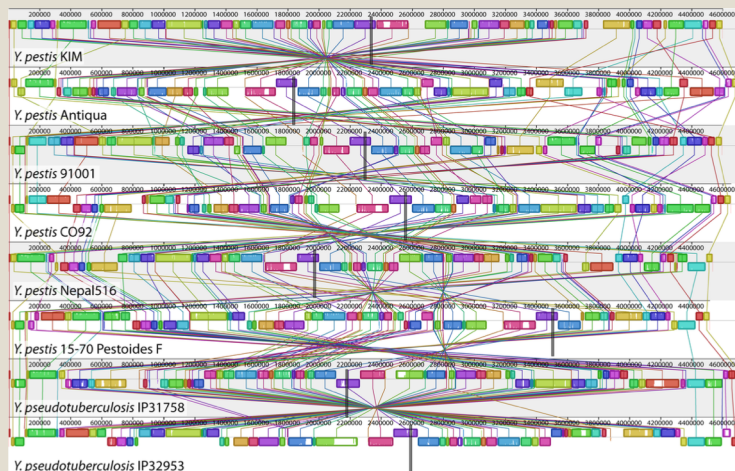
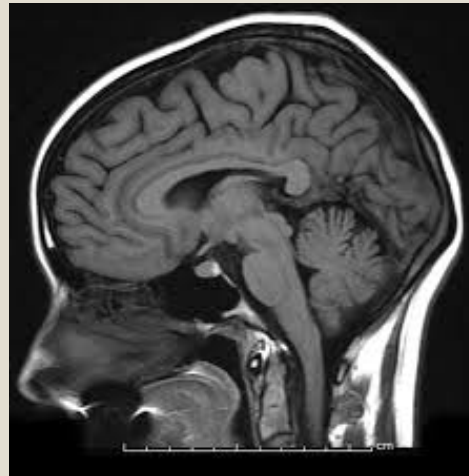




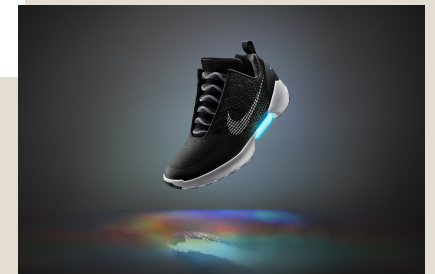
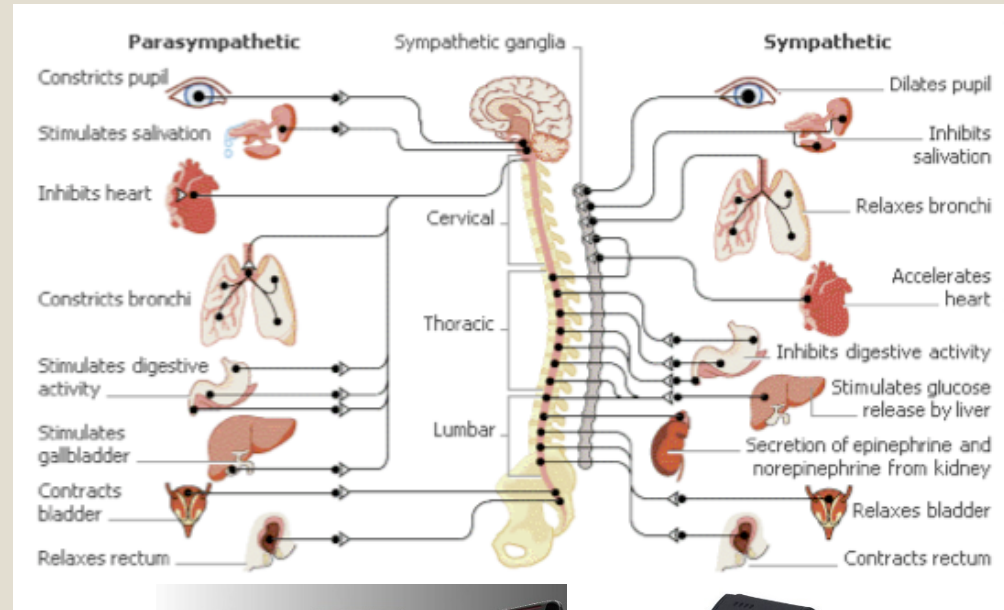
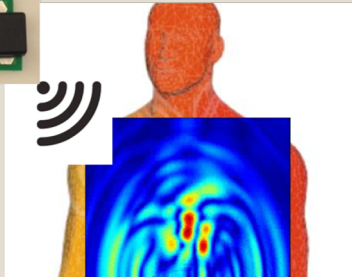
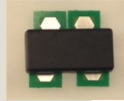
Estimating the Cost and Value of Data

