

❗ We need your help!

This document is being developed collaboratively. We hope to make it the best possible resource by leveraging the collective knowledge of government experts and other stakeholders. You can contribute by:

- Editing this document through github (<https://github.com/centerforgov/open-data-portal-requirements/blob/master/README.md>) or prose.io (<http://prose.io/#centerforgov/open-data-portal-requirements/edit/master/README.md>) (github account required)
- Creating a github issue (<https://github.com/centerforgov/open-data-portal-requirements/issues>) (github account required)
- Adding your comments to the discussion at the bottom of this page

Open Data Portal Requirements

Introduction

This document contains a set of sample requirements to help governments evaluate, develop (or procure), deploy, and launch an open data web site (portal).

Terms

The following terms are used throughout the requirements to indicate importance.

- **Must** denotes a requirement that is critical to the successful implementation of the open data portal. If this requirement is not met, it will significantly impair the use or management of the site.
- **May** denotes a requirement that is optional, but highly desirable for success.

Roles

The following roles are used throughout this document to approximate to whom each requirement is relevant. They are not intended to reflect a permission system, which might be implemented by a technology solution.

- A **consumer** is anyone who visits the portal to find, and access, and use data.
- Typically, a **publisher** is a government employee who is responsible for publishing or updating data and other content on the portal.
- Typically, an **administrator** is a government employee who manages the technical and configuration options of the portal, such as visual themes, metadata requirements, etc.

Generally, a publisher can do anything a consumer can do, and an administrator can do anything a publisher can do.

Data Catalog Features

A data catalog is a listing of available data, with each catalog entry corresponding to one or more data resources. It usually includes high-level information for each dataset, such as Title, Description, Release Date, Category(-ies), and Keyword(s).

Consumers

- must be able to search for datasets by one or more terms contained in dataset metadata
- must be able to browse or explore data by category

- must be able to download the catalog in a machine-readable format compatible with the common core metadata schema (<https://project-open-data.cio.gov/v1.1/metadata-resources/>) from a well-known address (typically /data.json)
- must be able to view a summary page for each dataset which details data resources, metadata, and other relevant documentation
- may be able to search for datasets by terms contained in the data

Publishers

- must be able to add, edit, or remove catalog entries
- may be able to mark catalog entries as private so they are not visible to the public

Administrators

- must be able to create, edit, and retire metadata categories
- must be able to configure default sorting for data catalog
- must be able to configure global settings for data catalog, including color theme, branding/logo, titling.
- may be able to configure additional global settings for data catalog, including custom Cascading Style Sheets (CSS), homepage display elements and layout, displayed catalog elements, and more.

Data Storage & Management Features

Consumers

- must be able to download stored data in bulk
- may be able to access stored data through a REST API

Publishers

- must be able to create new data resources
- must be able to replace or update existing data resources

Administrators

- must be able to manage permissions for other administrators and publishers

Data Presentation Features

Consumers

- must be able to preview data resources using an through their internet browser (for example, through a rows-and-columns grid or a map)
- may be able to visualize data resources in a graphical way, including maps, graphs, charts, etc.
- may be able to save visualizations with configuration settings for future viewing

Publishers

- may be able to create data visualizations and make them accessible through the data catalog or data resource page(s)

Administrators

- may be able to control whether consumer-created data presentations are visible to other consumers

Community Features

Consumers

- must be able to provide comments and feedback on individual data resources
- may be able to nominate data resources for public release
- may be able to provide comments and feedback on the data catalog

Publishers

- must be able to view consumer-nominated data resources
- may be able to update the status of consumer-nominated data resources
- may be able to respond to consumer-nominated data resources
- may be able to view and respond to consumer comments and feedback
- may be able to review and approve visibility of consumer comments and feedback

Administrators

- must track and analyze consumer feedback

Accounts and Profiles

Consumers

- may be able to register for an account with an email address
- may be able to create, edit, or remove profile information which may be accessible to other consumers

Publishers

- may be able to grant specific, registered consumers permission to access specific data resources that are owned by that publisher

Administrators

- must be able to manage permissions for other administrators and publishers
- must be able to disable consumer accounts (if consumer accounts are a feature of the portal)

Non-Role Requirements

The following requirements are general technical requirements which either don't require a specific role or apply to all roles.

- Data resources must be available for download in the format in which they were originally published on the portal
- API calls which are invalid or fail must return the appropriate HTTP response status code (404, 500, etc) and

status message

- The portal must be accessible through a city-provided internet domain name (e.g. `data.somecity.gov`)
- Transport Layer Security (TLS) must be used for account registration, logging in, and all actions performed by publishers and administrators
- Web pages must be accessible on large-screen devices (desktop computers, laptops, etc.)
- Web pages may be accessible on small-screen devices (mobile/smart phones, tablets, etc.)

How can we improve this content?

Note: Moderator approval is required to make anonymous comments visible to all readers.

13 Comments

labs.centerforgov.org

1 Login ▾

♥ Recommend

🔗 Share

Sort by Newest ▾



Join the discussion...



Luisa Helena · 6 months ago

I suggest include a service like a push, with information about new datasets ou versions

^ | ▾ · Reply · Share ›



Andrew Nicklin Mod ➔ Luisa Helena · 6 months ago

Thanks for the suggestion! Thinking about the roles, would it be a customer who provides a URL (hook) to which the push can be sent?

^ | ▾ · Reply · Share ›



Mark Headd · 8 months ago

"The portal must be accessible through a city-provided internet domain name (e.g. `data.somecity.gov`)"

Hmm, I'd suggest relaxing this a bit. There are some notable examples where government open data is published to a portal that does not have a city-provided domain name (e.g., opendataphilly.org). Additionally, there are a number of regional open data efforts underway where use of a generic, non-city domain name makes the most sense.

1 ^ | ▾ · Reply · Share ›



Andrew Nicklin Mod ➔ Mark Headd · 8 months ago

We can change this to a "may" requirement, although generally this should be manageable through DNS CNAME records and minor web server configuration. The regional case is interesting. I generally think it's important to have a layer of address abstraction in place, which helps solve some down-the-road portability issues. "nycopendata.socrata.com" is vendor lock-in; "data.cityofnewyork.us" is not.

^ | ▾ · Reply · Share ›