

Web Archival Survey of Global Data Portals

David Calano
CS 895: Web Archiving Forensics
Old Dominion University
November 28th, 2022

Who Cares About Statistical Data?

Data-gov wiki: Towards linking government data

L Ding, D DiFranzo, A Graves, JR Michaelis, X Li... - 2010 AAAI Spring 2010 - aaai.org
... In this paper, we discuss our efforts in developing the Data-gov Wiki - a website designed to host ... First, we provide an overview of the Data-gov Wiki. Second, we discuss strategies for ...
☆ Save 99 Cite Cited by 123 Related articles All 9 versions 88

[PDF] US government linked open data: semantic. data. gov

J Handler, J Holm, C Musialek... - IEEE Intelligent ..., 2012 - 222.252.30.203

Data. gov, the world's largest open government, datasharing website, and the use of linked

data in some of the site's community pages, innovation and startup activities outside the ...

☆ Save 99 Cite Cited by 148 Related articles All 7 versions Web of Science: 59 88

[HTML] Liberating data for public value: The case of Data. gov

R Krishnamurthy, Y Awazu - International Journal of Information ..., 2016 - Elsevier

Public agencies around the globe are liberating their data. Drawing on a case of Data.gov, we outline the challenges and opportunities that lie ahead for the liberation of public data. ...

☆ Save 99 Cite Cited by 55 Related articles All 5 versions Web of Science: 32

world bank data development indicators

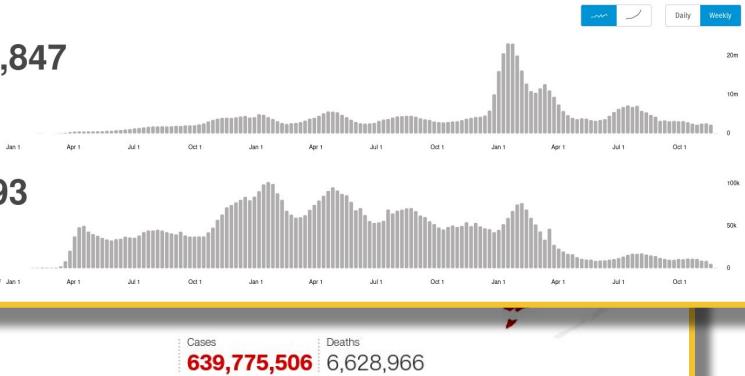
About 4,110,000 results (0.17 sec)

Globally, as of 5:54pm CET, 28 November 2022, there have been 637,404,847 confirmed cases of COVID-19, including 6,608,893 deaths, reported to WHO. As of 23 November 2022, a total of 12,959,275,260 vaccine doses have been administered.

Global Situation

637,404,847

confirmed cases



Reported cases and deaths

The figures below are based on data from the Johns Hopkins University Center for Systems Science and Engineering. These numbers are updated every 15 minutes but may differ from other sources due to differences in reporting times. For up-to-the-minute updates, follow our live coverage.

Location	Cases	...per 100K people	Deaths	...per 100K people
United States	98,571,212	30,030	1,079,207	329
India	44,673,078	3,269	530,614	39
France	36,577,224	54,544	155,089	231
Germany	36,373,165	43,753	157,495	189
Brazil	35,149,503	16,655	689,442	327
South Korea	26,959,843	52,138	30,413	59

<https://www.cnn.com/interactive/2020/health/coronavirus-maps-and-cases/>
<https://covid19.who.int/>

First Impressions?



Datamart Update Calendar Glossary API More

Search

Other UNSD Databases:



Popular statistical tables, country (area) and regional profiles	
Population	
• Population, surface area and density PDF CSV Updated: 18-Oct-2022	
• International migrants and refugees PDF CSV Updated: 18-Oct-2022	
• Population growth, fertility, life expectancy and mortality PDF CSV Updated: 18-Oct-2022	
• Population in the capital city, urban and rural areas PDF CSV	
• Afghanistan	
• Albania	
• Algeria	
• American Samoa	
• Andorra	
• Angola	
• Anguilla	
• Antigua and Barbuda	
• Argentina	
• Armenia	
• World	
• Africa	
• Northern Africa	
• Sub-Saharan Africa	
• Eastern Africa	
• Middle Africa	
• Southern Africa	
• Western Africa	
• Americas	
• Northern America	

<https://data.un.org/>

Vast and Infinite

32 databases - 60 million records

Other UNSD Databases:



Popular statistical tables, country (area) and regional profiles

Population

- Population, surface area and density
[PDF](#) | [CSV](#) Updated: 18-Oct-2022
- International migrants and refugees
[PDF](#) | [CSV](#) Updated: 18-Oct-2022
- Population growth, fertility, life expectancy and mortality
[PDF](#) | [CSV](#) Updated: 18-Oct-2022
- Population in the capital city, urban and rural areas
[PDF](#) | [CSV](#)

