



National Archives of the Netherlands
Ministry of Education, Culture and Science

Open Collection Data Policy at the National Archives of the Netherlands

Version 2.0
Date March 2017



This document is made available with a [Creative Commons Public Domain Mark](https://creativecommons.org/licenses/publicdomain/0.0/)

Version history

Version no.	Date	Status	Discussion	Contents
1.0	May 2015	Adopted in DR NA	DR NA	Previously adopted version. Incorporated by all RHCs.
2.0	January 2017			Policy evaluated and text fundamentally revised. Remarks from MT DIA and MT C&P included
2.0	March 2017	Adopted in DR NA		Evaluated and adopted version

Contents

	Version history	3
1	Executive summary	7
2	Introduction	8
2.1	Purpose of the policy	8
2.2	Scope	8
2.3	Strategic relationships with RHCs	9
2.4	Overview document	9
3	When is data open?	10
3.1	What is open data?	10
3.2	Open government data	10
3.3	Open vs public	11
3.4	Accessible and reusable	11
3.5	How open is open?	11
4	Legal and regulatory framework	12
4.1	Introduction	12
4.2	Copyright	12
4.3	Directive for reuse	12
4.4	Relationship Public Records Act and Reuse of Public Sector Information Act	12
4.5	Applicable sections from the Reuse of Public Sector Information Act	13
4.6	Interpretation of legislation and regulation	14
5	Vision and choices of the National Archives	17
5.1	Vision	17
5.2	The National Archives' choices	17
6	The National Archives' policy	19
6.1	Policy principles	19
6.2	Roles and responsibilities	20
6.3	Version management and revision process	20
7	Impact of this policy	21
8	Appendix: Overview of Creative Commons licence forms	22
8.1	Introduction	22
8.2	Features	22
8.3	The public domain (CCo and PDM)	23
9	Open Data terms A-Z	24
10	Sources	28

1 Executive summary

The National Archives has made metadata (such as inventories) and data (such as scans) from the collection available as open data for some time.

By determining policy, the National Archives is able to shape the processes and procedures for the open collection data. This policy document sets a framework. In this way, the National Archives can be accountable to internal and external stakeholders (directors, employees, customers, partners and citizens) with regard to open data and can also assign responsibilities within the organization. This also contributes to creating a support base for open collection data both within and beyond the National Archives.

For an explanation of the term open data, see subsection 3.1 What is open data?

For a vision on open data and the benefits of open data, see subsection 5.1 Vision.

For the principles of the Open Collection Data Policy of The National Archives, see section 6 The National Archives' policy.

For the impact of this policy on the National Archives, see section 7 Impact of this policy.

2 Introduction

More and more governmental and cultural institutions are making their information available in digital form. This creates new possibilities – for the institutions themselves but also for third parties – to use this information to develop new applications and websites, so that the general public is able to participate in new ways. In addition to making (presenting) information digitally available, it can also be made available as open data. Open data allows digital information to be reused without restrictions.

2.1 Purpose of the policy

The National Archives has made metadata (such as inventories) and data (such as scans) from the collection available as open data for some time.

By determining policy, the National Archives is able to shape the processes and procedures for the open collection data. This policy document sets a framework. In this way, the National Archives can be accountable to internal and external stakeholders (directors, employees, customers, partners and citizens) with regard to open data and can also assign responsibilities within the organization. This also contributes to creating a support base for open collection data both within and beyond the National Archives.

2.2 Scope

This document comprises the Open Data Policy part 1: Open Collection Data Policy, i.e. policy principles for:

- Metadata in the form of entries (EAD) and added entries, such as databases on the collection;
- Archive material: digital born or digitized, such as photos, maps;
- Metadata about legal caretakers, as stored in the Actors' register.

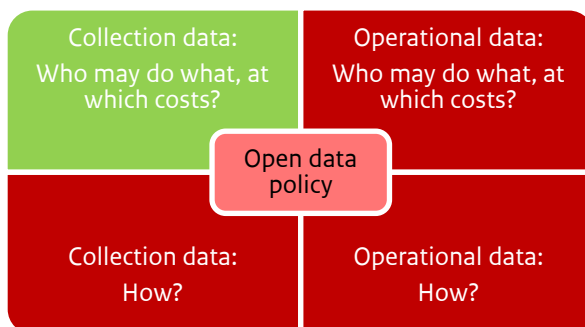
So-called management information falls outside the scope of this document. This includes visitor statistics, financial information and information about policy and the implementation of programmes and projects.

In addition, a distinction can be made between

- basic principles for policy and
- the practical or technical result and scope of these principles.

This document only addresses the basic principles for policy about ‘who may do what with the data and under which conditions’?

The scope of this framework may be visualized as follows:



The practical or technical result and scope of these principles will be further detailed later. This document does therefore not yet discuss Linked Open Data (as the technological ‘pinnacle of openness’). Nor does it discuss the National Archives’ expenses for making open data available. This is because the costs are closely related to the technical way in which it is provided.

2.3 Strategic relationships with RHCs

The Regional Historical Centres (RHCs) are important partners of the National Archives. Ideally, the RHCs and the National Archives share the vision on open data regarding the national collection. A shared vision means that the ‘archive consumer’ is not faced with different regimes. In order to develop a shared infrastructure, technical components must be shared more easily. The open collection data policy of the National Archives will therefore be discussed with the RHCs.

2.4 Overview document

Section 1 consists of the Executive summary.

