

2013 NDSA Web Archiving Survey Report Highlights

Nicholas Taylor ([@nullhandle](#))
Web Archiving Service Manager
[Stanford University Libraries](#)

[SAA Annual Meeting: Web Archiving Roundtable](#)

August 13, 2014

NDSA Web Archiving Survey Working Group



Jefferson Bailey
Internet Archive / Archive-It



Kristine Hanna
Internet Archive / Archive-It



Christie Moffatt
National Library of Medicine



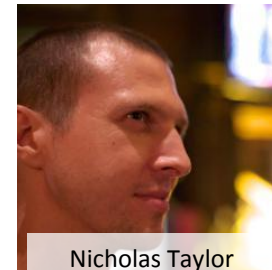
Abbie Grotke
Library of Congress



Cathy Hartman
University of North Texas



Edward McCain
University of Missouri



Nicholas Taylor
Stanford University

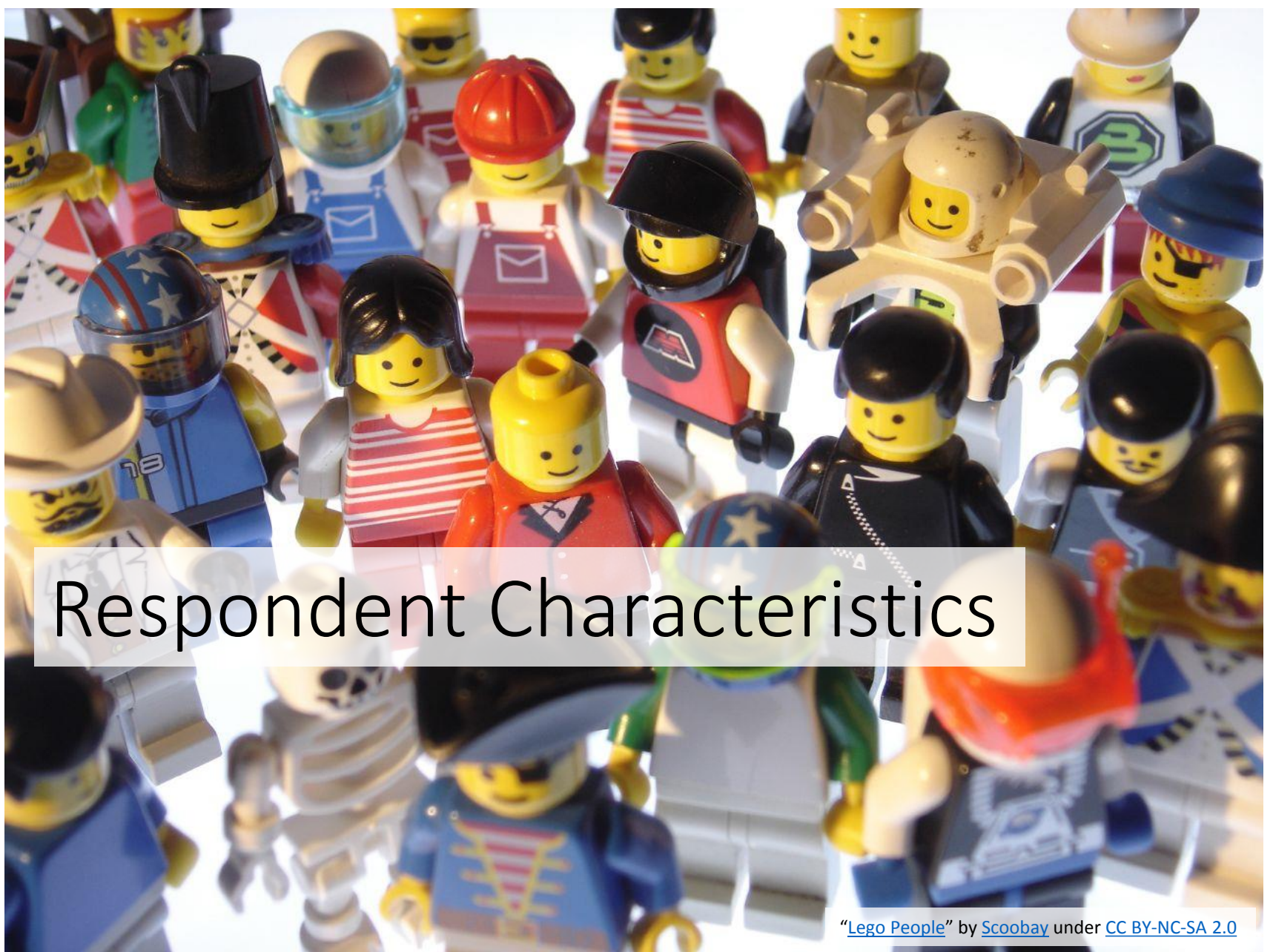
NDSA Web Archiving survey background

2011

- 78 respondents
- program info
- tools and services
- access
- policies

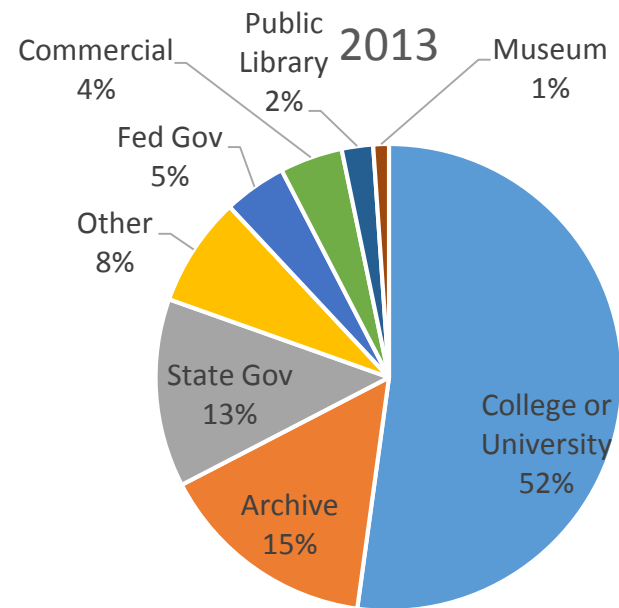
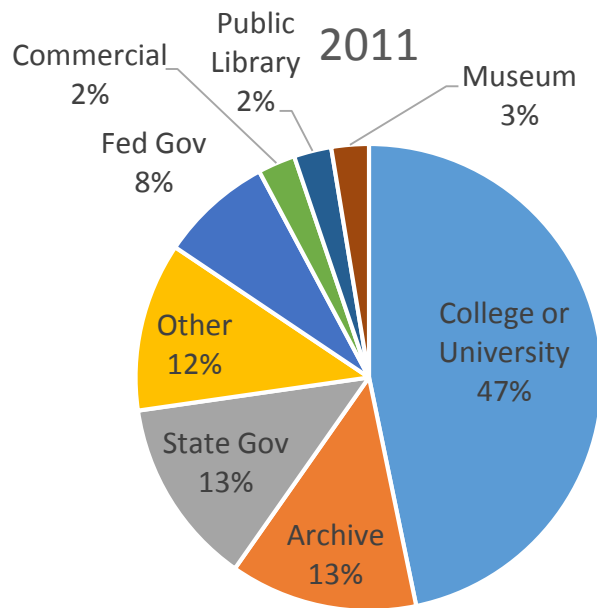
2013

