

Datasets (sections 11, 19 & 45)

Freedom of Information Act

Contents

Introduction.....	2
Overview.....	2
Scope and purpose of the dataset provisions	4
Providing a dataset in response to a request.....	5
How the duty arises	5
Datasets	7
Electronic form.....	7
Service or function.....	8
Factual information.....	8
Materially altered.....	11
Capable of re-use	12
Reasonably practicable.....	14
Licensing re-use: FOIA or RPSI?.....	16
Licensing re-use under FOIA	19
Specified licence	20
Fees	21
Publication schemes.....	24
Not appropriate	26
Model publication scheme	28
Complaints	28
Environmental information.....	29
Practical issues.....	30
Requests	30
Licensing re-use	30
Personal data	31
Copyright and licences	31
More information.....	32
Annex 1: Datasets flowchart	33

Introduction

1. The Freedom of Information Act 2000 (FOIA) gives rights of public access to information held by public authorities.
2. An overview of the main provisions of FOIA can be found in [The Guide to Freedom of Information](#).
3. This is part of a series of guidance, which goes into more detail than the Guide, to help public authorities to fully understand their obligations and promote good practice.
4. This guidance explains to public authorities the provisions regarding datasets that were added to FOIA by the Protection of Freedoms Act 2012. These create duties in relation to providing datasets in response to requests and making them available under a publication scheme.

Overview

- FOIA includes specific provisions relating to datasets in sections 11, 11A, 11B and 19.
- These provisions are about how information is released, rather than what information is released. They only relate to information that the public authority holds as a dataset, as defined. They are about the re-use of those datasets that the public authority 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, and there are no new exemptions from that duty.
- If a public authority 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 form so far as reasonably practicable.
- 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.

- A re-usable form means that the dataset is in a machine readable form and based on open standards.
- 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 public authority.
- If the dataset is a relevant copyright work and the public authority owns the intellectual property rights, the public authority must provide it under the terms of a specified licence. Datasets covered by Crown copyright or Parliamentary copyright are not relevant copyright works. If the dataset is covered by the Re-use of Public Sector Information Regulations 2015 (RPSI), then re-use licensing is dealt with under RPSI, not FOIA.
- The Open Government Licence (OGL) is the default licence for datasets that can be re-used without charge; other licences (the Non-Commercial Government Licence and the Charged Licence) are available where that is not appropriate.
- Under FOIA, a public authority may charge a fee for communicating the information and a fee for making the dataset available for re-use. There are fees regulations dealing with making the dataset available for re-use under FOIA. There is no re-use fee if the dataset is provided under the OGL or the Non-Commercial Government Licence.
- 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 include whether the information is exempt and the cost and effort of publishing in a re-usable form.
- Complaints that a public authority has not met its duties under the dataset provisions are dealt with by the Information Commissioner. The Information Commissioner also deals with complaints under RPSI.
- The dataset provisions do not apply to the Environmental Information Regulations but public authorities should also take account of regulation 6 of the EIR in relation to form or format.

- When dealing with a request involving a dataset, public authorities should first consider whether the information is exempt from disclosure under any FOIA exemption. In particular they should consider whether any personal data may be disclosed.
- Public authorities must also ensure that they own the copyright and database rights in a dataset before licensing it for re-use.

Scope and purpose of the dataset provisions

5. FOIA includes specific provisions relating to one type of information, namely 'datasets'. This term is defined in section 11(5) FOIA. The dataset provisions are in section 11 (means by which communication to be made), section 11A (release of datasets for re-use), section 11B (power to charge fees in relation to release of datasets for re-use), section 19 (publication schemes) and section 45 (Secretary of State's Code of Practice); the term 'dataset' has also been added to the list of definitions in section 84.
6. The dataset provisions were introduced by section 102 of the Protection of Freedoms Act 2012 (POFA), and they have been amended by regulation 21 of the Re-use of Public Sector Information Regulations 2015 (RPSI).
7. Broadly speaking, the dataset provisions are intended to ensure that public authorities make datasets available, either in response to a FOIA request or proactively under a publication scheme, in a way that allows them to be re-used. They create a duty to provide datasets in a re-usable form and under licensing conditions that permit re-use. However, if a dataset is covered by RPSI, then licensing for re-use is dealt with under RPSI, not FOIA.
8. The government has also issued a [Code of Practice](#) under section 45 of FOIA, dealing specifically with the dataset amendments. This Code supplements the main Section 45 Code of Practice; it does not replace it. Public authorities should consult the [Code of Practice](#) on datasets, as well as this guidance document, in order to understand their duties under the dataset provisions.

9. Public authorities' duties in relation to datasets are to do with the means of communicating information in response to requests, and making datasets available via publication schemes. The dataset provisions contain no additional right to obtain information that is not otherwise accessible under FOIA. They are about providing the information in a re-usable form and making it available for re-use, if it is a dataset.
10. Before the dataset provisions were introduced, FOIA provided a right for requesters to access information but it did not deal with its re-use. The dataset provisions include a requirement to provide datasets in a form that enables them to be re-used. However, making information available for re-use involves a consideration of not only the form or format of the information, but also of any relevant licensing conditions and charges. For many public authorities, it is likely that licensing and charging for the re-use of information, including datasets, will be dealt with under RPSI. In cases where RPSI does not apply to a dataset, licensing and charging is dealt with under the FOIA dataset provisions.
11. The dataset provisions are in line with the increasing trend towards Open Data. The [Open Data White Paper, Unleashing the potential](#), discussed this trend and set out government policy in this area. The aim of this policy is not only to promote greater transparency and accountability by releasing more data held by central government and public authorities, but also to enable and encourage developers to re-use that data in new products and applications. This in turn is expected to create business opportunities and drive economic growth. The dataset provisions were introduced as part of this policy.

