Model Architecture for National Archival Institutions (MARA)

A shared vision of Convent RHCs and the National Archives

Version 1.0

Date February 2016

Publishing details

Project name	Model Architecture for National Archival Institutions (MARA)
Location	Until agreed otherwise, this document will be managed by the National Archives
Contact person	Petra Helwig T +31 (0) 6 55 26 78 91 petra.helwig@nationaalarchief.nl Postbus 90520 | 2509 LM Den Haag
Author	Petra Helwig (National Archives) on behalf of: The Architecture Committee of the National Archives and RHCs
Together with	Rafaat Alebate (Historisch Centrum Overijssel) Roland Bisscheroux (Noord-Holland Archives) Rob Daems (National Archives / Ordina) Buddy Joe Groeneveld (National Archives) Gijsbert Kruithof (National Archives) Arjen van der Kuijl (National Archives) Jeroen van Luin (National Archives) Klaartje Pompe (Noord-Holland Archives) Yuri Riet (National Archives) Con Sadée (National Archives / Capgemini) Joost Salverda (Gelderland Archives) Martijn Smeets (National Archives / BiZZdesign) Peter Westerveld (Limburg Regional Historical Centre) Annelot Vijn (The Utrecht Archives) Kaj van Vliet (The Utrecht Archives) Ruud Yap (National Archives)
Version	1.0
Separate	Appendices: - Model Architecture for National Archival Institutions 1.0: Appendix: Objectives and Principles - Model Architecture for National Archival Institutions 1.0: Appendix: Models

Version history

Name	**Version**	**Date**	**Adopted by**	**Contents**
RHCs' Enterprise Architecture: "SOLL" 2015	1.0	April 2013	Convent	First version
Enterprise Architecture "Digital Tasks of Government (DTR)"	1.0	May 2014	Convent	Specifically aimed at DTR
Archival Institutions Reference Architecture	0.9	November 2015	DTR steering group	Combining and expanding previous architectures
Model Architecture for National Archival institutions	1.0	February 2016	Convent	Name change from ARA to MARA

More information about the connection between the various architecture documents can be found in Appendix IIRelation *to earlier architectures*.

Contents

Contents

1 Introduction 12

1.1 The most constant factor is change 12

1.2 Why architecture? 12

1.3 Objective of this document 12
1.3.1 Description of a shared vision 12
1.3.2 Assessment framework 12

1.4 Target group 13

1.5 Relation to previous documents 13

1.6 Further development 13

1.7 Management and responsibilities 13

1.8 Reader's guide 13
2 Scope and content of the Model Architecture for National Archival Institutions. 15

2.1 Contents 15

2.2 Model architecture or solution architecture? 15

2.3 For whom? 15

2.4 Scope of the Model Architecture for National Archival Institutions 15
2.4.1 Defining subjects 15
2.4.2 Scope of Business Architecture 15
2.4.3 Scope of Information Architecture 16
2.4.4 Scope of Technical Architecture 16
2.4.5 Security 16
2.4.6 Defining according to time: shorter term 16

2.5 Substantive relationships with other architecture 16
2.5.1 Coherence with NORA and "daughters" 16
2.5.2 Coherence with DUTO 16
2.5.3 Coherence with the Digital Heritage Network 16
2.5.4 Coherence with NCDD 16
3 Vision 17

3.1 Why is change necessary? 17
3.1.1 Accessibility 17
3.1.2 Technological developments 17
3.1.3 Efficiency 17
3.1.4 Security 17

3.2 Developments 17
3.2.1 Accessibility 18
3.2.2 Technological developments 19
3.2.3 Efficiency 19

3.3 Outline of the desired situation 19
4 Objectives and principles..21

4.1 Basic principles for archival institutions 21

4.2 Objectives 21
4.2.1 Objectives for the positioning of RHCs and the National Archives 21
4.2.2 Objectives for service provision and the exchangeability and findability of data and metadata 22
4.2.3 Objectives for the management and preservation of data and metadata 23

4.3 Principles 24
4.3.1 What are principles? 24
4.3.2 Principles for the positioning of RHCs and the National Archives 24
4.3.3 Principles for service provision and the exchangeability and findability of data and metadata 25
4.3.4 Principles for the management and preservation of data and meta data 26
4.3.5 Archival basic principles 28
5 Business architecture: services 29

5.1 Supervision 29

5.2 Giving advice 29

5.3 Presenting material 29

5.4 Making material available 29

5.5 Making data available for reuse 29

5.6 Management and preservation of outsourced archives 30
5.6.1 Why is this service provided? Unburdening legal caretakers 30
5.6.2 Precondition: Files must be closed before they are outsourced. 30

5.7 Facilitating viewing of outsourced archives 30
5.7.1 Precondition: As the legal caretaker must ensure searchability to give civil servants access they must therefore choose a suitable interface. 31
5.7.2 Precondition: The legal caretaker is responsible for providing authorization and authentication. 31
6 Business architecture: organization and main functions..........................32

6.1 Introduction 32

6.2 Recipients of services provided by an archival institution 32

6.3 Main functions of an archival institution 32

6.4 Use of shared services 33
6.4.1 Knowledge function: national knowledge network 33
6.4.2 Services provided to RHCs by the National Archives 34

7 Business architecture: processes. 35

7.1 WVI process landscape 35

7.2 Required adjustments 35
7.2.1 Primary processes for service provision 35
7.2.2 Supporting processes for service provision 36
7.2.3 Processes for implementing the knowledge function 36
8 Business/information architecture: information and data 37

8.1 Structure of an archive 37

8.2 Information from visitors and requests to be recorded 37
8.2.1 Visit history 38
8.2.2 Visitor 38
8.2.3 Request 38

8.3 Information about indexes to be recorded39

8.4 Other collections 39
9 Information architecture: functionality and applications 40

9.1 Introduction 40

9.2 Summary of required functionalities 40

9.3 Functionalities for long-term availability of archive records 40
9.3.1 Functionality: Digital storage site 40
9.3.2 Functionality: Management of manifestation-independent metadata 41
9.3.3 Functionality: Rendering of digital manifestation 42
9.3.4 Functionality: Rendering of digital manifestation in context 42

9.4 Functionality for enriching and connecting information 43
9.4.1 Functionality: Management of actors 43
9.4.2 Functionality: Enrichment of information 43
9.4.3 Functionality: Aggregation of information 43
9.4.4 Accessibility of the National Collection 44

9.5 Functionality for visibility 44
9.5.1 Public website 44
10 Standards 45

10.1 Introduction 45
I.1.1 What are standards? 45
I.1.2 What are open standards? 45

10.2 Standards to be used 45
11 Shared facilities: how this will be achieved 47

11.1 Solutions for the long-term availability of archive records 47
11.1.1 Solution: Digital storage site for transferred material 47
11.1.2 Solution: Management of manifestation-independent metadata 47
11.1.3 Solution: Rendering of digital manifestation 47
11.1.4 Solution: Rendering of digital manifestation in context 47
11.1.5 Solution: Management of actors 47

11.2 Solutions for recording metadata of transferred material 47
11.2.1 Exceptions to the separation of MOM and MAM 47

11.3 Solutions for recording metadata of outsourced material 48

11.4 Solutions for the aggregation of archives 48

11.5 Solution for viewing outsourced material at departments 48
11.5.1 Example: specific organization of architecture at the National Archives48

11.6 Solutions for visibility 49
11.6.1 Solution: Public website 49

I.2 Summary in a diagram 50
12 Technical infrastructure..51

12.1 What is technical infrastructure? 51

12.2 Objective of this chapter 51

12.3 Objective: Safeguarding business continuity 51
12.3.1 Description 51
12.3.2 Achieving the objective51

12.4 Data storage – e-Depot 52
12.4.1 Description 52
12.4.2 Requirements 52
12.4.3 Solution 52
12.4.4 Access to eBOM 52

12.5 Data storage: Hosting and housing 53
12.5.1 Description 53
12.5.2 Requirements 53
12.5.3 Precondition: government data centres53
12.5.4 Solution 53
12.5.5 Gradual transition through a temporary solution 54

12.6 Data storage: Back-up, contingency use and recovery 54
12.6.1 Description 54
12.6.2 Requirements 54
12.6.3 Precondition: government data centres55
12.6.4 Solution from 2017 55
12.6.5 Solution to the risks posed to business continuity 55

12.7 Data transport – zoning 55
12.7.1 Description 55
12.7.2 Requirements 55
12.7.3 Solution 55

12.8 Data transport: network 56
12.8.1 Description 56
12.8.2 Requirements 56
12.8.3 Solution for data connections 56
12.8.4 Solution for network equipment 56

12.9 Development and management of systems: DTAP 56
12.9.1 Description 56
12.9.2 Requirements 56
13 Security 57

13.1 What is information security? 57

13.2 Objective of this chapter 57

13.3 General 57

13.4 Security policy 57

13.5 Organization of information security 58

13.6 Management of business resources 58

13.7 Personal security 58

13.8 Physical security and security of the environment 59

13.9 Management of communication and operational processes 60

13.10 Access security 60
13.10.1 Management of access rights for users 60
13.10.2 Access control for networks 61

13.11 Procurement, development and maintenance of information systems 61

13.12 Incident management 62

13.13 Business continuity management 62

13.14 Compliance 62
14 References 63
15 Terms and abbreviations 64
Theme A Presentation of a digital archive item – integrity and authenticity 68

A.1 Introduction 68

A.2 What is a digital archive record? 68

A.3 What is necessary for a complete and unaltered reproduction of a digital record? 68
A.3.1 Essential characteristics must be preserved 69
A.3.2 The influence of viewing software 69
A.3.3 So how do we produce a version with integrity? 71
A.3.4 What does this look like? 71
A.3.5 Reuse? 72

A.4 Resulting principles 72
Theme B Reuse of archive material 73

B.1 What is reusable? 73

B.2 For authentic, complete and unaltered archive records, do not copy data but refer to it 73

B.3 Making data available for reuse 73

B.4 Resulting principles 73
Theme C Manifestations and metadata 74

C.1 The concept of "manifestations" 74

C.2 Principle: one-off storage of manifestation-independent metadata 74

C.3 Diagram of manifestation-dependent and manifestation-independent metadata 74

C.4 Structure of ISO 23081: indication of manifestation dependence and independence 75

C.5 Overview of metadata in the e-Depot application and collection management system 75

C.6 Resulting principles 76
Theme D Use of core registrations..77

D.1 Separate registration of actors: actor registers 77

D.2 Other authority files / core registrations 77

D.3 Resulting principles 78
Theme E Connection layer: aggregators versus Linked Data............................79

E.1 Why make a connection layer and how is this done? 79

E.2 Aggregators 79

E.3 Linked Data 79

E.4 Resulting principles 79
Theme F Implementation of the three-layer model 80

F.1 National Strategy for Digital Heritage and the three-layer model 80

F.2 Implementation of National Strategy in MARA 80

F.3 Searching and finding according to the three-layer model 81
F.3.1 Further details: 81
F.3.2 Searching and showing information in this model works as follows: 81
F.3.3 Searching and showing information from various manifestations 82
Theme G Further details on the technical infrastructure 83

G.1 The concepts housing and hosting 83
G.1.1 Description of housing 83
G.1.2 Description of hosting 83

G.2 Organization of ODC North 83

G.3 The concepts of back-up, recovery and contingency 84
G.3.1 Description of back-up 84
G.3.2 Description of recovery 85
G.3.3 Description of contingency 85

G.4 Concepts related to the network (zoning, data connections) 85
G.4.1 Description of zoning 85
G.4.2 Description of data connections 86

G.5 Rationale for gradual transition by means of a temporary solution 86
G.5.1 Business continuity: complexity of the DTR 86
G.5.2 ODC North's back-up and contingency location is not yet in order 87

G.6 Non-functional requirements 87
G.6.1 Reliability 87
G.6.2 Performance efficiency 88
Appendix I. Explanatory notes on architecture 90

I.1 Concept of "architecture" 90

I.2 What must be described? Business, information and technology 90

I.3 Methods for describing architectures Principles, models and coherence 91

I.4 Coherence 91
Appendix II. Relation to earlier architectures 93

II.1 WVI + DTR = MARA 93

II.2 Relation to DTR Architecture 93
II.2.1 Objective of DTR Architecture 93
II.2.2 Core points of the content of DTR Architecture 93
II.2.3 MARA in relation to DTR Architecture 93

II.3 Relation to WVI Architecture 94
II.3.1 Objective of WVI Architecture 94
II.3.2 Core points of the content of WVI Architecture 94
II.3.3 MARA in relation to WVI 94
Appendix III. Development and architecture sheets.....................................95

III.1 Architecture sheets that have already been developed 95

III.2 Subjects that have still to be developed further 95
Appendix IV. Services provided by the National Archives 96

IV.1 E-Depot services for RHCs 96

IV.2 Other services for RHCs 96

IV.3 Distribution of work between RHCs and the National Archives 96
IV.3.1 Implementing preservation activities 96
IV.3.2 Implementing destruction services on outsourced archives 97
Appendix V. Structure and organization of a national knowledge network. 98

V.1 The National Archives’ knowledge function 98

V.2 The National Archives’ Action Line 98

V.3 Safeguard as from 2016 99

V.4 Knowledge questions on the Knowledge & Innovation Agenda 99

V.5 Themes for knowledge platforms 99

1 Introduction

1.1 The most constant factor is change

The National Archives and the Regional Historical Centres (RHCs) are developing rapidly. There are various projects and programmes to cater to the changing wishes and requirements of users and new ways of providing services:

- The DTR (Digital Tasks of Government) programme was set up to ensure that the National Archives and the RHCs can continue to perform their normal duties, in particular as regards digital material from the Central Government, and that these institutions are ready for the future.
- Implementing the "outsourcing" service by departments is the main responsibility of the DWR Archive project.
- Implementing the e-Depot service to RHCs is partly covered by the DTR project but Archief2020 is also closely involved, as this project includes project pilots with RHCs regarding the transfer of digital material to the e-Depot.
- Archief2020 plays an important role in strengthening the archive sector as a whole and promoting the accessibility of collections.

Architecture – and therefore this document – aims to guide these changes.

1.2 Why architecture?

To achieve organizational goals within such a changing context it is essential that the organization, primary and supporting processes, technology, and "hard" infrastructure fit together seamlessly and work together to achieve their aims.

Architecture shows us what we want to achieve and what this means for the various domains (such as organization, information, infrastructure, security). Architecture makes it easier to see which projects contribute to which part of the final picture, where there is overlap and where alignment is required. It also guarantees that no elements of the final picture are forgotten.

Here is a fine description of the value of architecture: "Architecture gives you the luxury of only having to consider the same problem once".

There is a more detailed explanation of the various aspects of architecture in Appendix I *Explanatory notes on architecture*.

1.3 Objective of this document

1.3.1 Description of a shared vision

On the basis of the challenges facing the RHCs and the National Archives, this document sketches a broad outline of the shared vision and what is required in terms of organization, applications and information, infrastructure, and security to bring it about. This document is therefore a consolidation of insights and basic principles that have been gained in various places within the organization.

1.3.2 Assessment framework

As well as describing the shared vision, this Model Architecture for National Archival Institutions is also used in practice to provide starting projects with frameworks and guidelines. Furthermore, it serves as an assessment framework for all projects that are implemented in the various work packages within the framework of the DTR programme[1].

RHCs and the National Archives can also use this reference architecture as an assessment framework for projects within their own organization. This will help these projects to contribute to the subscribed shared vision.

1.4 Target group

This document is intended for:

- The management of the National Archives and Regional Historical Centres who want a framework for managing change;
- Programme and project leaders from the National Archives and Regional Historical Centres who are shaping changes;
- Architects, project architects or information analysts who are developing specific topics in greater detail.
- Others interested in the vision of the National Archives and the Regional Historical Centres, and the way in which this architecture is helping to shape it. This group includes colleagues from the Ministry of Education, Culture and Science who are involved in the Digital Heritage Network and colleagues from other archival institutions (other than the "Convent RHCs").

1.5 Relation to previous documents

This document follows on from two earlier architecture documents on DTR architecture and WVI architecture respectively. For readers familiar with these documents, the relationship between this and earlier documents is described in Appendix II Relation to earlier architectures.

1.6 Further development

The architecture is continuously being developed in more detail. Specific topics will lead to additions or extensions to this document, or will be described in the "architecture sheets" accompanying this document. Insights from the projects will be incorporated back into the architecture. MARA is therefore a living product.

Appendix III *Further development and architecture sheets* indicates which topics already feature in an architecture sheet and which topics will be explored in greater depth.

1.7 Management and responsibilities

This document is managed by the National Archives.

Content-related adjustments will be discussed by the architecture committee of the Regional Historical Centres and the National Archives. The Convent of RHCs is accountable for MARA.

The Enterprise Architect of the National Archives is responsible for delivery of MARA.

1.8 Reader's guide

This document is a rather lengthy text, but this is unavoidable given its nature. It is, after all, a reference work and a compilation of the knowledge and insights that have been obtained. It must be readable for an audience that does not have much prior knowledge, and requires information, while offering in-depth information to those who are already well acquainted with the general topic. In order to keep this document readable, it has been divided wherever possible into "digestible chunks" that can be read or skipped as desired.

[1] The decision that the architecture committee should function as architecture board for all DTR projects was taken by the DTR steering group on 3 October 2014.

The core of the text comes after Chapter 2 *Scope and content of the Model Architecture for National Archival Institutions*. The most important chapters for managers are chapters 3 *Vision* and 4 *Objectives and principles.*

These chapters are the introduction to the core text, which consists of the following chapters:

- 5 *Business architecture: services* describes the main functions and services to various target groups from each RHC and the National Archives;
- 6 *Business architecture: organization and main functions*. This describes the main functions required for these services and the shared services deployed;
- 7 *Business architecture*: *processes*.. This chapter gives a broad description of the processes required to produce these services;
- 8 *Business/information architecture*: *information and data* describes what is required in terms of functionalities and data, on the one hand to support processes and on the other hand because sustainable data storage is one of the core functions;
- 9 *Information architecture*: *functionality and applications* describes which functionality is required;
- 10 *Standards*. This is a description of which standards must be used for specific situations;
- 11 *Shared facilities: how this will be achieved*. This chapter contains concrete, shared solutions that are based on the frameworks and guidelines;
- 12 *Technical infrastructure* stipulates infrastructure requirements and the solutions that the National Archives has implemented;
- 13 *Security*. This chapter pays detailed attention to the security requirements and the way in which the National Archives meets them for the e-Depot.

After the core text, the document moves on to examine a number of themes. These provide a more in-depth examination and explanation of several aspects that are of specific interest to the architecture. *Architecture principles* have been defined for many of these aspects. The themes explain a number of principles.

Finally, an explanation is provided of the concept and objective of architecture (Appendix I*Explanatory notes on architecture*) and the document’s relation to earlier documents (Appendix II *Relation to earlier architectures*).

Appendix III *Development and architecture sheets* lists topics that have already been explored in separate architecture sheets. These have not been attached to this document.

Finally, there will be a short explanation of the organization of two shared services that all RHCs and the National Archives will make use of: Appendix IV *Services provided by the National Archives* and Appendix V *Structure and organization of a national knowledge network*..

There are two separate appendices to this document: one describes all objectives and principles, and the other describes all architectural elements, such as processes, functions and applications.

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.

2 Scope and content of the Model Architecture for National Archival Institutions.

2.1 Contents

The Convent has agreed upon MARA as the reference architecture for the National Archives and every RHC. It will include:

- Services
- Operational processes
- Information
- Technical infrastructure
- Security

insofar as they are shared by the RHCs and the National Archives.

2.2 Model architecture or solution architecture?

The title of this document is Model Architecture for National Archival Institutions (MARA). The term “model” indicates that a concrete architecture is not outlined for every point. For some parts, only the frameworks to which the specific architecture for RHCs and the National Archives must comply have been indicated. This architecture says *what* has to happen, but not *how*.

For other parts, the architecture is more specific and gives concrete details. This is mainly the case for components that all RHCs and the National Archives share with each other, such as the e-Depot. These components are explored in more detail in chapter 11 *Shared facilities: how this will be achieved*.

2.3 For whom?

MARA is the reference architecture for the National Archives and every RHC. However, other archival institutions may draw inspiration from this model and follow parts of it.

2.4 Scope of the Model Architecture for National Archival Institutions

2.4.1 Defining subjects

In MARA, we describe matters (subjects in business architecture, information architecture, infrastructure, and security) for which it has been agreed that it is necessary to choose solutions with a similar alignment. The Convent will decide how far collaboration and the shared alignment will go and agreements on this will be monitored.

A similarly-focused solution may be a shared application (such as the e-Depot) or an architecture for a shared service or functionality (such as a collection management system). In case of the latter, everyone can choose their own solution, but the preconditions will remain the same.

The Convent will decide on how far the collaboration and alignment will go.

2.4.2 Scope of Business Architecture

This is a description of what is shared with regard to the services, main functions and processes. This means that the primary processes are especially in scope and the supportive/business processes are so to a lesser extent. This is based on the generic processes that were developed in 2013[2].

[2] *Recruitment, management and provision. Guidelines for work processes for RHCs (2012).*

The background to this is that business operations are primarily a matter for each separate RHC, even though it never does any harm to "take a look at the neighbours". This architecture, however, mainly fits into the "digital infrastructure", the objective of which is the joint management of digital and non-digital archive material.

2.4.3 *Scope of Information Architecture*

In this document, the focus is on information that is managed for the primary process. The main question is which data on digital material and archive creators are managed by the RHCs and which functionality is required to manage this data. Information on business operations such as control information, financial information, personnel information etc. is not included here.

2.4.4 *Scope of Technical Architecture*

The technical information mainly concerns the e-Depot.

2.4.5 *Security*

Much attention will be paid to security requirements and the way in which they are met with regard to the e-Depot.

2.4.6 *Defining according to time: shorter term*

This reference architecture focuses on the relative short term, i.e. the next few years.

In the long term it will still be necessary to develop a vision for a number of fundamental questions concerning the changing nature of information. For example, is it possible or desirable in this digital era for archival institutions to fulfil their custodianship as they did in the past with "paper" and are now doing with the e-Depot by physically having digital and other material in their possession, stored in their own depots and e-Depots, and managed by their own employees?

This architecture scarcely addresses these long-term questions.

2.5 Substantive relationships with other architecture

2.5.1 *Coherence with NORA and "daughters"*

NORA and its "daughters" (including EAR and GEMMA) contain architectural principles at a high level. Although they will not be repeated in MARA, which is still to be developed, MARA must connect with them. If a principle identified in MARA is derived from a specific principle from NORA – a different architecture – this will be indicated.

2.5.2 *Coherence with DUTO*

The objectives and principles stated in MARA show a great many similarities with those in the system of standards for Long-term Accessibility (DUTO). MARA's objectives and principles, however, are mainly focused on the archival institutions.

2.5.3 *Coherence with the Digital Heritage Network*

The Digital Heritage Network (NDE), which also develops architectures, is run by the Ministry of Education, Culture and Science (OCW). Alignment will also be sought with this network in as far as possible. MARA can deliver input for NDE architecture.

2.5.4 *Coherence with NCDD*

The NCDD is working on research into a shared infrastructure for cultural heritage. Contact is being maintained with the NCDD in order to harmonize the work and alignments as much as possible.

3 Vision

3.1 Why is change necessary?

It is interesting to determine the real aspects that prompt or guide archival institutions to make desired changes to their organization and/or technologies. These aspects as known as drivers. The following important drivers for archival institutions have been identified:

- Accessibility requirements
- Technological developments
- Efficiency
- Security

Drivers of archival institutions
Long-term accessibility
Technological developments
Efficiency
Security

3.1.1 *Accessibility*

Traditionally, the objective of archival institutions is to make information accessible. Accessibility also means being easy to use.

Furthermore, it requires that the right information is acquired, stored properly and managed. After all, it is difficult to make information accessible if it is not stored properly.

3.1.2 *Technological developments*

An increasing amount of archival material is digital. Keeping archival material accessible permanently entails specific problems, such as file formats becoming obsolete. The nature of information is also changing: information is less likely to be linked one-on-one to a saved file in an application that is managed in one location by just one person – instead it is “everywhere” and “everyone’s”.

At the same time, digital developments also offer opportunities as material can be made available at more than one location. There are more possibilities to connect material to each other and to serve different target groups. Examples of this are linked data and big data.

3.1.3 *Efficiency*

Efficiency is also important to archival institutions. Not every RHC has to do everything themselves.

3.1.4 *Security*

Security is a different type of driver. This aspect does not initiate changes in archival institutions by itself. However, security is a hot topic when it comes to all forms of digital storage and accessibility. It also concerns privacy and ensures that everyone only has access to information that they are allowed to see. From the perspective of security, you do not undertake action to create new possibilities; instead you need to act to prevent undesirable situations from arising.

3.2 Developments

For each of these drivers there are developments or findings that trigger change.

3.2.1 Accessibility

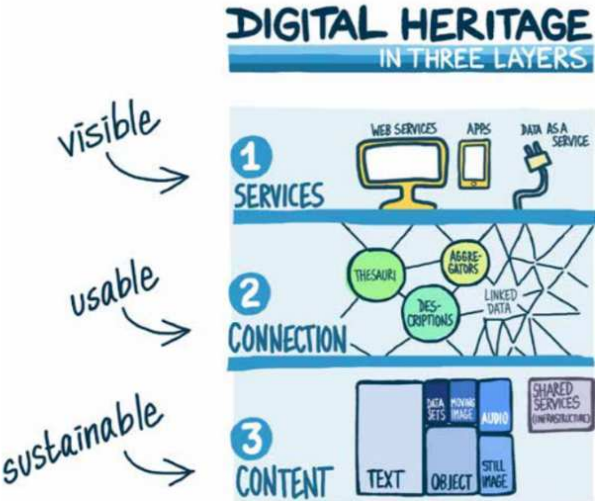
3.2.1.1 Larger reach of the collection and better response to demand: from silos to layers

Currently, archival and other material can often only be accessed through the website of the institution that manages the material in question. Furthermore, many archival and other institutions only provide visitors with material that they actually manage themselves. It is as if the collection is thought of first, and the visitor's request is only an afterthought.

A main cause of this is that the application landscape is dominated by "silos": systems that incorporate all parts of the chain, from archive management to accessibility. The storage, metadata management, search engine, and user interface functions that enable the public to search, view and order are therefore inextricably linked to each other. These silos are often organized around one specific object type. There will be a separate image bank, a library system and an system for managing the archive.

