Saving the Web
The Promise of a Belgian Web Archive
Welcome

Sara Lammens – KBR
Dr Karel Velle – State Archives Belgium
Keynote

Prof. Dr Niels Brügger – Aarhus University
The PROMISE project and a Belgian strategy

Friedel Geeraert – KBR / State Archives Belgium
Dr Peter Mechant – Ghent University
Alejandra Michel – CRIDS – UNamur
Dr Rolande Depoortere & Sébastien Soyez – State Archives Belgium
Sophie Vandepontseele – KBR
Introduction

Sophie Vandepontseele - KBR
Il est très souhaitable que les informations successives publiées sur des sites officiels publics belges restent accessibles à tous, si possible via un site internet d’archives lui-même officiel. **C’est utile en termes de sécurité juridique, pour les chercheurs et bien sûr pour les archivistes.**

It is highly desirable that the successive information published on Belgian official public websites should remain accessible to all, if possible via an official web archive. This is **useful in terms of local security, for researchers and of course for archivists.**

**Didier Coeurnelle**  
Conseiller stratégique communication  
Service Public Fédéral (SPF) Sécurité sociale
Testimonials

For the most part, historians cannot write histories of the 1990s unless they use web archives: with them, military historians will have access to the voices of rank-and-file soldiers on discussion boards; political historians, to blogs, the cut and thrust of websites, electoral commentary and beyond; and of course, social and cultural historians, to the voices of the people on a scale never before possible.

What happens if we do nothing?

Irremediable loss of information and historical sources

Historians will be unable to write about our history taking into account all sources
Starting late, is that positive?

Limitation of youth diseases

Benefit from the latest technologies

Take into account the current user needs

Participate in the development of new fields of research
The PROMISE project

Friedel Geeraert – KBR / State Archives Belgium
What is web archiving?

**Web archiving** is the process of **collecting** portions of the World Wide Web, **preserving** the collections in an archival format, and then serving the archives for access and use.

The PROMISE project

AIM
To develop a federal strategy for the preservation of the Belgian web

TIMING
July 2017 – December 2019

FUNDING
belspo
The PROMISE project goals

1. Identify best practices in the field of web archiving
2. Develop a Belgian web archiving strategy
3. Pilot the archiving of the Belgian web and provide access to the collections
4. Make recommendations for the implementation of a sustainable web archiving service
The PROMISE project

- Literature review
- Interviews with representatives of web archiving initiatives

Identify best practices in the field of web archiving
The PROMISE project

- Results ‘state of the art’
- Legal analysis
- Survey UGhent on user requirements
- Continuous dialogue State Archives & KBR

Develop a Belgian Web Archiving Strategy
The Belgian web: a history

1988
Creation of the .be domain

1994
129 registered .be domain names

2012
Creation of .vlaanderen and .brussels

2019
1.6 million .be 6400 .vlaanderen 7900 .brussels
Which extensions do Belgians choose for their domain name?

• Only 54.42% of Belgians choose the .be domain for their website.
• Other popular domain names are .com, .eu, .net etc.
• It is very important to keep the relevant content outside of the .be domain into account when developing selection policies.
The PROMISE project

- Pilot the archiving of the Belgian web and provide access to the collections
- Make recommendations for the implementation of a sustainable web archiving service
Conferences

InterPARES Trust
WEB ARCHIVING: BEST PRACTICES FOR DIGITAL CULTURAL HERITAGE
29-30 April 2018
Utrecht University, Utrecht, The Netherlands

Dag van de Nieuwste Geschiedenis
Journée de l’Histoire Contemporaine

IIPC
netpreserve.org
INTERNATIONAL INTERNET PRESERVATION CONSORTIUM

DLM FORUM

AAFB

DARIAH-EU
Digital Research Infrastructure for the Arts and Humanities

netwerk digitaal erfgoed

RESAW
A Research Infrastructure for the Study of Archived Web materials

Informatie aan Zee 2019

Benelux
Amsterdam
2018

promise
KBR

Rijksarchief
Archives de l'Espace
Staatarchief
www.arch.be

Ghent University

HE2B
HAUTE ECOLE BRUXELLES-BRABANT

Universiteit voor Antwerpen
Publications


Publications


The state of the art

Dr Peter Mechant – Ghent University
**BROAD CRAWLS**
(superficial capture)

