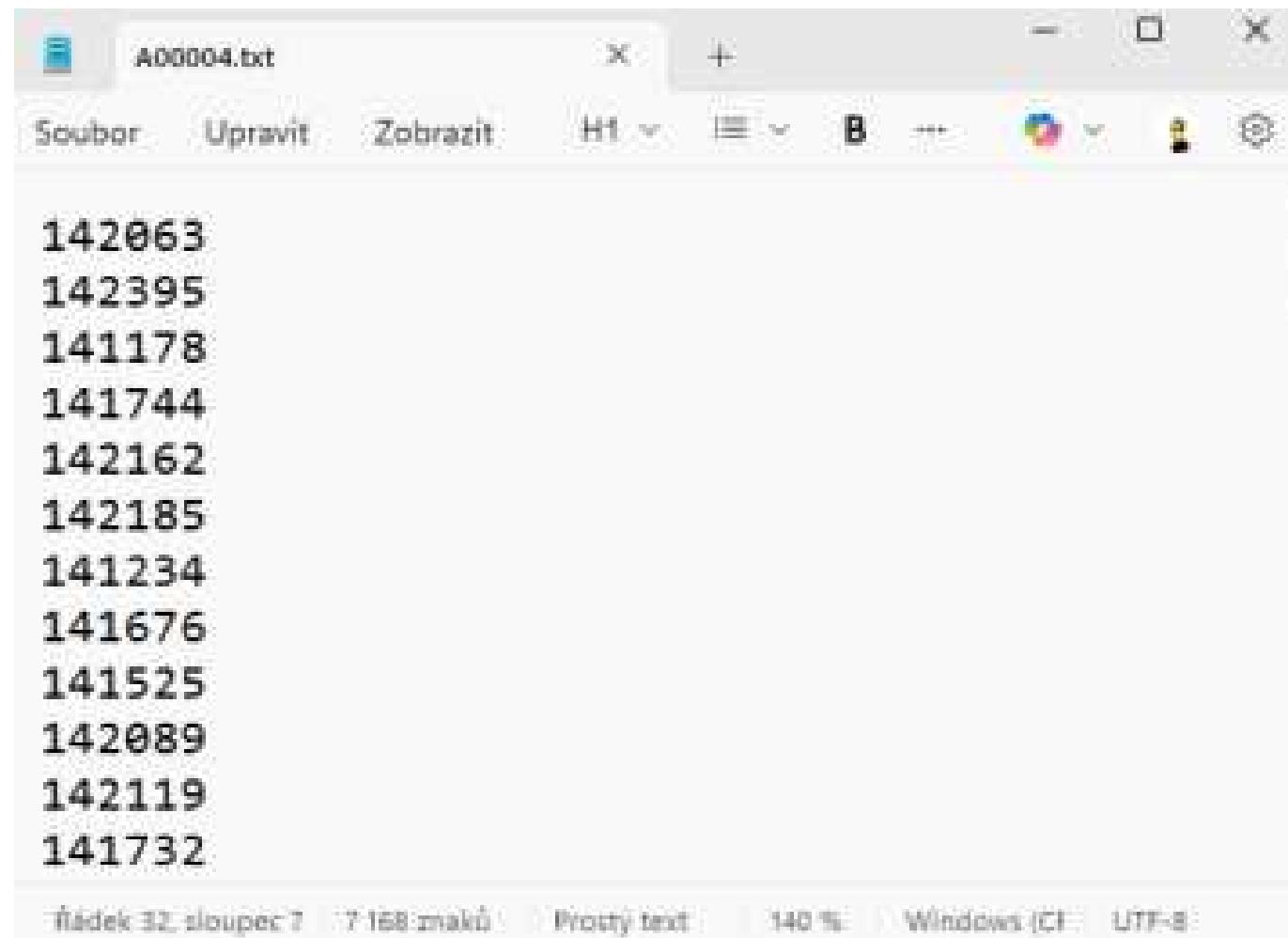
  **Findable**
Accessible
Interoperable
Reusable



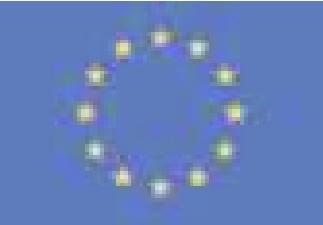
OPEN DATA



Reusable data example

A screenshot of a Microsoft Word document titled 'A00004.txt'. The document contains a single column of 14 IDs, each on a new line. The IDs are: 142063, 142395, 141178, 141744, 142162, 142185, 141234, 141676, 141525, 142089, 142119, and 141732. The document has a standard Word header with tabs for 'Soubor', 'Upravit', 'Zobrazit', and 'H1'. The status bar at the bottom shows 'Rámeček 32 sloupec 7 7168 znaků' and 'Windows (C) UTF-8'.