Section 2 Introduction describes the purpose and scope of this document, the connections to other policies and the purposes of the policy.

Section 3 When is data open? provides a further explanation of the term open data.

Section 4 Legal and regulatory framework is concerned with the main legal framework into which the open collection data policy should fit.

Section 5 Vision and choices of the National Archives reflects the National Archives’ vision on open data and the choices made within the legal framework.

Section 6 The National Archives’ policy comprises the adopted policy and provides the basis principles for dealing with open collection data.

The scope of this policy is described in section 7 Impact of this policy.

The appendices to this document clarify the various Creative Commons licencing forms and a number of terms. Appendix 10 contains an overview of the sources consulted.

3 When is data open?

3.1 What is open data?

Different definitions are used for the term 'open data'. The following is the most general: Open data is information that is readily available, accessible and reusable for everyone.

3.2 Open government data

The Dutch government states¹:

Open Data at the government is data that is both *accessible* and *reusable*. This means this is data that:

- Is paid for and generated from public funds with and for the performance of the public task;
- Is public;
- Is free of barriers such as the obligation to register oneself, costs or user restrictions;
- Is preferably computer readable and complies with the 'open standards' (no PDF, but XML or CSV);
- Is free of royalty such as copyright or other third-party rights.

Figure 1 Open data is accessible and reusable²

Accessible

Public task



Public



Free of barriers



Reusable

Computer readable



Royalty-free



When performing its public tasks, the government 'creates' and uses data. A large part of this data is public, which means that it is readily accessible to citizens, on the basis of the Government Information (Public Access) Act (*Wet Openbaarheid van bestuur*) or other legal regulation. Some ways of sharing sensitive information with citizens in the case of non-public data are to make it anonymous, create pseudonyms or aggregate it.

¹ National Audit Office, Trend report Open Data 2015 (adopted 30 March 2015, presented to the House of Representatives on 31 March 2015) p. 6.

² Original image source: National Audit Office, Trend report Open Data 2015 p. 6.

The essence of open data is that users can easily find the data and that it is reusable without any restrictions or conditions which must be met before the data can be used. Every user must be able to reuse the data without any limitations.

3-3 Open vs public

Public is not the same as open. For data to be labelled open it must be both accessible and reusable. Open is therefore a broader term than public³. The public nature of data is a prerequisite to making data available as open data.

3-4 Accessible and reusable⁴

Access to government data is the result of data collection for a public task (the data exists), the right to public access (I have access) and the availability on the internet without any barriers (I can reach it easily). Reusability of data is the right to freely use, copy, process and combine data with other data. This requires the data to be free from copyright and be presented in the form of a computer-readable file. The first guarantees a right to reuse (I am permitted to reuse it), the second the possibility to actually do so (I am able to reuse it). A PDF format is an example of a non-computer readable format because a user must copy the data by hand to another file in order to process it. In contrast, a CSV file⁵ is computer readable, which means that the data can immediately be processed.

3-5 How open is open?

The fact that data is open does not mean that reuse will be easy. Having to download documents one by one is cumbersome for example; it is therefore easier to present them as a bundle. If a data set is open but cannot be found, the information is not accessible, just as in a non-inventoried archive.

We, the National Archives, should therefore not only ask ourselves which data can and which cannot be open but also how easy we want to make it for the user to reuse this data. Facilitating reuse contributes to the actual use of data.

That means making agreements regarding technology and standards. A well-known example is the 5-star system devised by Tim Berners-Lee, inventor and founder of the World Wide Web, which provides a classification for the ease with which the user is able to reuse data:

★	Make your data available on the Web (in whatever format) under an open licence
★★	Make it available as structured data (such as Excel instead of an image of a table)
★★★	Use open, non-proprietary formats (publication formats that are not dependent on specific software, e.g. CSV instead of Excel)
★★★★	Use URIs (unique identifiers) to identify data so that people can point at your data
★★★★★	Link your data to other data to provide context

³ National Audit Office, Trend report Open Data 2015 p. 6.

⁴ Text cited from: National Audit Office, Trend report Open Data 2015 p. 6.

⁵ See list of terms.

4 Legal and regulatory framework

4.1 Introduction

This section is concerned with the legislation and regulation that provide a framework for an open data policy and addresses what this framework is. On occasion the legislation and regulations are ambiguous and interpretation is required.

4.2 Copyright

If government archive records are transferred to a records repository and there is copyright on these records, this copyright is transferred along with them. Copyright is transferable and can be given under licence. It is not possible to cancel, retract or allow it to expire. If the National Archives itself has copyright, it can unilaterally declare that it will not use its copyright against anyone. That means that others can hold the National Archives to that. Although the copyright is not implemented, the document remains a work protected by copyright.

When private archives are transferred/donated, the owner/copyright proprietor can reserve the copyright. In that case, the National Archives may not unilaterally state that it will not use its copyright against anyone.

4.3 Directive for reuse

On 13 June 2013, the European Parliament adopted the amendment of Directive 2003/98/EC regarding the reuse of public sector information. This amendment is intended to stimulate open data. The new directive obliges the Member States to make all public sector information available for reuse. An exception is made for information under legal restrictions or rights of third parties (such as privacy or copyright). The directive also applies to libraries, museums and archives.