- 92 respondents
- program info
 - staff time, metrics, skills, content concerns
- tools and services
- access and discovery
 - new discovery options
- policies
 - embargo, social media, robots.txt, resources






"Lego People" by Scoobay under [CC BY-NC-SA 2.0](https://creativecommons.org/licenses/by-nc-sa/2.0/)

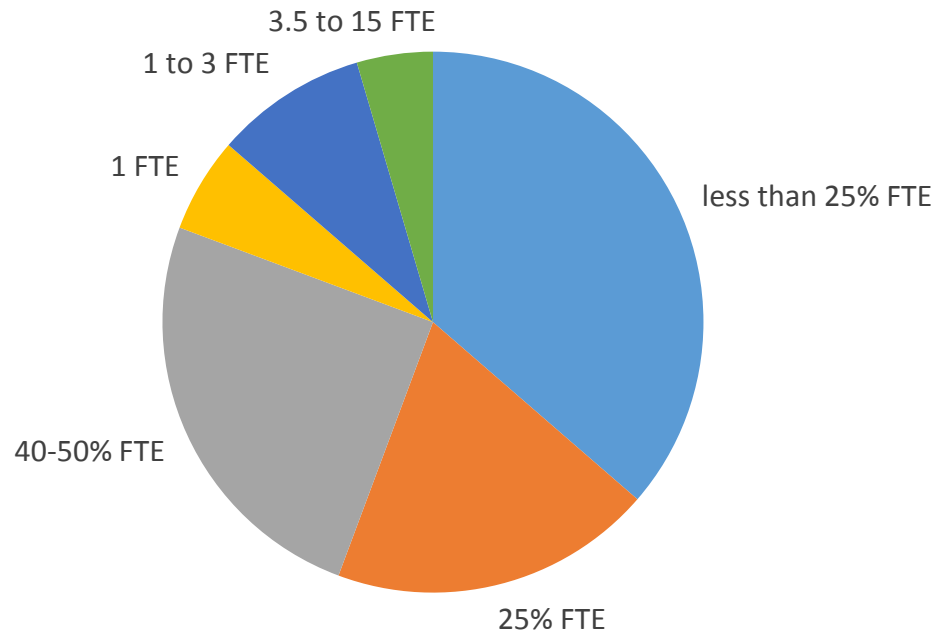
universities still make up most programs



SAA WebArch RT tops group affiliations

group	2011	2013
 IIPC netpreserve.org INTERNATIONAL INTERNET PRESERVATION CONSORTIUM	8%	7%
 NDSA	31%	33%
Web Archiving  Roundtable		45%

most programs are fractionally staffed

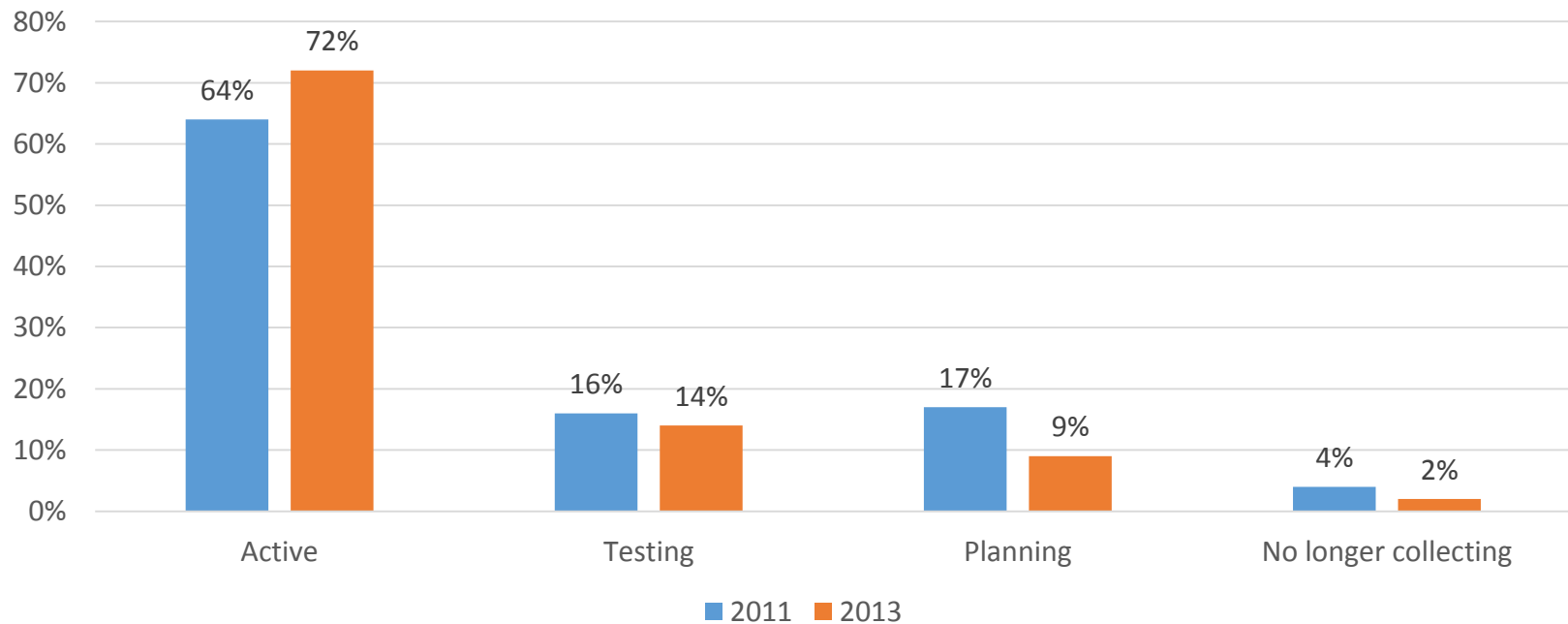




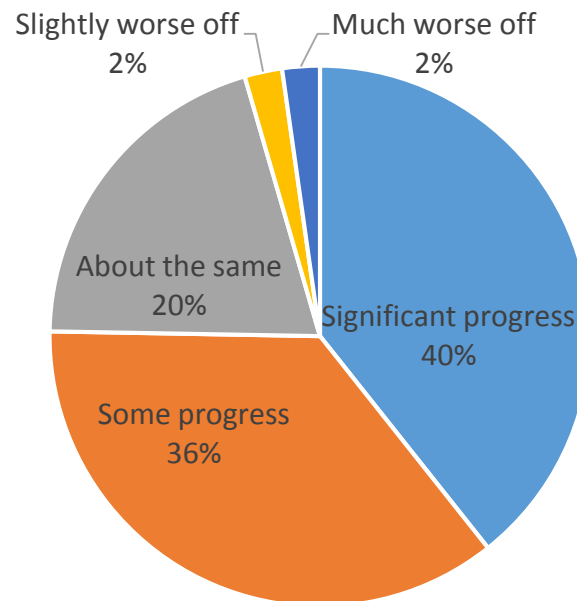
Maturity and Progress

["Apple Mouse Evolution"](#) by [raneko](#) under [CC BY 2.0](#)