- Afghanistan
- Albania
- Algeria
- American Samoa
- Andorra
- Angola
- Anguilla
- Antigua and Barbuda
- Argentina
- Armenia
- Aruba

- World
- Africa
- Northern Africa
- Sub-Saharan Africa
- Eastern Africa
- Middle Africa
- Southern Africa
- Western Africa
- Americas
- Northern America

Latin America & the Caribbean



<https://data.un.org/>

Inherent Sparsity

Core indicators on access to, and use of, ICT by households and individuals, latest available data (2010-2014) Search glossaries

Source: World Telecommunication/ICT Indicators Data | International Telecommunications Union

Download Explore

Country or area	Proportion of households with					Percentage of individuals				
	(HH1)		(HH2)		(HH3f)	(HH6)		(HH5)		Y
	Radio	Year of data	TV	Year of data	Fixed line teleph	Year of data	Internet access at home	Year of data	Computer	Y
1 Afghanistan	62.2	1	2011		
2 Albania	3.7		2012		
3 Algeria				
4 Andorra				
5 Angola				
6 Antigua and Barbuda	62.2	1	2011	99.6		2011		
7 Argentina		55.5	31.8		2013		
8 Armenia	3.7	2012	98.5	2012	73.7	13.9		2013		
9 Australia				
10 Austria				
11 Azerbaijan	99.6	2011	100.0	2011	58.8	...				
12 Bahamas				
13 Bahrain	31.8	2013	99.2	2013	24.3	...				
14 Bangladesh	13.9	2013	46.2	2013	3.1	2013	87.7	2013	5.7	2013

a. Not all indicators were available for all countries, so caution should be used in cross-country comparisons.

<http://data.un.org/DocumentData.aspx?id=374>

Paradoxes of Collaboration

coordinate location

14°S, 53°W
Wikimedia maps | Map data © OpenStreetMap c...
0 references

start time

26 February 2020	edit
▶ 1 reference	
+ add value	

number of deaths

0	edit
point in time	17 March 2020
▶ 3 references	
1	edit
point in time	19 March 2020
▶ 1 reference	
34	edit
point in time	23 March 2020
▶ 2 references	

Compare selected revisions

- (cur | prev) 03:10, 28 November 2022 MsynBot (talk | contribs) ... (3,329,840 bytes) (+20) .. (Updated item: add badge sitelink to redirect (Q70893996) to arwiki sitelink; see [Wikidata:Sitelink check script](#)) (undo) (restore)
- (cur | prev) 00:29, 28 November 2022 CovidDatahubBot (talk | contribs) ... (3,329,820 bytes) (+1,109) .. (Added reference to claim: [number of cases \(P1603\)](#): 35,149,503, Adding sources.) (undo) (restore)
- (cur | prev) 00:29, 28 November 2022 CovidDatahubBot (talk | contribs) ... (3,328,711 bytes) (+451) .. (Added qualifier: [point in time \(P585\)](#): 26 November 2022, Adding date.) (undo) (restore)
- (cur | prev) 00:29, 28 November 2022 CovidDatahubBot (talk | contribs) ... (3,327,883 bytes) (-3) .. (Updated item: Change P1603 rank to normal) (undo) (restore)
- (cur | prev) 00:29, 28 November 2022 CovidDatahubBot (talk | contribs) ... (3,327,886 bytes) (+1,109) .. (Added reference to claim: [number of deaths \(P1120\)](#): 689,442, Adding sources.) (undo) (restore)
- (cur | prev) 00:28, 28 November 2022 CovidDatahubBot (talk | contribs) ... (3,326,777 bytes) (+451) .. (Added qualifier: [point in time \(P585\)](#): 26 November 2022, Adding date.) (undo) (restore)
- (cur | prev) 00:28, 28 November 2022 CovidDatahubBot (talk | contribs) ... (3,326,326 bytes) (+375) .. (Created claim: [number of deaths \(P1120\)](#): 689,442, Adding death count from template) (undo) (restore)
- (cur | prev) 00:28, 28 November 2022 CovidDatahubBot (talk | contribs) ... (3,325,951 bytes) (-3) .. (Updated item: Change P1120 rank to normal) (undo) (restore)
- (cur | prev) 00:29, 27 November 2022 CovidDatahubBot (talk | contribs) ... (3,325,954 bytes) (+1,109) .. (Added reference to claim: [number of cases \(P1603\)](#): 35,149,503, Adding sources.) (undo) (restore)
- (cur | prev) 00:28, 27 November 2022 CovidDatahubBot (talk | contribs) ... (3,325,954 bytes) (+1,109) .. (Added reference to claim: [number of deaths \(P1120\)](#): 689,442, Adding sources.) (undo) (restore)
- (cur | prev) 00:28, 27 November 2022 CovidDatahubBot (talk | contribs) ... (3,325,954 bytes) (+1,109) .. (Added reference to claim: [number of cases \(P1603\)](#): 35,149,503, Adding sources.) (undo) (restore)

