

Susan Bokern

Vice President, ProQuest
Government & Historical Collections

GOVERNMENT CONTENT: PRESERVED AND REIMAGINED
APRIL 25, 2014



THE CENTER FOR RESEARCH LIBRARIES
GLOBAL RESOURCES COLLECTIONS FORUM

GOVERNMENT CONTENT: PRESERVED & REIMAGINED

DIFFERENT ROLES: WE CAN ALL WORK TOGETHER

- Different Roles we all play
 - Library of Congress
 - NARA
 - Government Agency Libraries
 - Government Printing Office
 - Intergovernmental Organizations
 - GOOGLE
 - Depository Libraries
 - Academic Libraries
 - National Libraries of countries around the world
 - Aggregators
 - Non Profits
- Aggregators: ProQuest
 - We have to create value or we don't have a role
 - We have to constantly keep updating our functionality and technology capabilities to meet emerging researcher needs and expectations

GOVERNMENT CONTENT: PRESERVED & REIMAGINED

GOVERNMENT'S ROLE

Role of government is to create and disseminate government information

Irony: while U.S. Government announces 'open data' initiatives, it also ceases creating content

- Agencies not producing reports any longer
- When the government takes down content/data from its web sites = it's gone
- U.S. Federal Government Sequestration = will this happen again?
- Some of the Statistical Reports Discontinued in 2013:

- **This Month in Futures Markets** Source: Commodity Futures Trading Commission
- **Forest Health Monitoring Reports** Source: Forest Service
- **U.S. Contributions to International Organizations: Report to the Congress** Source: Department of State
- **Financial News for Major Energy Companies** Source: Energy Information Administration (Department of Energy)
- **Long-Term Care Ombudsman Report [yr.] with Comparisons of National Data for [yr.]** Source: Department of Health and Human Services
- **Telephone Subscribership in the U.S.** (Approx. 55 p.) Source: Federal Communications Commission
- **Computers and Peripheral Equipment. Current Industrial Reports** Source: Bureau of Census



Libraries and Government Information
in the Age of Big Data



THE CENTER FOR RESEARCH LIBRARIES
GLOBAL RESOURCES COLLECTIONS FORUM

GOVERNMENT CONTENT: PRESERVED & REIMAGINED

GOVERNMENT'S ROLE

Each agency does its own thing

Cuts to Census Bureau, BLS, USDA, BEA

BLS discontinues:

- Measuring Green Jobs
- Mass Layoff Statistics
- International Labor Comparisons

USDA suspends:

- Agricultural Productivity in the U.S.
- County-Level Data Sets (“County Look-Up Tables”)
- Fertilizer Imports/Exports
Organic Handlers:
Procurement and Contracting
- Commodity Consumption by Population
Characteristics

NTIS ceases:

- World News Connection



USDA discontinues:

- Reports and datasets on dairy, cattle, vegetables, and 7 other commodities

Census Bureau delays:

- Planning for the 2020 decennial census
- Processing and release of 2012 Economic Census

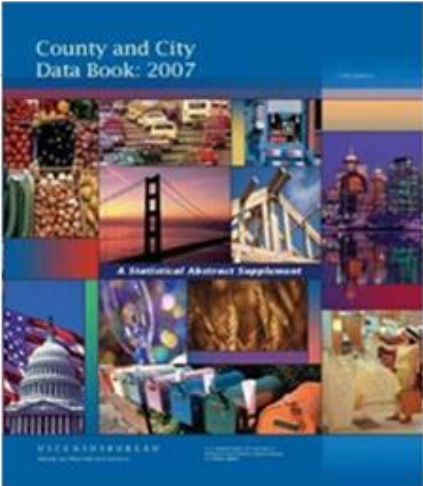
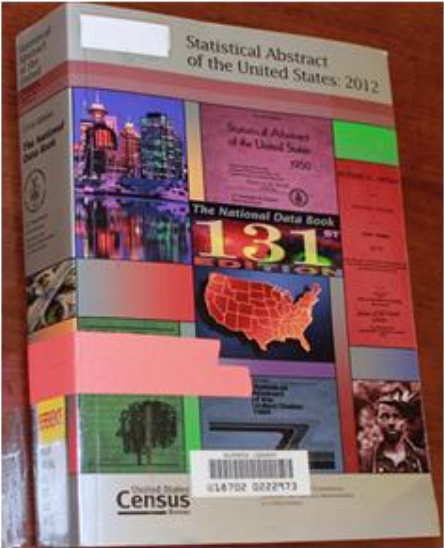
BEA curtails:

