



2017 IFLA NEWS MEDIA SATELLITE CONFERENCE, DRESDEN

Sächsische Landesbibliothek – Staats- und Universitätsbibliothek Dresden (SLUB)

16-18 August 2017

Frederick Zarndt

e-Legal Deposit Survey (2017)

- e-Legal Deposit Survey (2014)
- Survey of surveys
- e-Legal Deposit Survey (2017)
- e-Legal Deposit Survey (2017) preliminary results

WHAT IS LEGAL DEPOSIT?

Legal deposit is a legal requirement that a person or group submit copies of their publications to a repository, usually a library. The requirement is mostly limited to books and periodicals. Typically, the national library is one of the repositories of these copies.

Wikipedia contributors, "Legal deposit," Wikipedia, The Free Encyclopedia, https://en.wikipedia.org/w/index.php?title=Legal_deposit&oldid=787000576 (accessed August 16, 2017)

The German National Library is entrusted with the task of **collecting, permanently archiving, bibliographically classifying** and **making available** to the general public all German and German-language publications from 1913, foreign publications about Germany, translations of German works, and the works of German-speaking emigrants published abroad between 1933 and 1945.

Deutsche Nationalbibliothek website. http://www.dnb.de/EN/Wir/wir_node.html (accessed August 16, 2017)

- Survey of the born digital content legal deposit policies and practices
- Authors: Zarndt, McCain, Carner
- Respondents: 17 National Libraries
- 2015 International News Media Conference, NL Sweden, April 15-16, 2015;

Survey questions (focus on news):

- 1. Policies:** 3 questions about born-digital legal deposit laws or policies
- 2. Practices:** 6 questions about implementation of those laws and policies

Conclusions:

- Legal deposit laws vary widely from country to country
- Large counties with multiple governmental subdivisions may have a more difficult time enacting legislation
- Commerce regulation and practices as well as cost of resources are a challenge.
- Nordic countries have been leaders in the capture of digital content, while many others still make no legal provision for collecting digital content
- Overall, of the 16 countries surveyed, only 7 had policies that addressed the deposit of born-digital content.

SURVEY OF SURVEYS

- 17 surveys (2005 –2017) reviewed in order to optimize the quality of responses to the questions posed the 2017 e-Legal Deposit Survey
- Main categories:
 1. Audiovisual Preservation,
 2. Electronic Legal Deposit,
 3. Web Archiving,
 4. Digital Preservation of News,
 5. Preservation Standards and Best Practices,
 6. National / Federal Policies and Strategies for Preservation of Digital Heritage

SURVEYS 2015 – 2017

	Audiovisual Preservation	E-Legal Deposit	Web Archiving	Digital News Preservation	Preservation Standards & Best Practices	National Policies & Strategies
2005			IIPC Web Harvesting			
2006						
2007	IFLA AVMS		NL Netherlands			
2008	TAPE - EU		IIPC Member Profile			
2009		BL				
2010	IFLA AVMS					
2011		BL	NDSA		NDSA	
2012						
2013			IIPC-PWG NDSA			
2014				RJI Zarndt, Carner McCain		
2015						
2016	IFLA AVMS & IASA		NDSA		IFLA	UNESCO PERSIST

SURVEYS 2015 – 2017: WEB ARCHIVING



- **2005**: the International Internet Preservation Consortium (IIPC)'s web harvesting survey
- **2007**: The National Library of the Netherlands' Web Archiving User Survey
- **2008**: IIPC Member Profiles Survey
- **2013**: IIPC Preservation Working Group (PWG)'s survey
- National Digital Stewardship Alliance (NDSA) surveys:
 - **2011**
 - **2013**
 - **2016**

2005: IIPC Web Harvesting Survey

- Goal: identify and classify conditions that affect web harvesting
 - One response - from the Library of Congress - available
 - Three phases of harvesting identified: Acquisition, Parsing and Presentation
 - Rating system was: Easy, Difficult and Future
 - “Test Bed Taxonomy for Crawler” creates a “taxonomy of challenges for collecting web content.

2007: National Library of the Netherlands Web Archiving Survey

- Central question: “What should the contents and search options of the web archive look like?”
- Users preferred:
 - Full text searching
 - Hierarchical (taxonomic) presentation was also important
 - Anticipated collaboration with other organizations to expand their web archive collection and usage

2008: IIPC Member Profiles Survey

- Part 1: About You and Your Web Archiving Activities” & “Part 2: About Your IIPC Participation: Your Contributions and Expectations.”
- included three questions examining “Legal Issues and Policies,” with the results indicating that 15.6% of respondents have legal authority related to web archiving”

2013: International Internet Preservation Consortium (IIPC)

- Preservation Working Group conducted the survey called “IIPC PWG Survey on Web Archiving Practices”
 - Sent to 46 IIPC members
 - Received 25 complete responses
 - Focused on policy, access, preservation strategy, ingest, file formats and integrity
 - Survey indicated that long term preservation strategies were still lacking at many institutions

2011: National Digital Stewardship Alliance (NDSA) Web Archiving Survey

- Goal: to “better understand the landscape of web archiving activities in the United States, including identifying the organizations or individuals involved, the types of web content being preserved, the tools and services being used, and the types of access being provided.”
 - 77 respondents
 - 29 percent cultural heritage organizations
 - 22 percent government
 - 46 percent university communities
- **Survey takeaways:**
 - Recent entry of universities initiating web archiving programs
 - Lack of policies and permission guidelines
 - Inconsistent custodianship