A European directive must be implemented in national legislation and regulations. In the Netherlands, the Reuse of Public Sector Information Act was drafted for this purpose. The Reuse of Public Sector Information Act does *not* apply to archives. The underlying idea was that the Public Records Act 1995 (footnote) offers a sufficient framework. Archives are already tasked with making government information public.

However, the situation for archives has changed. The Public Records Act has been adapted in some areas to guarantee that the use of archive records as public sector information is in line with the Reuse of Public Sector Information Act. Through these adaptations, some sections from the Reuse of Public Sector Information Act *have* been declared applicable to archives.

4.4 Relationship Public Records Act and Reuse of Public Sector Information Act

The main amendments to the Public Records Act are listed below:

- New Section 2b Public Records Act
 - Subsection 1: 'Use' in the sense of Sections 14 and 17 of this Act [*the Public Records Act is referred to*] is also taken to mean 'reuse' in the sense of Directive 2003/98/EC [...]
 - Subsection 2: The Reuse of Public Sector Information Act does not apply to making records available that have been transferred to a repository for use, unless decreed otherwise by this law [*again taken to mean the Public Records Act*].
- Amended Section 17 Public Records Act (amendment in italics):
 - Subsection 1: The administrator of a repository makes available the archive records kept there to the requesting party to consult or use in compliance with the restrictions set on public access *and in accordance with Sections 5 and 6 of the Reuse of Public Sector Information Act*.

- In Section 19 of the Public Records Act regarding pricing, Section 9 (1) and (4) of the Reuse Act are declared applicable (amendment in italics):
 - If costs as referred to in Section 14 and Section 18 (6) are charged, the legal caretaker shall set rules for those costs. *Section 9 (1) and (4) of the Reuse of Public Sector Information Act applies accordingly to the costs charged.*

In summary: although the Reuse of Public Sector Information Act is in principle not applicable to repositories, some sections from the Act apply by referral from the Public Records Act, as follows:

Public Records Act		Reuse Act
Sect. 17 (use)	>>	Sect. 5 and 6 (reuse)
Sect. 19 (pricing)	>>	Sect. 9 (1) and (4)

4.5

Applicable sections from the Reuse of Public Sector Information Act

By above referral from the Public Records Act, the following sections from the Reuse Act apply to material stored in a repository:

- Section 5 Available formats
 - Subsection 1: The information available for reuse shall be provided in the way the information is present at the institution charged with a public task and where possible electronically, in an open and machine-readable format, together with the metadata, which must where possible comply with the formal open standards [...]
 - Subsection 2: An institution charged with a public task is *not obliged* to continue making and storing documents solely with a view to reuse.
- Section 6 Conditions
 - Subsection 1: The conditions for reuse are equal for similar categories of reuse.
 - Subsection 2: An institution charged with a public task shall not attach to the approval for reuse any licence terms that needlessly restrict the possibilities for reuse or by which competition is restricted.
- Section 9 Pricing
 - Subsection 1: For the reuse of information the institution entrusted with a public task shall charge at most the incremental costs of multiplication, provision and distribution.
 - Subsection 4: The institution entrusted with a public task shall give advance notice for any costs to be charged for reuse and for the factors which should be taken into consideration in calculating these costs. On request the institution shall provide insight into the calculation regarding a specific request for reuse.
- Section 9 Pricing: sect. 9 (2) and (3) are NOT applicable
 - Subsection 2: In derogation from the provisions in subsection 1, museums and libraries shall, for the reuse of information, charge at most the costs incurred for the collection, production, multiplication and distribution, conservation and settlement of rights, increased with a reasonable return on investment.
 - Subsection 3: In derogation from the provisions in subsection 1, an institution entrusted with a public task shall, to cover the costs of implementing the public task, charge for the reuse of information at most the costs incurred for the collection, production, multiplication and distribution, increased with a reasonable return on investment if so stipulated by law. The calculation of the costs takes place on the basis of objective, transparent and verifiable criteria.

The Explanatory Memorandum to the Reuse of Public Sector Information Act indicates that ‘for archives the choice was made to only allow the incremental costs to be charged.

It is not appropriate that the Public Records Act 1995 allows for the possibility that costs for use or reuse of archive records are linked to the production, multiplication, distribution, conservation and settlement of rights, increased with a reasonable return on investments.

In the case of repositories in the sense of the Public Records Act 1995 it also concerns reuse of archive records that in the context of performing the public task have already been produced or received by the public sector with public resources.’

4.6 Interpretation of legislation and regulation

Legislation and regulation does not provide all the answers to questions about what must, should and could be done, by using terms such as ‘equal treatment in equal situations’, ‘needless’ and ‘restrict’. The basic principles below are a legally reviewed interpretation of the legislation and regulation.

4.6.1 The information concerned:

In the case of archive institutions it is best to use the formal term ‘archive records’. Where it concerns archive records, the Reuse of Public Sector Information Act applies, unless one of the grounds for exception arises, such as not public or not royalty free. The National Archive’s photo collection is largely not considered an archive record in the sense of the Public Records Act. That also applies to a part of the private archive, because it has not been donated but ‘deposited’.

4.6.2 Is a private archive considered public sector information?

Yes. Private archives are also considered public sector information when they have become archive records by endowment.

4.6.3 Do I now have to digitize everything?

Information should be made available where possible in machine-readable format at the best level of accuracy. Metadata should also be included. This type of availability is primarily a best-effort obligation. Archives are not obliged to make their entire paper archive records available in digital form.