- Local Area Personal Income program
- Analysis of Foreign Direct Investment impacts

GOVERNMENT CONTENT: PRESERVED & REIMAGINED

GOVERNMENT'S ROLE

The Graveyard of Statistical Collections



USA Counties™



GOVERNMENT CONTENT: PRESERVED & REIMAGINED

GOVERNMENT'S ROLE

The Census Reform Act

113TH CONGRESS
1ST SESSION

H. R. 1638

To repeal the authority to conduct certain censuses, and for other purposes.

IN THE HOUSE OF REPRESENTATIVES

APRIL 18, 2013

Mr. DUNCAN of South Carolina (for himself, Mr. CHAFFETZ, Mr. HARRIS, Mr. JONES, Mr. PEARCE, Mr. RIBBLE, and Mr. SOUTHERLAND) introduced the following bill; which was referred to the Committee on Oversight and Government Reform, and in addition to the Committees on Agriculture and Appropriations, for a period to be subsequently determined by the Speaker, in each case for consideration of such provisions as fall within the jurisdiction of the committee concerned

A BILL

To repeal the authority to conduct certain censuses, and for other purposes.

1 *Be it enacted by the Senate and House of Representa-*
2 *tives of the United States of America in Congress assembled,*

3 **SECTION 1. SHORT TITLE.**

4 This Act may be cited as the "Census Reform Act
5 of 2013".

In the Crosshairs



THE CENTER FOR RESEARCH LIBRARIES
GLOBAL RESOURCES COLLECTIONS FORUM

GOVERNMENT CONTENT: PRESERVED & REIMAGINED

GOVERNMENT'S ROLE

Data Collection is Costly

- The 1960 census cost \$520 million (in 1990 dollars)
- The 1990 census cost \$2.6 billion.
- The 2010 census cost \$7.5 billion (in 1990 dollars)

A 1300% increase in 50 years!

GOVERNMENT CONTENT: PRESERVED & REIMAGINED

GOVERNMENT'S ROLE

Federal Staffing Levels Are Flat

Table 469. Governmental Employment and Payrolls: 1982 to 2010

[Employees in thousands (15,841 represents 15,841,000), payroll in millions of dollars (23,173 represents \$23,173,000,000). Data are for the month of October through 1992. Beginning with the 1997 survey, data are for the month of March. Covers both full-time and part-time employees. Local government data are estimates subject to sampling variation; see Appendix III and source.]

Type of government	1982	1987	1992	1997	2000	2005	2007	2008	2009	2010
EMPLOYEES (1,000)										
Total	15,841	17,212	18,745	19,540	20,876	21,725	22,116	22,462	22,668	22,607
Federal (civilian) ¹	2,848	3,091	3,047	2,807	2,899	2,720	2,730	2,769	2,824	3,066
State and local.....	12,993	14,121	15,698	16,733	17,976	19,004	19,386	19,693	19,844	19,599
Percent of total.....	82	82	84	86	86	87	88	88	88	87
State.....	3,744	4,116	4,595	4,733	4,877	5,078	5,200	5,270	5,346	5,326
Local.....	9,249	10,005	11,103	12,000	13,099	13,926	14,186	14,423	14,498	14,274
Counties.....	1,824	1,963	2,253	2,425	(NA)	(NA)	(NA)	(NA)	(NA)	(NA)
Municipalities.....	2,397	2,493	2,665	2,755	(NA)	(NA)	(NA)	(NA)	(NA)	(NA)
School districts.....	4,194	4,627	5,134	5,675	(NA)	(NA)	(NA)	(NA)	(NA)	(NA)
Townships.....	356	393	424	455	(NA)	(NA)	(NA)	(NA)	(NA)	(NA)
Special districts.....	478	529	627	691	(NA)	(NA)	(NA)	(NA)	(NA)	(NA)
PAYROLLS (mil. dol.)										
Total	23,173	32,669	43,120	49,156	58,166	71,599	78,583	83,268	85,281	86,643
Federal (civilian) ¹	5,959	7,924	9,937	9,744	11,485	13,475	14,427	15,472	15,106	16,238
State and local.....	17,214	24,745	33,183	39,412	46,681	58,123	64,156	67,796	70,175	70,404
Percent of total.....	74	76	77	80	80	81	82	81	82	81
State.....	5,022	7,263	9,828	11,413	13,279	16,062	17,789	18,726	19,425	19,579
Local.....	12,192	17,482	23,355	27,999	33,402	42,062	46,368	49,070	50,750	50,825
Counties.....	2,287	3,270	4,698	5,750	(NA)	(NA)	(NA)	(NA)	(NA)	(NA)
Municipalities.....	3,428	4,770	6,207	7,146	(NA)	(NA)	(NA)	(NA)	(NA)	(NA)
School districts.....	5,442	7,961	10,394	12,579	(NA)	(NA)	(NA)	(NA)	(NA)	(NA)
Townships.....	370	522	685	869	(NA)	(NA)	(NA)	(NA)	(NA)	(NA)
Special districts.....	665	959	1,370	1,654	(NA)	(NA)	(NA)	(NA)	(NA)	(NA)