The objective is to enable websites and apps to provide access to a multiplicity of sources, including material kept in a different archive. This makes it possible to respond better to requests from users. It also increases the reach of the collection, as it is available at more than one location.

The challenge is to move from a system of silos to a system of layers (functional decomposition). In any case, the presentation layer (website, apps) must be disconnected from the systems in which the information is managed. It must also be possible to create semantic relationships between objects.



1 **Source: National Strategy for Digital Heritage**

This looser structuring of the applications has a number of important advantages. First of all, parts can be replaced separately from each other: a new viewer would then have no consequences for the existing choice for a collection management system (or vice versa). This makes the archive service less dependent on specific market actors. Applications on the same layer could converge. For example, it would be possible to use the same viewer or a related collection management system for different types of documents. A third advantage is that this type of architecture would make it possible to disclose sources at a level that transcends organizations (in relation to each other). This would create a full picture of the various collections.

Technological possibilities have also raised users' expectations regarding accessibility and user-friendliness: "Google can also do that…".

3.2.1.2 Accessibility of outsourced records for the legal caretaker

Material that is outsourced must remain accessible for the original legal caretaker: civil servants must be able to access it if they are entitled to do so. The fact that outsourced material can only be consulted on the basis of authorizations places demands on how accessibility is organized.

3.2.1.3 Open Data

Accessibility of information is also receiving more attention in legislation, as shown by developments regarding open data. In addition to end-users of material and ready-made presentations, a growing group is interested in actively using the presented material. They, the “crowd”, add information themselves, enriching or reusing raw data to create their own end products. This makes new demands on presenting and providing access to data or information.

3.2.2 Technological developments

Archival institutions are managing increasing amounts of digital material. This requires appropriate infrastructure, which is expensive and complicated to build and maintain. Each institution had its own depot for storing paper records, but they now work together on storing digital records. Examples include sharing knowledge and joint use of an e-Depot.

It is often difficult for the original legal caretakers (the persons who have to transfer the archive document) to store digital archive records for a long time. Even records that are not eligible for permanent storage must sometimes be kept for decades.

RHCs and the National Archives want to be able to provide services to their legal caretakers to help with the long-term accessibility of their records. Their concrete intention is that legal caretakers must be able to “outsource” their records to archival institutions. Responsibility for material that has not yet been statutorily transferred remains with the original legal caretaker, but the archival institution will provide long-term accessibility.

3.2.3 Efficiency

Cooperation is increasing between RHCs and the National Archives. They are building an infrastructure together to keep digital records available. As it is not efficient for everyone to reinvent the wheel separately, the objective is for RHCs to be able to use this infrastructure. RHCs will then be able to deliver services to their “community”.

Other components related to the e-Depot are being examined to see if collaboration would be expedient.

The objective is to use shared infrastructure components wherever this is expedient. Important components include the e-Depot and accompanying storage as well as generic parts of a public website, for example.

The objective is to jointly implement a knowledge function.

3.3 Outline of the desired situation

In the short term – a few years – the desired situation is as follows:

The National Archives maintain an e-Depot where digital archive records can be managed and stored. Files are stored securely and if necessary it will be possible to undertake preservation activities, such as conversion or migration. Metadata of archive records or other collections will be managed in one or more collection management systems. All RHCs use the same management system for the National Collection, which is also managed by the National Archives.

There is a place where digital archive records are presented (“rendered”) so that a complete, unaltered and authentic version of the archive records is available. A generic building block is used to present the digital archives (stored in the e-Depot) so that they look the same on the websites of all the RHCs and the National Archives.

The National Archives function as a service organization for the RHCs. The services of the National Archives' service organization mainly consist of providing secure storage for digital archive records and the accompanying manifestation-independent metadata. The service organization also offers RHCs the opportunity to implement preservation activities themselves on digital archive records to ensure that they are maintained at all times. RHCs can use the National Archives' infrastructure, in particular the e-Depot.

The archive records transferred to the National Archives or RHCs can be "harvested" by aggregators. These harvesters generate factual and searchable cross-institution collections. Harvesters may also play a role in the semantic connection of collections, so that collections of varying natures, such as archive records and, for example, books and newspapers, can be searched using just one search query.

The archive records are therefore uniquely and persistently identifiable, so that it is always possible to refer to them.

Data of archive records can be downloaded as *open data*.

To a certain extent, the RHCs and the National Archives wish to convey that they are part of a greater whole. One way in which this is expressed is that the entire *National Collection* can be searched in the same way on every public website of an RHC.

Departments can outsource digital archive records to the National Archives and local authorities can do likewise with some RHCs. The National Archives or the RHCs then take care of the management and preservation of these records on behalf of the original legal caretakers. They also ensure – with the agreement of the original legal caretaker – that records are destroyed after the expiry of the conservation period.

Departments and municipal authorities that have outsourced digital archive records still want to be able to search them. Archival institutions therefore present outsourced data to the original archive creators in a common format as far as this is possible. They are responsible for managing the access rights to outsourced records, making these records searchable, and for authorization and authentication when making them available.

4 Objectives and principles

NB The images in this chapter are small, but are easy to read on a screen if they are enlarged.

4.1 Basic principles for archival institutions

An architecture does not operate in isolation but is aimed at facilitating or helping the "business" to reach its objectives. Therefore this chapter looks at a number of basic principles (objectives and guiding principles) that are independent of the organization of processes and technologies.

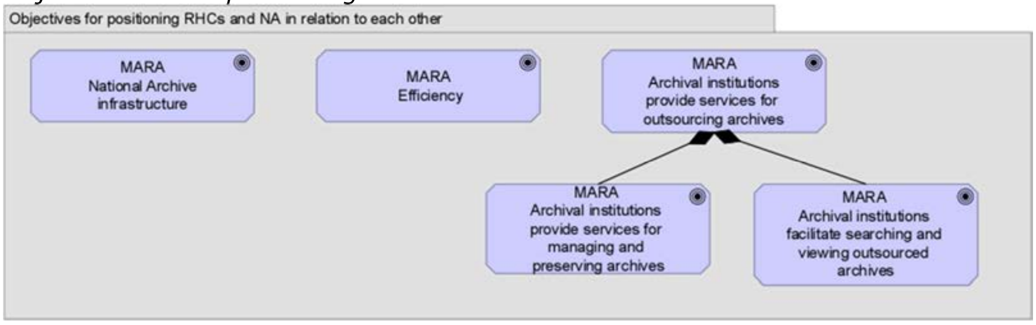
4.2 Objectives

The diagrams below state the objectives identified when working on the architecture. Objectives may contribute to the aforementioned objectives or consist of sub-objectives. The diagrams contain a description of the objectives. In the separate appendix *Model Architecture for National Archival Institutions: objectives and principles,* each objective is explained in the light of:

- Documentation: what does the objective mean?
- Rationale: why is this an objective?
- Implications: what are the consequences of the objective?

NB Undoubtedly it is possible to come up with many more objectives and principles. We have limited ourselves here to a number of objectives and principles that are currently important, which actually steer decisions in a specific direction and which are important within our context.

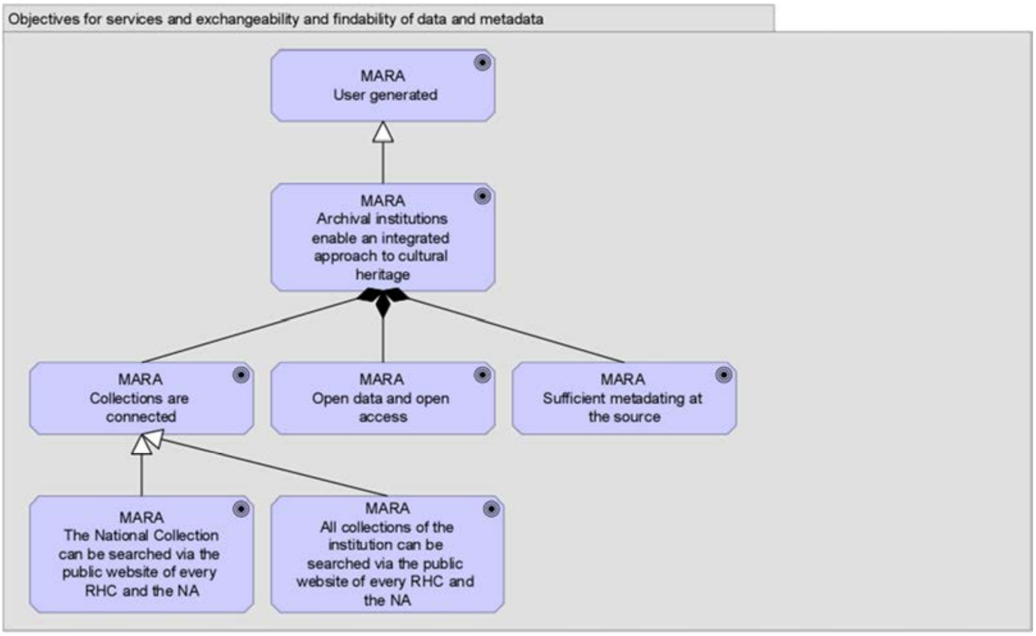
4.2.1 *Objectives for the positioning of RHCs and the National Archives*


Objectives for positioning RHCs and NA in relation to each other
MARA National Archive infrastructure
MARA Efficiency
MARA Archival institutions provide services for outsourcing archives
MARA Archival institutions provide services for managing and preserving archives
MARA Archival institutions facilitate searching and viewing outsourced archives

Name	Documentation
MARA National Infrastructure	
MARA Efficiency	
MARA Archival institutions provide services for outsourcing archives	Archival institutions can support the outsourcing of archives In doing so they take over management from legal caretakers, without the statutory transfer of the archive.
MARA	Part of the service for outsourcing is the provision of storage and preservation

Archival institutions provide services for the management and preservation of archives	services.
MARA Archival institutions facilitate searching and viewing outsourced archives	Legal caretakers must be able to view the outsourced material, for example if this is once more necessary for a work process. Within the current possibilities, archival institutions cannot fully facilitate making these archives available, due mainly to the management of authorizations and authentication. Archival institutions facilitate availability, for example, by making the outsourced data and metadata available to a search engine.

4.2.2 *Objectives for service provision and the exchangeability and findability of data and metadata*

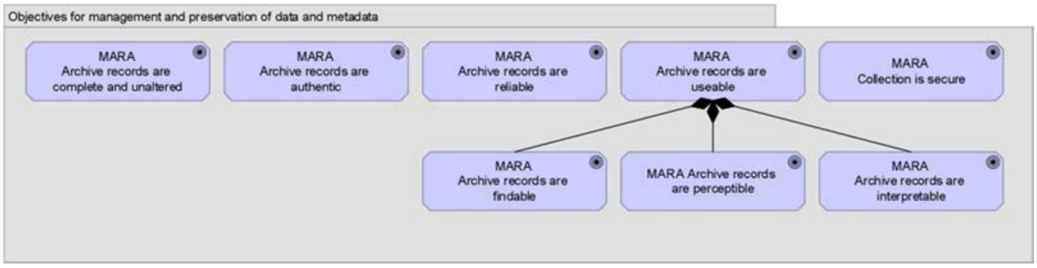

Objectives for services and exchangeability and findability of data and metadata
MARA User generated
MARA Archival institutions enable an integrated approach to cultural heritage
MARA Collections are connected
MARA Open data and open access
MARA Sufficient metadating at the source
MARA The National Collection can be searched via the public website of every RHC and the NA
MARA All collections of the institution can be searched via the public website of every RHC and the NA

Further details:

Name	Documentation
MARA User-centred	Archival institutions present transferred material according to the wishes of the various user types.
MARA archival institutions facilitate an integrated approach to cultural heritage	Archival institutions facilitate an integrated approach to all forms of cultural heritage. They facilitate users' receiving coherent (possibly cross-institution) information from various collections, such as archives, books, newspapers, AV material, etc.
MARA Collections are connected	Archival institutions encourage the connection of elements from their archival collection to each other and to other forms of cultural heritage.
MARA Open data and open access	- Use of standards - Data are exportable - Rights to data and metadata are known

MARA Sufficient metadata at the source	Archival institutions ensure or encourage that material is provided with sufficient metadata at the source.

4.2.3 *Objectives for the management and preservation of data and metadata*


Objectives for management and preservation of data and metadata
MARA Archive records are complete and unaltered
MARA Archive records are authentic
MARA Archive records are reliable
MARA Archive records are useable
MARA Collection is secure
MARA Archive records are findable
MARA Archive records are perceptible
MARA Archive records are interpretable

Further details:

Name	**Documentation**
MARA Archive records have integrity	ISO 15489: Integrity – the record is complete and unaltered http://archiefwiki.org/wiki/Integriteit: A characteristic of an archival item or element is that when it is consulted, its form, content and structure are exactly the same form, content and structure as when it was received or drawn up.
MARA Archive records are authentic	ISO 15489: Authenticity: the record is what it purports to be and was created by the person purported to have created it; http://archiefwiki.org/wiki/Authenticiteit: A characteristic of an archival item is that its integrity is established thanks to the controllable manner of archive creation, delivery, storage, and consultation.
MARA Archive records are reliable	ISO 15489: Reliability – the information in the record is accurate and can be relied on;
MARA Archive records are usable	ISO 15489: Usability – the record can be located, retrieved, presented and interpreted. It is composed of findable, perceptible and interpretable material.
MARA Archive records are findable	ISO 15489: Usability – the record can be *located,* retrieved, presented and interpreted.
MARA Archive records are perceptible	ISO 15489: Usability – the record can be located, *retrieved*, *presented* and interpreted.
MARA Archive records are interpretable	ISO 15489: Usability – the record can be located, retrieved, presented and *interpreted*.
MARA All of the institution’s collections can be searched through the public website of every RHC and the National Archives.	= including newspapers, photos, library material, etc.

4.3 Principles

4.3.1 What are principles?

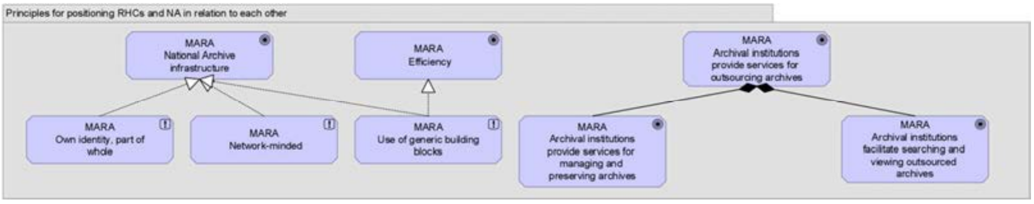
A principle is a normative characteristic of all systems in a given context. It conveys a conviction regarding the way in which the desired situation can be reached.

Below is a list of the principles derived from each objective. The division into categories is sometimes arbitrary: some principles contribute to more than one objective. The use of "generic building blocks", for example, contributes to both efficiency and service provision.

This chapter only mentions the principles (the objectives were described in the previous paragraph). In the separate appendix *Model Architecture for National Archival Institutions: objectives and principles,* each principle and objective is explained in the light of the documentation, rationale and implications.

The "light green" principles come from NORA. Some NORA principles are also important for achieving MARA objectives.

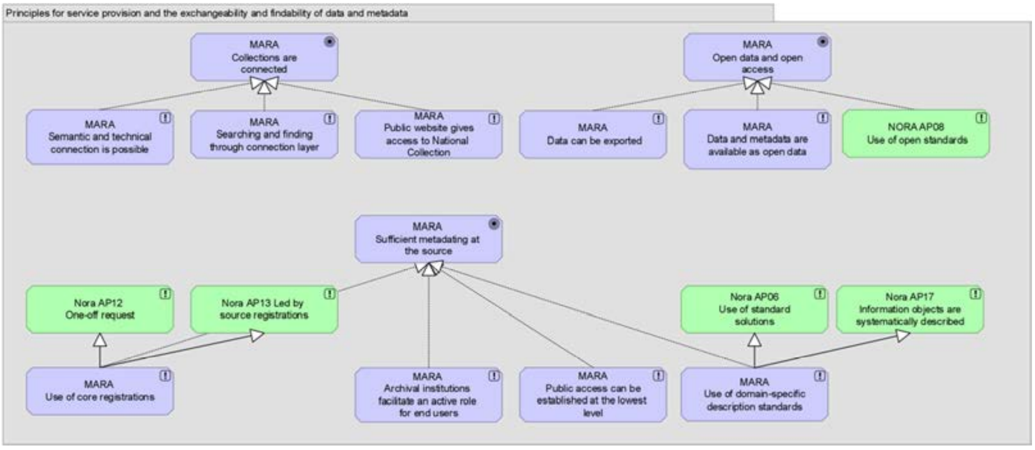
4.3.2 Principles for the positioning of RHCs and the National Archives


Principles for positioning RHCs and NA in relation to each other
MARA National Archive infrastructure
MARA Efficiency
MARA Archival institutions provide services for outsourcing archives
MARA Own identity, part of whole
MARA Network-minded
MARA Use of generic building blocks
MARA Archival institutions provide services for managing and preserving archives
MARA Archival institutions facilitate searching and viewing outsourced archives

Further details:

Name	Documentation
MARA Own identity, part of the whole	The RHCs and the NA retain their own identities. With regard to searching the National Collection, the RHCs and the NA consider it important to be recognizable as a "part of a greater whole".
MARA Network concept	Perhaps this should be renamed "the national archive infrastructure".
MARA Use of generic building blocks	Generic building blocks are used to support archival institutions' shared services/functionalities.

4.3.3 *Principles for service provision and the exchangeability and findability of data and metadata*

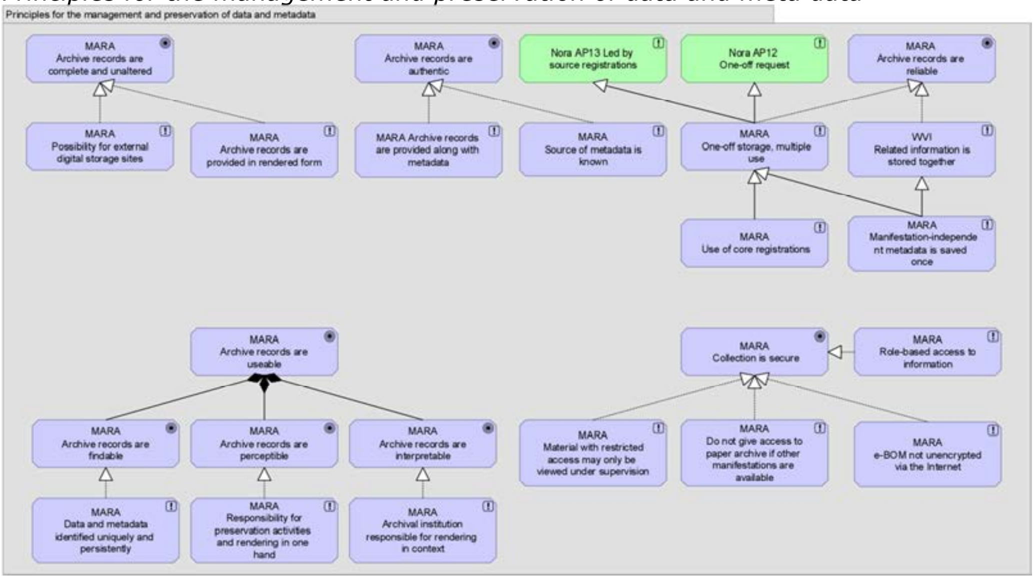

Principles for service provision and the exchangeability and findability of data and metadata
MARA Collections are connected
MARA Open data and open access
MARA Semantic and technical connection is possible
MARA Searching and finding through connection layer
MARA Public website gives access to National Collection
MARA Data can be exported
MARA Data and metadata are available as open data
NORA AP08 Use of open standards
MARA Sufficient metadating at the source
Nora AP12 One-off request
Nora AP13 Led by source registrations
Nora AP06 Use of standard solutions
Nora AP17 Information objects are systematically described
MARA Use of core registrations
MARA Archival institutions facilitate an active role for end users
MARA Public access can be established at the lowest level
MARA Use of domain-specific description standards

Further details:

Name	Documentation
MARA Semantic and technical connection is possible	It is possible to relate elements from collections to each other semantically and connect them to each other technically in order to present them in a coherent manner.
MARA Searching and finding through a connection layer	Search interfaces on a public website do not have a direct interaction with a collection management system. Each search query is conducted through an aggregator, so that it is always possible to connect metadata from several collection-managing applications to each other. This may involve several collections within one institution or one type of collection across several institutions (such as the National Collection).
MARA Public website gives access to the National Collection	The public website of each RHC and the NA gives access to the National Collection.
MARA Data can be exported	Data can be exported from each system where data (e.g. digital objects) or metadata are stored.
MARA Data and metadata are available as open data	If possible – within the applicable limits, statutory or otherwise – collection-managing institutions will make data and metadata available as open data via open access and in a standardized manner.
MARA Use of core registrations	When describing (metadating) archival and other material, maximum use is made of already existing and preferably shared core registrations/authority files. In other words: it is better to refer to than describe. We distinguish two main entities that have separate registrations: - Archives (= NEN-ISO 23081 Record) - Legal caretakers/archive creators (= NEN-ISO 23081 Actor) NB We do not distinguish "basic registrations" for the other entities mentioned in NEN-ISO 23081, such as Mandate and Activities.
MARA Archival institutions facilitate an	Archival institutions facilitate an active role for end users. Explanatory note: end users may play an active role in, for example, describing

active role for end users	and making connections between cultural heritage objects, i.e. crowdsourcing.
MARA Public access can be established at the lowest level	Public access can be indicated at every level (at the item level or even lower). Public access is inherited. If a level – for example, a file – has been designated as "public", it may contain public items that are not subject to limitations. If a level has been designated as "limited public access" this also applies to all underlying items. If both open access and limited access records are in one level, it is indicated at that level that there is no single definition: derivative code.
MARA Use of domain-specific description standards	If standards are available for the description of entities, we use these standards insofar as possible. This applies both to the description of objects and the export of the descriptions.

4.3.4 *Principles for the management and preservation of data and meta data*


Principles for the management and preservation of data and metadata
MARA Archive records are complete and unaltered
MARA Archive records are authentic
Nora AP13 Led by source registrations
Nora AP12 One-off request
MARA Archive records are reliable
MARA Possibility for external digital storage sites
MARA Archive records are provided in rendered form
MARA Archive records are provided along with metadata
MARA Source of metadata is known
MARA One-off storage, multiple use
WVI Related information is stored together
MARA Use of core registrations
MARA Manifestation-independe nt metadata is saved once
MARA Archive records are useable
MARA Collection is secure
MARA Role-based access to information
MARA Archive records are findable
MARA Archive records are perceptible
MARA Archive records are interpretable
MARA Material with restricted access may only be viewed under supervision
MARA Do not give access to paper archive if other manifestations are available
MARA e-BOM not unencrypted via the Internet
MARA Data and metadata identified uniquely and persistently
MARA Responsibility for preservation activities and rendering in one hand
MARA Archival institution responsible for rendering in context

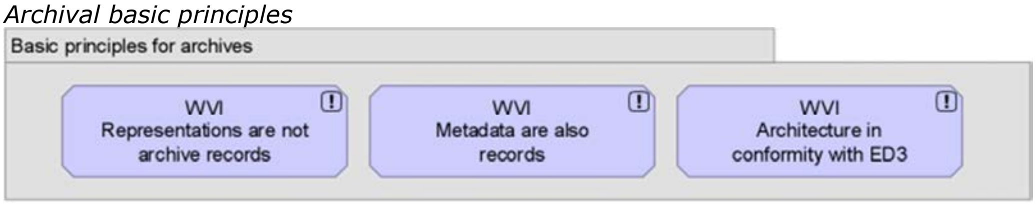
Further details:

Name	**Documentation**
MARA Possibility of external digital storage places	Transferred or outsourced material may be stored in physically separate environments. In this context, external means: not under the direct responsibility of the archival institution. It may be considered desirable to not physically transfer data to the e-Depot but to store and physically manage it in a different environment. For instance, it is conceivable that digital map material is transferred to the National Archives but physically stored in the Land Registry Office.
MARA archive records are made available after they are "rendered"	For archive records: The archival institution offers an incorruptible version of archive records. An archival item is created anew each time from data in addition to a rendering mechanism (plus the hardware and system software that the rendering software runs on). For a complete and unaltered version: - The archival institution may refer to a place containing a rendering application. This is possible for simple files, such as Acrobat Reader for PDFs , but not for all files.

	- The archival institution may refer to a place where the archival record can be viewed.
MARA archival records are made available along with metadata.	Archive records are provided in context, together with the accompanying metadata that are required to safeguard their authenticity.
MARA Source of metadata is known	It is known who has attached which metadata. For example, each time a description of an archival item or a relation is added, it is known who has done this and whether they are e.g. a visitor or an archival institution.
MARA One-off storage, multiple use	Ideally, all data are saved once (in one system) and used several times (through different applications or functions).
MARA Use of core registrations	When describing (metadating) archival and other material, maximum use is made of already existing and preferably shared core registrations/authority files. In other words: it is better to refer to than describe. We distinguish two main entities that have separate registrations: - Archives (= NEN-ISO 23081 Record) - Legal caretakers/archive creators (= NEN-ISO 23081 Actor) NB We do not distinguish "basic registrations" for the other entities mentioned in NEN-ISO 23081, such as Mandate and Activities.
WVI Saving related information together	All information that can be linked to a specific hierarchical level of an archive should also be saved at that level.
MARA Saving manifestation-independent metadata once	Substantive metadata are separated into information about sites and logistics concerning materials. Substantive metadata do not contain any information about the physical storage site. One component forms the primary source for the manifestation-independent metadata of all archives. The manifestation-dependent metadata is maintained in an appropriate location (e.g. technical metadata in the e-Depot).
MARA Data and metadata are uniquely and persistently identified	All information objects have a unique, persistent and readable web address/code. Each managed archival item has a universal and persistent identifier. An archival item is here: the place (URL) where it is presented, i.e. not the file. It is not enough for the identifier to be unique within one system; there must be a universal unique identifier.
MARA Responsibility for preservation activities and rendering in the hands of one party	Preservation activities, such as conversion and migration, are carried out by the manager of the external digital repository (and thus not necessarily by the party to which the material has been transferred). The institution performing data storage is also responsible for providing a rendering-option (viewer) to display the data object.
MARA Limited public access material is only accessible under supervision	Limited public access material may only be viewed under the supervision of the archival institution and may not be distributed or multiplied without authorization.
MARA Do not make paper	In general, if an archive has been digitalized or is available on a microchip, the

records available if there are alternative manifestations	paper manifestation is not made available for viewing.
MARA eBOM not unencrypted via Internet	There is no access to limited access material via the Internet.
MARA Role-based access to information	Further details: Not everyone needs access to all information. Access to information must be connected to a specific role. This means that information must be regrouped in line with the various roles.