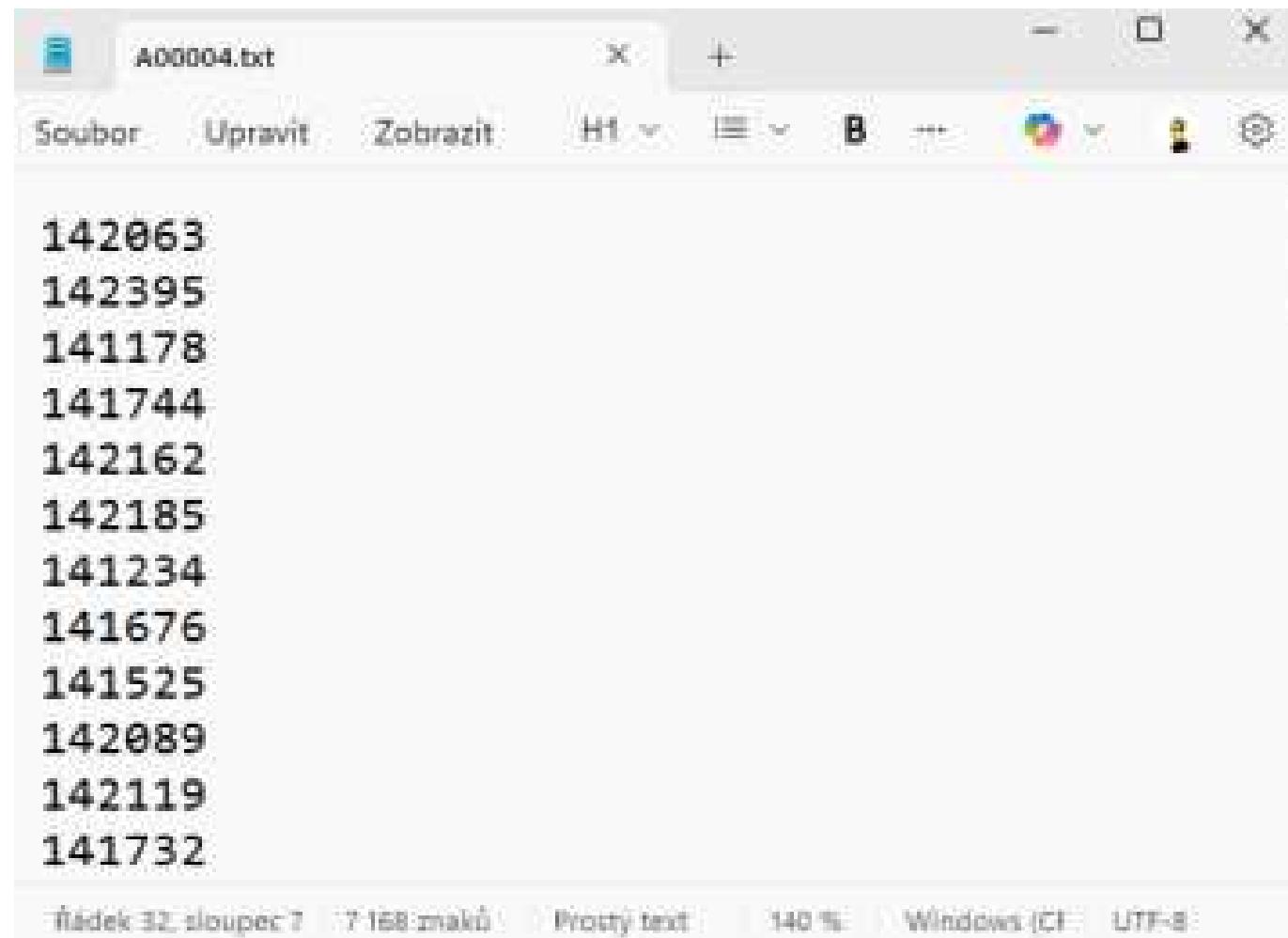
 **Findable**
 **Accessible**
 **Interoperable**
 **Reusable**



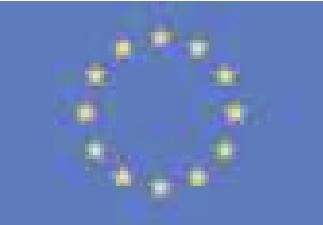
OPEN DATA



Reusable data example

A screenshot of a Microsoft Word document titled 'A00004.txt'. The document contains a single column of 14 numerical IDs, each on a new line. The IDs are: 142063, 142395, 141178, 141744, 142162, 142185, 141234, 141676, 141525, 142089, 142119, and 141732. The document has a standard Word header with tabs for 'Soubor', 'Upravit', 'Zobrazit', and 'H1'. The status bar at the bottom shows 'Rámeček 32 sloupec 7 7168 znaků' and 'Windows (C) UTF-8'.

