

Open data policy

Contents

1. Overview

- a. Intended audiences
- b. Feedback on this policy
- c. Document owners
- d. Other information policies
- e. Reuse of this policy

2. Context

- a. What is open data?
- b. Why open data?**
- c. Our open data goals**
- d. Our open data commitments**

3. Publishing open data

4. Using open data
5. Data use Code of Conduct

Appendix 1: Open data workflows

Appendix 2: Data release table template

Appendix 3: Open data sets



Overview

This document describes our policy on open data.

It includes our definition of open data.

It describes our best practice for how we create and publish datasets that are available for open reuse.

It details the several steps we take to ensure this best practice.

The policy also includes our Code of Conduct for those that wish to use our data.

Additionally, this document also describes how we utilise data that is already open and available.

Intended audiences

This document is for anyone interested in how we produce, publish and use open data.

Document history

This is the first issue of this policy. It was first published in September 2015. This policy is based on version 1.0 of the generic Open Data Policy (<https://github.com/OpenDataServices/OpenDataPolicy>)

Feedback on this policy

Feedback should be provided via ocf@oxfordhsire.org

Document owners

This policy is owned by Oxfordshire Community Foundation.

Other information policies

- Oxfordshire Community Foundation is registered with the Information Commissioner
- Our open data policy sits alongside policies around Confidentiality, Equality and Diversity, Social Responsibility, Website Terms of Use and staff terms and conditions.

Reuse of this policy

This policy is published under Attribution-ShareAlike 4.0 International (CC BY-SA 4.0) <http://creativecommons.org/licenses/by-sa/4.0/>

About Oxfordshire Community Foundation

Company / Charity number(s): 1151621

Oxfordshire Community Foundation is registered with the Information Commissioner: ZA140352

Registered office: Oxfordshire Community Foundation, 3 Woodin's Way, Oxford, Oxfordshire. OX1 1HD



Context

What is open data?

We subscribe to the following definition of open data:

Open data is data that can be freely used, reused and redistributed by anyone.

It is subject only, at most, to the requirement to attribute and sharealike.

Open data are the building blocks of open knowledge.

Open knowledge is what open data becomes when it's useful, usable and used.

Why open data?

We are committed to a strong community and voluntary sector in Oxfordshire and the United Kingdom. We believe that open data can assist with this ambition, through enabling transparency of our actions, in the form of open data.

Our open data goals

When publishing data about our work, we have three main aims:

That our data is useful.

We wish to publish data that brings value to those building a stronger community and voluntary sector.

That our data is meaningful.

We wish to publish data that adds to the ongoing dialogue and discussion about the sector.

That our data is accessible.

We wish to publish data in an open and accessible way, and provide regular narrative and understanding to the stakeholders.

In this context, we aim to operate an open data policy that is robust and practical, leading to greater engagement in the issues faced by the community and voluntary sector in Oxfordshire and the United Kingdom.

Our open data commitments

When considering open data at Oxfordshire Community Foundation, we commit to the following:

We will publish data openly

- We will publish data on key aspects of our activities in an open and accessible manner.
- We will update this data regularly and often.
- We will provide documentation and descriptions of the data we publish, including a licence.

We will use open data

- We will analyse, utilise and discuss our own data
- We will utilise other openly available datasets to provide insight into the voluntary sector of Oxfordshire and the United Kingdom.



We will support those using our data

When making our data openly available, we maintain a set of expectations, known as our Code of Conduct. If you utilise open data published by us, we request that you consider this.

We will support the sector with open data

In our work with voluntary sector organisations across Oxfordshire, we will provide advice and guidance in terms of publishing and using open data.

This policy provides information relevant to each of these commitments. We describe the key aspects, actions and mechanisms that we use to deliver our open data policy.

Approved By:
Jayne Woodley
April 2017



Publishing open data

Our principles and best practices

The act of providing open data is to publish and share. We understand that this involves responsibility and due diligence.

When we publish data openly, our aim is to ensure it is of a quality to be accessed, used and understood. In doing so, we place the following expectations on our data publication.

It can often be the case that whilst data is made available openly, it is poorly structured, out-of-date, bad quality amongst other aspects. It is the intention of this policy to mitigate against such factors, via these criteria.

Note: these best practices should be read alongside our open data workflows and checklists. Together, these describe our expectations and methodologies.

Privacy

Principle:	Our open data will respect privacy .
Best practice:	We will always ensure our open data is free from identifiers that could be linked to an individual person. We will not disclose any data or information that has been collected privately.
See also:	Data privacy (Appendix 1)

Quality

Principle:	Our open data will be comprehensive for the subject.
Best practice:	We will always quality assure our data, in terms of the level of completeness and readiness for publication. We will not knowingly publish data that is incomplete for the relevant focus and/or time period.
See also:	Data quality criteria (Appendix 1)

Meaningful and manageable

Principle:	Our open data will be relevant and succinct for the subject.
Best practice:	We will always consider the size, scope and spread of our data - to make it useful for those who may want to access it. We will not publish open data that is overly large, or not provide logics, lookups or additional materials.



