The challenges of digital preservation as a public mission

Marco de Niet
Digital Heritage Netherlands

PREFORMA Innovation Workshop
Padua, 7 March 2017

"Digital preservation is the defining challenge of this generation."

(Nick Poole, CollectionsTrust/CILIP)

"The time to act was yesterday."

(Richard S. Whitt, Santa Clara High Tech)

Personal profile

- Involved in digital preservation work since 1995

- First head of Digital Preservation Research department at the National Library of the Netherlands

- Board member of the Dutch National Coalition on Digital Preservation

- Founder of Dutch Cultural Coalition on Digital Preservation

- Member of Steering Committee of Unesco's PERSIST

Underlying assumptions of Unesco's PERSIST

- Digital documentary heritage has critical importance for humanity as it has become the <u>primary means</u> of knowledge creation and expression.

- We have not yet been able, publicly or privately, to find <u>sustainable economic solutions</u> to achieve 'continuity of content',

- It is a <u>shared responsibility</u> of public and private parties to enhance the sustainability of the Information Society.

- The main parties to act on this responsibility are <u>Governments, Memory Institutions and ICT industry</u>.

- Unesco has the <u>convening power</u> on a global level to bring these three parties to the table.
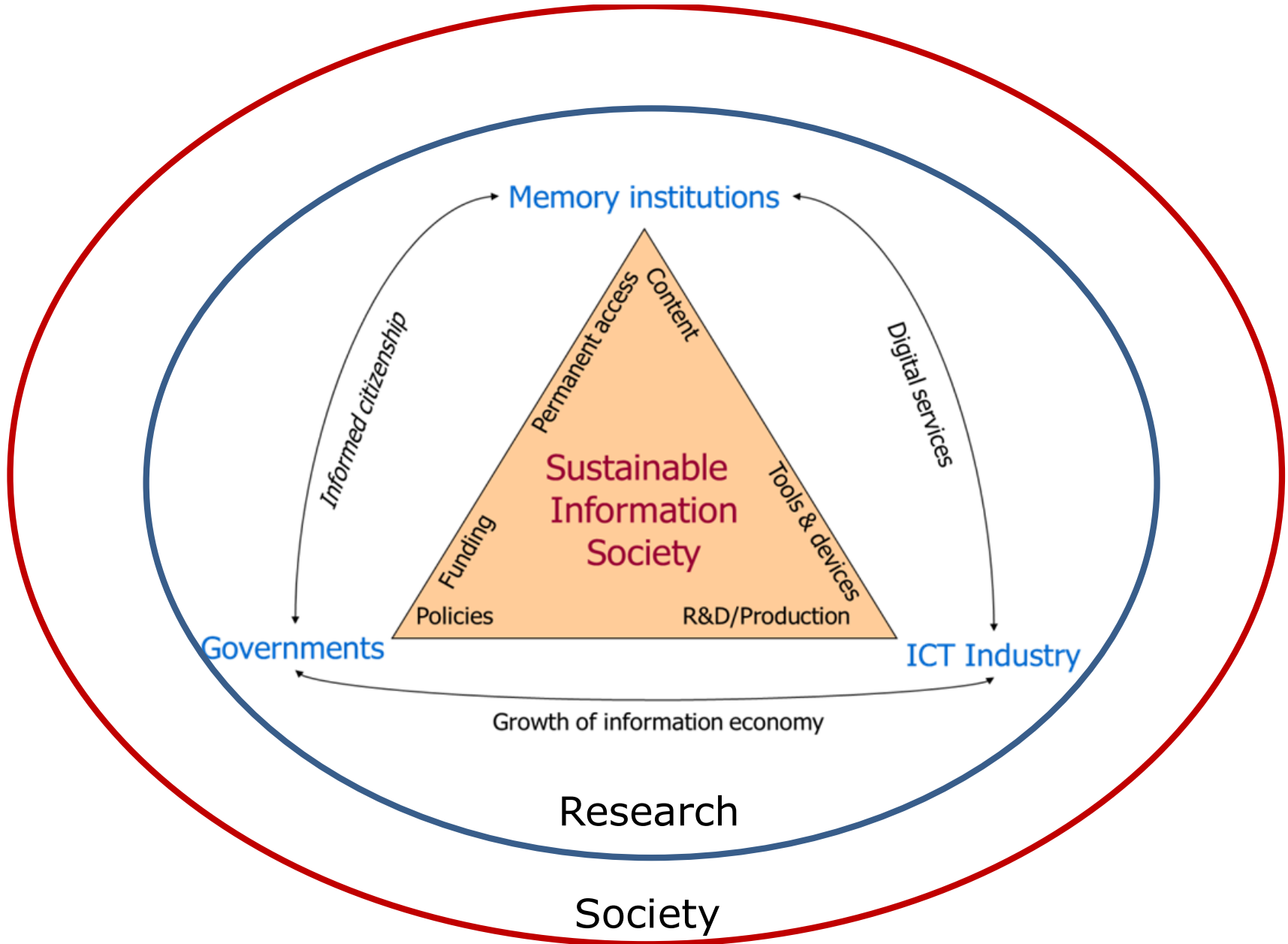
Cultural heritage institutions

2 issues:
We're not there yet...
What do we preserve?

