A Bibliographic Scan of Digital Scholarly Communications Infrastructure

David W. Lewis

Coalition for Networked Information, Fall Taskforce Meeting
Washington, D.C.
December 9, 2019

PDF of slides at: https://tinyurl.com/**y3cwmte5**

© 2019 David W. Lewis. This work is licensed under a Creative Commons Attribution-NonCommercial 4.0 International License.

Agenda

1. Bibliographic scan or “map” of digital scholarly communications infrastructure
2. How much should open scholarly communications infrastructure cost

What was the Bibliographic Scan?

Part of the Mapping the Digital Scholarly Communication Infrastructure project funded by the Andrew W. Mellon Foundation.

Literature review attempted to identify projects and products that make of digital infrastructure that supports scholarly communications. What we called “Scholarly Communication Resources” or SCRs.

- for-profit and not-for-profit
- open and proprietary

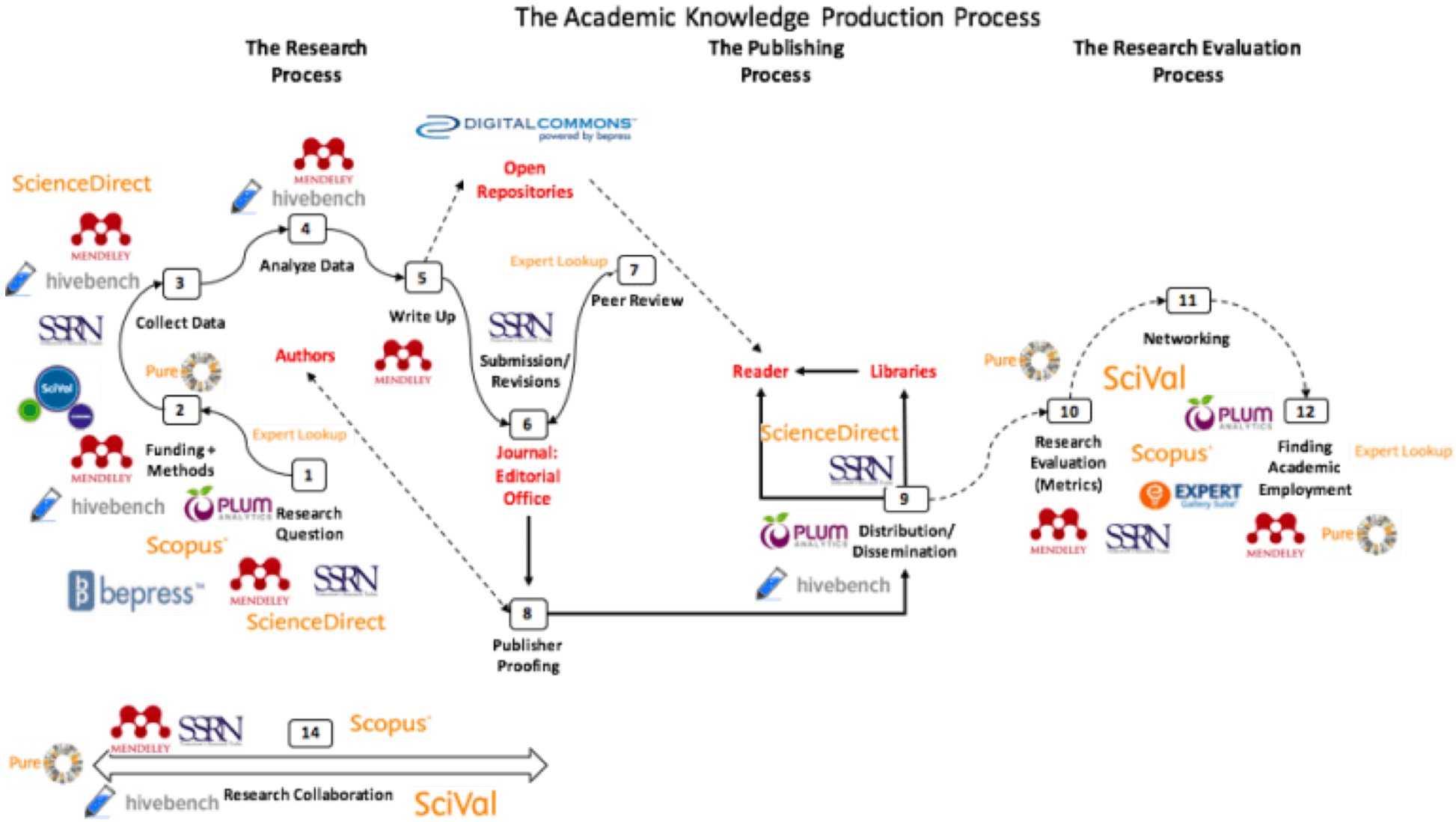
What are SCRs?

