

# Considerations for developing a standard for storing electrophysiology data in HDF5



MOUSEK, R. *Considerations for developing a standard for storing electrophysiology data in HDF5*. Stockholm, 2013.

---

The INCF Program on Standards for Data Sharing has a working group that is developing a standard for storing electrophysiology data in HDF5. The most important requirement of such a standard is to accommodate the common types of data used in electrophysiology and also the metadata required to describe them. Neuroshare defines four data types: analog signals, segments, neural events and experimental events; as well as some metadata. A standard needs to efficiently store these data types, and probably also imaging data and some kinds of data generated in the data processing chain. Further, a standard way of storing the metadata must be specified. The set of metadata required to describe electrophysiology data is difficult to determine a priori because the types of experiments are so varied. So, a flexible mechanism must be used which allows referencing and specifying values for currently existing ontologies and also accommodates information not currently systematized. Techniques to include post-experiment annotations of data, and for relating different data parts, are also required. So far, the working group entertains two approaches towards defining a standard, which may eventually be merged. One, currently named Pandora, defines a generic data model that can be used with HDF5 or other storage back-ends. Due to the generic nature, the data model can be used to store various kinds of neuroscience data. The other proposal, called epHDF, defines domain specific schemata for storing electrophysiology data in HDF5. For any approach, a suite of test data sets to help evaluate a proposed standard is needed, and tools to allow validating data files are desirable.

27.08.2013

Stockholm