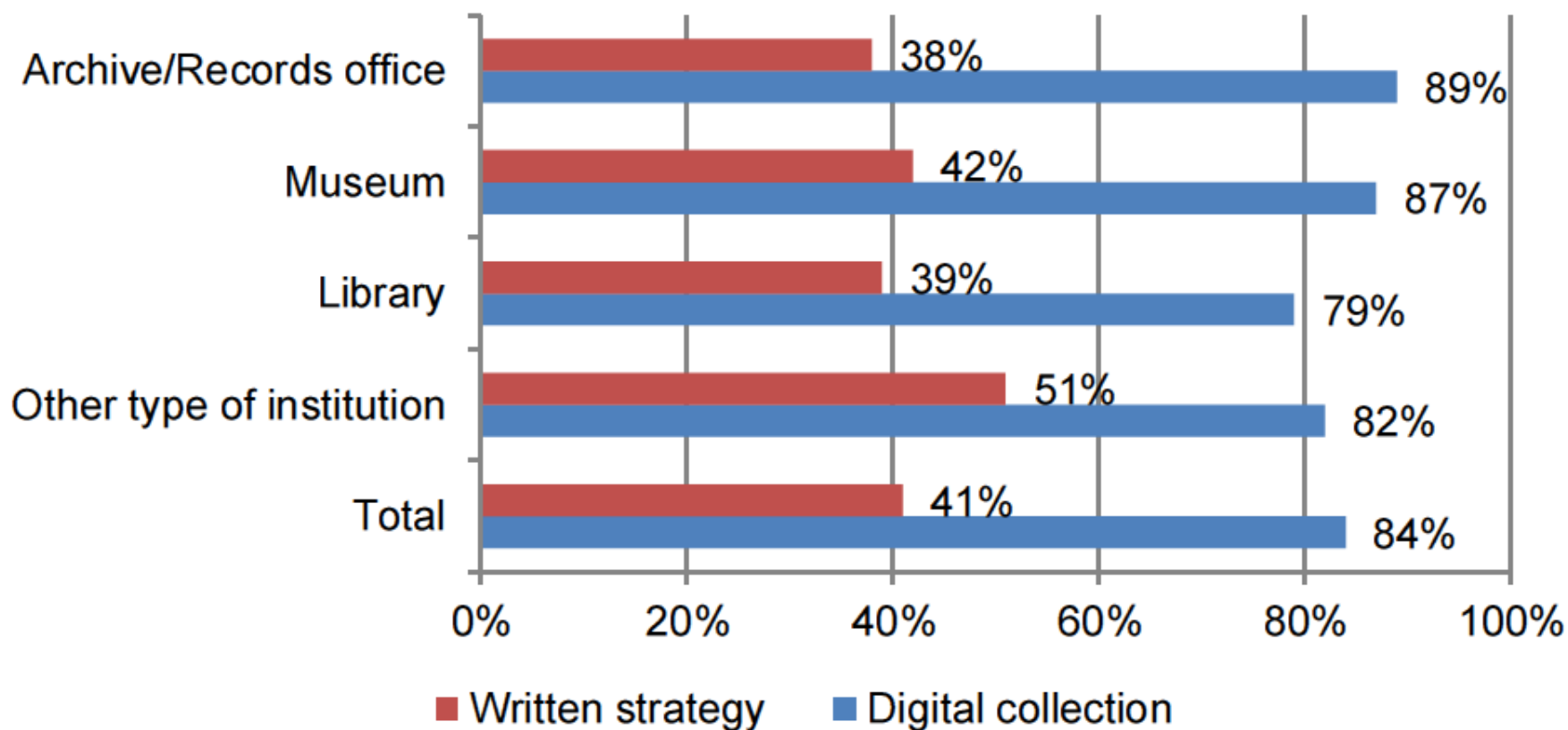**Figure 3.2:** Digital Collection (n=993) and Written digitisation strategy (n=835)
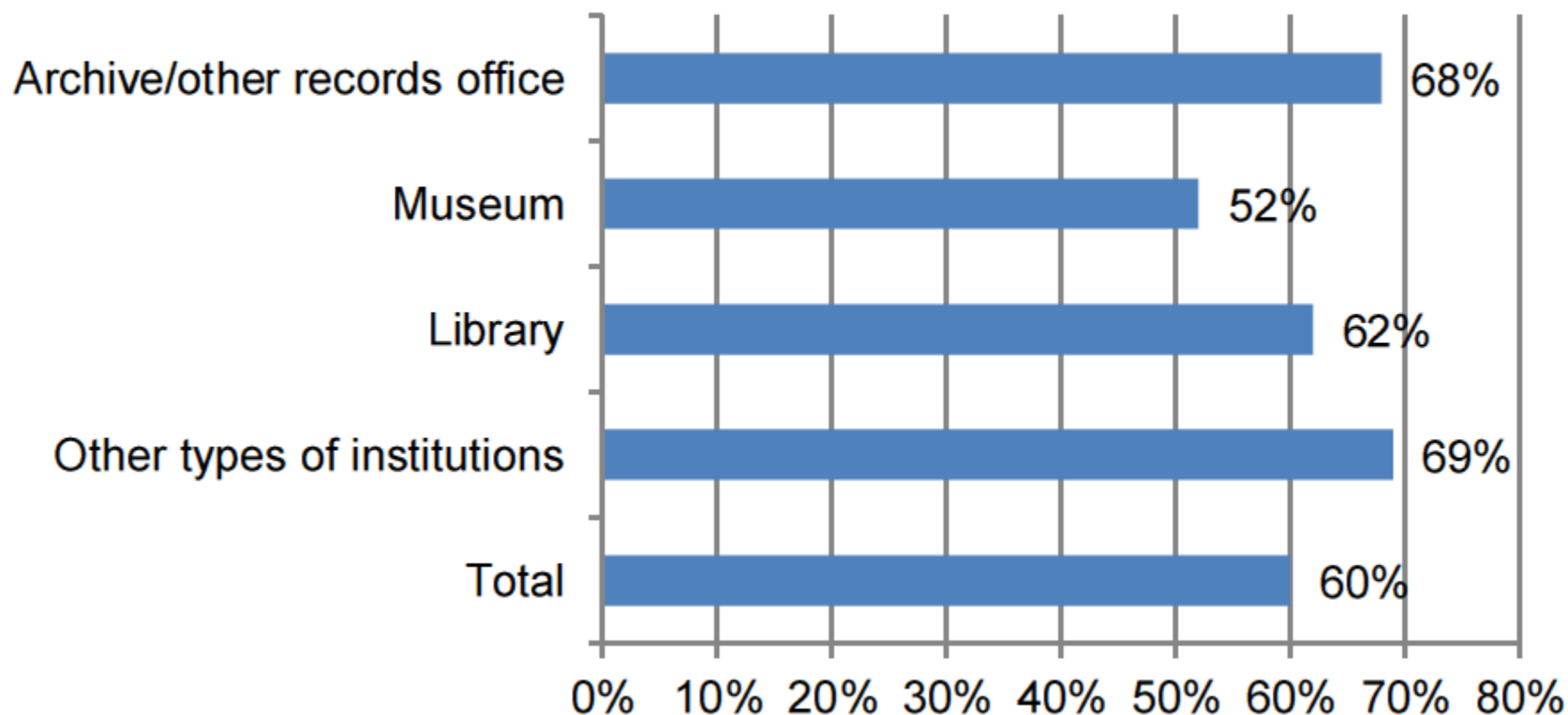
Figure 3.3.1: Does your organisation collect born digital heritage? (n=830)

