

## Aiyetan\_2019 ML-MSM Meeting - Abstract Submission Form

**Title: Towards A Universal Multiscale Data Representation and Knowledge Integration**

**PI(s) Grant: N/A**

**Institution(s):**

1. Frederick National Laboratory for Cancer Research, Frederick MD 21701
2. School of Systems Biology, George Mason University, Manassas VA 20110
3. Intelligent Systems Engineering, Indiana University, Bloomington, IN 47405
4. Opto-Knowledge Systems Inc., Torrance, CA 90502
5. Fluidigm Corporation, San Francisco, CA 94080 (Current Affiliation)

**Grant Number: N/A**

**Title of Grant: N/A**

**Abstract Authors:**

Paul Aiyetan<sup>1,2, \*</sup>

Kedar Narayan<sup>1</sup>

Randy Heiland<sup>3</sup>

Samuel Friedman<sup>4</sup>

Corinne Zeitler<sup>1</sup>

Paul MacKlin<sup>3</sup>

Andrew Quong<sup>1,5</sup>

\*Corresponding Author

Paul Aiyetan, [paul.aiyetan@nih.gov](mailto:paul.aiyetan@nih.gov)

### **Abstract Text**

The biomedical and translational research landscape is a superhighway of information exchange, and data is the currency of exchange. But beyond the modes or forms of transmission is the data model. A data model is an abstract construct that organizes elements of data. It standardizes how they relate to one another and to the properties of real-world entities. Many data models do exist – to some extent addressing a specific need of a specialized agent or human inquiry and oftentimes described within the purview of a specific field of knowledge. More often than not, these describe events in single time scales.

However, for a holistic view of biological processes, a description of events at multiple time scales and space is a given. While addressing complexity, the best of efforts at describing biological processes has for so long been boxed into specific and almost independent research niches – birthing siloed situations of discovery endeavors. It is here posited with some demonstrable evidence, that a universal framework provides a basis for integration of diverse arrays of data that constitute the different time scales that may explain a biological phenomenon.

We here illustrate three sample disparate biological datasets, described within a common data model and framework implemented at the Frederick National Laboratory for Cancer Research (FNLCR). Datasets include those derived from digital cell line assays, transcriptome sequencing

assays, proteome sequencing assays, light microscopy, and focused ion-beam scanning electron-microscopy (FIB-SEM). This paradigm not only simplify a somewhat complex picturesque but is proposed to facilitate advances in multiscale model development – a more holistic inquisition into biological processes.