The Institute of Quantitative Social Science

Gary King





The Social Science Data Revolution

The Last 50 Years:

- Survey research
- Aggregate government statistics
- In depth studies of individual places, people, or events
- The Next 50 Years: Spectacular changes due to. . .
 - Much more of the above
 - Shrinking computers & the growing Internet
 - Analogue-to-digital transformation of government records
 - The replication movement: academic data sharing (e.g., Dataverse)
 - Advances in methods
 - The march of quantification: through academia, professions, government, & commerce (MoneyBall, SuperCrunchers, The Numerati)
 - The end of the quantitative-qualitative divide:
 - Diverse perspectives remain but, with text, video, field notes etc. as data,
 - We now all benefit from the same technologies
 - Advantages of collaboration are growing



IQSS

Social scientists need vastly improved infrastructure.

Harvard's response: IQSS -- a hybrid organization

- A traditional research center: facilitating research
- Part of the Harvard administration: building infrastructure
- (Synergies: turn administrative issues into quasi-research projects)





Need Bigger, Faster Computers than you can afford?

IQSS Research Computing Environment (RCE):

- > 1300 shared computing cores
- 450 active users
- A persistent environment, sign on from anywhere
- Growing via collaborative research project (with Harvard CTO Jim Waldo and SEAS) to extend our platform to the Amazon cloud





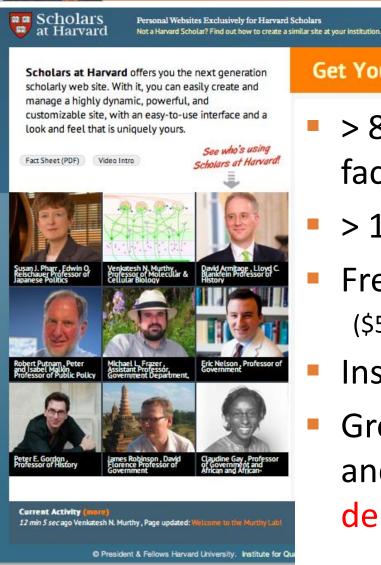
Consulting and Training

- Short courses: R, Stata, SAS, Numeric Data Resources, Atlas.ti, GIS, & many IQSS technologies
- Research consulting: one-on-one help with data management and analysis
- Survey Consulting: help with question wording, sampling, analysis, nonstandard data collection
- Recently (re)grown with help from FAS





OpenScholar: Faculty & Student Web sites



Get Your Site Now €

- > 800 scholar sites, including 300 by faculty
- > 120 project sites
- Free at scholar.harvard.edu (\$5-20K saved per site)
- Installation at > 50 other Universities
- Growing via collaboration (with HUIT and HPAC) to extend OpenScholar to department sites





Dataverse Network



Harvard's home for social science research

Creating, sharing, and preserving scientific knowledge about human societies.

IQSS Dataverse Network

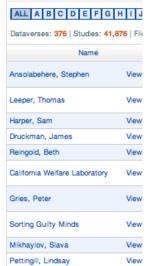
Search User Guides Report Issue

Welcome to the IQSS Dataverse Network!

Access the world's largest collection of social science research dat across or browsing through one of the virtual data archives (called 'You may also create a dataverse of your own, backed up in perpet Murray Archive, which may easily be customized to appear as if it in fact is served by the IQSS Dataverse Network. This means that a visibility, and access control for the data devolve to you, but all the guarantees, and software and hardware upgrades and maintenance IQSS. Learn more about the Dataverse Network Project ...

All Dataverses





Your own easy-to-use virtual archive

- Our installation: ~ 1000 dataverses,
 41,000 studies, 671,000 files
 - > 15 installations in other Universities
- Growing via collaboration (with the Harvard Library) to expand infrastructure for diverse projects across humanities, social sciences, & sciences





Zelig: Everyone's Statistical Software

- Easy, standardized solution to R's power and diversity
- ≈100 statistical methods in common format
- > 100,000 users worldwide
- (Optional GUI through Dataverse, no R knowledge required)
- Large network effects: students learning from each other
- Growing via (open source) contributions

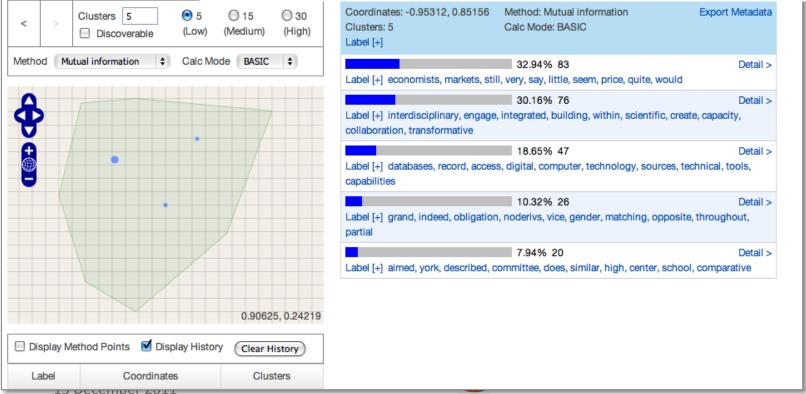




Program on Automated Text Analysis

- Text
 Clustering

 Set Name: NSFC
 Size: 252 Files
 Description: NSF complete corpus, updated
- "Computer-assisted reading", a new way to organize and make sense of huge amounts of unstructured text.
 - Advising with existing approaches available now
 - New approaches under development, for public use in 2012





Grants Administration

- 37 active grants
- 30 40 proposals submitted annually
- Pre- and post-award support
- Working closely with department staff
- Social scientists from FAS and other schools come for help & stay to interact on research & teaching





IQSS Activities Build a Community of Scholars

Administrative Services

- Support for PIs, Programs and Centers
- Sponsored Research Administration

Research/Affiliated Programs

- Center for Geographic Analysis
- Program on Survey Research
- Program on Text Research
- RWJS in Health Policy Research
- Positive Political Economy
- Data Privacy Lab
- Workshops and Conferences
- Program on Methods



Technology Development

- OpenScholar
- Dataverse Network
- Zelig
- Text Clustering Tool
- Research Computing



Technology Services

- Research and Technology
 Consultants
- Technical Training
- Research Computing Environment (RCE)
- Repository for Social
 Science Data (Dataverse)
- Faculty Web Sites (OpenScholar)