SCRs are defined as organized groups building, offering, or using tools, platforms, and services in ways that enable active engagement and participation in the scholarly communication and publishing process, broadly defined.

Content is not included.

What are SCRs?

- OJS is a platform used to produce and publish content;
- HathiTrust is a hybrid that includes both a platform through which content can be stored, accessed, and preserved, and also content (objects);
- Directory of Open Access Journals (DOAJ) is a service that indexes Open Access Journals that meet specific criteria (infrastructure);
- JSTOR is a platform that provides content that already exists and is used to conduct research (content).



Alejandro Posada and George Chen, "Inequality in Knowledge Production: The Integration of Academic Infrastructure by Big Publishers," ELPUB 2018, Jun 2018, Toronto, Canada. 10.4000/proceed- ings.elpub.2018.30, hal-01816707, https://hal.archives-ouvertes.fr/hal-01816707/document

General Services (14)

Researcher Tools
- Reading (3)
- Writing (4)
- Annotation (2)
- Reference Managers (5)
- Collaboration (5)

199 projects
65 profit
134 non-profit

Discovery (32)

Repositories (18)
- Data (2)
- Tools (10)
- Directories (3)
- Service Providers (3)
- Subject (7)

Publishing (8)
- Data (3)
- Journals (21)
- Monographs (10)
- Meetings (2)
- Peer Review (3)
- Production (13)
- Math Notation (4)

Preservation (14)

Assessment (8)
- CRIS (5)

General Services

Project	Organization	NP or P
Common Workflow Language		**NP**
Creative Commons		**NP**
Crossref		**NP**
DataCite		**NP**
GetFTR (Get Full Text Resource)		**NP**
Handle System		**NP**
Initiative for Open Citations (I4OC)		**NP**
International Image Interoperability Framework (IIIF)		**NP**
Open Science Framewrk	**Center for Open Science**	**NP**
ORCID		**NP**
Research Organization Registry		**NP**
SeamlessAccess.org (RA21)		**NP**
Sherpa/Romeo		**NP**
RightLink Author	**Copyright Clearance Center**	**P**

Researcher Tools

Project	Organization	Function	NP or P
Readium		Reader	NP
Rebus Ink		Reader	NP
Scholarcy		Reader	P
Jupyter Notebooks		Writing	NP
Stencila	Code for Science & Society	Writing	NP
Wax	Collaborative Knowledge Foundation (Coko)	Writing	NP
Authorea		Writing	P
Hypothes.is		Annotation	NP
PaperHive		Annotation	P
Zotero		Reference Managers	NP
Citationsy		Reference Managers	P
EndNote	Clarivate Analytics	Reference Managers	P
ReadCube		Reference Managers	P
RefWorks	ProQuest	Reference Managers	P
Humanities Commons	MLA	Collaboration	NP
Academia.edu		Collaboration	P
Mendeley	Elsevier	Collaboration	P
ResearchGate		Collaboration	P
Synapse Platform	Sage	Collaboration	P

Discovery NP

Project	Organization	NP or P
Blacklight		NP
Directory of Open Access Journals (DOAJ)		NP
Enhanced Network Monographs		NP
Get the Research	ImpactStory	NP
Grobid		NP
Meta Science	Chan Zuckerberg Initiative	NP
OAIster	OCLC	NP
Open Access Button		NP
Open Knowledge Maps		NP
OpenCitations		NP
Research Organization Register		NP
ROAD: the Directory of Open Access Scholarly Resources	ISSN International Center	NP
SHARE	Center for Open Science (COS)	NP
Unpaywall	ImpactStoryc	NP
Virtual International Authority File (VIAF)	OCLC	NP
WorldCat	OCLC	NP

Discovery Profit

Project	Organization	NP or P
1findr	1Science/Elsevier	P
DeepDyve		P
Dimensions	Digital Science	P
Google Scholar	Google	P
IRIS.AI		P
Kopernio	Clarivate Analytics	P
Morressier		P
Open Research Library	Knowledge Unlatched	P
Paper Digest		P
SciCrunch		P
Scopus	Elsevier	P
SemanticScholar		P
Web of Science	Clarivate Analytics	P
Yewno Discover		P

Repositories

Project	Organization	Function	NP or P
Avalon Media System		Repository	NP
CONTENTdm	OCLC	Repository	NP
DSpace	DuraSpace/Lyrasis	Repository	NP
EPrints		Repository	NP
Fedore Repository	DuraSpace/Lyrasis	Repository	NP
Hyku (Hydra in a Box)		Repository	NP
Islandora		Repository	NP
Mukurtu		Repository	NP
Omeka		Repository	NP
OSF Preprints	Center for Open Science	Repository	NP
Samvera (formerly Hydra)		Repository	NP
TIND Indtitutional Repository		Repository	NP
Zenodo	OpenAire	Repository	NP
Digital Commons	Bepress/Elsevier	Repository	P
Figshare	Digital Science	Repository	P
Slideshare	Linkedin	Repository	P
Vimeo		Repository	P
YouTube		Repository	P
Dryad Digital Repository		Data	NP
TIND Data Management		Data	NP

Repository Tools, Etc.