4.3.5 *Archival basic principles*


Basic principles for archives
WVI Representations are not archive records
WVI Metadata are also records
WVI Architecture in conformity with ED3

5 Business architecture: services

This chapter describes the services that the RHCs and the National Archives can provide. They are summed up in the table below:

Service	**Clients**
Supervision	Legal caretaker (before transfer)
Giving advice	Legal caretaker (before transfer)
Presenting material	Public (archive consumer)
Making material from archival institutions available	Public (archive consumer)
Making data available for reuse	Public (re-user)
Management and preservation of outsourced archives	Legal caretaker (before transfer)
Facilitating viewing of outsourced archives	Legal caretaker (before transfer)

Not every archive service provides all of these services. The National Archives, for example, do not supervise and not every RHC provides the “Management and preservation of outsourced archives” service.

Below is a short explanation of the services.

5.1 Supervision

This service consists of monitoring legal caretakers’ compliance with the Public Records Act.

5.2 Giving advice

This service involves advising legal caretakers on matters such as compliance with the Public Records Act.

5.3 Presenting material

This service consists of making specific presentations, such as web exhibitions.

5.4 Making material available

This service is about making transferred archival material available to archive consumers.

An important aspect of this service is the basic principle that RHCs and the National Archives provide access to more material that just that of their own institution. In concrete terms, the RHCs and the National Archives have agreed to give access to all archival material from all RHCs and the National Archives on their public websites.

5.5 Making data available for reuse

Making data available for reuse is a new service. NB This is not the same as making archival records available: see more about this in Theme B Reuse of archive material.

5.6 Management and preservation of outsourced archives

This involves managing and preserving data and metadata of outsourced records for legal caretakers.

This is a new service, whereby legal caretakers can relinquish the management of archives that have not yet been statutorily transferred to the archival institution that will subsequently take care of its storage and preservation.

NB For more information about this service please see *National Archives PDC for departments* and *RHCs PDC for legal caretakers.*

5.6.1 Why is this service provided? Unburdening legal caretakers

There are two problems facing legal caretakers who have to take care of digital records that must be kept for a long time (for example, seven years or longer, or even permanently). First of all, digital records takes up storage space. Second, it is difficult for legal caretakers to keep the records permanently accessible, while this is a core task of archival institutions. File formats become obsolete and e.g. Document Management Systems (DMSs) are not made to show these file formats.

Archival institutions may offer legal caretakers the opportunity to outsource their archival records to their institution. The National Archives provide this service for legal caretakers from the central government, while RHCs provide this service to affiliated legal caretakers from the local authorities. The archival institutions take care of the management and preservation of these archive records. This management and preservation consists of:

- Providing storage for information objects;
- Adapting (descriptive) metadata, for example after the expiry of a period in which public access was restricted;
- Implementing preservation activities.

The original legal caretaker will remain responsible at all times for the outsourced archive.

There are two options for outsourced records:

- Either they are not eligible for permanent storage and are destroyed after a specific (long) period of time; or
- They are selected for permanent storage and are therefore formally transferred to the archival institution at some point in time.

It must also be possible for the legal caretaker to state that they no longer wish to make use of the "outsourcing" service and to resume looking after the records themselves.

5.6.2 Precondition: Files must be closed before they are outsourced.

The e-Depot is used to store records and make them available. Outsourced records are physically located in the e-Depot.

Reasons: The e-Depot is not a document management system. It does not have any functionality for processing or changing documents. This means that it is only suitable for housing closed records. If a civil servant wishes to adapt a record for use in a new work process, they can do so by downloading a file and processing it. The outsourced document will remain unaltered.

5.7 Facilitating viewing of outsourced archives

Civil servants may sometimes need a record for a work process. The objective is to ensure that it does not matter to a civil servant whether a record has been outsourced or is being managed by its own organization. They must be able to search for and find it in their work environment.

The RHCs and the National Archives may be able to facilitate search, find and display but they cannot regulate this entirely.

NB For more information about this service please see *National Archives PDC for departments* and *RHCs PDC for legal caretakers.*

5.7.1 *Precondition: As the legal caretaker must ensure searchability to give civil servants access they must therefore choose a suitable interface.*
It is possible that an interface enables metadata from several sources, such as a DMS and the e-Depot, to be searched through. As every legal caretaker is in a different situation, the National Archives and RHCs will not provide such a search interface. It may be possible to export metadata from the e-Depot so that it can be indexed by a search engine. It will also be possible to export an "up-to-date manifestation" from the e-Depot.

5.7.2 *Precondition: The legal caretaker is responsible for providing authorization and authentication.*
Outsourced records are not yet accessible to a large audience (this is not possible until they have been transferred). Civil servants may only see outsourced records if they are entitled to do so. The method for establishing who has access to which information and how this is monitored is the task of the legal caretaker.

6 Business architecture: organization and main functions

6.1 Introduction

This chapter describes which organizations and processes are needed to shape the vision. It mainly focuses on the relationships between the various institutions and their responsibilities.

6.2 Recipients of services provided by an archival institution

The following picture is based on the roles that can be distinguished[3]. These roles can be seen as parties that receive services from an archival institution. The distinguished roles are those of:

- Legal caretaker (before transfer)
  These are persons or institutions that transfer or outsource archive records. This role may view outsourced material.

- Archival institutions
  The RHCs and the National Archives.

- Other archival institution
  Also the RHCs and the National Archives. This has been included to show that one archival institution may advise another one, and that an archival institution may dispose of material to another archival institution.

- Archive consumer
  Persons interested in archive records. The difference with an archive re-user is that for the role of archive consumer it is essential that archival institutions provide a complete and unaltered version of the archive record.

- Archive re-user
  These are people or parties (public, publicly financed or private) who wish to reuse data or metadata in order to develop an end product for a specific market or to provide a service. Examples of reuse are: using an image to make a picture postcard, or using data to build your own application.
  Data is also made available to this role (in addition to complete and unaltered archive records). You can find more information about the difference between use and reuse of archive material and reuse of data in Theme B Reuse of archive material.

Organizations or people may fulfil different roles. For example, a municipal council may fulfil the role of archive creator, but also that of archive consumer if civil servants want to see their own material again. In general an RHC fulfils the role of archival institution, but may also occasionally fulfil the role of archive consumer if it organizes an exhibition and borrows archive material from another institution for this purpose.

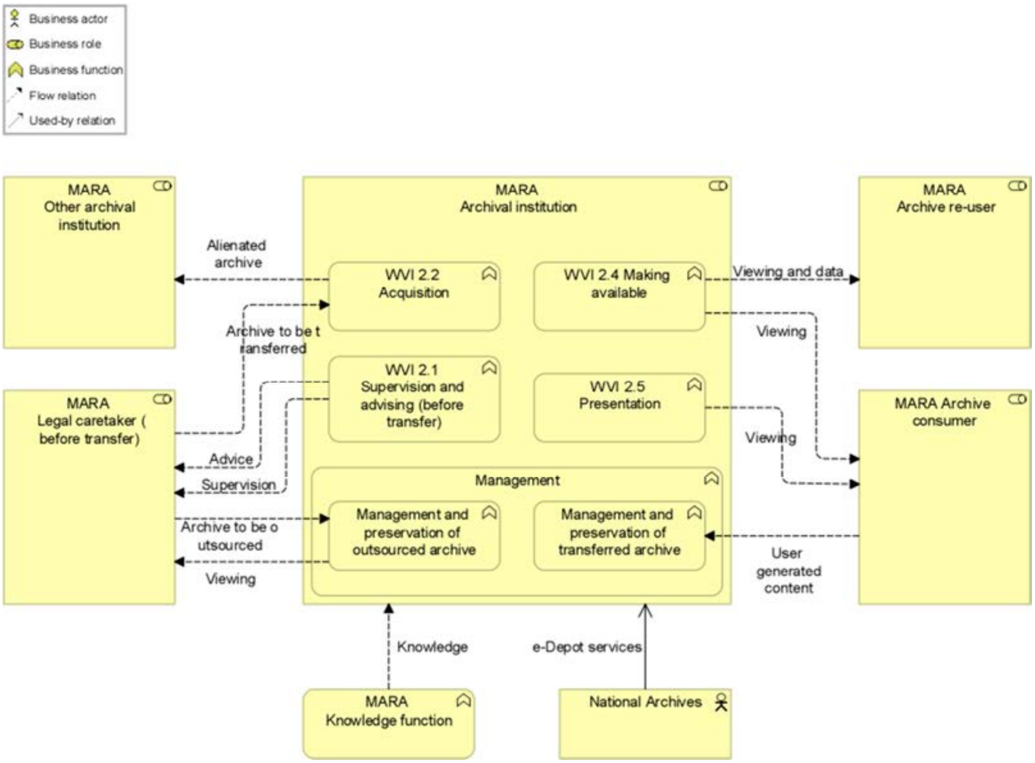
6.3 Main functions of an archival institution

The diagram below shows the main functions of an archival institution (National Archives and RHCs). They correspond to the services to be delivered as described in the previous chapter. These functions are:

[3] With regard to WVI architecture, the role of “Archive re-user” is new.

- Supervising and advising
- Acquisition
- Making archival material available
- Presentation
- Management; this also includes:
    - Management and preservation of outsourced archives
    - Management and preservation of transferred archives.

Not every archival institution conducts every function because not every archival institution provides every service. The National Archives, for example, do not supervise, and not every RHC may support the outsourcing of archives.



6.4 Use of shared services

To perform these functions, archival institutions use two types of services:

- Knowledge, supplied by a national knowledge network
- The National Archives' e-Depot services

6.4.1 Knowledge function: national knowledge network

The knowledge function will be performed by a national knowledge network. This will not only supply knowledge to the RHCs and the National Archives, but to the entire archive sector.

Four action lines will be used to set up a national knowledge network with regard to the archive function:

1. Set-up, implementation and creation of an annual *knowledge and innovation agenda* for the archive function;
2. Set-up of a national knowledge network using resources such as the creation of *knowledge platforms* (learning circles), a business newsletter, reinforcement of

knowledge sharing and safeguarding through the website, regularly organized knowledge sessions, and more interactive knowledge developed together;

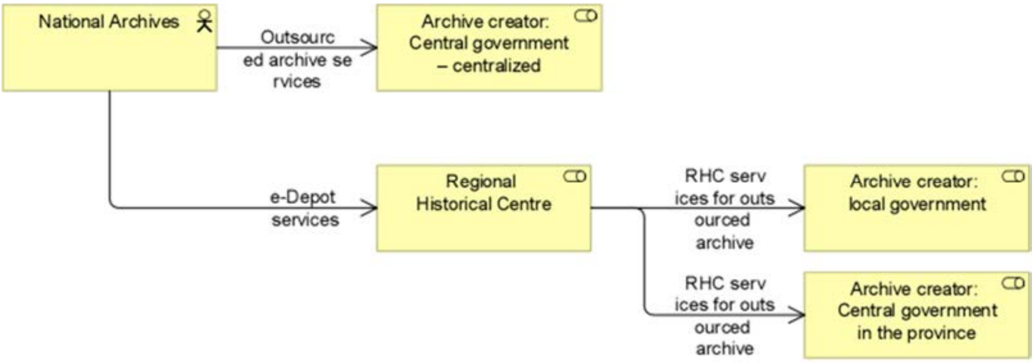
3. Connecting *external knowledge partners* (sectoral institutions, universities etc.) to the national knowledge network;
4. Reinforcing *knowledge management* within the National Archives and providing *best practices* in this area to archival institutions.

There is more information on setting up the knowledge network inAppendix V *Structure and organization of a national knowledge network*..

6.4.2 *Services provided to RHCs by the National Archives*

The National Archives will become a service organization for the RHCs. This means that the National Archives provide a number of services to RHCs, namely the e-Depot services, which, in turn, enable the RHCs (and the National Archives) to deliver the services described in chapter 5 *Business architecture: services* to their legal caretakers.

The diagram below shows this for outsourcing:


National Archives
Outsourc ed archive se rvices
Archive creator: Central government – centralized
e-Depot services
Regional Historical Centre
RHC serv ices for outs ourced archive
Archive creator: local government
RHC serv ices for outs ourced archive
Archive creator: Central government in the province

It has been agreed that the National Archives will not provide services to legal caretakers in the province of South Holland. On the basis of the network approach, the North Holland Archives will service local legal caretakers in the province of South Holland. The province of South Holland will, however, in its capacity as a legal caretaker, receive services from the National Archives.

With regard to the management of decentralized national archives the distinguishing criterion is whether storage is decentralized or centralized. If a decentralized government organization's files are also stored in a decentralized manner, they will be managed by an RHC. If a decentralized government organization (such as the Ministry of Justice) wants to make use of central storage, the National Archives will manage this. This means that a legal caretaker in the province will approach the National Archives for the outsourcing of closed files from a centralized storage site. This also means that if a decentralized government organization centralizes its storage, the National Archives will take over management tasks for the decentralized national archives for this organization from the RHC at some point.

There is more information on the National Archives' services in Appendix IV *Services provided by the National Archives*.

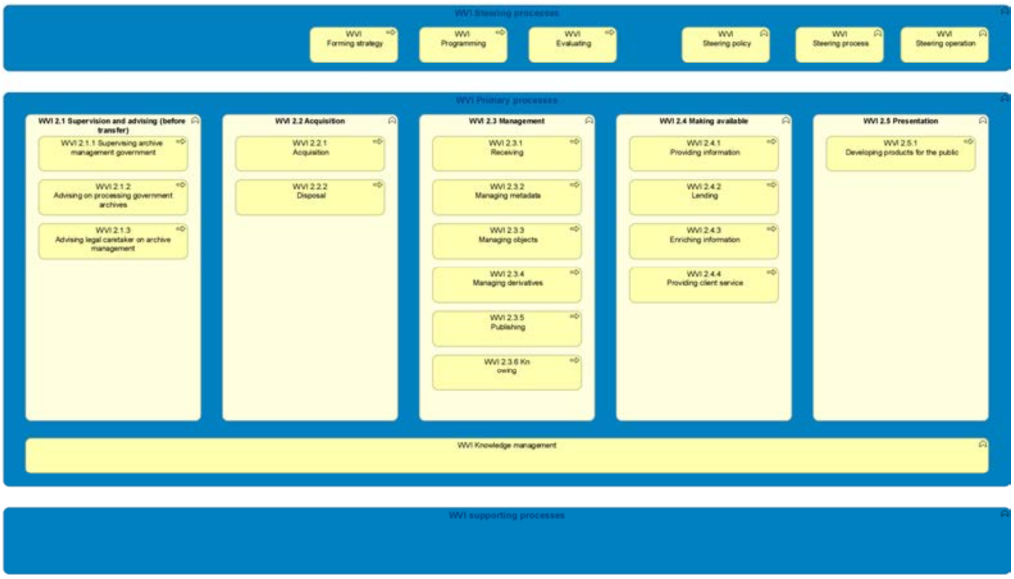
7 Business architecture: processes.

7.1 WVI process landscape

In 2013 the working group on WVI processes[4] developed generic processes for RHCs and the National Archives[5]. The processes identified in this WVI working group are described in detail in the separate appendix *Model Architecture for National Archival Institutions: models.*

This architecture focuses on the primary processes – see image:

Process landscape of RHCs as drawn up in working group on WVI processes


WVI Steering processes
WVI Forming strategy
WVI Programming
WVI Evaluating
WVI Steering policy
WVI Steering process
WVI Steering operation
WVI Primary processes
WVI 2.1 Supervision and advising (before transfer)
WVI 2.1.1 Supervising archive management government
WVI 2.1.2 Advising on processing government archives
WVI 2.1.3 Advising legal caretaker on archive management
WVI 2.2 Acquisition
WVI 2.2.1 Acquisition
WVI 2.2.2 Disposal
WVI 2.3 Management
WVI 2.3.1 Receiving
WVI 2.3.2 Managing metadata
WVI 2.3.3 Managing objects
WVI 2.3.4 Managing derivatives
WVI 2.3.5 Publishing
WVI 2.3.6 Kn owing
WVI 2.4 Making available
WVI 2.4.1 Providing information
WVI 2.4.2 Lending
WVI 2.4.3 Enriching information
WVI 2.4.4 Providing client service
WVI 2.5 Presentation
WVI 2.5.1 Developing products for the public
WVI Knowledge management
WVI supporting processes

7.2 Required adjustments

There are reasons for reviewing the process descriptions. First of all, other processes have been added, such as processes for providing services for the outsourcing of archives. Now that organizations are really connected to the e-Depot there will be more insight into processes related to the management and maintenance of digital material. The processes for putting the "knowledge function" in place have not yet been described.

This review has not yet taken place. Therefore the "WVI processes" are still taken as the starting point in this document. More detailed descriptions have been provided where expedient.

7.2.1 Primary processes for service provision

Generic processes for the management of outsourced material have not yet been developed. This involves matters such as metadata management, providing material to the legal caretaker, and destroying outsourced material with a limited storage period.

[4] Working group on Preparing Implementation; set up in 2011-12 to prepare for the implementation of the e-Depot.
[5] *Acquisition, management and provision. Guidelines for work processes for RHCs (2012).*

7.2.2 *Supporting processes for service provision*

If an archival institution is going to provide certain services, such as supporting outsourced archives, this will have an impact on the organization. For example, the following functions must be set up in it:

- Contract management
- Financing
- Relation management
- User consultation

The way in which these functions are set up may vary per RHC. This is why these functions have not been worked out in more detail in MARA.

Reaching an agreement on RHC-wide frameworks for these functions is worth considering. In that case these frameworks will also be incorporated into MARA.

7.2.3 *Processes for implementing the knowledge function*

The implementation of a shared knowledge function still requires further elaboration. This involves questions such as:

- How and by whom is which knowledge collected?
- How is the knowledge managed?
- How is the knowledge shared and made available?
- What are the agreements regarding finance and organization?

These points will also be included in MARA, once they have been worked out in detail.

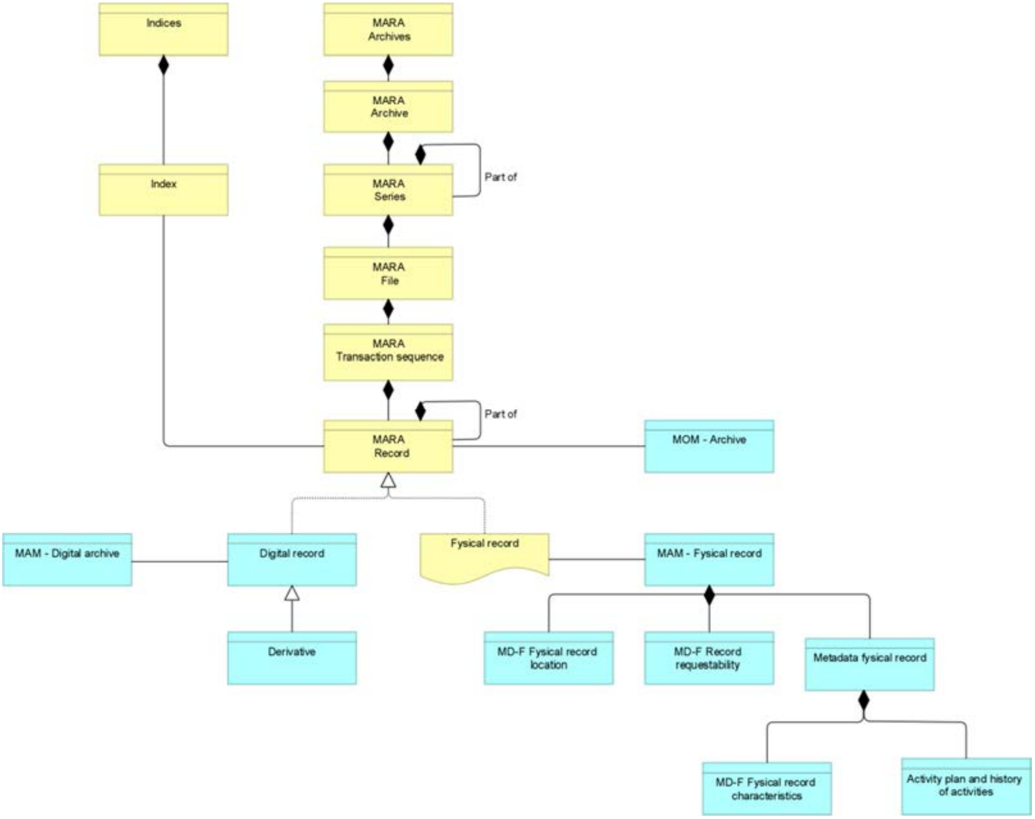
8 Business/information architecture: information and data

NB The images in this chapter are small, but are easy to read on a screen if they are enlarged.

8.1 Structure of an archive

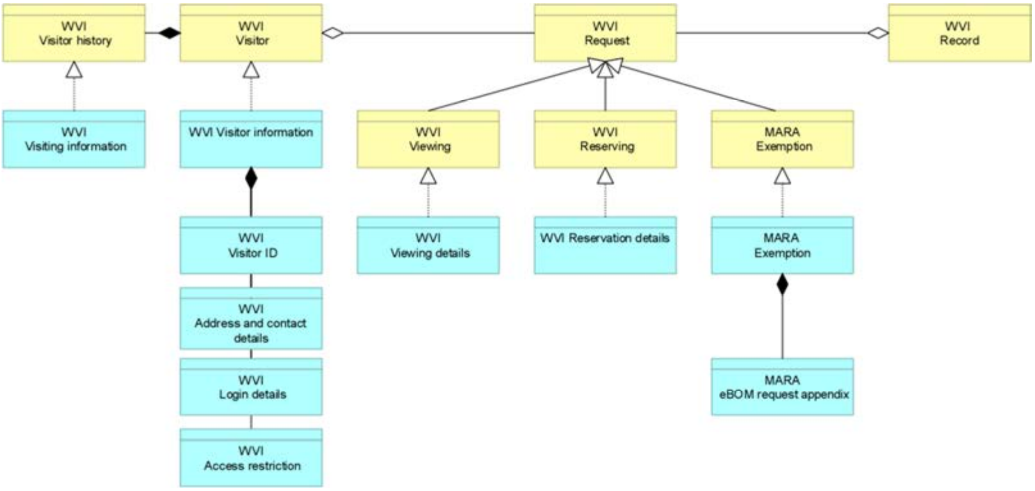
The following diagram shows that an archive (as defined in NEN-ISO 23081) is a hierarchical structure consisting of series and sub-series, that in turn consist of files. An item is the smallest separate unit of archive records that is managed as an entity. Items may contain components, like an e-mail with appendices. However, the components of an item are managed as a single entity in the system.

There may be several manifestations of one item, such as a physical item, a scan, a transcription etc. They may be analogue or digital.



8.2 Information from visitors and requests to be recorded

The following diagram shows the relationship between data from/about visitors, data from requests and the archive item.



8.2.1 *Visit history*

List of occasions on which a visitor visited the archival institution’s study room. A new record is made whenever the client visits the study room. The following information is registered:

- Date and time in
- Date and time out (assuming that it is noted when the visitor leaves, for example by returning an access pass).

8.2.2 *Visitor*

A visitor is someone who physically visits the archival institution’s study room, requests items online, places orders etc. insofar as this cannot be done anonymously. (NB Naturally, it is also possible to view public digital items online anonymously; no data is kept of anonymous “archive consumers”.)

8.2.3 *Request*

This is information regarding which visitors have requested certain items, both digital and physical. Historical information is kept so that it is possible to trace which visitor has requested items if something goes missing. The list contains references to the unique IDs of visitors and archive items.

NB It may be useful to record the following information on physical archive items:

- A reference to the abstract archive item
- A reference to the concrete item, e.g. a microfilm or paper record.

NB The application component that manages this list must perform a regular clean-up.

- **Viewing:** A list of items that the visitor has requested in the past. This can be combined with the reservations list.
- **Reservation:** A list of items that the visitor has requested in the past. This can be combined with the reservations list.
- **Dispensation (exemption):** A dispensation indicates that a visitor is entitled to view a specific limited public archive item. Dispensations contain:
    - A reference to the visitor
    - A reference to the archive item or items
    - The period for which the dispensation is valid

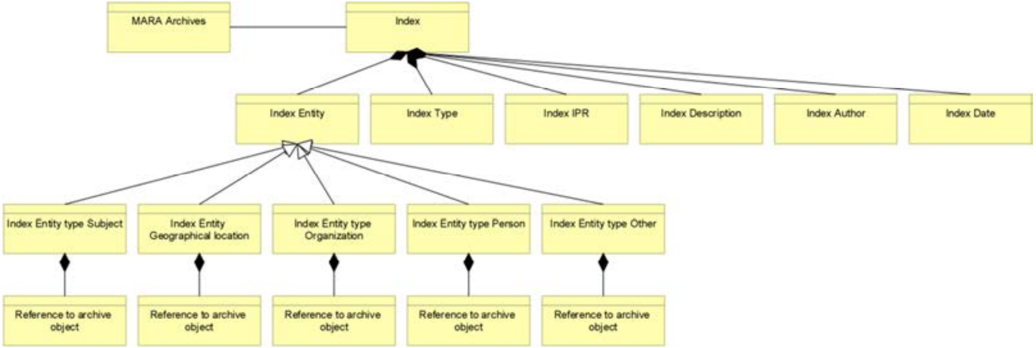
A dispensation is always granted for the highest level, i.e. if it states that a visitor has a dispensation for a certain level, they automatically have a dispensation for all items in the hierarchies under this level.

8.3 Information about indexes to be recorded

The diagram below shows the most common forms of indexes (specialised finding aids).

The most common types of entities are:

- Subject
- Geographical location
- Organization
- Person


MARA Archives
Index
Index Entity
Index Type
Index IPR
Index Description
Index Author
Index Date
Index Entity type Subject
Index Entity Geographical location
Index Entity type Organization
Index Entity type Person
Index Entity type Other
Reference to archive object
Reference to archive object
Reference to archive object
Reference to archive object
Reference to archive object

8.4 Other collections

Archival institutions do not just manage archives. It is possible to distinguish different types of collections. The most common are:

- Archive, transferred and outsourced
- Maps
- Library material
- Newspapers
- Museum objects
- Photographs
- Audio-visual material

Archival institutions manage the physical objects, digital manifestations and the metadata belonging to the aforementioned categories.