2016: National Digital Stewardship Alliance (NDSA) Web Archiving Survey

- **Goals:** similar to 2011 and 2013 but with additional questions about program details,
- The survey consisted of 31 questions organized around five topics:
 - background information about the organization
 - current state of their web archiving program
 - tools and services used
 - access and discovery systems
 - policies including type of content, capture and availability
- 104 responses, an increase of 13% from 2013 survey

SURVEY OF SURVEYS: LESSONS LEARNED

- *Keep it simple!* We chose to streamline the queries and potential responses.
- *Use professional tools:* survey software (SurveyGizmo) that allowed for the use of question skip logic, letting respondents skip irrelevant questions, based on previous answers.
- *Use multiple angles to approach respondents:* personal email, mailing lists, conference.
- *Use language people understand:* translate the survey



www.surveygizmo.com/s3/3651847/2017-digital-e-legal-deposit-survey

Open until the end of August 2017

2014 & 2017 RESPONDENTS



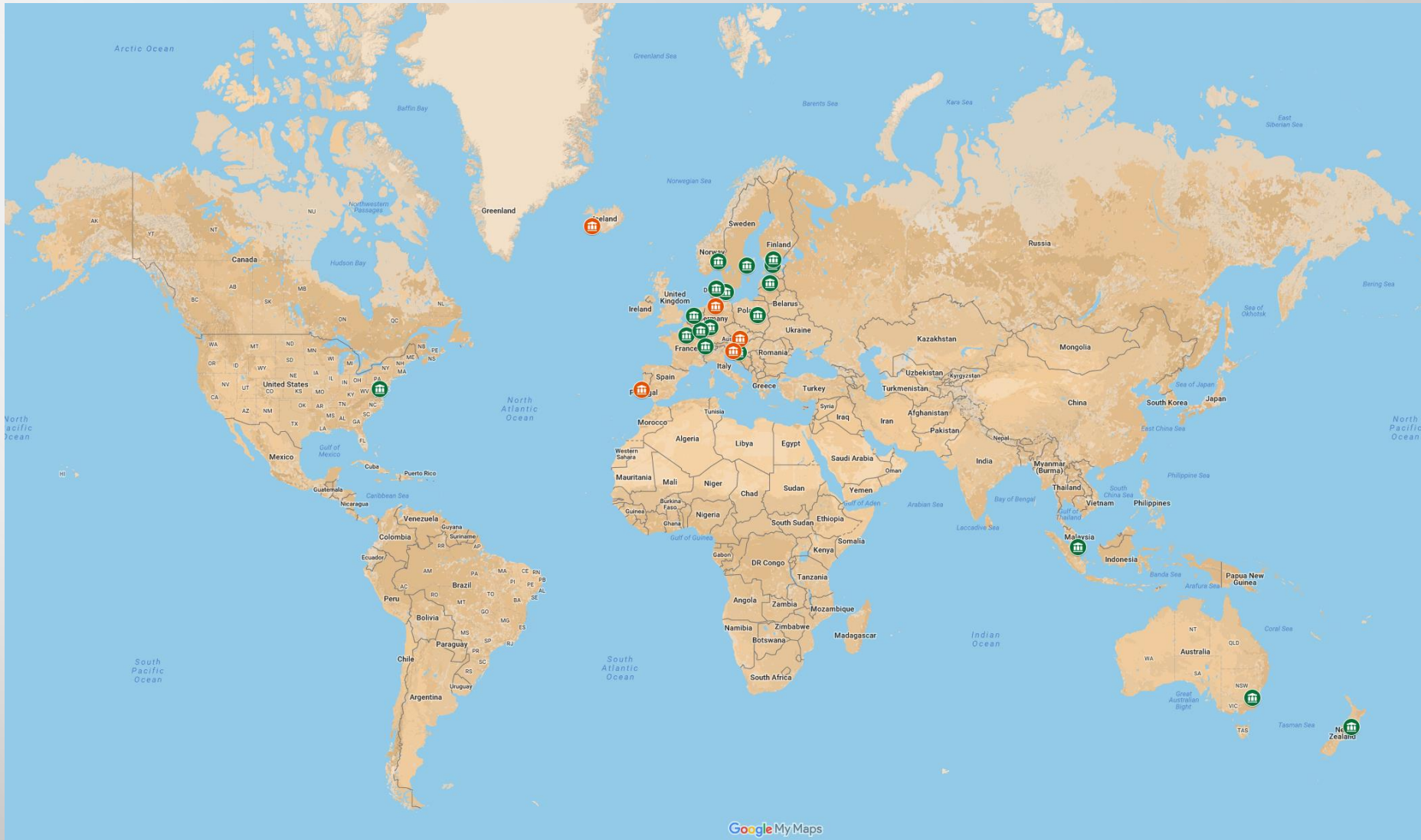
2014

Australia: National Library of Australia
Croatia: Nacionalna i sveučilišna knjižnica u Zagrebu
Denmark: Statsbiblioteket (Aarhus)
Estonia: Eesti Rahvusraamatukogu
Finland: Kansalliskirjasto
France: Bibliothèque nationale de France
Germany: Deutsche Nationalbibliothek
Latvia: Latvijas Nacionālā bibliotēka
Luxembourg: Bibliothèque nationale de Luxembourg
New Zealand: National Library of New Zealand
Norway: Nasjonalbiblioteket
Poland: Biblioteka Narodowa
Singapore: National Library Board
Sweden: Kungliga biblioteket - Sveriges nationalbibliotek
Switzerland: Schweizerische Nationalbibliothek / Bibliothèque nationale suisse
The Netherlands: Koninklijke Bibliotheek
United States: Library of Congress