Archive/other records office	68%
Museum	52%
Library	62%
Other types of institutions	69%
Total	60%

Figure 6.2: Estimation of the percentage of the total annual expenditures on creating/acquiring, maintaining, enhancing and preserving digital collections. Incidental costs compared to structural costs. (n=675)
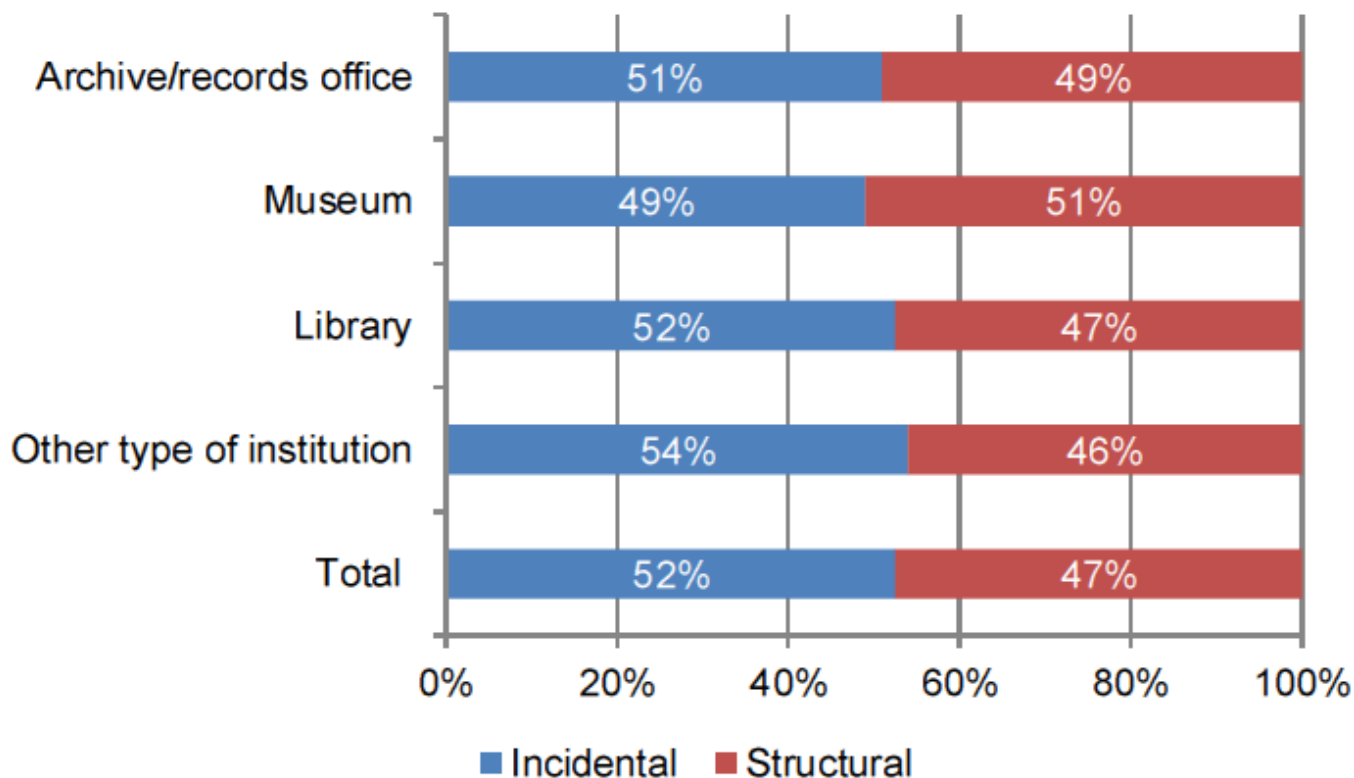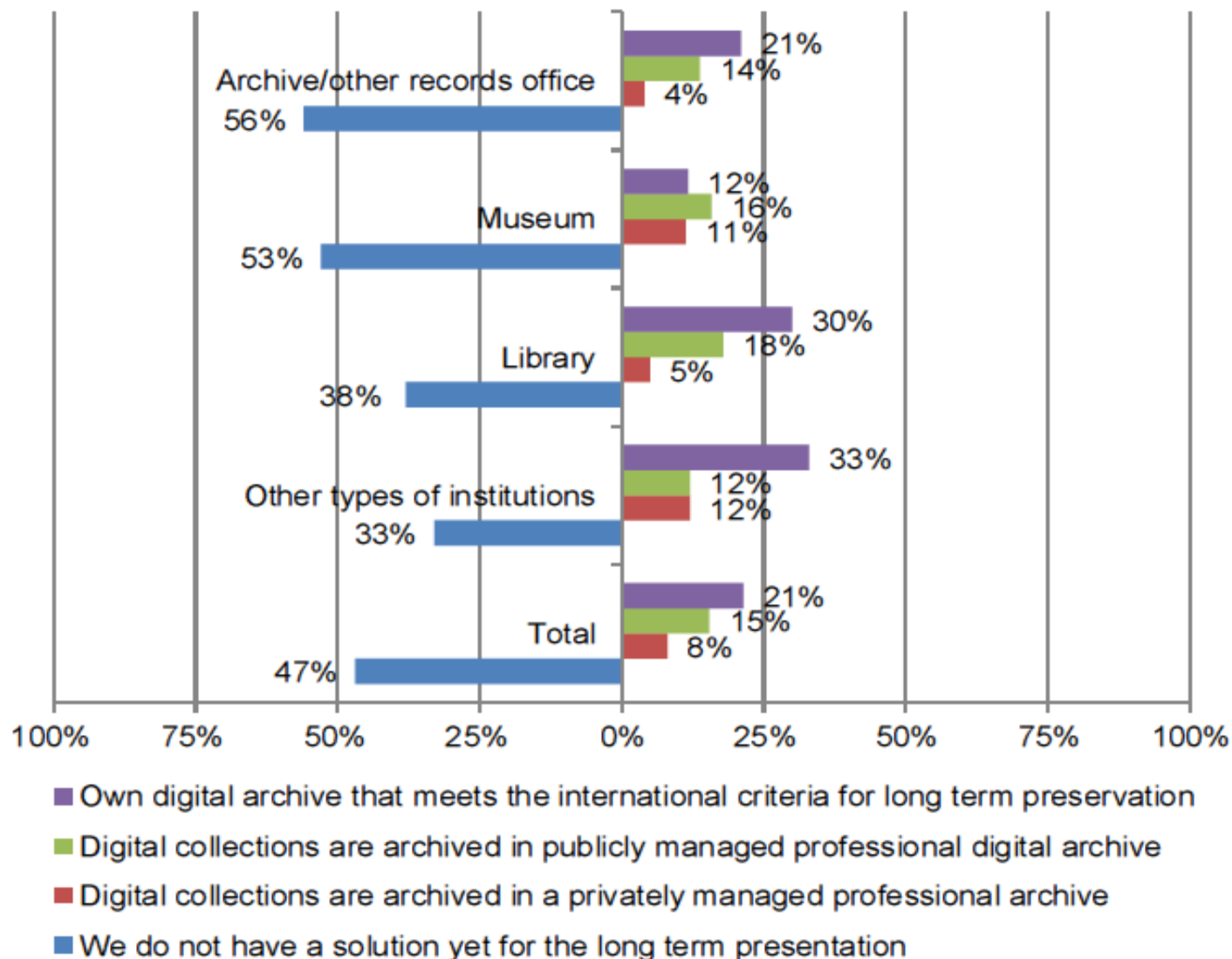
Figure 5.2.1.: Are (parts of) your digital collections stored in digital archives that have been set up according to international standards for digital preservation? (n=722)

