SensoMatrix

A NEUROIMAGING-INSPIRED DATA MANAGEMENT SYSTEM FOR BODY-SENSING

Atousa Assadi[1], Niroigen Suntharam[1], Christine Rogers[2], Najmeh Khalili-Mahani[1,2]

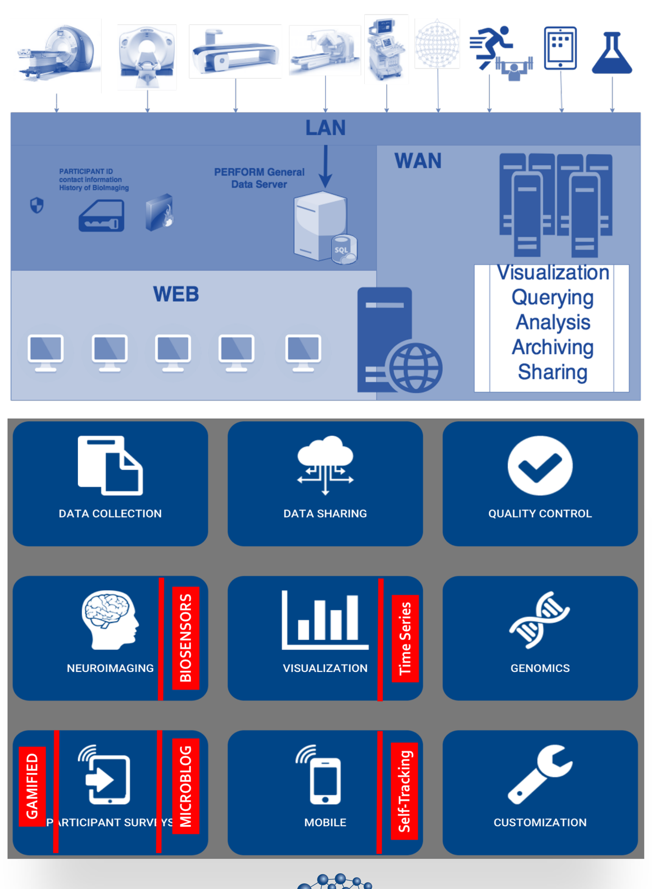
1. PERFORM Centre and Department of Electrical Engineering, Concordia University, Montreal, QC, Canada;
2. McGill Centre for Integrative Neuroscience, McGill University, Montreal, QC, Canada

INTRODUCTION: Rapid evolution of pervasive computing and body-sensing technologies that propel the Internet of Health, provide novel opportunities for clinical and epidemiologist researchers to gather ecological data from real-life experiences of patients [1]. Yet, a lack of open-source data management tools limits researchers in their ability to integrate data from various manufacturers of different biosensors. Also, the proprietary nature of many software development kits (SDKs) significantly reduces the scale, equity and inclusivity of the data collection process--keeping such systems available to few major industry players or wealthier academic institutes. We propose a data management systems to integrate heterogeneous data from various body worn sensors.

OBJECTIVE: For more than 30 years, large-scale neuroimaging studies have relied on open-source software for studying the life-span predictors of healthy aging or neurodevelopment. Population-based multi-centre bioimaging studies have taken advantage of data management ecosystems such as LORIS [2] (Fig 2A) that provides a web-based interface to help researchers collaborate globally as they rely on this software to:

- Receive data through secure networks in an ongoing basis;
- Keep a harmonized database from multiple sources/centres;
- Integrate, visualize and quality-control data online;
- Store derivative data including all analytical metadata; and
- Analyse data in standardized pipelines, with provenance tracking.

SensoMatrix leverages the existing modules in LORIS, to incorporate data from various body sensors for population studies with various objectives (Fig 2B). The wealth of knowledge and experience in the field of neuroimaging, can help create the blue-print of a community-driven approach to integration of data from various biosensors into clinical research.



(A)

(B)

Figure 1: (A) Longitudinal Online Research and Imaging System (LORIS), is an OPEN SOURCE web-based data and project management software for neuroimaging research studies, and facilitates storing and processing behavioural, clinical, neuroimaging and genetic data. LORIS makes it easy to manage large datasets acquired over time in at different locations in a large multi-site study. (B) SensoMatrix builds on LORIS, and focuses on data obtained using ecological momentary assessment. Specifically, it addresses the need for incorporating data collected from wearable biosensors, as well as mobile applications (such as games) into LORIS, by adapting existing data-formats, and visualization tools for electrophysiological data (EEG).

METHODS FROM NEUROIMAGING: Clinical neuroimaging research thrives on 3 principles: 1) an agreed-upon statistical template of "normal" which is developed by averaging hundreds of data acquired under similar conditions; 2)Analytical pipelines that keep track of all signal processing steps that model every single dataset with respect to that template; 3) an agreed-upon data format (digital imaging and communication in medicine-DICOM), and automatic conversion scripts to bundle hundreds of DICOM files into a single multi-dimensional file. In designing the SensoMatrix data format, we are inspired by MINC2.0 [3], and Brain Imaging Data Structure (BIDS) [4].