 **Findable**
 **Accessible**
 **Interoperable**
 **Reusable**



Reusable data example

A screenshot of a Microsoft Word document window titled 'A00004.txt'. The document contains a list of 15 IDs, each on a new line. To the right of the list, the letters 'F.A.I.R' are written in large green capital letters. The IDs are: 142063, 142395, 141178, 141744, 142162, 142185, 141234, 141676, 141525, 142089, 142119, and 141732.

```
142063
142395
141178
141744
142162
142185
141234
141676
141525
142089
142119
141732
```

A vertical list of four characteristics, each preceded by a large green checkmark. The characteristics are: Findable, Accessible, Interoperable, and Reusable.

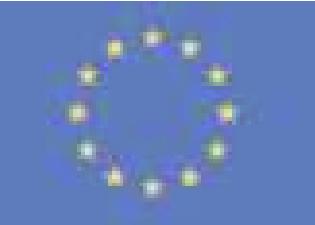
- Findable
- Accessible
- Interoperable
- Reusable

(META)DATA



Data behind data

- Data contains experimental results
- Metadata contains information about data:
 - Sample name and description
 - Measurement conditions
 - Instrument information
 - Operator information
 - Other information
- Included in data file or separated
- **Important for reproducibility!**

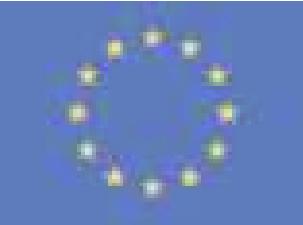
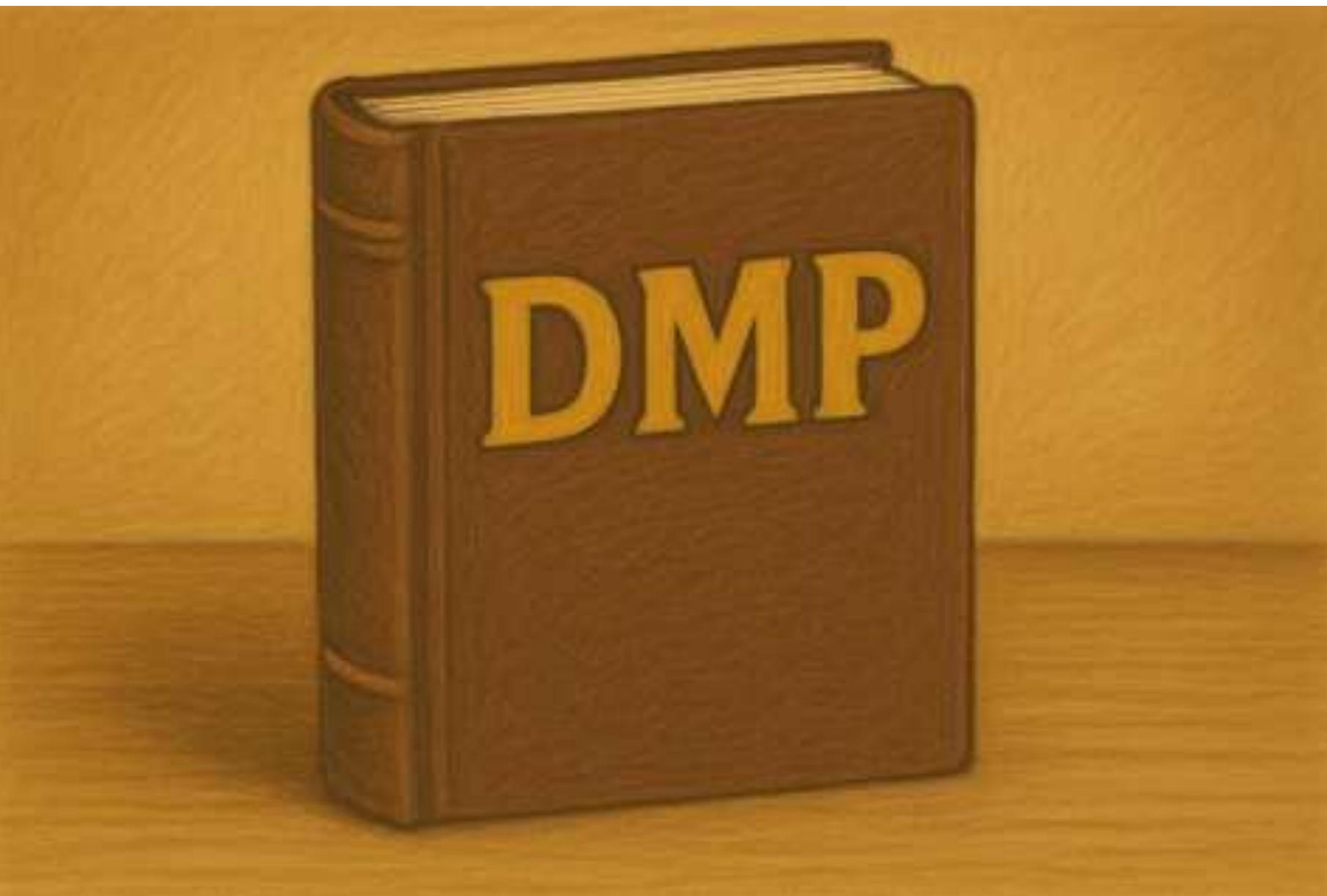