See also:	Data segmentation (Appendix 1)
-----------	--------------------------------

Encoded

Principle:	Our open data will be interoperable .
Best practice:	We will always provide codes and lookups for our data, particularly in terms of administrative geography We will not publish data that involves jargon or acronyms that are not documented.
See also:	Data codes checklist (Appendix 1)

Format

Principle:	Our open data will be presented in an open and standard format .
Best practice:	We will publish data in common, accessible and standard formats such as CSV, XML. We will not publish open data in bespoke, redundant or proprietary formats.
See also:	Data formats checklist (Appendix 1)

Licence

Principle:	Our open data will be appropriately licenced .
Best practice:	We will always issue an open licence with our open datasets. Our default is a Creative Commons Attribution-ShareAlike 4.0 International (CC BY-SA 4.0) We will not publish data that is subject to a restrictive licence
See also:	Open licence (Appendix 1)

Documentation

Principle:	Our open data will always be well documented .
------------	---



Best practice:	We will always provide notes and guidance to accompany our datasets. We will always detail the nature, scope and purpose of our open datasets in a release document, available to all. We will not purposefully provide data that is poorly described, or requires sector knowledge to comprehend.
See also:	Data release table (Appendix 2)

Available

Principle:	We will publish data in open, accessible and consistent ways.
Best practice:	We will always publish data in a consistent method, making it accessible to all. We will not publish data with passwords or access restrictions, or in places that are not signposted.
See also:	Open data publishing channels (Appendix 1)

Updated

Principle:	Our open data will be timely .
Best practice:	We will always provide regular and timely updates to relevant open datasets. We will not miss updates to our relevant datasets, or let our data go "stale".
See also:	Data publication schedule (Appendix 1)

Feedback

Principle:	We welcome feedback and discussion about our open data.
Best practice:	We will always make it clear how to provide feedback on our open data, and any resultant actions. We will not publish data without a feedback mechanism.
See also:	Open data feedback mechanisms (Appendix 1)



Data use Code of Conduct

Oxfordshire Community Foundation encourages others to access, use and discuss our open data. We strive towards a strong community and voluntary sector in Oxfordshire, and value the contributions and insights that can be gleaned through use of data.

When doing so, we would hope the following basic Code of Conduct is observed:

- **Accessing our data**

Much of the datasets published by Oxfordshire Community Foundation are succinct and easily available for download. When accessing our data, we request that you do not place unnecessary burden on our servers by making repeated data requests over a short period of time.

- **Attribution**

When using our data, we request that our licence is observed. When producing any material that uses our data, please ensure an attribution to Oxfordshire Community Foundation is included.

- **Derivation**

When making use of our data, always state any steps that were made to undertake calculations or analysis that are not present in the source.

- **Violations**

When using Oxfordshire Community Foundation data, you must not:

- make an application that pretends to be from Oxfordshire Community Foundation organisation;
- present the data in a misleading or incorrect manner or to misrepresent or change the data;
- use the name Oxfordshire Community Foundation or the Oxfordshire Community Foundation website for party political purposes;
- use the data in or to support a criminal or illicit activity;
- use the data on an application to inflame or make comments that are racist, sexist or homophobic, or which promote or incite violence or illegal activity.

- **Discussion**

We encourage discussion of our data, and the uses. In doing so, particularly in our online forums, we request you are respectful of others.

- **Feedback**

If you spot any mistakes, errors or points for clarification, please feedback via our designated channels. We also encourage requests and ideas for new data that Oxfordshire Community Foundation may publish. Again, please do so via our feedback channels.



Appendix 1: Oxfordshire Community Foundation open data workflows

The following checklists are used by Oxfordshire Community Foundation in the preparation, publication and update of open data. These are linked to our open data best practices, detailed in our open data policy. Over time, Oxfordshire Community Foundation will update and enhance these.

When preparing any data for publication, Oxfordshire Community Foundation would always undertake the following:

Consideration: Privacy	Action
Does the data contain names of individuals?	If yes, then remove
Does the data contain any unique identifiers that can be used to retrieve personal information from external systems?	If yes, then remove

Consideration: Quality	Action
For data that is aggregated, check the results for any outliers	Check source data / reports to validate figures.
For any dataset, check for any blank fields or zero results.	If numerous, check source data. If acceptable, then record in data release table.

Consideration: Segmentation	Action
For any dataset, consider the overall physical file size	If over 10MB, then check contents and consider further segmentation
For any dataset, check the column headers and data labels are legible.	If not, provide lookup file and note in data release table
For aggregated datasets, check that aggregations are explained and logged.	Ensure these are documented in data release table
For any dataset, check that time periods used are in accordance with common standards (eg: financial quarters, calendar months)	If there is a bespoke date range, then detail in data release table



Consideration: Data Codes	Action
For geographic areas used within datasets, provide the code alongside the name.	Applicable to: Local authority Ward Clinical Commissioning Groups
Provide and/or signpost data users to the latest lookup of any codes used.	In the case of administrative geographic regions, refer to authoritative sources such as Ordnance Survey, Office for National Statistics and the NHS.
When using internal / organisation name specific codes, ensure that a lookup and/or explanation is provided.	Log this in the data release table.