9 Information architecture: functionality and applications

9.1 Introduction

This chapter lists the required functional components/services (broadly outlined). The following chapter looks at the specific solutions used for the component.

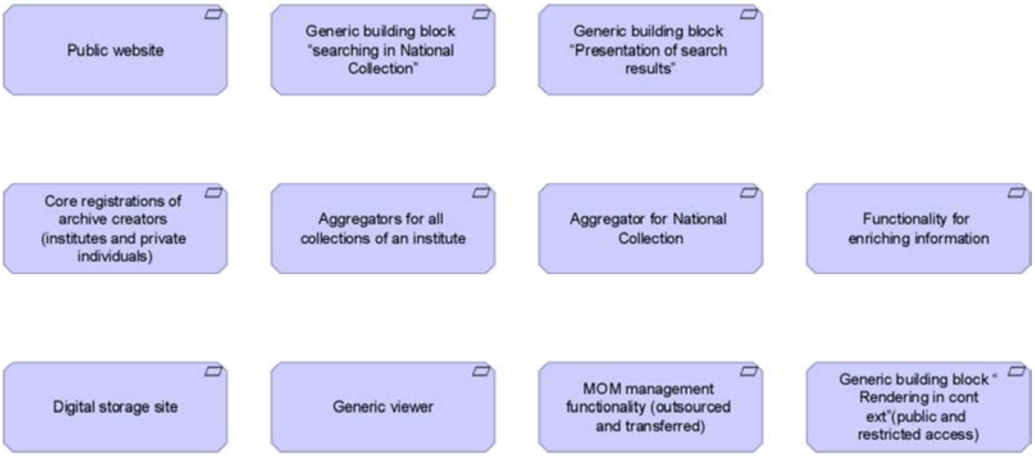
Requirements and restrictions are given for each functionality where applicable. A requirement is a functionality that must be performed by a system; a restriction is a limitation in the way in which a system is created.

9.2 Summary of required functionalities

NB Categorized according to the NDE three-layer model; see Theme F *Implementation of the three-layer model*)

- Functionality for long-term availability of archive records:
  - Digital storage site
  - Management of manifestation-independent metadata
  - Rendering of digital manifestation
  - Rendering of digital manifestation in context (public access and limited access)
- Functionality for enriching and connecting information
  - Management of actors
  - Enrichment of information (for example, indexes)
  - Aggregation of information, including accessibility of National Collection
- Functionality for visibility
  - Public website, into which generic building blocks have been integrated

An illustration of this:


Public website
Generic building block "searching in National Collection"
Generic building block "Presentation of search results"
Core registrations of archive creators (institutes and private individuals)
Aggregators for all collections of an institute
Aggregator for National Collection
Functionality for enriching information
Digital storage site
Generic viewer
MOM management functionality (outsourced and transferred)
Generic building block " Rendering in cont ext"(public and restricted access)

9.3 Functionalities for long-term availability of archive records

9.3.1 Functionality: Digital storage site

This entails maintaining the material that is stored here. This function provides for the storage of digital originals (born-digital material) and other manifestations, such as representations in A2A format[6], scans and any derivatives (presentation examples, such as JPEGs).

The functionality can automatically implement preservations activities on a specific set of material, for example, as a result of a trigger, such as passage of time, or offer possibilities that enable officials to conduct these activities. Preservation activities include implementing conversions or migrations of file formats.

This function may supply an "active" manifestation of a digital archive item (the bitstream) plus the accompanying metadata required to be able to show (render) this example. This function does not itself provide rendering.

The digital storage site also plays a role in bit preservation (physical storage), recovery and contingency use. The actual bit preservation is the responsibility of the physical storage media (servers and tapes), and the digital storage site checks if this is being done properly.

This function entails the management of manifestation-dependent metadata, such as metadata on file formats and other information required to present the records.

9.3.1.1 Requirement: Open Data interface

The digital storage site must have an interface through which public data can be accessed using open standards.

9.3.2 *Functionality: Management of manifestation-independent metadata*

This function entails the management of manifestation-independent metadata, i.e. metadata that does not depend upon its manifestation (paper, digital or otherwise). This may include information such as the archive creator (NB through a reference; the data on the archive creator is stored in a register of actors), period, subject, public access, etc. This component also knows the existing manifestations of archive items, e.g. paper, microchip, digital.

This function acts as a guide, telling you what you have and where it is.

- There are two possible options for digital:
  - The component knows the storage sites:
    The component knows the references to locations where digital manifestations are stored, so it knows that manifestations exist, but it does not know which ones (e.g. whether they are TIFFs, JPEGs, etc.): this is the responsibility of the Digital Storage site. This component may ask the Digital Storage Site to supply information on the digital manifestations present and then pass on this information.
    This means that this component must know the various digital or analogue storage sites containing manifestations. In this case, this function must have access to the services of the Physical Metadata function and the Digital Storage Site.
  - The component knows the manifestations:
    The component knows the references to specific manifestations. It therefore has persistent references to the manifestations, such as TIFFs, JPEGs, etc.
- For physical: this component knows that physical manifestations exist, but not how many and of what kind (e.g. paper or microfiche): this is the responsibility of the Physical Metadata Component.

[6] The A2A data model is a generic metadata format that is used to exchange and disclose different sources containing historical personal data. For more information see http://www.den.nl/standaard/386/.

9.3.2.1 Requirement: Reference to core registrations
This functionality must be able to refer to information in core registrations for all fields.

9.3.2.2 Requirement: Interface for export

9.3.2.3 Restriction: Export via accompanying metadata standard

9.3.3 Functionality: Rendering of digital manifestation
NB See Theme A *Presentation of a digital archive item – integrity and authenticity* for more information on showing a complete, unaltered and authentic archive item.

This entails ensuring a "perceptible" version of the manifestation, e.g. a record displayed on a website.

This functionality takes a bitstream and accompanying presentation metadata and can then call up an application to show this bitstream.

All institutions with a Digital Storage Site that contains certain material must ensure that it is "rendered" somewhere. The rendered record can then be shown on a presentation website – provided with metadata, for example. You can compare this to embedding a video on YouTube: YouTube has the files and ensures that they can be shown in a viewer that you can, for instance, fast forward.

9.3.3.1 Requirement: Providing the rendered manifestation with a permanent and persistent identifier
To enable reuse, it is necessary to provide the rendered manifestation with a permanent and persistent identifier.

9.3.3.2 Restriction: Rendering of publicly available material through the web
The rendered version of publicly available material must be accessible through the web.

9.3.3.3 Restriction: Rendering of material that is restricted for the public not through the web
The rendered version of material that is restricted for the public must not be accessible through the web. For reasons of security, this application or the component performing this function may only be used in the archival institution (e.g. in the study room). Logging in to a website that can be accessed from a distance is not sufficiently secure, as users would be able to take photos or print the screen.

9.3.4 Functionality: Rendering of digital manifestation in context
There must be a functionality to show the manifestation in context, i.e. in combination with its accompanying metadata. After all, archival institutions want to provide a complete, unaltered and authentic version.

9.3.4.1 Restriction: Can be incorporated into public websites of archival institutions
This component must be constructed so that it can be easily incorporated into the public website of an archival institution.

9.3.4.2 Restriction: Rendering of publicly available material in context through the web
The rendered version of publicly available material must be accessible through the web.

9.3.4.3 Restriction: Rendering of material that is restricted for the public not through the

web
The rendered version of material that is restricted for the public must not be accessible through the web. For reasons of security, this application or the component performing this function may only be used in the archival institution (e.g. in the study room). Logging in to a website that can be accessed from a distance is not sufficiently secure, as users would be able to take photos or print the screen.

9.4 Functionality for enriching and connecting information

9.4.1 Functionality: Management of actors

NB See Theme D *Use of core registrations*. for more information on using core registrations in general and managing actors in particular.

This concerns a functionality for managing private and institutional archive creators. NB It is possible that this functionality could be performed by more than one application: for example, one application for managing institutional archive creators and another one for managing private archive creators.

9.4.1.1 Requirement: Joint use by multiple institutions
This is ideally a core registration that can be used by several institutions. It is necessary to make agreements on the management of the registered information.

9.4.1.2 Restriction: Management in accordance with ISAAR / EAC

9.4.2 Functionality: Enrichment of information

Users must have the opportunity to enrich information, for example by transcribing or providing extra metadata.

9.4.2.1 Requirement: When enriching information, users must be able to use ontologies or thesauri

9.4.2.2 Requirement: It must be known at all times who has supplied which information
Rationale: An end user or re-user must be able to assess the value of the information. For example, it is helpful if it is known whether information has been validated by a specific organization, such as an archival institution. The value attached to information added by “the public” is a consideration that the user or re-user must be able to make for themselves. In other words, the metadata must also be provided with metadata.

Implications:

- A process is required to validate information from third parties
- Storage options are required for information added by third parties

9.4.3 Functionality: Aggregation of information

This is a functionality to physically collate metadata in order to make it searchable. It involves:

- Bringing together several collections from one institution
- Bringing together collections from several institutions. A specific requirement concerns making the National Collection searchable.

Various solutions are possible for this. Aggregators are the obvious short-term solution. See Theme E *Connection layer: aggregators versus Linked Data*.

9.4.3.1 Requirement: Generic building block “Searching in the National Collection”

9.4.4 Accessibility of the National Collection

9.4.4.1 Description

The National Archives and the RHCs all manage parts of the National Collection. The National Collection consists largely of archive material from Central Government organizations that are or were housed in the provinces, such as law courts and offices of the Directorate-General for Public Works and Water Management. The aim is for the National Collection to be “integrally searchable” (see master plan).

9.4.4.2 Elaboration:

The term “integrally searchable” can be interpreted in a variety of ways. It has been interpreted here (made SMART) as follows:

- There is a search box on the website of every RHC and the National Archives that enables users to search the entire National Collection. This search box has the same functionality for every RHC (e.g. the same possibility to fill in search terms, filter results, etc.).
- The concept of the “National Collection” is a theoretical term. In the actual items, or their metadata, it is usually not possible to see whether or not an item belongs to the National Collection. Furthermore, “National Collection” is not a term that appeals to the public. It has therefore been chosen to ensure that the entire digital collection of the RHCs can be searched via the search box, on the basis of the line of thought that if you can search through everything, then the National Collection is sure to be included.

9.5 Functionality for visibility

9.5.1 Public website

Every archival institution has its own public website. This is the website through which every archival institution communicates with the users of the collection. Services are provided through the public websites, such as searching in the collection, showing the collection, reserving, ordering, and enriching.

9.5.1.1 Restriction: Must depend on a specific framework as little as possible

9.5.1.2 Restriction: Make it possible to use generic building blocks

10 Standards

10.1 Introduction

Overall, open standards must be used as much as possible for exchanging data. A "comply or explain" policy applies to the use of standards. This means that RHCs and the National Archives must comply with these standards in principle.

I.1.1 What are standards?

Standards are agreements set out in specification documents. An interface or exchange format (file) of an IT system must be implemented in full compliance with the relevant specification document. Only then will it be possible to exchange data faultlessly with other systems that support the standard[7].

I.1.2 What are open standards?

Open standards are available to the public. The standard's specifications may be applied, handled and used free of license rights. The term is mainly used for hardware and software because they use many closed standards for which one must request a licence in order to view the specifications[8].

The objective of open standards is exchangeability between various information systems or, even more important, to increase data collection and to enable humanity to record information in a future-proof format. A side effect of open standards is more freedom of choice regarding (and therefore less dependency on) suppliers [9].

10.2 Standards to be used

The following tables show the available standards.

The following table shows standards related to the semantics, syntax or transport of elements from a **collection type**.

	General & image	Archive	Archive (dig.-born)	Libraries	Museum
Semantics	Dublin Core, VRA Core	ISAD(G)	NEN-ISO 23081	ISBD	
Syntax	METS	EAD (+ national guidelines)	TMLO	MODS, MARC21XML	LIDO
Transport	OAI-PMH, Open Search, CMIS				

The following table shows standards related to the semantics, syntax or transport of elements from **core registrations**. These standards therefore apply to the recording and exchange of information (about an entity in the core registration) between the collection management system and the core registrations.

	General & image	Archive creators	Function	People	Geo
Semantics		ISAAR (CPF)	ISDF		GeoNames

[7] Source: https://www.forumstandaardisatie.nl/open-standaarden/over-open-standaarden/

[8] Source: https://nl.wikipedia.org/wiki/Open_standaard

[9] Source: https://nl.wikipedia.org/wiki/Open_standaard

Syntax	MADS	EAC-CPF	EAC-F (in development)	A2A	
Transport	OAI-PMH, CMIS				

11 Shared facilities: how this will be achieved

This chapter looks at how the required functionalities described in the previous chapter will be implemented. It is mainly a summary of the programme architecture of the DTR programme, specifically work package 1 of this programme in which these solutions will be implemented and activities related to Archief2020.

11.1 Solutions for the long-term availability of archive records

11.1.1 Solution: Digital storage site for transferred material

The e-Depot is used for this purpose. Being researched: collaboration with other data managers, such as Beeld & Geluid or the Land Registry.

11.1.2 Solution: Management of manifestation-independent metadata

- The RHCs and National Archives use one or more collection management systems for descriptive metadata of *transferred* archive material and other material (e.g. photos, books etc.). The Management of manifestation-independent archive metadata function will be assigned to the relevant collection system.
- The e-Depot is used for the descriptive metadata of *outsourced* archive material.

11.1.3 Solution: Rendering of digital manifestation

- For material publicly available in the e-Depot: e-Depot rendering framework
- Ideally, the same solution is used to render both material in the e-Depot with restricted public access and material that is openly available to the public.
- If material is stored at a digital storage site other than the e-Depot, this is the responsibility of the organization managing this digital storage site. For example: if audio-visual material is stored in Beeld & Geluid, then Beeld & Geluid is responsible for supplying a rendering function for it.

11.1.4 Solution: Rendering of digital manifestation in context

This functionality is not yet available. A generic building block is being developed for it.

11.1.5 Solution: Management of actors

This functionality is not yet available. For institutional actors it may be possible to further develop the existing register of actors, "Tasks and Organization of the Central Government" (TOCO) which is only used to a limited extent. For private archive creators it may be possible to develop the biographical portal.

11.2 Solutions for recording metadata of transferred material

We assume that the e-Depot application provides a framework for the "Digital Storage Site" component, and therefore is

- the location where the digital manifestation is stored, and
- the location where manifestation-dependent metadata for digital material is stored.

11.2.1 Exceptions to the separation of MOM and MAM

11.2.1.1 <u>"Public access" field</u>

The "public access" field is MOM but will also be stored in MAM. There is a technical reason for this, namely the two logical installations of the e-Depot application. See chapter 12.4 Data storage – e-Depot.

The criterion for whether or not to copy something is the public access coding. If this is only stored in the collection management system (CMS), communication is required for every record to request whether or not a record is publicly available. This is not desirable from the perspective of performance considerations: it must be possible to make copies independently of the CMS.

11.3 Solutions for recording metadata of outsourced material

It is not yet clear what the management of metadata of outsourced material actually entails. The estimate is that mainly the public access and possibly the storage period fields will be adjusted. Following on from these principles, the obvious thing to do is to store manifestation-independent metadata in the collection management system. This would require a transformation from TOPX to EAD. As it is not yet possible to assess what this would entail, while at the same time it has been estimated that metadata of outsourced material is only changed to a limited extent and in general there is only one manifestation of outsourced material, it has been decided in the meantime to store manifest-independent metadata of outsourced material in the e-Depot.

The implication is that there is no full overview of all the managed (transferred or outsourced) material in the collection management system (CMS): the CMS does not yet know the outsourced material. The implication is that the functionality concerning changing metadata , reporting items to be destroyed and removing these items (with the agreement of the manager) must be built into the e-Depot application.

11.4 Solutions for the aggregation of archives

APE is used as an aggregator. Users will be able to search for and find archive material through an APE widget that has not been built yet.

Explanatory note: searches and results are currently available throughwww.archivesportaleurope.eu. Users should not be aware that they have left the archival institution’s website. Furthermore, searching, finding and showing archive material, including archive material of the central government, should look the same on every RHC website. For these reasons, generic building blocks (cf. “widget”) are being built to make it possible to enter a search query and display search results. A generic building block is being developed to present displayed material in context.

All RHCs implement these building blocks in their public websites.

11.5 Solution for viewing outsourced material at departments

This paragraph describes how the National Archives make outsourced material available to legal caretakers.

NB This concerns the services provided by the National Archives to the departments. This does not automatically mean that RHCs will choose the same solution for their service for allowing local authorities to view outsourced material. Strictly speaking, this means that this paragraph is not part of this architecture. However, it has been included as an example and to provide inspiration.

11.5.1 Example: specific organization of architecture at the National Archives

This paragraph describes how the National Archives make outsourced material available to legal caretakers.

NB Strictly speaking, the following paragraph is not part of MARA, in the sense that it is not a generic solution that all RHCs must follow when making outsourced material available. It can, however, serve as an example.

When making outsourced material available, it is necessary to decide at the personal level whether or not the material will be issued, depending on whether the person in question is entitled to see it.

An Access Control List (ACL) has been included in the metadata to make it possible to determine access rights. The ACL contains *groups* that may see the document. Another requirement for outsourcing is synchronizing metadata. The National Archives/e-Depot do not do anything with this metadata regarding their content.

There is a search engine (in the case of the departments: Autonomy) which has access to a ministry's data through an interface. Not all data may be found by the search engine. A security tag indicates whether or not data may be found by a search engine.

The search engine may index items it has access to, including the ACL, which is contained in the metadata.

A civil servant looking for a specific item will use the search engine, which "knows" whether or not the item may be shown to the civil servant. The search engine has a *group service* that can decide whether or not a specific user is part of a group. Items are picked up by the search engine. There must be no direct communication between the civil servant and the e-Depot.

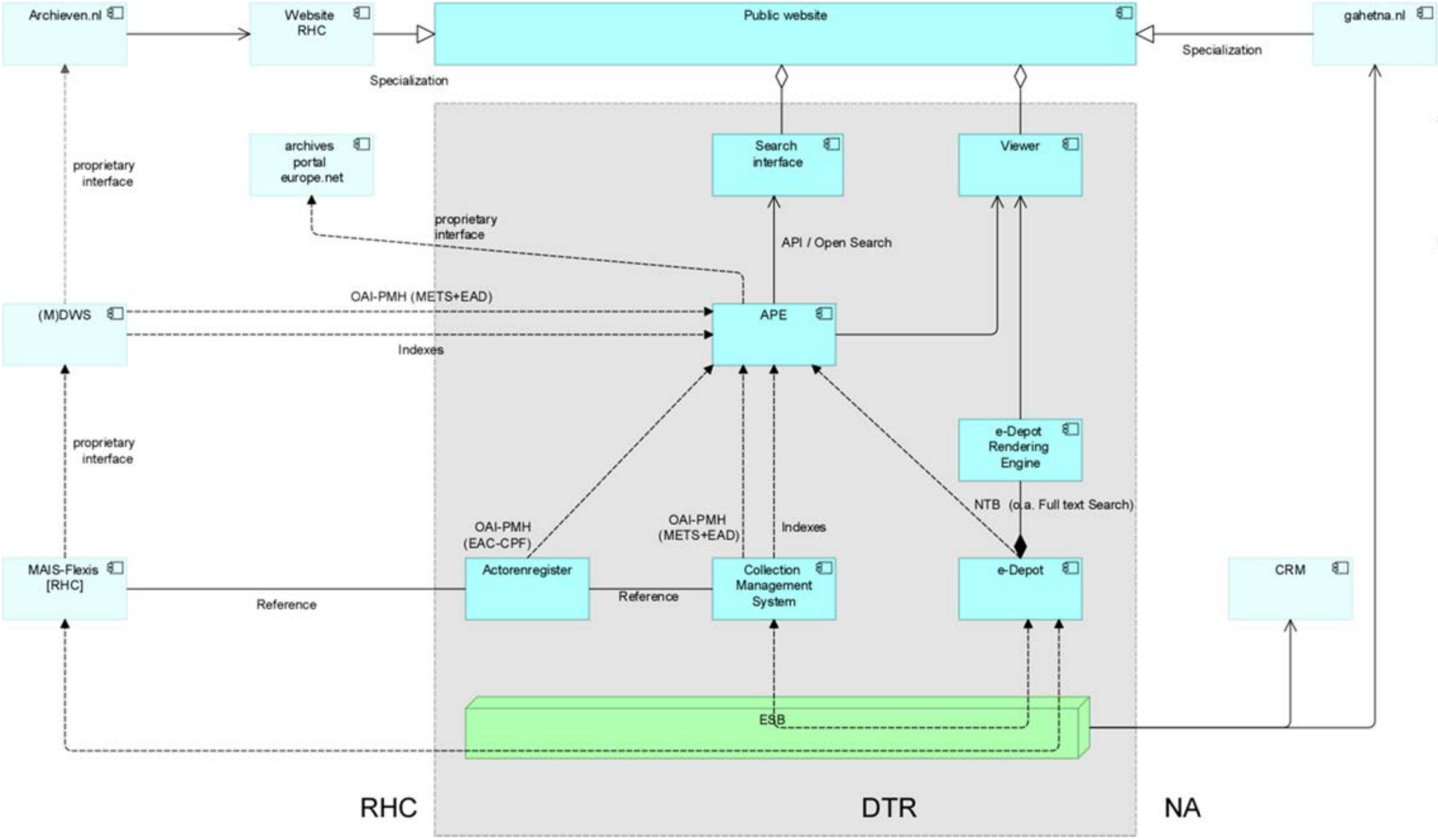
This implies that rendering cannot be done by the e-Depot, but occurs in the civil servant's environment. This solution is therefore not entirely in accordance with the principle of providing incorruptible and authentic archive material, but it is the maximum that is technically possible at the moment.

11.6 Solutions for visibility

11.6.1 Solution: Public website

Every RHC and the National Archives have their own public website. This is bound by certain requirements (such as the use of generic building blocks); see the previous chapter.

I.2 Summary in a diagram



12 Technical infrastructure

12.1 What is technical infrastructure?

“Technical structure” is understood as the collection of facilities that are required for the storage and transport of digital data. It includes all physical and technical resources that move, distribute, route and record the (photo) electrical signal (as data carrier). In concrete terms, this refers to all network, server and storage equipment, and accompanying software, configurations and data connections, server space with standard and emergency power supplies, and back-up and contingency facilities.

12.2 Objective of this chapter

This chapter is also a reference and gives a number of concrete details. The requirements are generic and apply to all RHCs and the National Archives. When referring to data that is physically managed by the National Archives, as is the case with the e-Depot, the solution will also be described. It is the task of the RHCs and the National Archives to organize a solution for the other data that meets the requirements.

The following topics will be dealt with:

- Objective – to guarantee business continuity
- Data storage – e-Depot
- Data storage – infrastructure
    - Hosting and housing
    - Back-up, contingency use and recovery
- Data transport – infrastructure
    - Zoning (including firewalls)
    - Network (data connections)
- Development and management of systems
    - DTAP

For each topic the following will be provided:

- Short description
- Requirements
- Preconditions (if any)
- Solution

12.3 Objective: Safeguarding business continuity

12.3.1 Description

IT systems must have such a reliable level of availability and reparability that the National Archives and RHCs are not hindered in the performance of their core tasks, including the services provided by the National Archives to the RHCs, and the services provided by the RHCs to their clients.

12.3.2 Achieving the objective

With regard to the shared applications, such as the e-Depot, the National Archives will ensure a reliable infrastructure that provides services according to the agreements. This means creating housing and/or hosting for IT systems as well as organizing back-up and contingency facilities and creating recovery facilities. These topics will be dealt with below.

12.4 Data storage – e-Depot

12.4.1 Description

Archive material is stored in the e-Depot, from where it can be requested later. When requesting data, there is an important distinction concerning the type of data. This concerns the public accessibility of data, which indicates whether data can be accessed freely by the public or if access is limited.
In addition to distinctions with regard to the public accessibility of data, the ability to ingest data and access at a specific speed also play a role. This speed must be acceptable within the set agreements and may not be impaired if the system is burdened more heavily due to an increase in access or ingest.

12.4.2 Requirements

In other words, the e-Depot must:

- be able to provide requested archive data in an authorized manner (security);
- have and maintain the ability to perform in case of an increase in requests for data (performance).

12.4.3 Solution

The e-Depot has set up two logical e-Depot structures;

- one logical structure contains all the material (e-Depot everything);
- one logical structure contains only material accessible to the public (e-Depot Public Access);

See the diagram below. When data is ingested into the e-Depot, the metadata of all the material is included in the “everything” structure. The metadata of the material open to the public ends up in the “public access” structure. Only the public access database can be accessed through the Internet. As the link to restricted public material simply does not exist in the “public access” structure, this means that restricted access material can never be disclosed via the Internet.

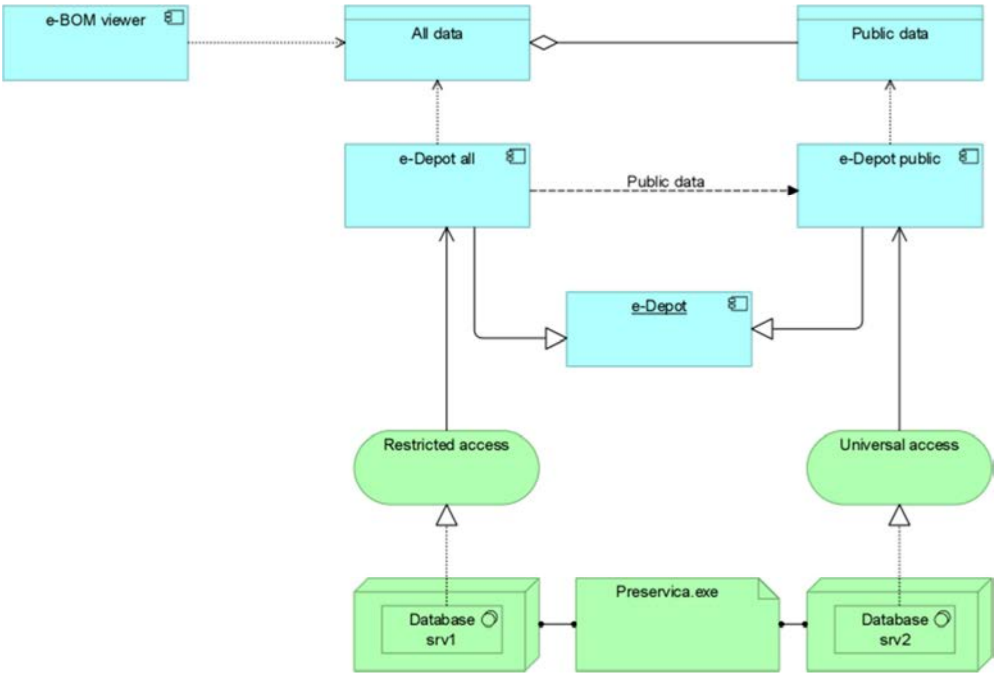
The second advantage of this configuration is that the access side (public) is not burdened by heavy ingests (as these go into the “everything” structure).

