

On December 7-8, 2007, thirty open government advocates gathered in Sebastopol, California and wrote a set of **eight principles of open government data**.

This page annotates the original 8 principles (https://public.resource.org/8_principles.html) and links to additional principles found around the web.

Larry Lessig on Open Government Data Principles



David Orban interviews Larry Lessig at the conclusion of the workshop.

Further Reading

There are many definitions of “open” and this is but one. The 2007 working group’s definition sits at the unique intersection of open government and open data and has United States sensibilities.

For a broader notion of open data, see the Open Definition (2005) (<http://opendefinition.org/>).

See the resources at the right, and continue reading below for annotated principles of open government data and other principles found around the web.

- 2005 The Open Definition (<http://opendefinition.org/>)
- American Association of Law Libraries: Public Information on Government Websites (<http://www.aallnet.org/main-menu/Advocacy/access/aellcorevalues.pdf>)
- 2007 Association of Computing Machinery Recommendation on Open Government (<http://www.acm.org/public-policy/open-government>)
- 2009 Open Government Data: The Book (<http://opengovdata.io/>), by Joshua Tauberer
- Sunlight Foundation Open Data Policy Guidelines (<http://sunlightfoundation.com/opendataguidelines/>) (updated in 2014!)
- White House Memorandum on Open Data Policy (M-13-13) (<http://www.whitehouse.gov/sites/default/files/omb/memoranda/2013/m-13-13.pdf>)
- 2013 Best-Practices Language for Making Data “License-Free” (<http://theunitedstates.io/licensing/>)
- living* OKF Working Group on Open Government Data (<http://opengovernmentdata.org/>)
- living* U.S. State/Local Open Data Policies (<http://sunlightfoundation.com/policy/opendatamap/>) mapped by the Sunlight Foundation

The 8 Principles of Open Government Data

The following is from the 8 principles (https://public.resource.org/8_principles.html) and the group’s wiki work following their meeting. New annotations are in white boxes.

Government data shall be considered open if it is made public in a way that complies with the principles below:

1. Complete

All public data is made available. Public data is data that is not subject to valid privacy, security or privilege limitations.

While non-electronic information resources, such as physical artifacts, are not subject to the Open Government Data principles, it is always encouraged that such resources be made available electronically to the extent feasible.

“Bulk data” means that an entire dataset can be acquired. Even the simplest of applications, such as computing the sum of line items, requires access to the *entire* dataset. This principle also implies that bulk data should be made available *before* “APIs” are created because APIs typically only return small slices of the whole data.

7 Additional Principles

Here are some additional principles of open data that the working group did not consider but might have:

• Online & Free

Information is not meaningfully public if it is not available on the Internet at no charge, or at least no more than the marginal cost of reproduction. It should also be findable.

This principle appears in...

Open Definition (2005) (<http://opendefinition.org/>) (“Access”)
Sunlight Foundation’s Principles for Transparency in Government (February 2009)
Sunlight Foundation Open Data Policy Guidelines (2012) (<http://sunlightfoundation.com/opendataguidelines/>) (“Require Public Information To Be Posted Online”, “Create A Public, Comprehensive List Of All Information Holdings”)

• Permanent

This principle also appears in...

[Sunlight Foundation Open Data Policy Guidelines \(2012\)](http://sunlightfoundation.com/opendataguidelines/)
(<http://sunlightfoundation.com/opendataguidelines/>) ("Publish Bulk Data")

2. Primary

Data is as collected at the source, with the highest possible level of granularity, not in aggregate or modified forms.

If an entity chooses to transform data by aggregation or transcoding for use on an Internet site built for end users, it still has an obligation to make the full-resolution information available in bulk for others to build their own sites with and to preserve the data for posterity.

This principle also appears in...

[White House M-13-13 \(2013\)](http://www.whitehouse.gov/sites/default/files/omb/memoranda/2013/m-13-13.pdf)
(<http://www.whitehouse.gov/sites/default/files/omb/memoranda/2013/m-13-13.pdf>) ("Complete")

3. Timely

Data is made available as quickly as necessary to preserve the value of the data.

This principle also appears in...

[White House M-13-13 \(2013\)](http://www.whitehouse.gov/sites/default/files/omb/memoranda/2013/m-13-13.pdf)
(<http://www.whitehouse.gov/sites/default/files/omb/memoranda/2013/m-13-13.pdf>)

4. Accessible

Data is available to the widest range of users for the widest range of purposes.

