Contribution ID: **95** Type: **POSTER&PITCH**

How to make Biomedical Imaging Datasets AI-ready?

Monday 4 November 2024 15:00 (1 hour)

The vast amount of observations needed to train new generation AI models (Foundation Models) necessitates a strategy of combining data from multiple repositories in a semi-automatic way to minimize human involvement. However, many public data sources present challenges such as inhomogeneity, lack of machine-actionable data, and manual access barriers. These issues can be mitigated through the consequent adherence to the FAIR (Findable, Accessible, Interoperable, Reusable) data principles, as well as state-of-the-art data standards and tools. In the poster, we highlight the inhomogeneity of the schema definitions in the field, provide helpful tips on what could improve the AI-readiness of data and inspect example data sources which implement the most novel concepts in working with data and metadata in the machine-actionable fashion.

Please specify ”other”

In addition, please add 3 to 5 keywords.

Artificial Intelligence, Fair Data Point, Bioimaging, Data harmonization

Please specify ”other”

For whom will your contribution be of most interest?

Data professionals who provide and maintain data infrastructure

Please assign yourself (presenting author) to one of the following groups.

Researchers

Primary author: DVORETSKII, Stefan (HMC Hub Health, DKFZ)

Co-authors: Mr MOORE, Josh (German BioImaging e.V.); KULLA, Lucas (DKFZ); NOLDEN, Marco (DKFZ); Mr SCHADER, Philipp (HMC Hub Health, DKFZ)

Presenter: DVORETSKII, Stefan (HMC Hub Health, DKFZ)

Session Classification: Poster Session B

Track Classification: Connecting research data: 4. Metadata annotation and management