

## **Report 5: Environmental/Geospatial informatics.**

**Convener:** Marie Lynn Miranda

### **Brief History:**

- New technologies are generating terabytes of data, as result of substantial investments by NIH and other research agencies, as well as private sector entities
- Underinvestment in methods for Data architecture and statistical and mathematical modeling – therefore unable to fully organize, architect, and analyze these data
- Abundant data from non-traditional sources such as industry, transport sector, marketing, insurance etc. is relevant to NIEHS

### **Discussion Highlights:**

- Need perspectives from wide range of disciplines including expertise outside of health sciences (astronomers, engineers, economists, electrical engineers etc.)
  - Examples are professions that look for patterns in large datasets
- Abundant data, need improvement in computations, hardware, software and training program
- Spatial-temporal technologies can be used to construct objective estimates of exposure in prospective framework
  - These issues would be essential to Critical Developmental windows.
- Think about non-traditional sources of data- commercial sources (insurance companies, grocery stores, advertising agencies) i.e. public private partnerships- to predict behaviors, exposures, SES among other critical variables.
  - Think about sectors with overlapping interests- transport, utilities, public civil engineering, flood control etc.
  - How to sort through/take advantage of industries/sciences that may have already developed relevant methods
- Important to attend to confidentiality issues
  - Start a discussion regarding different standards for commercial/industry use of data vs. public health research use.

### **Recommendations:**

- Workshop to bring together people from different government agencies, industries/sectors as well as researchers from different disciplines.
  - Discuss what data exist, how can it be linked, how can be leveraged in support of EH research.
  - Who is looking at/exploiting large datasets?
- What are the most relevant exposures that can be characterized using environmental informatics,
  - Built environment, Air pollution, diet, physical activity
- Develop a meta-data registry with an associated wiki-like criticism system for entering, validating and sharing for QA/QC.
- Invest in the pipeline of technical staff and researchers trained in Environmental Informatics
- Invest in statistical and mathematical modeling/methods development

**Discussion Participants:**

Balshaw, Boyles, Bradfield, Cory-Slechta, Drew, Fargo, Fasman, London, McAllister, Miranda, Umbach, Wright.