Providing a dataset in response to a request

How the duty arises

12. Section 11 of FOIA deals with how information is provided to a requester (or "applicant") in response to a request. It applies to the information that the public authority is going to release; it does not apply to any information that is exempt from disclosure. Therefore, before applying section 11, the public authority must first have decided what information should be released and what, if any, is exempt under a FOIA exemption, as it would do with any FOIA request.

13. The dataset provisions add a new subsection 11(1A) after subsection 11(1):

(1A) Where—

(a) an applicant makes a request for information to a public authority in respect of information that is, or forms part of, a dataset held by the public authority, and

(b) on making the request for information, the applicant expresses a preference for communication by means of the provision to the applicant of a copy of the information in electronic form,

the public authority must, so far as reasonably practicable, provide the information to the applicant in an electronic form which is capable of re-use.

14. These provisions are about the duty to provide a dataset in a re-usable form when responding to a request. The requester does not have to specify when submitting their request that they want a re-usable dataset. If the information that the requester wants happens to be held in a dataset (as defined in the provisions), and the requester expresses a preference for receiving the information in electronic form, then the duty to provide the dataset in a re-usable form is triggered. If the requester has asked to receive the information in hard copy, then there is no duty to provide it in an electronic form capable of re-use.
15. If the requester has submitted their request electronically eg by email, then the public authority should treat this as in effect expressing a preference to receive a copy of the information in electronic form (unless of course the requester has specifically asked for a hard copy). They should not seek to avoid the duty to provide a dataset in a re-usable form on the basis that the requester has not literally asked for a copy of the information in electronic form.
16. Where the requester has asked for information that forms part of a dataset, then the authority has a duty to provide that information in a re-usable form. So, if the requester has asked for part of a dataset, the public authority is not obliged to provide the whole of the dataset; they only have to provide the information that has been requested. Depending on the

circumstances, it may nevertheless be easier for the authority or more helpful to the requester to provide the whole dataset, provided the other information is not exempt.

17. In order to carry out the duty in section 11(1A) it is important to understand certain key terms, in particular what is meant by:

- "a dataset"
- "an electronic form which is capable of re-use"
- "so far as reasonably practicable"

Datasets

18. The term "dataset" is defined in subsection 11(5) of FOIA. This definition contains a number of elements.

Electronic form

(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—

19. This paragraph shows that the term only applies to information that is held in electronic form. If the public authority only holds the requested information in another form, eg in hard copy, then it is out of scope; the public authority is not required to turn the hard copy information into an electronic dataset.

20. This paragraph also defines a dataset as "a collection of information" which implies that it comprises several data elements, rather than, for example, a single figure.

21. The subsection then goes on to list the criteria that define 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. If all or most of the collection does not meet all of the following criteria, it is not a dataset for FOIA purposes.

Service or function

22. The first criterion is about the purpose for which the authority holds the information:

(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...

23. The wording here means that the information is to do with the services that the public authority provides or the functions it is carrying out. The services and functions of a public authority ultimately relate to its purpose or the reason for its existence, which are usually derived from statutory powers or duties that are given to that authority. Furthermore, the information must have been collected "for the purpose of providing the public authority with information" in connection with these services and functions. This suggests that it is likely to be management information, or information that the public authority itself needs in order to provide services and carry out functions. This would be consistent with the Open Data policy aim of providing greater transparency about the work of public authorities. The phrase "in connection with" should be interpreted broadly. It is likely that this part of the definition will encompass most of the information held by public authorities, but the wording also suggests that there can be some information held by a public authority that does not meet this criterion.

Factual information

24. The next part of the definition specifies that "all or most of the information in the collection" must be "factual". This does not simply mean numeric information. It could include, for example, not only figures for expenditure by departments, but also a list of the addresses of properties. Factual information is quantitative rather than qualitative information.

Example

A local authority conducts a survey of visitors to a tourist attraction that it runs, to gather information for business planning. The survey includes questions on mode of transport used to visit the attraction, and the first part of the home

postcode of visitors. The answers to both of these questions would be factual information.

The survey also includes a question asking for any comments, in free text. The answers to this would not be factual information. They cannot be measured or compared in an objective way.

If the survey had asked for respondents to tick a box showing how likely they were to return (eg very likely; likely; unlikely), then the results of this would be factual information. They would show the fact that x% of respondents said they were very likely to return etc.

However, if the majority of the information was factual, then the results of the survey as a whole would still constitute a dataset, provided the other criteria in section 11(5) are met.

- Not the product of analysis or interpretation

25. The term “factual information” is then qualified by two further criteria; firstly,

(b) is factual information which—

(i) is not the product of analysis or interpretation other than calculation ...

26. This suggests that the definition is limited to ‘raw’ data that has been produced or obtained by the public authority, rather than value-added data that has been produced by analysis or interpretation.