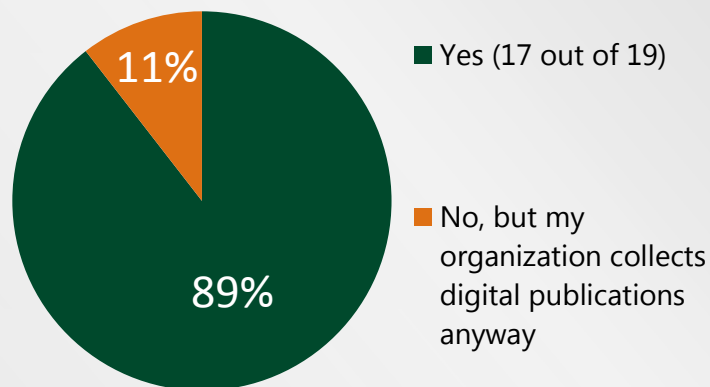
2017 (so far)

Australia: National Library of Australia
Austria: Österreichische Nationalbibliothek
Croatia: Nacionalna i sveučilišna knjižnica u Zagrebu
Denmark: Det Kgl. Bibliotek
Estonia: Eesti Rahvusraamatukogu
Finland: Kansalliskirjasto
France: Bibliothèque nationale de France
Germany: Deutsche Nationalbibliothek
Germany: Staats- und Universitätsbibliothek Hamburg
Iceland: National and University Library of Iceland
Latvia: National Library of Latvia
New Zealand: National Library of New Zealand
Norway: National Library of Norway
Portugal: Fundação para a Ciência e Tecnologia
Portugal: Arquivo.pt
Singapore: National Library Board Singapore
Slovenia: Narodna in univerzitetna knjižnica
Switzerland: Schweizerische Nationalbibliothek / Bibliothèque nationale suisse
The Netherlands: Koninklijke Bibliotheek
United States: Library of Congress

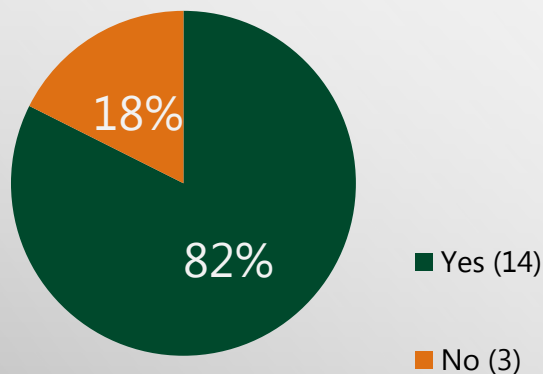
2014 & 2017 RESPONDENTS



Does your country / state have a legal deposit law?



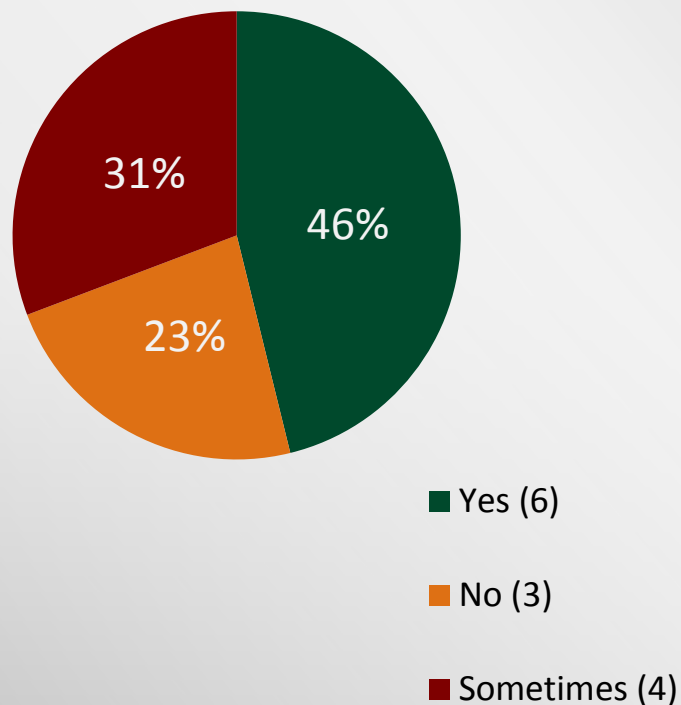
Does the legal deposit law cover digital works?



Comments:

- Legal deposit law revised since 2014:
 - Norway (2016)
 - Estonia (2017)
 - Croatia: new Law on Libraries expected
- State Law for all publications (Germany)
- Voluntary deposits: publishers encouraged to deposit digital items even though it is not required by law (Singapore)

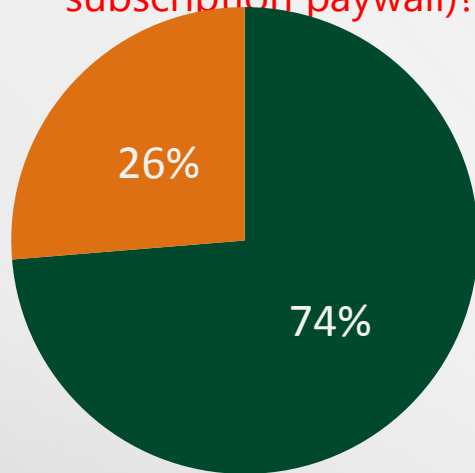
Do the laws of your country / state **require** publishers to legally deposit digital works? In this case we mean that publishers **MUST** send digital works to one or more legal deposit authorities.