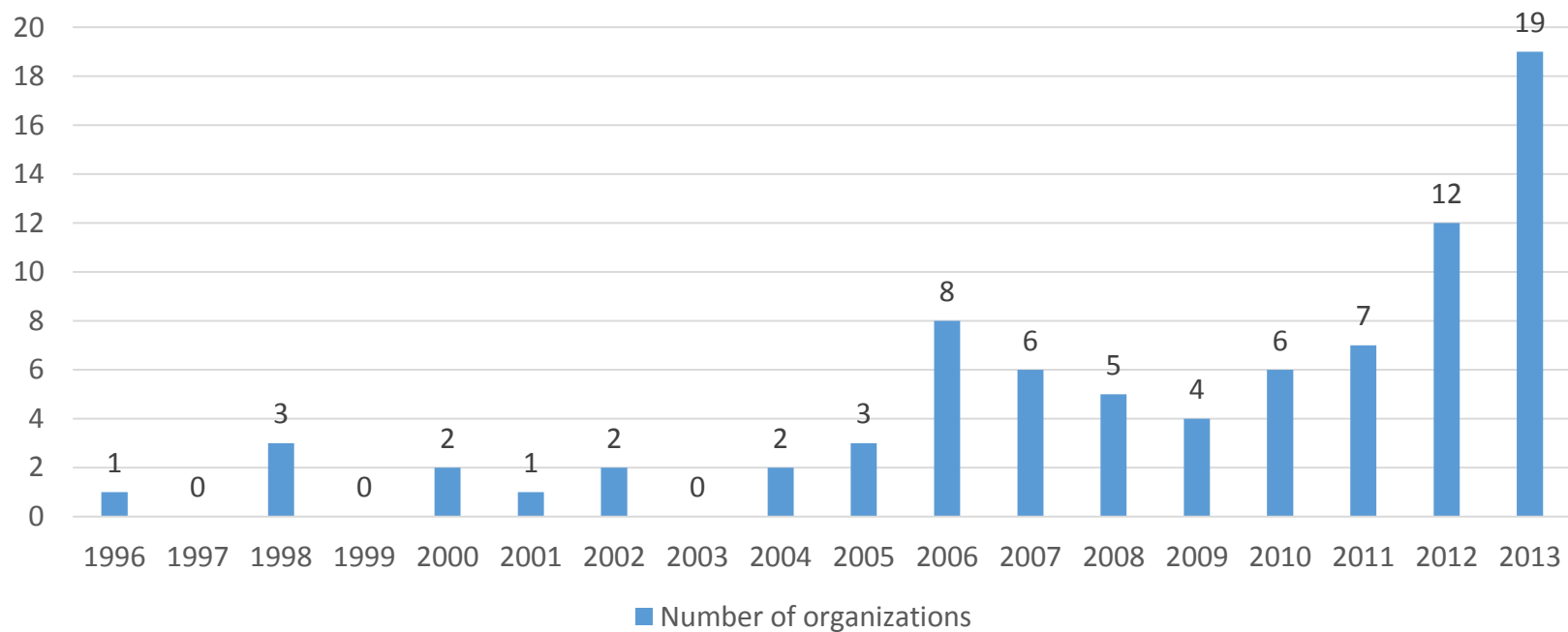
programs have matured slightly since 2011



