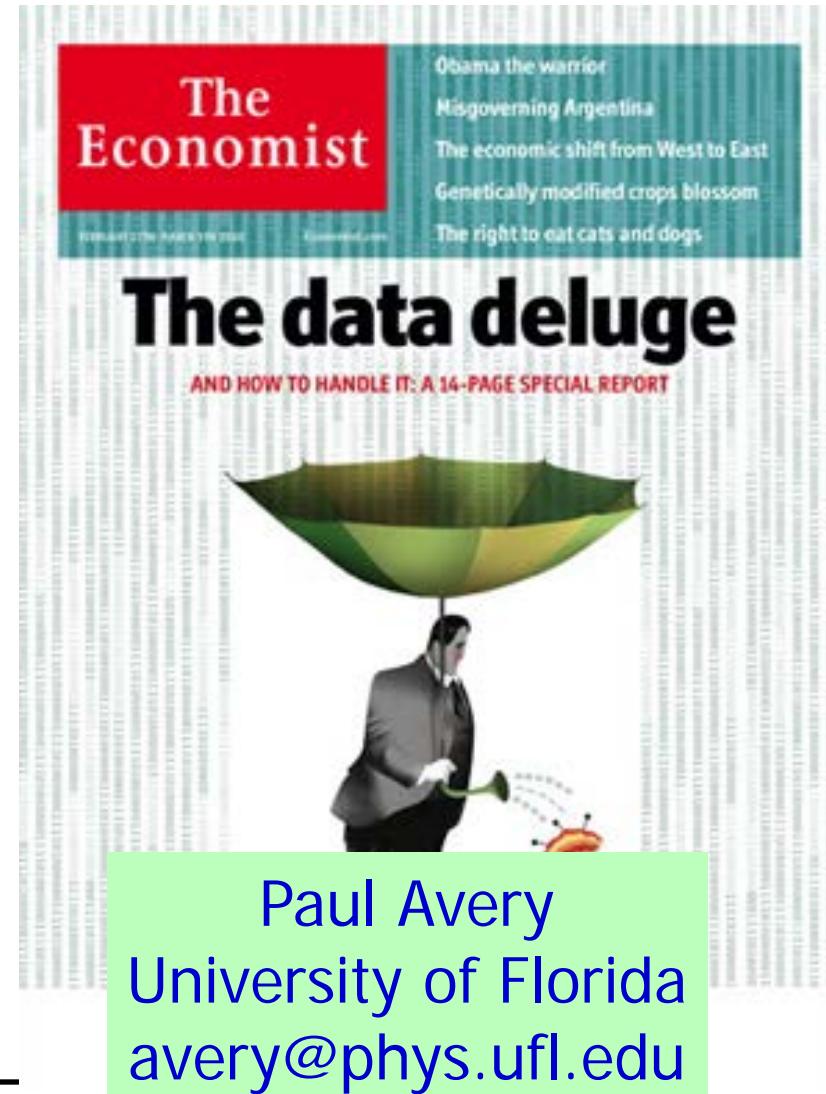


# Research Computing Day: Spring 2012

## Data

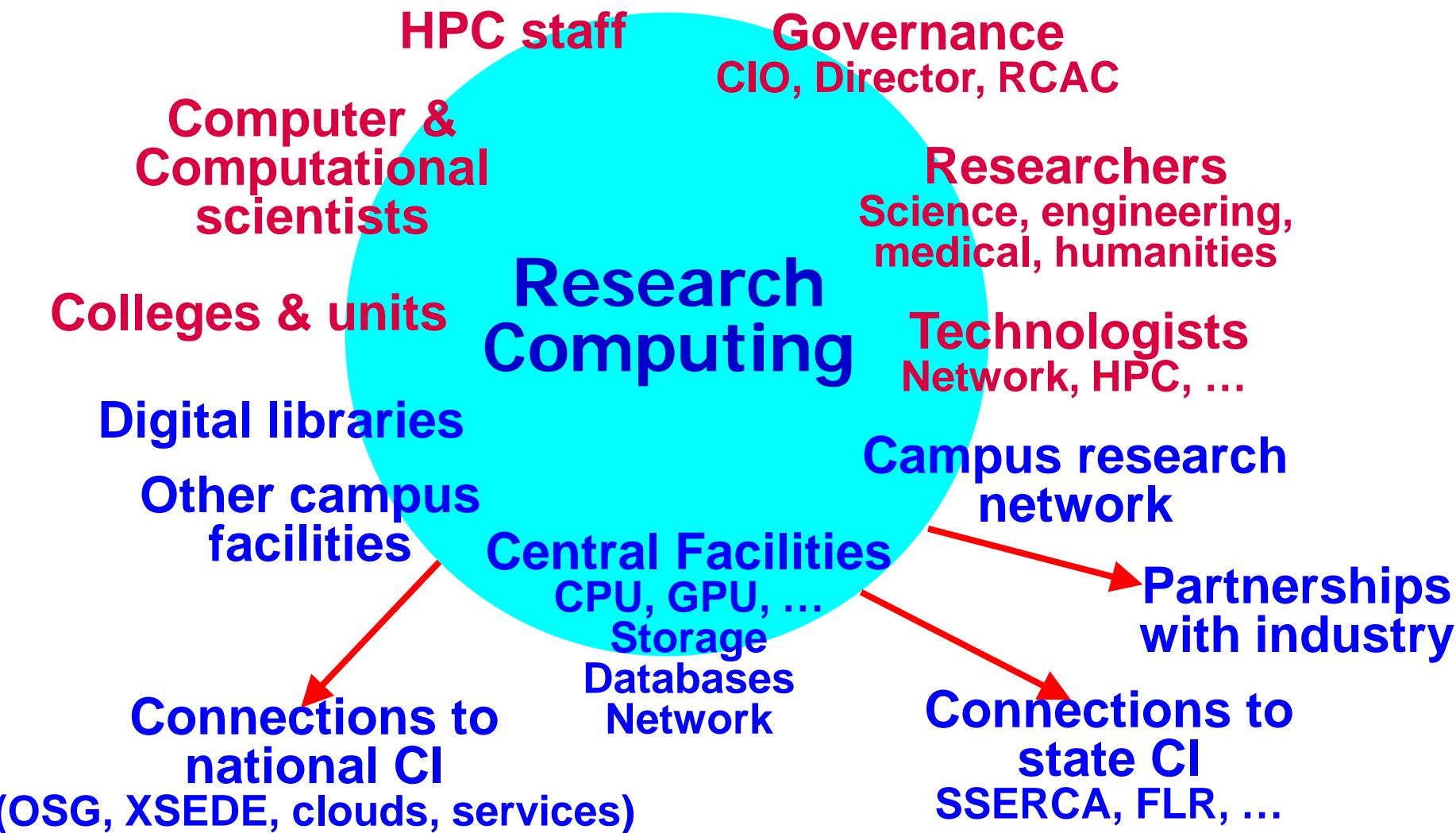


Paul Avery  
University of Florida  
avery@phys.ufl.edu

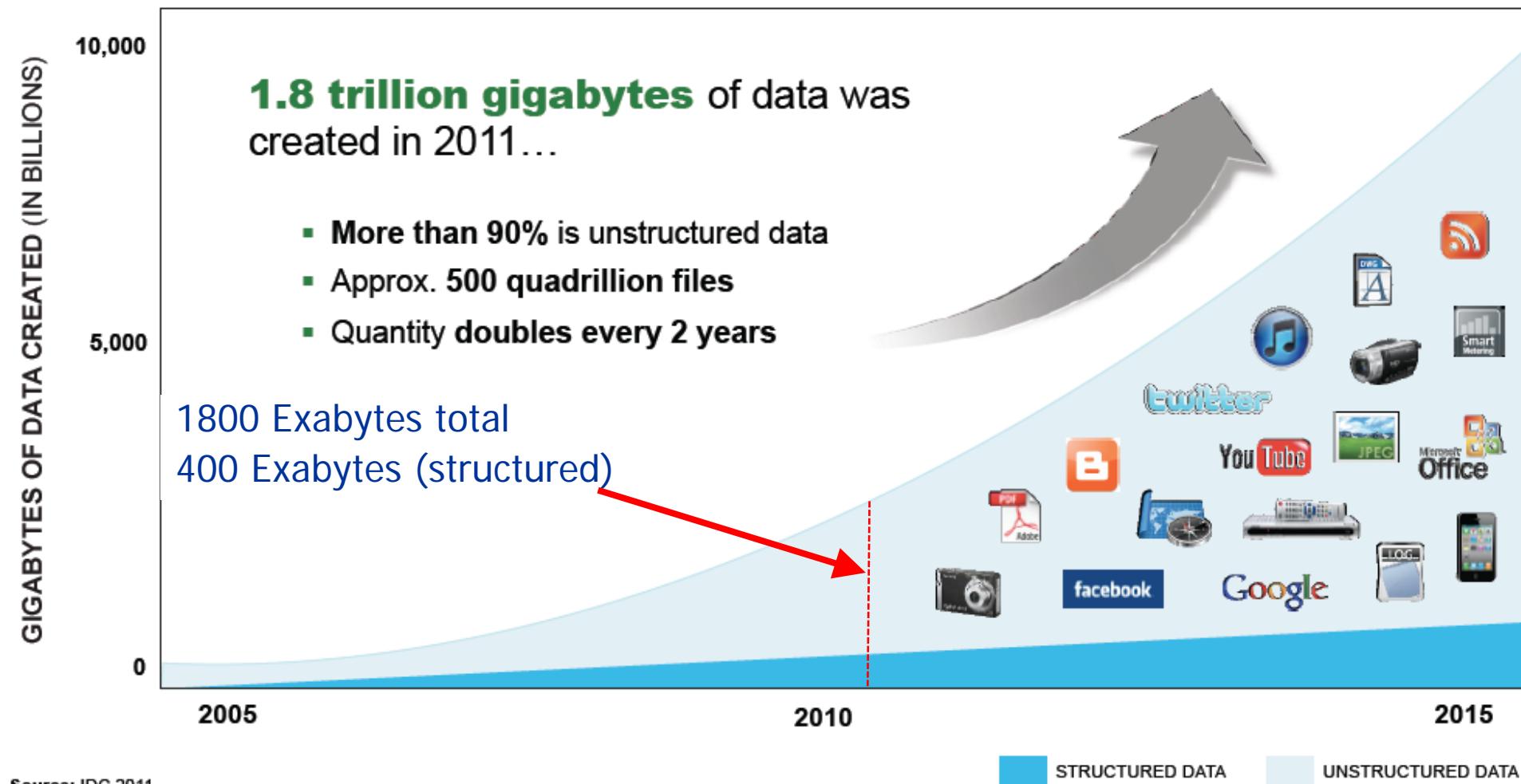
# Research Computing at UF

- 2011: UF administration made a major commitment to build solid infrastructure for Research Computing
  - ◆ Resources from Provost, VP Research, CIO
  - ◆ \$1.7M/year of *steady funding* for 5 years (was \$300K)
  - ◆ Matching program to encourage faculty investment
  - ◆ Encouragement to build computational biology infrastructure
- 2011 matching program (\$300K) was very successful