Wendy J. Nilsen, PhD
Program Director,
Smart and Connected Health,
Directorate for Computer & Information Science & Engineering
National Science Foundation

Biomedical Data



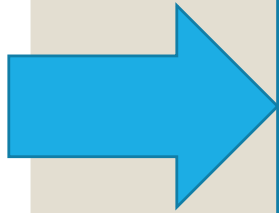
Biomedical Data also



Biomedical Data too



21st Century Data Opportunities and Challenges

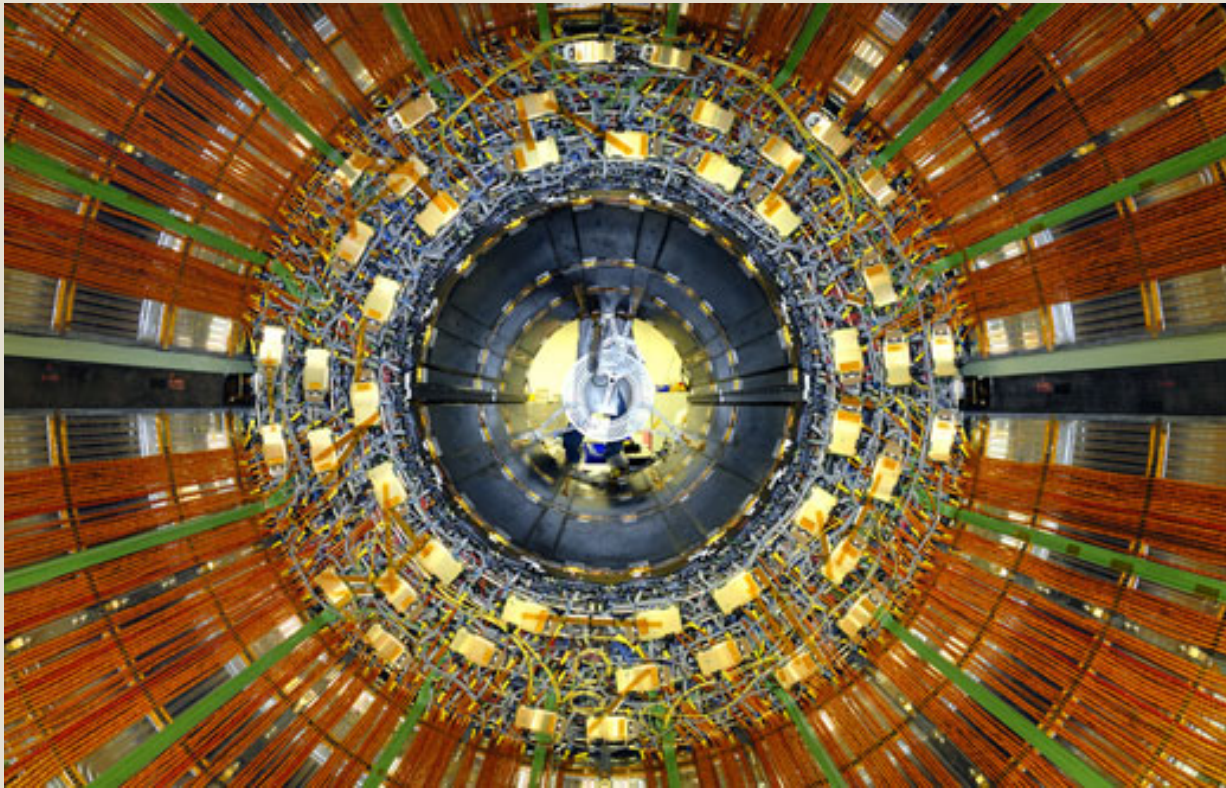


- Usable
 - Machine readable
 - Metadata
 - Accessible
 - Low cost to entry

- Massive Data
 - Validity
 - Irregular intervals
 - Missing data
 - Biases

- Analytics
 - Data linkage
 - Data fusion
 - Missing data
 - Data analytics
 - Dynamic data models
 - Computer Vision and Imaging
- Privacy and Security
 - Deidentified
 - Secure

Large Hadron Collider – Example of Hard Choices



[This Photo](#) by Unknown Author is licensed under [CC BY-SA-NC](#)

Is all data useful?

- What are the hallmarks for useful data?
 - Quantity
 - Reproducibility
 - High cost data (e.g., rare group samples)
 - Does the data already exist

What do data users want?

- Machine readable
- Accessible
- Low cost (both in time and money)
- High quality (meta data, indications of validity and reproducibility)

Determining which data is the best data?

- Use – analytics
- Crowdsourced value
- Feedback from users
- Diverse community governance
- REPEAT and REPEAT!



THANK YOU!

Wendy Nilsen, NSF Smart and Connected Health
703-292-2568
wnilsen@nsf.gov