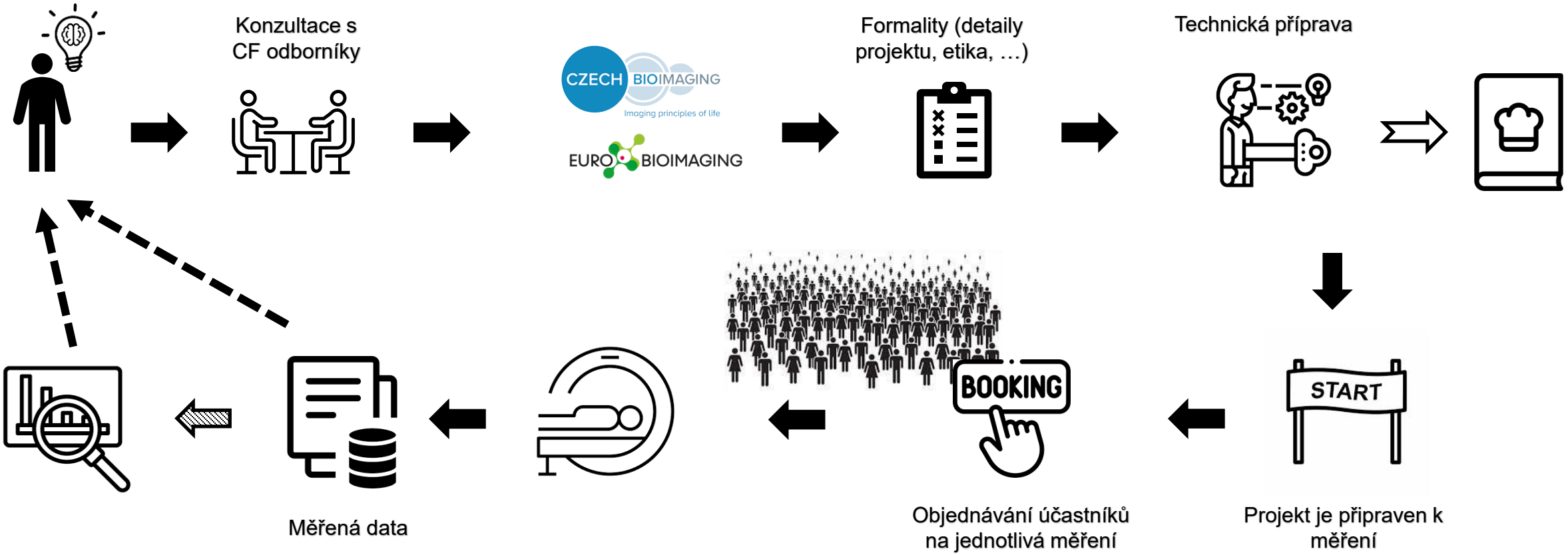
Data management pro uživatele CF MAFIL

Michal Mikl

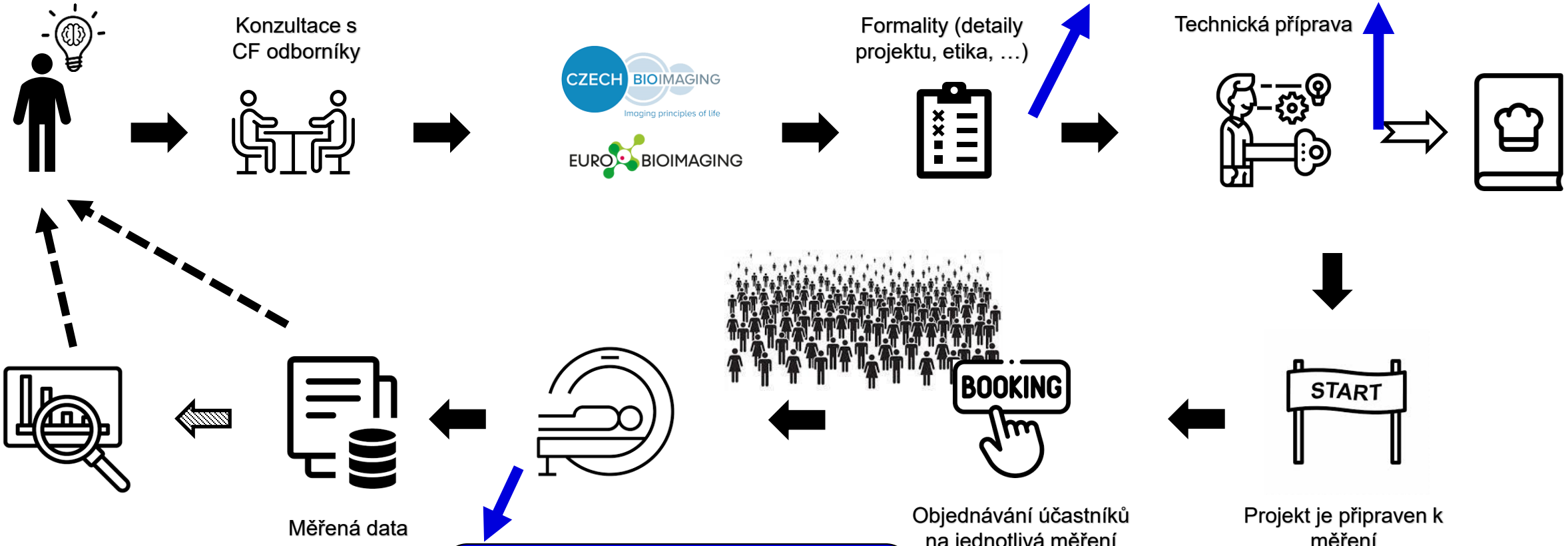
Data v MAFIL – životní cyklus projektu