Sample comments:

- Digital works published on the internet (public electronic network) must be made accessible to the library for download (even if behind a paywall). Publisher's do not need to 'send' them in. Digital publishing on a physical media (e.g. DVD) is subject to deposit.
- In 2016 our library welcomed long-anticipated changes to the copyright law. For the first time in its history, the Library could at last collect electronic publications under the legal deposit provisions of the law. Legal deposit provisions were extended to cover the online publishing landscape. This includes all national print and electronic books, journals, magazines, newsletters, reports, sheet music, maps, websites and public social media.
- The legal deposit law from the 80sin general does NOT cover digital works. However, there is decree-law from 2006, that extends the legal deposit to also include MsC and PhD theses in digital format. Our organization also manages the network of repositories that preserve the theses.
- Our law permits us to take a copy, and, if we need it, to require the publisher's assistance in doing this. This contrasts with the situation for physical format items, where the obligation is on the publisher to deposit.

Do the laws of your country / state require cultural heritage institutions (libraries) to harvest websites and webpages that are publicly available (not behind a subscription paywall)?

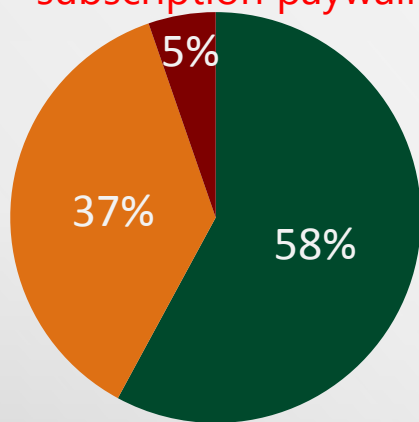


■ Yes (14) ■ No (5)

Sample comments:

- "Require" isn't quite the right word - we have the right to copy, but the intention of the legislation was to be selective rather than comprehensive in digital collecting.
- Under the legal deposit provisions in the Act, the National Library requests the delivery of online material through the process known as web crawling or web harvesting. This process uses harvesting robots to initiate requests to the web servers delivering online content using the HTTP protocol 'Get' request process.
- Not yet, but we do harvest government websites currently. We are reviewing our legal deposit law to extend it to include websites.

Do the laws of your country / state require cultural heritage institutions (libraries) and publishers of websites and webpages to cooperate in order to preserve digital works when these works are behind a subscription paywall?



■ Yes (11)

■ No (7)

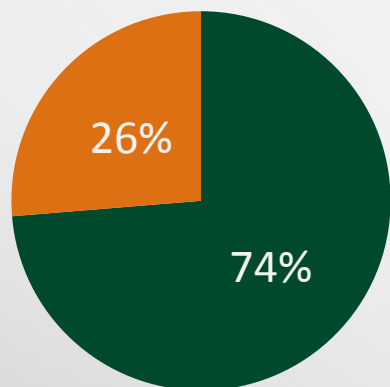
■ I don't know (1)

Sample comments:

- Publishers are required to make this material accessible. There are no penalties outlined for non-compliance. We have not actively pursued this as most material of interest is not behind a paywall.
- Recent changes allow this to happen, but the National Library has been focusing online books and serials and harvesting open access websites. Work to explore how to harvest material behind a paywall will commence in 2018.

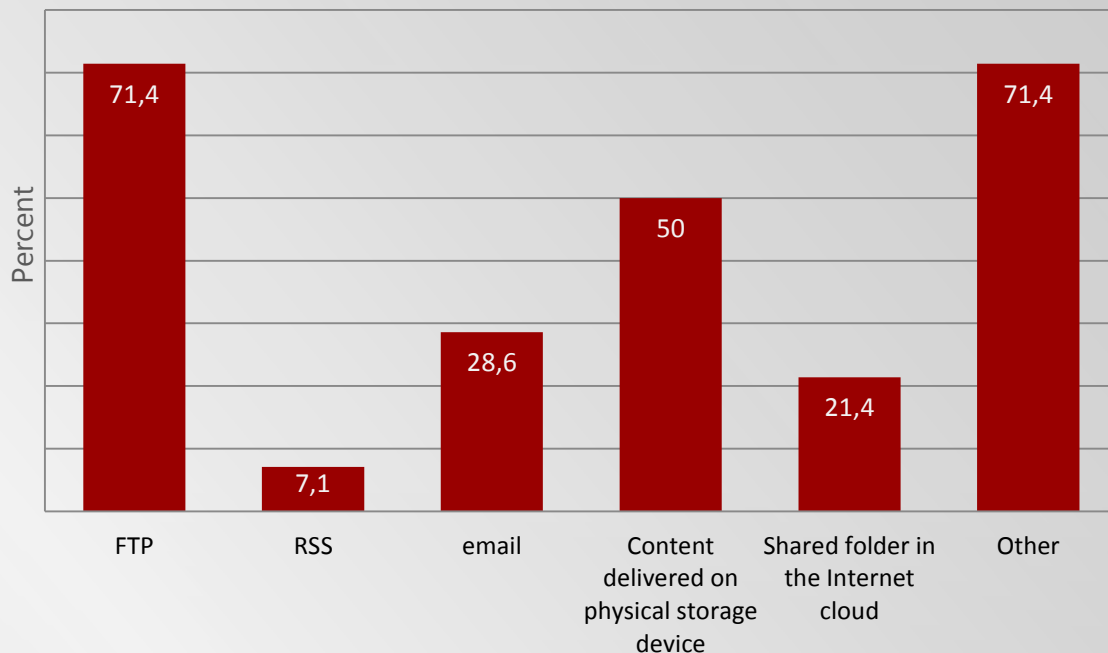
Does your library receive digital works from publishers?