NA Not available. ¹ Includes employees outside the United States.

Source: U.S. Census Bureau, Governments Division, "Annual Survey of Public Employment and Payroll Summary Report," January 2012.
<http://www.census.gov/govs/apes/>.



LEVIATHAN

Libraries and Government Information
in the Age of Big Data



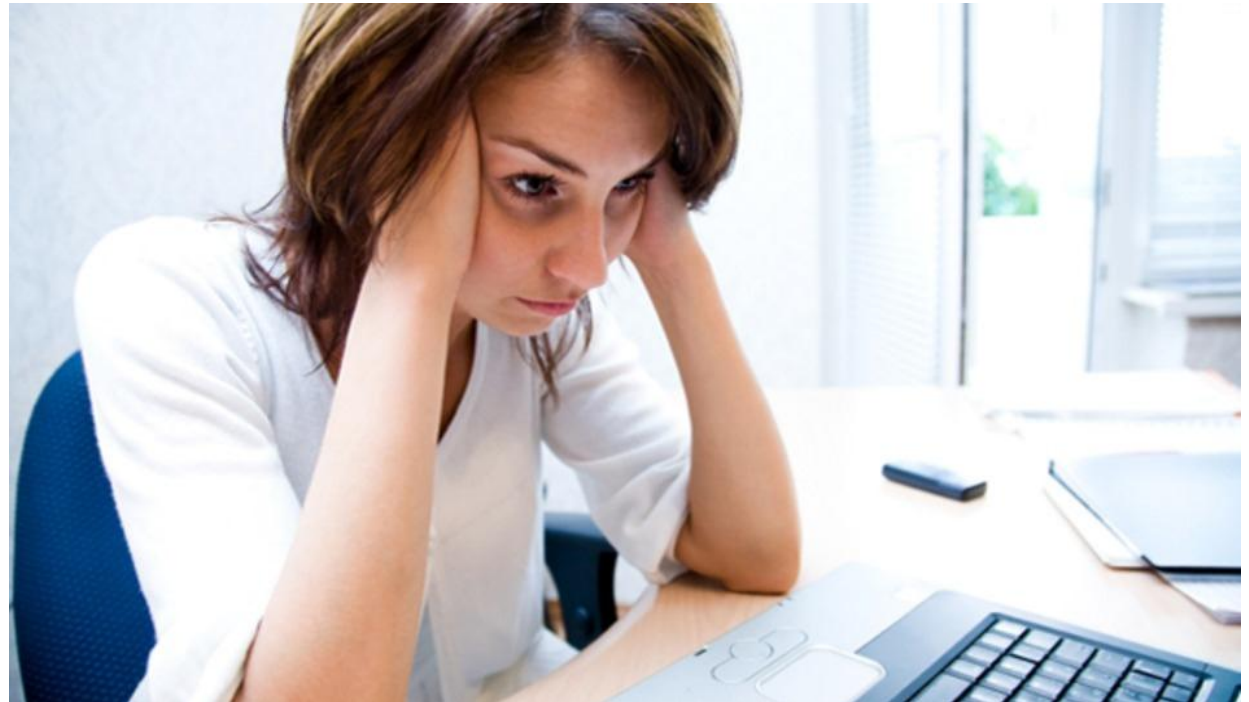
THE CENTER FOR RESEARCH LIBRARIES
GLOBAL RESOURCES COLLECTIONS FORUM

GOVERNMENT CONTENT: PRESERVED & REIMAGINED

PROQUEST'S ROLE: GOVERNMENT CONTENT IN CONTEXT OF RESEARCHER NEEDS

- **Our focus:** is on researchers and the research workflow processes – to help them find, gather, store, save, create, share, and output their research
- Additional value comes from creating the discovery, metadata, search interface, tools – to meet researcher needs