343
point in time 3 April 2020
edit
▶ 1 reference
reference URL <https://coronavirus.jhu.edu/map.html>
+ add reference

241
point in time 2 April 2020
edit
▶ 1 reference

359
point in time 3 April 2020
edit
▶ 1 reference
reference URL <https://covid.saude.gov.br/>
+ add reference

<https://www.wikidata.org/wiki/Q86597695>

Data Archive != Archived Data

The screenshot illustrates a comparison between two versions of the UNdata Explorer page, highlighting the difference between a data archive and an archived data page.

Left Side (Nov 2009 Capture):

- The title bar indicates "421 captures" and the date range "26 Nov 2009 - 16 Nov 2022".
- The main content shows a sidebar with links to various statistics databases (Commodity Trade Statistics Database, Demographic Statistics Database, Energy Statistics Database, Environment Statistics Database) and a section for FAOSTAT (Food and Agriculture Organization). The FAOSTAT section includes categories like Crops, Fertilizers by Product, and Food Supply - Crops Primary Equivalent, each with "Preview" and "View data" buttons.
- A prominent red box highlights an error message: "Error: Http status = 0".

Right Side (Oct 2018 Capture):

- The title bar indicates "29 AUG 2018 2019" and the date range "2017 - 2019".
- The main content shows a message: "Hrm. The Wayback Machine has not archived that URL." Below this, a yellow box says "This page is available on the web!" and "Help make the Wayback Machine more complete! Save this url in the Wayback Machine".
- A red box highlights the message: "The Wayback Machine has not archived that URL."

<https://web.archive.org/web/20180829235347/http://data.un.org/Explorer.aspx>

A State of Suspended Animation

The screenshot shows the CIA Library homepage. At the top, there's a navigation bar with links for HOME, ABOUT CIA, CAREERS & INTERNSHIPS, OFFICES OF CIA, NEWS & INFORMATION, LIBRARY, and KIDS' ZONE. The main content area is titled "Library". It features a large image of an open book. On the left sidebar, there are sections for "Publications", "Center for the Study of Intelligence", "Freedom of Information Act Electronic Reading Room", "Kent Center Occasional Papers", "Intelligence Literature Reports", "Related Links", and "Video Center". The main content area includes sections for "Featured Links" (The World Factbook, World Leaders, Studies in Intelligence, Video Center), "Publications by Category" (Resources, International Relations, Intelligence History, Cold War), and "Additional Resources" (Center for the Study of Intelligence, Intelligence Literature Suggested Reading List, Reports, Freedom of Information Act Electronic Reading Room (FOIA), Kent Center Occasional Papers). A footer at the bottom indicates the page was posted on April 15, 2007, and last updated on May 01, 2019.

<https://web.archive.org/web/20190715001358/https://www.cia.gov/library/>

The screenshot shows a 404 Error page. The title "404 Error" is highlighted with a green box. The main message is "The page you requested doesn't exist." Below it, a sub-message says "CIA.gov has changed ...". The page explains that if the user arrived via a bookmark or favorite link, they should update it. It also suggests using the search form or links below to find the information sought. A list of related links is provided, including Contact Us, News & Stories, CIA Careers, The World Factbook, World Leaders, Center for the Study of Intelligence, FOIA, and Spy Kids. The page ends with a "RETURN TO HOMEPAGE" button and a curl command output in a terminal window.

```
curl -iL https://www.cia.gov/library/
HTTP/2 404
accept-ranges: bytes
content-type: text/html
etag: "e79e260b868a7a4f9d13bf1ab20b9524:1669213195.241892"
last-modified: Wed, 23 Nov 2022 14:19:55 GMT
x-akamai-transformed: 9 339616 0 pmb=mNONE,1
```

<https://cia.gov/library/>

<https://cia.gov/library/>

Persistent Consequences

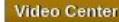
Library

The CIA releases millions of pages of documents each year and frequently releases public interest on this website. The Library contains a wealth of information, from current publications to basic references, reports and maps.

Featured Links

[The World Factbook](#)

[Studies in Intelligence](#)



Publications by Category

[Resources](#)

[International Relations](#)

[Intelligence History](#)



Loading...

<http://www.cia.gov/library/publications/resource/the-world-factbook> |
19:33:28 July 06, 2019

Got an HTTP 301 response at crawl time

Redirecting to...

<https://www.cia.gov/redirects/ciaredirect.html>

Home » redirects » CIA Site Redirect

CIA Site Redirect

CIA.gov has changed its Web address ...

