

FREEDOM OF INFORMATION

A guide to Datasets and Re-use January 2014



This briefing provides guidance on the CCGs legal responsibility under Section 102 of the Protection of Freedoms Act 2012 which adds new provisions to the Freedom of Information Act (FOIA) (in particular sections 11 and 19) regarding datasets.

The new provisions are about how information is released, rather than what information is released. They only relate to information that the CCG holds as a dataset, which is a defined term in the new provisions. They are about the re-use of those datasets that the CCG provides in response to a request or under a publication scheme. There is no new duty to provide any information in response to a FOIA request that was not previously accessible.

The main duties of the new provisions are:

- To provide dataset information in a re-usable format (normally CSV format)
- Allow commercial re-use of dataset information via an Open Government Licence
- Publish any datasets requested and subsequent updates on the CCG's Publication Scheme.

Datasets

A dataset is a collection of factual information in electronic form to do with the services and functions of the authority that is neither the product of analysis or interpretation, nor an official statistic and has not been materially altered.

The FOIA defines datasets as:

(5) In this Act “dataset” means information comprising a collection of information held in electronic form where all or most of the information in the collection—

The term datasets only applies to information that is held in electronic form. If the CCG only holds the requested information in another form, such as hard copy, then the CCG is not required to turn the hard copy information into an electronic dataset. The term “a collection of information” also implies that it comprises several data elements, rather than, for example, a single figure.

The FOIA lists criteria that defines whether a collection of electronic information constitutes a dataset. The collection is a dataset if “all or most” of it meets all of these criteria.

- (a) has been obtained or recorded for the purpose of providing a public authority with information in connection with the provision of a service by the authority or the carrying out of any other function of the authority...
- (b) is factual information which—
- i. is not the product of analysis or interpretation other than calculation ...

ii. is not an official statistic (within the meaning given by section 6 (1) of the Statistics and Registration Service Act 2007) ...

(c) remains presented in a way that (except for the purpose of forming part of the collection) has not been organised, adapted or otherwise materially altered since it was obtained or recorded.

Re-usable Format

If a CCG is providing information that constitutes a dataset and the requester has expressed a preference to receive the information in electronic form, the public authority must provide it in a re-usable format so far as reasonably practicable.

The Explanatory Notes to the Protection of Freedoms Act 2012 say that “a re-usable format is one where the information is available in machine-readable form using open standards which enables its re-use and manipulation” (note 389). The ability to re-use and manipulate the data depends on its being machine readable and based on open standards.

“Machine-readable” means that the data must be structured. It is not enough that the data is in a document that can be read by a user on a computer screen. It must be possible for software to extract specific elements of the data. This is possible if the data is in a structured form such as a spreadsheet, but it is not normally possible if the data is contained in the text of an unstructured document in Word or PDF format.

The Explanatory Notes to the Protection of Freedoms Act 2012 show that the term “capable of re-use” means not only that the data is machine-readable but also that it is available in a format based on open standards, rather than a proprietary format. For example, a spreadsheet that is made available in a CSV (comma separated value) format uses an open standard, whereas a Microsoft Excel spreadsheet does not, even though that particular product is widely used. However, the Excel spreadsheet can be saved as a CSV file.

Reasonably Practicable

The duty to provide the information in an electronic form capable of re-use is not absolute. Section 11(1A) says that the public authority must provide it in this form, “so far as reasonably practicable”. The FOIA goes on to say:

(2) In determining for the purposes of this section whether it is reasonably practicable to communicate information by particular means, the public authority may have regard to all the circumstances, including the cost of doing so.

Factors that can affect whether it is reasonably practicable to provide the dataset in re-usable form include the time and cost of conversion, technical issues and the resources of the CCG. Issues may include:

- To convert it to an open format may involve significant expenditure on purchasing additional software or paying an expert to write a program.
- It may be a worthwhile investment, in order to make the dataset available and re-usable in the spirit of open data.
- If the scope of the request is limited and the CCG is planning to publish more comprehensive information as open data, the CCG may decide that it is not reasonably practicable to convert the limited information that has been requested to an open format.
- If a data set is heavily redacted, although the requestor is still entitled to the information remaining it may have limited value and the cost of conversion of this information may not be reasonably practicable.
- Technical issues may make it impractical to convert the information

Re-use and Licensing

Apart from the dataset provisions, the FOIA gives a right of access to information but no right to re-use it. However, the dataset provisions create a duty under section 11A(2) to provide datasets that are relevant copyright works under a licence that permits re-use:

(2) When communicating the relevant copyright work to the applicant, the public authority must make the relevant copyright work available for re-use by the applicant in accordance with the terms of the specified licence.

For this duty to apply, four conditions must be satisfied:

11A Release of datasets for re-use

- (1) This section applies where—
- a. a person makes a request for information to a public authority in respect of information that is, or forms part of, a dataset held by the authority,
 - b. any of the dataset or part of a dataset so requested is a relevant copyright work,
 - c. the public authority is the only owner of the relevant copyright work, and
 - d. the public authority is communicating the relevant copyright work to the applicant in accordance with this Act.

If the dataset is a relevant copyright work, the public authority must provide it under the terms of a specified licence. In order to license a relevant copyright work for re-use, the

CCG must be the only owner of the copyright or the database right in that work. If a third party owns the copyright or database rights in the dataset, CCG should say when it makes the dataset available that it is not licensing re-use and tell the requester who owns the rights, if it knows this. If a third party owns the rights in part of the dataset, the public authority should make it clear that it is not licensing that part for re-use.