1. The national domain (top-level domain crawls)

2. Other websites that are considered interesting

---

**SELECTIVE CRAWLS**
(complete capture)

1. Themes

2. Events

3. Emergencies
<table>
<thead>
<tr>
<th>Country</th>
<th>Institution</th>
<th>Broad crawls</th>
<th>Selective: thematic</th>
<th>Selective: events</th>
<th>Selective: other</th>
</tr>
</thead>
<tbody>
<tr>
<td>Netherlands</td>
<td>Nat. Library</td>
<td>No</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td>France</td>
<td>Nat. Library</td>
<td>Yes (top-layers only)</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes (emergencies)</td>
</tr>
<tr>
<td>UK</td>
<td>British Library</td>
<td>Yes (non-print legal deposit web)</td>
<td>Yes (open UK web archive)</td>
<td>Yes (open UK web archive)</td>
<td>Yes (emergencies- open UK web archive)</td>
</tr>
<tr>
<td>Luxembourg</td>
<td>Nat. Library</td>
<td>Yes</td>
<td>No</td>
<td>Yes</td>
<td>No</td>
</tr>
<tr>
<td>Denmark</td>
<td>Royal Danish Library</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes (emergencies, research projects, videos)</td>
</tr>
<tr>
<td>Portugal</td>
<td>Arquivo.pt</td>
<td>Yes</td>
<td>No</td>
<td>Yes</td>
<td>No</td>
</tr>
<tr>
<td>Ireland</td>
<td>Nat. Library</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>No</td>
</tr>
<tr>
<td>Canada</td>
<td>Libr. &amp; Arch. Canada</td>
<td>No (in preparation)</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes (emergencies, risk of disappearing)</td>
</tr>
<tr>
<td>Canada</td>
<td>Nat. Libr. &amp; Arch. Quebec</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
<td>No</td>
</tr>
<tr>
<td>Switzerland</td>
<td>Nat. Libr.</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
<td>No</td>
</tr>
<tr>
<td></td>
<td>Facebook</td>
<td>Twitter</td>
<td>Youtube</td>
<td>Instagram</td>
<td>Flickr</td>
</tr>
<tr>
<td>---------------------</td>
<td>--------------</td>
<td>---------</td>
<td>---------</td>
<td>-----------</td>
<td>--------</td>
</tr>
<tr>
<td>France (Nat. Libr.)</td>
<td>Not anymore</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td>Denmark (Roy. Libr.)</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>No</td>
</tr>
<tr>
<td>Luxembourg (Nat. Libr.)</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>No</td>
</tr>
<tr>
<td>UK (British Library)</td>
<td>Yes</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td>Ireland (Nat. Libr.)</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
<td>No</td>
<td>Yes</td>
</tr>
<tr>
<td>UK (Nat. Arch.)</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td>Library and Archives (Canada)</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>No</td>
</tr>
</tbody>
</table>
Suggest websites to be preserved

International websites are welcome!

Websites (1 per line)

www.site.com

Description of websites

Your email

Submit
Participate and contribute

Saving websites from being lost is not only a service to the general public but a service to site owners in the first place. If you have a website or create a new one, let us know about it. If you see a website you think should be saved for the future, tell us about it. Suggestions, questions comments on special collections, topics, data, or any other subject, we are all ears!