LABORATORY LEVEL



Data Management Plan

- Written set of rules and principles
- Determines how data is handled
 - Data formats
 - Data storage, sharing and backup rules
 - Metadata contents and formats
 - Responsible person
- Data loss prevention
- Easy onboarding of new employees
- Everyone knows exactly their duties

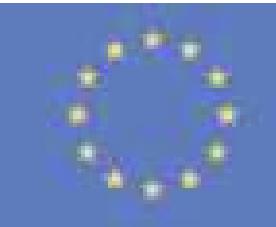


PUBLICATION LEVEL



Funding and Good practice

- **Typically defined by project funder**
- Good practice – open access cloud
 - Data are publicly available
 - F.A.I.R. principles are fulfilled
 - Free of charge service
- Requires to set rules on a team level
 - Project/Team Data management plan
 - Responsible person
- **Unacceptable to restrict access to data**
(data upon request, data upon REASONABLE request)



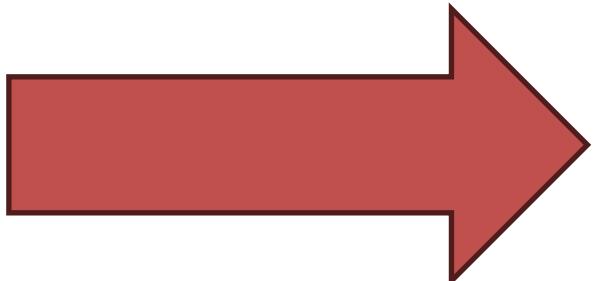
PUBLIC DATA REPOSITORY



Simple and effective

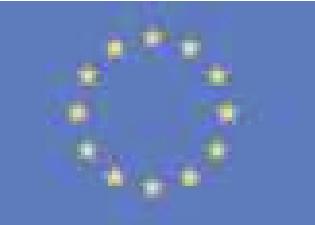
YOUR F.A.I.R Responsibility

- Interoperability
- Reusability



REPOSITORY F.A.I.R Responsibility

- Findability
- Accessibility

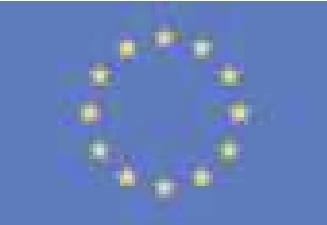


PUBLIC DATA REPOSITORY



How to do it?

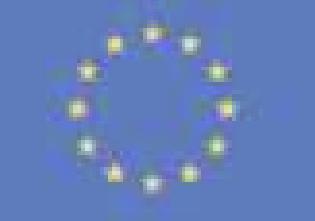
- **Register on Zenodo.org or similar**
 - Zenodo is reliable public-accessible scientific data repository running under the CERN facility
 - Registration can be done on <http://zenodo.org> (simple email registration & verification)
 - Any standard online repository which can fulfill FAIR requirements is suitable, not just Zenodo
- **Prepare all the published data**
 - The data must meet the requirements on **interoperability and reusability!**
 - Use open formats usable with any computer!
(if your instrument is using closed format, export data to open format and then upload both)
 - If you are uploading graph, do not forget to add the datasets for it, etc.!



PUBLIC DATA REPOSITORY

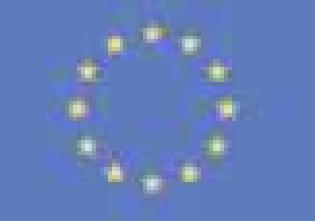


Practical example





Q&A





APPROACH

Thank you

Jakub Navářík
jakub.navarik@upol.cz

The Approach project receives funding from the European Commission's Horizon Europe Research Programme under Grant Agreement Number 101120397

The Accelerator project receives funding from the European Commission's Horizon Europe Research Programme under Grant Agreement Number 101087318