NB Physically there is only one storage: only metadata (such as references to files) are saved twice, the documents are only saved once. The extra costs for double storage of metadata are negligible.

The implication is that a copy session must be effected from the full database to the public database. Various solutions are possible for this.

12.4.4 Access to eBOM

To make limited access material available, access is granted from an RHC to the “everything” installation, which houses limited access material. Access can be protected by stating exactly who is granted access based on IP address.



12.5 Data storage: Hosting and housing

12.5.1 *Description*

Housing makes it possible to incorporate one's own IT systems (servers) into a Data Centre. A Data Centre provides an optimally-secured system environment.

Hosting or dedicated server hosting means using a hired service. The main advantage of dedicated server hosting is that the entire infrastructure is outsourced (both hardware and software by "virtualization")

Another version of housing and hosting is the cloud facility. With "cloud services" clients only receive functions. In principle, a data centre can place servers all over the world without this causing any nuisance to the client.

For more information, see Theme G Further details on the technical infrastructure.

12.5.2 *Requirements*

Availability of infrastructure (on an annual basis):
- On the basis of unplanned unavailability of 22 hours = 99.479% uptime
- On the basis of planned and approved unavailability of 48 hours = 99.452% uptime

Total availability of infrastructure (on an annual basis):
- On the basis of unplanned and planned unavailability of 70 hours = 99.202% uptime

For a description of these requirements, see G.6.1 Reliability.

12.5.3 *Precondition: government data centres*

At the central government level, the cabinet has decided on four government data centres (ODCs), connection to which is compulsory.

12.5.4 *Solution*

The National Archives will receive services from ODC North, related to hosting and housing. For a description of this solution, see G.2 Organization of ODC North.

12.5.5 *Gradual transition through a temporary solution*

In view of the scope and complexity of the DTR programme it is not desirable to migrate to a government data centre during the programme as this is also a complex project.

The National Archives will retain the existing server space within their building for the duration of the DTR programme in order to be able to fulfil the obligations ensuing from the programme. The National Archives will start migrating to an ODC immediately after rounding off the programme (2017). For further explanation, see G.5Rationale for gradual transition by means of a temporary solution.

12.6 Data storage: Back-up, contingency use and recovery

12.6.1 *Description*

Stored data must not be lost. Therefore, data is saved several times at different locations. In the event of an emergency at one location, it is always possible to recover a complete set of data (back-up), or if the situation is even worse, services and data can be made available at a contingency location. For a more detailed explanation, see G.3 *The concepts of back-up, recovery and contingency*.

12.6.2 *Requirements*

Back-up:

- A full set of data is present in at least three physical locations.
  Rationale: recovery in the event of a disaster. If a serious emergency occurs at one of the locations, it will take a few weeks or months before the site is up and running again. During this time, it is not desirable for the data to be available at only one location.
- There must be two sets of data at at least 5 km distance from each other.
  Rationale: to prevent a disaster taking out two locations simultaneously. NB 70 km is the maximum distance for synchronized storage, as the latency (delay) of glass fibre cables is too great above this distance.
- There must be a complete set of data on at least two different technologies.
  Rationale: there may be an error in the technology, or the supplier of a specific technology may go bankrupt. If the data is only stored in one technology, a “challenge” may arise.

Contingency:

- One set of data must be stored at at least 50 km distance from the other two locations.
  Rationale: if it is further away, there is less risk of a power cut, natural disaster etc.

Data recoverability:

- RPO (Recovery Point Objective) = 7 days for the e-Depot.
- RPO (Recovery Point Objective) = 1 day for other data.

Recoverability of the archive management system:

- RTO (Recovery Time Objective) = 5 days for the archive management system’s physical collection (CMS).

Recoverability of the entire National Archives IT facilities (all data and infrastructure):

- MTPOD (Maximum Tolerable Period of Down time) = 6 months.

For a more detailed explanation, seeG.6.1 Reliability.

12.6.3 *Precondition: government data centres*
At the central government level, the cabinet has decided on four government data centres (ODCs), connection to which is compulsory.

12.6.4 *Solution from 2017*
From 2017, data services, including back-up and contingency, will only be facilitated by ODCs. In concrete terms, this means the following:

- primary data in an ODC (>50 km away from ODC North);
- back-up in ODC North, location 1;
- back-up in ODC North, location 2 (> 5 km away from location 1)

12.6.5 *Solution to the risks posed to business continuity*
To guarantee business continuity, the National Archives have commenced preparations for connection to an ODC by first putting their contingency facility and back-up in place in an ODC. For a description of this solution, see G.5 Rationale for gradual transition by means of a temporary solution.

12.7 Data transport – zoning

12.7.1 *Description*
The network infrastructures of the various data are compartmentalized into zones. A zone is a defined network of IT facilities where data may be exchanged freely. Data exchange with other zones is protected by firewalls. The primary objective of zoning is to isolate risks so that threats and incidents in one zone cannot spread to other zones. The National Archives uses two zones:

- A Demilitarized Zone (DMZ; semi-trusted zone): a border zone or linking zone between zones with differing confidentiality. This is used for web hosting, data and file exchange, and authentication links via reverse proxy.
- Internal National Archives (production environment; trusted zone): Controlled zone for application and database servers and systems. Logically, the DMZ and the Internal National Archives network zone are separated by a firewall.

For a more detailed explanation, seeG.4.1 Description of zoning.

12.7.2 *Requirements*
Network zoning is structured in accordance with the National Archives’ security requirements, which were drawn up in conformity with the BIR. For a more detailed explanation, see13.8 Physical security and security of the environment.

12.7.3 *Solution*
Two DMZs have been set up for communication with the outside world:

- DMZ – Haagse Ring (and Diginetwerk Rijksweb and Gemnet) for ministries and municipal authorities
- DMZ – Internet for RHCs and other clients.

12.8 Data transport: network

12.8.1 Description

Information exchanges take place both to and from the National Archives using network connections.

12.8.2 Requirements

Communication between the outside world and the National Archives must be conducted securely. We therefore stipulate requirements for these data connections.

12.8.3 Solution for data connections

We stipulate the following requirements for data connections involving communication between various external networks from outside to the inside (between the outside world and the National Archives):

- HTTPS and FTPS with TLS 1.0, TLS 1.1 or TLS 1.2 (or more recent);
- optionally, Diginetwerk (and the connected linking networks Gemnet and Haagse Ring)
- use of specific SSLCipherSuite configuration, whereby the Cipher Order is enforced;
- the e-Depot's SSL configuration must use 2048-bits Diffie-Hellman parameters.

For a more detailed explanation, seeG.4.2 Description of data connections.

12.8.4 Solution for network equipment

The National Archives use a virtualized infrastructure for servers and storage. All virtual servers and storage operate in a single pool in the National Archives network which contains all physical servers and storage hardware. All e-Depot tenants operate on this hardware alongside other applications.

12.9 Development and management of systems: DTAP

12.9.1 Description

MARA is based on "production environments". It is usually preferable or necessary to have several environments for the further development/production of an IT service. The number of environments has a direct influence on the required system capacity, license, management efforts, and realization/operation costs. There may also be variations, depending on specific wishes, with regard to the hosting of environments.

12.9.2 Requirements

- A DTAP environment must be available for each application, regardless of where it is hosted;
- There are minimally logically separated systems for Development, Testing and/or Acceptance and Production (DTAP);
- DTAP servers are kept in separate VLANs.

13 Security

13.1 What is information security?

The objective of information security is to chart risks and reduce them to an acceptably low level. Risks are inventoried in the following areas:

- Availability:
  Availability concerns the extent to which information and information systems are available at the right times to persons authorized to consult the information or to use information systems.
- Integrity:
  Integrity concerns the correctness and completeness of information in an information system.
- Confidentiality:
  Confidentiality concerns the extent to which access to information is limited to authorized persons.

13.2 Objective of this chapter

To outline the way in which it is ensured that information is secure. This involves organization/processes, systems, applications and infrastructure.
This chapter is also a reference and offers a number of solutions. The requirements are generic and apply to all the systems. The solution for the e-Depot is also described. It is up to the RHCs and the National Archives to install a secure solution for the other systems also.

13.3 General

This chapter is based on the BIR (Information Security Baseline of the Government of the Netherlands)[10]
The following aspects are distinguished:

- Security policy
- Organization of information security
- Management of business resources
- Personal security
- Physical security and security of the environment
- Management of communication and operational processes
- Access security
- Procurement, development and maintenance of information systems
- Incident management
- Business continuity management
- Compliance

There is also an Information Security Baseline for Municipal Councils (BIG) and an Interdepartmental Information Security Baseline (IBI) for the Association of Provincial Authorities (IPO). The BIR, de BIG and IBI are all based on the ISO 27001 standard for information security. They are all structured in the same way.

The National Archives' information security is based on BIR. Below follows an explanation of how each aspect of this is structured.

13.4 Security policy

The National Archives has an established information security policy. This policy is based on the information security policy of the Ministry of Education, Culture and Science (OCW). The National Archives has its own security policy because OCW's

[10] *Baseline Informatiebeveiliging Rijksdienst – Tactisch normenkader*, 1 December 2012

policy only concerns documents in their active phase, such as the Government Information (Public Access) Act (WOB), VIR and VIR-BI. This uses classifications such as "departmentally confidential", "state secret" etc. and does not mention anything about archives, which only distinguish between open access and restricted access (see article 15 paragraph 1 of the Public Records Act). Lighter or heavier security measures can be taken, depending on the restrictions to access.

The RHCs apply their own security policy, based on the policy of the affiliated legal caretakers. In other words, they follow BIG and IBI. The RHCs follow the information security level stipulated in them, namely departmental confidentiality. This must be aligned with the other RHCs and the National Archives' policy on the National Collection. No additional risk classification is necessary (13.6). It is still unclear how to deal with very restricted secret information (appointment of the mayor and research concerning the Public Administration (Probity Screening) Act (Bibob).

13.5 Organization of information security

Objective: establishing roles, responsibilities and authorizations.

Required:
Every system has an owner, who decides on changes to a system etc.

Solution:
At the e-Depot, the following roles can be distinguished:

- CIO: establishes information security
- e-Depot system owner: takes decisions on changes to be made to the system. The National Archives' CIO acts as the e-Depot system owner.
- Information owner: decides on the information in the system. At the e-Depot, an owner is designated for each tenant. The information owner is also responsible for conducting or ordering a risk classification of the information, as set out in paragraph 13.6 Management of business resources.

13.6 Management of business resources

A classification system is needed to classify information. Information within the National Archives must be classified according to this classification system.

Concrete details:
The National Archives use the nation-wide "five point scale" for availability, integrity and confidentiality. This means that for each type of information, the risk level of each of these three aspects is classified as extremely low, low, mid, high or extremely high. This classification is used to determine whether compliance with BIR offers a sufficient level of information security or whether it is necessary to conduct further analysis, on the basis of which it may be necessary to take additional measures.
The classification of information in the e-Depot/CMS/ESB is described in TopX.

13.7 Personal security

Objective: to ensure that employees and hired staff understand their responsibilities and are suitable for their positions in order to reduce the risk of theft, fraud, or abuse.

In concrete terms this means that:

- The archival institution requires a relevant Certificate of Good Conduct from all its personnel. A confidentiality statement will be required from hired personnel.
- Measures to raise awareness are taken regularly.

- When an employment contract comes to an end, access rights are revoked and business resources are taken back.

Concrete details:

- The National Archives require a relevant Certificate of Good Conduct from all its personnel. A confidentiality statement will be required from hired personnel. Contracts with some organizations that are worked with stipulate that the personnel of these companies sign a confidentiality statement
- Government officials must take an oath of office or make a binding promise (General Public Service Regulations paragraph 3).
- The National Archives organize regular meetings on information security. New employees are made familiar with the "ten golden rules" of information security.
- The National Archives currently do not have a confidential function, e.g. (see BIR) a function whereby someone has direct access to the minister or comes into contact with state secrets in a structural manner. The National Archives, in conformity with the Public Records Act, does not contain any state secrets. The minister decides which functions are confidentiality functions. A screening by the General Intelligence and Security Service (AIVD) is only possible for confidentiality functions. It is currently being considered whether it is possible to designate confidentiality functions within the National Archives, so that a screening can be requested for services (namely ingest of NATO archives, which do not fall under the Public Records Act, and for outsourced archives, which may contain state secrets).

13.8 Physical security and security of the environment

Objective: to prevent unauthorized physical access, damage or disruptions to the organization's premises and information.

Required:

- Zoning must be applied in the buildings. Accessing zones is only possible with authorization.
- If services are outsourced to third parties (e.g. in case of external housing or hosting) the outsourcing party will remain responsible.

The current structure within the National Archives (without outsourcing):

- Zone 0:
  - Public space: no authorization is necessary.
- Zone 1:
  - Office space. Access with a *Rijkspas* [smart card for central government] (b= visitor, e= external, p = personnel) or accompanied by someone with a *Rijkspas.*
  - Study room. Access with study room pass bearing a photo. Proof of identity must be shown to obtain this card. Proof of identity is registered in the system (CRM).
- Zone 2:
  - Physical depots. Access with specially authorized *Rijkspas.* The director will decide on authorization after being advised by the head of department.
  - Server space. Access with specially authorized *Rijkspas.* The head of the I&S department or the head of the Business Operations department will decide, after receiving advice from the other head.
- Zone 3:
  - Known as the "safe". Archive records with extremely restricted public access are stored in this safe. Access with specially authorized

  - *Rijkspas.* The director will decide on authorization after being advised by the head of department.
- Zone 4:
  - The safe within the safe. Only two people simultaneously may access the safe and they must both show their specially authorized *Rijkspas.* The director will decide on authorization after being advised by the head of department.

NB In the long term, zones 3 and 4 will probably not be necessary as the access restrictions are being revaluated.

In the event of outsourcing, the National Archives stipulate that the third party must at least comply with ISO 27001, and preferably with BIR. The party responsible for the management and maintenance of the e-Depot (Ordina) complies with ISO 27001.

13.9 Management of communication and operational processes

Objective: to safeguard the correct and secure operation of IT facilities.

Required:
- Procedures are written down (therefore not just in people's minds);
- Changes may not be implemented without the consent of the system owner;
- The "four eyes" principle applies to removing records. This applies to both archive material and the accompanying metadata.

Solution:
- e-Depot: if there are archive records that are meant to be permanently stored in the e-Depot, two people are required to initiate the workflow to remove these records permanently.
- NB This is also a requirement for Statistics Netherlands (CBS).
- The person acting as manager works with two accounts: one for implementing everyday work (ingest, changes to metadata) and one for implementing management tasks such as changing the configuration.

13.10 Access security

Objective: To manage access to information.

13.10.1 Management of access rights for users

Required:
- Users are registered;
- The authorizations of all users are known;
- All users are checked regularly, for example annually , to ascertain whether they still need all of their authorizations;
- There must be a password policy, stating:
  - How complex passwords must be;
  - How passwords are issued to users;

Solution:
- The e-Depot application registers users in the OpenLDAP application in accordance with the LDAP (Lightweight Directory Access Protocol) protocol. LDAP is a network protocol that describes how to approach data from directory services using e.g. TCP/IP. All users who are defined in OpenLDAP can be given access to the e-Depot.
- Users' rights are established in accordance with RBAC, or Role-Based Access Control. A user has one or more roles. Each role has one or more authorizations.

- The login data for the e-Depot's production environment is provided by the service organization through two separate channels: the application's login names and URLs are sent by email, while the password is sent by text message.

13.10.2 Access control for networks

Required:

- Only identified and authorized equipment may be directly connected (i.e. without having to go through firewalls) to a trusted network.
- Implication: Bring Your Own Device (BYOD) may only be connected to a non-trusted zone;
- Logically, the production environment is separated from other environments (such has Development, Testing, and Acceptance environments);
- Sensitive systems that require a very high level of confidentiality are kept in a separate, isolated computer environment that is not connected to any network.

Solution:

- Netword Access Control (NAC) is implemented. NAC is a hardware solution to ensure that only identified and authorized equipment can be connected to a trusted network.
- The National Archives' visitors network is separated from the network on which the e-Depot application runs to ensure that it is not possible to access this application through a BYOD.
- Logically, the production environment and the Training and Demonstration (TED) environment are separated from the other (OTA) environments;
- Implication: as the National Archives do not have systems for extremely high levels of confidentiality, state secrets are not archived there;
- The networks on which the e-Depot operates are protected in conformity with BIR (Information Security Baseline for Central Government). As BIR and BIG (Information Security Baseline for Municipal Councils) share the same basic principles, it is highly likely that the e-Depot also complies with BIG, although this has not been explicitly tested.

13.11 Procurement, development and maintenance of information systems

Objective: To ensure that security is an integral part of information systems.

Required:

- Requirements for security measures should also be included in company standards for new information systems or extensions of existing information systems;
- Updates and patches for vulnerabilities where there is a high risk of abuse resulting in extensive damage will be made as soon as possible, within one week at the latest.

Solutions:

- In the National Archives' software development and management processes it is necessary to include security requirements as acceptance criteria when drawing up proposals for change (a change proposal will not be accepted if it does not meet the security criteria);
- Every Wednesday evening, there is a maintenance window when it is possible to conduct updates and patches.

13.12 Incident management

- Objective: To ensure that information security incidents are reported in time so that corrective measures can be taken and the impact of the incident can be minimized.

Required:

- There is a procedure for reporting and following up security incidents;
- An emergency plan has been drawn up.

Solutions:

- The National Archives have a procedure for reporting and following up security incidents; The Service Bureau is a permanent contact point for clients of the National Archives, also in the event of a disaster;
- The National Archives have an emergency plan; This plan is known within the service bureau and directs the implementation of this emergency plan.

13.13 Business continuity management

Objective: Preventing disruptions to business activities.

Required:

- A business continuity plan has been drawn up. It is tested regularly and brought to people's attention. The business continuity plan focuses on:
- The continuity plan pays minimal attention to:
    - Identification of essential procedures for business continuity;
    - Information to be secured (acceptability of loss of information);
    - Priorities and order of recovery and reconstruction;
    - Documentation of systems and processes;
    - Knowledge and expertise of personnel to start up processes again.

Solutions:

- The National Archives do not have a business continuity plan at this point in time (Q4 2015). Such a plan is planned for 2016.

13.14 Compliance

Objective: Compliance: ensuring that laws, legal, contractual and regulatory obligations and security requirements as described above are complied with.

Required:

- It must be known which laws and rules are applicable;
- A policy on the use of facilities;
- It is necessary to report on information security during the Planning & Control cycle;
- Information systems are regularly checked for compliance with the implementation of security standards.

Solutions:

- The National Archives' information security policy contains the applicable legislation and regulation.
- There is a policy on the use of facilities by employees of the National Archives;
- A report is drawn up annually on the basis of a self-evaluation for the Ministry of Education, Culture and Science;
- The National Archives is currently (Q4 2015) implementing active monitoring of information systems that incorporates vulnerability scanning.

14 References

Brief description in this document	**Full reference**
DTR WP1 programme architecture	DTR WP1, *Programma Architectuur DTR WP1*, October 2015
Open Access report.	K. van der Heiden and I. Zandhuis, *Open toegang tot archiefcollectie Nederland. Verkenningen en aanbevelingen Archief2020,* April 2014; online: http://archief2020.nl/file/110/download?token=p5t0PvOEXay-_1vYCbcg5yCtWOanFcpOOKlNVzgFlqU
DTR Architecture	National Archives, *Enterprise Architectuur "Digitale Taken Rijk (DTR)"*, May 2014
National Archives PDC for RHCs	National Archives, *Producten en Diensten Catalogus Archiefdiensten voor Regionaal Historische Centra*, March 2015.
National Archives PDC for departments	National Archives, *Producten en Diensten Catalogus Archiefdiensten voor departementen*, Version 1.0 July 2015
RHCs PDC for legal caretakers	RHCs working group, *Producten en Diensten Catalogus Archiefdiensten voor zorgdragers*, Version 0.6 August 2015
National Strategy	Digital Heritage Network, *Nationale Strategie Digitaal Erfgoed. Een initiatief van het Netwerk Digitaal Erfgoed*, March 2015; online at https://www.rijksoverheid.nl/documenten/publicaties/2015/03/09/nationale-strategie-digitaal-erfgoed
WVI processes	Working group on Preparation of e-Depot Implementation for RHCs, Work processes section:*Verwerven, beheren en beschikbaar stellen. Handboek werkprocessen voor RHC's*, Version 1.0 April 2013
WVI Architecture	Working group on Preparation of e-Depot Implementation for RHCs, Architecture section, *Enterprise Architectuur RHC's: "SOLL" 2015,* Version 1.0 April 2013

15 Terms and abbreviations

used in this document.

Term / abbreviation	**Further details**
API	Abbreviation of Application Programming Interface. An API is a collection of definitions through which computer programmes can communicate with each other. A programme can call on the functionality of another programme through the "outside" of that program, without any knowledge being required about the "inside" of the programme being called on.
Archive	An archive is a collection of documents, collected by an organization, family or person (the archive creator). Keeping these records may be compulsory by virtue of the Public Records Act, or desirable for other reasons. An archive may consists of agendas and minutes, files, letters, deeds, maps, drawings, photos, etc.
Archive component	An archive component is a unit of archive records that belong together within the archive as a whole. For example, the series of personnel files, a databank or the series of management board minutes.
Archive records	Documents (records), regardless of their form, that have been received or drawn up by governmental bodies and which are intended to be kept by these bodies due to their nature (i.e. being related to the organization's tasks).
Archive service	The archive service is the manager of the archives. On the grounds of the Public Records Act 1995, archive services that manage government archives are obliged to manage archives correctly and make them accessible to the public.
Archival institutions	In this document: RHCs that are affiliated to the Convent of RHCs and the National Archives.
Archive access	An archive access contains an overview of descriptions of archive items that are part of that archive (this overview is called an inventory).
Digital archive components	Archive components that can only be consulted with the help of operating or application software, such as emails and information in databases.
DTR	Abbreviation of: Digital Tasks of Government (*Digitale Taken Rijk*). The DTR programme is aimed at setting up a digital infrastructure for the national government's archives, and where possible for other governments as well.
DWR archive	DWR is an abbreviation of: Digital Government Workplace (*Digitale Werkplek Rijk*). Part of this project consists of creating a facility for departmental and other archive creators to outsource material to the

	National Archives, i.e. deposit it with the National Archives before transfer or destruction under public records legislation.
EAD	Encoded Archival Description. An international XML standard for marking archive inventories.
Ingest	The process by which a delivered set of digital archive documents (including corresponding substantial and technical metadata) is checked and after approval is included and registered in the e-Depot.
Intellectual management	Management of manifestation-independent metadata.
Institutional archive creator	An archive creator that is an institution or institute, such as a political party or trade union. In contrast to natural persons, these archive creators often have legal predecessors and successors.
Inventory	The overview of descriptions of archive records is hierarchically structured. Categories add structure to the information and are part of the description.
Linking	Referring from one's own system to corresponding data in an external system. An example of this is linking an archive creator's data to the biographical portal in a collection management system. If data in the core registration change, the data in the collection management system change with it (because they are not stored in the collection management system).
Metadata	Information about information. Metadata are the characteristics of physical and electronic information. Metadata are crucial to reliable records management. Archival metadata relate to archival documents (context, content and structure as well as their management over time). They record status, format and location and also document the activities of the archive system.
OAI-PMH	Abbreviation of Open Archives Initiative Protocol for Metadata Harvesting. OAI-PMH is a protocol intended to collect the metadata of entities in archives. OAI-PMH is a low-threshold mechanism to bring about the exchange between repositories of archives.
ODC	Abbreviation of Government Data Centre (*Overheids Data Centrum*). There are four ODCs: • The 'The Hague Square Kilometre' in Rijswijk, which is intended to serve SSC-ICT Haaglanden, the Dutch central government's IT service centre; • ODC North in Groningen; • Two hired twinning locations at Equinix in Amsterdam; • The Tax Authorities' Quintax complex in Apeldoorn.
Derivation	An easy, one-off way of copying data from an external system to one's own system. An example of this is deriving a title description from a core registration in a collection management system. The title description is then (also) saved in the collection management system. If the data in the

	core registration change, the data in one's own system will not change along with it.
Ontology	A strict and exhaustive scheme for a specific subject domain, usually with a hierarchical structure, that contains all relevant quantities and their relations as well as the rules to which the quantities and relations comply within that domain[11].
Transferred archives	Archives that have been placed in the National Archives' e-Depot by a legal caretaker and of which legal caretakership is taken over by the National Archives or an RHC. There is a twenty-year period for transferring an archive that falls under the Public Records Act. It is also possible to transfer archives earlier or to suspend transfer.
RDA	Abbreviation of Resource Description and Access. RDA is a standard for title descriptions that is used in the library sector (follow-up to FOBID).[12]
Record	In this document, this is synonymous with archive item.
RHC	Regional Historical Centres (RHCs) are a "public body" on the grounds of the Joint Regulations Act (WGR). The central government, one or more municipal councils, a province or another institution are represented in this type of public body. RHCs keep national archives that are intrinsically connected to the implementation of central government tasks in the province. They may also store local and regional archives and collections. In this document, RHCs refer to the RHCs that are part of the Convent of RHCs, previously known as the National Archives in the provinces.
National archives	Archives that are set up by organizations within central government, such as ministries, high councils of state, executive agencies and a number of inspection services.
Taxonomy	
Thesaurus	
TOCO	Abbreviation of Tasks and Organization of Central Government (*Taken en Organisatie Centrale Overheid*). TOCO is the current name of an application that could be used as a core registration for institutional creators.
Outsourced archive	Outsourced archives are archives that have not yet been transferred under public records legislation: legal caretakership still lies with the original creator. Storage and preservation are the responsibility of the archival institution that the archive has been outsourced to.
UUID	Abbreviation of Universally Unique Identifier. A UUID is a way to ensure that a piece of information always has a fixed "address" and can therefore always be found. UUID is an ID used in software architecture, standardized by the Open Software Foundation (OSF) as part of the Distributed Computing

[11] Source: http://nl.wikipedia.org/wiki/Ontologie_(informatica)
[12] Source: http://www.fobid.nl/ontsluiting-rda

	Environment (DCE).
Widget	A widget is a simple, graphical object or element with an uncomplicated, specific and often-used function that can be activated or deactivated by a user to supplement or set up an existing user interface. Source: http://nl.wikipedia.org/wiki/Widget.
WVI Architecture	Architecture drawn up in 2012 by a working group of various RHCs and the National Archives. It is available at http://www.noord-hollandsarchief.nl/content/downloads/files/WVI%20Enterprise%20Architectuur.pdf.
Legal caretaker	Organization tasked with the care for an archive. Legal caretakers are responsible for setting up and managing archive records to support their public duties. They must make funds and resources available for this purpose.