MUNI | CEITEC

Data v MAFIL

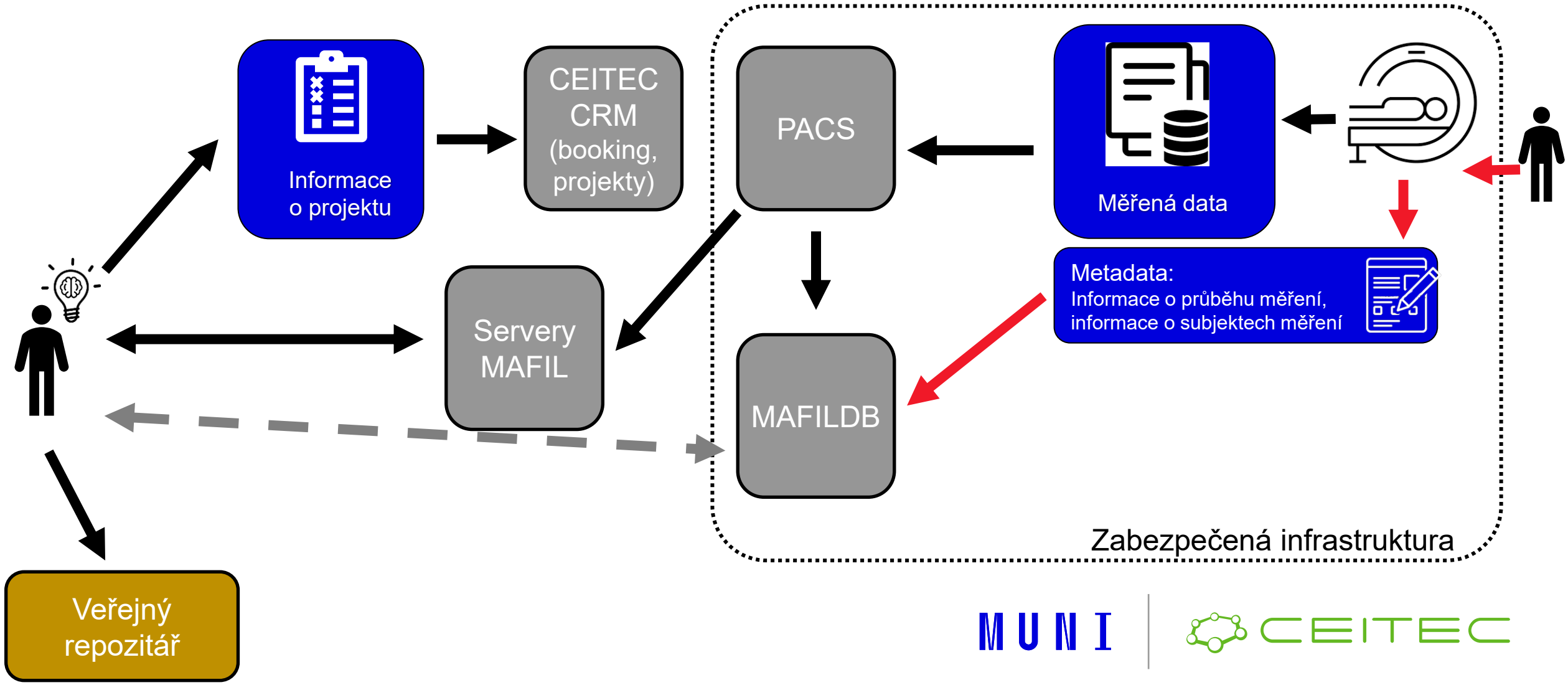
Metadata:
Řešitel, financování,
Poděkování, přístup k datům
Měřící postup, použité přístroje