Finding
content,
making it
useful to
create better
research
outcomes



GOVERNMENT CONTENT: PRESERVED & REIMAGINED

PROQUEST'S ROLE

1. Process begins with market research to uncover research needs
 - Surveys, advisory boards, focus groups, discussions with librarians and faculty
2. Create an acquisitions strategy to develop a collection, preserve the content or data, and keep the technology up-to-date for users to access it over time
3. Have to know where to obtain the content – which source locations
 - i.e., Executive Branch Documents (1789-1932) – 100 source libraries
4. Establish Criteria...varies
 - comprehensive to include “all” content:
 - i.e., Congressional, 1789-current, Parliamentary House of Commons Papers (1715-current)
 - selective, yet includes complete series of government publications
 - i.e., Statistical - 10,000 titles, 1M Tables
 - all relevant data from government statistical and data offices worldwide
 - i.e., International Datasets
 - comprehensive index that links to government content
 - i.e., Government Publications Index, GPI; NTIS index



GOVERNMENT CONTENT: PRESERVED & REIMAGINED

PROQUEST'S ROLE: EXPERTISE & ADDING VALUE

100M+ Government Pages Archived and Accessible Online



Colonial State Papers



Documents on British Policy Overseas



ProQuest Legislative Insight



ProQuest Statistical Insight



ProQuest® Statistical Abstract of the United States



ProQuest Statistical Abstracts of the World



ProQuest International Datasets



ProQuest Congressional Basic



House of Commons Parliamentary Papers



ProQuest History Vault



ProQuest Historical Newspapers



ProQuest Congressional Hearings Digital Collection



ProQuest Congressional Research Digital Collection



NTIS



ProQuest Executive Branch Documents (1789-1932)



Digital National Security Archive

100+ Professional Staff Expertise in government content/data

Value is the expertise of the staff: many with 35 years of service in...

- content acquisitions - worldwide
- authoritative bibliographic records
- editorial subject knowledge in abstracting and indexing
- legislative history research = professionally researched histories
- statistical/data expertise, why we could continue to create the
 - Statistical Abstract of the U.S.

GOVERNMENT CONTENT: PRESERVED & REIMAGINED

NEW ROLES FOR PROQUEST: EXCITING!

IT'S NOT JUST ABOUT THE DOCUMENT ANYMORE!

It's about...

- enabling researchers to access and work with the materials according to the technology available at the most current time: now and into the future
- providing tools to researchers to improve their workflow processes in the content gathering, storing and sharing stages
- simplifying researcher workflows = saving them time
- helping researchers make connections they couldn't see before:
 - = VISUALIZATION
 - = ANALYTICS
 - = TEXT MINING, DATA MINING
- offering cloud-based services with new business models



GOVERNMENT CONTENT: PRESERVED & REIMAGINED

NEW ROLES FOR PROQUEST: ADDING VALUE TO CONGRESSIONAL CONTENT

1 **Vote Report- 2011 Senate Vote No. 161**
 Vote Date: October 12, 2011 Congress-Session: 112 -1
 Chamber: U.S. Senate Bill: 112 H.R. 3080
 Bill Title: United Statesy09Korea **Free Trade**
 Agreement Implementation Act
 The Senate passed H.R. 3080, the United States-**Korea Free Trade** Agreement Implementation Act.
[Citation/Abstract](#) [Permalink](#)

2 **Vote Report- 2011 House Roll No. 783**
 Vote Date: October 12, 2011 Congress-Session: 112 -1
 Chamber: U.S. House of Representatives Bill: 112 H.R. 3080
 Bill Title: United Statesy09Korea **Free Trade**
 Agreement Implementation Act
 The House passed H.R. 3080, the United States-**Korea Free Trade** Agreement Imple
[Citation/Abstract](#) [Permalink](#)