Project	Organization	Function	NP or P
OpenDOAR Directory of Open Access Repositories		Directory	NP
Registry of Open Access Repositories (ROAR)		Directory	NP
ROARMAP		Directory	NP
4Science		Service Provider	P
Atmire		Service Provider	P
discoverygarden		Service Provider	P
OpenAire Broker	OpenAire	Tool	NP
OpenAIRE Content Provider Dashboard	OpenAire	Tool	NP
OpenAire Explore Portal	OpenAire	Tool	NP
OpenAIRE Mining Service	OpenAire	Tool	NP
OpenAIRE ScholeXplorer	OpenAire	Tool	NP
OpenAIRE ScholeXplorer	OpenAire	Tool	NP
OpenAire Usage Statistics	OpenAire	Tool	NP
Permissions Checker	Open Access Button	Tool	NP
ShareYourPaper.org	Open Access Button	Tool	NP
1foldr	1Science/Elsevier	Tool	P

Repositories
Subject

Project	Organization	Function	NP or P
ArXiv		Repository — Subject	NP
ASAPbio		Repository — Subject	NP
Earth and Space Science Open Archive (ESSOAr)		Repository — Subject	NP
HathiTrust		Repository — Subject	NP
Open Humans		Repository — Subject	NP
Protocols.io		Repository — Subject	P
SSNR	Elsevier	Repository — Subject	P
Sci-Hub		Repository — Subject	?

Publishing

Project	Organization	Function	NP or P
OpenEdition		Publishing	NP
Prose Mirror		Publishing	NP
PubPub		Publishing	NP
PubSweet	Collaborative Knowledge Foundation (Coko)	Publishing	NP
Scalar		Publishing	NP
Knowledgr		Publishing	P
Overleaf	Digital Science	Publishing	P
River Valley Technologies		Publishing	P
PREreview		Peer Review	NP
Publons	Clarivate Analytics	Peer Review	P
Data2paper	JISC	Data	NP
Dataverse		Data	NP
Shiny		Data	NP

Publishing Journals

Project	Organization	NP or P
DEIP		NP
Dokieli		NP
Janeway		NP
Libero		NP
Lodel		NP
Open Journal Systems (OJS)	Public Knowledge Project (PKP)	NP
Phenom Screener		NP
Publications Router (JISC)	JISC	NP
Texture		NP
Transpose		NP
AMS (Journal & Article Management System)	MDPI	P
F1000 (Faculty of 1000)	Science Navigation Group	P
JAMS (Journal & Article Management System)	MDPI	P
Literatum	Atypon	P
Phenom Reviewer		P
ScholarOne		P
Scholastica		P
ScienceOpen		P
SilverChair		P
Ubiquity Press		P
UNSILO		P

Publishing Monographs and Meetings

Project	Organization	Function	NP or P
Editoria	Collaborative Knowledge Foundation (Coko)	Monographs	NP
Fulcrum		Monographs	NP
HIRMEOS Project		Monographs	NP
Manifold Scholarship		Monographs	NP
Open Monograph Press	Public Knowledge Project (PKP)	Monographs	NP
Public Access Submission System (PASS)		Monographs	NP
Quire		Monographs	NP
Vega Publishing		Monographs	NP
Pressbooks		Monographs	P
Rua		Monographs	P
OSF Meeitng	Center for Open Science	Meetings	NP
SciForum	MDPI	Meetings	P

Publishing Other

Project	Organization	Function	NP or P
Electric Book		Production	NP
epub.js		Production	NP
Fidus Writer		Production	NP
Hy-phen		Production	NP
Hyphenopoly		Production	NP
Hyphr		Production	NP
INK	Collaborative Knowledge Foundation (Coko)	Production	NP
Lens		Production	NP
Open Typesetting Stack	Public Knowledge Project (PKP)	Production	NP
paged.js		Production	NP
Pandoc		Production	NP
Vivliostyle		Production	NP
xPub	Collaborative Knowledge Foundation (Coko)	Production	NP
xSweet	Collaborative Knowledge Foundation (Coko)	Production	NP
KaTeX		Math Notation	NP
le-tex Transpect		Math Notation	NP
MathJax		Math Notation	NP
Tectonic		Math Notation	NP

