Time is the Thief of Memory:

book and journal publishing,
open access, and
long-term digital preservation

Alicia Wise,
Executive Director

CLOCKSS

@wisealic
awise@clockss.org

Changing World

There's been a dramatic transformation in scholarly communications in the recent past.

And we must be more mindful of energy consumption and our environmental impact.

Disruptions happen

Knowledge that takes generations to build can be lost in an instant or erode gradually and out of sight.

Unforeseen events, catastrophic accidents, and stewardship gaps can erase access to both traditional and digital publications.

CLOCKSS in numbers

49 million journal articles and **436 thousand books** - XML, video, software, images, audio, excel, LaTeX, powerpoint, HTML, PDFs, RTF, and much more

Published by **440 publishers** in **48 countries**

Supported by **300 libraries** in **14 countries**

Content is stored in **12 nodes** hosted by **major world universities,** including Edinburgh University and Humboldt University in Europe

64 triggered journals, accessed **334,945** times in the last 12 months

300 SUPPORTING LIBRARIES

435 PARTICIPATING PUBLISHERS

12 SELF-HEALING NODES

436,240 BOOKS
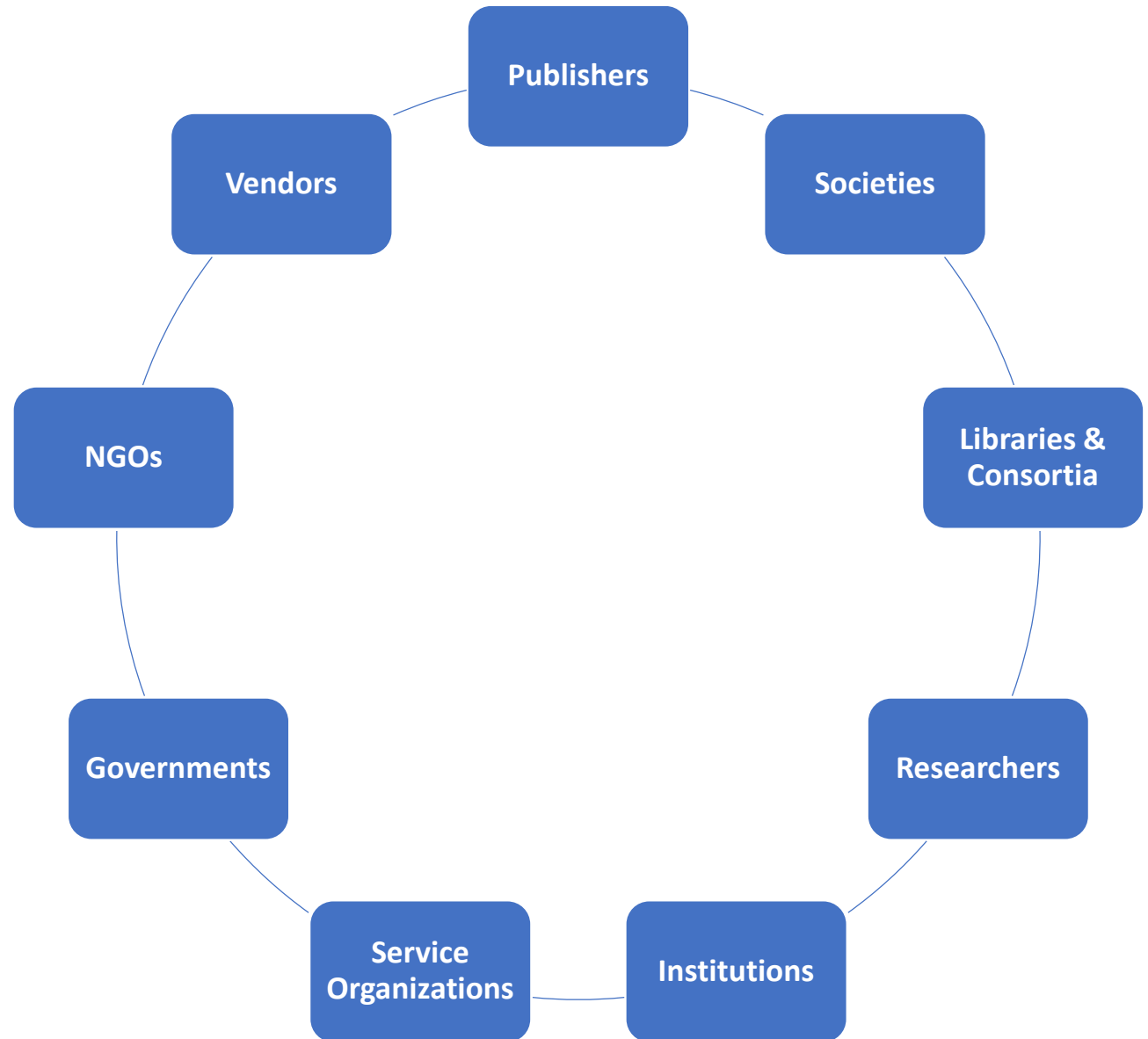
49 MILLION JOURNAL ARTICLES