The [Explanatory Notes](#) to the Protection of Freedoms Act 2012 say:

“Examples of the types of datasets which meet the definition, though not a comprehensive list, will include datasets comprising combinations of letters and numbers used to identify property or locations, such as postcodes and references; datasets comprising numbers and information related to numbers such as spend data; and datasets

comprising text or words such as information about job roles in a public authority” (note 394).

27. The examples given in the Explanatory Notes are of data that has been collected or produced by the authority, without further analysis or interpretation; the data records the postcodes of properties or the amount spent by a department or the job roles within the authority.
28. Furthermore, the phrase “other than calculation” means that, for example, if expenditure data has been collected at the level of sections within the departments of a public authority, and it is then added up to show expenditure by department and by the authority as a whole, or even to show the percentage of the total expenditure by each department, this does not take it out of the definition, since totals and percentages are produced by calculation and they are inherent in the data itself.
29. However, it is likely that a table in a report proposing how resources should be allocated to departments in future years, based on the authority’s policies and priorities, would not be factual information because it is produced by analysis and interpretation, rather than simply recording and making calculations with factual data. The analysis and interpretation depend on factors that are not inherent in the data itself.
30. We do not consider that quality checking the information, eg ensuring that entries are recorded consistently or correcting errors that may have occurred in recording them, would constitute analysis or interpretation so as to take the dataset out of the definition. These processes are part of the authority’s normal work of collecting management information, which is likely to involve checking the accuracy or completeness of the information collected, rather than further analysing or interpreting it.
31. This does not mean that an inaccurate or incomplete dataset is outside the definition or exempt from disclosure. Even if the information that has been requested is inaccurate or incomplete, the public authority must still release it unless a FOIA exemption applies. If the public authority does have to release incomplete or inaccurate information in response to a FOIA request, because that is the information that it holds, then it would be sensible for it to issue a caveat or other information to put the release into context at the same time.

- Official statistics

32. The second qualification to the term “factual information” concerns official statistics:

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

33. Official statistics, as defined in the Statistics and Registration Service Act 2007, are not datasets within the meaning of the dataset amendments to FOIA. However, the underlying raw data used to produce official statistics could fall within the definition of a dataset, assuming that the other criteria were satisfied.

Materially altered

34. The final part of the definition is that “all or most of the information in the collection”,

(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.

35. To meet the definition, the information must remain presented in a way that has not been materially altered. This implies that the criterion is whether the *presentation* of the information has been altered, not whether the *information* itself has been altered. This means that if the information remains the same but the way that it is presented is changed, it can fall outside the definition. However, the phrase “or otherwise materially altered” suggests that any such change, ie organisation or adaptation, would have to be significant. A minor change, such as reordering the columns in a spreadsheet, is unlikely to represent a material alteration in the way that the information is presented. The intention of the subsection seems to be to define a dataset as a collection of raw data that is presented in essentially the same way that it was organised when the public authority originally obtained or recorded it.

36. The phrase “except for the purpose of forming part of the collection”, must refer to the information in the dataset, since it is the information that forms part of the collection. To paraphrase the legislation in simple terms, a dataset is a collection of information where the way that the information is presented has not been materially altered since the information was first collected, apart from the work involved in putting the information into the dataset. The process of adding information to a dataset does not constitute a material alteration to the presentation of that dataset.
37. When releasing a dataset in response to a FOIA request, the public authority may need to redact information that is exempt under one of the FOIA exemptions, for example personal data that is exempt under section 40. Redacting information in this way will not take the dataset out of the definition, even if the redactions are substantial. This is because the definition of a dataset in section 11(5) refers to the dataset that is held by the public authority ie the original, unredacted version, and not to a redacted version that can be released in response to a FOIA request.
38. There may be situations where, in order to answer a request, a public authority extracts data from various sources and compiles a new table or spreadsheet, possibly involving calculations such as totals and percentages. Whilst still held for the purposes of FOI this new table or spreadsheet is not a dataset for the purposes of these provisions. The dataset provisions are about raw data. The new table or spreadsheet is not data that has been obtained or recorded to provide the authority with information in connection with its services or functions; the presentation of the data on which it is based has been materially altered since it was originally obtained or recorded.

Capable of re-use

39. The public authority has a duty to provide the requested information in an electronic form “which is capable of re-use”.
40. 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.

41. "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.
42. 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...Regulation 2 of RPSI contains the following definitions, which are helpful in interpreting the phrase 'capable of re-use':
 - 'machine readable format' means a file format structured so that software applications can easily identify, recognize and extract specific data, including individual statements of fact, and their internal structure;
 - 'open format' means a file format that is platform-independent and made available to the public without any restriction that impedes the re-use of documents;
 - 'formal open standard' means a standard which has been laid down in written form, detailing specifications for the requirements on how to ensure software interoperability;
43. A dataset made available in a CSV format is capable of re-use, but it is possible to improve the re-usability of data beyond that level. The [Open Data White Paper, Unleashing the potential](#) refers (at pages 24-25) to the Five Star Scheme developed by Sir Tim Berners-Lee and indicates that the Government intends to adopt it as a measure of the usability of its Open Data. There is no mandatory requirement that public authorities achieve a certain level on this scheme in order to comply with the dataset provisions, but it can serve as a tool for assessing

the re-usability of data. The Five Star Scheme also refers to the issue of open licences; the requirements of the dataset provisions in relation to licensing are discussed below in this guidance document.