MINC2.0

- An extensible header format that includes both data acquisition and analysis history.
- Self-documenting metadata using human-readable, descriptive variable and attribute names.
- Support for high dimensionality and arbitrary coordinate systems, including irregularly sampled dimensions.
- Straightforward support for new modalities.
- Data portability and platform independence.

EEG-BIDS

- An extension of BIDS, used to keep pre- and post-processed neuroimaging data in a standard directory structure.
- BIDS-EEG [5] places data from various electrodes with .edf format, into a folder that also contains the meta-data (.json), timing of events (.tsv), electrode names (.tsv), and electrode coordinates (obtained from MRIs, .json)

EXAMPLE APPLICATION: An example application elucidates user requirements. We investigated the impact of computer games playing on stress response in seniors [6]. A multi-session experiment (Fig 2A) was conducted to evaluate physiological, and behavioral response of seniors to cognitive games. Data was collected from a single device with different sensors (E4). In the future, we will add Kinect and portable EEG to the protocol, each with multiple channels, recording data at different sampling rates per sensor type and providing raw or preprocessed data in different file formats. We require to organize data from each session in a single dataset with a meta-data description to keep track of information shown in (Fig 2B).

(A) **Epochs**

15-20 min	5 min	15 min	5 min	15 min	5 min	5 min
Prepare	Encode	Game1-A	Stroop	Game1-B	recall	end
Prepare	Encode	Game2-A	Stroop	Game2-B	recall	end
Prepare	Encode	Game3-A	Stroop	Game3-B	recall	end

Multi-sensors: Movement, Pulse, Temperature, Skin conductance

(B)

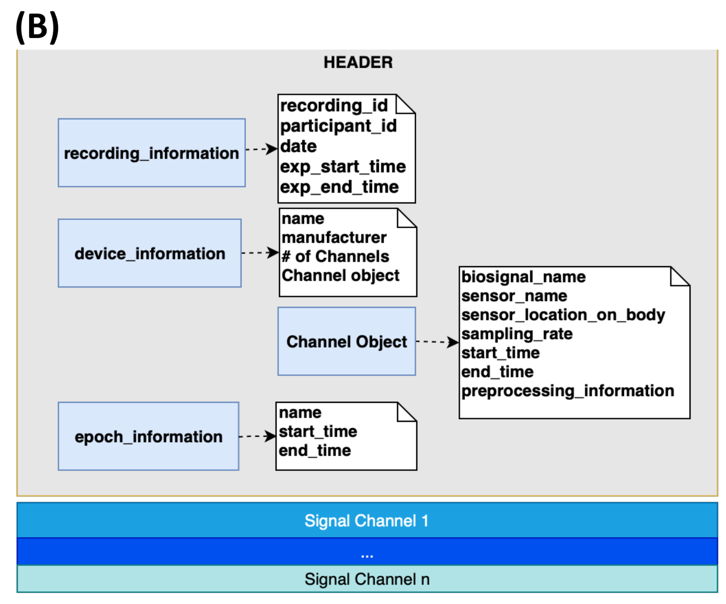


Figure 2: (A) Example of a longitudinal multi-sensor, multi-epoque experiment; (B) Exemplar meta-data in the header of an object oriented data structure notation.

SensoMatrix FRAMEWORK DESIGN SPECIFICATIONS include the following:

1. Data collection and annotation is a human task => an **ACTOR** must initiate data conversion, and prepare customized data entry and data conversion procedures.
2. Quality control depends on ACTOR's ability to view the raw and processed signals => a **VISUALIZATION and TAGGING** system is necessary to enable ACTORS to ensure the quality or flag errors in the stored signal
3. Reproducibility and continuity of longitudinal studies requires streamlined processing => a **STANDARD DATA FROMAT** to codify complex data into comprehensive representation, with sufficient meta-data that describes the details of data acquisition and pre- and post-processing.
4. Automation reduces human error => **PIPELINES** to facilitate supervised batch data-entry and analysis are desirable.

Based on these specifications we have designed and prototyped SensoMatrix Framework (Fig 3) to include a web-based interface to allow ACTORS to take data from various sources, and create device- and study-specific meta-data templates that will be used to convert biosensor data into JSON files organized similarly to a MINC file. Once the template is available for one sample, it can be used to batch-import all data from that particular device into a visualization tool, that can be used for aligning the data temporally, tagging segments of data that are of interest (or are to be excluded), and performing simple statistical analyses on data from one or multiple channels.