4.6.4 Is attribution permitted?

The use of licences is permitted, but governments may not impose conditions that obstruct the availability of public sector information. Repositories are included under governments. The obstruction prohibition is taken strictly. Attribution may be requested but only if there are sound reasons and if it is not on a structural basis.⁶

4.6.5 Is a request for registration permitted?

It is not permitted to register parties requesting reuse of data. Registration of requests to provide open data can be considered a condition of use in advance. Recital 26 of Directive 2003/98/EC regarding the reuse of public sector information determines that the conditions may impose the least possible restrictions on reuse. In the situation in the Netherlands this is considered a non-desirable condition of

6 See the Explanatory Memorandum to the Reuse of Public Sector Information Act p. 15: ‘With regard to imposing the condition to reference sources, some organizations consider this advisable because any person can ascertain where the information comes from on which the developed product is based. Other organizations do not reference sources in order to not be associated with random end products for reuse. When referencing sources is a condition, the starting principle is also that this should be exception rather than rule.’

use. 'Free ordering' of high resolution scans is an undesirable way of making this material available as open data because re-users would first have to register⁷.

4.6.6

May costs be charged for the provision of data?

The Directive for reuse allows archival institutions to charge on the incremental costs which they have to incur to comply with a request. Translated to open data this means that only the costs for data may be charged that are demonstrably linked to a request. For data (such as high resolution scans) that is already available, in general no additional costs need to be incurred. Therefore no prices may be included in the pricing regulation for ordering scans that have already been made.⁸ Depreciation costs of e-depots or the costs for data storage may not be counted as incremental costs.

This means the following:

- Metadata (entries, inventories, indices, ...) must be provided free of charge, insofar as they are exempt from rights and are digitally available;
- Collection data (scans of archive records, photos, ...):
 - If data is already available then this data will be made available free of charge in the form in which it is present (this is in the highest resolution). This is in line with what the amendment of Directive 2003/98/EC means with 'making available at the best level of accuracy'. In addition, there are no 'incremental costs' for this material;
 - Digital born material is supplied free of charge.

4.6.7

Are exclusivity agreements possible?

Exclusivity agreements, for example where a market player digitizes material in exchange for the exclusive right to exploit this material, should be avoided where possible. A separate regulation will be drafted for digitizing cultural resources (exclusivity will apply for a maximum of 10 years). However, an exclusivity agreement is not possible in view of Section 14 of the Public Records Act: 'anyone is authorized ... to consult an archive record, make images, etc.'

4.6.8

Summary

Question	Interpretation of legislation and regulation	Implication >> basic policy principles
What exactly is public sector information? Does this include a private archive (transferred to a repository)?	Yes. A private archive is also considered public sector information when it has become an archive record by endowment.	The policy applies to both institutional archives and private archives.
Sect. 6 of the Reuse Act stipulates that licences may not needlessly restrict reuse. Is the obligation to register a restriction?	Yes, the obligation to register is a restriction.	Material must be available where possible without an obligation to register. Registration is only demanded if it is absolutely

⁷ See the Explanatory Memorandum to the Reuse of Public Sector Information Act p. 15: 'Conditions of use that are in no event desirable relate for example to the way in which the supplied information may be used, such as [...] the registration of the re-user prior to providing data.'

⁸ See the Explanatory Memorandum to the Reuse of Public Sector Information Act p. 11: 'Regarding archives, the choice was made to only charge incremental costs. It is not appropriate that the Public Records Act 1995 allows for the possibility that costs for use or reuse of archive records are linked to the production, multiplication, distribution, conservation and investments. Repositories, in the sense of the Public Records Act 1995, are also mainly concerned with the reuse of archive records that have, in the context of performing the public task, already been produced or received by the public sector with public resources.'

		necessary.
And is the requirement of source reference a restriction?	Yes, the requirement of source reference is a restriction.	Material must be available where possible without the obligation to reference sources. Source reference is only demanded if it is absolutely necessary.
May costs be charged for the provision of data?	Only incremental costs, which means costs that are a direct result of the request for availability. That does not include the costs for storage, management, etc.	Two types of costs are conceivable: costs for scanning on demand and costs for data traffic.

5 Vision and choices of the National Archives

5.1 Vision

Open Collection Data contributes to achieving the National Archives' mission: "We serve everyone's right to information and provide insight into our nation's past. To that end, we bring the general public into contact with the collection as much as possible." The National Archives uses open data to make more information available faster, better and appropriate for reuse. The National Archives aims to make as much data in the current collection as possible available as open data, with linked open data being the technological high point. Note: linked open data falls outside the scope of this policy framework. To achieve its aim, the National Archives has made a number of choices which are elaborated on in more detail in subsection 5.2 ff.