44. Public authorities should also take account of the [Public Data Principles](#) drawn up by the Public Sector Transparency Board. Following these principles will help to improve the re-usability of datasets.

Reasonably practicable

45. 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".
46. Consideration of what it is reasonably practicable must relate to how the information is provided (ie in a re-usable form), not to whether it should be provided. This is an amendment to section 11, which is about the means of communicating information to the requester. Any issues about the cost of answering the request, as distinct from the cost of converting the dataset to a re-usable form, are not relevant, since they are dealt with already under section 12 of FOIA.
47. There is no definition of "reasonably practicable" in the dataset provisions and no specified cost limit. However, it should be remembered that the provisions are an addition to section 11 of FOIA, which deals with the means by which the public authority communicates information to the requester. They introduce an additional requirement (ie to provide a dataset in re-usable form) where the preferred means of communication are electronic. Section 11(2) refers to the concept of 'reasonably practicable' in relation to communicating information by particular means:

(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.

Example

The case of [Independent Parliamentary Standards Authority \[IPSA\] v the Information Commissioner and Ben Leapman, EA/2012/0242 29 April 2013](#), did not involve datasets, but it did touch on the issue of whether it was reasonably practicable to provide the requested information (in this case, copies of invoices) in the form requested. IPSA argued that it was not reasonably practicable to provide the information in the requester's preferred form because of the anticipated cost of having to do the same in response to other similar requests in future.

The First-tier Tribunal said at paragraph 28 that they "had no doubt whatsoever that the true construction of section 11 allowed only regards to 'all the circumstances' of a particular application".

The First-tier Tribunal described this as an ancillary issue in that case, but their comments suggest that the cost of providing datasets in a re-usable form in the future cannot be used to argue that it is not reasonably practicable to do so in any particular case.

48. The question in relation to datasets is whether it is reasonably practicable to provide the information in a re-usable form. The time and the cost involved can be relevant factors. In some cases it will be straightforward and involve no expenditure to convert a dataset to an open format such as CSV. If the dataset is held as a relational database, rather than as a single table, then this may be a more complex operation. If the dataset is very large or held in a proprietary system, to convert it to an open format may involve significant expenditure on purchasing additional software or paying an expert to write a program. The size and resources of the public authority can be relevant here; a large public authority may be better able to undertake this work than a very small one, but on the other hand a large public authority with a national coverage may be handling sizeable datasets that would be expensive to convert.
49. Even if converting the dataset would involve significant expenditure, the public authority may decide that this is a worthwhile investment, in order to make the dataset available

and re-usable in the spirit of open data. On the other hand, if the scope of the request is very limited, and the public authority is planning to publish more comprehensive information as open data, the public authority may decide that it is not reasonably practicable to convert the limited information that has been requested to an open format.

50. There may also be a situation in which the dataset has been heavily redacted, eg to remove personal data that is exempt from disclosure under FOIA, and what is left may have limited informative value. The requester is still entitled to receive this under FOIA, but if to convert it to a re-usable form would involve substantial cost and effort there may be a case for saying that it is not reasonably practicable to do so.
51. Apart from the cost considerations, there may also be cases where technical issues make it impracticable to convert the data from a proprietary to an open format.
52. What is reasonably practicable is a decision for the public authority to make, taking account of all the circumstances of the case, including the purpose of the legislation, which is to promote open data. If the public authority decides that it is not reasonably practicable to provide the information in a re-usable form, the requester can ask the authority to review its decision and then, if they are not satisfied with the authority's review, complain to the Information Commissioner.

Licensing re-use: FOIA or RPSI?

53. As well as requiring public authorities to provide datasets in a form that is technically capable of re-use, the dataset provisions also create a duty, in certain specified circumstances, to provide them under a licence permitting re-use.

11A Release of datasets for 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.

54. In most cases public authorities will have to deal with licensing for re-use of datasets under RPSI, rather than under FOIA. RPSI deals with the re-use of public sector information in general; the FOIA provisions on licensing only apply in a situation where the dataset is not covered by RPSI. This section explains how to decide whether FOIA or RPSI applies. The flowchart in Annex 1 also shows how RPSI relates to the process of releasing datasets for re-use.
55. The duty to licence the re-use of datasets released in response to a FOIA request is in section 11A. Subsection 11A(1) sets out the conditions in which it applies, but subsections 11A(1A) and 11A(1B) set out exceptions to this, where RPSI applies instead.

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.

(1A) But if the whole of the relevant copyright work is a document to which the Re-use of Public Sector Information Regulations 2015 apply, this section does not apply to the relevant copyright work.

(1B) If part of the relevant copyright work is a document to which those Regulations apply—

(a) this section does not apply to that part, but

(b) this section does apply to the part to which the Regulations do not apply (and references in the following provisions of this section to the relevant copyright work are to be read as references to that part).

56. The four conditions in section 11A(1) are as follows. All of these conditions must be met for the duty in section 11A(2) to apply:

- The requester must have made a request for information that is (or is part of) a dataset that the authority holds. The requester does not have to have specifically referred to a 'dataset' in their request.
- The dataset or any part of it is a 'relevant copyright work'. If someone requests information and that information is held as a dataset (or a part of one) and any of that dataset (or any of that part) is a relevant copyright work, then this condition is satisfied. The term 'relevant copyright work' is explained below.
- The public authority is the only owner of the relevant copyright work. The public authority cannot licence the re-use of a dataset if it does not own the intellectual property in that dataset.
- The public authority is providing the dataset to the requester, ie the information is not exempt from disclosure under one of the FOIA exemptions.