- Own digital archive that meets the international criteria for long term preservation
- Digital collections are archived in publicly managed professional digital archive
- Digital collections are archived in a privately managed professional archive
- We do not have a solution yet for the long term presentation

Figure 5.1.1: Does your organisation have a written Digital Preservation Strategy, which is endorsed by the management of your organisation? (n=724)
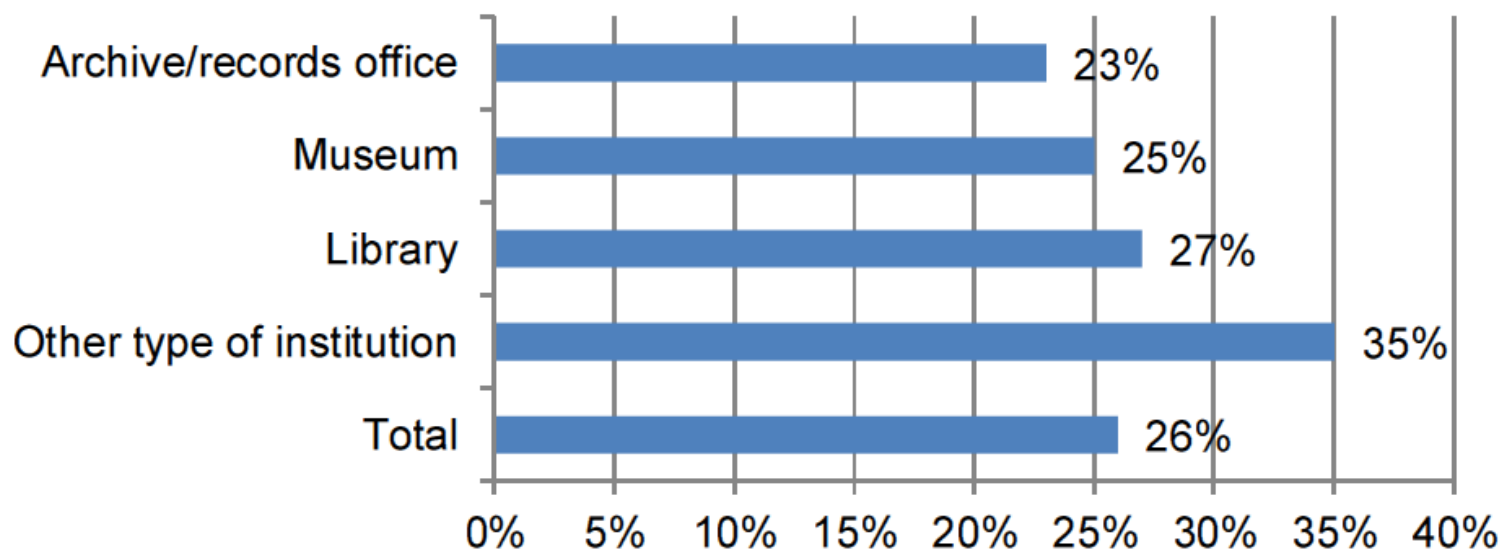
Conclusion 1:

After 20 years of digitisation the majority of cultural heritage institutions in Europe do not yet have strategies and operational solutions for digital curation.

The UNESCO/PERSIST

Guidelines for the selection of digital heritage for long-term preservation

By the UNESCO/PERSIST Content Task Force

March 2016

Authors:

Sarah CC Choy (Legislative Council Archives, Hong Kong)
Nicholas Crofts (ICOM CIDOC)
Robert Fisher (Library and Archives Canada)
Ngian Lek Choh (National Library Board of Singapore)
Susanne Nickel (Eskilstuna stadsmuseum)
Clément Oury (ISSN – International Centre)
Katarzyna Ślaska (National Library of Poland)

Chair of the Content Task Force: Ingrid Parent (University of British Columbia)
Secretary of the Content Task Force: Julia Brungs (International Federation of Library Associations and Institutions - IFLA)

Endorsed by:

Question:

a. Who thinks that public memory institutions should archive all information in the world (web, deep web, social media, research data, public sector data, institutional systems, intranets, private collections etc.)?

b. Who thinks that they should not archive any digital information at all for long term purposes?

c. If neither: give a estimate on a scale from 10 (= all information) to 0 (no information), how much of the digital information that is being created purposefully will end up in the digital archives of memory institutions for long term preservation purposes?