GLOBAL SUCCESSION PLAN

Building a sustainable future

Safeguarding scholarly publications – traditionally a library function – is a different job than it once was – and it is continuing to evolve.

Dynamic solutions require rethinking relevant issues, looking beyond national frameworks, and forging new roles and relationships in the scholarly communications ecosystem to achieve more than any stakeholders could individually.

Partners

https://doaj.org/preservation/

JASPER:
JOURNALS
PRESERVED

- Open access journals that do not charge APCs and that are registered with DOAJ and have no preservation policy, i.e. **7,000** titles

- Three archiving options for project (and more to come!)

JASPER: JOURNALS PRESERVED

Titles using Open Journal Systems (OJS), the open source publishing tool developed and distributed freely by PKP

PKP PN

PKP

PUBLIC
KNOWLEDGE
PROJECT

JASPER: JOURNALS PRESERVED

Titles registered in DOAJ and providing content with detailed metadata

CLOCKSS

JASPER: JOURNALS PRESERVED

Titles registered in DOAJ but with no metadata

INTERNET ARCHIVE

JASPER: JOURNALS PRESERVED

Reporting about the archiving progress is publicly accessible

Keepers Registry

https://keepers.issn.org/

Journals archived include scholar-led journals from Croatia, Finland, Ireland, Poland & Sri Lanka:

Open Access Books are even more complicated!

PEERING INTO THE JUNGLE

Challenges in determining preservation status of open access books

Mikael Laakso

Hanken School of Economics
Helsinki, Finland
mikael.laakso@hanken.fi
0000-0003-3951-7990