Theme A Presentation of a digital archive item – integrity and authenticity

A.1 Introduction

End users of archival material must be able to rely on receiving a complete, unaltered and authentic archive record from the archival institution. This theme describes what is required to accomplish this.

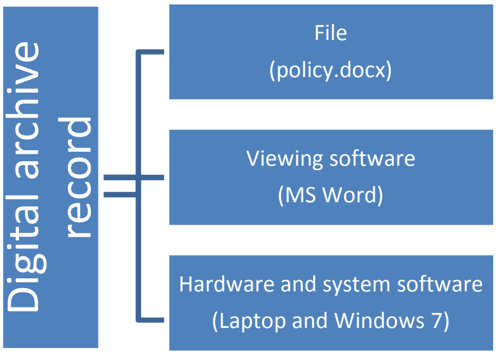
A.2 What is a digital archive record?

Imagine, a civil servant writes a policy document, using word processing software, such as Microsoft Word. When they are satisfied they save the text and send it to their superior, who opens the document.

The implicit expectation is that the receiver sees the document in exactly the same way as the writer. The record is what is shown on the screen. To reproduce a digital document the following elements are necessary:

- Data, a file for example;
- Application software for viewing the data;
- An environment in which the application software operates.

An illustration of this:



To keep an archive record you would have to keep the PC with the viewing software and file open. Naturally, this is not what happens: usually, the file (policy.docx) is the only thing that is saved. If the digital document is required, one finds a computer and the file is opened again in the viewing software. In other words:

> You cannot keep a digital record, you can only recreate it.

Furthermore, the situation outlined above is a simple one, involving only one file. A more complication situation is conceivable if the record shown on the screen does not correspond to just one file, but is, for example, the result of a query about several databases and a text file.

A.3 What is necessary for a complete and unaltered reproduction of a digital

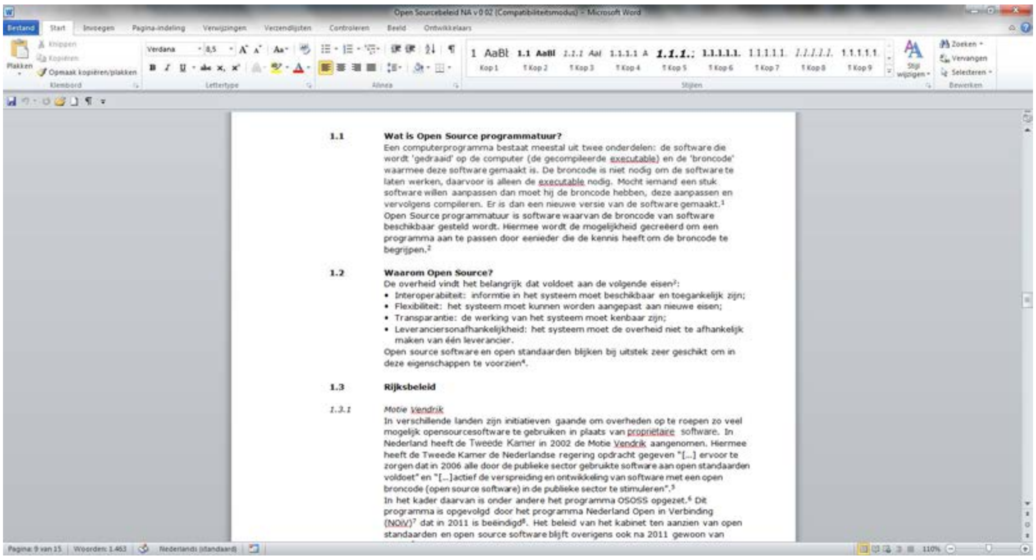
record?

A.3.1 *Essential characteristics must be preserved*

In the example above, we assume that the file and viewing software deliver a complete and unaltered digital archive record. This means that the *essential characteristics* (such as content, lay-out, paging, images, possibly font etc.) of the original version of the record (in the environment of the civil servant who typed the document) are identical to those in the recreated version (in the environment of the superior who opened the file). If both persons use the same version of MS Word, for example, there is a good chance that this will be the case.[13]

A.3.2 *The influence of viewing software*

See the examples below, in which a file that was originally made in MS Word is opened in two software applications: first in MS Word, and on the right in Notepad, an auxiliary programme.



1.1 **Wat is Open Source programmatuur?**

Een computerprogramma bestaat meestal uit twee onderdelen: de software die wordt 'gedraaid' op de computer (de gecompileerde executable) en de 'broncode' waarmee deze software gemaakt is. De broncode is niet nodig om de software te laten werken, daarvoor is alleen de executable nodig. Mocht iemand een stuk software willen aanpassen dan moet hij de broncode hebben, deze aanpassen en vervolgens compileren. Er is dan een nieuwe versie van de software gemaakt.[1]
Open Source programmatuur is software waarvan de broncode van software beschikbaar gesteld wordt. Hiermee wordt de mogelijkheid gecreëerd om een programma aan te passen door eenieder die de kennis heeft om de broncode te begrijpen.[2]

1.2 **Waarom Open Source?**

De overheid vindt het belangrijk dat voldoet aan de volgende eisen[3]:

- Interoperabiteit: informtie in het systeem moet beschikbaar en toegankelijk zijn;
- Flexibiliteit: het systeem moet kunnen worden aangepast aan nieuwe eisen;
- Transparantie: de werking van het systeem moet kenbaar zijn;
- Leveranciersonafhankelijkheid: het systeem moet de overheid niet te afhankelijk maken van één leverancier.

Open source software en open standaarden blijken bij uitstek zeer geschikt om in deze eigenschappen te voorzien[4].

1.3 **Rijksbeleid**

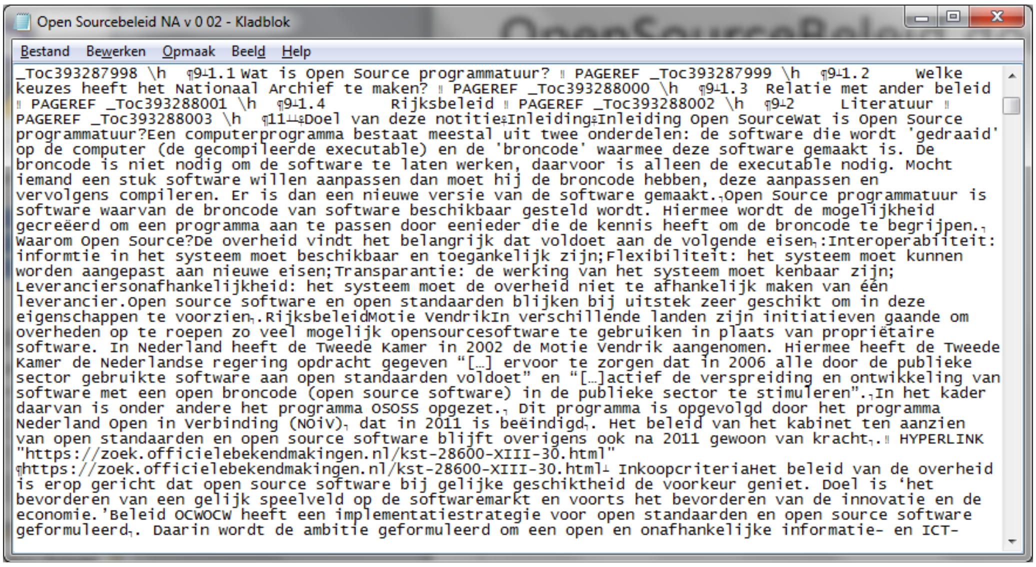
1.3.1 *Motie Vendrik*

In verschillende landen zijn initiatieven gaande om overheden op te roepen zo veel mogelijk opensourcesoftware te gebruiken in plaats van proprietaire software. In Nederland heeft de Tweede Kamer in 2002 de Motie Vendrik aangenomen. Hiermee heeft de Tweede Kamer de Nederlandse regering opdracht gegeven "[...] ervoor te zorgen dat in 2006 alle door de publieke sector gebruikte software aan open standaarden voldoet" en "[...]actief de verspreiding en ontwikkeling van software met een open broncode (open source software) in de publieke sector te stimuleren".[5]
In het kader daarvan is onder andere het programma OSOSS opgezet.[6] Dit programma is opgevolgd door het programma Nederland Open in Verbinding (NOiV)[7] dat in 2011 is beëindigd[8]. Het beleid van het kabinet ten aanzien van open standaarden en open source software blijft overigens ook na 2011 gewoon van

Figure 2 File opened in MS Word

Now we will open the same file in the auxiliary programme Notepad:

[13] But this cannot be assumed: in the word processing package WordPerfect, for example, the tab setting was not embedded in the text file, but in WordPerfect's settings. It was therefore possible that someone made a table in a text, but when someone else opened the text in WordPerfect, and had adjusted their tab settings, the numbers in the table shifted over the columns. This is not an authentic reproduction of a record!

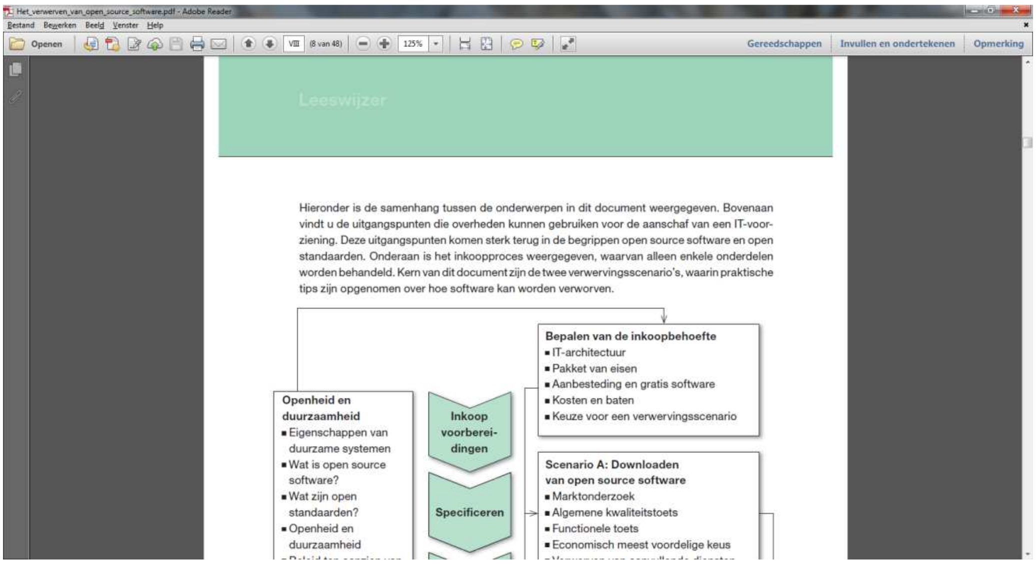


```
_Toc393287998 \h  ¶9↓1.1 Wat is Open Source programmatuur? ‼ PAGEREF _Toc393287999 \h  ¶9↓1.2     Welke
keuzes heeft het Nationaal Archief te maken? ‼ PAGEREF _Toc393288000 \h  ¶9↓1.3  Relatie met ander beleid
‼ PAGEREF _Toc393288001 \h  ¶9↓1.4        Rijksbeleid ‼ PAGEREF _Toc393288002 \h  ¶9↓2     Literatuur ‼
PAGEREF _Toc393288003 \h  ¶11↓↓‡Doel van deze notitie‡Inleiding‡Inleiding Open SourceWat is Open Source
programmatuur?Een computerprogramma bestaat meestal uit twee onderdelen: de software die wordt 'gedraaid'
op de computer (de gecompileerde executable) en de 'broncode' waarmee deze software gemaakt is. De
broncode is niet nodig om de software te laten werken, daarvoor is alleen de executable nodig. Mocht
iemand een stuk software willen aanpassen dan moet hij de broncode hebben, deze aanpassen en
vervolgens compileren. Er is dan een nieuwe versie van de software gemaakt.┐Open Source programmatuur is
software waarvan de broncode van software beschikbaar gesteld wordt. Hiermee wordt de mogelijkheid
gecreëerd om een programma aan te passen door eenieder die de kennis heeft om de broncode te begrijpen.┐
Waarom Open Source?De overheid vindt het belangrijk dat voldoet aan de volgende eisen┐:Interoperabiliteit:
informtie in het systeem moet beschikbaar en toegankelijk zijn;Flexibiliteit: het systeem moet kunnen
worden aangepast aan nieuwe eisen;Transparantie: de werking van het systeem moet kenbaar zijn;
Leveranciersonafhankelijkheid: het systeem moet de overheid niet te afhankelijk maken van één
leverancier.Open source software en open standaarden blijken bij uitstek zeer geschikt om in deze
eigenschappen te voorzien┐.RijksbeleidMotie VendrikIn verschillende landen zijn initiatieven gaande om
overheden op te roepen zo veel mogelijk opensourcesoftware te gebruiken in plaats van propriëtaire
software. In Nederland heeft de Tweede Kamer in 2002 de Motie Vendrik aangenomen. Hiermee heeft de Tweede
Kamer de Nederlandse regering opdracht gegeven "[…] ervoor te zorgen dat in 2006 alle door de publieke
sector gebruikte software aan open standaarden voldoet" en "[…]actief de verspreiding en ontwikkeling van
software met een open broncode (open source software) in de publieke sector te stimuleren".┐In het kader
daarvan is onder andere het programma OSOSS opgezet.┐ Dit programma is opgevolgd door het programma
Nederland Open in Verbinding (NOiV)┐ dat in 2011 is beëindigd┐. Het beleid van het kabinet ten aanzien
van open standaarden en open source software blijft overigens ook na 2011 gewoon van kracht┐.‼ HYPERLINK
"https://zoek.officielebekendmakingen.nl/kst-28600-XIII-30.html"
¶https://zoek.officielebekendmakingen.nl/kst-28600-XIII-30.html↓ InkoopcriteriaHet beleid van de overheid
is erop gericht dat open source software bij gelijke geschiktheid de voorkeur geniet. Doel is ‘het
bevorderen van een gelijk speelveld op de softwaremarkt en voorts het bevorderen van de innovatie en de
economie.’Beleid OCWOCW heeft een implementatiestrategie voor open standaarden en open source software
geformuleerd┐. Daarin wordt de ambitie geformuleerd om een open en onafhankelijke informatie- en ICT-
```

Figure 3 The same file, opened in Notepad

Although most of the text has been preserved when opened in Notepad, few people will regard this reproduction as a complete and unaltered version of the original archive record.

Saving it in an open file format such as PDF does not help either, as the example below demonstrates:



Hieronder is de samenhang tussen de onderwerpen in dit document weergegeven. Bovenaan vindt u de uitgangspunten die overheden kunnen gebruiken voor de aanschaf van een IT-voorziening. Deze uitgangspunten komen sterk terug in de begrippen open source software en open standaarden. Onderaan is het inkoopproces weergegeven, waarvan alleen enkele onderdelen worden behandeld. Kern van dit document zijn de twee verwervingsscenario's, waarin praktische tips zijn opgenomen over hoe software kan worden verworven.

Openheid en duurzaamheid
▪ Eigenschappen van duurzame systemen
▪ Wat is open source software?
▪ Wat zijn open standaarden?
▪ Openheid en duurzaamheid
Inkoop voorbereidingen
Specificeren
Bepalen van de inkoopbehoefte
▪ IT-architectuur
▪ Pakket van eisen
▪ Aanbesteding en gratis software
▪ Kosten en baten
▪ Keuze voor een verwervingsscenario
Scenario A: Downloaden van open source software
▪ Marktonderzoek
▪ Algemene kwaliteitstoets
▪ Functionele toets
▪ Economisch meest voordelige keus

Figure 4 PDF file, opened in PDF reader

Now we will open the same PDF file in the auxiliary programme Notepad:

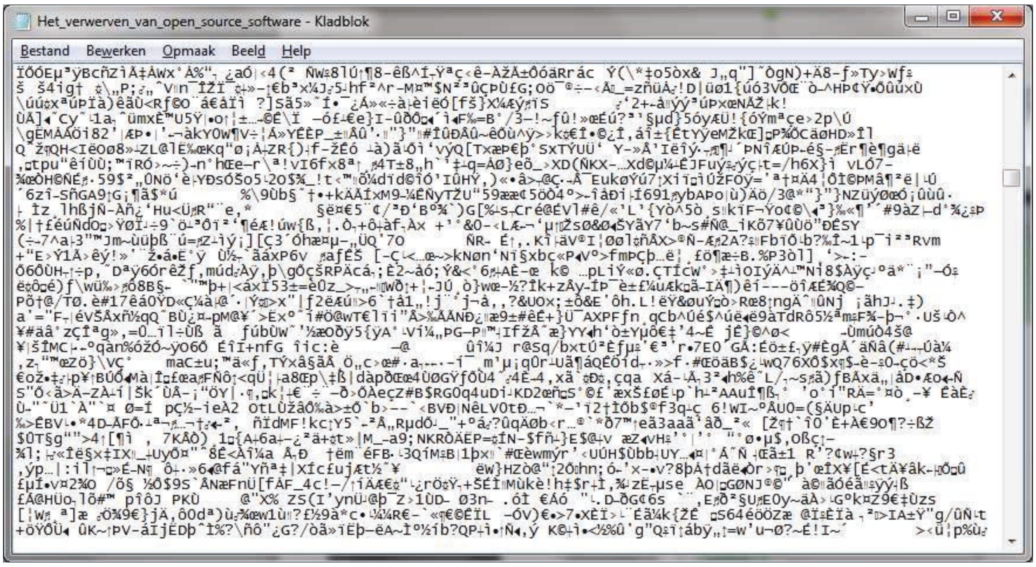


Figure 5 PDF file, opened in Notepad

Now nothing remains of the text: the recreation of the archive record has not produced a complete and unaltered version.

This shows that the playback software used influences the integrity of the archive record.

A.3.3 *So how do we produce a version with integrity?*

Integrity is a base value for archival institutions. The mission of archival institutions is to provide authentic, unaltered and complete versions of the analogue and digital archive records in their care to their users. In view of the above, there are roughly two options:

- Option 1: The archival institution supplies the file and refers the user to easily available viewing software. For example: *"You can download the PDF file here. You can open this file with Acrobat Reader, which can be downloaded here."* The archival institution then knows (due to research) that this combination will produce a faithful and complete reproduction.
- Option 2: The archival institution will take care of rendering itself and will refer users to a place where this archive record can be viewed. For example: "*You can read this digital document here*." The file will then be made available online and in a browser in rendering software.

Option 1 is possible for extensively used and well-known formats such as PDF and many formats for images, such as JPEG, for which it is easy to obtain reliable and free viewers.

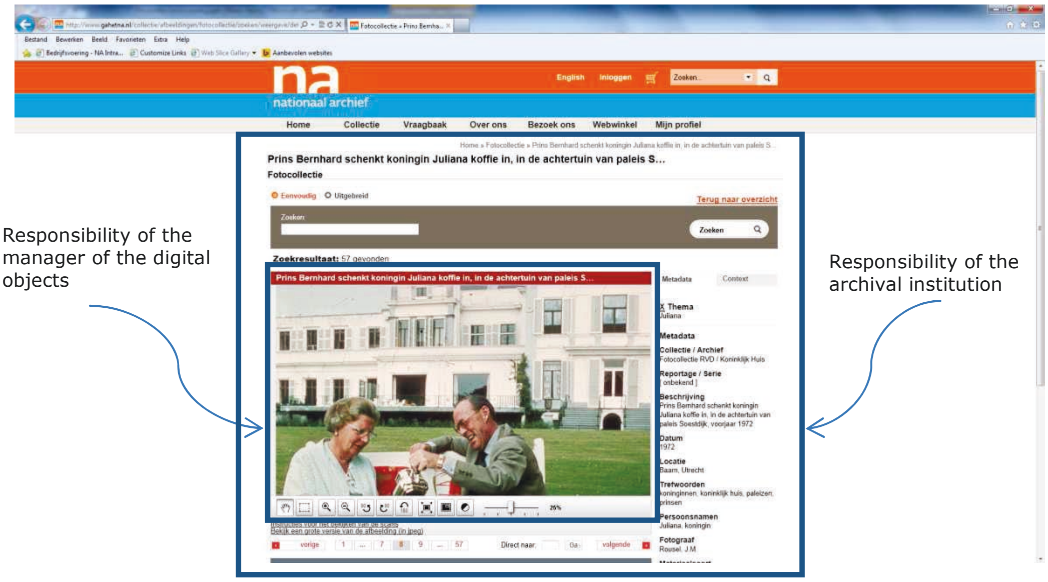
Option 2 provides more certainty regarding the integrity of the reproduction (because less is left to the user). This is actually the only option for more difficult file formats, for which there are no easily available free viewers, or for records that do not correspond to one file but are generated "on the fly" by combining information from various sources, such as a query about several databases.

A.3.4 *What does this look like?*

You can compare it to embedding a video on YouTube: YouTube provides a viewer that shows the video file. All of this (the video, shown in a viewer) can be shown on your own website. From the perspective of the user, the film is on the website, but YouTube is providing the reproduction in the background.

If AV material from the National Archives or an RHC is at Beeld & Geluid, then Beeld & Geluid will provide a reproduction that can be shown on a National Archives or RHC website, along with its metadata.

Example:



Home » Fotocollectie » Prins Bernhard schenkt koningin Juliana koffie in, in de achtertuin van paleis S...

Prins Bernhard schenkt koningin Juliana koffie in, in de achtertuin van paleis S...

Fotocollectie

Zoekresultaat: 57 gevonden

Prins Bernhard schenkt koningin Juliana koffie in, in de achtertuin van paleis S...

X Thema
Juliana

Metadata

Collectie / Archief
Fotocollectie RVD / Koninklijk Huis

Reportage / Serie
[onbekend]

Beschrijving
Prins Bernhard schenkt koningin Juliana koffie in, in de achtertuin van paleis Soestdijk, voorjaar 1972

Datum
1972

Locatie
Baarn, Utrecht

Trefwoorden
koninginnen, koninklijk huis, paleizen, prinsen

Persoonsnamen
Juliana, koningin

Fotograaf
Rousel, J.M

Instructies voor het bekijken van de scans
Bekijk een grote versie van de afbeelding (in jpeg)

Responsibility of the manager of the digital objects

Responsibility of the archival institution

Figure 6 Rendering of archive material

In the example given above, the party who actually physically stores the digital objects (the data is stored in their database) supplies the viewer on which the data can be shown. The archival institution can then make a building block in which the images are shown, equipped with the necessary metadata.

The archival institution, or a third party, such as a thematic website, can place the rendered image on its own website.

A.3.5 *Reuse?*

NB the above text concerns a situation in which the integrity of the displayed material is the most important factor. The method described can be used to show a digital document in a different context, such as a thematic website, without impairing integrity: there is certainty that a good rendering has been used and the metadata originate from the archival institution.

This does not affect the fact that a re-user may also download only the *data* of the archive record (the bitstream, e.g. the JPEG or PDF file), find a viewer and show the item on the organization’s own website. Perhaps the idea is simply to show a nice picture. In such a case, the archival institution cannot guarantee the integrity of the resulting archive item.

A.4 Resulting principles

- The data holder is responsible for the integrity of the reproduced archive item.

Possible interpretation:

- The data holder provides a rendering option;
- The data holder refers to rendering software

Theme B Reuse of archive material

B.1 What is reusable?

Material can be reusable in two ways.

- The emphasis is on integrity and authenticity
  The first way is where "everyone" "knows" where an archive record is. The "signposts" can be reused: all kinds of parties may draw up indexes, accesses etc. on the basis of information that the archives have made available. The objective is always to produce a digital archive record that is authentic, complete and unaltered, such as a specific deed. The authentic, complete and unaltered deed can be viewed on the site of the archival institution whose collection it is part of. The end user can trust that the version at the archival institution is the authentic one. This does not mean that the deed cannot be viewed elsewhere, but that the archival institution provides an authentic, complete and unaltered version. Examples of such reuse are the various "aggregators" whose objective is to form a portal to as much cultural heritage as possible (e.g. APE). This sort of reuse is a kind of "new road leading to the same Rome".
- Reuse of data
  In the second case, the re-user is often not interested in making a signpost to the same archive item in a new way; instead they want to do something different with the actual material, such as presenting it on their own website (in which case the archival institution will not guarantee authenticity and integrity) or using a scan of a map to print on products. This type of reuse is less often about the archive item bearing witness as "a product of government actions", but more about its intrinsic worth e.g. its being an attractive picture.

B.2 For authentic, complete and unaltered archive records, do not copy data but refer to it

Authenticity and integrity are essential for archival institutions. Archival institutions ensure that their material is complete and unaltered and presented together with its metadata. Other institutions may combine "signposts" to the information, addresses, and harvests, perhaps combined with other information, and make all this searchable through their own channels. This makes it possible, for example, to build thematic websites.

The means that locations where complete and unaltered archive items can be viewed must have a persistent and unique address.

B.3 Making data available for reuse

There are also situations in which a re-user is not primarily interested in authentic and integral material, but wants to do other things with it, or only wants to reuse the metadata. Therefore we also provide data of records (archive material and metadata) as *open data*.

B.4 Resulting principles

- Every data holder has an interface to export data according to open standards.
- Data and metadata and locations where faithful and intact archive items can be viewed have a persistent and unique address.

Theme C Manifestations and metadata

C.1 The concept of "manifestations"

One archive record may have several different appearances. These are known as "manifestations". An archive record may originally be a paper document (a physical manifestation), but there may also be a scan of it (a digital manifestation) that may also be transcribed (another digital manifestation). Some metadata on an archive record concern all manifestations belonging to that item, such as the archive creator, the period to which the item pertains, and the author. We refer to such metadata as "manifestation-independent metadata". Other metadata only concerns one manifestation. Metadata may vary, depending on whether it concerns a physical or a digital manifestation. For example, a physical manifestation has physical measurements and is made from a specific material. A digital manifestation does not have this kind of information: unlike physical manifestations, the file format is important here. Lending information is only found in physical manifestations: scans of born-digital material may be viewed but now borrowed.