For this question by "receive" we mean that publishers initiate the transmission of digital works to the legal deposit authority (library). In tech speak, the publisher "pushes" the works to the authority (library).



■ Yes (14) ■ No (5)

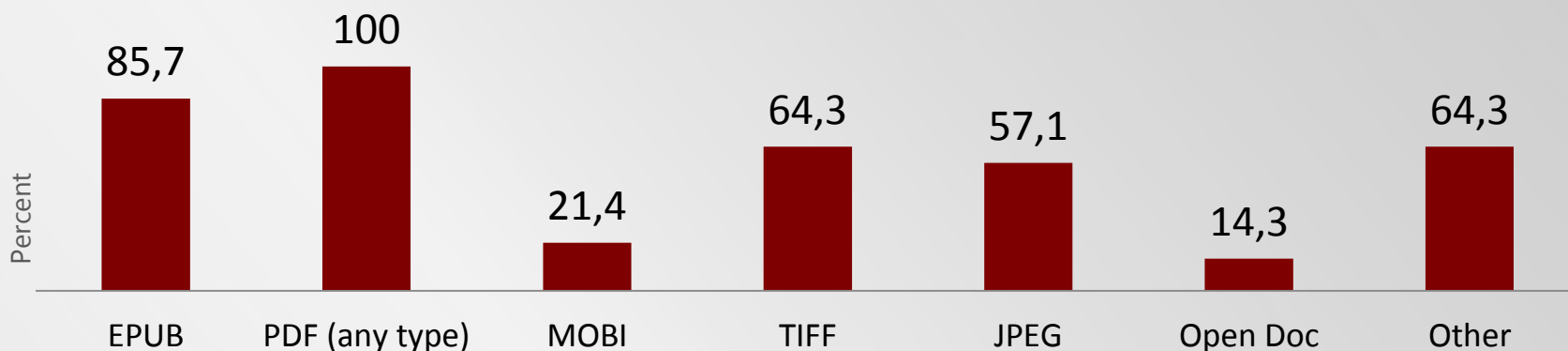
If publishers "push" digital works to libraries, how do you receive them?



Sample comments:

- OAI-PMH (metadata with direct links to files, so actually more pull than push); web form with upload.
- Publishers upload the file in preferred formats via our deposit website.

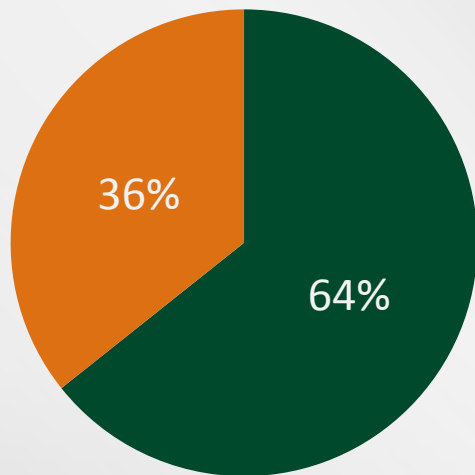
In what format(s) does your library accept digital works?



Other:

- MP3, MP4, PDF 2000 (e-legal deposit newspapers), PNG,
- EPUB, PDF or mobi files for **books, journals, magazines, newsletters** and **music scores** (epub is preferred); PDF, GeoPDF, TIFF, or GeoTIFF files for **maps**; Word documents – not accepted; for **cover art**: publishers can upload JPG, JPEG, TIF or TIFF cover images with an RGB colour profile (max 250 MB); e-deposit system accepts a CMYK colour profile but the files cannot be currently display online.
- ZIP for HTML and packages with attachments (research data)
- XML (e.g. JATS/NLM-DTD), HTML

Does your library offer a batch or bulk legal deposit service to publishers?

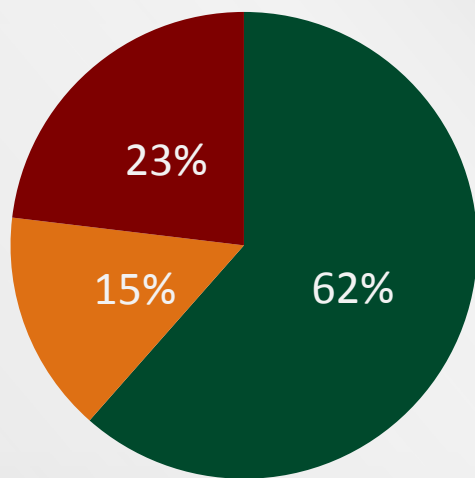


■ Yes (9) ■ No (5)

Sample responses:

- Bulk deposit methods are available upon agreement/negotiation with publishers.
- Only upon request. Usually the library is the one requesting.
- If the library receives older vintage newspapers, the publishers are offered digitised files in return. New newspapers are available to the publishers on the library's website (the same applies to regional libraries).
- Publishers can deliver in bulk to us if they find it inconvenient to dispatch items frequently.
- We already announced the possibility to serve our publishers as digital archive but they seem not to be interested or have another solutions.