Top 20 most visited global websites (1 March 2017)

1. Google.com
2. Youtube.com
3. Facebook.com
4. Baidu.com
5. Yahoo.com
6. Wikipedia.org
7. Google.co.in
8. Qq.com
9. Sohu.com
10. Google.co.jp
11. Taobao.com
12. Tmall.com
13. Live.com
14. Amazon.com
15. VK.com
16. Twitter.com
17. Instagram.com
18. 360.cn
19. Sina.com.cn
20. LinkedIn.com

Source: Alexa

den Kenniscentrum Digitaal Erfgoed

Top 20 most visited global arts website (1 March 2017)

1. Youtube.com
2. Facebook.com
3. Twitter.com
4. IMDB.com
5. BBC.co.uk (global: 106)
6. Spotify.com
7. ESPN.com
8. CNN.com
9. Deviantart.com
10. BBC.com/news
11. Wikipedia.org
12. Foxnews.com
13. Livejournal.com
14. Scribd.com
15. Pandora.com
16. Xfinity.com
17. NDTV.com
18. AZlyrics.com
19. Hulu.com
20. Urbandictionary.com

Should the content of these cultural information sources be part of selection policies of public memory institutions?

Internet Archive: 282; Smithsonian: 6.502; Rijksmuseum: 78.529; Vatican library: 298.241

Source: Alexa

YOUTUBE SITE TRAFFIC

EVERY MONTH
1 BILLION unique visitors

6 BILLION hours of video watched

THAT'S NEARLY **1 HOUR** for every person on Earth.

EVERY DAY
3x INCREASE in users subscribing
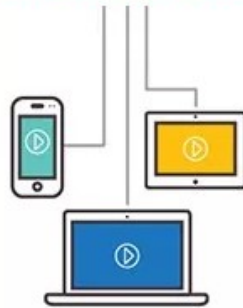
4x INCREASE in channel follows

4 BILLION videos are viewed worldwide

EVERY MINUTE
100 HOURS of video are uploaded to YouTube

Internet total storage capacity
2014 estimates:
1 million Exabytes

Youtube
2015 estimates: 1 Exabyte

Amazon
2014 estimates: 410 Tb

Library of Congress
2014 estimates: 75 Tb

Sound and Vision, Hilversum
2016 estimates: 15 Pb

Conclusion 2:

No single cultural heritage institution is able
to archive the largest cultural websites
on the internet.

Conclusion 3:

The often used narrative

*"If we (=CHI) don't act now,
digital information will get lost"*

does not suffice, as a large portion of
digital information will not be curated
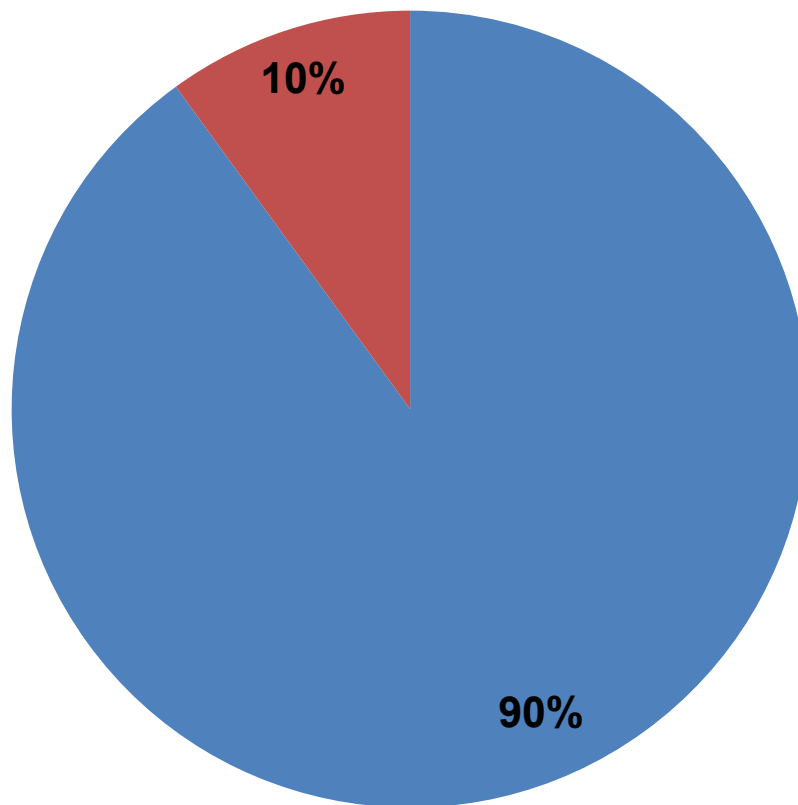for long term purposes.

We need new narratives!