57. The key terms used here are defined in section 11A(8) as follows:

- A 'relevant copyright work' is a copyright work¹ or a database² subject to database right³, but it excludes a relevant Crown work or a relevant Parliamentary work.
- A 'relevant Crown work' is one where the Crown is the copyright owner or the owner of the database right
- A 'relevant Parliamentary work' is one where the House of Commons or the House of Lords is the copyright owner or the owner of the database right.

58. The exclusion of 'relevant Crown works' from the definition of 'relevant copyright work' means that they are not subject to the duty in section 11A(2) to make the dataset available for re-use under a specified licence. This duty would not apply to

¹ As defined in the Copyright Designs and Patents Act 1988 section 1(2)

² As defined in the Copyright Designs and Patents Act 1988 section 3A

³ As defined in Part 3 of the Copyright and Rights in Databases Regulations 1997 S.I. 1997 No.3032

public authorities (such as government departments) whose datasets are 'relevant Crown works'. However, such datasets would already be available for re-use by authority of the Controller of HMSO.

59. If the four conditions in section 11A(1) are satisfied, then the next question for the public authority is: does RPSI apply to the dataset, or to the relevant part of the dataset?
60. RPSI applies to public sector bodies, as defined in regulation 3 of RPSI. Many FOIA public authorities are also RPSI public sector bodies, but some are not, for example schools, colleges and universities (although university libraries are covered by RPSI). The public authority therefore needs to establish whether it is a public sector body as defined in regulation 3 RPSI.
61. RPSI only applies to information that is part of the 'public task' of the public sector body. The public task comprises the core role and functions of the body as defined in legislation or established through custom and practice. As the section above on [Service or function](#) makes clear, datasets are information recorded or obtained in connection with providing services and carrying out functions, so if a dataset meets that criterion it is likely that it will form part of the public task of the public sector body, in terms of RPSI.
62. If RPSI does apply, then, because of the exceptions in section 11A(1A) and (1B), the licensing of the dataset for re-use must be dealt with under RPSI. The following sections of this guidance, which deal with licensing and re-use fees under FOIA, are not relevant. Instead, public authorities should consult our brief [Guide to RPSI](#) and the detailed [guidance produced by the National Archives](#).
63. If RPSI does not apply, then the licensing of the dataset for re-use must be dealt with under FOIA (provided the four conditions in FOIA section 11A(1) are met). The following sections explain the FOIA provisions on licensing datasets for re-use and charging re-use fees.

Licensing re-use under FOIA

64. In order to license a relevant copyright work for re-use, the public authority 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, the public authority 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.

65. Although the duty in section 11A(2) to make datasets available for re-use under a specified licence only applies where the dataset is a relevant copyright work and the public authority is the only owner of that work, the duty under section 11(1A) to make datasets available in a re-usable form in response to a request is not limited in this way. This means that a public authority still has to make a dataset available in a re-usable form in response to a request, even if it is not required to license re-use of that dataset under these provisions. Providing a dataset in a re-usable form, rather than for example a proprietary format, may at least make it easier for people to read that information, even though it is not licensed for re-use. If the public authority is not licensing re-use under these provisions it should explain this to the requester.
66. There are occasions, however, where there is no duty on the public authority to make a dataset available for re-use under section 11A(2) but it is still important that they inform the requester that the dataset (which must in any event be provided in a re-usable format) may still be re-used. For example, the duty in section 11A(2) would not apply to public authorities (such as government departments) whose datasets are 'relevant Crown works', but such datasets would already be available for re-use by authority of the Controller of HMSO. There may also be datasets which are not subject to any copyright, in which case there would be no restrictions on the requester's re-use of the dataset and the public authority should inform the requester of this.

Specified licence

67. Section 11A(2) says that public authorities must make the dataset available under a specified licence. According to section 11A(8), the specified licence is one that is specified by the Secretary of State in a code of practice issued under section 45. 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. It is also the default licence for Crown copyright works.

68. The OGL does not require any registration or charge and it allows re-use for both commercial and non-commercial purposes. In order to use the OGL public authorities should include a statement with the dataset to show that it is being made available under the OGL, as well as a link to the OGL and an attribution statement which those using the dataset should cite in their products and applications.
69. The OGL is the default licence for public authorities to use for datasets that can be re-used without charge. 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. The Charged Licence has been developed for situations where the public authority makes a charge for cost recovery and a reasonable rate of return for use and re-use of the information. A public authority will need to have a good reason for making a dataset available for re-use under a licence other than the OGL.
70. Public authorities should consult the [UK Government Licensing Framework](#) for detailed information on the OGL and how to use it, and for information on the other licences.

Fees

71. In considering whether a public authority can charge a fee for providing a dataset it is necessary to consider both,
 - any fee under section 9 of FOIA for providing the information; and
 - any fee for making the dataset available for re-use under a specified licence.
72. A request for a dataset is a request for information under FOIA and a public authority may charge a fee for this, as for any other request, in accordance with section 9 of FOIA and regulation 6 of the [Freedom of Information and Data Protection \(Appropriate Limit and Fees\) Regulations 2004](#) ("the Fees

Regulations”). Under regulation 6(3) of the Fees Regulations, the costs that the public authority may take into account in calculating the fee include the costs of complying with any obligation under section 11(1) of FOIA as to the means or form of communicating the information. This does not include the cost of providing a dataset in a re-usable form, since the obligation to do this is in section 11(1A), not section 11(1).