  - ◆ Strong participation by UF computational biologists
- 2012 matching program begins TODAY
  - ◆ See Erik Deumens or other Research Computing staff for details

# Research Computing at UF



# The Data Challenge



From Mike Olson, Cloudera

# Today's Focus is on Data

- Data increasingly seen as the most challenging issue in meeting computing needs
  - ◆ Scale (petabyte – exabyte), access, management, curation, etc.
  - ◆ NSF solicitations
  - ◆ New Internet2 data services (e.g., box.net)
- “Big Data” initiative inaugurated by White House
  - ◆ Webcast March 29
  - ◆ \$200M in new initiatives
  - ◆ Six agencies and departments are involved

[www.whitehouse.gov/sites/default/files/microsites/ostp/big\\_data\\_press\\_release\\_final\\_2.pdf](http://www.whitehouse.gov/sites/default/files/microsites/ostp/big_data_press_release_final_2.pdf)

Today's agenda will focus on data issues

# Agenda

9:00 – 9:15	Welcome (P. Avery)
9:15 – 10:00	Research Computing Infrastructure at UF (E. Deumens)
10:00 – 10:15	Break
10:15 – 11:00	Data Management in Biology (Rich LeDuc, Indiana University)
11:00 – 11:45	Data Management 101 (H. Norton, R. Milan)
11:45 – 1:00	Lunch
1:00 – 1:45	Data Lifecycle Management (H. Norton, R. Milan)
1:45 – 2:30	Poster presentations (research groups)
2:30 – 2:45	Break
2:45 – 3:30	Q & A session (RC and Library staff)