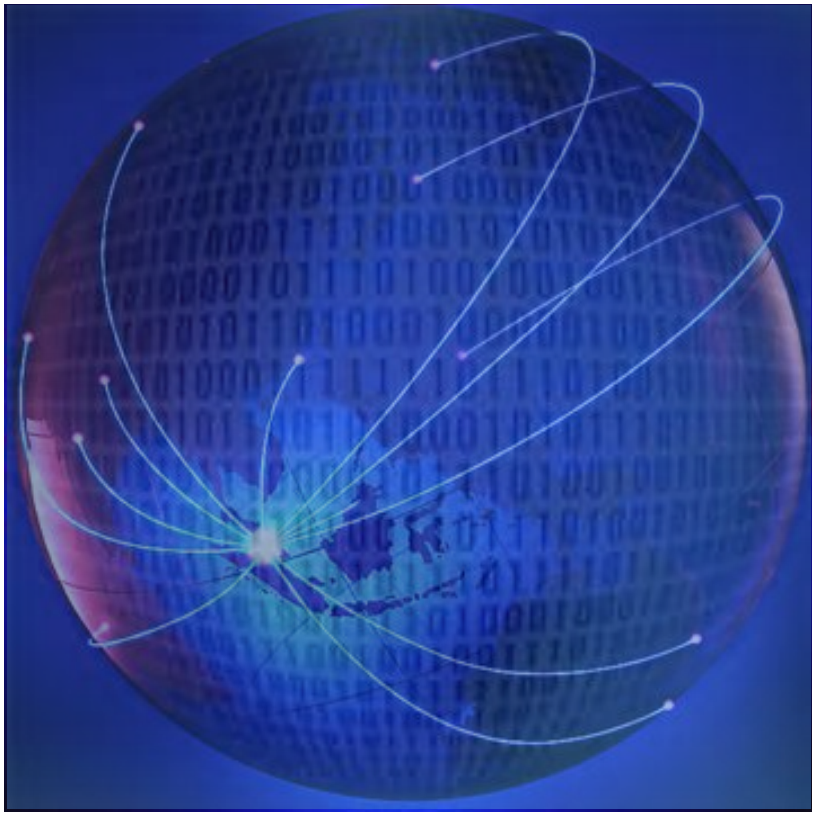
History Behind Data Management and Policy

W.A. Shelton

Motivation

- Data is being generated both experimentally and computationally at a very rapid pace
- The way science will be done in the future may be quite different than today
- Multi-disciplinary scientific teams initiated through a world-wide virtual scientific community
- This will cultivate the exchange of ideas through assimilation of data to solve important scientific problems
- It's being legislated

