Human Computer Interaction and Data Science Focus Group

Data Science Symposium

March 5, 2014

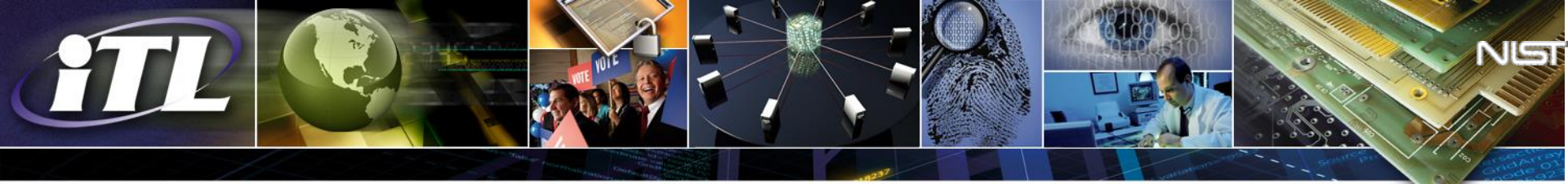
Human Computer Interaction (HCI) and Data Science

Data science technologies are changing the way we analyze data and make decisions. We are moving from a process of data analysis in order to better understand and confirm our models and theories to a data-driven discovery process. Data science allows us to use data to drive and develop our models while accessing much more heterogeneous and unstructured data. For data scientists, decision makers, as well as the public, the human computer interaction is crucial—users must be able to interact with both the discovery process and the output of the analysis. The user interfaces and the workflows must be intuitive and easy to use, in the specific contexts to enable innovations but not mislead users.

HCI and Data Science Focus Group

The purpose of this Focus Group was to begin to identify challenges that arise from human interaction with data science technologies. Among the questions we considered were:

- What are the current major data science needs for your users/customers? What role has HCI played in addressing these needs?
- How do you know your system is successful?
  - How do you measure success? As a developer? As a user?
- What HCI functionalities would be useful to you/your users/customers?
- How do you currently integrate the HCI with data science technology?
  - At what stages do your users/customers interact with the data and analysis?
- How do you measure the impact of the HCI on the decision-making process?
- What are the challenges? What can NIST do to help?



The Users

- Data Scientists
  - 80% of time is handling data, not analyzing
- Programmers
  - Don’t have subject matter expertise
- Executives
  - Don’t understand data analytics
- End Users
  - Technology Transition-how do you get them to adopt the tools?

The Human Interaction

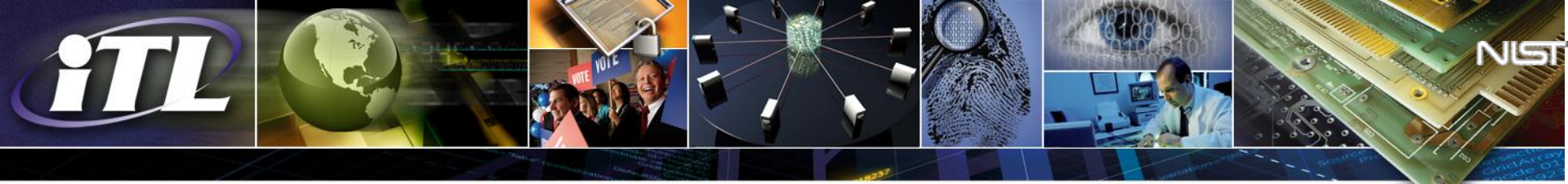
- Mobile devices
- Displays and projectors
- Real time interaction with the visualization
- Information overload
  - How do humans perceive the “big data”?
- Next generation interaction
  - E.g., haptics, speech
- Accessibility
  - Multiple languages

Functional Needs

- Assessing the accuracy of data
- Views for different consumers
  - Customizing for the user
- Presentation/summarization generation
  - Understandability, fit for purpose
- Data handlers: data cleaning and transformation
  - “Smart” assistants
  - No coding
  - E.g., Stanford Data Wrangler, Google Refine
- Ease of use in configuring and applying benchmarks

What NIST can do: Standards, Measures and Metrics

- Standardized vocabularies
  - between application specialists and developers
- Making analyses reproducible
- Best practice user interface guidelines for the analysis process
- Standardized formats for data
- Metrics
  - To measure effectiveness of human interaction at different stages during analysis
  - Ratio of data handling to analysis
  - Develop framework to identify quantitative metrics for different goals, e.g., for recommender systems: click throughs



Comments and Suggestions

Contact: sharon.laskowski@nist.gov

Thank you!