Accelerating Translation Medicine: Providing "Out-Of-The-Box" Machine-Readability

bringing Pharmaceuticals, Implanted Medical Devices, and
Cancer Immuno-therapies to Market Sooner and more Cost-Effectively

Linguistic Technology Systems (LTS) has a team of software developers that have designed a unique data integration platform, which is part of our new cloud-computing framework, to enable heterogeneous biomedical data (e.g. genomic/proteomic/transcriptomic, cellular, image biomarkers, and clinical) to be rendered machine readable as an out-of-the-box solution so as to optimize machine-learning algorithms. In so doing, we are addressing one of the most significant obstacles in machine learning for computational therapeutics: the inability to *efficiently merge data from different sources into one common data space* in order to render the data machine readable.

What we offer is a crucial precursor to machine learning by enabling variegated data structures to be subsumed under one Common Data Model. In this way, machine learning would be optimized because algorithms working within one single data space can then operate on a wide spectrum of parameters derived from many different datatypes, something which has not been done heretofore.

Our data integration platform that uniquely offers "out-of-the box" machine readability represents a significant improvement over the traditional *ad hoc* (and unduly cumbersome) way of setting up data parameters for machine learning. That is, unlike the traditional approach which requires for each group of parameters their own data acquisition and computational protocol to be set up, our approach provides data models insuring that every group of parameters needed for a given machine learning algorithm has a *unified* computational framework, thereby eliminating the need to set up parameters on an *ad-hoc* basis — which is both time consuming and costly. In addition, having an integrated protocol reduces the risk of important data points (e.g. a region of interest on a bioimage such as confocal microscopy) being *excluded* from simulations and/or machine learning, which often occurs when a data integration protocol for testing/modeling/simulation is rendered incomplete because of extant difficulties in integrating heterogeneous data types.

Such bottlenecks in computational therapeutics, caused by problems with preparing variegated data types to make it machine readable, unfairly delay (and increase the costs of) the bringing of pharmaceuticals, immunotherapies and implanted medical devices to market. Our software engineering technology is designed to address this critical problem in computational therapeutics so that the robustness in the development of life-saving therapeutics is matched by an equally robust data integration platform so as to optimize machine learning.

In addition to biomedical use-cases, our data-integration/database engineering protocols and new cloud infrastructure have applications in Scientific Publishing, by integrating data sets with publications so that the data is findable, accessible, interoperable, and reusable.