Data must be made available on the Internet so as to accommodate the widest practical range of users and uses. This means considering how choices in data preparation and publication affect access to the disabled and how it may impact users of a variety of software and hardware platforms. Data must be published with current industry standard protocols and formats, as well as alternative protocols and formats when industry standards impose burdens on wide reuse of the data.

Data is not accessible if it can be retrieved only through navigating web forms, or if automated tools are not permitted to access it because of a robots.txt file, other policy, or technological restrictions.

This principle also appears in...

[Open Definition \(2005\)](http://opendefinition.org/) (<http://opendefinition.org/>) ("Access", "Absence of Technological Restriction")
[White House M-13-13 \(2013\)](http://www.whitehouse.gov/sites/default/files/omb/memoranda/2013/m-13-13.pdf)
(<http://www.whitehouse.gov/sites/default/files/omb/memoranda/2013/m-13-13.pdf>) ("Accessible")

5. Machine processable

Data is reasonably structured to allow automated processing.

The ability for data to be widely used requires that the data be properly encoded. Free-form text is not a substitute for tabular and normalized records. Images of text are not a substitute for the text itself. Sufficient documentation on the data format and meanings of normalized data items must be available to users of the data.

The Association of Computing Machinery's [Recommendation on Open Government](http://www.acm.org/public-policy/open-government) (<http://www.acm.org/public-policy/open-government>) (February 2009) stated this principle another way: "Data published by the government should be in formats and approaches that promote analysis and reuse of that

Data should be made available at a stable Internet location indefinitely and in a stable data format for as long as possible.

This principle appears in...

[AALL: Public Information on Government Websites \(2007\)](http://www.aallnet.org/main-menu/Advocacy/access/aelliccorevalues.pdf)
(<http://www.aallnet.org/main-menu/Advocacy/access/aelliccorevalues.pdf>)
[Sunlight Foundation Open Data Policy Guidelines \(2012\)](http://sunlightfoundation.com/opendataguidelines/)
(<http://sunlightfoundation.com/opendataguidelines/>) ("Create Permanent, Lasting Access To Data")

• Trusted

The Association of Computing Machinery's [Recommendation on Open Government](http://www.acm.org/public-policy/open-government) (<http://www.acm.org/public-policy/open-government>) (February 2009) stated, "Published content should be digitally signed or include attestation of publication/creation date, authenticity, and integrity." Digital signatures help the public validate the source of the data they find so that they can trust that the data has not been modified since it was published. Since provenance is for originally-published documents, it is *not* a reason to prevent the public from modifying government documents.

This principle appears in...

[ACM Recommendation on Open Government \(2009\)](http://www.acm.org/public-policy/open-government)
(<http://www.acm.org/public-policy/open-government>)

• A Presumption of Openness

The presumption of openness rests on laws like the Freedom of Information Act, procedures including records management, and tools such as data catalogs.

Sunlight Foundation's [Open Data Policy Guidelines](http://sunlightfoundation.com/opendataguidelines/) (<http://sunlightfoundation.com/opendataguidelines/>) state, "Setting the default to open means that the government and parties acting on its behalf will make public information available proactively and that they'll put that information within reach of the public (online), with low to no barriers for its reuse and consumption. . . . Setting the default to open is about living up to the potential of our information, about looking at comprehensive information management, and making determinations that fall in the public interest."

This principle appears in...

[Sunlight Foundation Open Data Policy Guidelines \(2012\)](http://sunlightfoundation.com/opendataguidelines/)
(<http://sunlightfoundation.com/opendataguidelines/>) ("Set the Default to Open", "Create A Portal Or Website Devoted To Data Publication Or Policy", "Create Binding Regulations Or Guidance For Implementation", "Create New Legal Rights Or Other Mechanisms")
[White House M-13-13 \(2013\)](http://www.whitehouse.gov/sites/default/files/omb/memoranda/2013/m-13-13.pdf)
(<http://www.whitehouse.gov/sites/default/files/omb/memoranda/2013/m-13-13.pdf>) ("Public")

• Documented

Documentation about the format and meaning of data goes a long way to making the data useful.

The American Association of Law Libraries's [Principles & Core Values Concerning Public Information on Government Websites](http://www.aallnet.org/main-menu/Advocacy/access/aelliccorevalues.pdf) (<http://www.aallnet.org/main-menu/Advocacy/access/aelliccorevalues.pdf>) (March 24, 2007) noted that it is as important for users to know the data is current as for the data itself to be current. Their principles state, "Government websites must provide users with sufficient information to make assessments about the accuracy and currency of legal information published on the website."

This principle appears in...