strong perceptions of progress since 2011



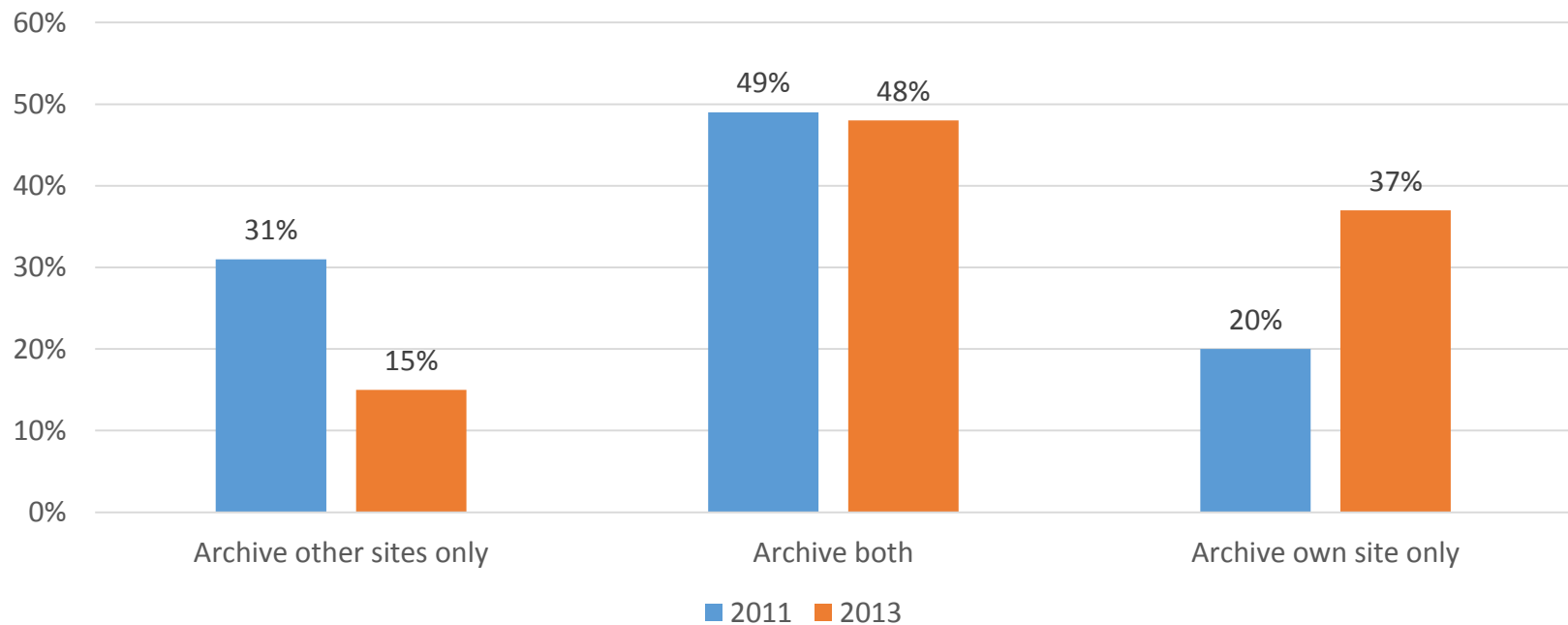
many new programs since 2011



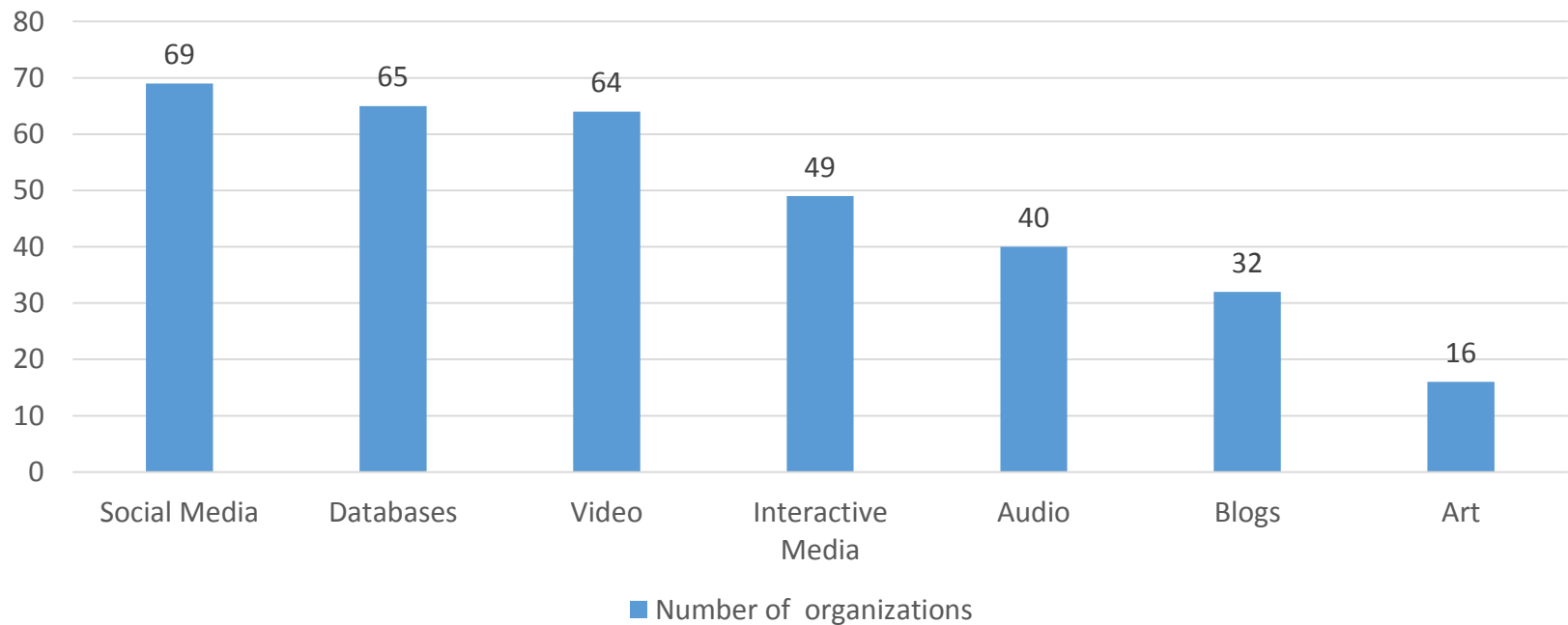
Archiving Focus

"Ant Farm Media Van v.08 (Time Capsule) in Bellewether at Southern Exposure" by [Steve Rhodes](#) under [CC BY-NC-SA 2.0](#)