What type of access do you provide to e-legal deposit digital content?



■ Onsite only (8)

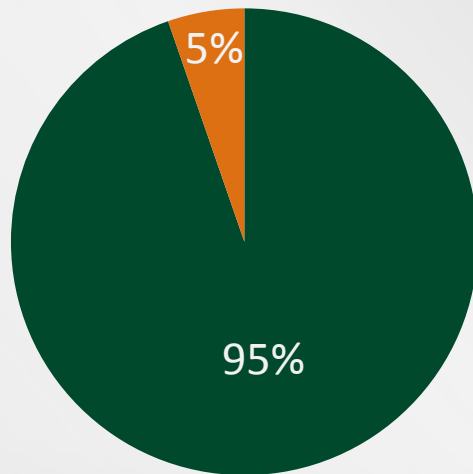
■ Onsite and offsite after an embargo period (2)

■ Onsite and offsite immediately (3)

Sample responses:

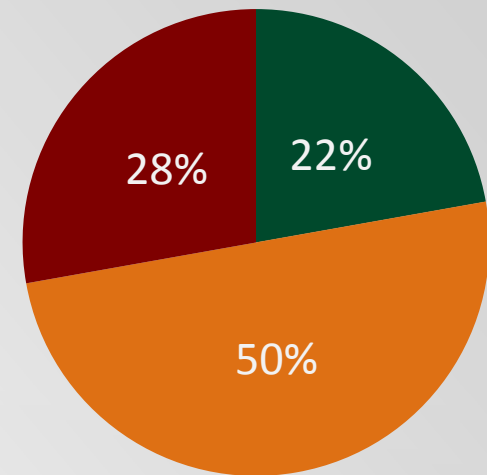
- According to the legal deposit law it is onsite. But the right holders can grant us the right to give offsite access and for these publications we offer offsite access.
- Access depends on negotiations with publishers since we don't have a legal deposit.
- By the law publisher has right to assign the type of access. All above types are possible.
- Onsite-only access is currently limited to 2 dedicated PCs, with no download/upload capability. Print only.

Does your library harvest websites and pages?



■ Yes (18) ■ No (1)

If your library harvests websites and pages, does this include those behind a paywall?

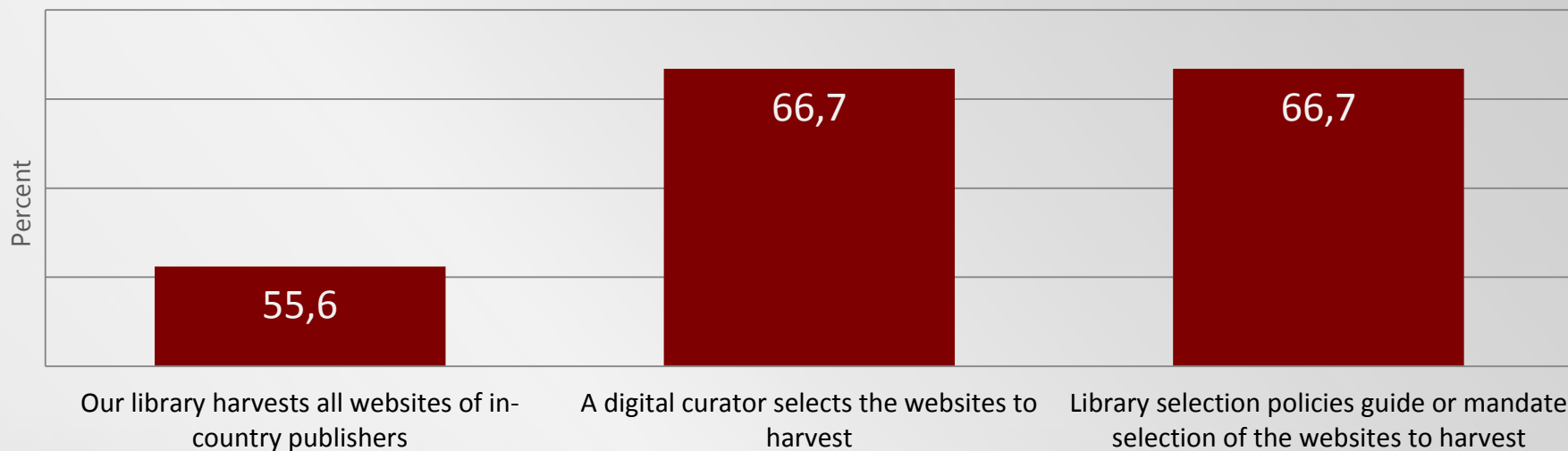


■ Yes (4)

■ No (9)

■ For selected websites only (5)

If your library harvests websites and webpages, what criteria are used to decide if born digital works from a particular publisher should be preserved?