Suggest a website
Fields marked with an * are required

PLEASE ADD THIS WEBSITE TO THE LUXEMBOURG WEB ARCHIVE *

YOUR EMAIL ADDRESS *

FIELD OF INTEREST
Society & Politics

☐ I AM THE OWNER OF THIS WEBSITE

COMMENT OR DESCRIBE YOUR SUGGESTION
<table>
<thead>
<tr>
<th>Country</th>
<th>Institution</th>
<th>Access method</th>
<th>Who has access</th>
</tr>
</thead>
<tbody>
<tr>
<td>The Netherlands</td>
<td>National Library</td>
<td>Open &amp; freely accessible online</td>
<td>Physical access on location</td>
</tr>
<tr>
<td>The Netherlands</td>
<td>National Archive</td>
<td>Yes (for websites with an 'open' status)</td>
<td>Yes (for websites with a 'restricted' or 'offline' status)</td>
</tr>
<tr>
<td>France</td>
<td>National Library</td>
<td>No</td>
<td>Yes (but also from within the 26 partner libraries)</td>
</tr>
<tr>
<td>Luxembourg</td>
<td>National Library</td>
<td>No</td>
<td>Yes</td>
</tr>
<tr>
<td>UK</td>
<td>British Library</td>
<td>Yes (for the UK web archive)</td>
<td>Yes (for the legal deposit UK web archive and JISC domain dataset)</td>
</tr>
<tr>
<td>UK</td>
<td>National Archives</td>
<td>Yes</td>
<td>No</td>
</tr>
<tr>
<td>Denmark</td>
<td>Royal Danish Library</td>
<td>Yes (only for researchers conducting research on a Ph.D-level or above)</td>
<td>Yes (only for researchers)</td>
</tr>
<tr>
<td>Portugal</td>
<td>Foundation for Science &amp; Technology</td>
<td>Yes</td>
<td>No</td>
</tr>
<tr>
<td>Ireland</td>
<td>National Library</td>
<td>Yes</td>
<td>No</td>
</tr>
</tbody>
</table>
“A common feature of most web archiving backed by legal deposit legislation is some sort of restrictions on the access afforded to the end user of the archive” (Webster, 2017: p. 180).

Types of access:

- Open and freely accessible online + physical access on location:
  - For everyone
  - For certain profiles and/or for certain content only

- Only physical access on location

- No access
<table>
<thead>
<tr>
<th>Country</th>
<th>Institution</th>
<th>Search options</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>URL</td>
</tr>
<tr>
<td>The Netherlands</td>
<td>National Library</td>
<td>Yes</td>
</tr>
<tr>
<td>The Netherlands</td>
<td>National Archive</td>
<td>No</td>
</tr>
<tr>
<td>France</td>
<td>National Library</td>
<td>Yes</td>
</tr>
<tr>
<td>UK</td>
<td>British Library</td>
<td>Yes</td>
</tr>
<tr>
<td>UK</td>
<td>National Archives</td>
<td>Yes</td>
</tr>
<tr>
<td>Denmark</td>
<td>Royal Danish Library</td>
<td>Yes</td>
</tr>
<tr>
<td>Portugal</td>
<td>Foundation for Science and Technology</td>
<td>Yes</td>
</tr>
<tr>
<td>Ireland</td>
<td>National Library</td>
<td>Yes</td>
</tr>
</tbody>
</table>
Webarchive file formats

- WARC
- DAFF
- ARC
Webarchive playback technology

- The OpenWayback... 55.6%
- Solr 11.1%
- Viewerproxy 5.6%
- warcprox 5.6%
- SHINE 5.6%
- WebRecorder 5.6%
- NutchWAX 11.1%
Legal analysis

Alejandra Michel – CRIDS – UNamur
Legal context of web archiving

Article 10 ECHR

Web archives

Web archives facilitate research and access to information for the public.

Right to information
Heritage value
Legal context of web archiving

“The maintenance of Internet archives is a critical aspect [for enhancing public’s access to news and facilitating the dissemination of information] and the Court considers that such archives fall within the ambit of the protection afforded by Article 10”

ECHR (4th section), case Times Newspaper LTD v. the United Kingdom (Nos 1 and 2), 10 March 2009, §27.
Legal aspects of web archiving

However, **even if** web archives are protected by Article 10 ECHR, they **raise** many legal issues **at different stages**:

- Selection
- Collecting
- Archiving
- Long-term preservation
- Access
Legal aspects of web archiving

- Copyright Sui generis right on databases
- Privacy law Data protection law
- Illegal contents
- Definition of the “national web”
- Missions of national cultural heritage institutions
- Probative value of web archives
Any questions?
Strategy & Options

Sébastien Soyez – State Archives Belgium
Strategy based on OAIS-model
### Metadata schema : OCCLC choice (14)


