



Yes We Scan: US government access policies and the revived attempt to 'Reboot .Gov'

David Vaile, Executive Director

Cyberspace Law and Policy Centre
UNSW Law Faculty

April 2009

<http://cyberlawcentre.org/unlocking-ip/2009/>



Outline

- Malamud, GPO, and the committee
- Reboot .Gov
- Rural Internetification
- Library of USA, Publishing Academy
- Fed Flix
- So What?



Once upon a time in the USA...

US: home of Creative Commons and other 21stC advances for the public domain, as well as early examples of effective public information platforms.

Government 'commercial' decisions or inertia last century negated much of the earlier promise of initiatives such as the Government Printing Office.

Result: key information lost to public access, or inaccessible for want of resources.

Public access to US legislation, cases, and regulatory material has been limited, compared to other initiatives e.g. the Free Access to Law movement.

A patchwork of progress or regression in related fields of public information:

- Standards
- Safety
- Skills for public information creation.

The new digital-friendly Obama regime in Washington presented a fertile opportunity

- renewed push by long-time activists (YesWeScan.org), and a coalition of supporters
- break the paralysis, and leap-frog to the front of the field.
- This paper looks past the early promise and past failures, focuses on the key elements of the new proposals.



Carl Malamud and the GPO

What is the GPO?

- US Government Printing Office
- ‘opened its doors the day Abraham Lincoln was inaugurated’
- ‘an agency which has worked to “Keep America Informed” for 148 years’
- Publishes key material such as Federal Register.

Who is Carl Malamud? (Public.Resource.Org)

- Online activist: ‘08, publish 32M pp. primary legal materials, videos, photos.
- 90s: fought to place the databases of the US on the Internet.
- 80s: fought to make standards governing ‘Net *open standards* avail. to all.
- 2008 campaign to become head of the GPO, on a platform of revitalizing the GPO and ‘rebooting .gov’

Who was Augustus Giegengack?

- Malamud’s hero, ‘the greatest public printer since Ben Franklin’
- A working printer and regular leather apron man inspired to public service by the election of FDR, ran arly grass roots campaign to be appointed Public Printer of the United States, head of the Government Printing Office.



The Committee to Reboot .Gov

Honorary Chairs

- **Lawrence Lessig**
law.stanford.edu
- **David Halperin**
campusprogress.org
- **Paul Vixie**
isc.org
- **Tim Bray**
tbray.org
- **Tim O'Reilly**
oreilly.com
- **Tim Stanley**
justia.com
- **Cory Doctorow**
boingboing.net
- **Brewster Kahle**
archive.org
- **Paul Jones**
ibiblio.org
- **Joichi Ito**
creativecommons.org
- **Thomas Bruce**
law.cornell.edu
- **Ellen Miller**
sunlightfoundation.com
- **Tim Wu**
law.columbia.edu
- **Jerry Goldman**
oyez.org
- **Gary Bass**
ombwatch.org
- **Mark Frauenfelder**
boingboing.net
- **James Grimmelman**
nyls.edu
- **Fred von Lohmann**
eff.org
- **Pamela Samuelson**
law.berkeley.edu
- **Marshall T. Rose**
public.resource.org
- **Adrian Holovaty**
everyblock.com
- **Paul Ohm**
law.colorado.edu
- **Jed Sundwall**
jedsundwall.com
- **Ted Lemon**
nominum.com
- **Marc Rotenberg**
epic.org
- **Nicholas Negroponte**
laptop.org
- **Aaron Swartz**
aaronsw.com
- **David Gold**
spindlelaw.com

Reboot .GOV (1)

What People See Today:

Very dense 1-page PDF files or unformatted text, not suitable for reading on-line. Navigation a primitive browse/search. No notification or other mechanisms.

Aim: A more visual, usable Federal Register (75k pp/yr), groundwork for a fundamental change in how US distributes Official Journals of government, and how citizens participate in the rulemaking process.