Open Collection Data contributes to achieving our mission in a variety of ways:

- By the online presentation on other platforms the collection's reach is enlarged as much as possible. This is demonstrated by visitor statistics: 120,000 visitors a month to our own website compared with 12 million visitors a month to Wikipedia.
- Other research institutes and government institutions can connect directly to the information of the National Archives or download it as larger data sets. Private persons can download an index in one go and search it themselves without being dependent on how the search screen on the public website presents it. Professionals and private persons can use the enormous amount of image material for their publications or projects. Because of this possibility, the National Archives' materials emerge in unexpected locations, such as on Wikipedia.
- On the basis of the metadata of the National Archive's collection, third parties can compile collections (aggregations) of metadata. This allows searches across several collections (such as archives from different archive institutions, or archives and newspapers). One example is the Archives Portal Europe (APE). This allows the National Archive's collection to be found quicker, more easily and more often.
- The open collection data of the National Archives can be used for innovative purposes. By making open collection data available, third parties are able to develop additional channels, such as applications or products. Two examples are *openarchieff.org* and *tijdbalk.nl*
- On *openarchieff.org* heritage institutions with an API collection are automatically sent a tweet several times a day. The National Archives has four twitter streams. Three of them (cycling, football and Anefo) are derived from the Anefo collection. One is generated from the open data map collection 4.VEL.
- *Tijdbalk.nl* is the winner of the National Archive Award in the Open Culture competition. It is used to easily compile a timeline from a number of open culture data sets including the Anefo collection and Open Images from the Netherlands Institute for Sound and Vision. Users can also add their own photos to a timeline through *Flickr.com*

5.2 The National Archives' choices

Taking into account the statutory obligations from the previous section, the National Archives has, on a number of points, made policy decisions that fit in with the vision outlined above. These are explained in more detail below.

5.2.1 Handling third-party data

Material held by the National Archives but which is the property of third parties is treated where possible as material from the own collection, unless other binding agreements apply (such as for the Spaarnestad collection).

5.2.2

Target groups and relation to re-users

- The National Archives does not indicate specific target groups for open data. Citizens, companies, journalists, other governments, etc. and/or intermediaries have access to the same data under the same conditions.
- The National Archives does not maintain specific relationships with re-users, i.e. no different to those for ordinary visitors to the reading room or the website.

5.2.3

Copyright and licences

Government archives may contain material subject to copyright. That could be a report, but also a letter from a citizen or an entry made by a private individual. The National Archives can decide how to deal with copyright held by the National Archives itself.

The following principles were chosen:

- If the copyright on the open collection data is held by the National Archives, it will not implement this right. The data is provided with a CCo waiver. This is in accordance with current practice at the National Archives;
- In the event that copyright is vested in the open data collection, with a third party being the entitled party, but this party has declared not to implement this right, the data will be provided with a CCo waiver. This is in accordance with current practice at the National Archives;
- At the time that a third party possibly wants to implement its right, the data does not meet the definition of open data. It is therefore not advisable to present the data or make it available online for reuse at that time;
- At the time there is no copyright vested in the open collection data of the National Archives, there is no choice and the data will be made available under a Public Domain Marker. This applies to both metadata and data such as scans.

5.2.4

Source references

The basic principle should be that source referencing is the exception rather than the rule. The National Archives has opted to follow the line as stated below:

- The National Archives carefully considers whether conditions (such as attribution) should be set for reuse when making open data available. The basic principle here is that there are as few barriers as possible. That is why a Public Domain Mark is used where possible, or a CCo waiver. That means that the use of licences will be scaled down further.
- If necessary, data is marked with an appropriate 'higher' creative commons licence CC-BY ("attribution") or CC-BY-SA ("share alike", see appendix) so that everyone is able to verify the origin of the information on which the product is based.

5.2.5

Charging on of incremental costs

- The National Archives will not charge the incremental costs resulting from data traffic (when data is downloaded) on to the user. That relieves it from the need to request registration (as no invoices need to be sent).
- Material that is located in the collection in analogue form but is digitized for instance by scanning on demand will be made available at incremental cost (charging on the costs incurred to comply with the request). The rates shall be clearly published on the National Archives' website.

6 The National Archives' policy

6.1 Policy principles

The National Archives' vision drawn up in the previous sections, the obligations arising from law and regulation and the choices made result in the following policy:

General

- Open data where possible or necessary;
 - In acquiring a private archive, agreements must be made about the availability in the form of open data;
- When is it necessary?
 - When it concerns public sector information. Archive records are public sector information;
- When is it possible?
 - Material is free from restrictions:
 - No restrictions on public access;
 - Royalty-free and copyright-free;
 - If there are no non-standard agreements with other parties, such as in the case of the Spaarnestad collection.

Target groups and relation to re-users

- The National Archives does not distinguish in type of re-user (nor between commercial and non-commercial use);
- The National Archives does not enter into a specific relation with re-users or groups of re-users.

Pricing

- Metadata (entries, inventories, indices, ...) are provided free of charge, insofar as they are royalty-free;
- Collection data (scans of archive records, photos, ...) is provided
 - In the highest resolution available;
 - Digital born material is supplied free of charge;
 - Material that is the property of third parties is treated where possible as material from the own collection, unless other binding agreements apply.
- Material that is located in the collection in analogue form but is digitized for instance by scanning on demand will be made available at incremental cost (charging on the costs incurred to comply with the request).

Copyright and licences

- Is there no copyright? Then it is made available with a Public Domain Marker. This applies to both metadata and data such as scans.
- Is there copyright on the open collection data held by the National Archives? This right should not be implemented: make it available with a CCo waiver.
- Does a third party have copyright, but has it stated it won't implement its right? It should be made available with a CCo waiver.
- Does a third party have copyright and does it possibly want to implement its right? It is not open data.

Source references

- The National Archives sees source referencing as an obstacle that should be imposed as little as possible;
- To be considered on a case-by-case basis. If it is really necessary, use the appropriate Creative Commons Licence (CC-BY or CC-BY-SA);
- In all other cases open collection data should be made available under a Public Domain Mark or CCo waiver.