73. As explained in the section on [Licensing re-use: FOIA or RPSI?](#), if the dataset is covered by RPSI, then the public authority must license it for re-use under RPSI. In that case, any fee which it charges for allowing re-use must be determined according to the charging provisions in regulation 15 of RPSI. Alternatively, if the dataset is not covered by RPSI, then it must be licensed for re-use under FOIA and any re-use fee must be determined according to the FOIA dataset provisions.
74. Section 11A(3) says:

(3) The public authority may exercise any power that it has by virtue of regulations under section 11B to charge a fee in connection with making the relevant copyright work available for re-use in accordance with subsection (2).

75. The regulations made under section 11B give public authorities a statutory power to charge a fee for making a dataset available for re-use. However, the [Code of Practice](#) on datasets encourages public authorities to make their datasets available for re-use under the OGL, and there is no charge for re-use under this licence. When public authorities make datasets available for re-use under the OGL, they will not make a charge under the regulations made under section 11B or other statutory power.
76. Section 11B says that the Secretary of State may issue fees regulations “in connection with making relevant copyright works available for re-use under section 11A(2) ...” and specifies the scope of these regulations. The regulations referred to in section 11B are the Freedom of Information (Release of Datasets for Re-use) (Fees) Regulations 2013 No. 1977 (referred to here as the “Datasets Re-use Fees Regulations”). They do not replace the Fees Regulations. They deal with fees for making datasets that are relevant copyright works available for re-use, whereas the Fees Regulations deal with fees for communicating the information to the requester.

77. Regulation 2 of the Datasets Re-use Fees Regulations set out how the re-use fee is to be calculated:

2 ...

(2) The total fee shall not exceed the sum of—

- (a) the cost of collection, production, reproduction and dissemination of the relevant copyright work; and
- (b) a reasonable return on investment.

(3) A fee shall be determined, so far as is reasonably practicable,—

- (a) in accordance with the accounting principles applicable to the public authority from time to time, and
- (b) on the basis of a reasonable estimate of the demand for a relevant copyright work over the appropriate accounting period.

(4) A fee charged by a public authority for a relevant copyright work shall not include the cost of an activity mentioned in paragraph (2)(a) if that cost has been included in any other fee charged under the Act to the same applicant by that authority in respect of that work...

78. In a situation where a public authority is charging a fee for re-use under FOIA, then, so far as is reasonably practicable, it must calculate the charge according to the relevant accounting principles and a reasonable estimate of demand. Furthermore, the total fee cannot exceed the cost of collecting, producing, reproducing and disseminating the dataset, plus a reasonable return on investment. These considerations will mainly be relevant to public authorities that operate on a commercial basis in order to cover their costs.
79. The Datasets Re-use Fees Regulations also say that a public authority shall, so far as reasonably practicable, establish standard fees for re-use. In any case it must explain the basis for the re-use fee if the requester asks it to do so.
80. The effect of regulation 2(4) is to prevent “double charging”. If the public authority is charging the requester a fee under regulation 6 of the Fees Regulations, it cannot charge the

requester for the same expenditure under the Datasets Re-use Fees Regulations as well.

81. Public authorities may have powers under other legislation (ie other than FOIA) to charge a fee for making relevant copyright works available for re-use, and section 11A(4) preserves these powers. Furthermore, regulation 2(8) of the Datasets Re-use Fees Regulations makes clear that if a public authority has the power to charge a fee for re-use under other legislation, it cannot charge a fee under these Regulations; it must use its power under the other legislation instead.
82. It is expected that in most cases public authorities will use the OGL to make datasets available for re-use, in which case the issue of making a charge for re-use under section 11B will not arise. If the public authority is going to make a charge for re-use, it should first check whether it has powers under other legislation to do so. If not, it can make a charge under the Datasets Re-use Fees Regulations. It must give the requester a re-use fees notice, and it is not obliged to provide the dataset for re-use until the fee has been received. If the public authority is also issuing a fees notice under section 9 of FOIA (ie a fee for communicating the information in response to the FOIA request), then it may combine this with the re-use fees notice. The provisions regarding fees notices are in section 11A(5)-(7) of FOIA.

Publication schemes

83. In addition to the duty to make datasets available in response to requests, there are also new provisions regarding datasets in section 19 of FOIA, which deals with publication schemes. In general terms, the public authority's publication scheme must include a requirement to publish datasets and make them available for re-use:

(2A) A publication scheme must, in particular, include a requirement for the public authority concerned—

(a) to publish—

(i) any dataset held by the authority in relation to which a person makes a request for information to the authority, and

(ii) any up-dated version held by the authority of such a

dataset,

unless the authority is satisfied that it is not appropriate for the dataset to be published,

(b) where reasonably practicable, to publish any dataset the authority publishes by virtue of paragraph (a) in an electronic form which is capable of re-use,

(c) subject to subsections 2(AA) and 2(AB), where any information in a dataset published by virtue of paragraph (a) is a relevant copyright work in relation to which the authority is the only owner, to make the information available for re-use in accordance with the terms of the specified licence.

