



Data publication

Division for Public Administration and Development Management (DPADM), United Nations Department of Economic and Social Affairs (UNDESA)

Content

- Data publication checklist
- OGD publication overview
- Data catalogue overview
- Data catalogue features and the use of Application Programming Interface (APIs)
- Data catalogues - customization
- Best practice for publishing data on the Internet

Data publication checklist

Great you arrived here! You have already followed the steps for data prioritization and preparation for publication (Learning Material N. 4.14.a).

Your data should now be prepared for publication. Please double-check:

- Datasets have been identified and prioritized
- Data has been anonymized (personal data was removed)
- Dataset is described with metadata
- Data is structured and in a format that is machine-readable
- Data is openly licensed

Now you are ready to publish your data!

OGD publication overview

There are several methods to publish data on the Internet. Established best practices for the publication of OGD include publishing your data:

- Via a data catalogue (data portal)
- Via an API (Application Programming Interface)
- Via your existing website
- Via a 3rd party website
- Via FTP-Servers (File Transfer Protocol)
- As Torrents (BitTorrent)
- As Linked Open Data (LOD) using the Resources Description Framework (RDF)

Data catalogue overview

Data catalogues provide access to publicly available data. Catalogues have the ability to **store data** in their own management system (and provide users with basic visualization options), and/or **simply manage metadata** along with links to the complete dataset stored on other servers. Main vendors include:

- CKAN www.ckan.org
- DKAN www.nucivic.com/dkan
- Junar www.junar.com
- Open Government Platform <http://ogpl.github.io>
- Socrata www.socrata.com
- OpenDataSoft www.opendatasoft.com

Data catalogue features *(example of CKAN)*



Publish & find datasets

Publish datasets via import or through a web interface. **Search** by keyword or filter by tags. See **dataset information** at a glance. Full change **history** lets you easily undo changes or view old versions.



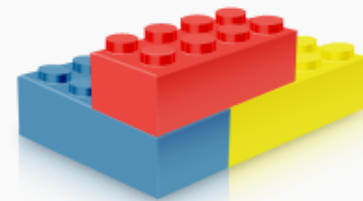
Store & manage data

Store the raw data and metadata. **Visualise** structured data with interactive tables, graphs and maps. Get statistics and usage metrics for your datasets. Search **geospatial** data on a map by area.



Engage with users & others

Federate networks with other CKAN nodes. **Theme** with CSS or integrate with a CMS. Build a **community** with extensions that allow users to comment on and follow datasets.



Customise & extend

Use the **API's** rich programming interface, and benefit from over 60 **extensions** including link checking, comments, and analytics. CKAN's **Open Source** licence allows you to download and run it for free.

Screenshot from: www.ckan.org/features

Please see the feature tour: www.ckan.org/features

Data catalogues features for data access and storage

- Provide data in a structured format such as Spreadsheets, CSV, XML, and JSON.
 - For convenience of specialized users, data can also be provided formats such as STATA or shapefiles (GIS)
- On-site (internal) dataset storage, scalable to client needs
- External links for datasets stored on separate, public servers
- Custom metadata fields
- Dublin Core/DCAT support (DCAT is an RDF-based standard for publishing metadata)
- User access rights management: can restrict access to “embargoed” datasets to specific user groups
- Manual, bulk and automated data uploading
- Data interface to external/agency systems

Data catalogues features relating to user experience

- Searching: full-text search, faceted search, filtering and sorting of results
- Geographic searching (search for data in a defined geographic area)
- On-site data preview
- Data visualization via charts and maps
- Engagement: ability to “like” and/or comment on datasets; suggest new datasets; social media links to Facebook, Google+, Twitter, etc.
- Ability for user to save visualizations and/or embed in blogs and other sites
- Multi-language support
- Mobile support

Data catalogues features and the use of Application Programming Interface (API)

- Search/query data catalogue via API
- API access in multiple formats (CSV, XML, JSON)
- Dataset additions and updates via API (for data providers)
- Update metadata via API (for data providers)

Data catalogues - customization

- Custom branding and theming
- Custom home pages and landing pages
- Integrates with external content management system
- Extensibility: add additional or custom features via modules
- Analytics: provide statistics on page views and downloads

Thank you!

Division for Public Administration and Development Management (DPADM) of the United Nations Department of Economic and Social Affairs (UNDESA).

OGD Project and OGDCE Guidelines
publicadministration.un.org/en/ogd

United Nations Publication. Copyright © United Nations, 2017. All rights reserved.