6.2 Roles and responsibilities

- The management team approves the open data policy on advice of the members of the management team.
- The Chief Information Officer initiates the development and implementation of the open data policy and reports to the management team as and when necessary.
- Line managers implement the open data policy where appropriate in their field of responsibility and issue advice to the Chief Information Officer about the impact of the open data policy on their fields of responsibility.

6.3 Version management and revision process

6.3.1 Version management

The open collection data policy is evaluated and where necessary revised annually or as frequently as the Chief Information Officer deems appropriate.

6.3.2 Revision process

After a round of revisions the management team will re-adopt the open collection data policy.

7 Impact of this policy

This document partly provides the rationale for what is already current practice (such as making our entries available). In addition, a number of amendments are proposed as a result of changing legislation and regulation. The main implications are:

- No costs will be charged on to the user for data that is already available and has already been financed. Existing high resolution scans should be made available as open data. The archive records of the National Archives which are digitized according to plan (e.g. in work package 3 of the Digital Tasks National Archives programme) are made available as open data, free of charge and in the highest available resolution. Only costs demonstrably related to a request (such as for scanning on demand) may be charged as incremental costs on to the user. Processes and systems must be set up to include this option.
- Licence references in the entries, indices and data (scans) must be checked and adapted in many places (e.g. from CC-BY to CC-o or to Public Domain Mark). The basis principle is that the use of licences will be scaled down further where possible.
- Adaptation to process of acquiring private archives. Private archives are treated in much the same way as the national collection in accordance with this policy. When transferring a private archive, priority must be given to the National Archives' policy to make the archive (metadata and data) available as open data (obviously subject to publication restrictions). Agreements must be reached with the transferring party about this.
- Financial consequences: because the National Archives provides free downloads there is a loss of income as is shown in the 2014 overview below.⁹

2014 information		
Source:	Number of sales	Total revenue in Euros
archive/image bank/download	1176	1176
photo/image bank/download	815	4082.5
photo/image bank/print	10	117.5
photo/photo on demand/download	277	3047
maps/image bank/download	365	1825
Grand total	2643	10,248

Disclaimer: this English version is a translation of the original in Dutch for information purposes only. In case of a discrepancy, the Dutch original will prevail.

⁹ With thanks to the Association for the National Archives.

8 Appendix: Overview of Creative Commons licence forms¹⁰

8.1 Introduction

Creative Commons (CC) is originally an American project that promotes open content. It aims to make creative works more freely available than is possible under traditional copyright, so that these works may be copied and distributed more easily or that others can work on them further. The project offers various free licences that copyright holders can use when distributing information to prevent problems that could arise under current copyright legislation.

Creative Commons comprises a number of international projects to translate the licences from English and American law to other languages and legal systems.

Creative Commons uses features to characterize its licences because it has ensured that the licences are easy to understand for both people (the Commons Deed) and machines (the metadata) and that the licences are enforceable in court (the legal code, the actual licence).

8.2 Features

The following four features are used in combinations and form the different Creative Commons licences.

- Attribution (BY) means that copying, distributing, showing and performing the work and derived works is permitted on the condition that the original author is named.
- Non-commercial (NC) means that copying, distributing, showing and performing the work and derived works is permitted for non-commercial purposes.
- No Derivate Works (ND) means that copying, distributing, showing and performing the work is permitted but not changes to the work.
- Share Alike (SA) means that distributing derivated works is only permitted under an identical licence.

Code	Name Title
CC-BY	Attribution
CC-BY-SA	Attribution, share alike
CC-BY-ND	Attribution, no derivate works
CC-BY-NC	Attribution, no commercial works
CC-BY-NC-SA	Attribution, no commercial works, share alike
CC-BY-NC-ND	Attribution, no commercial works, no derivate works

In addition to the different CC licences, there is also CCo. CCo is intended for copyright owners to release work into the public domain, where certain legal systems effectively prevent the transfer of rights from an owner, even if the author wishes to.

¹⁰ Source: <http://creativecommons.org/licenses/> and http://nl.wikipedia.org/wiki/Creative_Commons

8.3 The public domain (CCo and PDM)¹¹

An explanation of the public domain and the two tools offered by Creative Commons to contribute to it can be found below.

- [the public domain](#)
- [the CCo public domain statement](#)
- [the Public Domain Mark](#)

8.3.1 *The public domain*

The public domain is a term used to indicate that works no longer fall under the protection of copyright and related rights. There is no entitled party in this situation and therefore no approval is required to use the works. Most works in Europe ‘fall’ in the public domain on 1 January of the 71st year after the death of the creator, in light of the fact that copyright generally expires 70 years after the death of the creator. Strictly speaking, it is not possible in the Netherlands to place works directly in the public domain before the copyright has expired. According to the Dutch Copyright Act it is not possible to relinquish *all* copyrights, neighbouring rights or databank rights. The entitled party can, however, declare it will not exercise its rights in any way. In practice this means that the works enter the public domain. Creative Commons is a great advocate of this possibility and has therefore drafted a “[CCo public domain statement](#)”, making this way of relinquishing rights easier. In addition, Creative Commons developed the [Public Domain Mark](#) to mark works that are no longer protected by copyright or related rights. These works no longer require a licence; the mark is designed to easily recognize these works.