CIA.gov is now encrypted, except for our Electronic Reading Room, to assure visitor confidentiality. As a result, the Web address for pages and documents in our site has changed from http:// to https://. In addition, CIA Careers has moved to a new location within the Web site

Please use the links or the search form below to find the information you seek.

- [CIA.gov home page](#)
- [Contact Us \(by postal mail, by fax, by phone, or via our e-mail form\)](#)
- [CIA Careers \(find a job and submit a resume\)](#)
- [The World Factbook \(our premier country reference\)](#)
- [World Leaders \(formerly listed as Chiefs of State\)](#)
- [Site Map \(links to major sections of the Web site\)](#)

Thank you for visiting CIA.gov.
The CIA Web Team

Error Code: 301 - Moved Permanently

Posted: May 09, 2007 06:55 AM

Last Updated: Nov 30, 2010 01:54 PM

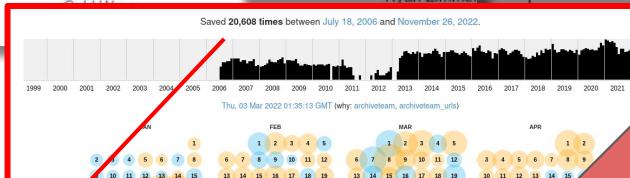
ir the Study of
ce

of
on Act
c Reading

ter
nal Papers
ce Literature

Links
inter

Saved 20,608 times between July 18, 2006 and November 26, 2022.



<https://web.archive.org/web/20190706174759/https://www.cia.gov/redirects/ciaredirect.html>



3rd-Party Adaptations

Contents

- Introduction**
- Geography**
- People and Society**
- Environment**
- Government**
- Economy**
- Energy**
- Communications**
- Transportation**
- Military and Security**
- Transnational Issues**

Introduction

Background

Globally, the 20th century was marked by: (a) two devastating World Wars; (b) the Great Depression of the 1930s; (c) the end of vast colonial empires; (d) rapid advances in science and technology, from the first airplane flight at Kitty Hawk, North Carolina (US) to the landing on the moon; (e) the Cold War between the Western alliance and the Warsaw Pact nations; (f) a sharp rise in living standards in North America, Europe, and Japan; (g) increased concerns about environmental degradation including deforestation, energy and water shortages, declining biological diversity, and air pollution; (h) the onset of the AIDS epidemic; and (i) the ultimate emergence of the US as the only world superpower. The planet's population continues to explode; from 1 billion in 1820 to 2 billion in 1930, 3 billion in 1960, 4 billion in 1974, 5 billion in 1987, 6 billion in 1999, and 7 billion in 2012. For the 21st century, the continued exponential growth in science and technology raises both hopes (e.g., advances in medicine and agriculture) and fears (e.g., development of even more lethal weapons of war).

Tip

Visit the Definitions and Notes page to view a description of each topic.

Definitions and Notes

Geography

<https://www.cia.gov/the-world-factbook/countries/world/>

Go to file Add file <> Code + About

World Factbook Country Profiles in JSON - Free Open Public Domain Data - No API Key Required :)

government, government, people, people, economy, economy, africa, africa, religion, religion, oceania, oceania, gis, gis, scraping, scraping, publicdomain, publicdomain, europe, europe, CC-0 license, CC-0 license, 869 stars, 869 stars, 343 contributors, 343 contributors

Readme

auto-update week 46

- africa auto-update week 46 10 days ago
- antarctica auto-update week 45 17 days ago
- australia-oceania auto-update week 46 10 days ago
- central-america-n-caribbean auto-update week 46 10 days ago
- central-europe auto-update week 46 10 days ago
- east-southeast-asia auto-update week 46 10 days ago
- europe auto-update week 46 10 days ago
- meta auto-update week 46 10 days ago
- middle-east auto-update week 46 10 days ago
- north-america auto-update week 46 10 days ago
- oceania auto-update week 46 10 days ago
- south-asia auto-update week 46 10 days ago
- south-america auto-update week 46 10 days ago
- world auto-update week 46 10 days ago

MANIFEST.md

LICENSE.md

MONGO.md

Up



Italy :: Europe

Introduction

Background:
Italy became a nation-state in 1861 when the regional states of the peninsula, along with Sardinia and Sicily, were united under King Victor EMMANUEL II. An era of parliamentary government came to a close in the early 1920s when Benito MUSSOLINI established a Fascist dictatorship. His alliance with Nazi Germany led to Italy's defeat in World War II. A democratic republic replaced the monarchy in 1946 and economic revival followed. Italy is a charter member of NATO and the European Economic Community (EEC) and its subsequent successors the EC and the EU. It has been at the forefront of European economic and political unification, joining the Economic and Monetary Union in 1999. Persistent problems include sluggish economic growth, high youth and female unemployment, organized crime, corruption, and economic disparities between southern Italy and the more prosperous north.