Metadata:
Průběh měření
(logy, poznámy, provázanost)

MUNI | CEITEC

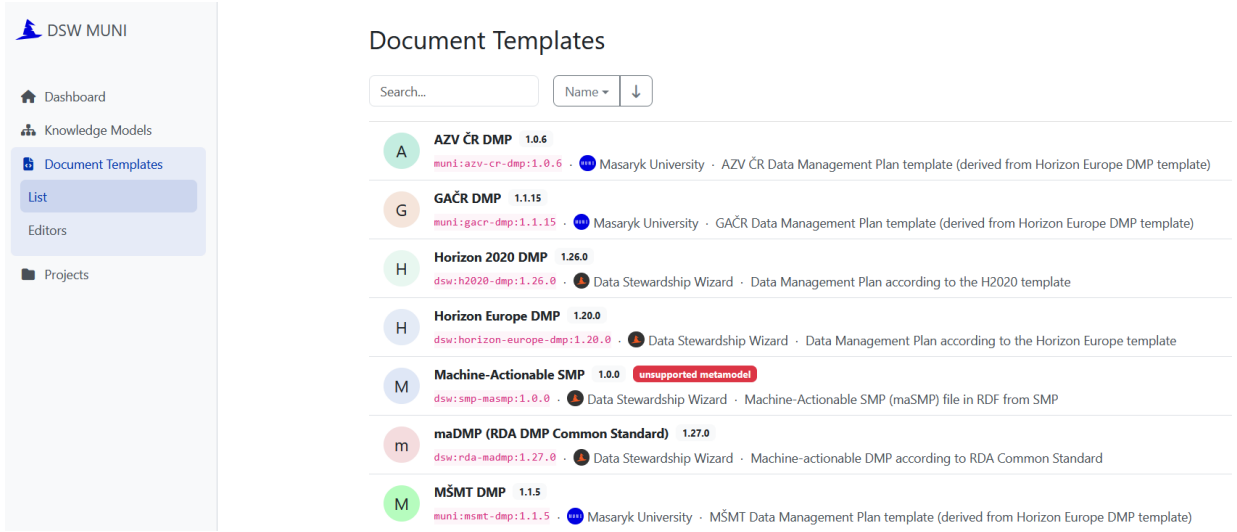
Data v MAFIL – zdroje a toky dat / metadat