Preservation

Project	Organization	Function	NP or P
Academic Preservation Trust (APTrust)		Preservation	NP
Archive-It	Internet Archive	Preservation	NP
BitCurator Consortium	Educopia Institute	Preservation	NP
Chronopolis		Preservation	NP
CLOCKSS		Preservation	NP
Digital Preservation Network		Preservation	NP
DuraCloud	DuraSpace/Lyrasis	Preservation	NP
LOCKSS		Preservation	NP
MetaArchive Cooperative	Educopia Institue	Preservation	NP
Portico	Ithaka	Preservation	NP
Software Preservation Network		Preservation	NP
Archivematica	Artefactual Systems	Preservation	P
AtoM (Access to Memory)	Artefactual Systems	Preservation	P

Assessmnet

Project	Organization	Function	NP or P
COUNTER		Asessment	NP
Depsy	Our Research	Assessment	NP
ImpactStory Profiles	Our Research	Assessment	NP
Paperbuzz	Our Research	Assessment	NP
Retraction Watch		Assessment	NP
Altmetric	Digital Science	Assessment	P
PlumAnalytics	Elsevier	Assessment	P
RedLink		Assessment	P
DSpace-CRIS		Assessment — CRIS	NP
VIVO	DuraSpace/Lyrasis	Assessment — CRIS	NP
Converis	Clarivate Analytics	Assessment — CRIS	P
Pure	Elsevier	Assessment — CRIS	P
Symplectic Elements	Digital Science	Assessment — CRIS	P

How much should open digital scholarly communication infrastructure cost?

Until we can answer this question we will not know how much we need to budget in our own institutions or how much support we need from foundations and governments.

	Source	Entity	Staff	Budget
OCLC	GuideStar 2017	501©3	964	$213,096,462
Wikimedia Foundation	GuideStar 2017	501©3	230	$89,973,967
Ithaka Harbors	GuideStar 2017	501©3	388	$88,355,880
Lyrasis	GuideStar 2017	501©3	56	$77,916,237
Internet Archive	GuideStar 2017	501©3	150	$17,811,981

		Source	Entity	Staff	Budget
1	Crossref	GuideStar 2017	501©3	24	$7,467,519
2	Center for Open Science	GuideStar 2017	501©3	83	$6,994,472
3	OpenAIRE	estimate from web			$4,248,795
4	ORCID	GuideStar 2017	501©3	13	$3,909,207
5	Creative Commons	GuideStar 2017	501©3	11	$3,600,414
6	Digital Public Library of America	GuideStar 2017	501©3	17	$3,298,694
7	Organization A	e-mail 2018		15.5	$2,700,000
8	DuraSpace (Dspace, Fedora, VIVO, DuraCloud)	2018 AR	501©3	9.5	$1,905,145
9	Project A	Mapping	501©3	0	$1,567,463
10	Project B	Mapping	501©3	8	$1,550,000
11	Hypothes.is Project	GuideStar 2017	501©3	13	$1,349,133
12	Organization B		LLC	3	$1,200,000
13	ArXiv	Wikipedia 2017	Cornell		$826,000
14	Public Knowledge Project	2017 AR	Simon Fraiser	16	$826,000
15	Educopia Institue	GuideStar 2017	501©3	10	$756,716
16	Project C			6	$715,000
17	Organization C	e-mail 2018		8	$700,000
18	Organization D			3	$700,000
19	Organization E	e-mail 2018		8	$676,000
20	Dryad Digital Repository	GuideStar 2017	501©3	6	$663,206
21	Organization F		MOU	3	$530,000
22	Organization G	e-mail 2018		3	$500,000
23	Code for Science and Society	GuideStar 2017	501©3	2	$377,430
24	Project D	Mapping		9	$300,000
25	Organization H	e-mail 2018		2	$220,000
26	Organization J		under 501©3	1.1	$200,000
27	Islandora	2018 AR		2	$155,900
28	Organization K		MOU	1	$92,000

	Staff	Budget
Total	253.1	$40,561,575
Mean	9.4	$1,502,281
Median	8	$756,716

Average Organization Budget	Nmber of Organizations in Sample	Number or Organizations Minimum	Number or Organizations Maximum	Total Budget Minimum	Total Budget Maximum
$7,500,000	2	4	8	$30,000,000	$60,000,000
$5,000,000	4	8	12	$40,000,000	$60,000,000
$2,500,000	5	8	15	$20,000,000	$37,500,000
$1,250,000	3	10	15	$12,500,000	$18,750,000
$750,000	6	15	15	$11,250,000	$11,250,000
$500,000	4	15	20	$7,500,000	$10,000,000
$250,000	4	15	20	$3,750,000	$5,000,000
Total	28	75	105	**$125,000,000**	**$202,500,000**
Average Organizaton Budget				$1,666,667	$1,928,571
North American share at 60%				$75,000,000	$121,500,000

Questions/Comments

David W. Lewis
dlewis@iupui.edu

PDF of slides at: https://tinyurl.com/**y3cwmte5**

© 2019 David W. Lewis. This work is licensed under a Creative Commons Attribution-NonCommercial 4.0 International License.