8.3.2 *The CCo public domain statement*

CCo is not a licence, but a document with which the party entitled to copyright can indicate whether he or she relinquishes all copyright. In contrast to a licence, no infringement can be made on a work made available with a CCo statement. By declaring CCo applicable to a work, this work is placed in the public domain by relinquishing all rights in as far as permitted under law. CCo is thus intended for works that are still protected by copyright but where the entitled party wishes to relinquish these rights.

Works made available under CCo can therefore be used by everyone for all purposes and attribution is not an obligation.

The [legal text](#) of the CCo public domain statement has been translated into Dutch linguistically, which means that the text has not been adapted to specific Dutch legislation.



8.3.3 *The Public Domain Mark*

The Public Domain Mark is not a licence, but a method to mark works that are no longer protected by copyright or that cannot be protected by copyright. Many cultural institutions such as museums, libraries and archives are curators of works in the public domain. This mark makes it easy to class them as public domain works. These works are more recognizable and traceable because of the Creative Commons’ technological infrastructure. The Public Domain Mark is intended for works that are not or no longer protected by copyright and related rights.

Public domain marked works may therefore be used by everyone for all purposes. However, we want to point out that in some countries moral rights are infinite.



¹¹ Source: <http://creativecommons.nl/publiek-domein/>

Open Data terms A-Z¹²

Term	Further details
API	An API (Application Programming Interface) gives people and computers online access to a database so that the data can be used directly for apps and websites. When the data in the database changes, it automatically changes in the linked website or app. In this way all data is always identical and up to date. For example, Dutch Railways makes its database with up-to-date travel information available online through an API. Different apps retrieve this information directly from the API and so provide the option for travellers to plan their trip on the basis of up-to-date information.
APP	An app (or application) is a software program that is used by a smartphone, tablet or computer. Apps can be downloaded or used via a website. It is easy to add additional functions to a device using an app. For example, a phone can indicate if it will rain (e.g. using the weather radar), or a tablet can edit photos (e.g. using iPhoto).
Big Data	The term big data indicates data sets that are so large that the average database software is unable to retrieve, store, manage and analyse the data. Because big data comes in such large volumes and in addition is highly complex, it requires a cost-effective and innovative way of information processing. In light of the rapid technological developments, fewer data sets will qualify as big data, because database software is now able to handle increasing amounts of data and this capacity will only continue to grow in future.
Source code	The source code (or source text) of a computer program is the text written by a programmer (see also Programmer) in a computer language, with which he or she gives the computer specific instructions. Using this source code, a programmer builds a software programme.
Bulk data	Data is available as bulk when an entire data set is available with one press of the button or by a machine-automated request.
CCo	With a Creative Commons licence (CC) the entitled party gives permission to others to distribute its data, share it with others and in some cases even to modify it. A CCo statement (CC zero) indicates that copyright has been completely relinquished or that it concerns a data set to which copyright is not or no longer attached.
Code	A piece of text written in a computer language that gives the computer specific instructions to run a software program (see also Source code).
(APP) Challenge	A challenge promotes the development of sustainable, valuable apps. These are complex and extensive apps that contribute significantly to society. The monetary prizes for a challenge are larger than for a competition but therefore also fewer in number. With challenges, developers are supported in building the apps and their positioning in the market.
(APP)	Many developers regularly participate in app competitions. The goal of these competitions is to

¹² Open State Foundation red., *About open data* (web publication www.openstate.eu/nl/overopendata) and *The value of Open Data – Choices and effects of open data strategies for public organizations* (<http://www.rijksoverheid.nl/documenten-en-publicaties/rapporten/2012/04/26/de-waarde-van-open-data-keuzes-en-effecten-van-open-datastrategieen-voor-publieke-organisaties.html>)

Competition	build as many different applications as possible using certain data, for example to quickly and efficiently examine what the possibilities of the data are. Because participants stand a good chance of winning a (small) cash prize, there is a lot of interest in these types of competition. That is a good thing because the more participants take part, the more new applications are created.
CSV	CSV (Comma Separated Values) is an example of an open format (see also Open Standard). A CSV file is a specially formatted file in which rows and columns are stored. The rows are divided by the various lines and the columns are divided by a comma between the various values on a line. This simple file type is used to transport data between different databases of different file formats (export and import).
Data	Data is information that can be interpreted, exchanged and saved. When a measurement is taken or facts are registered, data is produced. When this data is subsequently interpreted, information is created.
Data set	A data set is an organized collection of data. It can comprise files, such as text files, tablet information, geo information or digital photos. Combined these files form a data set.
Geo data	Geo data is data which includes a spatial element, i.e. a reference to a location on earth. It could be information on a building, information from a satellite or land surveying information. Geo data can be combined with maps, aerial photos and other geo data to create compound maps. This is how new, combined insights into places on earth are created.
Granularity	The extent to which data is aggregated and can be traced to individual objects. Data is highly granular when each record in a data set can be traced to an individual object. If data is only presented in groups, it becomes less granular. The least granular is a data set that only reproduces statistics over the entire population.
Hackathon	A Hackathon is an event at which programmers and other people involved (designers, researchers, project managers, etc.) work closely together on a software project in a short period of time. Hackathons may take one day to a week and have a specific topic on which the participants focus. It could be a computer language or a certain application, but also an API (see API). Some hackathons are held for purely educational or social purposes, but they are usually aimed at creating useful and innovative software. The Open State Foundation organizes two major hackathons a year in the Netherlands. In 2011 the focus was on Apps for the Netherlands, Apps for North-Holland and Apps for Amsterdam, in 2012 Apps for Democracy and 2013 Apps for Europe and the Public Broadcasting Hackathon.
JSON	JSON (JavaScript Object Notation) is an example of an open format (see also Open standard). It is an often-used and simple format for data exchange and it is computer-language independent. Compared with other formats, JSON is lightweight, flexible and not too extensive.
Linked Data	Users and re-users of data often want to combine information from different sources with each other to create new insights and added value. Linked Data is a method to do this via the internet. It could be seen as a cloud of linked data sets.
Metadata	Metadata gives information about the character of data. In simple terms, it is data describing data. For example, the title of a data set, a description of the contents of a set, the frequency with which the data set is updated, an explanation or a web address where the data can be downloaded. It is also possible to indicate the file format of the data (e.g. JSON, CSV or XML) or the file size.