3 **Vote Report- 2011 House Roll No. 771**
 Vote Date: October 11, 2011 Congress-Session: 112 -1
 Chamber: U.S. House of Representatives Bill: 112 H.Res. 425
 Bill Title: PROVIDING FOR CONSIDERATION OF
 THE BILL (H.R. 3079) TO IMPLEMENT THE UNITED
 STATES-PANAMA **TRADE** PROMOTION AGREEMENT
 The House agreed to H. Res. 425, to provide for the consideration of the Senate amendment to H.R. 2832, to

● PAC Contributors
★ Individual Contributor
● Member

Member Profile— doc view

Bill Number: Retrieve Full Text Version
Bill Profile: Retrieve Bill Tracking Report
Permalink: [HTTP://testbed2.congressional.bowker.com/congressional/docview/159.d60.1121099007831?accountid=137327](http://testbed2.congressional.bowker.com/congressional/docview/159.d60.1121099007831?accountid=137327)

Sponsor

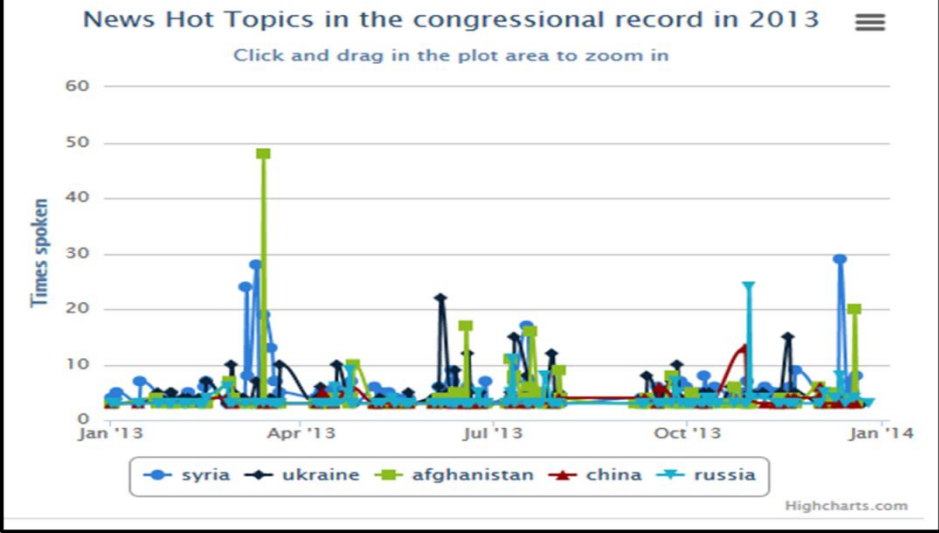
Bill Sponsor: Representative Eric Cantor (R -VA)

Vote Sponsor: Representative Sander M. Levin (D -MI)

Vote Result

Vote Outcome: The House passed H.R. 3080, the United States-**Korea Free Trade** Agreement Implementation Act.

Total Tally: AYE: 278 NAY: 151
Democrat Tally: AYE: 59 NAY: 130
Republican Tally: AYE: 219 NAY: 21
Independent Tally: AYE: 0 NAY: 0



GOVERNMENT CONTENT: PRESERVED & REIMAGINED

NEW ROLES FOR PROQUEST: SOLVING COMMON STATISTICAL AND DATA PROBLEMS



I need to know how many or how much.



I need multi-year trends, deep time series.



I want to compare two countries, states, companies, industries, etc.



I want to manipulate the data.



LEVIATHAN

Libraries and Government Information in the Age of Big Data



GOVERNMENT CONTENT: PRESERVED & REIMAGINED

NEW ROLES FOR PROQUEST: MAKING INTERNATIONAL GOVERNMENT DATA DISCOVERABLE, EASY-TO-FIND & USE

Data from governments around the world. Saves time in the researcher workflow process in finding the data.