The Open Government Licence (OGL) is the default licence for datasets, other licences (the Non-Commercial Government Licence and the Charged Licence) are available where that the OGL is not appropriate.

There is a Code of Practice under section 45 of FOIA which deals specifically with the dataset provisions. The Code encourages public authorities to use the Open Government Licence (OGL) for datasets that can be re-used without charge. Furthermore, the UK Government Licensing Framework recommends the OGL as the default licence for public sector information.

There may however be circumstances in which it is not appropriate. The Non-Commercial Government Licence is available for situations where information may only be re-used for non-commercial purposes. A CCG will need to have a good reason for making a dataset available for re-use under a licence other than the OGL.

Publication Scheme

Under its publication scheme a public authority should publish datasets that have been requested, and any updated versions it holds, unless it is satisfied that it is not appropriate to do so. Factors that can be taken into account when deciding whether it is not appropriate to include whether the information is exempt and the cost and effort of publishing in a re-usable form. Some things to consider would be:

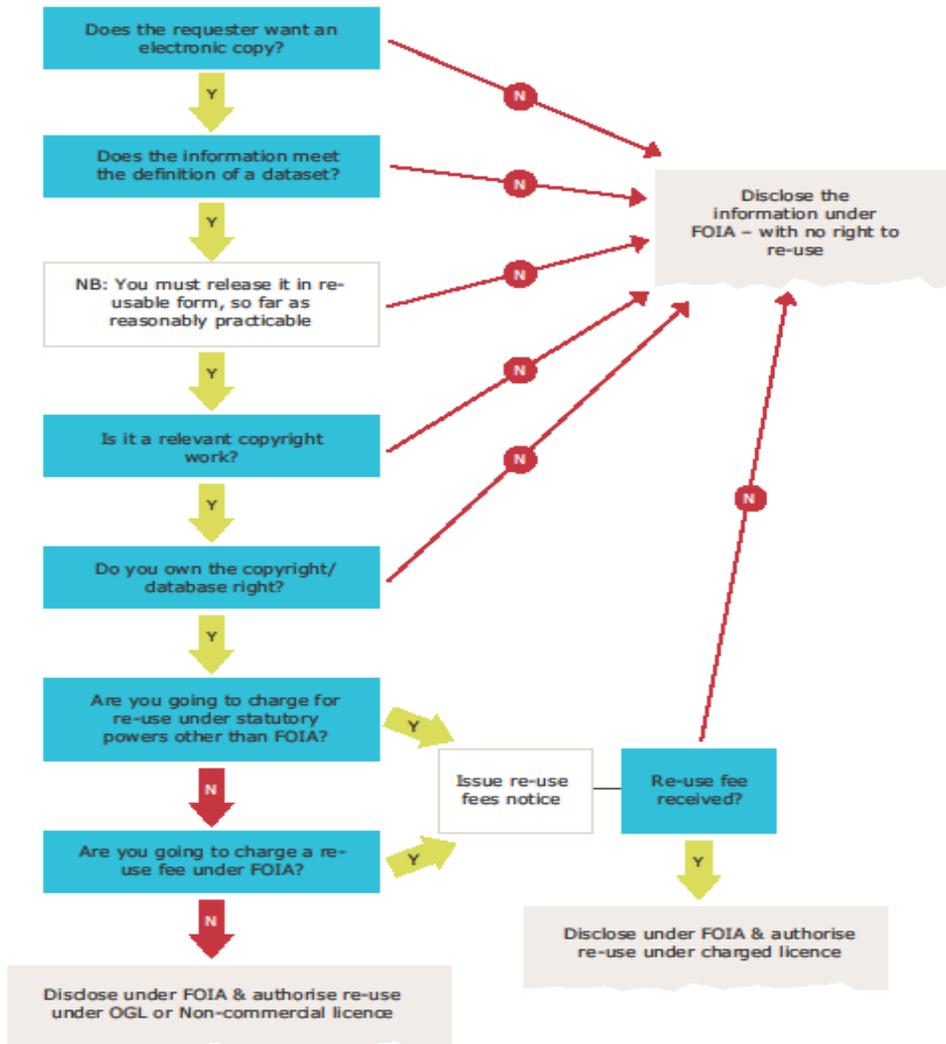
- There may be no real benefit, in terms of promoting transparency, in continuing to publish the dataset routinely.
- The cost or the technical issues involved in making the dataset available under a publication scheme
- Is there a public interest in providing the information routinely under the Publication Scheme.

Section 19(2A)(a)(ii) extends the requirement to “any up-dated version held by the authority of such a dataset”. This does not mean that once the CCG has published a dataset under its publication scheme, it is required to keep it updated purely for the purposes of the publication scheme. However, if the CCG actually holds an updated version of the dataset, then it is required to publish that updated version. The dataset provisions do not say anything about how frequently the published version should be updated and the requirement to publish updated versions is also limited by the “not appropriate” qualification.

Complaints

If a requester considers that the public authority has not dealt with their request in accordance with section 11, 11A, 11B or the Datasets Re-use Fees Regulations, for example because the public authority has not provided a dataset in a re-usable form, they should ask the authority to review its decision. If the requester is not satisfied with the outcome of the internal review, they may apply to the Information Commissioner for a decision under section 50 of the FOIA.

Datasets Flowchart



Datasets (sections 11, 19 & 45)
20130809
Version: 1.0

33