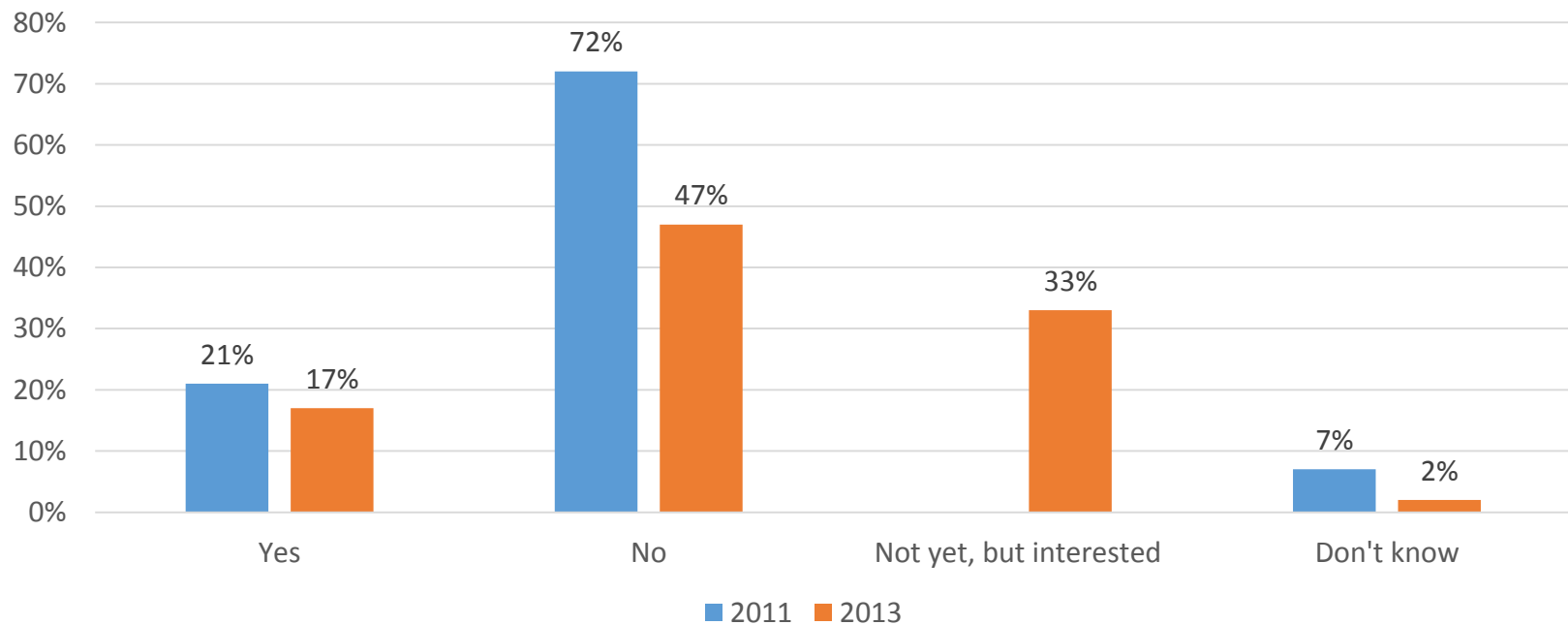
more programs are only self-archiving

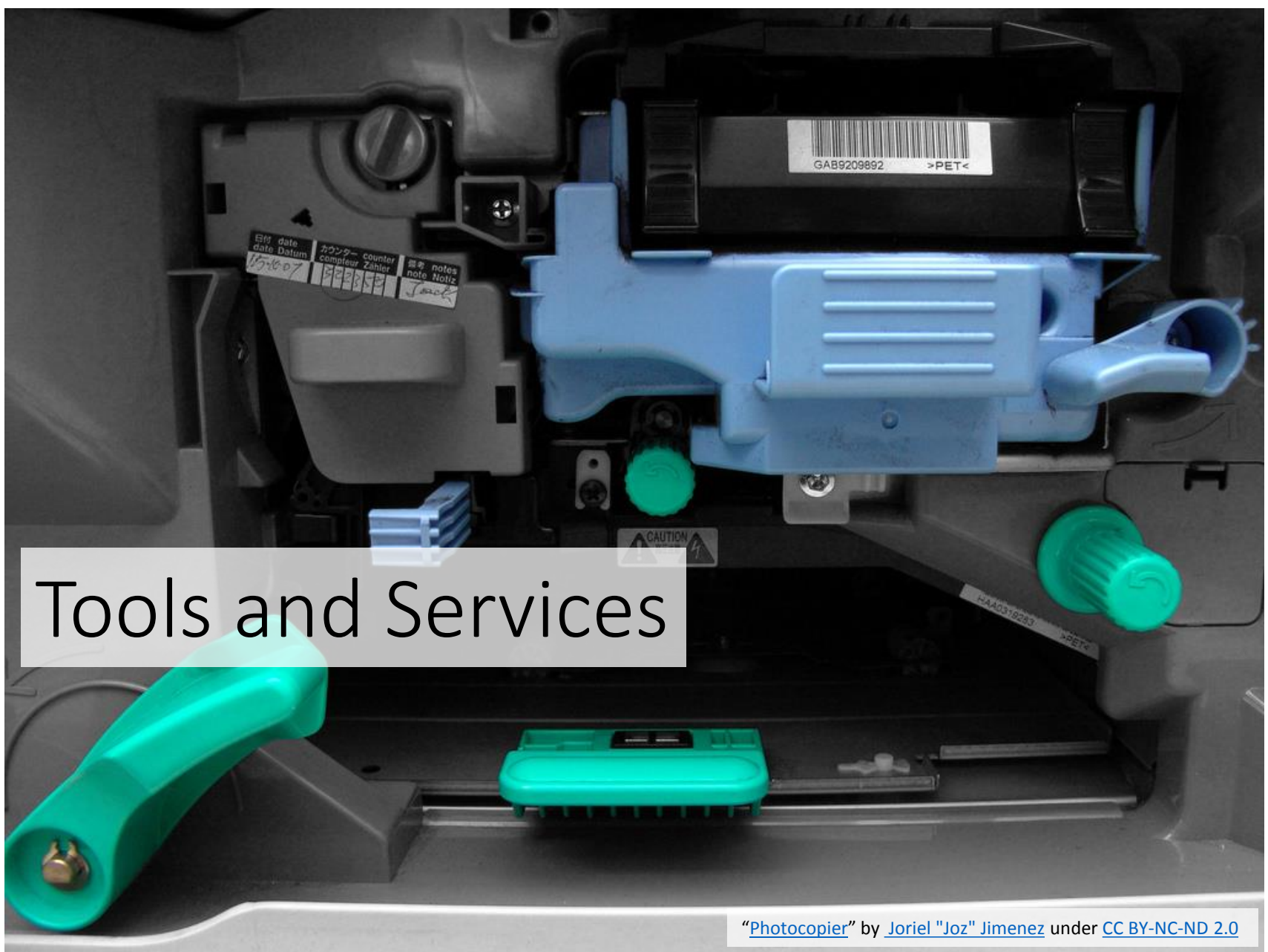


concern about social media, databases, video



untapped interest in collaboration

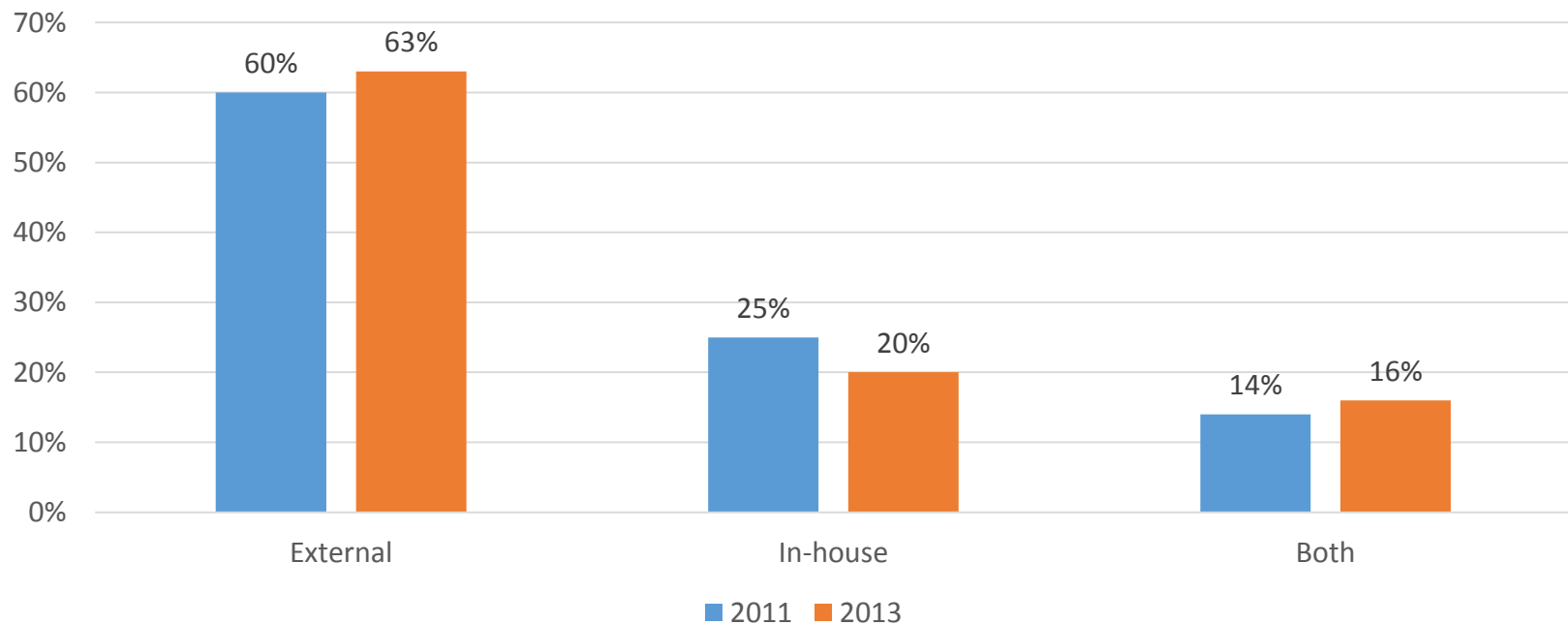