C.2 Principle: one-off storage of manifestation-independent metadata

An important principle of this new architecture is that manifestation-independent material is only saved once and not each time anew for every manifestation. If there are several manifestations, only information that is relevant for the manifestation in question will be stored. This prevents double or unnecessary storage of the same data. This principle is a direct interpretation of the NORA AP13 principle: source registrations are leading.

Three types of metadata can be distinguished:

- Manifestation-independent metadata
  Examples:
  Information on legal caretakers/archive creators, openness, substantive description.
- Manifestation-dependent metadata for physical material
  Examples:
  Information on measurements, physical condition, physical location (depot, case, shelf), lending information, information on who has seen which item.
- Manifestation-dependent metadata for digital material
  Examples:
  Information on file format, checksums, information on who has seen which item.

It is important to make a distinction in the architecture between the functionalities that manage this metadata. This does not exclude the possibility that several functionalities are performed by the same application.

C.3 Diagram of manifestation-dependent and manifestation-independent metadata

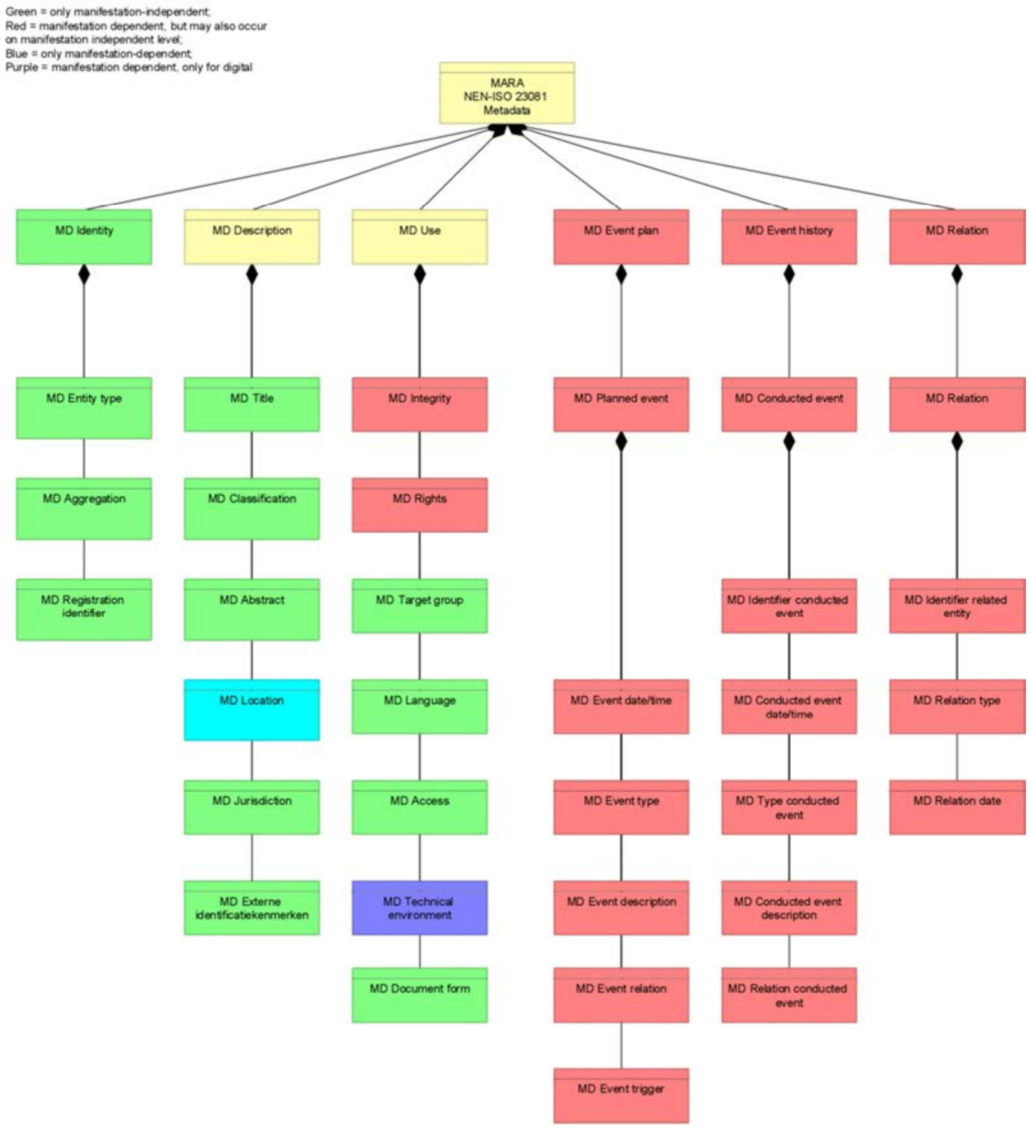
The organization of data architecture is based on the standard for metadata information NEN-ISO 23081. The following diagram shows the structure of NEN-ISO 23081 for metadata. This standard describes which metadata must be saved and stipulates that a metadata scheme must comply with this structure. It does not, however, itself provide a metadata scheme. This standard distinguishes the following main groups of metadata:

- Identity
- Description

- Use
- Event plan
- Event history
- Relationships

Theses main categories are then divided into a number of sub-categories. The diagram below shows this structure. In the detailed explanation the Dutch and English explanations of NEN-ISO 23081 have been adopted.

C.4 Structure of ISO 23081: indication of manifestation dependence and independence



C.5 Overview of metadata in the e-Depot application and collection management system

	Transferred archives		Outsourced archives	
	Physical	**Digital**	**Physical**	**Digital**

Collection Management System	**MOM**			
	- Description: ISAD(G), … - Reference to archive creator (Actor register) - Public access - Copyright - Reference to manifestations 1…n [UUID] - Reference to presentations [UUID]			
	MAM		**MAM**	
	- Physical storage site - Size - Material - Availability - Consultability - Material condition - Restoration status		- Physical storage site - Size	
e-Depot application		**MOM (copy)**		**MOM**
		- Public access (Security tag)		*ToPX*: - Description - Official - Date - Etc.
		MAM		MAM
		- Creation data - Application - File format - Resolution - Etc.		- Public access - Creation data - Application - File format - Resolution - Confidentiality - Retention period - Encryption and access rights - Agency-specific - Etc.

C.6 Resulting principles

- Manifestation-independent metadata is only saved once

Theme D Use of core registrations.

D.1 Separate registration of actors: actor registers

Under the current situation, all information about the archive creator is usually stored as part of the archive's metadata. The NEN-ISO 23081 standard for metadata refers to this as the one-entity model: all information is saved in the "entity", a record or archive item.

If there are several archives from the same archive creator, this leads to a lot of information on the archive creator being saved more than once. The archival institution often needs information about the archive creator, and not only in relation to archival items.

For that reason, the architecture described here proposes setting up a separate register with all information about archive creators and only including a reference to the information on the archive creator in the archive item's metadata. This makes information on archive creators generally applicable, prevents duplications and double storage, and reduces the risk of error. This is also better aligned to NEN-ISO 23081, which is based on five entities (see diagram below) but this is still some way off.

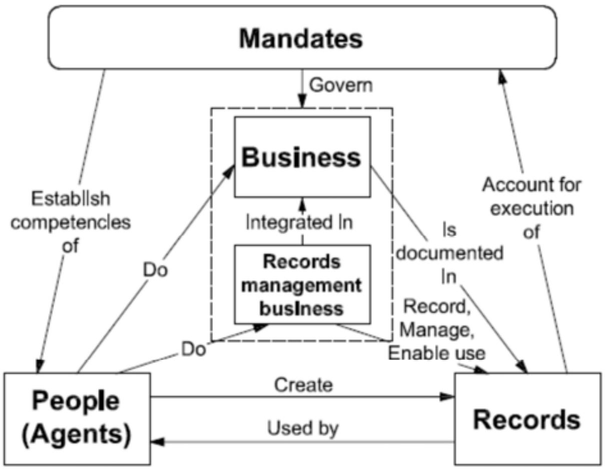


Figure 1: NEN-ISO 23081 entity model

It is possible that several actor registers will complement each other, for example one for institutional archive creators and one for private archive creators. Ideally, these core registrations could be shared by several archival institutions and perhaps by the actual archive creator. This would prevent duplications and would also make it possible to use the core registration of archive creators find out in which institutions and archives the archive creator is present.

D.2 Other authority files / core registrations

The objective is to make maximum use of references to well-defined terms elsewhere and to make as little use as possible of free text fields when describing archive or other material.

Example:

It is preferable to refer to a place where “The Hague” is defined (such as DBPedia) instead of entering “The Hague” as a string. This makes it easier to relate material to each other and connect them semantically.

Example:
Using a thesaurus can ensure that a user searching for ‘mill’ will also find a hit if the access/index contains the word ‘open trestle’ because an open trestle post mill is a type of windmill. In that case, the last term has been recorded in a thesaurus.

D.3 Resulting principles

- Maximum reference to core registrations/authority files
- One or more actor registers
- When using core registrations:
    - Stable organization
    - Making agreements on the management of shared data

Theme E Connection layer: aggregators versus Linked Data

E.1 Why make a connection layer and how is this done?

The objective is to connect different collections to each other. On one hand, this concerns collections from one archival institution (such as archive items, newspapers, photos, books), on the other hand this concerns collections that transcend institutions, such as archive collections from several institutions.

End users must also be able to add information.

A connection layer like this requires:

- A solution for connecting collection elements to each other semantically, e.g. it must be possible to establish a relationship to an archive item from a book.
- A solution for connecting data from several different sources to each other technically.

E.2 Aggregators

We are provisionally using aggregators for technical connections. This is “proven technology”. One example is the Archives Portal Europe (APE). Aggregators harvest metadata from different sources, bring them together, enrich them if necessary and make the collected data fully searchable.

The disadvantage of aggregators is that semantic information must be stored in a fixed structure containing a fixed number of fields to be filled in. If you know anything additional about a collection item, you cannot enter it in the structure because there is no box for it, and if you lack certain knowledge you have to leave an empty space. It is also necessary to make agreements on the formats for data exchange (e.g. an XML scheme) and technology (e.g. OAI-PMH) when using aggregators for data collection or it will not be possible to harvest.

E.3 Linked Data

It may be possible to use Linked Data to solve the disadvantage of the fixed structure of descriptive fields. If using Linked Data there is no more need for a fixed format. There is more flexibility as mini-items can be recorded in so-called triples.

Linked Data also has potential disadvantages: it is more difficult to check who has made certain links or is the author of little nuggets of knowledge because they are so small. A traditional archive description in EAD format has just one author of all fields in the description instead of an author for every field in the description.

Although using Linked Data to aggregate collections certainly has potential, it still requires a great deal of research and thought.

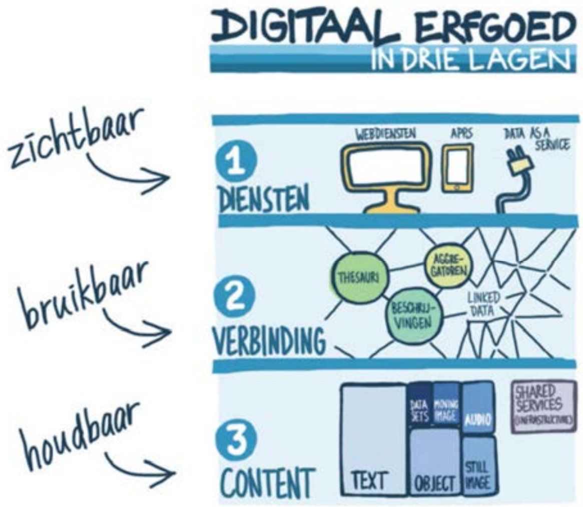
E.4 Resulting principles

- It is always known who has added metadata, such as a relation.

Theme F Implementation of the three-layer model

F.1 National Strategy for Digital Heritage and the three-layer model

A three-layer model has been developed in the National Strategy for Digital Heritage. The objective is to enable cultural heritage from different domains (libraries, science, archives…) to be connected to each other so that it is easier for end users to find relevant information.



The objective is to move “from silos to layers”.

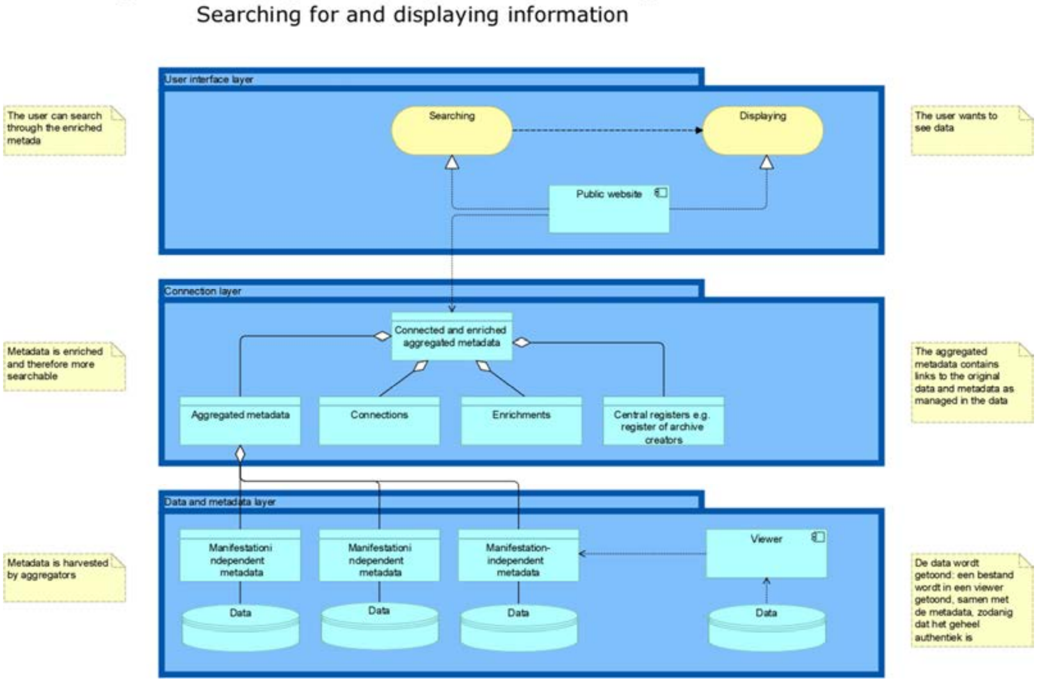
F.2 Implementation of National Strategy in MARA

The three-layer model is a conceptual model, a way of thinking The most important point is that there is a middle layer between the top layer and bottom layer, where collections from both inside and outside the collection can be connected to each other.
Two types of connection are required:

- A conceptual connection
- A technical connection

The basic premise for the architecture is that information for end users can be found via an aggregator. The search tool is not directly linked to the applications/systems in which the information is contained. NB There is an exception to this: information that can only be viewed by a (very) restricted group of users, e.g. only on-site.

F.3 Searching and finding according to the three-layer model


Searching for and displaying information
User interface layer
The user can search through the enriched metada
Searching
Displaying
The user wants to see data
Public website
Connection layer
Metadata is enriched and therefore more searchable
Connected and enriched aggregated metadata
The aggregated metadata contains links to the original data and metadata as managed in the data
Aggregated metadata
Connections
Enrichments
Central registers e.g. register of archive creators
Data and metadata layer
Metadata is harvested by aggregators
Manifestationi ndependent metadata
Manifestationi ndependent metadata
Manifestation-independent metadata
Viewer
De data wordt getoond: een bestand wordt in een viewer getoond, samen met de metadata, zodanig dat het geheel authentiek is
Data
Data
Data
Data

F.3.1 *Further details:*

- The lowest level, 'data and metadata at institutions', consists of the data (e.g. digital objects) and metadata available at institutions and archival institutions. This data layer ensures authenticity and integrity, partly through rendering: the "right" file in the "wrong" rendering engine will not produce a complete and unaltered archive record.
- The middle – connecting – layer ensures that this data and metadata are compiled and can be requested, exchanged and used. Establishing automatic and semi-automatic links between objects and collections creates opportunities to put together presentations and develop joint products and services.
  Connection layer:
  - This contains aggregations of data that transcends owners (e.g. archival institutions, archive creators).
  - Only metadata are aggregated, no data. The metadata do, however, contain references to data.
  - Data can also be enriched by persons other than the data manager. Examples are the use of thesauri, further access, information on archive creators etc.
- The top (presentation) layer consists of the public interfaces, i.e. the products and services.

F.3.2 *Searching and showing information in this model works as follows:*

The user searches, for example through a widget, a set of metadata that was harvested from different sources and has perhaps been enriched with additional information. This metadata may contain references to central registers, in this case, the register of archive creators. If the user wishes to access the information found, the data – which is managed by an archival institution – is shown in a viewer made available by the institution (actually, the data holder), including metadata to guarantee authenticity and integrity.

F.3.3 *Searching and showing information from various manifestations*

A user searches on the basis of metadata or content, not on the basis of the type of manifestation. A users may be interested in […].

Here is an example:
A user looks for information about a sailor called "Jansz" on the VOC ship "Rembrandt". He types "Rembrandt" and "Jansz" into the search box. The archive may contain:
An original archival item;
A scan of this archival item;
A transcription of the scan;
Information about "Jansz" from this archival item in A2A format.

A historian has a translation of the transcription into modern Dutch on their website. The manifestation-independent metadata also contains references to all these manifestations. The user then receives the answer:
"Click here if you would like to request the original". (NB In practice it will not be possible to request the original as this is also a scan; this is just to give an idea.) The user will then be asked to login and reserve the item.
"Click here if you would like to see the scan." The system will then show the item.
"Click here if you would like to see the transcription."

Theme G Further details on the technical infrastructure

G.1 The concepts housing and hosting

G.1.1 Description of housing

Housing makes it possible to incorporate one's own IT systems (servers) into a Data Centre. A Data Centre provides an optimally-secured system environment. This security concerns physical access to the data centre as well as to the redundantly equipped facility for the power supply (double-sided closed electricity), climate control (heating, ventilation and air-conditioning (HVAC)) and connection to the Internet (glass fibre) for the systems. Housing therefore means hiring a space with basic facilities where you are responsible for the server hardware and software.

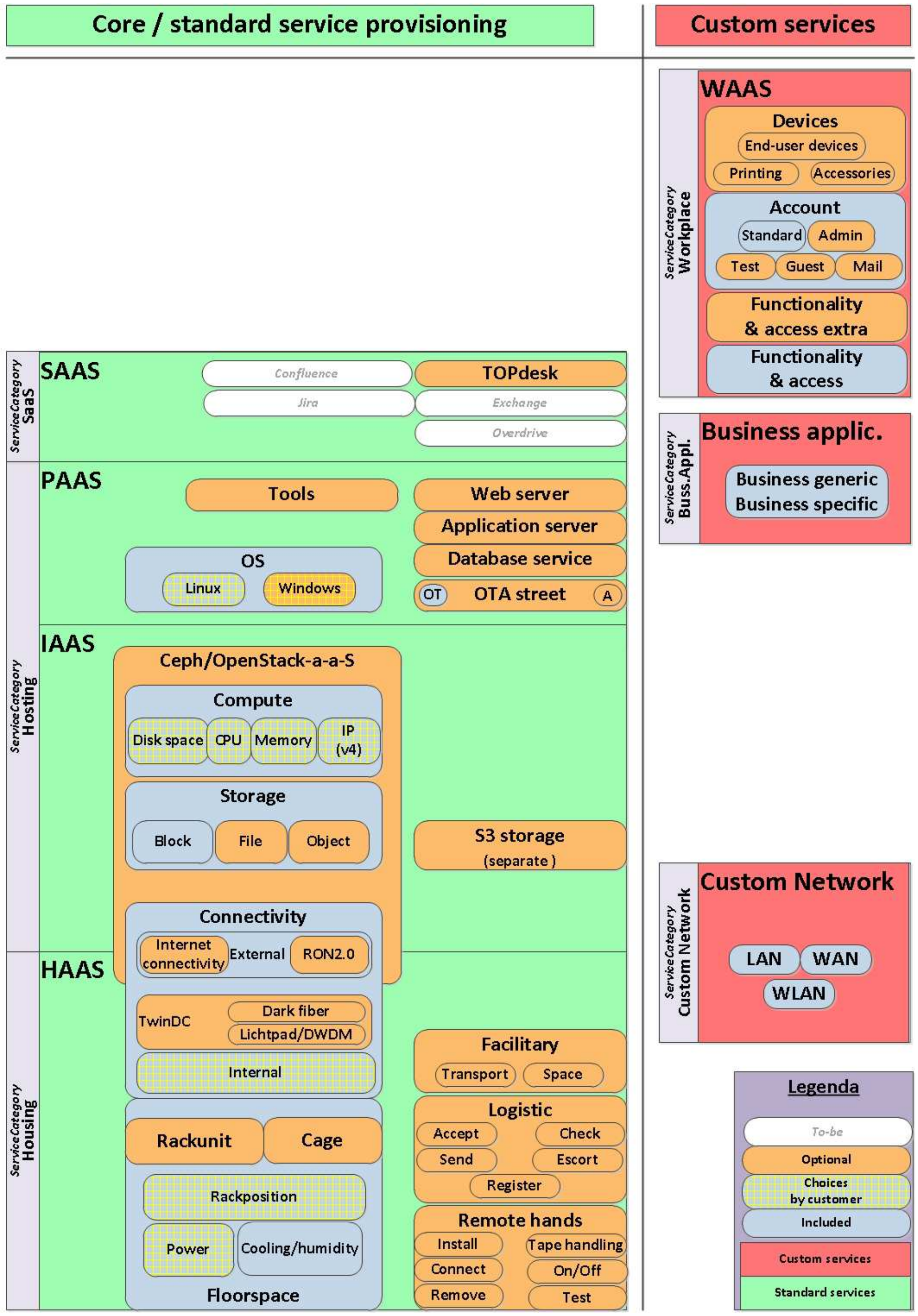
G.1.2 Description of hosting

Hosting or dedicated server hosting means using a hired service. The main advantage of dedicated server hosting is that the entire infrastructure is outsourced (both hardware and software, through "virtualization") In addition to the space and accompanying housing facilities, the physical infrastructure is also provided as a service. Agreements are made about the provision of hosted services, such as availability, scaling up and down, the type of management etc.

Another version of housing and hosting is the cloud facility. With "cloud services" clients only receive functions. In principle, a data centre can place servers all over the world without this causing any nuisance to the client. The main difference between hosting and cloud in obtaining functions is the ability of these functions to work together. Cloud-based services have been developed to integrate better with each other than hosted services. Typical examples of this are the Google Apps such as Gmail, Google Docs, Agenda and YouTube. An Internet connection and browser (client interface) are sufficient to enable integrated use. These are known as SAAS (Software-as-a-Service) services. Other possible cloud options include PAAS (Platform-as-a-Service); environments for developing or testing applications (OS software, e.g. Java runtime, webserver, database, etc.) and IAAS (Infrastructure-as-a-Service); managing the entire infrastructure environment (OS, VMs, firewalls, load balancers, storage, networking, etc.).

G.2 Organization of ODC North

ODC North has chosen a platform based on OpenStack and CEPH for hosting. OpenStack is a powerful and robust open source software platform for building and using a Cloud infrastructure. In Government Data Provision, housing involves providing floor surface, cupboard space, basic cables, and facilities (in accordance with the definition of the Housing Steering Group Programme for the Consolidation of Data Centres for ICCIO). The adjoining diagram gives an overview of the services provided by ODC North.



7 **Organization of ODC North**

G.3 The concepts of back-up, recovery and contingency

G.3.1 Description of back-up

A back-up or reserve copy is a copy of data on a data carrier or within an application so that they can be recovered if they are damaged. These copies are a preventative measure to protect important information in case the original carrier is lost or damaged. If necessary, a back-up can be reinstalled on a similar original carrier.

There are several ways to make a back-up. They fall roughly under two methods:

- back-ups of the entire hard disk, known as image back-ups;
- back-ups at the file level

Image back-ups are mainly suitable for quickly making an identical replacement disk in the event of a catastrophic error on a hard disk Back-ups at the file level are used to reinstall files without influencing the rest of the system if there are errors in one or more files (e.g. accidental deletion of all or part of a file folder). It is also possible to reinstall older versions of files insofar as there are back-ups, so that they can be compared to the current version.

G.3.2 Description of recovery

When we use the general term recovery, we specifically mean data recovery. This involves restoring the original data by:

- recovering files using recovery software;
- reinstalling back-up data.

To recover data, processes are set up and agreements are made with involved parties (business and IT) regarding procedures.

G.3.3 Description of contingency

In the event of an enormous catastrophe in which the production environment is permanently lost (major fire, bomb attack etc.) the National Archives will make services (including the e-Depot) and data available at a contingency location. The choice of this type of contingency would determine recovery time in the event of a catastrophe. The Collection Management System will be at least partly available within a week. Depending on the choice of contingency facility it may take up to a maximum of 6 months before all e-Depot services and data are once more available (in the event that the contingency facility only consists of a script and it is necessary to find and set up a location and restore all e-Depot data from a back-up).

Making data available at another location and ensuring that it is operational within the agreed-upon time guarantees business continuity. The National Archives are responsible for ensuring that a full set of data is present in at least three physical locations. There must be two sets of data at at least 5 km distance from each other. The data at the primary location must be stored at least 50 km away from the other two locations. The data must be up-to-date (back-up) and the agreed upon primary processes must be able to operate from the contingency location. This makes demands on the organization of the contingency facility (technical infrastructure) and harmonization (processes) with the primary location.

G.4 Concepts related to the network (zoning, data connections)

G.4.1 Description of zoning

The network infrastructures of the various data are compartmentalized into zones. A zone is a defined network of IT facilities where data may be exchanged freely. Data exchange with other zones is conducted through defined interfaces. The primary objective of zoning is to isolate risks so that threats and incidents in one zone cannot spread to other zones. Relevant zones are:

- A Demilitarized Zone (DMZ; semi-trusted zone): a border zone or linking zone between zones with differing confidentiality. This is used for web hosting, data

and file exchange, and authentication links via Reverse Proxy. A firewall provides a logical separation between the DMZ and the "outside" world. The National Archives distinguish two DMZs:
  - DMZ with Internet (HTTPS/FTPS via TLS), for communication with RHCs (using Digilinking)
  - DMZ with the "Haagse Ring" (linking network connected to Rijksweb Diginetwerk) for communication with departments, among others.
- Internal National Archives (production environment; trusted zone): controlled zone for application and database servers and systems. Logically, the DMZ and the Internal National Archives network zone are separated by a firewall.