<table>
<thead>
<tr>
<th>URL</th>
<th>Title</th>
<th>Collector</th>
<th>Genre/Form</th>
<th>Description</th>
<th>Extent</th>
<th>Rights</th>
<th>Date</th>
</tr>
</thead>
<tbody>
<tr>
<td><a href="http://www.fr.mark%D0%B1%D0%B5%D0%B3%D0%BE%D0%BD%D0%B8%D1%8F.be">http://www.fr.markбегония.be</a></td>
<td>Fondation Médocale Reine Elisabeth</td>
<td>Archives de l'État en Belgique</td>
<td>Site web organisation</td>
<td>Archivage des collections</td>
<td>1.4</td>
<td>1.4</td>
<td>1.4</td>
</tr>
<tr>
<td><a href="http://www.fr.mark%D0%B1%D0%B5%D0%B3%D0%BE%D0%BD%D0%B8%D1%8F.be">http://www.fr.markбегония.be</a></td>
<td>Fondation Prince Laurent</td>
<td>Archives de l'État en Belgique</td>
<td>Site web organisation</td>
<td>Archivage des collections</td>
<td>1.4</td>
<td>1.4</td>
<td>1.4</td>
</tr>
<tr>
<td><a href="http://www.fr.mark%D0%B1%D0%B5%D0%B3%D0%BE%D0%BD%D0%B8%D1%8F.be">http://www.fr.markбегония.be</a></td>
<td>Stichting Koninklijke Prins Albert Stichting</td>
<td>Archives de l'État en Belgique</td>
<td>Site web organisation</td>
<td>Archivage des collections</td>
<td>1.4</td>
<td>1.4</td>
<td>1.4</td>
</tr>
<tr>
<td><a href="http://www.fr.mark%D0%B1%D0%B5%D0%B3%D0%BE%D0%BD%D0%B8%D1%8F.be">http://www.fr.markбегония.be</a></td>
<td>De Belgische Senaatskamer</td>
<td>Archives de l'État en Belgique</td>
<td>Site web organisation</td>
<td>Archivage des collections</td>
<td>1.4</td>
<td>1.4</td>
<td>1.4</td>
</tr>
<tr>
<td><a href="http://www.fr.mark%D0%B1%D0%B5%D0%B3%D0%BE%D0%BD%D0%B8%D1%8F.be">http://www.fr.markбегония.be</a></td>
<td>Kamer van volksvertegenwoordigers</td>
<td>Archives de l'État en Belgique</td>
<td>Site web organisation</td>
<td>Archivage des collections</td>
<td>1.4</td>
<td>1.4</td>
<td>1.4</td>
</tr>
<tr>
<td><a href="http://www.fr.mark%D0%B1%D0%B5%D0%B3%D0%BE%D0%BD%D0%B8%D1%8F.be">http://www.fr.markбегония.be</a></td>
<td>Commissie permanent de contrôle de la Commission permanente des services de police et justiciables</td>
<td>Archives de l'État en Belgique</td>
<td>Site web organisation</td>
<td>Archivage des collections</td>
<td>1.4</td>
<td>1.4</td>
<td>1.4</td>
</tr>
<tr>
<td><a href="http://www.fr.mark%D0%B1%D0%B5%D0%B3%D0%BE%D0%BD%D0%B8%D1%8F.be">http://www.fr.markбегония.be</a></td>
<td>Commissie permanent de contrôle de la Commission permanente des services de renseignements et de sécurité</td>
<td>Archives de l'État en Belgique</td>
<td>Site web organisation</td>
<td>Archivage des collections</td>
<td>1.4</td>
<td>1.4</td>
<td>1.4</td>
</tr>
<tr>
<td><a href="http://www.fr.mark%D0%B1%D0%B5%D0%B3%D0%BE%D0%BD%D0%B8%D1%8F.be">http://www.fr.markбегония.be</a></td>
<td>Commissie permanent de contrôle de la Commission permanente des services de renseignements et de sécurité</td>
<td>Archives de l'État en Belgique</td>
<td>Site web organisation</td>
<td>Archivage des collections</td>
<td>1.4</td>
<td>1.4</td>
<td>1.4</td>
</tr>
<tr>
<td><a href="http://www.fr.mark%D0%B1%D0%B5%D0%B3%D0%BE%D0%BD%D0%B8%D1%8F.be">http://www.fr.markбегония.be</a></td>
<td>Parlement der Deutschen Reichs au Parlement de la République fédérale d'Allemagne</td>