Alicia Wise

CLOCKSS
California, USA
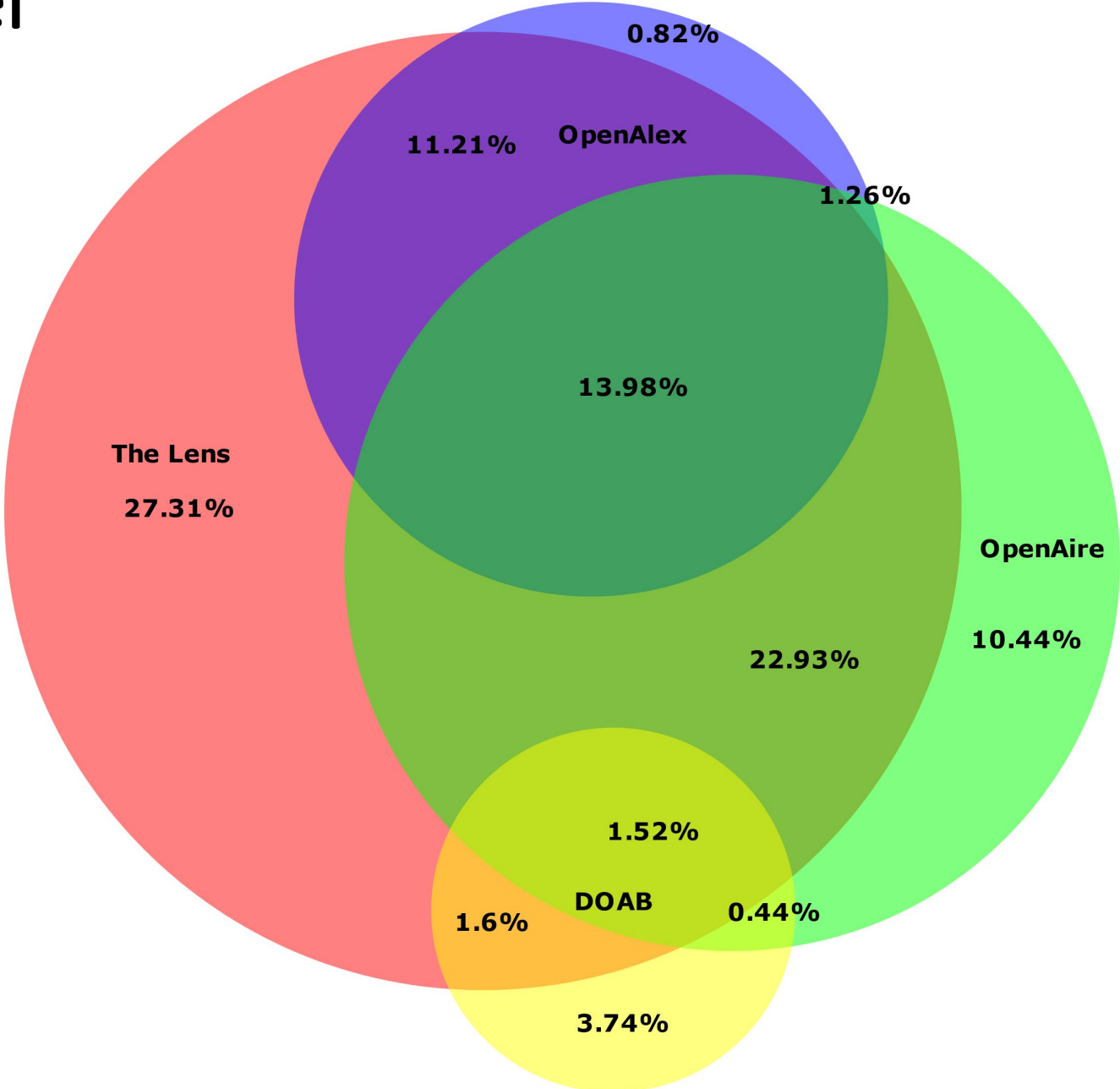awise@clockss.org
0000-0002-7898-3428

Ronald Snijder

OAPEN Foundation
The Hague, The Netherlands
r.snijder@oapen.org
0000-0001-9260-4941

Content distribution over major data sources

Challenges include

- *Definitions -* When is a book an academic book, and when is it open access?
  - Different indexing services have different levels of data quality, despite narrowing down searches as much as possible still a lot of "noise" in the data.
  - Interpretation/tagging of Open Access status also varies across services

- *Data management* - Physical extraction of metadata to represent the "global bookshelf" of academic OA books
  - Putting together the dataset a varied mix of REST APIs, JSON dumps, CSV files.
  - Many datasets needlessly large for investigations of this kind, hopefully more varied ways to access data like this in the future.

- *Unique identifiers -* Taming the wilderness of identifier metadata describing OA books
  - A single book can be assigned an ISBN (or nine ISBNs), a DOI, both or none.
  - Preservation services still dominantly ISBN-based at least when it comes to public book preservation data, an expansion into also including DOIs would for many purposes be beneficial.

Preliminary results

Existing coverage of content in international preservation services (match either by ISBN or exact book title match)

	DOAB	WorldCat	The Lens	OpenAlex	OpenAire	Scielo
CLOCKSS	22,5 %	6,6 %	3,3 %	4,2 %	8,1 %	0,0 %
Portico	31,2 %	31,3 %	8,6 %	11,3 %	22,3 %	8,7 %
LOCKSS	22,5 %	1,2 %	0,1 %	0,1 %	0,2 %	0,0 %
Livro Aberto	0,0 %	0,0 %	0,0 %	0,0 %	0,0 %	0,0 %
Found in at least one of the above preservation services	**45,7 %**	**32,6 %**	**9,7 %**	**12,5 %**	**24,7 %**	**8,7 %**

Are our digital collections safe for the long term?

Working with one digital preservation service is not enough to weather future uncertainties.