Networks are divided into VLANs and routing between VLANs is conducted via a combined router/firewall. Communication between internal networks, between internal networks and the DMZ, and between the DMZ and the Internet is usually kept closed and is only opened when necessary. Separate ports or port ranges for entire subnets can be opened from secure to insecure and between secure networks; opening from insecure to secure is only possible for specific hosts.

The network is monitored and managed so that attacks, disruptions or errors can be discovered and recovered and the network's reliability never falls below the agreed-upon minimum level.

G.4.2 Description of data connections

Information exchanges take place both to and from the National Archives using network connections. The following standards have been set for data connections:

- By default, the National Archives enable organizations to use the e-Depot via an Internet connection with Hypertext Transfer Protocol Secure (HTTPS) and FTP on SSL (FTPS) with Transport Layer Security TLS 1.0, TLS 1.1 or TLS 1.2 (or more recent);
- The National Archives enable organizations to use the e-Depot via the Diginetwerk (and connected linking networks Gemnet and Haagse Ring) with HTTPS and FTP using TLS 1.0, TLS 1.1 or TLS 1.2 (or more recent);
- A specific SSLCipherSuite configuration in the SSL configuration is used, whereby the Cipher Order within the CipherSuite is enforced;
- The e-Depot's SSL configuration uses 2048-bits Diffie-Hellman parameters.

This configuration produces the 'A' score (on the American A-to-F scale, A being the highest) in SSL Labs' checking tool.

We distinguish the following types of connections:

- From central government legal caretakers to the National Archives
- From RHCs to the National Archives
- From local authorities to the National Archives

G.5 Rationale for gradual transition by means of a temporary solution

The are two reasons to gradually migrate to ODC North:

G.5.1 Business continuity: complexity of the DTR

In the period leading up to the migration to ODC North, there is a risk to the continuity of business operations and the services provided by the National Archives. A risk analysis will be conducted. Not only will the risks be mapped out, but migration measures will be presented for various scenarios. These measures must comply with the requirements that apply up to the transition to ODC North. They mainly concern non-functional requirements in the area of availability, performance and security, with choices related to storage, back-up, recovery, and contingency facilities.

G.5.2 ODC North's back-up and contingency location is not yet in order

A possible scenario is that the National Archives will retain their existing server space within their building for the duration of the DTR programme so that they can fulfil the obligations ensuing from the programme. This means the following: housing and hosting the hardware and software for which the National Archives are responsible will be conducted on the National Archives' premises. The National Archives will start migrating to ODC North immediately after completing the programme (2017).

ODC's contingency facility is not yet optimally arranged. This is marked as a risk that requires measures to be taken. ODC North will serve as a back-up and contingency location. It may at first serve only as a housing facility.

G.6 Non-functional requirements

Non-functional requirements involve a variety of quality characteristics. The National Archives' systems must comply with the requirements specified for these quality characteristics. This concerns the requirements that can be divided into the main characteristics of reliability and performance efficiency in accordance with ISO standard 25010 for IT product quality (software and systems).

G.6.1 Reliability

This concerns requirements set for a system, product or component that performs specified functions for a specified period of time. The quality characteristics of reliability, error-resistance, and recoverability are also specified.

- Availability

With a 7x24 hour service window, the availability of the technical infrastructure depends on unavailability.

- A maximum of 0.251% unplanned unavailability (22 hours) on an annual basis for the most important services. This means 22 hours of unavailability that cannot be planned for regardless of the cause are accounted for on an annual basis.
  - Availability = 99.479% uptime (100% minus unplanned unavailability of 22 hours).
- It is still the assumption that a window of four hours is planned, announced and approved once every month. This amounts to 12 x 4 = 48 hours (0.548%).
  - Availability = 99.452% uptime (100% minus planned and approved unavailability of 48 hours).

The total availability of the infrastructure depends on unplanned and planned unavailability. This means that 70 hours of planned and unplanned unavailability on an annual basis are taken into account for the availability calculation.

- Total availability = 99.202% uptime (100% minus planned and unplanned unavailability of 70 hours).

- Recoverability

In the event of an emergency that disrupts the entire National Archives' IT facilities, conditions will be set for the recoverability of data and infrastructure.

- Data may be lost for a maximum of 24 hours. This means that full back-ups must be made and stored off-site at least once a day. A period of one week applies to the e-Depot. This means that new data or data that was altered in the period prior to the emergency will NOT be saved because a back-up that was saved no more than 24 hours earlier (7 days earlier in the case of the e-Depot) will be recovered. It is possible to limit the impact by procedure and/or having the data saved by the organizations and systems that supplied it.

- RPO (Recovery Point Objective) = 7 days for the e-Depot.
- RPO (Recovery Point Objective) = 1 day for other data.

- It may take a maximum of 5 days before the most important data and infrastructure (servers, data connections) are once more available to the archive management system. The minimum service level has still to be established. The archive management system must be once more available with an as-yet-to-be-decided minimum functionality for users both within and outside the National Archives within 5 days. This will be achieved by means of a contingency location or the restored National Archives' IT facilities. This does not include making work places available.
    - RTO (Recovery Time Objective) = 5 days for the archive management system's physical collection (CMS).

In the event of an emergency in which all the National Archives facilities fail, it may take a maximum of 6 months before all data and required infrastructure (servers, data connections) are once more available and services are fully restored. This will be achieved by using a contingency location or restoring the National Archives' IT facilities. This does not include making work places available.

- MTPOD (Maximum Tolerable Period of Down time) = 6 months.

G.6.2 *Performance efficiency*

This involves performances related to the amount of resources used under the stated conditions. The quality characteristics of speed, resource use and capacity have been specified.

- Capacity

The storage capacity of the technical infrastructure has been calculated for the period up until 2030 on the basis of the Financial Explanatory Notes, version 2.0, dated 15 January 2013. This concerns the available Net Storage Capacity. The growth figures are in conformity with the Financial Explanatory Notes, version 2.0, dated 15 January 2013.

- Scenario II
- National archives central collection, see Figure 14, scenario II on page 38
- RHC's decentralized collection, see Figure 13 on page 41.

These figures are important for deciding the capacity of scalability of both the facility part (delivery service space) and IT part (delivering the National Archives' working technical infrastructure) in the long term.

They include m^2 of floor surface, floor load, energy use, required cooling capacity.

- Net Storage Capacity to be available per
    - 31-12-2016 = 1.6 PB (petabyte)
    - 31-12-2029 = 26.4 PB (petabyte)

  Disaggregated per year:
    - 31-12-2013 0.6 PB
    - 31-12-2014 0.9 PB
    - 31-12-2015 1.3 PB
    - 31-12-2016 1.6 PB
    - 31-12-2017 2.1 PB
    - 31-12-2019 3.0 PB
    - 31-12-2020 3.6 PB
    - 31-12-2021 4.4 PB
    - 31-12-2022 5.3 PB
    - 31-12-2023 6.4 PB
    - 31-12-2025 9.8 PB
    - 31-12-2026 12.3 PB
    - 31-12-2027 15.8 PB
    - 31-12-2028 20.2 PB
    - 31-12-2029 26.4 PB

- 31-12-2018 2.5 PB 31-12-2024 7.9 PB

Appendix I. Explanatory notes on architecture14

I.1 Concept of "architecture"

An architecture is a coherent, consistent collection of principles, differentiated into basic premises, rules, standards and guidelines.

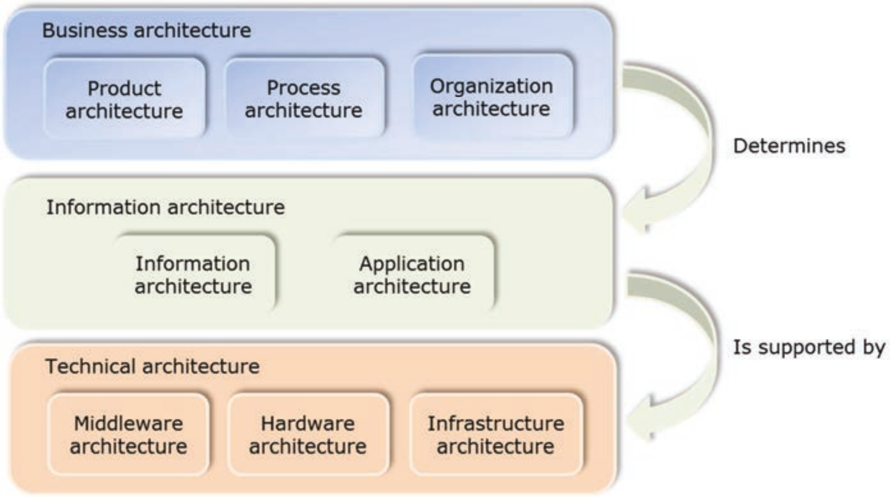
An architecture can describe the current situation ("Ist") or a future or desirable situations ("Soll" or "Target").

I.2 What must be described? Business, information and technology

When describing architecture, a distinction is often made between business architecture, application or information architecture, and technical architecture.

An *enterprise architecture* focuses on all three layers for a specific "enterprise". "Enterprise" refers to a unit that shares specific aspects of processes, information or infrastructure. This type of unit may consist of various organizations, such as RHCs and the National Archives.

See the diagram below.



A *business architecture* describes the organizational contours required to archive the business objectives. A business architecture describes:

- Which products and services make it possible to achieve the business objectives (product architecture);
- The processes that are necessary to be able to supply these products and services (process architecture);
- The organizational structure that is necessary to steer these processes (organizational architecture).

Information architecture describes the information that is necessary and the route of information flows. This may concern all information flows, both automated and non-automated, between people and between applications, both concerning business operation applications and processes and supporting applications and processes. An information architecture describes:

[14] This chapter has been largely taken from the WVI Architecture.

- Which information is necessary for an organization to function (information architecture);
- The applications that ensure that information is distributed (application architecture).

Technical architecture concerns the technology on which applications operate and information is saved. Technical architecture relates to:

- Hardware
- Network components
- Standard software that is required to share information between applications.

I.3 Methods for describing architectures Principles, models and coherence

Architectures are usually described in terms of principles and models.

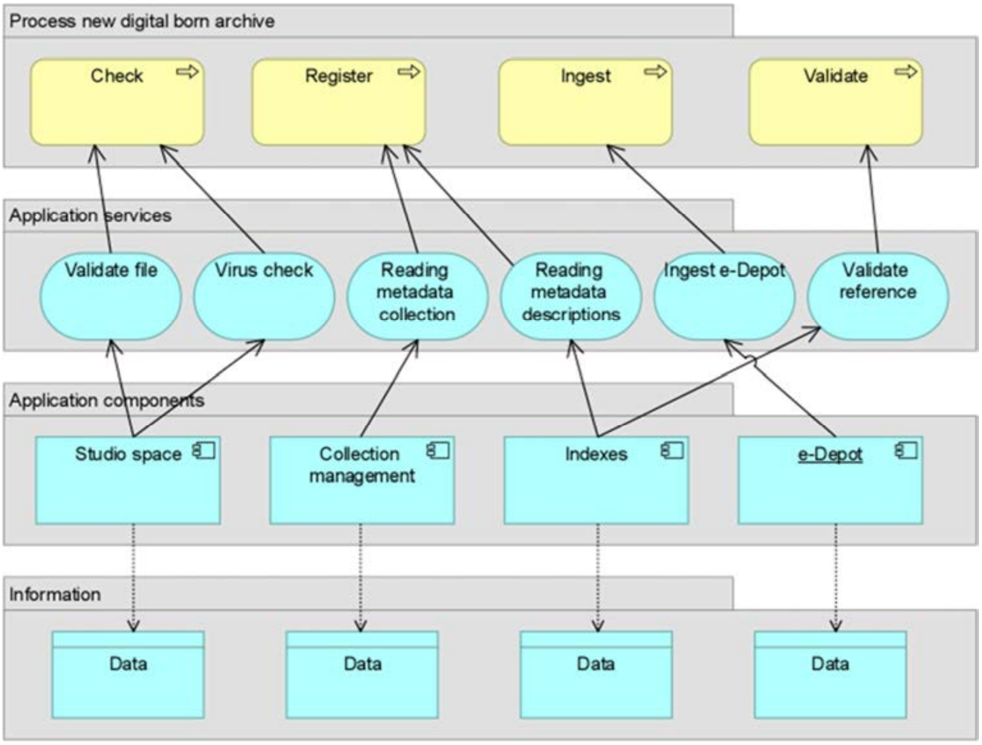
Architecture principles are directive agreements that convey a conviction on the way in which the desired situation can be brought about. Each of the architecture principles will be described using:

- Description: which principle is being used?
- Rationale (motivation): why is this principle being used?
- Implications (consequences): what are the consequences for the architecture?

Models are also used. To describe models in as standard a way as possible, this report uses Archimate, the standard architecture model language used by The Open Group. This makes it easy to extend the models drawn up in this report at a later date and to make depth layers in the descriptions of the architecture. Finally, it is possible to make a connection with other organizations using this shared architectural language.

I.4 Coherence

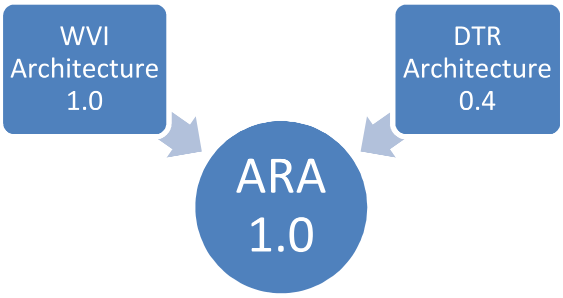
Naturally, *coherence* between the architecture elements is of crucial importance. The whole must be consistent. Processes within process architecture use application services. These services are provided by application components. Application components, in turn, manage part of the information. See the example below:


Process new digital born archive
Check
Register
Ingest
Validate
Application services
Validate file
Virus check
Reading metadata collection
Reading metadata descriptions
Ingest e-Depot
Validate reference
Application components
Studio space
Collection management
Indexes
e-Depot
Information
Data
Data
Data
Data

Appendix II. Relation to earlier architectures

II.1 WVI + DTR = MARA

This document follows on from two earlier documents: WVI Architecture and DTR Architecture. Below you can read more on the objective of these earlier architectures and how the contents of both documents was processed and supplemented.



II.2 Relation to DTR Architecture

II.2.1 Objective of DTR Architecture

DTR Architecture played a role in giving substantive direction to the DTR programme and various projects within it.

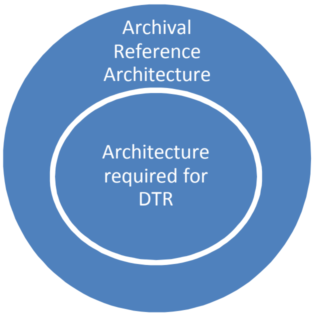
II.2.2 Core points of the content of DTR Architecture

In DTR Architecture, a great deal of attention is paid to building relationships between different (archive and other) collections using a connection layer. Attention is also paid to aggregators and core registrations/authority files. The basic principles described in this document will be adapted in full in MARA but the title of DTR architecture will no longer be used: there will be no second version of a document bearing the title
DTR Architecture

II.2.3 MARA in relation to DTR Architecture

The National Archives and RHCs have a long-term collaboration relationship which extends past the duration of the DTR programme.

What happens in the DTR is more narrow than the scope of MARA. For example: up until now, processes for registering visitors have not fallen under DTR’s scope, even though all archival institutions are familiar with it. The same applies to dealing with information that does not fall under the National Collection but which belongs to the tasks of archival institutions. The architecture required for DTR is therefore part of MARA. This is illustrated in the diagram below:


Archival Reference Architecture
Architecture required for DTR

II.3 Relation to WVI Architecture

II.3.1 Objective of WVI Architecture

The primary objective of WVI Architecture is to support preparations for the implementation of the e-Depot. In the Convent, it was decided that all RHCs should use the e-Depot for the National Collection. This gave rise to the question of how the relationship between the e-Depot and other components, such as an archive management system was and how it should be. Another question was what the implementation of the e-Depot would mean for the various work processes. The work processes have been elaborated on in a separate document: WVI Processes.

II.3.2 Core points of the content of WVI Architecture

A number of principles formulated in WVI Architecture continue to apply, such as the “two-entity model” and “separation of manifestation-independent metadata from manifestation-dependent metadata”. WVI Architecture is less concerned with other aspects, such as access. There is also no attention paid to the management of other collections managed by the archival institutions.

II.3.3 MARA in relation to WVI

The still-applicable principles from WVI Architecture will be adopted in MARA.

In MARA, WVI Architecture has been mainly extended with regard to the following points:

- The possibility that archives will preserve material that has not yet been transferred under public records legislation has been taken into consideration;
- Greater focus will be placed on the presentation layer;
- Attention will be paid to creating a semantic and technical connecting layer;
- It meets the desire of several archival institutions to use what are known as “authority files” or core registrations.
- The possibility of “enriching” information, by the archival institution or by the “crowd” will be considered;

Appendix III. Development and architecture sheets

The architecture is continuously being developed in more detail. Specific topics will lead to additions or extensions to this document, or will be described in the "architecture sheets" accompanying this document. Insights from the projects will be incorporated back into the architecture. MARA is therefore a living product.

III.1 Architecture sheets that have already been developed

Title	**Date**	**Description**
Description of link between the e-Depot and CMS (*Koppelvlakbeschrijving e-Depot – CBS*)	Version 1.1, October 2015	This document describes the way in which the link between the e-Depot application and a Collection Management System (CMS) should be given shape.
Technical appendix to link between the e-Depot and CMS (*Koppelvlakbeschrijving e-Depot – CBS*)	Version 1.0, October 2015	Technical appendix to aforementioned document.

III.2 Subjects that have still to be developed further

- Architecture concerning outsourced archive material;
- Handling user-generated content (e.g. metadating by "the crowd");
- Dealing with several storage sites for transferred archive material, such as audio-visual archive material for Beeld & Geluid, material in DANS;
- Dealing with several instances of the e-Depot;
- Using one search query to search in several collections (archives, photos, newspapers, audio-visual material) of one's own institution.

Appendix IV. Services provided by the National Archives

IV.1 E-Depot services for RHCs

RHCs can use the National Archives' infrastructure. This means that the RHCs can use both the e-Depot and several help applications. This enables them to implement the "Management and preservation of outsourced archives" and "Management and preservation of transferred archives" functions.

In concrete terms, this means that the services give RHCs

- access to the e-Depot application through a web interface;
- access to related stand-alone applications.

The National Archives PDC for RHCs presents products and services that are part of the NA's services related to the e-Depot for outsourced and transferred archive material. These services relate to both governmental material and non-governmental material. The Products and Services Catalogue specifies the Service Organization's services to the RHCs. A cost model is attached to this catalogue.

IV.2 Other services for RHCs

In addition to e-Depot services, the National Archives provide RHCs with additional services, such as training and advice related to e-Depot services. For more information, see National Archives PDC for RHCs.

The core of the services portfolio consists of basic services used by every RHC. RHCs may decide not to take optional services, to implement them themselves or to receive them from the Service Organization.

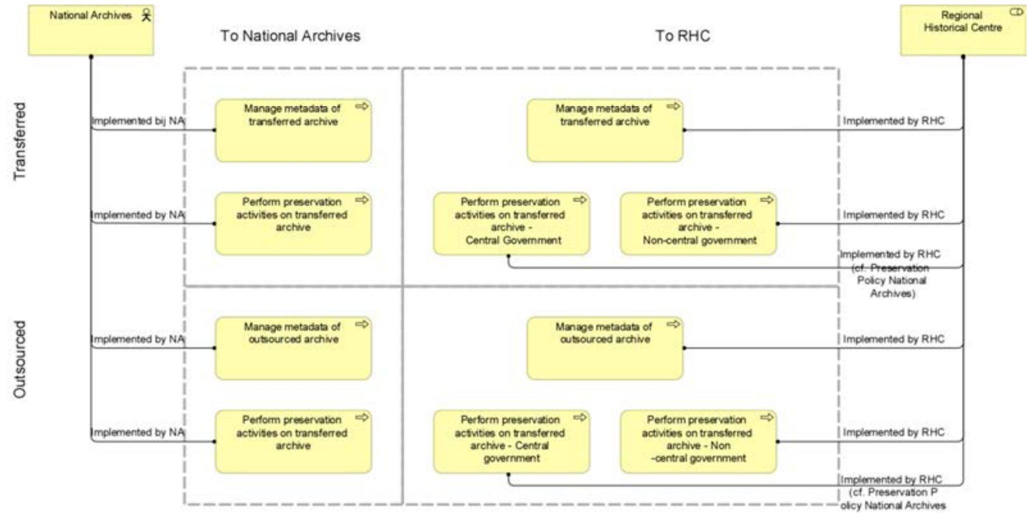
NB Other advisory services fall under the "Advice" service as described in section 5 *Business architecture: services*.

IV.3 Distribution of work between RHCs and the National Archives

The RHCs manage the archive records stored in the e-Depot. The RHCs implement the National Archives' preservation policy for the National Collection. The RHCs remain responsible for managing and providing access to all archive records that have been transferred to them. Authorities that have outsourced material remain responsible for "their" archive records.

IV.3.1 Implementing preservation activities

The diagram below shows the allocation of tasks related to preservation activities:


National Archives
To National Archives
To RHC
Regional Historical Centre
Transferred
Outsourced
Implemented bij NA
Implemented by NA
Implemented by NA
Implemented by NA
Manage metadata of transferred archive
Perform preservation activities on transferred archive
Manage metadata of outsourced archive
Perform preservation activities on transferred archive
Manage metadata of transferred archive
Perform preservation activities on transferred archive - Central Government
Perform preservation activities on transferred archive - Non-central government
Manage metadata of outsourced archive
Perform preservation activities on transferred archive - Central government
Perform preservation activities on transferred archive - Non -central government
Implemented by RHC
Implemented by RHC
Implemented by RHC (cf. Preservation Policy National Archives)
Implemented by RHC
Implemented by RHC
Implemented by RHC (cf. Preservation P olicy National Archives

IV.3.2 *Implementing destruction services on outsourced archives*

After the retention period for outsourced archives has expired, the National Archives and RHCs will destroy this material after obtaining permission from the legal caretaker.

Appendix V. Structure and organization of a national knowledge network.

V.1 The National Archives' knowledge function

In 2011, the Archive Vision assigned the National Archives the task of strengthening the knowledge function in the archives sector. The Archief2020 programme implemented an innovation programme for the 2012-2016 period. To continue this task and programme, the National Archives appointed a quartermaster for the knowledge function and knowledge broker in late 2014. To align the knowledge function with the needs of the target group (archival institutions and authorities), talks were held with key figures in the archive sector. These outcomes, combined with the results of studies of relevant internal documents, such as the National Archives' Plan for the Future, were processed in the Project Plan to Strengthen the Knowledge Function in the Archive Sector (2015). In this plan, the National Archives elaborated on its function as a hub, in conformity with agreements made within the framework of initiatives outlined in the National Strategy for Digital Heritage (NDE). Possibilities for connecting the entire heritage network in the area of knowledge are explored in this position paper.

V.2 The National Archives' Action Line

The National Archives base their basic principles on the idea of a network: the sector will only be able to effectively deal with the enormous challenges it is facing (such as digitalization and open access) if it works together as a whole to develop and share knowledge. This is why we do not speak of a National Archives' Knowledge Centre, but rather of a National Knowledge Network. The National Archives emphatically wish to be an active participant in this. Four lines of action have been developed in the project plan:

- Action line 1: Knowledge and Innovation Agenda

Introduction of a Knowledge and Innovation Agenda to be shared and jointly implemented by the archive sector. This will make the relevant knowledge issues clear to the archive sector, shape the required knowledge development and establish the joint implementation of this agenda. This agenda will be updated annually on the basis of further discussions on the mission and the course to be taken.

- Action line 2: Knowledge sharing and provision

This includes structuring discernible facilities for sharing knowledge and making it available, such as a knowledge bank, a broker's function, good website facilities, a newsletter and setting up knowledge platforms for information sharing. A helpdesk will be set up for knowledge questions and a business newsletter is currently being prepared. These platforms are laying a basis for the network and providing space for further knowledge development (mainly concerning themes from the K&I Agenda) in collaboration spaces and knowledge files that are to be developed. They are separate from the National Archives' website, where formal and established information can be found. Another important aspect is that knowledge platforms working online have an obligation to organize knowledge sessions as experience teaches that knowing each other and meeting face-to-face are important factors that influence the success of a community.

- Action line 3: Building a knowledge network

Knowledge is created and multiplied by social interaction. This involves setting up a knowledge network (based on a growth model) that is responsible for the exchange of experiences, ideas, new developments in the sector, points of view and thoughts between stakeholders in the archive sector (via the aforementioned platforms) and beyond. This concerns other heritage domains, universities and specialized knowledge institutes such as NIOD (National Institute for War Documentation), Beeld & Geluid, IISG (International Institute of Social History) etc.

- Action line 4: Organization development

Proposals for organizational HRM measures focus on the introduction of knowledge management. In the first instance this will be within the National Archives, but they will also serve as a model for a knowledge function outside of it. Themes include how to steer towards acquiring knowledge, safeguarding knowledge, sharing knowledge internally and developing competencies of knowledge workers.

V.3 Safeguard as from 2016

An important task is the safeguarding of the achievements of the Archief2020 innovation programme that ended in 2016. Accrued knowledge, experience, products and cooperation will be embedded in the National Archives’ knowledge function. In early 2016, the project plan will be evaluated and new activities and accents may be introduced.

V.4 Knowledge questions on the Knowledge & Innovation Agenda

1. Policy framework and guide to valuation and selection for central and decentral governments
2. Structuring principles of central government organizations (DUTO)
3. Guide and best practices for digital depot facilities
4. Policy plan and preservation tools
5. Policy framework for public access in a digital environment
6. Open access
7. Open collection data policy
8. Exploration of e-discovery/big data for valuation and selection
9. Guide and aids for private archives and documentation of society
10. Guide to archive documents outside the DMS
11. Towards the front of the chain (SIO, independent expert)
12. Metadata
13. Quality system for digital archive management
14. Chain automation
15. Digital service provision

V.5 Themes for knowledge platforms

1. Organization of information management for governments
2. Quality systems
3. Valuation & Selection
4. Preservation & Permanence
5. Metadata
6. Access
7. Public access
8. Digital acquisition & the e-Depot
9. Conservation (paper)
10. Presentation & Service Provision
11. Professionalization & increasing expertise