[AALL: Public Information on Government Websites \(2007\)](http://www.aallnet.org/main-menu/Advocacy/access/aelliccorevalues.pdf)
(<http://www.aallnet.org/main-menu/Advocacy/access/aelliccorevalues.pdf>)
[Sunlight Foundation Open Data Policy Guidelines \(2012\)](http://sunlightfoundation.com/opendataguidelines/)
(<http://sunlightfoundation.com/opendataguidelines/>) ("Require Publishing

data.” The most critical value of open government data comes from the public’s ability to carry out its own analyses of raw data, rather than relying on a government’s own analysis.

As part of this, the use of unique, numeric identifiers for entities mentioned in the data can help connect the data to other relevant information.

This principle also appears in...

[ACM Recommendation on Open Government \(2009\)](http://www.acm.org/public-policy/open-government)
(<http://www.acm.org/public-policy/open-government>)

[Sunlight Foundation Open Data Policy Guidelines \(2012\)](http://sunlightfoundation.com/opendataguidelines/)
(<http://sunlightfoundation.com/opendataguidelines/>) (“Mandate The Use Of Unique Identifiers”)

[White House M-13-13 \(2013\)](http://www.whitehouse.gov/sites/default/files/omb/memoranda/2013/m-13-13.pdf)
(<http://www.whitehouse.gov/sites/default/files/omb/memoranda/2013/m-13-13.pdf>) (“Accessible”)

Metadata Or Other Documentation”)

[White House M-13-13 \(2013\)](http://www.whitehouse.gov/sites/default/files/omb/memoranda/2013/m-13-13.pdf)

(<http://www.whitehouse.gov/sites/default/files/omb/memoranda/2013/m-13-13.pdf>) (“Described”)

- **Safe to Open**

The Association of Computing Machinery’s [Recommendation on Open Government](http://www.acm.org/public-policy/open-government) (<http://www.acm.org/public-policy/open-government>) (February 2009) stated, “Government bodies publishing data online should always seek to publish using data formats that do not include executable content.” Executable content within documents poses a security risk to users of the data because the executable content may be malware (viruses, worms, etc.).

This principle appears in...

[ACM Recommendation on Open Government \(2009\)](http://www.acm.org/public-policy/open-government)
(<http://www.acm.org/public-policy/open-government>)

6. Non-discriminatory

Data is available to anyone, with no requirement of registration.

Anonymous access to the data must be allowed for public data, including access through anonymous proxies. Data should not be hidden behind “walled gardens.”

This principle also appears in...

[Open Definition \(2005\)](http://opendefinition.org/) (<http://opendefinition.org/>) (“No Discrimination Against Persons or Groups”, “No Discrimination Against Fields of Endeavor”)

[Sunlight Foundation Open Data Policy Guidelines \(2012\)](http://sunlightfoundation.com/opendataguidelines/)
(<http://sunlightfoundation.com/opendataguidelines/>) (“Remove Restrictions For Accessing Information”)

[White House M-13-13 \(2013\)](http://www.whitehouse.gov/sites/default/files/omb/memoranda/2013/m-13-13.pdf)
(<http://www.whitehouse.gov/sites/default/files/omb/memoranda/2013/m-13-13.pdf>) (“Accessible”)

- **Designed with Public Input**

The public is in the best position to determine what information technologies will be best suited for the applications the public intends to create for itself. Public input is therefore crucial to disseminating information in such a way that it has value.

This principle appears in...

[Association of Government Accountants’ Recovery and the Transparency Initiative \(Annual CFO Survey\) \(2009\)](http://www.granthornton.com/staticfiles/GTCom/Public%20sector/CFO%20survey/Annual%20CFO%20Survey.pdf)
(<http://www.granthornton.com/staticfiles/GTCom/Public%20sector/CFO%20survey/Annual%20CFO%20Survey.pdf>)
[Sunlight Foundation Open Data Policy Guidelines \(2012\)](http://sunlightfoundation.com/opendataguidelines/)
(<http://sunlightfoundation.com/opendataguidelines/>) (“Build On The Values, Goals, And Mission Of The Community And Government”)

7. Non-proprietary

Data is available in a format over which no entity has exclusive control.

Proprietary formats add unnecessary restrictions over who can use the data, how it can be used and shared, and whether the data will be usable in the future. While some proprietary formats are nearly ubiquitous, it is nevertheless not acceptable to use only proprietary formats. Likewise, the relevant non-proprietary formats may not reach a wide audience. In these cases, it may be necessary to make the data available in multiple formats.

This principle also appears in...