Geography

Location:
Southern Europe, a peninsula extending into the central Mediterranean Sea, northeast of Tunisia

Geographic coordinates:
42 50 N, 12 50 E

Map references:
Europe

Area:
total: 301,340 sq km
land: 294,140 sq km
water: 7,200 sq km
note: includes Sardinia and Sicily

Area - comparative:
almost twice the size of Georgia; slightly larger than Arizona

Land boundaries:
total: 1,836.4 km
border countries (6): Austria 404 km, France 476 km, Holy See (Vatican City) 3.4 km, San Marino 37 km, Slovenia 218 km, Switzerland 698 km

Coastline:
7,600 km

Maritime claims:
territorial sea: 12 nm
continental shelf: 200-m depth or to the depth of exploitation

Climate:

Contents

- 1. Introduction
- 1.1 Background
- 2. Geography
- 2.1 Location
- 2.2 Geographic coordinates
- 2.3 Map references
- 2.4 Area
- 2.5 Area - comparative
- 2.6 Land boundaries
- 2.7 Coastline
- 2.8 Maritime claims
- 2.9 Climate
- 2.10 Terrain
- 2.11 Elevation
- 2.12 Natural resources
- 2.13 Land use
- 2.14 Irrigated land
- 2.15 Population distribution
- 2.16 Natural hazards
- 2.17 Environment - current issues
- 2.18 Environment - international agreements
- 2.19 Geography - note
- 3. People and Society
- 3.1 Population
- 3.2 Nationality
- 3.3 Ethnic groups
- 3.4 Languages
- 3.5 Religions
- 3.6 Agro-industry
- 3.7 Dependency ratios
- 3.8 Median age
- 3.9 Population growth rate
- 3.10 Birth rate

<https://github.com/factbook/factbook>

<https://factbook.github.io/it>

Missing Archival Data

The screenshot shows a web interface for searching archival data. At the top, there's a date range selector showing 'JUN 12 2021' and a 'Select Date Range' button. Below it is a 'Station Search' bar with a placeholder 'Ex: Airport'. There are three view options: 'List View' (selected), 'Summary View', and 'Map View'. A status bar at the top indicates 'File Count/Limit: 10000/1000 Total File Count: 669196' and a 'Bulk Download' button.

Two data entries are listed:

- BUFFALO NIAGARA INTERNATIONAL, NY US (72528014733.csv)**
 - FILE SIZE: 9.33 MB
 - 104 DATA TYPES: Show Data Type
 - PERIOD OF RECORD: 2014-01-01 to 2014-12-31
 - STATION: BUFFALO NIAGARA INTERNATIONAL, NY US
 - Download + Select buttons (highlighted with a red box)
- DES MOINES INTERNATIONAL AIRPORT, IA US (72546014933.csv)**
 - FILE SIZE: 8.07 MB
 - 103 DATA TYPES: Show Data Type
 - PERIOD OF RECORD: 2019-01-01 to 2019-12-31
 - 1 STATION: DES MOINES INTERNATIONAL AIRPORT, IA US
 - Download + Select buttons (highlighted with a red box)

A large red arrow points from the 'Download' button of the Buffalo entry to a modal window. This modal has a title '72528014733.csv' and a message 'Failed - No file'.

<https://web.archive.org/web/20220812130920/https://www.ncei.noaa.gov/access/search/data-search/global-hourly>

Missing Live Data



Health Topics ▾

Countries ▾

Global Health Observatory

By category > Global influenza virological surveillance

Influenza virologic surveillance Data by country

External content

This content is provided from a WHO source other than the Global Health Observatory system. The link below will open a new

window containing the requested resource.

– External content



This page cannot be found

The page or file you are trying to access cannot be found. This is because the web address is incorrect or the file has been moved or deleted. In 2020, we migrated our web content to a new system so some older content may no longer be available online or at the same place.

Please try the following:

- Return to the [WHO home page](#).
- Check that the web address is correct.
- If you still encounter problems, please report these using the [feedback page](#).

Data analysis
By theme
By category
By indicator
Metadata
About the Observatory
Search

By category > Global influenza virological surveillance > Influenza virological surveillance data since 1995

<https://apps.who.int/gho/data/node.main.173?lang=en>

https://www.who.int/globalatlas/loginManagement/autoLogins/funnel_login.asp

Temporal Archive Issues

INTERNET ARCHIVE
WayBack Machine

<https://www.ncei.noaa.gov/data/gsom/>
26 captures
12 Dec 2016 - 1 Jun 2022

Index of /data/gsom

Name	Last modified	Size	Description
Parent Directory	-	-	
access/	2022-05-26 20:31	-	
archive/	2017-10-10 16:18	-	
doc/	2020-06-23 09:57	-	