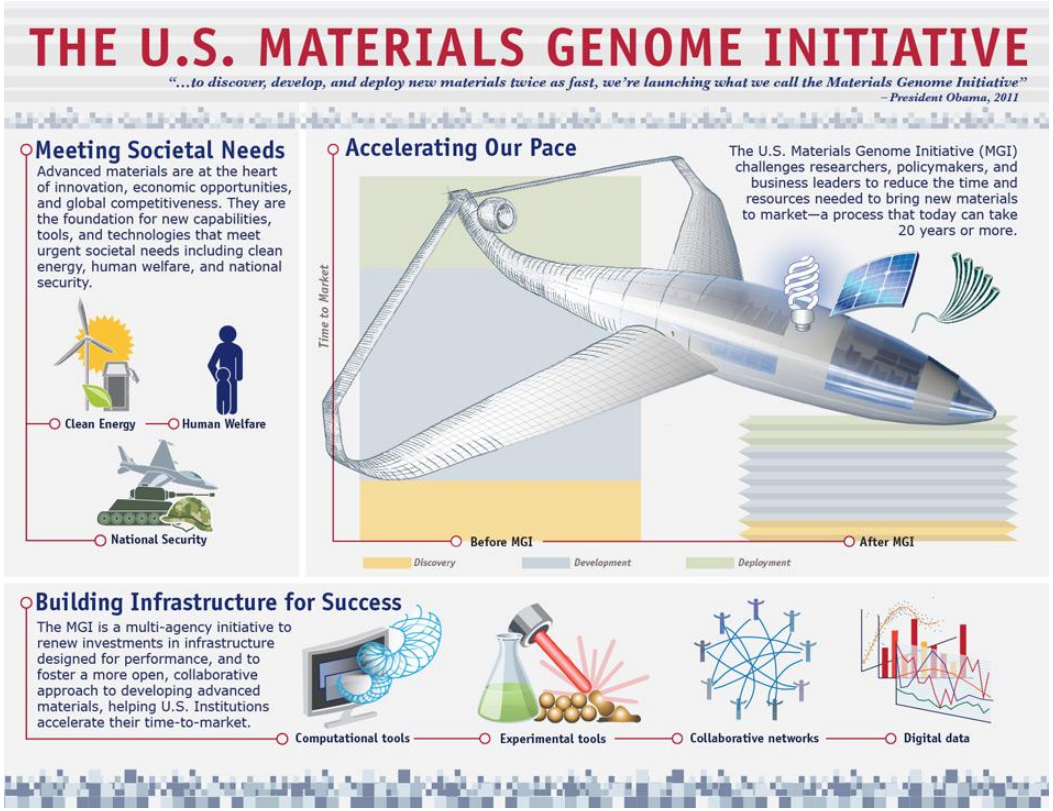


When did it all begin ?

- Senator Richard Shelby (R-Ala.) inserted a provision into the FY 1999 Omnibus Appropriations Bill (Public Law 105-277). “To require Federal awarding agencies to ensure that all data produced under an award will be made available to the public through the procedures established under the Freedom of Information Act.”
- Federal Research Public Access Act of 2006 (Sen. John Cornyn (R-TX) and Sen. Joseph Lieberman (D-CT))
- Federal Research Public Access Act of 2009 (Sens. Cornyn (R-TX) and Lieberman (D-CT))
- Federal Research Public Access Act of 2010 (Rep. Mike Doyle (D-PA))

And then came
"The Materials Genome Initiative"

Presidential Initiative: $1/2 billon dollars



- DOE,DOD, NIST & NSF
- doubling the speed an reducing the cost of discovering, developing, and deploying new advanced materials

Top 4 list

- Incentivizing open data and access to tools
- Structuring public-private partnerships
- Driving innovation across computation, data informatics and experimentation
- Moving the community to a different cultural norm

Call to Action

- Data-pre-competitive sharing, access, informatics
- Feedback to OSTP and Federal agencies

2013 Office of Science & Technology Policy (OSTP)

- Federal Research Public Access Act (FRPAA) of 2012 (Sens. Cornyn (R-TX), Wyden (D-OR), and Hutchinson (R-TX) and Reps. Doyle (D-PA), Yoder (R-KS), and Clay (D-MO))
- The Fair Access to Science and Technology Research Act (FASTR) of 2013 (Sens. Cornyn (R-TX), Wyden (D-OR). Reps. Doyle (D-PA), Lofgren (D-CA) and Yoder (R-KS))
- February 22, 2013: OSTP Director John Holdren has directed Federal agencies with more than $100M in R&D expenditures to develop plans to make the published results of federally funded research freely available to the public within one year of publication and requiring researchers to better account for and manage the digital data resulting from federally funded scientific research.

What's is this data ?

- Without any reference or description the data may not very useful beyond the person who generated the data
- We not only need a data plan we need ways to describe our data so that not ourselves but our future students, collaborators, etc. can access and be able to use our data
- This is where the workshop begins
  - Semantics and Ontologies for Physical Sciences
  - Curation & Provenance, Data Management, etc.
- The aim is to educate and to help provide direction for LA-Sigma Data management plan