1106-62-21 **Susan Holmes*** (susan@stat.stanford.edu), Statistics Department, Sequoia Hall, 390 Serra Mall, Stanford, CA 94305. *Statistically relevant metrics for heterogeneous data.*

Finding the right distance or dissimilarity solves difficult statistical problems. This talk will provide a survey of mining heterogeneous biological data including networks, trees, images and heteroscedastic variables using weighted dissimilarities and locally defined distances. Carefully tailored "distances" can incorporate prior information on data structure such as hierarchical dependencies between rows of a data matrix or the graph of correlations between the column-variables. Links to differential geometry are useful in incorporating localized information for these complex data structures. Distances are central to the statistical endeavor and enable generalizations of the notions of variance decomposition, nearest neighbor classification and clustering. I will show examples of how these generalizations prove useful in integration of multiple sources of information in the context of the study of the human microbiome. (Received September 15, 2014)