<https://web.archive.org/web/20220601162542/https://www.ncei.noaa.gov/data/gsom/>

```
curl -s "http://web.archive.org/cdx/search/cdx?url=https://www.ncei.noaa.gov/data/gsom/archive/gsom&match=tar.gz"
```

https://web.archive.org/web/20170509114427/https://www.ncei.noaa.gov/data/gsom/archive/gsom-latest.tar.gz
https://web.archive.org/web/20170102232048/https://www.ncei.noaa.gov/data/gsom/archive/gsom_latest.tar.gz
https://web.archive.org/web/20170102232048/https://www.ncei.noaa.gov/data/gsom/archive/gsom_latest.tar.gz

INTERNET ARCHIVE
WayBack Machine

<https://www.ncei.noaa.gov/data/gsom/archive/>
9 captures
25 Dec 2016 - 15 Feb 2022

Index of /data/gsom/archive

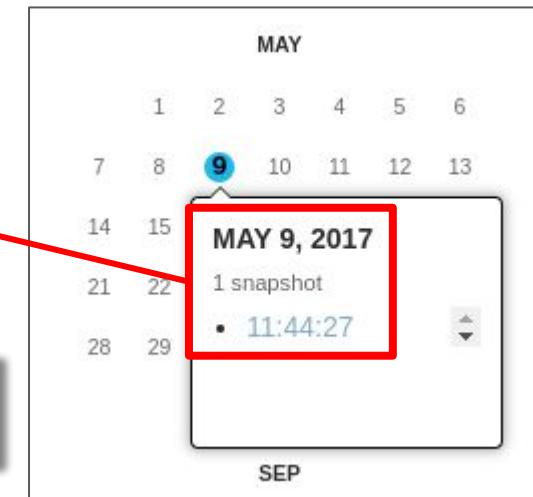
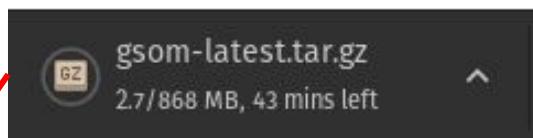
Name	Last modified	Size	Description
Parent Directory	-	-	
gsom-latest.tar.gz	2022-02-08 01:21	1.0G	

<https://web.archive.org/web/20220215030230/https://www.ncei.noaa.gov/data/gsom/archive/>

gsom-latest.tar.gz
gsom_latest.tar.gz

```
curl -s "http://web.archive.org/cdx/search/cdx?url=https://www.ncei.noaa.gov/data/gsom/archive/gsom&match=tar.gz"
```

https://web.archive.org/web/20170509114427/https://www.ncei.noaa.gov/data/gsom/archive/gsom-latest.tar.gz
https://web.archive.org/web/20170102232048/https://www.ncei.noaa.gov/data/gsom/archive/gsom_latest.tar.gz
https://web.archive.org/web/20170102232048/https://www.ncei.noaa.gov/data/gsom/archive/gsom_latest.tar.gz



<https://web.archive.org/web/20170509114427/https://www.ncei.noaa.gov/data/gsom/archive/gsom-latest.tar.gz>

Impractical Archives

The screenshot shows a Google search results page for 'dataset education'. The results are filtered by 'Topic' and 'Free'. The first result is a detailed dataset from CEICdata.com:

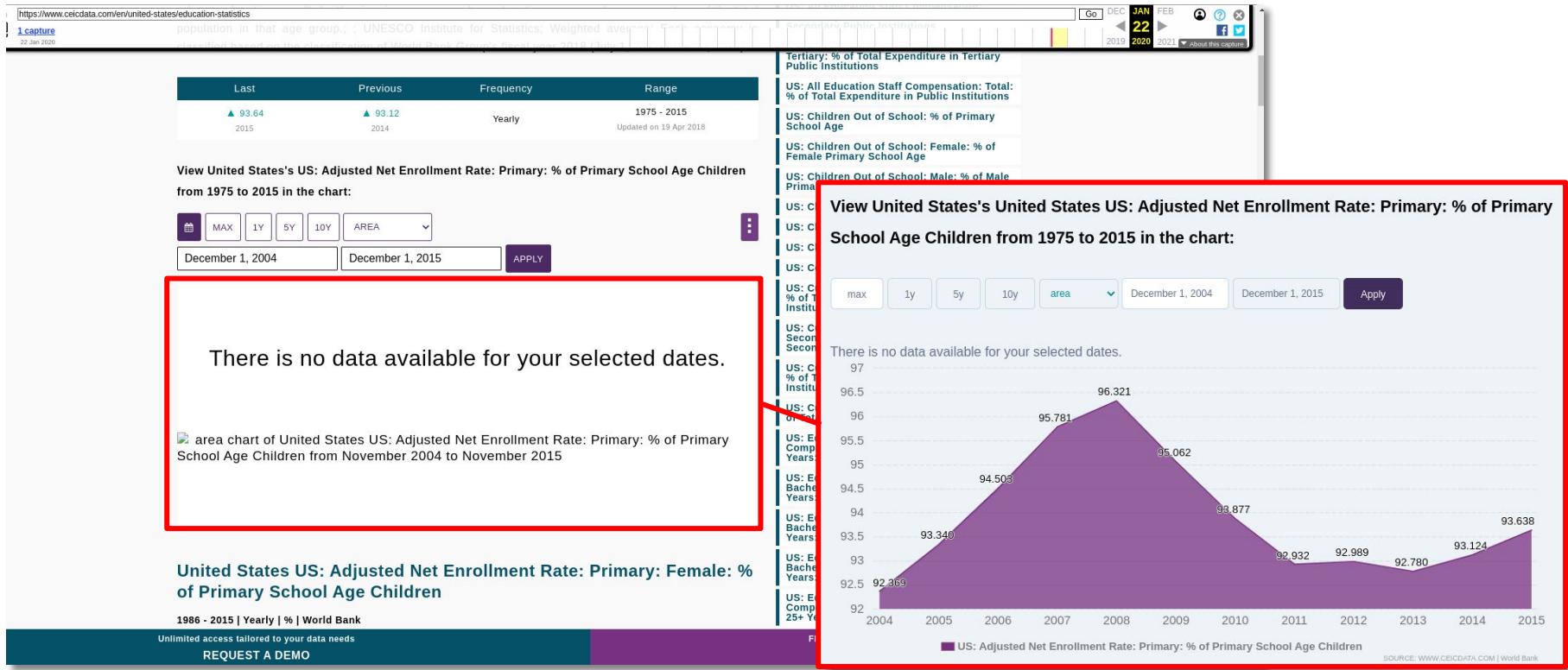
- United States US: Educational Attainment: Doctoral or Equivalent: Population 25+ Years: Male: % Cumulative**
- Explore at:** [United States | Economic In...](#)
- Dataset provided by:** CEICdata.com
- License:** Attribution 4.0 (CC BY 4.0)
- Time period covered:** Dec 1, 2013 - Dec 1, 2015
- Area covered:** United States
- Variables measured:** Education Statistics
- Description:** US: Educational Attainment: Doctoral or Equivalent: Population 25+ Years: Male: % Cumulative data was reported at 2.066 % in 2015. This records a decrease from the previous number of 2.316 % for 2014. US: Educational Attainment: Doctoral or Equivalent: Population 25+ Years: Male: % Cumulative data is updated yearly, averaging 2.207 % from Dec 2013 to 2015, with 3 observations. The data reached an all-time high of 2.316 % in 2014 and a record low of 2.066 % in 2015. US: Educational Attainment: Doctoral or Equivalent: Population 25+ Years: Male: % Cumulative data remains active status in CEIC and is reported by World Bank. The data is categorized under Global Database's USA – Table US.World Bank: Education Statistics. The percentage of population ages 25 and over that attained or completed Doctoral or equivalent.; UNESCO Institute for Statistics; ;

Other datasets listed include:

- United States US: School Enrollment: Tertiary: Male: %... ceicdata.com
- United States US: School Enrollment: Secondary: Private... ceicdata.com
- Laos LA: Preprimary Education: Duration ceicdata.com
- United States US: Educational Attainment: Doctoral or... ceicdata.com
- Israel IL: School Enrollment: Preprimary: Male: % Gross ceicdata.com
- United States US: All Education

<https://datasetsearch.research.google.com/search?src=0&query=education&docid=L2cvMTFyaHp0NXI3Xw%3D%3D>

Our Old Friend, JavaScript



<https://web.archive.org/web/20200122125450/https://www.ceicdata.com/en/united-states/education-statistics>

Federal Bureau of Inaccessibility

The screenshot shows the homepage of the Federal Bureau of Investigation Crime Data Explorer. The URL is <https://crime-data-explorer.fr.cloud.gov/pages/home>. The page features a large map of the United States with the state of Virginia highlighted. To the left, there's a sidebar with navigation links like 'Home', 'Crime Data', 'NIBRS Estimation Data', 'Hate Crime', etc., and a 'Resources' section with links to 'About the data', 'CDE General FAQ's', 'FBI UCR Program', etc. The main content area includes a map of the US with Virginia highlighted, a 'Trend of Violent Crime from 2011 to 2021' section with dropdown menus for 'From' (2011), 'To' (2021), and 'Crime Select' (All Violent C...), and a 'Resources' section with various links.

<https://crime-data-explorer.fr.cloud.gov/pages/explorer/crime/crime-trend>

<https://web.archive.org/web/20221123093558/https://crime-data-explorer.fr.cloud.gov/pages/explorer/crime/crime-trend>

Finding Workable Balances

This is a public dashboard showing how Federal agencies are performing on the Open Data Policy. [Learn more](#)

Current Milestone: The milestone selected is still in progress. The status of each field will be updated as frequently as possible, but won't be final until the milestone is complete.