FR uses SGML, predecessor to XML. GPO makes raw SGML feeds available as products. These offer opportunity to leapfrog the current system.

- The current system has no **navigation**. One should be able to easily go backwards and forwards in time, by agency, by code section, and by topic.
- Current search capabilities are primitive and do not take advantage of **metadata**, such as official thesaurus of index terms, available in raw feeds.
- FR is time-sensitive: proposed meetings, when rules become effective, deadlines for bidding. A **calendar** and newsfeed-based **notification** mechanism should be an integral part of the user interface.
- No **permanent, human readable URLs** one can link to. Instead, you navigate by page number or search: impossible to build a blog entry that links to FR. One should be able to easily build a link to a particular page or article.

Reboot .GOV (2)

- Documents signed, site use a secure link with a valid certificate, and Secure DNS. People should be certain they see **authentic** government document.
- **Links:** when section of the Federal Register references *U.S. Code*, Code of Federal Regulations or other law, one should be able to click and see it.
- FR incorporates by reference many technical standards. One should be able to **read the incorporated documents** as part of FR.
- Current FR is not “**search engine optimized.**” Keywords, dates, authors, and other important info built into the headers of these files...
- One should be able to see how a proposed regulation would change existing Code of Federal Regulations: a “redline” of **current versus proposed.**
- Law suits are based on past events. Code of Federal Regs and all related compilations put into Source Code Control, so you can navigate **forwards and backwards in time** to see what regulations or laws were in effect.
- There are a large number of **advanced visualization techniques**, from logic maps to tree diagrams, to sort through the mass of information.
- For each article, a way to provide **feedback:** a link to the existing Regulations.Gov site, an agency-specific forum, or just email address.

Reboot .GOV (3)

3 Levels of Access

These examples have to do with usability and navigation of the site hosting the Federal Register, as well as visual appearance. However, private groups digest Official Journals and build systems to supplement what the government does, and in many cases are much better. Need to provide access at 3 levels:

1. At the lowest level, **bulk data** in the original SGML, as well as PDF and HTML versions should be available as signed zip files and at no charge. The **current \$17,000 price** tag has a huge chilling effect on innovation and is not appropriate for works of government.
2. An API should allow any **blogger** to embed a document, a part of a document, or a document stream into their own site, much as they embed a YouTube video or playlist.
3. The same API outside developers use can then be used by the government to build a better web site, **Federal Register 2.0**.

Providing an API and bulk data is a way to hedge our bets. Even if the government can't build a better web site, we can count on the public doing so. As a public printer, there is no justification for the U.S. government withholding raw feeds of public domain data.



Reboot .GOV (4)

The 4-Year Goal: *America's Operating System*

- The Federal Register system represents ... the official publications of the executive branch. A large stream of other documents come from the legislature and judiciary, forming a collection of **primary legal materials** that make up “America’s Operating System,” rules that govern our society.
- A goal of the new administration should be to make America’s Operating System open source, guaranteeing **a complete and current archive of all primary legal materials in the US are freely available on Internet.**
- This goal is partly about democracy, allowing citizens to see the rules that govern our society, but America’s Operating System is also about innovation, guaranteeing that any scholar or entrepreneur can download our legal materials and develop new and more effective ways of presenting, practicing, communicating, and learning about the law.

What We Will See in 4 Years:

- Making all primary legal materials available will change the way the law is practiced in the United States, making the rules of our society available to all.
- Reaching beyond the beltway to incorporate feedback into the rulemaking, legislative, and judicial processes will make government more transparent, relevant, and timely.



Rural Internetification

RURAL INTERNETIFICATION ADMINISTRATION - Grassroots NBN?

Aim: to bring high-speed broadband to 98% of rural Americans, as Rural Electrification Administration did for electricity last century:

- Repurposing the Amateur Radio League
- Modifying radio spectrum policy
- Injecting capital into rural co-ops [not gov't utility ownership/sell-back.]