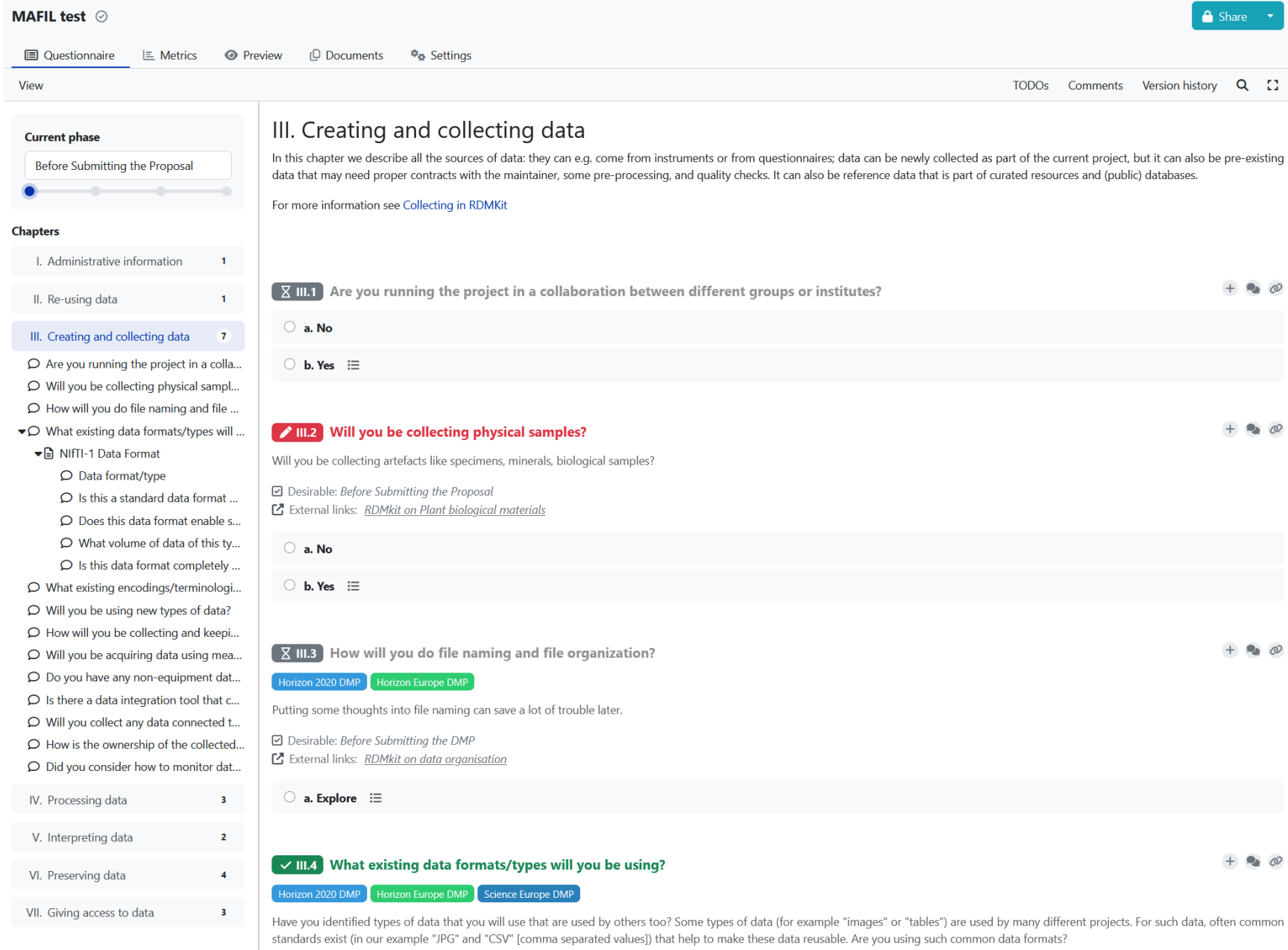


MUNI | CEITEC

Novinky v MAFIL

- Do konce roku 2026 se plánuje nový booking systém
  - Otazník nad podobou systému evidence projektů
- Zapojení do vývoje repozitáře pro medicínská obrazová a fyziologická data (MINDR-CZ)
- Na CEITEC MU se řeší příprava DMP (data management plan) šablon v nástroji DSW
  - Michaela Vaňharová (+ Michal Mikl, Tomáš Slavíček, Jan Fousek, …)




MAFIL test
Share
Questionnaire
Metrics
Preview
Documents
Settings
View
TODOs
Comments
Version history
Current phase
Before Submitting the Proposal
Chapters
I. Administrative information 1
II. Re-using data 1
III. Creating and collecting data 7
Are you running the project in a colla...
Will you be collecting physical sampl...
How will you do file naming and file ...
What existing data formats/types will ...
NIfTI-1 Data Format
Data format/type
Is this a standard data format ...
Does this data format enable s...
What volume of data of this ty...
Is this data format completely ...
What existing encodings/terminologi...
Will you be using new types of data?
How will you be collecting and keepi...
Will you be acquiring data using mea...
Do you have any non-equipment dat...
Is there a data integration tool that c...
Will you collect any data connected t...
How is the ownership of the collected...
Did you consider how to monitor dat...
IV. Processing data 3
V. Interpreting data 2
VI. Preserving data 4
VII. Giving access to data 3
III. Creating and collecting data
In this chapter we describe all the sources of data: they can e.g. come from instruments or from questionnaires; data can be newly collected as part of the current project, but it can also be pre-existing data that may need proper contracts with the maintainer, some pre-processing, and quality checks. It can also be reference data that is part of curated resources and (public) databases.
For more information see Collecting in RDMKit
III.1 Are you running the project in a collaboration between different groups or institutes?
a. No
b. Yes
III.2 Will you be collecting physical samples?
Will you be collecting artefacts like specimens, minerals, biological samples?
Desirable: Before Submitting the Proposal
External links: RDMkit on Plant biological materials
a. No
b. Yes
III.3 How will you do file naming and file organization?
Horizon 2020 DMP
Horizon Europe DMP
Putting some thoughts into file naming can save a lot of trouble later.
Desirable: Before Submitting the DMP
External links: RDMkit on data organisation
a. Explore
III.4 What existing data formats/types will you be using?
Horizon 2020 DMP
Horizon Europe DMP
Science Europe DMP
Have you identified types of data that you will use that are used by others too? Some types of data (for example "images" or "tables") are used by many different projects. For such data, often common standards exist (in our example "JPG" and "CSV" [comma separated values]) that help to make these data reusable. Are you using such common data formats?