"…in the real world, content preserved at three independent services is probably safe against anything short or a major Coronal Mass Ejection… if they happen, society is likely to have bigger problems than rereading old academic papers."

-David S.H. Rosenthal, "Where Did the Number 3 Come From?" DSHR blog, 14 June 2022

The scholarly record is vulnerable, but damage doesn't have to be irrevocable

Publishers have an important role to play by ensuring they preserve materials with certified repositories (at least one and ideally three or more)

Journals, books, conference proceedings, databases, videos, software, metadata, and other content and materials that are preserved in dark archives can be retrieved intact to inform future generations of researchers.

CLOCKSS

https://www.clockss.org

Working Together to Safeguard Knowledge

Can we count you in?

Supporting Slides

Community Owned and Governed

Scholarly publishers and research libraries share governance, strategy, policymaking, costs, and skills for community-governed and supported CLOCKSS.

When you join CLOCKSS, you become part of a global community dedicated to developing and sharing preservation solutions and services.

1

Global

Our twelve nodes are located at major research institutions with expertise in digital preservation to provide exceptional redundancy and resilience. Your authors' works will endure even if CLOCKSS and/or your organization doesn't.

2

Stanford University

HUMBOLDT-UNIVERSITÄT ZU BERLIN

NII Inter-University Research Institute Corporation / Research Organization of Information and Systems
National Institute of Informatics

Australian National University

RICE UNIVERSITY

UNIVERSITAS CATTOLICA SACRI CORDIS JESU
UNIVERSITÀ CATTOLICA del Sacro Cuore

OCLC

IU

VIRGINIA

UNIVERSITY OF ALBERTA

THE UNIVERSITY of EDINBURGH

3

Dependable

Designed by David S. H. Rosenthal and colleagues at Stanford University, CLOCKSS holds the **highest certification score ever recorded** to a preservation service by the **Center for Research Libraries**. It keeps up the technology and tools needed to transfer and use the archived content and ensure that the metadata is healthy.

CRL Center *for* Research Libraries
....
GLOBAL RESOURCES NETWORK

Innovative

The world of research is changing, and so are we – to ensure that a broader spectrum of knowledge **remains discoverable, retrievable, accessible, and usable across systems and time**. CLOCKSS is evolving new features and services.

4

JASPER: JOURNALS PRESERVED

THE ANDREW W. **MELLON** FOUNDATION

5

Collaborative

Our dedicated team is **friendly and easy to work with**.

Global LOCKSS Network

- Post-cancellation access
- Local/regional preservation
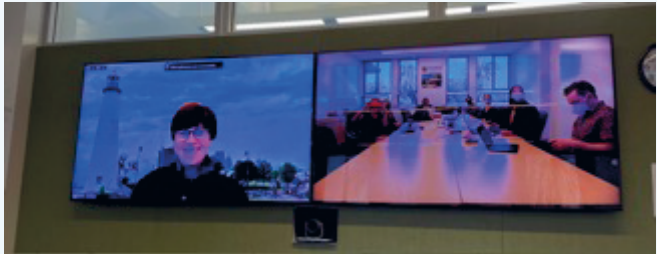

LOCKSS
Software

5

Collaborative

CLOCKSS

- Long-term preservation of licensed academic content
- Global
- In partnership with publishers

