Help! I’m an Accidental Government Information Librarian

presents

Fugitive hunters

Community-Based
Digital Collection Development
of
Born-Digital
Government Information

with James A. Jacobs and James R. Jacobs
Part 1

Introduction, background, terms, scale, focus

Notes, links, slides, etc. on FGI

FreeGovInfo.info/fugitives

• You can find links to the forum, the report, my speaker notes and more at freegovinfo.info/fugitives
What is a Fugitive?

Any “document” that, according to the 44 USC definition, *should* be part of the depository system, but that is not in the system.

Government Document: “*Informational matter*” published at Government expense or as required by law
Example
(not a new problem)

a “Committee Print”
Why a Fugitive?

- Administrative errors (GPO)
- Agency non-compliance with requirements (desktop publishing)
- Agencies publish with private sector and place publication under copyright (privatization)
- Agencies use a Title 44 exemption for publications “which must necessarily be sold in order to be self-sustaining"
- Agencies distribute information via agency web sites without notifying the FDLP

Examples
from DiMario (1997)

- Scientific and technical documents
  (70% of tech docs: 55,000 in 1996)
- *Journal of the National Cancer Institute* (Oxford University Press)
- Court decisions
- Federal Election Commission financial disclosure statements
- CRS reports
Examples
from Bower (1989)

• Air quality benefits of alternative fuels.
• Greenhouse effect, sea level rise and coastal wetlands.
• Reagan administration regulatory achievements
• A Report to the Secretary on the homeless and emergency shelters
• Report, Task Force on Women in the Military.

Not a new problem.
Past Strategies.

• **Institutional.** The Library of Congress ran a service called DocEx from 1946-2004. Libraries could subscribe to this service and receive fugitives tracked down in D.C. (See [Thomas Shaw, 1966](#))

• **Technical.** GPO microfiches copies of documents printed elsewhere.

• **Individual.** Librarians captured fugitives by scouring agency newsletters and press releases and local newspapers and by cultivating agency contacts.

• **Legal.** Attempts to revise Title 44
Cynthia Bower study

- “No one knows”
- fugitive documents outnumber depository documents by an average ratio of eight to one
- 43% of docs in American Statistics Index
Scope

- About 50% of the universe of Federal printing
- 78% of NIH's publications

Scope
What is in FDsys?

• **Congress.** More than half of the [FDsys Collections](#) are explicitly Congressional.

• **Courts.** Opinions of some, but not all, federal appellate, district, and bankruptcy courts, dating back to 2004.

• **Executive.** Big gaps. (Apparently only 56 of 246 “government authors” in Fdsys are agencies.)
Summary

• “Fugitives” are not a new problem.
• The number of fugitives is very, very large.
• The information content of fugitives is important
• There is no fool-proof strategy/solution to the fugitive problem, yet.
So…

In absence of a universal solution, every library can contribute to a loosely-coupled, decentralized strategy/solution.
Technical Notes


Example: Keystone XL Pipeline Project Final Supplemental Environmental Impact Statement (SEIS)
Technical Notes


To preserve this “Page,” you would need to preserve 13 Files (URLs).

- The HTML file that is this page at its URL, plus the other files that comprise the contents of this page:
  - 7 images files
  - 3 javascript files
  - 2 css style sheet files
Technical Notes


To preserve this “Title,” you would need to preserve

- 94 PDF files that comprise all the 11 volumes and all the parts of the Environmental Impact Statement.
Technical Notes

2. Which version, edition, or copy?

Example: Executive Order 13662 (Blocking Property of Additional Persons Contributing to the Situation in Ukraine)

• The GPO Authenticated version:

![Executive Order 13662 image]
2. Which version, edition, or copy?

Example: Executive Order 13662 (Blocking Property of Additional Persons Contributing to the Situation in Ukraine)

- A total of ten different URLs have (apparently) the same information and metadata about the content.
Technical Notes

3. dot-gov, dot-mil, or beyond…?
ABCs of Fugitive hunting
Needles…
...or Haystacks
Link rot

Drops, oceans & reservoirs

Flickr waterfall picture by discordia1967. That's actually me at Hanakapi`ai falls in Kauai :)

Flickr photo by Elle Is Oneirataxic. Attribution-NonCommercial-ShareAlike 2.0 Generic Creative Commons license


Web Archiving

EEMs
EEMs

- Everyday Electronic Materials
- Serendipitous fugitive collection
- Currently tracking 8 agencies (from lostdocs.freegovinfo.info) (HT Laura Lind!)
- Collecting the Web 1 drop at a time!
- CNI report on EEMs bit.ly/cni-eems
Increasing analytic capacity of state and local law enforcement agencies: moving beyond data analysis to create a vision for change

English.


1 online resource (20 pages) : color illustrations

Contributors

CONTRIBUTOR


Subjects

UNITED STATES

Law enforcement > United States > Data processing.

Data mining in law enforcement.

Bibliographic information

PUBLICATION DATE

2012

http://searchworks.stanford.edu/view/10355660
Harvesting the ocean

Harvesting Websites since 2007 with Archive-it:

- Fugitive agency sites (GAO, EPA, NACA/NASA, Sourcebook of Criminal Justice Statistics, NOAA Deepwater Horizon Archive, Keystone EIS, FRUS, NBII etc)
- FOIA reading rooms and FOIA’d documents
- CRS Reports
- https://archive-it.org/organizations/159
Creating FDLP reservoirs

Creating FDLP reservoirs (II)

Things everyone can/should do:

- Keep track of your favorite agency’s publications/data. Make sure those urls are in WayBack Machine!
- Submit fugitives to GPO’s lostdocs form (and lostdocs.freegovinfo.info!)
- Save documents to local Web servers and/or upload to the Internet Archive (“seeding the cloud”)
- Build Web-harvested collections that your local community wants/needs
- Join the "Everyday Electronic Materials" Zotero group and help us test out a newer, faster, more automatic fugitive document workflow!! (zotero.org/groups/everyday_electronic_materials)
- #FDLP IRC Channel to discuss fugitives and digital collection development

This is YOUR FDLP. Participate!!
“...let us save what remains: not by vaults and locks which fence them from the public eye and use in consigning them to the waste of time, but by such a multiplication of copies, as shall place them beyond the reach of accident.”

— Thomas Jefferson, February 18, 1791