

Towards metadata descriptors to support interoperability and data reuse in circadian and sleep science

Tuesday 10 October 2023 10:50 (20 minutes)

ABSTRACT

Introduction

Meaningful metadata are essential for the description, retrieval and reuse of data, in particular in multi-centric cooperative research projects. For the Integrative Human Circadian Daylight Platform (iHCDP, <https://ihcdp.org/>), a collaborative, transnational project between the University of Basel (Switzerland), the Technical University of Munich and the Max Planck Institute for Biological Cybernetics in Tübingen (Germany), we are developing the Circadian Data Hub (CDH). The CDH enables data upload and exchange to support interactivity and reuse of study data within each iHCDP module and across the platform. Expressive metadata are needed to describe the data stored in the CDH. The aim was to identify the expressive metadata needed to describe the data stored in the CDH.

Methods

The CDH is designed according to the FAIR principles formulated by Wilkinson et al. 1 to make data findable, accessible, interoperable and reusable. In several workshop meetings we have harmonized aspects of the data collection effort across institutions.

To create a catalog of metadata, we first integrated information from all data which are currently being collected within the iHCDP project teams. For this purpose, a spreadsheet for data entry was completed, where study modality, variable names, units, devices, sampling methods and frequencies were entered. We then clustered the same and similar information together. From these metadata clusters an initial set of variables for the CDH metadata descriptor was created. Next, we developed a pilot data collection tool in the openBIS system [2]. When uploading data to the CDH, this data collection tool collects information about the data collected in a particular project.

Results

We created a pilot project in the openBIS system. To build the project structure and to consider the specific data sets and variables in circadian and sleep studies we created objects for projects and for participants. For each of these objects we implemented a metadata descriptor with general and project specific metadata. General metadata for projects include: project ID, project title and total number of participants. Participant metadata are participant ID, associated project ID, age, sex and gender. As project-specific metadata, we implemented the variables we created from the metadata clusters, as Boolean variables in the metadata descriptor for projects. We grouped the collected information according to how the data was collected and into metadata clusters. For example, one cluster is: data collected by study personnel - physical/vital signs. In this cluster, we have the following variables: core body temperature, blood pressure, heart rate, weight and height.

Discussion

With the implemented metadata clusters and the variables, projects in the CDH can be described and other researchers could reuse the data in the future. The metadata descriptor can be revised at later times to respond to the addition of novel data collection modalities.

Conclusion

We developed a novel metadata descriptor, which captures the data collected within a specific project for use within the CDH.

Funding

VELUX Stiftung, Switzerland (Proj. No. 1636)

References

[References are in the Comments field.]

Please assign your contribution to one of the following topics

Data interoperability through harmonised metadata and interoperable semantics

Please specify "other" (stakeholder)

In addition please add keywords.

Metadata, FAIR principles, Harmonization, Sleep science, Circadian studies

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

Researchers

Primary author: Mrs LINDOERFER, Doris (Chronobiology & Health, TUM Department of Sport and Health Sciences (TUM SG), Technical University of Munich, Munich, Germany)

Co-authors: Dr GARBAZZA, Corrado (Centre for Chronobiology, Psychiatric Hospital of the University of Basel, Basel, Switzerland); Prof. CAJOCHEN, Christian (Centre for Chronobiology, Psychiatric Hospital of the University of Basel, Basel, Switzerland); Dr MÜNCH, Mirjam (Centre for Chronobiology, Psychiatric Hospital of the University of Basel, Basel, Switzerland); Prof. SPITSCHAN, Manuel (Chronobiology & Health, TUM Department of Sport and Health Sciences (TUM SG), Technical University of Munich, Munich, Germany)

Presenter: Mrs LINDOERFER, Doris (Chronobiology & Health, TUM Department of Sport and Health Sciences (TUM SG), Technical University of Munich, Munich, Germany)

Session Classification: Parallel Track 2

Track Classification: Facilitating connectivity of research data: Data interoperability through harmonised metadata and interoperable semantics