<td>Archives de l'État en Belgique</td>
<td>Site web organisation</td>
<td>Archivage des collections</td>
<td>1.4</td>
<td>1.4</td>
<td>1.4</td>
</tr>
<tr>
<td><a href="http://www.fr.mark%D0%B1%D0%B5%D0%B3%D0%BE%D0%BD%D0%B8%D1%8F.be">http://www.fr.markбегония.be</a></td>
<td>Parlement des Pays-Bas</td>
<td>Archives de l'État en Belgique</td>
<td>Site web organisation</td>
<td>Archivage des collections</td>
<td>1.4</td>
<td>1.4</td>
<td>1.4</td>
</tr>
<tr>
<td><a href="http://www.fr.mark%D0%B1%D0%B5%D0%B3%D0%BE%D0%BD%D0%B8%D1%8F.be">http://www.fr.markбегония.be</a></td>
<td>Nederlandse Kamer van volksvertegenwoordigers</td>
<td>Archives de l'État en Belgique</td>
<td>Site web organisation</td>
<td>Archivage des collections</td>
<td>1.4</td>
<td>1.4</td>
<td>1.4</td>
</tr>
<tr>
<td><a href="http://www.fr.mark%D0%B1%D0%B5%D0%B3%D0%BE%D0%BD%D0%B8%D1%8F.be">http://www.fr.markбегония.be</a></td>
<td>Commissie permanent de contrôle de la Commission permanente des services de renseignements et de sécurité</td>
<td>Archives de l'État en Belgique</td>
<td>Site web organisation</td>
<td>Archivage des collections</td>
<td>1.4</td>
<td>1.4</td>
<td>1.4</td>
</tr>
<tr>
<td><a href="http://www.fr.mark%D0%B1%D0%B5%D0%B3%D0%BE%D0%BD%D0%B8%D1%8F.be">http://www.fr.markбегония.be</a></td>
<td>Commissie permanent de contrôle de la Commission permanente des services de renseignements et de sécurité</td>
<td>Archives de l'État en Belgique</td>
<td>Site web organisation</td>
<td>Archivage des collections</td>
<td>1.4</td>
<td>1.4</td>
<td>1.4</td>
</tr>
<tr>
<td><a href="http://www.fr.mark%D0%B1%D0%B5%D0%B3%D0%BE%D0%BD%D0%B8%D1%8F.be">http://www.fr.markбегония.be</a></td>
<td>Commissie permanent de contrôle de la Commission permanente des services de renseignements et de sécurité</td>
<td>Archives de l'État en Belgique</td>
<td>Site web organisation</td>
<td>Archivage des collections</td>
<td>1.4</td>
<td>1.4</td>
<td>1.4</td>
</tr>
<tr>
<td><a href="http://www.fr.mark%D0%B1%D0%B5%D0%B3%D0%BE%D0%BD%D0%B8%D1%8F.be">http://www.fr.markбегония.be</a></td>
<td>Commissie permanent de contrôle de la Commission permanente des services de renseignements et de sécurité</td>
<td>Archives de l'État en Belgique</td>
<td>Site web organisation</td>
<td>Archivage des collections</td>
<td>1.4</td>
<td>1.4</td>
<td>1.4</td>
</tr>
<tr>
<td><a href="http://www.fr.mark%D0%B1%D0%B5%D0%B3%D0%BE%D0%BD%D0%B8%D1%8F.be">http://www.fr.markбегония.be</a></td>
<td>Commissie permanent de contrôle de la Commission permanente des services de renseignements et de sécurité</td>
<td>Archives de l'État en Belgique</td>
<td>Site web organisation</td>
<td>Archivage des collections</td>
<td>1.4</td>
<td>1.4</td>
<td>1.4</td>
</tr>
<tr>
<td><a href="http://www.fr.mark%D0%B1%D0%B5%D0%B3%D0%BE%D0%BD%D0%B8%D1%8F.be">http://www.fr.markбегония.be</a></td>
<td>Commissie permanent de contrôle de la Commission permanente des services de renseignements et de sécurité</td>
<td>Archives de l'État en Belgique</td>
<td>Site web organisation</td>
<td>Archivage des collections</td>
<td>1.4</td>
<td>1.4</td>
<td>1.4</td>
</tr>
</tbody>
</table>
**SELECTION**