Consideration: Formats	Action
For spreadsheets and tabular data, release in standard open formats.	Release as: Open Document Format for spreadsheets (.odf) Comma Separated Format for flat files (.csv)
When working with other data standards and systems, ensure that the format is open and accessible.	Consider XML, JSON or RDF formats as open. Check with standard or publication organisation.
Avoid publishing data in closed, proprietary and formats that make the data inaccessible.	



Open data licence

When publishing data, we will always ensure a relevant licence is provided. Our default licence is a Creative Commons Attribution-ShareAlike 4.0 International (CC BYSA 4.0)

Consideration: Licence	Action
Does the dataset fit within the default licence for [organisation name]?	If yes, then ensure this is in the data release table.
See: http://creativecommons.org/licenses/by-sa/4.0/	If no, then select alternative, and document accordingly.
Does the data contain any information that is derived from other sources?	If yes, then detail these sources in the data release table. If there may be an issue with these derivations, then seek advice.

Consideration: Schedule	Action
For datasets that are updated periodically, ensure this takes place within acceptable timeframe. Guidance: Monthly updates - two weeks after period ends Quarterly updates - one month after period ends Annual updates - three months after period ends	If data publication is outside of these thresholds, update/add to data release table
Ensure that relevant older data can be accessed after an update - that it is not deleted or destroyed.	For ongoing statistics, ensure the new time period data is made available alongside other periods. Where data must be overwritten, document in data release table.

Consideration: Publishing Channels	Action
When creating data files, check the file name is logical, descriptive and devoid of any inconsistencies. eg: volunteers_statistics.csv rather than volunteersstats(1)-EDITED.csv	Use the shared naming convention wherever possible.



When hosting data files, always ensure that the end URL is accessible, and free from any security barriers, passwords or blocks.	If there is an issue in terms of accessing the URL to the data file, seek advice.
--	---

Consideration: Feedback	Action
When an "Issue" is posted, respond accordingly - even if acknowledgement.	Where the issue can be progressed, respond accordingly. If no action can be applied immediately, respond accordingly.
When feedback is received via settings such as face-to-face meetings or workshops, consider how best to add this to existing feedback.	Where appropriate, create a new Issue for the relevant dataset, attributing the source of the observation / remark.
When comments are made about our usage of open data, respond according to the Code of Conduct.	

Attribution checklist

When using datasets published by other organisations, we will always ensure the following attribution considerations are provided within the context of any material we produce.

Attribution	Comment
Name of the dataset utilised	eg: Adult Learning Centres
Publisher of dataset	eg: Manchester City Council
Source URL (from where the data can be retrieved)	eg: http://www.manchester.gov.uk/site/custom_scripts/getmetadata.php?guid=a65c8dfb-7371-4db9-a3b3-a9e4be72a493
Date retrieved	DD/MM/YYYY
Notes on usage	Any notes on actions undertaken that result in the source data being changed or modified.



Appendix 2: data release table template

Example: Volunteer data

Data consideration	Oxfordshire Community Foundation response
Dataset name	Grant data
Brief description	Information on the grants awarded by Oxfordshire Community Foundation following each round (approx. quarterly)
Responsible person / team <i>Who is the authority for this data?</i>	Data manager
Privacy <i>What privacy considerations should be taken?</i>	Names and personal details are not included.
Legality <i>Does the data contain any derived data?</i>	No
Format <i>How will be the data be made available</i>	As a spreadsheet, saved in a CSV format
Segmentation <i>How will the data be segmented?</i>	One file, with the following reports: Grants awarded to each project Grants awarded from each fund
Dates <i>What date period are used in the data, if any?</i>	The project activity will have a start and end date. Each grant round will be identified by a month.



Structured <i>Does the data contain any inherent jargon or interpretation?</i>	No.
Interoperable <i>How can other datasets be referenced?</i>	Groups charity number will be included Beneficiary area will be identified by LSOA code
Availability <i>How will the data be disseminated?</i>	Our open data registered with 360 Giving
Documentation <i>How will the data (and processes) be described?</i>	On our data website pages.
Updated <i>What is the update timetable?</i>	Quarterly - two weeks after the end of each grants panel
Licensing <i>What licence will be applied?</i>	Creative Commons, Attribution http://creativecommons.org/licenses/by/4.0/
Feedback <i>How will feedback be progressed?</i>	Established Oxfordshire Community Foundation feedback processes
Additional Notes	



Appendix 3: Opendata sets

List of datasets published by |organisation name|. This list will be updated and maintained.

Dataset	Description	First published	Update(s)
Funding data	Details of funding provided by us to other organisations		Quarterly