(2AA) If the whole of the relevant copyright work is a document to which the Re-use of Public Sector Information Regulations 2015 apply, subsections (2A)(c) and (2B) to (2F) do not apply to the relevant copyright work.

(2AB) If part of the relevant copyright work is a document to which those Regulations apply—

(a) subsections (2A)(c) and (2B) to (2F) do not apply to that part, but

(b) those provisions do apply to the part to which the Regulations do not apply (and references in the following provisions of this section to the relevant copyright work are to be read as references to that part).

84. Public authorities have a duty under section 19(1)(b) of FOIA to publish information in accordance with their publication scheme. The requirement in section 19(2A)(a) to include datasets is in line with the Open Data aim of encouraging public authorities to make datasets routinely available, rather than only on request.
85. The scope of this requirement is potentially very wide. Under section 19(2A)(a)(i) it applies not only to datasets that the public authority has actually released under FOIA, but to any dataset that has been requested. Furthermore, the public authority must publish any updated version of the dataset that it actually holds. However, the requirement is also qualified in that the public authority does not have to publish a dataset as

part of its publication scheme if it is satisfied that it is “not appropriate” to do so.

Not appropriate

86. The public authority therefore has to decide whether or not it is appropriate to publish a requested dataset and any updated versions. The term “not appropriate” is not defined in the dataset provisions, but there are a number of factors that may assist in making this decision.
87. As noted above, the requirement extends to datasets that have been requested, not just to those that have been released. In some cases the public authority may not have released the information requested because it was exempt from disclosure under one of the exemptions in FOIA. This would be a good reason for concluding that it would not be appropriate to publish it under the publication scheme, if the circumstances that led to the application of exemptions previously remain the same. The public authority should nevertheless be aware that a decision not to disclose information in response to a request is a decision taken at a point in time, and circumstances may change such that the information can be published in the future.
88. Given the wide range of requests that a public authority can receive, there may be cases where a request covers a very narrow area of information. While the requester is entitled to ask for this and to receive it unless it is exempt under FOIA, there may be no real benefit, in terms of promoting transparency, in continuing to publish the dataset routinely. In particular this may be the case if the information would give a partial or even misleading impression and the public authority routinely publishes more comprehensive and informative datasets. On the other hand, if the authority frequently receives requests for a particular dataset, this may indicate that it is appropriate to make it routinely available.
89. The cost or the technical issues involved in making the dataset available under a publication scheme could also be relevant to whether it is appropriate to do so.
90. In deciding whether it is appropriate to publish a dataset under its publication scheme, the public authority should also take account of the public interest in transparency and accountability. This is recognised in section 19(3) of FOIA:

(3) In adopting or reviewing a publication scheme, a public authority shall have regard to the public interest—

(a) in allowing public access to information held by the authority, and

(b) in the publication of reasons for decisions made by the authority.

91. 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 public authority 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 public authority actually holds an updated version of the dataset that it has published, 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; this is determined by what the public authority actually holds. The requirement to publish updated versions is also limited by the “not appropriate” qualification, as the following example shows.

Example

A public authority holds both annual and monthly figures relating to a service it provides. It routinely publishes a dataset showing the annual figures. It receives a FOIA request for the latest five month’s figures. It makes a dataset containing these figures available to the requester in a re-usable form under the terms of the specified licence, in accordance with sections 11(1A) and 11(A)(2).

The public authority then has to decide whether to publish the five month’s figures and updated versions of them under its publication scheme, in accordance with section 19(2A)(a).

The public authority may decide that it is not appropriate to keep on publishing the five month’s figures since it will continue to publish annual figures under its publication scheme.

92. When publishing datasets as part of a publication scheme, public authorities are required to publish them in a re-usable form where reasonably practicable.
93. If the dataset in the publication scheme is covered by RPSI, then licensing it for re-use and any re-use charges must be dealt with under RPSI. This is the effect of sections 19(2AA) and (2AB). If it is not covered by RPSI, then it must be dealt with under FOIA, under the terms of section 19(2A): the public authority must make datasets that are relevant copyright works available under the terms of a specified licence, and it may charge a fee for re-use of datasets in a publication scheme. The provisions about licensing and fees in section 19(2A) reflect those in section 11A, relating to making datasets available in response to requests. The defined terms in section 19 (eg "relevant copyright work") have the same meaning as in section 11A.

Model publication scheme

94. The dataset provisions in section 19 are framed in terms of publishing datasets that have been requested under FOIA. They do not make any reference to other datasets that have not been requested or introduce any requirement as to what other datasets should be routinely published.
95. Most (if not all) public authorities will use our Model Publication Scheme in order to meet the duty under section 19(1) to have a publication scheme. This is supported by a series of definition documents for the main sectors and template guides to information for smaller authorities. These set out the types of information that we would expect public authorities to publish routinely.
96. Our [Guide to FOI](#) includes a section on 'What information do we need to publish?' which gives more information about publication schemes and links to the Model Scheme, the definitions documents and the template guides.

Complaints

97. The Information Commissioner deals with complaints under both FOIA and RPSI.

98. 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, 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.
99. RPSI requires public sector bodies to have an internal complaints procedure. In a case where RPSI applies, if the requester considers that the public authority has not complied with the requirements of RPSI, they should complain to the public sector body. If the requester is not satisfied with the outcome they can complain to the Information Commissioner
100. The Information Commissioner will deal with complaints under FOIA or RPSI as appropriate. Further information on how we deal with complaints is available in our [Guide to FOI](#) and our [Guide to RPSI](#).