Open Access	Open access means unlimited online access to scientific research. It is primarily intended for scientific articles, but it is increasingly being used for theses, chapters from books and scientific studies published according to open access norms.
Open Data	<p>This is data that is publicly accessible, machine readable and free for use/reuse and distribution/redistribution. The data may be derived from governments or from businesses. It is important that in technological and legal terms the data can be used, reused and distributed as accessibly and practically as possible.</p> <p><i>Technologically open</i> means that the data is available in an open format that is also readable for a machine (computer). Computers can retrieve and process the data in applications such as apps and websites (see also API). In addition, it is important that the data can also be downloaded in bulk (see Bulk data). For direct use the data should also be in open standard (see also JSON and CSV).</p> <p><i>Legally open</i> means that a licence applies to the data that explicitly gives permission to use/reuse and distribute/redistribute the data commercially and non-commercially (see also CCo and Open Data licences).</p>
Open Data licences	It is important for open data that there are no restrictions on use or reuse. This is regulated in a licence such as the open data licence (ODbl) or a CCo statement (see CCo). Data to which the statement applies may be freely used/reused, distributed/redistributed and edited.
Open Source	Open source is an open source model. It means that source materials (the source) are accessible to everyone and may be adapted to create a specific end product. By using someone else's product design, a re-user is able to create a product that is entirely his/her own. Where software design is concerned, the source material consists of a source code (see also Source code) that is readily available. Everyone can (technologically) and may (legally) read, adapt and distribute this source code.
Open Standard	Digital information can easily be exchanged across the globe by using open, non-software specific standards. It improves communication between governments, businesses and citizens.
Programmer	A programmer (or developer) is a person who writes software in a computer language to give the computer specific instructions. Programmers know one or more computer languages (e.g. PHP, Python or Java).
Raw data	Raw (or primary) data is collected without processing or other forms of manipulation. Only when this raw data has been processed and interpreted, it is termed information.
Rawness	<p>The extent to which data is used for disclosure.</p> <p>Raw data is a one-on-one copy of the data present in the system, with identifying information such as the citizen service number, name and address removed. When it comes to processed data, the organization has made an effort to present it in a different way than it was used in the source systems.</p>
Semantic form of information exchange	<p>Data can be released in accordance with a semantic structure or standard.</p> <p>Data that can be released according to a semantic structure is in its simplest form a .csv, a database or an .xls. In its most advanced form the data is structured in such a way that it can be automatically related to other data sets. A standard is a formally established structure and is in general also recognized and used by other organizations.</p>
Scraping (and	Data scraping is a technology with which a computer program copies certain data from certain

Scrapers)	files (such as a web page or database) and stores it in a new file. The computer programs involved are called scrapers. Scrapers help to order data from other files.
Software	Software is the common term for computer programs. Besides computers and tablets, devices such as telephones, televisions and cars increasingly contain software.
Web service	A web service is a generally accepted technology to exchange information between two computerized information systems.
Government Information (Public Access) Act (request)	The Government Information (Public Access) Act (<i>Wet openbaarheid van bestuur</i>) provides for the right of citizens to information from the government. In principle, the data collected by the government is public, unless the Government Information (Public Access) Act or other legislation stipulates that it is not appropriate to make the requested information public. Anyone can request information from the government about any administrative matter. A request under the Government Information (Public Access) Act should be lodged with an administrative body. The Government Information (Public Access) Act also applies to an organization or service that falls under the responsibility of an administrative body.
XML	XML (EXtensible Markup Language) is an example of an open format (see also Open standard). XML is characterized by the extensive possibilities to clarify data, describe it in detail and link parts to international standards.

10 Sources

- Directive 2003/98/EC regarding the reuse of public sector information
Online (1 May 2015): <http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=COM:2011:0877:FIN:NL:PDF>
- Proposal regarding Public Records Act and Reuse of Public Sector Information Act
Online: <http://www.tweedekamer.nl/downloads/document?id=d75032f8-food-4d36-8238-26a8ba85058d&title=Voorstel%20van%20wet.pdf>
(retrieved 1 May 2015)
- Explanatory Memorandum to the Reuse of Public Sector Information Act
Online: <http://www.tweedekamer.nl/downloads/document?id=e12ede51-1436-44dc-970f-d05435a68b1e&title=Memorie%20van%20toelichting.pdf>
(retrieved 1 May 2015)
- National Audit Office, Trend report Open Data 2015 (adopted 30 March 2015, presented to the House of Representatives on 31 March 2015)
Online via
http://www.rekenkamer.nl/Publicaties/Onderzoeksrapporten/Introducties/2014/03/Trendrapport_Open_data
(retrieved 1 May 2015)