ProQuest International Datasets
Search All Content | Compare Countries | Find Sub-National Data | **Data Providers** | Upload Data

POWERED BY DataMarket

Enter keywords to search for data. Example: China USA GDP

Explore data by subject

- Banking and finance (3,219)
- Criminal justice and law (1,600)
- Economic conditions (4,789)
- Education (1,897)
- Food and agriculture (5,687)
- Government (3,576)
- Health (2,389)
- Industry (6,778)
- International relations (4,520)
- Labor (3,245)
- Natural resources and environment (3,565)
- Population and income (5,250)
- Science and technology (1,420)
- Society and culture (1,390)

Featured datasets

People living with HIV/AIDS
Units: Number

Year	North Africa	South Africa	Sub-Saharan Africa (developing only)
1990	~1M	~1M	~2M
1995	~1.5M	~2M	~10M
2000	~2M	~3M	~18M
2005	~2.5M	~4M	~22M

Source: World Bank
[Go to data](#)

GOVERNMENT CONTENT: PRESERVED & REIMAGINED

NEW ROLES FOR PROQUEST: MAKING INTERNATIONAL GOVERNMENT DATA USABLE

Data from governments around the world.

Saves time in the researcher workflow process in finding the data.

Cost to start a business (% of income per capita)

Units: % of income per capita

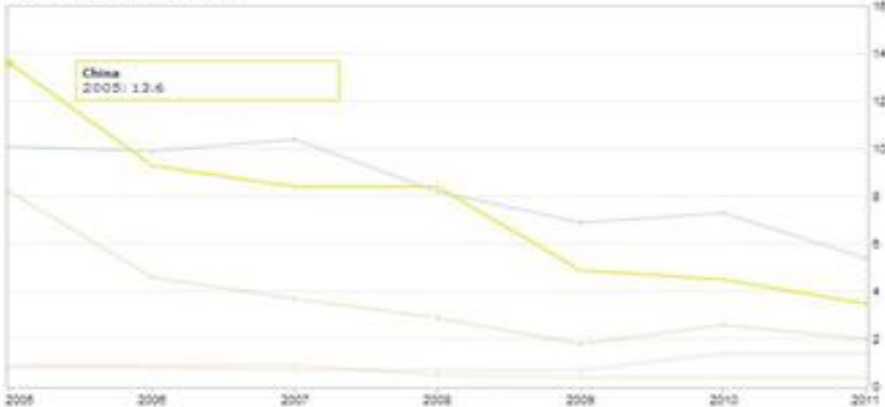
Year	Brazil	Canada	China	Russian Federation	United States
2005	10.1	0.9	13.6	8.2	0.8
2006	9.9	0.9	9.3	4.6	0.8
2007	10.4	0.9	8.4	3.7	0.7
2008	8.2	0.5	8.4	2.9	0.7
2009	6.9	0.4	4.9	1.8	0.7
2010	7.3	0.4	4.5	2.6	1.4
2011	5.4	0.4	3.5	2.0	1.4

Source: World Bank (citing: World Bank, Doing Business Project (<http://www.doin>

[Show detailed information](#)

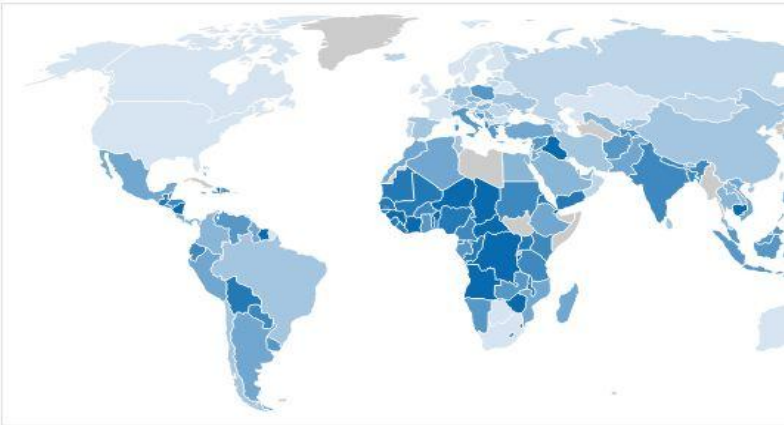
Cost to start a business (% of income per capita)