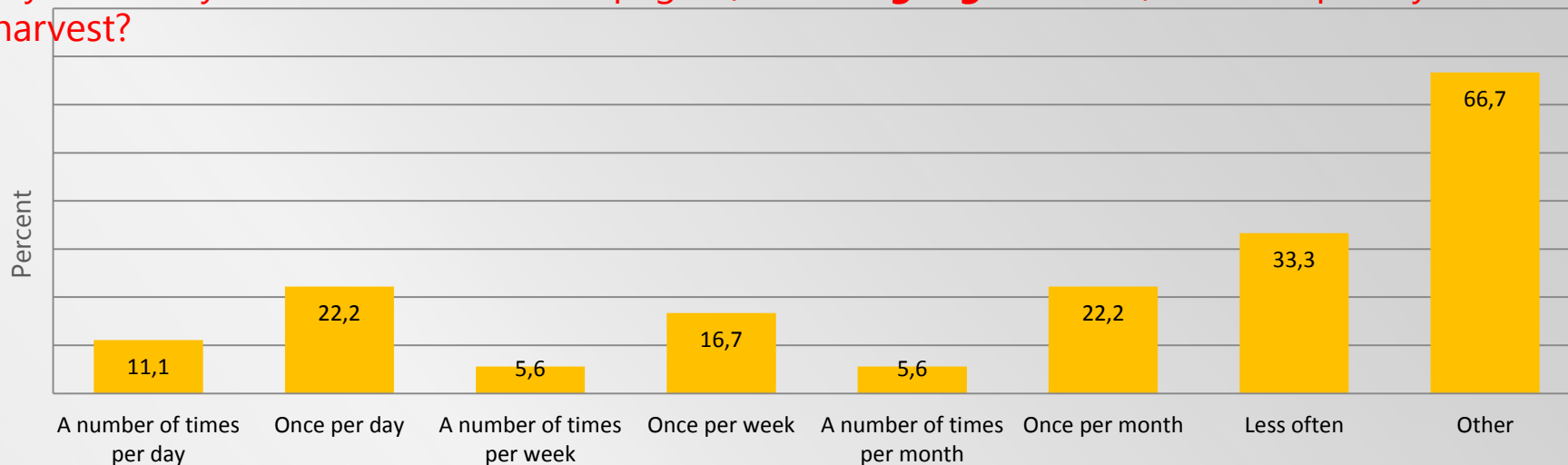


Sample comments:

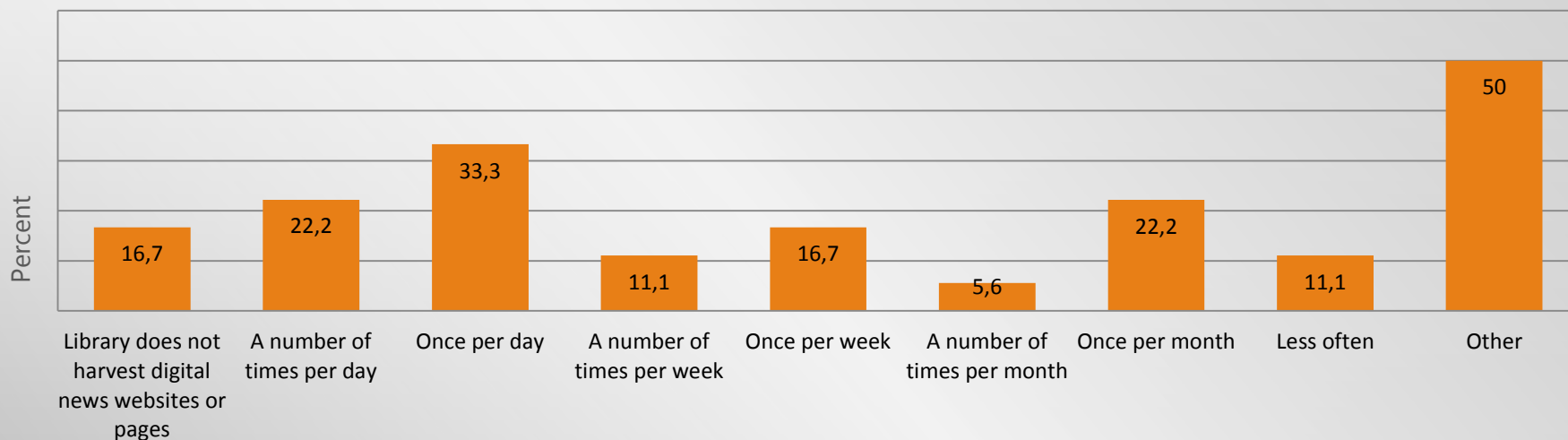
- The main criteria is national author, national language or published nationally. For the thematic collection we have about 1375 websites that we harvest on a regular basis. For the domain based harvesting the number of seed URLs is 117,000.
- Although we are entitled by law to harvest websites behind paywall, in reality we so far haven't requested access to any protected website.
- We accept user suggestions.
- Selective / Thematic crawling, e.g. elections.