Local & Regional Preservation Services

- Alabama DPN
- Cariniana
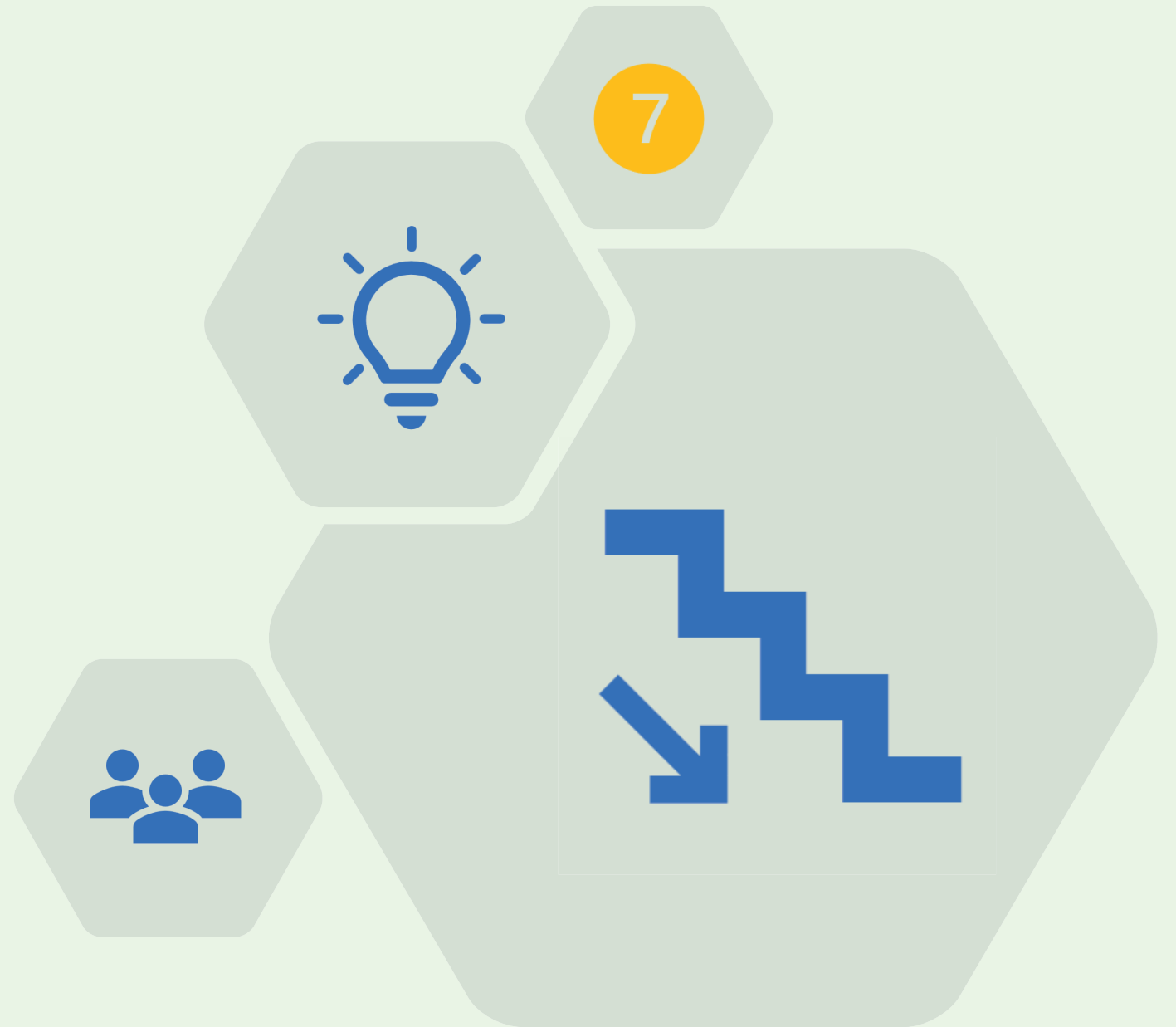- Canadian Govt Info
- COPPUL
- PKP-PN

6

Open & Transparent

CLOCKSS is built on highly secure yet open technologies, standards, and protocols, and freely shares information about the technology it uses.

A unique feature is that **after publications are triggered, they are openly accessible to everyone** under Creative Commons licenses.

Affordable

CLOCKSS is mindful: the money we steward belongs to libraries and publishers. We are efficient and **keep our costs down while providing a best-in-class service**.

CLOCKSS has the highest accreditation score of any digital preservation service ever awarded by the Center for Research Libraries. We are not-for-profit and a joint venture between academic publishers and research libraries. Our mission is to safeguard knowledge.

300 SUPPORTING LIBRARIES

435 PARTICIPATING PUBLISHERS

12 SELF-HEALING NODES

436,240 BOOKS

49 MILLION JOURNAL ARTICLES

GLOBAL SUCCESSION PLAN

Seven reasons why CLOCKSS stands out among long-term digital preservation archives

1. Community-owned

2. Global

3. Dependable

4. Innovative

5. Collaborative

6. Open & Transparent

7. Affordable

CLOCKSS