Podpora FAIR / open data na MAFIL

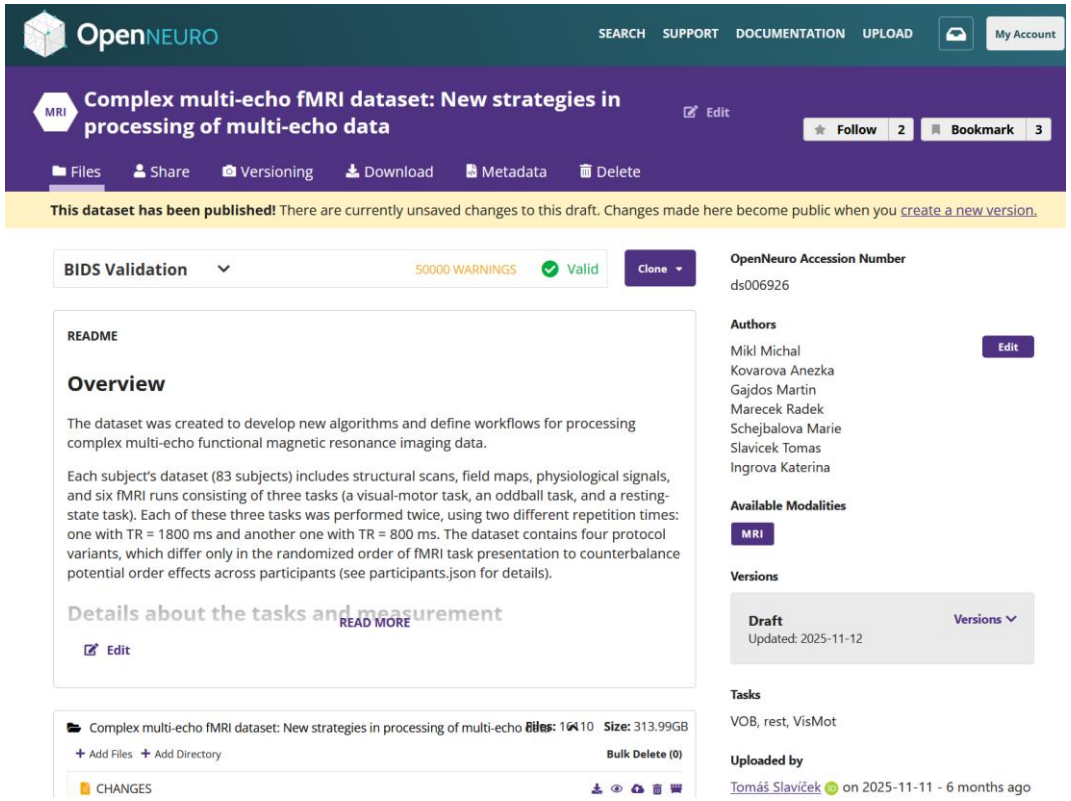
- Michaela Vaňharová – data steward pro primární kontakt s uživateli v obecnější rovině
- Tomáš Slavíček – data curator pro řešení specifických situací v datasetech, techničtější rovina
- Martin Kojan, Kateřina Ingrová – specifická podpora práce s daty