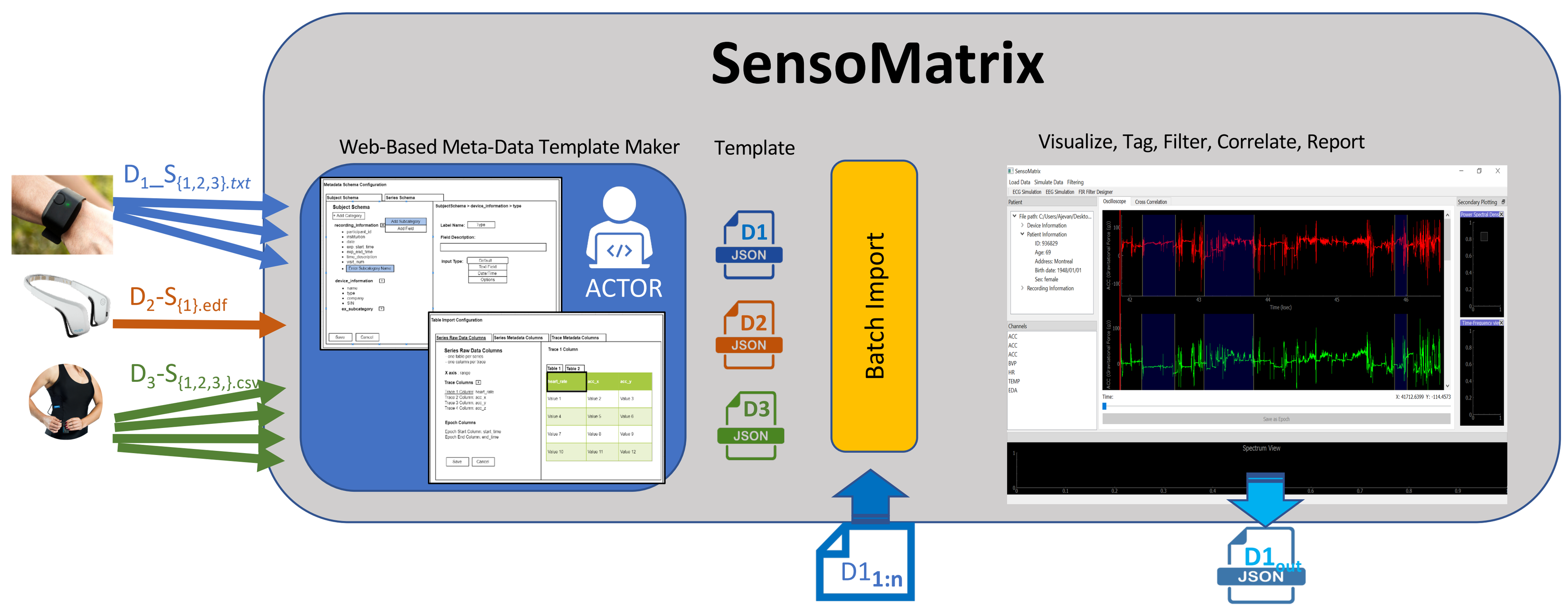


Figure 3: A schematic representation of SensoMatrix. SensoMatrix extends the BIDS-EEG implementation in LORIS, by providing a user interface to create device-specific header files (Fig 2B), and by providing a visualization tool that enables users to synchronize data from different sensors, with different sampling rates, and data formats. A prototype of SensoMatrix is currently available in https://github.com/sensomatrix/sensocore.

FUTURE DIRECTIONS The aim of this poster is to draw attention to a scientific need to follow the lead of communities such as OPEN BCI and OPEN NEURO in order to advance the application of biosensors in population health studies. Many wearable technologies exist but they are often tied to a data-service provided by their manufacturer. The first step in making these technologies accessible and inclusive is to standardize data representation, visualization and sharing.

REFERENCES: **[1]** Insel TR. Digital Phenotyping: Technology for a New Science of Behavior. ***JAMA*** 2017; **[2]** Das, S, et al LORIS: a web-based data management system for multi-center studies. ***Frontiers in Neuroinformatics*** 2012; **[3]** Vincent, R, MINC 2.0: A Flexible Format for Multi-Modal Images, ***Frontiers in Neuroinformatics*** 2016; **[4]** Gorgolewski, K et al, The brain imaging data structure, a format for organizing and describing outputs of neuroimaging experiments, ***Scientific Data*** 2016; **[5]** Pernet CR et al, EEG-BIDS, an extension to the brain imaging data structure for electroencephalography, ***Scientific Data*** 2016; **[6]** Khalili-Mahani N et al, Reflective and Reflexive Stress Responses of Older Adults to Three Gaming Experiences In Relation to Their Cognitive Abilities: Mixed Methods Crossover Study, ***JMIR*** 2020; **[4]**;

Funding for this work is provided by Concordia University.

MEDIA HEALTH
Concordia University PERFORM Centre
MCIN MCGILL CENTRE for INTEGRATIVE NEUROSCIENCE