

Statipedia: a platform for collaboration across statistical agencies

For ICSP Innovation group meeting, 23 June 2010

Peter B. Meyer

Office of Productivity and Technology, BLS
and

James A. Buszuwski

Office of Compensation and Working Conditions, BLS



... with lots of advice from others.

This research does not represent the
Bureau itself, only the named authors.

www.bls.gov

Vision: work together on methodology online

- Tools and workspaces for statistical staff across agencies
 - With source material for economists and statisticians
 - Available to new employees
 - Not for the general public
- Web tools
 - Wikis
 - Version control system for source code
 - Search engine
 - Blogs, statistical languages, and more

OMB on OpenGov

OMB directive (Dec 8, 2009)

“The Plan should include proposals to use technology platforms to improve collaboration among people within and outside your agency.”

At DOL:

- OpenDOL web site invitation: “Share your ideas”
- Visitors could click “agree” or “disagree”
- This proposal tied for 3rd out of 130+ proposals

Model government collaboration platform

After 9/11, the U.S. intelligence agencies reviewed their communication tools and processes.

- They added web tools and work spaces to their network
 - an interagency wiki: “Intellipedia”
 - blogs, tagging, search, instant messaging, videos
- This made it easier to find and use information across the 16 agencies.
- We met the Intellipedia founders; they presented at BLS.

Other government collaboration platforms

- EPA's blogs and wikis
- DOD's Techipedia and forge.mil
- OMB's MAX
- OECD's WikiProgress.org
- British and Canadian civil service common platforms
- Eurostat publications wiki

In future: DOL's wikis, GSA's apps.gov, OMB's SCOP

➔ These models support Open Gov concepts

Examples of wiki uses

- Repository for definitions
- Topical explanations:
 - Seasonal adjustment methods
 - Occupation and industry classification
 - Translation and harmonization of data series
 - “Green jobs,” “high tech jobs”
 - “Systemic risks”
 - How to measure and reduce such risks?
 - And fast discussion as at intelligence agencies

Search capability

If it's easy to search **all** the sites, then

→ Users find specialists by their contributions!

➤ Wiki pages, blogs, code

➤ knowledge is easier to discover

→ Communities emerge

➤ User groups

➤ Collaborations

➤ Standards develop

Other tools for analysts

- Blogs for projects and persons
- Statistical computing languages
- Geographic information systems
- Videos for training
- Repository of documents
- **Source code control**
 - Stores source code of computer programs
 - Tracks relationships among files (E.g. these 10 go together to make one program)
 - Supports study and fixing of bugs
 - Eases working together over distances
 - Possible host: GSA's apps.gov

Administrative principles

■ Scale up

- ▶ Per capita costs fall and benefits go up (as on Web)
- ▶ Broad open audiences → emergent communities

■ Meet open technical standards

- ▶ E.g. HTML, TEX, wiki-text links, hooks/extensions
- ▶ So new components fit in easily
- ▶ Enable extension by interested government staff

■ Enable and encourage participation

Norms to encourage online

- Anchor discussions to sources, evidence, theory
 - ▶ This requires a technically rich work environment
 - Enable drilling down to sources
 - Support hyperlinks, equations, footnotes, source code
 - Address broad/open audiences on topics
 - ▶ Welcome open cross-talk
 - ▶ Not organization-specific
 - E.g. “green jobs” both labor and environmental
- ➔ Such practices encourage scientific community and reproducible results

Virtual Communities

- Likely examples:
 - Productivity (BLS, BEA, Fed, DOE, USDA, NIST, ...)
 - Seasonal adjustment (Census & X12 users)
 - Survey methods
 - Information technology
- These virtual communities can have:
 - common pools of source material
 - mutual awareness
 - potential for review of one another's work

Savings over time from platform

Reduced costs for software:

- Development and maintenance
- Acquisition, approval/certification

How?

- Less duplication of effort across agencies
- Costs are amortized over more organizations
- Specialization benefits a larger audience

Long run effect -- better knowledge management

It will be easier to:

- find specialists and experts
- prepare for retirements and turnover
- integrate new staff
- reproduce and confirm scientific results
- reduce overwhelming flows of email
(Less “dissemination”; more “discoverability”)

Moving forward: a pilot project

- Proposal: actively try shared web tools and workspaces for scientific collaboration across agencies
 - Joint with partner agencies
 - Usable experimental platform
 - Evaluate/report in 1 year on users and content
 - Budget goal for agencies: zero
- Purpose: evaluate benefits, costs, and pitfalls of a permanent Statipedia

Pilot project: outline

- An ideal pilot project:
 - develops a wiki of statistical & survey methods
 - covers topics of joint interest
 - meets the OMB OpenGov collaboration directive
 - offers new services to the staff

- Cross agency participation requires:
 - a host (EPA, OECD, DOL, SCOP, GSA)
 - users with authorization

- The project would not:
 - involve production processes
 - use sensitive information, e.g. sensitive PII

Benefits to the Statistical Community

By participating in a productive, high-morale pilot project, ICSP agencies:

- make it more likely that future joint systems serve their needs
 - ✓ prepare for new ways of sharing code, data, and methodology in a safe work environment
- unlock value from one another
 - ✓ work smarter, not harder
 - ✓ liberate methodology!