Echoes of the techno-egalitarianism of the early Internet:

“Amateur radio is based on the bargain that, if an individual demonstrates **technical competence** and a **commitment to public service**, they are allowed additional transmission privileges.

By updating those transmission privileges, for example allowing a Class 1 Wireless Engineer to access **unused portions of the television spectrum** in rural areas for use in civic, emergency, and other cooperative systems, one could quickly begin to deploy a new class of protocol-based communications solutions.”



The Library of the USA

- A book series and **public works job program** to create an archival series of curated documents drawn from our cultural institutions.
- Full-quality **masters** of the books and research materials made available for other publishers to draw on.
- The program would employ the GPO master printers and would recruit writers, archivists, artists, and other creative workers through a national call for participation.
- Producing high end books taps into some of the talented resources at the GPO, **preserving jobs** by shifting printing towards the high end. Complements the high-volume low end digital publication of much other routine public information.
- If the books are fundamental in nature ... the Library of the U.S.A. ... would **sell well** in bookstores and on-line venues such as Amazon, recovering many of the costs of production.
- By posting high-quality PDF files of the finished product as well as full-resolution imagery used in production, we create an invaluable on-line **resource** for future generations, and for other publishers to build upon.



The US Publishing Academy

- GPO should expand current training programs such as the Institute for Federal Printing and combine them with current workforce development efforts to create a national academy similar to the National Mine Academy and the National Fire Academy
- Training its own workforce, the government, and the local schools in the art, craft, and science of publishing.
- A program of education ... range from the traditional crafts such as bookbinding, design, and typesetting to their modern equivalents of high-speed photocomposition, advanced print-on-demand services, and the modern languages of PostScript, CSS and HTML.
- A program of testing and certification can provide an objective measurement of the skills and knowledge attained, and the materials developed for certification programs and curricula can be offered on-line for others to use without restriction, helping fuel similar efforts throughout the country.



FedFlix—No Late Charges in the Public Domain!

Stock Footage for a Remix Era

- While mainstream media struggles with the onslaught of the Internet, one bright spot has been the explosion of **documentary filmmakers, independent news producers, vloggers, mashup and remix artists**, and others who incorporate audio and video into their works.
While video production has blossomed, clearing copyright on film clips or photos is the hardest part of making a multimedia production.
The government can change that.
- With public domain materials, one does not have to ask for permission. The public domain is an essential common resource, allows both non-commercial and commercial world to thrive. Scholars writing books often need to use photographs, and long complained **clearing rights is next to impossible**. Travel writers ... complain how hard it is to clear rights, even public resources such as photographs from Smithsonian.
- US has no consistent program for posting audio, video, and photo materials online. We demonstrated with our **FedFlix** program (588 films uploaded) how easy it is to quickly change that, and **create a common resource** by posting films and photos. The federal government should up the ante, and adopt a goal of posting 5,000 films and 50,000 photographs next year, coordinated by an agency such as GPO.



Why is this interesting?

- Specific-purpose campaigns where you have direct election to a public office – very different US system.
- Vision draws on historical stories, does not claim to be totally novel.
- Grass roots involvement in creation of platforms, not just central government mega-programs
- Recognises potential for skilled individuals and businesses to be stakeholders/contributors/republishers, not mere ‘users’.
- A mix of both grass roots and central programs, working together – not in opposition, part of the same mixed system.
- Recognises the role of mixed and hybrid business models, and the value-adding that greater access confers on government material.
- Both production *and* consumption of content, in all media, as normal acts to be supported – dual role, not just ‘consumer’.
- Central role for government to catalyse, support, but not control?
- Focus on practical implementation, not legislation/legal instruments



David Vaile, Executive Director

Cyberspace Law and Policy Centre
UNSW Law Faculty

<http://cyberlawcentre.org/unlocking-ip/2009/>