Selected: Evidence Act Milestone 10 - December 31st 2021 ▾

CFO Act Agencies	Crawl Summary												
	Last Crawl	Last Modified	Public Datasets	Valid Metadata	Programs	Bureaus	Public Datasets	Restricted Datasets	Non-public Datasets	Datasets with downloads	Total Download URLs	Working Download URLs	
Department of Agriculture	15-Nov-2021 01:06:38 EST	29-Oct-2021 17:11:19 EDT	2261	100%	32	17	93.2%	6.6%	0.1%	94.4%	5004	67.6%	
Department of Commerce	02-Oct-2021 04:52:58 EDT	10-Mar-2021 10:50:31 EST	38960	100%	27	10	100.0%	0.0%	0.0%	100.0%	27170	5.9%	
Department of Defense	15-Nov-2021 02:35:26 EST	26-Sep-2018 13:19:43 EDT	177	100%	6	9	100%	0.0%	0.0%	1.0%	188	3.7%	
Department of Education	15-Nov-2021 02:19:03 EST	29-Sep-2021 12:00:07 EDT	619	100%	3	10	94.7%	0.3%	5.0%	85.8%	2606	79.2%	
Department of Energy	15-Nov-2021	09-Nov-2021	450	100%	12	5	58.2%	1.3%	40.4%	93.6%	718	39.0%	

Department of Education

<http://www.ed.gov>

Evidence Act Milestone 10 - December 31st 2021

Selected: Evidence Act Milestone 10 - December 31st 2021 ▾

Automated Metrics ⓘ

These metrics are generated by an automated analysis that runs every 24 hours until the end of the quarter at which point they become a historical snapshot

data.json	
● Expected Data.json URL	http://www.ed.gov/data.json (From USA.gov Directory)
● Resolved Data.json URL	https://www2.ed.gov/data.json
● Number of Redirects	
● HTTP Status	200
● Content Type	application/json
● Valid JSON	Valid
● Detected Data.json Schema	federal-v1.1
● Datasets with Valid Metadata	100% (627 of 627)
● Valid Schema	Valid
● Datasets	627
● Number of Collections	2
● Number of datasets not in a collection	625
● Datasets with Distribution URLs	86.0% (539 of 627)
● Datasets with Download URLs	59.3% (372 of 627)

<https://web.archive.org/web/20220610035834/https://data.gov/metrics.html>

<https://web.archive.org/web/20220804223950/https://dashboard.data.gov/offices/detail/49018/2021-12-31>

Limitations of Walled Gardens



THE WORLD BANK
IBRD • IDA

Microdata Library

Data Catalog Collections Citations Terms of use About

Home / Central Data Catalog / SIEF / MDG_2018_HDCT-ML_V01_M



Human Development Cash Transfer Program Impact Evaluation - Midline Household Survey 2018

Madagascar, 2018

GET MICRODATA

Reference ID MDG_2018_HDCT-ML_v01_M

Producer(s) World Bank

To request access to licensed datasets, please login to the website to continue.

Log in

Email Address

Password

Login **Register** **Forgot password**

Information you provide by registering or login to our catalog will be used in compliance with the terms of the World Bank's Privacy Policy.

(SIEF)

JSON

RELATED DATASETS

DONATE

INTERNET ARCHIVE
Wayback Machine

CREATED ON
Nov 23, 2022

LAST MODIFIED
Nov 23, 2022

PAGE VIEWS
79

DOWNLOADS
1

Hrm.

The Wayback Machine has not archived that URL.

archive.today
webpage capture

<https://microdata.worldbank.org/index.php/catalog/4776> search

search examples:
• microdata.worldbank.org for all snapshots from the host
• *worldbank.org for list of subdomains
• https://microdata.worldbank.org/index.php/catalog/4776 for exact url
• https://microdata.worldbank.org/index.php/catalog/4776 for url prefix

No results

You may want to

- archive this url
- search this url in Google Cache
- search this url in Archive.Org
- search this url in WebCite
- search this url in Megalodon

<https://microdata.worldbank.org/index.php/catalog/4776>

Conclusions

- Disparate and dynamic data requires integrative tooling
- Need for robust systems to limit the downsides of siloed data
- Automation of data gathering and preservation
- Finding balance between accessibility and usability
- Temporal importance of live and archived data
- Beware of long-lasting design consequences

Geographic area	(HH4)	Year of data
	Computer	
Afghanistan	...	
Albania	...	
Algeria	...	
Andorra	...	
Angola	...	
Antigua and Barbuda	...	
Argentina	51.0	3 2011
Armenia	49.2	2013
Australia	82.6	2011
Austria	81.0	2014
Azerbaijan	60.3	2014
Bahamas	...	
Bahrain	94.6	2014
Bangladesh	5.7	2013