Environmental information

101. The dataset provisions relate to FOIA, not to the Environmental Information Regulations (EIR). There are no corresponding amendments to the EIR. Public authorities should be aware of this, given that many of the datasets that requesters are likely to want to re-use will contain environmental information. However, if a public authority receives a request for an environmental dataset, and the requester has asked for it in a re-usable form, the authority should consider its duties under regulation 6 of the EIR relating to making information available in a particular form or format. The Commissioner considers that the term "a particular form or format" in regulation 6(1) of the EIR can encompass a re-usable form. Public authorities should make the environmental dataset available in a re-usable form, so far as reasonably practicable. Although there is no duty to license re-use under the EIR, we would also encourage public authorities to make the dataset available under a licence that permits re-use.
102. Information that is accessible under the EIR may also be the subject of a re-use request under RPSI, if the public authority is covered by RPSI.

103. Public authorities should also be aware of their duties to make spatial data available under the [INSPIRE regulations](#).

Practical issues

104. The dataset provisions do not provide any additional right to obtain information that is not otherwise accessible under FOIA, nor do they add any exemptions. They concern the form in which information is made available, both in response to requests and through a publication scheme, and the conditions for re-use.

Requests

105. When a public authority receives a request for information that it holds in the form of a dataset, it should first consider whether any of the information is exempt from disclosure, as it would with any other FOIA request. If the information is exempt, then the FOIA duties relating to datasets do not arise. Guidance on all the FOIA exemptions is available in our [guidance index](#).

106. If the information is not exempt from disclosure, then the public authority should check whether it is or forms part of a dataset as defined in section 11(5). If it does not meet that definition, then the FOIA duties relating to datasets do not arise. If it is a dataset, FOIA requires public authorities to provide it in a form capable of re-use.

Licensing re-use

107. The public authority's obligations to do with licensing re-use (and charging a re-use fee) depend on whether the public authority and the dataset are covered by RPSI. If the public authority is also a public sector body under RPSI⁴ and the dataset is not excluded from RPSI⁵, then licensing and re-use fees are dealt with under RPSI. If the public authority is not a RPSI public sector body, or it is covered but the dataset is excluded from RPSI, then licensing and re-use fees are dealt with under FOIA.

⁴ Public sector bodies are specified in regulation 3 of the Re-use of Public Sector Information Regulations 2015 S.I. 2015 no. 1415

⁵ The types of information excluded from RPSI are listed in regulation 5 of the Re-use of Public Sector Information Regulations 2015 S.I. 2015 no. 1415

108. It is therefore important that public authorities establish whether they are a public sector body under RPSI, and if so what information they hold that is covered by RPSI.
109. The easiest way for a public authority to comply with the licensing requirements of both FOIA and RPSI is to make information available for re-use under the Open Government Licence, where appropriate.

Personal data

110. Given that many datasets will contain personal data, public authorities should in particular take account of the exemption in section 40(2) of FOIA, where disclosure of personal data would contravene DPA principles. Public authorities can redact personal data that is exempt under section 40(2). However, as more datasets are released for re-use, the possibilities of analysing the dataset, for example by a pivot table, and of combining datasets from different sources and publishing the results will increase. This is part of the rationale behind the open data agenda, of which the dataset provisions are a part. However, this also increases the possibility that individuals can be identified from analysing and combining apparently anonymous datasets. It is difficult for public authorities to know what other information will be available, but they should be aware that there may be an increased risk of 're-identification'. Our [Anonymisation: managing data protection risk code of practice](#) addresses this problem and provides advice to data controllers, including public authorities.

Copyright and licences

111. The duty to make relevant copyright works available for re-use under the terms of a specified licence means that public authorities must be able to establish whether their datasets meet the definition of a relevant copyright work. In particular, they need to know whether they own the intellectual property in the datasets they hold, since they cannot license re-use if they are not the only owner. This suggests that public authorities should carry out an audit of copyright and database rights, in particular in relation to the datasets they consider are most likely to be requested. In practice it may be difficult for public authorities to know who owns the intellectual property rights in all of their information, but if the public authority uses an electronic document and records management system

(EDRMS) it may be possible to extract this information from the metadata.

More information

112. This guidance has been developed drawing on ICO experience. Because of this it may provide more detail on issues that are often referred to the Information Commissioner than on those we rarely see. The guidance will be reviewed and considered from time to time in line with new decisions of the Information Commissioner, Tribunals and courts.
113. It is a guide to our general recommended approach, although individual cases will always be decided on the basis of their particular circumstances.
114. If public authorities need any more information about this or any other aspect of freedom of information, they should contact us; see our website www.ico.org.uk.

Annex 1: Datasets flowchart

This flowchart shows the decisions and the steps which the public authority must take when releasing information that it holds as a dataset.

It is intended to provide practical guidance for public authorities in applying the dataset provisions. It supplements this guidance document. It is not possible to give a full explanation of the process within the flowchart itself. Public authorities should read the guidance document for definitions of the terms used here and a full explanation of the dataset provisions.

The starting point for the flowchart is that the public authority has received a request and has identified information which is not exempt and which it is going to release under FOIA.