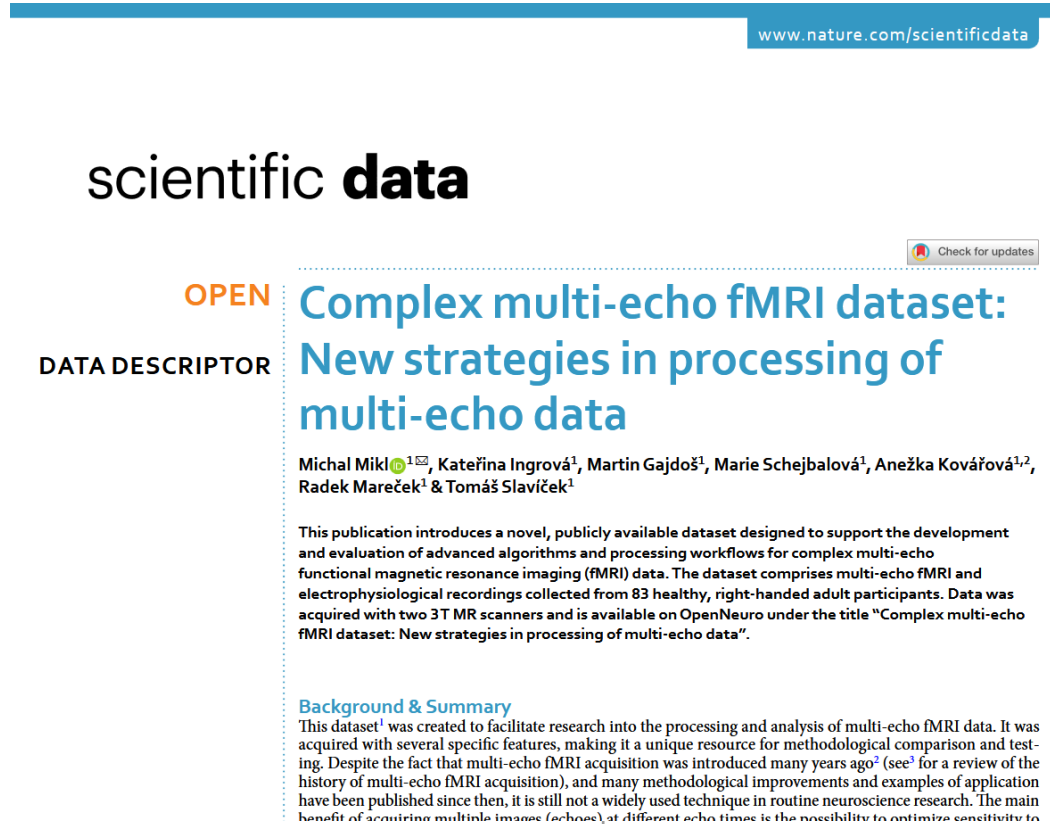
Anonymizace dat v MAFIL

- Na přelomu 2025/2026 hotová první technická implementace anonymizace v MAFILDB a PACS a začátkem 2026 otestování na vybraných projektech z roku 2015
- Provedena kontrola MR projektů z let 2015 a 2016 za MAFIL, nyní budeme oslovovat jednotlivé řešitele/kontaktní osoby s možností jejich kontroly / doplnění informací před propvedením anonymizace

Další zkušenosti s publikací dat

– Dataset z projektu ELICIT + článek v časopise Scientific Data





Zkušenosti z recenzního řízení

- Postřehy k úpravě datasetu (pojmenování v BIDS)
- Doplnění informací o datech z jednotlivých ech
- Zveřejnění kódu (dkriptů) pro kontrolu kvality a preprocesing

1. When complex-valued data are present, the magnitude images should have 'part-mag' in the filenames. The omission of 'part-mag` from filenames is only valid if magnitude data are the only data available for that acquisition.

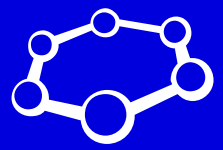
Thank you for this suggestion. You are right, that 'part-mag' in the filenames is correct and more intuitive. We changed the filenames accordingly.

2. I believe the T2starw suffix is not the correct choice for your anatomical images. I think the MEGRE suffix is what is recommended for this kind of acquisition.

Thank you for the comment. We changed the suffix to MEGRE.

4. It would make it easier to work with the data if the acquisition entity was used. For example, the following labels could be useful:
- task-rest_acq-tr800 and task-rest_acq-tr1800 instead of task-rest1 and task-rest2.
- task-rest1_acq-scanner1 and task-rest1_acq-scanner2 instead of placing the information about the scanner in the participants.tsv file. I think the differences in parameters between scanners merits a distinct acq entity.
- task-rest_acq-tr800+scanner1, task-rest_acq-tr1800+scanner1, task-rest_acq-tr800+scanner2, and task-rest_acq-tr1800+scanner2.

Thank you for the suggestion of use of the acquisition entity. This can help to distinguish between the tasks and repetition times. Therefore, we implemented it to the filenames structure. Regarding the scanner version, we checked the BIDS specification (https://bids-specification.readthedocs.io/en/stable/longitudinal-and-multi-site-studies.html#option-2a-collate-sites-at-subject-level) and selected different approach (option 2a at this webpage) which we consider clearer. We incorporated scanner type (A/B) into the subject names. And the scanner type is in the participants.tsv table too.

MUNI |  CEITEC