<table>
<thead>
<tr>
<th></th>
<th>Option A – Selective crawl</th>
<th>Option B – Broad crawl (sample)</th>
</tr>
</thead>
<tbody>
<tr>
<td>A.1 (AGR)</td>
<td>≈ 650 URLs</td>
<td>B.1/B.2/B.3 AGR/KBR (1%, 10% or 100% &lt; 1 million URLs)</td>
</tr>
<tr>
<td>A.1 (KBR)</td>
<td>920 URLs &amp; 1400 web pages</td>
<td></td>
</tr>
<tr>
<td>A.2 (AGR)</td>
<td>1,400 URLs</td>
<td></td>
</tr>
<tr>
<td>A.3 (AGR)</td>
<td>300 URLs</td>
<td></td>
</tr>
</tbody>
</table>

**COLLECTING** (once or multiple times / year)

- **Option A** – Selective crawl
  - Server for Collecting
    - AGR
    - KBR
  - XML/OCLC Descriptive Metadata
  - WARC Web Archive + Technical Metadata

- **Option B** – Broad crawl (sample)
  - WARC (AGR & KBR)
  - Separate Tools & Servers

**QUALITY CONTROL**

- Option A – Manual control
- Option B – Semi-automatic control

**CRAWLING TOOL**

- XML/OCLC Descriptive Metadata
- WARC Web Archive + Technical Metadata

**WEB ARCHIVE** (WARC + Technical Metadata)
DATA MANAGEMENT

STORAGE

INGEST

TRANSFER TOOLS

Option A
Shared tools

Option B
Separate tools

Option C = Secure Storage

Option B
Separate Servers

Option A
Shared Servers

DATA MANAGEMENT

XML/EAD
Descriptive Metadata

XML/MARC21
Descriptive Metadata

SYRACUSE
STORAGE (Option B with C)

Web Collection AGR

Web Collection KBR

LTP (collections AGR/KBR)

Secure Storage

ACCESS (Option A)

Web Collection AGR/KBR

Web Collection KBR

BELGIAN WEB ARCHIVE

KBR & AGR (XML/OCLC)

+ if necessary, KBR (XML/MARC21)

+ if necessary, AGR (XML/EAD)

python

Wayback Machine

USERS
The envisioned scenarios

Dr Rolande Depoortere – State Archives Belgium
Sophie Vandepontseele – KBR
BELGIAN WEB
≈ 1 million URLs

B.3
100 % URLs

B.2
10% URLs

B.1
1% URLs

Selective crawls (KBR)
(archiving complete websites or pages)

Selective crawls (AGR)
(archiving complete websites)

Broad crawls (AGR+KBR)
(archiving sample .BE web)

4 SELECTED SCENARIOS

SCENARIO 1 = FULL
A.1 A.2 A.3 A.1 B.3

SCENARIO 2 = MEDIUM
A.1 A.1 B.2

SCENARIO 3 = BASIC
A.1 A.1

SCENARIO 4 = FULL (OUTSOURCING)
A.1 B.3
COST CALCULATION METHOD

INFRA

STORAGE

TOOLS

SELECTION
QUALITY CONTROL
STORAGE
PRESERVATION PLAN
ADMIN

COLLECTING
INGEST
DATA MANAGEMENT
ACCESS

HR
ARCH
LIB
IT
SCENARIO 1 = FULL (AGR+KBR)

- **A.1** 650 URLs (full) AGR
- **A.2** 1400 URLs (full) AGR
- **A.3** 300 URLs (full) AGR
- **A.1** 920 URLs (full) + 1400 pages KBR
- **B.3** 1.000.000 URLs (sample) AGR+KBR