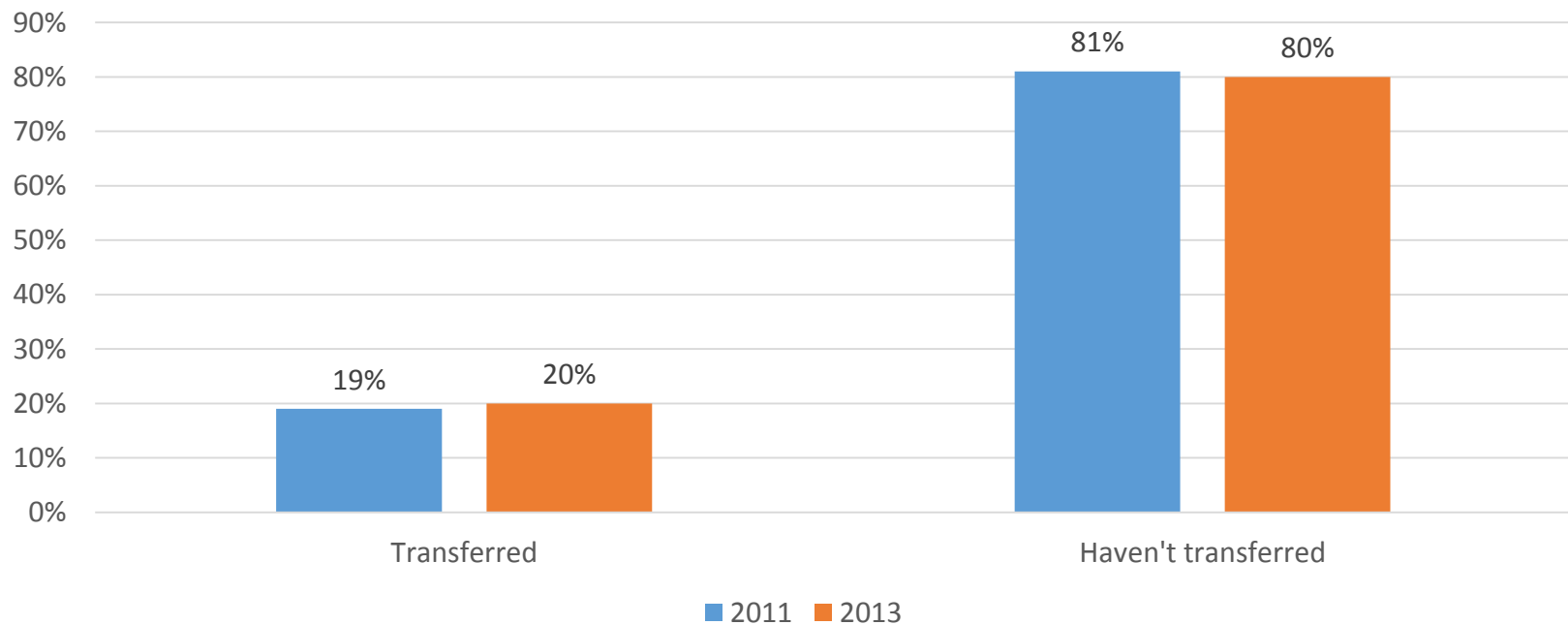
Tools and Services

"Photocopier" by [Joriel "Joz" Jimenez](#) under [CC BY-NC-ND 2.0](#)

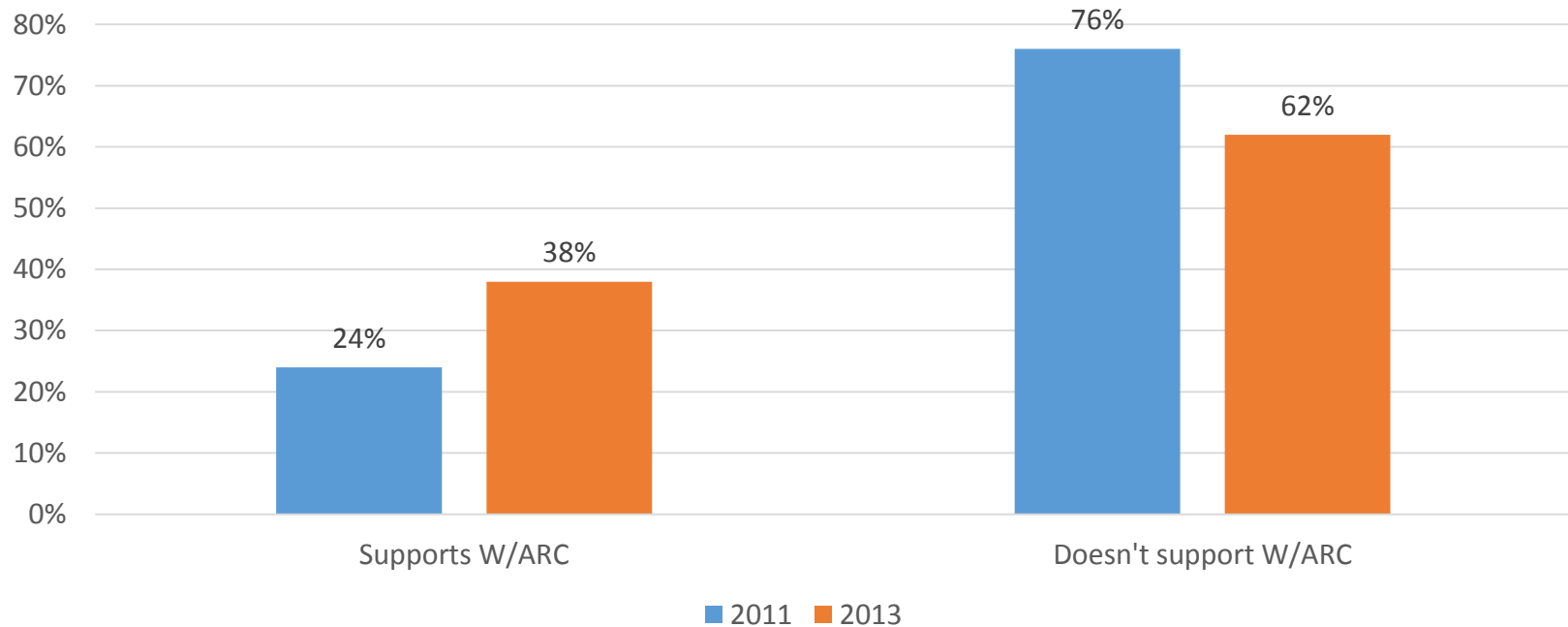
web archiving as a service still most popular