[Sunlight Foundation Open Data Policy Guidelines \(2012\)](http://sunlightfoundation.com/opendataguidelines/)
(<http://sunlightfoundation.com/opendataguidelines/>) (“Mandate Open Formats”)

[White House M-13-13 \(2013\)](http://www.whitehouse.gov/sites/default/files/omb/memoranda/2013/m-13-13.pdf)
(<http://www.whitehouse.gov/sites/default/files/omb/memoranda/2013/m-13-13.pdf>) (“Accessible”)

8. License-free

Data is not subject to any copyright, patent, trademark or trade secret regulation. Reasonable privacy, security and privilege restrictions may be allowed.

Because government information is a mix of public records, personal information, copyrighted work, and other non-open data, it is important to be clear about what data is available and what licensing, terms of service, and legal restrictions apply. Data for which no restrictions apply should be marked clearly as being in the public domain.

Requiring attribution to the government, even though attribution might be reasonable in other contexts, would constitute a major policy shift in the United States with significant legal implications for the press. The Creative Commons CC0 public domain dedication can make a work license-free.

This principle also appears in...

[Sunlight Foundation Open Data Policy Guidelines \(2012\)](#)
(<http://sunlightfoundation.com/opendataguidelines/>) ("Remove Restrictions On Reuse Of Information")

[Best-Practices Language for Making Data "License-Free" \(2013\)](#)
(<http://theunitedstates.io/licensing/>)

in weaker form: [Open Definition \(2005\)](#) (<http://opendefinition.org/>)
("Redistribution", "Reuse")

in weaker form: [White House M-13-13 \(2013\)](#)
(<http://www.whitehouse.gov/sites/default/files/omb/memoranda/2013/m-13-13.pdf>) ("Reusable")

Compliance must be reviewable.

This principle also appears in...

[Association of Government Accountants' Recovery and the Transparency Initiative \(Annual CFO Survey\) \(2009\)](#)
(<http://www.grantthornton.com/staticfiles/GTCom/Public%20sector/CFO%20survey/2009%20AGA%20Survey.pdf>)

[Sunlight Foundation Open Data Policy Guidelines \(2012\)](#)
(<http://sunlightfoundation.com/opendataguidelines/>) ("Create Or Appoint Oversight Authority")

[White House M-13-13 \(2013\)](#)
(<http://www.whitehouse.gov/sites/default/files/omb/memoranda/2013/m-13-13.pdf>) ("Managed Post-Release")

Definitions

"public" means:

The Open Government Data principles do not address what data should be public and open. Privacy, security, and other concerns may legally (and rightly) prevent data sets from being shared with the public. Rather, these principles specify the conditions public data should meet to be considered "open."

"data" means:

Electronically stored information or recordings. Examples include documents, databases of contracts, transcripts of hearings, and audio/visual recordings of events.

While non-electronic information resources, such as physical artifacts, are not subject to the Open Government Data principles, it is always encouraged that such resources be made available electronically to the extent feasible.

"reviewable" means:

A contact person must be designated to respond to people trying to use the data.

A contact person must be designated to respond to complaints about violations of the principles.

An administrative or judicial court must have the jurisdiction to review whether the agency has applied these principles appropriately.

About the 2007 Workshop

Participants: Carl Malamud (Public.Resource.Org), Tim O'Reilly (O'Reilly Media), Greg Elin (Sunlight Foundation), Micah Sifry (Sunlight Foundation), Adrian Holovaty (EveryBlock), Daniel X. O'Neil (EveryBlock), Michal Migurski (Stamen Design), Shawn Allen (Stamen Design), Josh Tauberer (GovTrack.us), Lawrence Lessig (Stanford), Dan Newman (MapLight.Org), John Geraci (outside.in), Edwin Bender (Inst. for Money), Tom Steinberg (My Society), David Moore (Participatory Politics), Donny Shaw (Participatory Politics), JL Needham (Google), Joel Hardi (Public.Resource.Org), Ethan Zuckerman (Berkman), Greg Palmer (NewCo), Jamie Taylor (MetaWeb), Bradley Horowitz (Yahoo), Zack Exley (New Organizing Institute), Karl Fogel (Question Copyright), Michael Dale (Metavid), Joseph Lorenzo Hall (UC Berkeley), Marcia Hofmann (EFF), David Orban (Metasocial Web), Will Fitzpatrick (Omidyar Network), Aaron Swartz (Open Library).

The meeting was coordinated by Tim O'Reilly of O'Reilly Media and Carl Malamud of Public.Resource.Org, with sponsorship from the Sunlight Foundation, Google, and Yahoo.