**SERVERS / TOOLS**
- HERTRIX
- Python
- WayBack Machine

**STORAGE** 114 TB

**ARCH**

**LIB**

**IT**

**5,76 FTE**

**HR** 273.840 €

**INFRA** 205.266 €

**TOTAL ANNUAL COST** 479.106 €
SCENARIO 2 = MEDIUM (AGR+KBR)

- **650 URLs** (full) AGR
- **920 URLs** (full) + 1400 pages KBR
- **100,000 URLs** (sample) AGR+KBR

**SERVERS / TOOLS**

**IT**

**STORAGE**

- **75 TB**

**ARC**

**LIB**

**IT**

**4.76 FTE**

**HR**

- **222,151 €**

**INFRA**

- **136,464 €**

**TOTAL ANNUAL COST**

- **358,615 €**
SCENARIO 3 = BASIC (AGR+KBR)

- **AGR**:
  - 650 URLs (full)
  - Additional 1400 pages

- **KBR**:
  - 920 URLs (full)

**STORAGE**: 71 TB

**SERVERS / TOOLS**:
- ARC
- LIB
- IT

**HR**: 222,151 €

**INFRA**: 129,752 €

**TOTAL ANNUAL COST**: 351,903 €
COMPARISON OF SCENARIOS – TOTAL ANNUAL COST

Scenario 1 - Full (AGR+KBR)  479,106 €
Scenario 2 - Medium (AGR+KBR)  358,615 €
Scenario 3 - Basic (AGR+KBR)  351,903 €
COMPARISON OF SCENARIOS – COST DISTRIBUTION AGR & KBR

<table>
<thead>
<tr>
<th>Scenario</th>
<th>Full (AGR+KBR)</th>
<th>Medium (AGR+KBR)</th>
<th>Basic (AGR+KBR)</th>
</tr>
</thead>
<tbody>
<tr>
<td>AGR</td>
<td>37%</td>
<td>24%</td>
<td>23%</td>
</tr>
<tr>
<td>KBR</td>
<td>63%</td>
<td>76%</td>
<td>77%</td>
</tr>
</tbody>
</table>
### COMPARISON OF SCENARIOS – COST DISTRIBUTION AGR & KBR (details)

<table>
<thead>
<tr>
<th>Sc. 1</th>
<th>AGR</th>
<th>KBR</th>
<th>AGR + KBR</th>
</tr>
</thead>
<tbody>
<tr>
<td>TOT</td>
<td>175.884 €</td>
<td>303.222 €</td>
<td>479.106 €</td>
</tr>
<tr>
<td>PERS</td>
<td>2,45 ETP</td>
<td>3,31 ETP</td>
<td>273.840 €</td>
</tr>
<tr>
<td>INFRA</td>
<td>50.612 €</td>
<td>154.654 €</td>
<td>205.266 €</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Sc. 2</th>
<th>AGR</th>
<th>KBR</th>
<th>AGR+KBR</th>
</tr>
</thead>
<tbody>
<tr>
<td>TOT</td>
<td>85.599 €</td>
<td>273.016 €</td>
<td>358.615 €</td>
</tr>
<tr>
<td>PERS</td>
<td>1,45 ETP</td>
<td>3,31 ETP</td>
<td>222.151 €</td>
</tr>
<tr>
<td>INFRA</td>
<td>12.016 €</td>
<td>124.448 €</td>
<td>136.464 €</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Sc. 3</th>
<th>AGR</th>
<th>KBR</th>
<th>AGR+KBR</th>
</tr>
</thead>
<tbody>
<tr>
<td>TOT</td>
<td>82.243 €</td>
<td>269.660 €</td>
<td>351.903 €</td>
</tr>
<tr>
<td>PERS</td>
<td>1,45 ETP</td>
<td>3,31 ETP</td>
<td>222.151 €</td>
</tr>
<tr>
<td>INFRA</td>
<td>8.660 €</td>
<td>121.092 €</td>
<td>129.752 €</td>
</tr>
</tbody>
</table>
Any questions?