2. Governments

**Memory institutions look at governments
to create legal frameworks
to provide them with a mandate
for digital preservation**
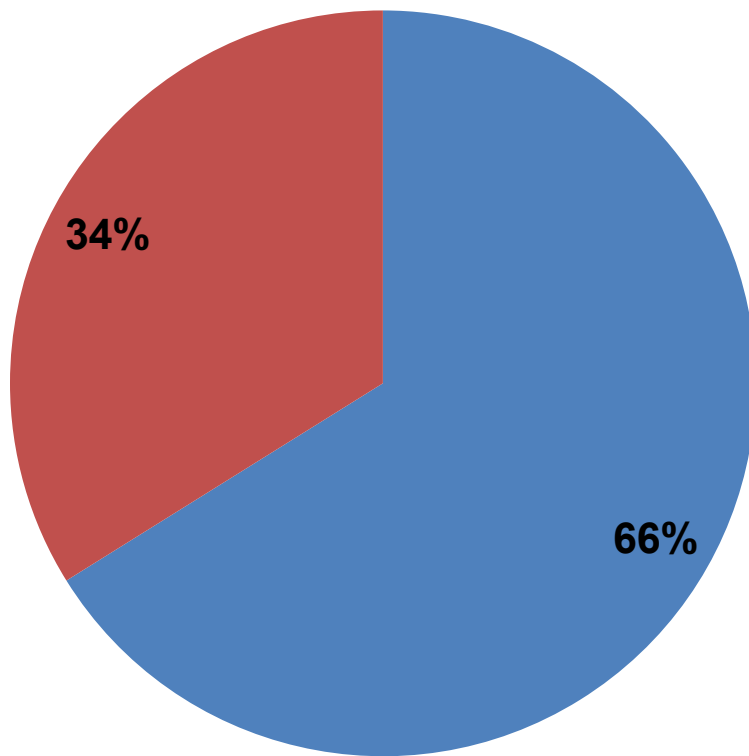
**Are digital policies also
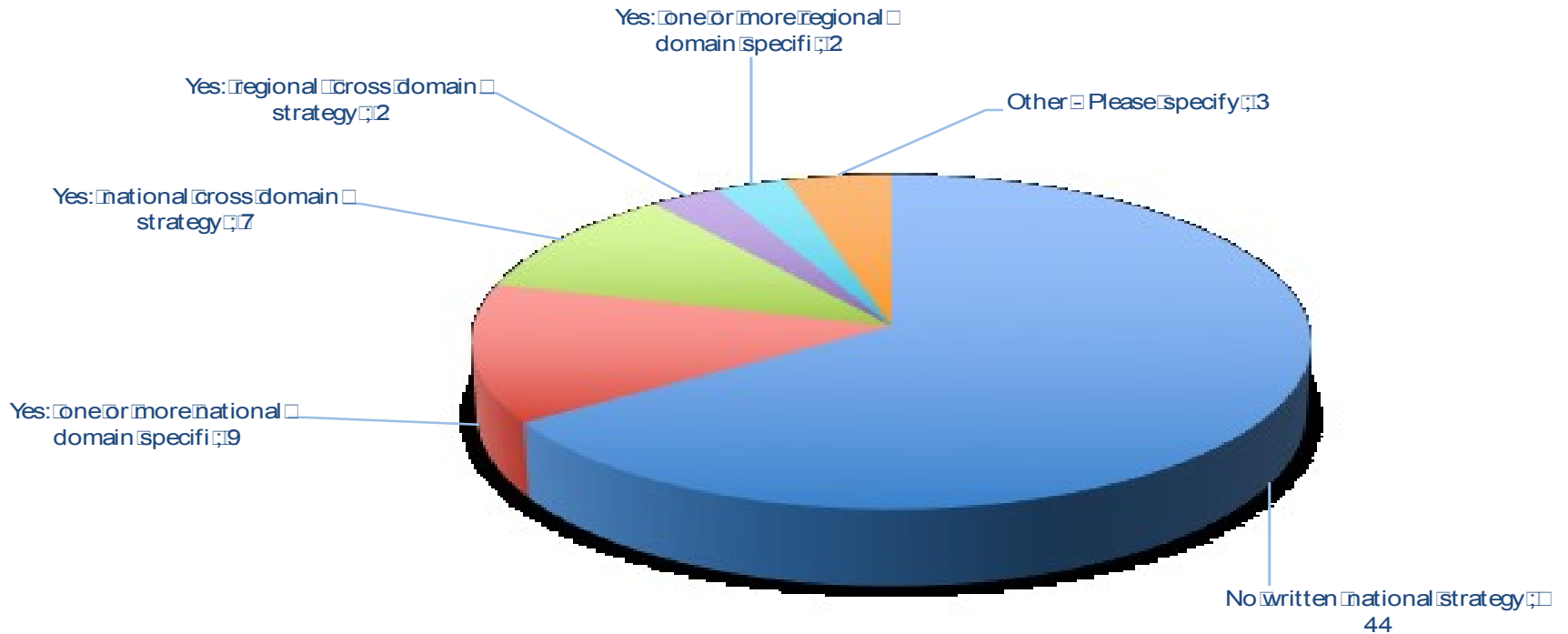a governmental task? (n= 57)**

■ Yes   ■ No

10%

90%

Is there in your country legislation on digital preservation? (n= 62)

■ Yes ■ No

34%

66%

Written strategy (N = 67)

- Yes: one or more regional domain specifi ; 2
- Yes: regional cross domain strategy ; 2
- Other - Please specify ; 3
- Yes: national cross domain strategy ; 7
- Yes: one or more national domain specifi ; 9
- No written national strategy ; 44

No written national strategy

Yes: one or more national domain specifi

Yes: national cross domain strategy

Yes: regional cross domain strategy

Yes: one or more regional domain specifi

Other - Please specify

PERSIST

45 respondents described the biggest challenges re. a national digital preservation strategy as:

- lack of leadership
- lack of knowledge
- lack of funds
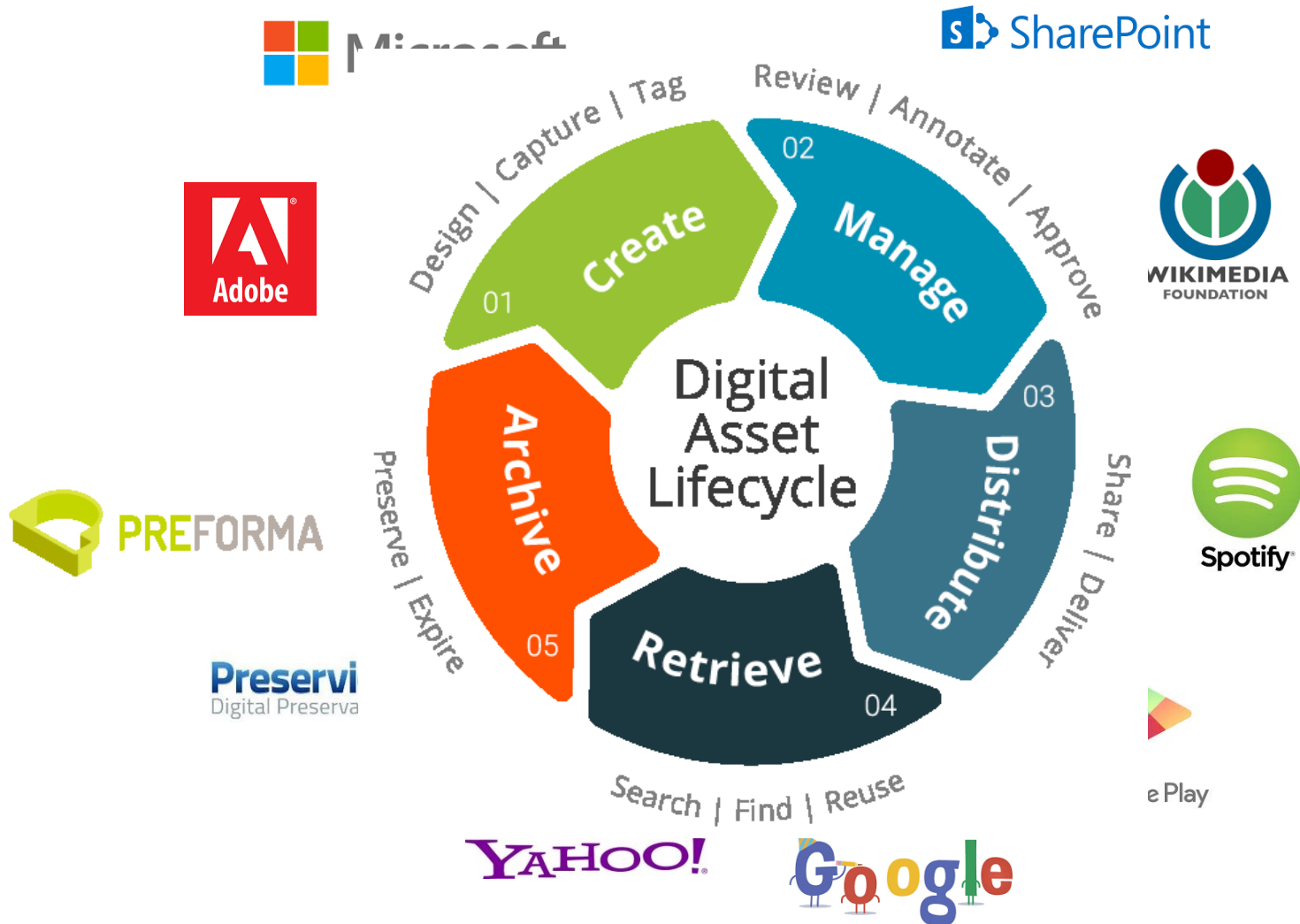- lack of consensus between domains/institutions

Conclusion 4:

Apart from the national legal frameworks
(e.g. archival law, library deposit law)
the majority of governments
do not yet actively support
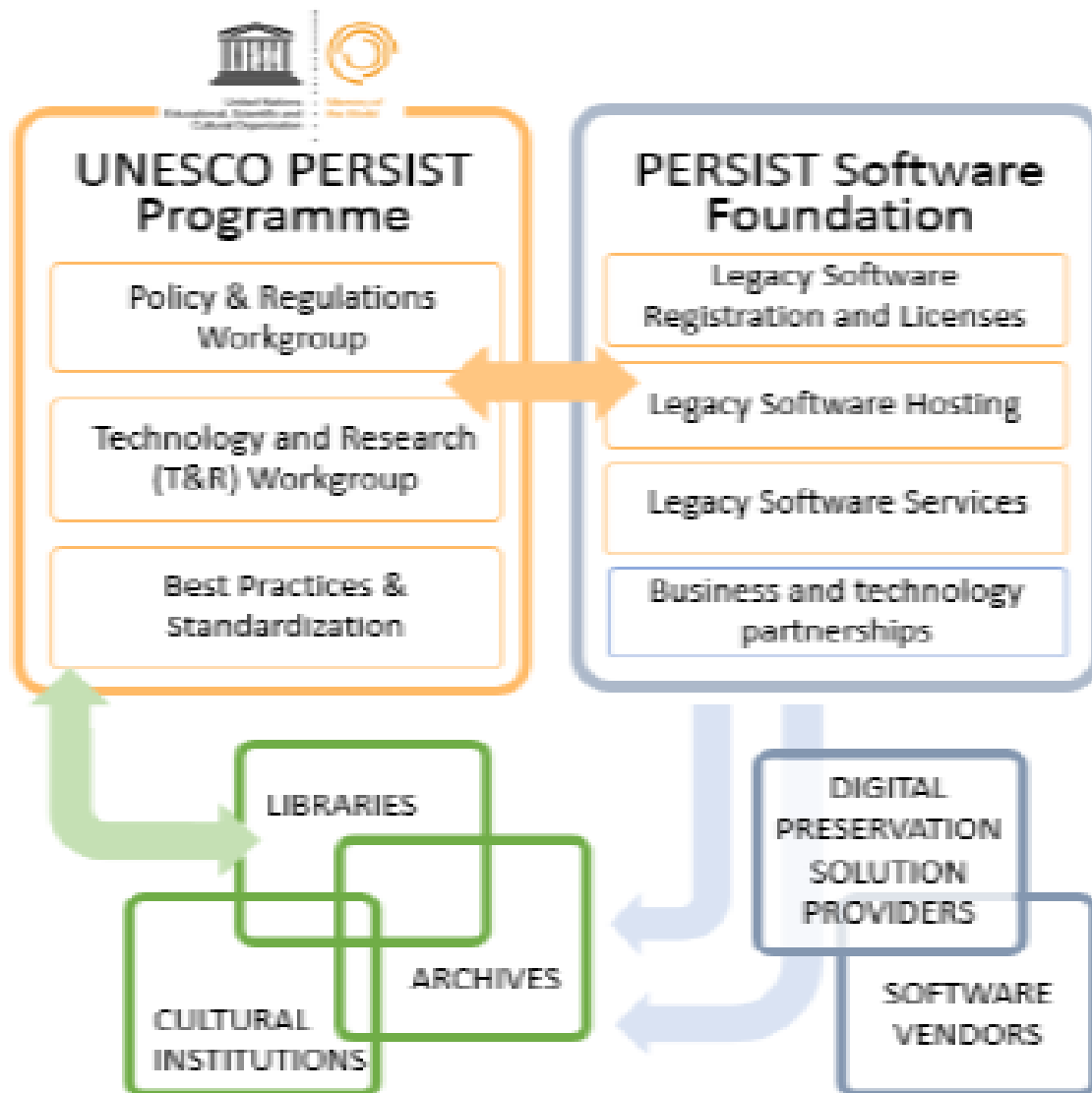digital preservation
at a crossdomain level.