Units: % of income per capita



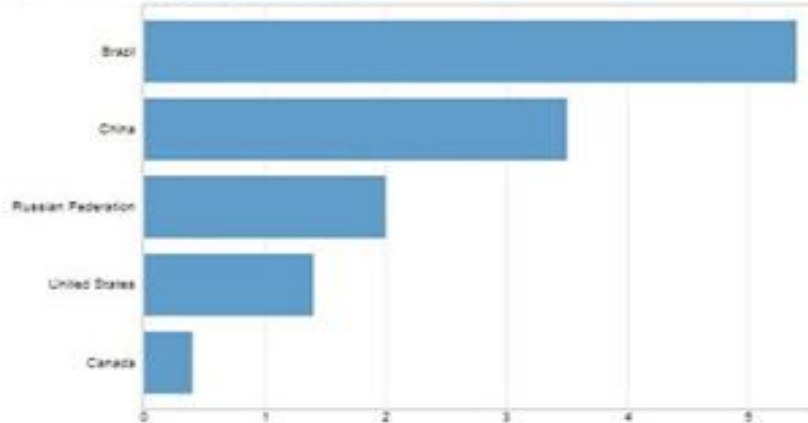
Cost to start a business (% of income per capita)

Units: % of income per capita Year: 2011



Cost to start a business (% of income per capita)

Units: % of income per capita Year: 2011



Providing a variety of display, analytical, and export options

GOVERNMENT CONTENT: PRESERVED & REIMAGINED

NEW ROLES FOR PROQUEST: MAKING INTERNATIONAL GOVERNMENT DATA DISCOVERABLE, EASY-TO-FIND & USE

Promoting the discovery of high-value content on hard-to-navigate government sites.

Examples:

- India parental preference for male vs female children, by sex and state, 2005-2006
- India Wages in Cotton Mills [monthly wage for lowest-paid operative, selected cities, from 2010]
- India women's mean age at effective marriage in major states, [by urban-rural residence, 2005-2011]



GOVERNMENT CONTENT: PRESERVED & REIMAGINED

NEW ROLES OF GOVERNMENT

Again, the irony: while it announces 'open data' initiatives, it ceases creating content

- Should they preserve all their content? Where?
 - Preservation is a new concept to them.
- Should they be creators of published reports with analysis?
- Should they create robust web sites?
- Or, should they just provide the raw data?
- Where and how should they store researcher data sent back to governments in 'open access models'?
- How is that content made available for additional research?



GOVERNMENT CONTENT: PRESERVED & REIMAGINED

NEW ROLES OF GOVERNMENT

Some challenges:

- Terms of obtaining content from governments worldwide is evolving –
 - some agencies require licenses and some international ones seek royalties
- Aggregator costs are increasing significantly - acquiring, digitizing, adding value
- Government entities need to adopt practices that encourage aggregators to obtain a reasonable ROI
- For print/microform content, they need to reconsider embargo periods before digitized content is put on their web sites for free, or aggregators will not be able to afford to digitize it.
- For 'born digital' content, they need to make the content/data more accessible.
- If we can't partner effectively and efficiently, the **outcome:**
 - less digitization, making fewer government documents available for research and teaching.
- Open to new business models.

GOVERNMENT CONTENT: PRESERVED & REIMAGINED

WHAT I THINK ABOUT?

- Government Open Access / Open Data
- Perceived Free Content
 - i.e., Google (\$400 Billion Market Cap = someone is paying)
- Curation = who is digitizing what content?
- Preservation = in what forms do we keep content? For the future?
- Migration of Print/Microform to Digital
 - Costs are enormous to change content into various digital formats
- New Technologies: new start ups, non-profits
- Libraries becoming publishers & mini-aggregators
- Disappearing Government Documents librarian positions = folding into 'Reference'
- Libraries & space considerations – trade offs
- Libraries & ROI's to justify spend



GOVERNMENT CONTENT: PRESERVED & REIMAGINED

NEW ROLES IN WORKING TOGETHER

New roles for us all!

Over time, the role of librarians in government aggregation and preservation will become clearer, as libraries

- determine which activities can best be done in-house
- determine which activities can be carried out most effectively and efficiently by aggregators

Let's build that bridge together...to solve the challenges in supporting research.



GOVERNMENT CONTENT: PRESERVED & REIMAGINED

NEW ROLES IN WORKING TOGETHER

- Thank You!
- Questions and comments
- susan.bokern@proquest.com
- 301-968-3820

Off the Beaten Path: Govdoc Frontiers

Leviathan:

Libraries and Government
Information in the Age of Big Data

Chicago • April 24-25, 2014

Robert E. Lee
Director of Strategic Partnerships
and Online Publishing
East View Information Services