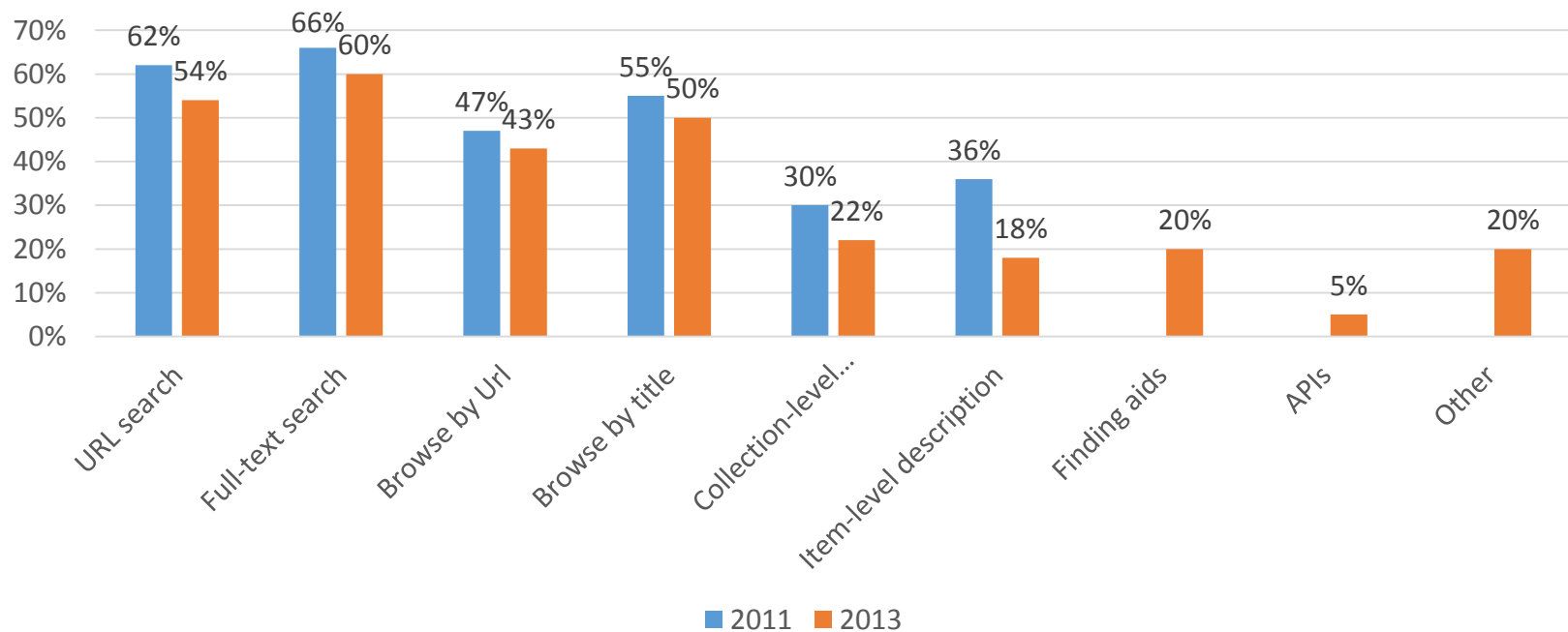
data not transferred from service provider



increased use of tools supporting W/ARC



less granular descriptive metadata

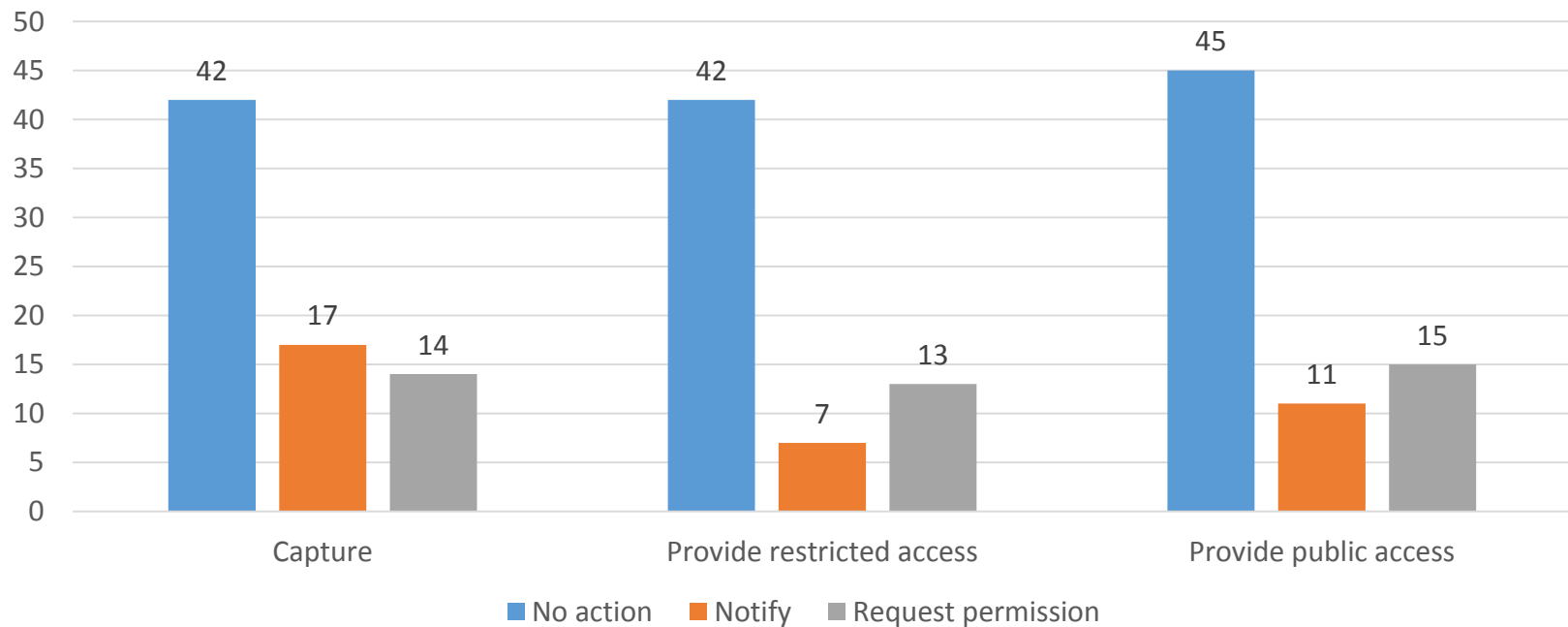


The background of the slide is a close-up photograph of three wooden crates. Each crate has a black line-art label. The first crate on the left has two upward-pointing arrows. The middle crate has a label of a glass, with the text 'Handle With Care' partially visible below it. The third crate on the right has a label of an open umbrella, with the text 'Keep Dry' partially visible below it. A semi-transparent white rectangular box is overlaid on the middle crate, containing the title text.