2017 E-LEGAL DEPOSIT SURVEY: PRELIMINARY RESULTS

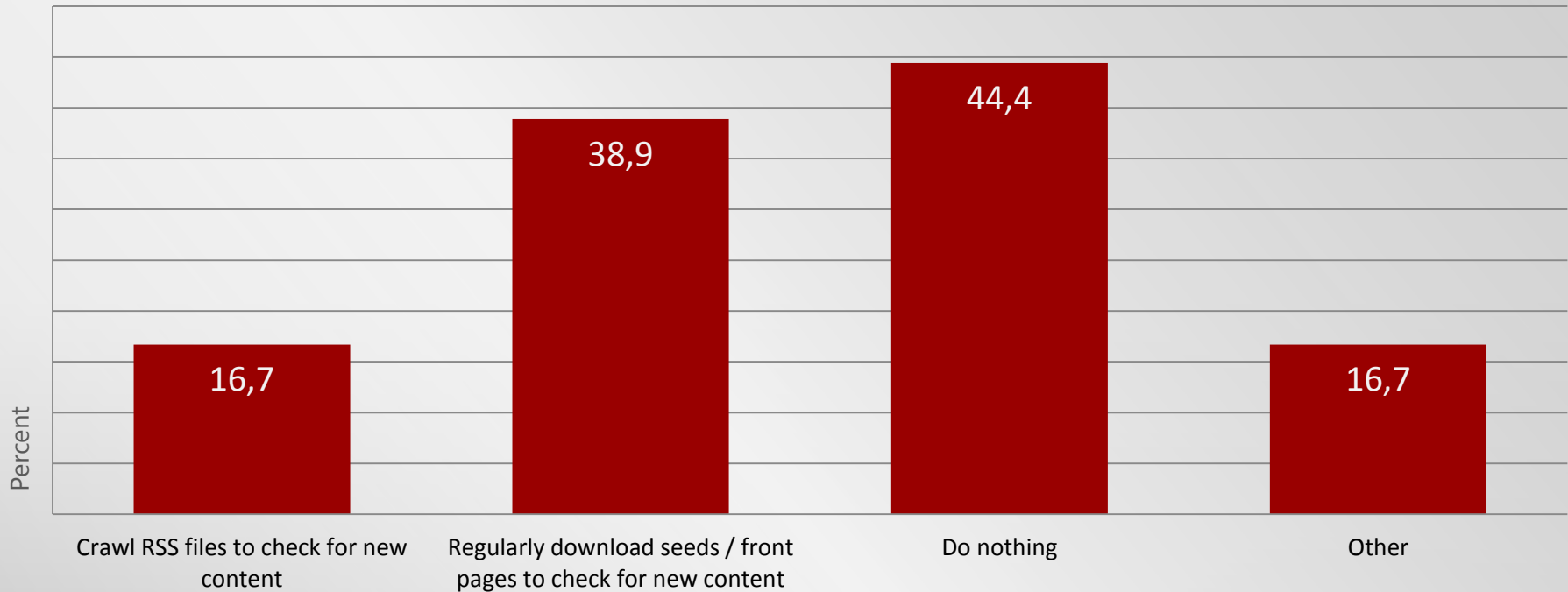
If your library harvests websites and pages (**excluding digital news**), how frequently does it harvest?



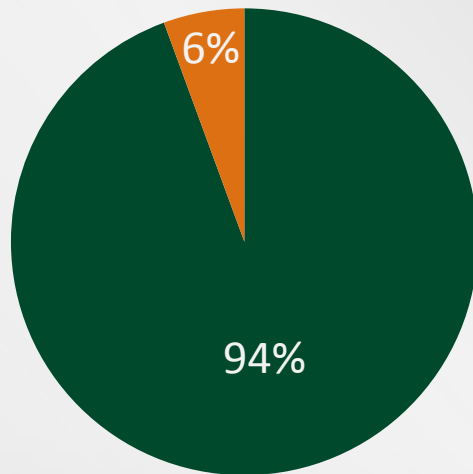
If your library harvests **digital news websites and pages**, how frequently does it harvest?



Depending on the publisher, born digital content published on the web may be updated several times in an hour, day, or week. What methods does your library use to capture updated page?

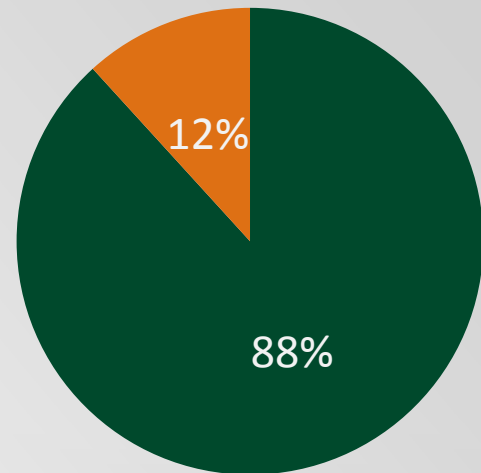


Does your library require preservation of its digital content?



■ Yes (17) ■ No (1)

At your library digital preservation is ...



■ Mandatory for all digital works and websites (15)

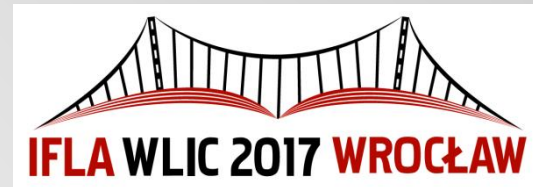
■ Automatic but not mandatory (publisher or the library can choose not to preserve certain content) (2)



www.surveygizmo.com/s3/3651847/2017-digital-e-legal-deposit-survey

Open until the end of August 2017

[LINK TO IFLA PAPER](#)



Session 074, IASE Conference Room

Sunday, 20 August 2017
12:30 - 13:30