3. ICT Industry

We need an holistic view on information management

3.1 Advocate for open preservation

| Strategy	Benefit	Open	Affiliate	Charter
Communicate the requirements and opportunities of open digital preservation approaches and solutions to policy makers, funders, and other strategic bodies	Request responses to consultations Request external communications			✓ ✓

Source: OPF Strategic Plan 2015-2018

den Kenniscentrum Digitaal Erfgoed

DCH-RP

A Roadmap for Preservation of Digital Cultural Heritage Content
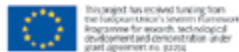
Take further actions in identified major areas of the roadmap

Harmonise data storage and preservation

Consolidate mature requirements for preservations in the DCH environment

Improve interoperability

Implement tools, developed and tested in earlier stages, in selected e-Infrastructures facilitating interoperability aspects in the technical and semantic field.

Establish conditions for cross-sector integrations

Fill in remining gaps in cross-sector integration according to needs identified in earlier stages.

Governance models for infrastructure integration

Offer mature business model for preservation services for different types of institutional settings

Horizon 2020 Draft Work Programme 2018 - 2020 Future and Emerging Technologies

FETFLAG-01-2018: Preparatory Actions for new FET Flagships
This topic aims at launching coordination and support actions (CSA) to prepare new candidate FET Flagships (FET Flagships are science- and technology-driven, large-scale, multidisciplinary research initiatives built around a visionary unifying goal)

(1) ICT and Connected Society / ICT for Social Sciences and
      Humanities:
New ICT technologies and approaches for the understanding of large-scale complex socio-technical systems and their interaction and interdependencies or for collecting, <u>preserving</u>, studying and promoting Europe's unique cultural heritage and exploiting these to achieve major societal or economic benefits.

Conclusion 5:

It is really difficult to get the major,
most influential ICT companies at the table.

There is a good vibe in the area of open source and
connecting to research infrastructures.

However, we do need to engage ourselves constantly
with the ICT-industry/ICT-community
(e.g. through their user groups) as in the end,
DP is also a matter of sustainable business.

To summarize (1):

We need stronger narratives to broaden our impact and make DP a societal issue (incl. business)

Narrative 1: Selection policies
We need to make it more clearer what type of content (global/national/local) we intend to curate for future use in order to explain our position in the (highly commercialized) information society.

Narrative 2: Added value
We need to make it clearer what our added value is with the information we curate for long term purposes:
- Legal frameworks (democracy, transparancy)
- Securing conditions for access (copyright, human rights)
- Safeguarding authenticity when it is crucial ('as intended')
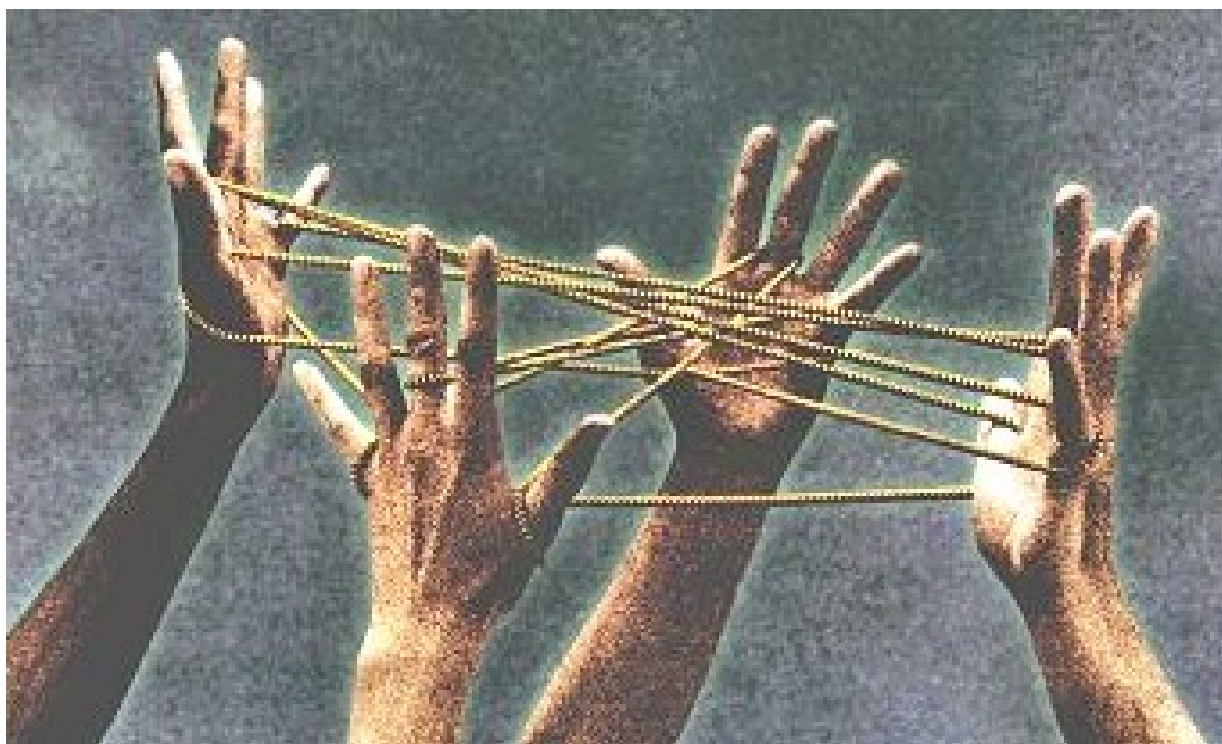- Promoting and improving standardisation

To summarize (2):

We should more explicitly address our internal contraints (e.g. still dominance of an analog mindset; lack of business attitude to achieve economies of scale; missions that are mostly framed in a local and/or national context).

The institutional paradigm will only suffice for local content. The national paradigm works mostly in a legal context. For all other content curation we need stronger international alliances.

Share your knowledge on digital culture!

www.den.nl