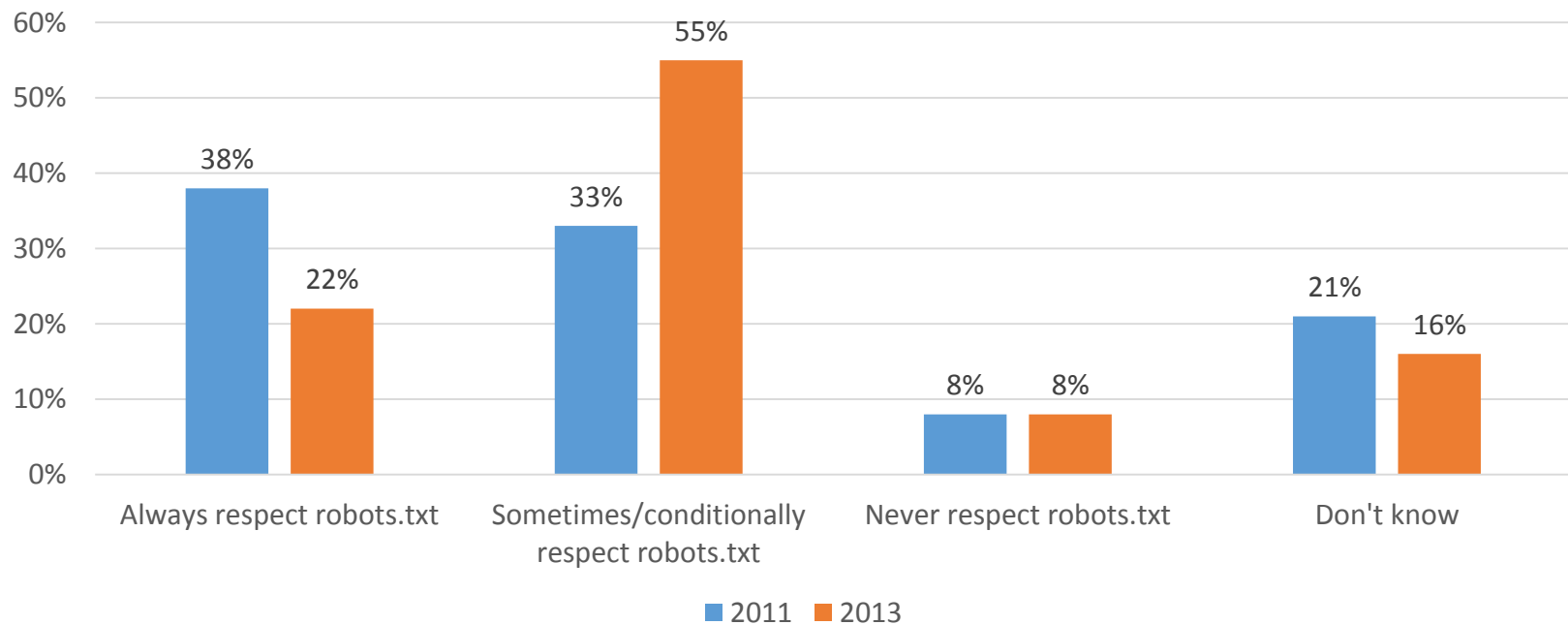
Archiving Policies

"Handle With Care" by [ServInt](#) under [CC BY-NC-ND 2.0](#)

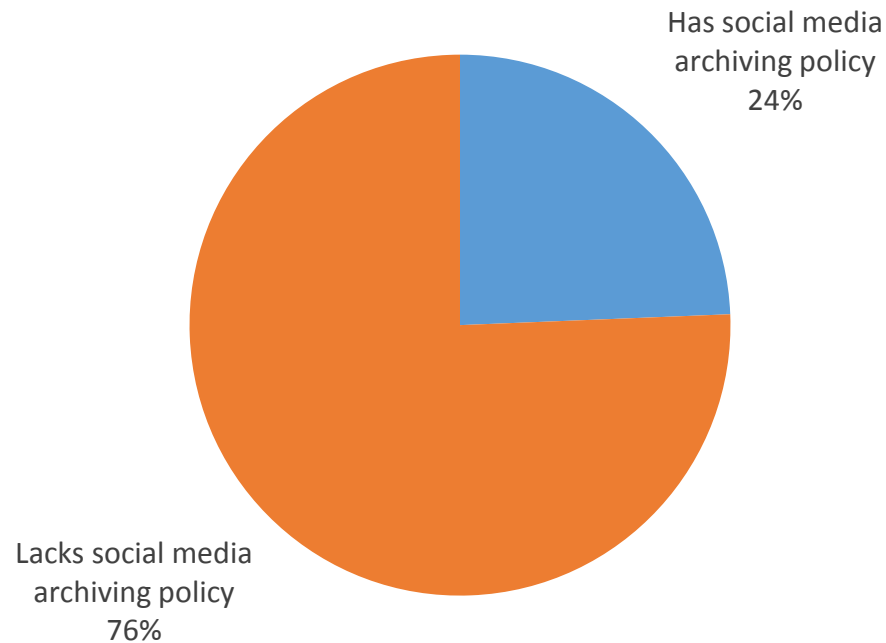
most don't notify or seek permission



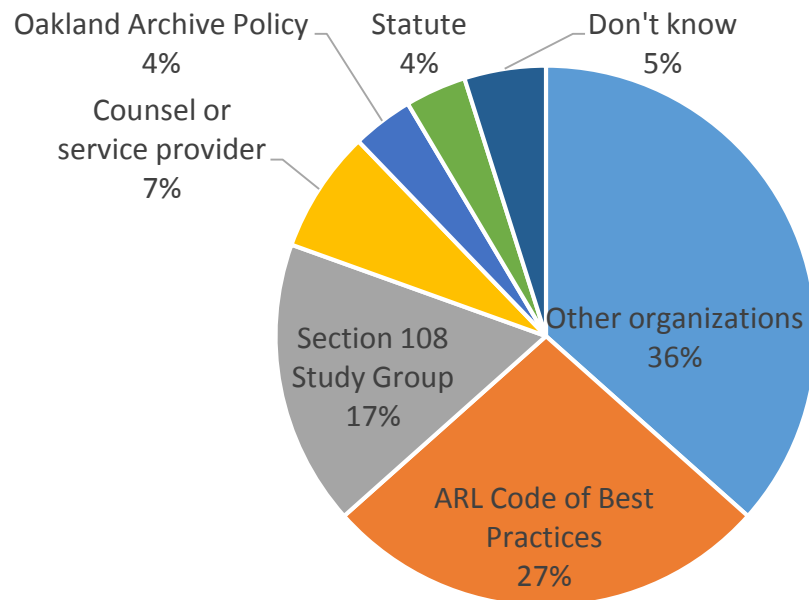
more conditional handling of robots.txt



social media archiving policies are uncommon



policies based on community practices



takeaways and questions for SAA WebArch RT

- for individual organizations:
 - if you're only self-archiving, what's on your [roadmap](#)?
 - how are you [preserving](#) your web archive data?
 - how do you [describe](#) and [enable discovery](#) of web archives?
 - how do you handle [robots.txt](#)?
 - what are your plans for [social media archiving policy](#)?
- for the group:
 - what is [this group](#) (vs. IIPC, NDSA) best equipped to do?
 - what kind of [collaboration](#) are you interested in?

THANK
YOU

Nicholas Taylor
[@nullhandle](#)

["Thank You"](#) by [vistamommy](#) under [CC BY 2.0](#)