

A Snapshot of the U.S. Web Archiving Landscape through the 2013 NDSA Survey Report

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

[Archive-It Partner Meeting](#)

November 18, 2014

"Barrage balloon manufacture..." by [Alfred T. Palmer](#) under [public domain](#)

NDSA Web Archiving Survey Working Group



Jefferson Bailey
Internet Archive / Archive-It



Kristine Hanna
Internet Archive / Archive-It



Christie Moffatt
National Library of Medicine



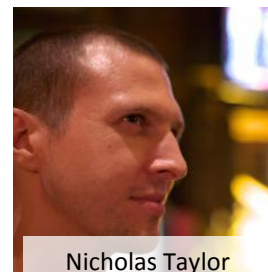
Abbie Grotke
Library of Congress



Cathy Hartman
University of North Texas



Edward McCain
University of Missouri



Nicholas Taylor
Stanford University

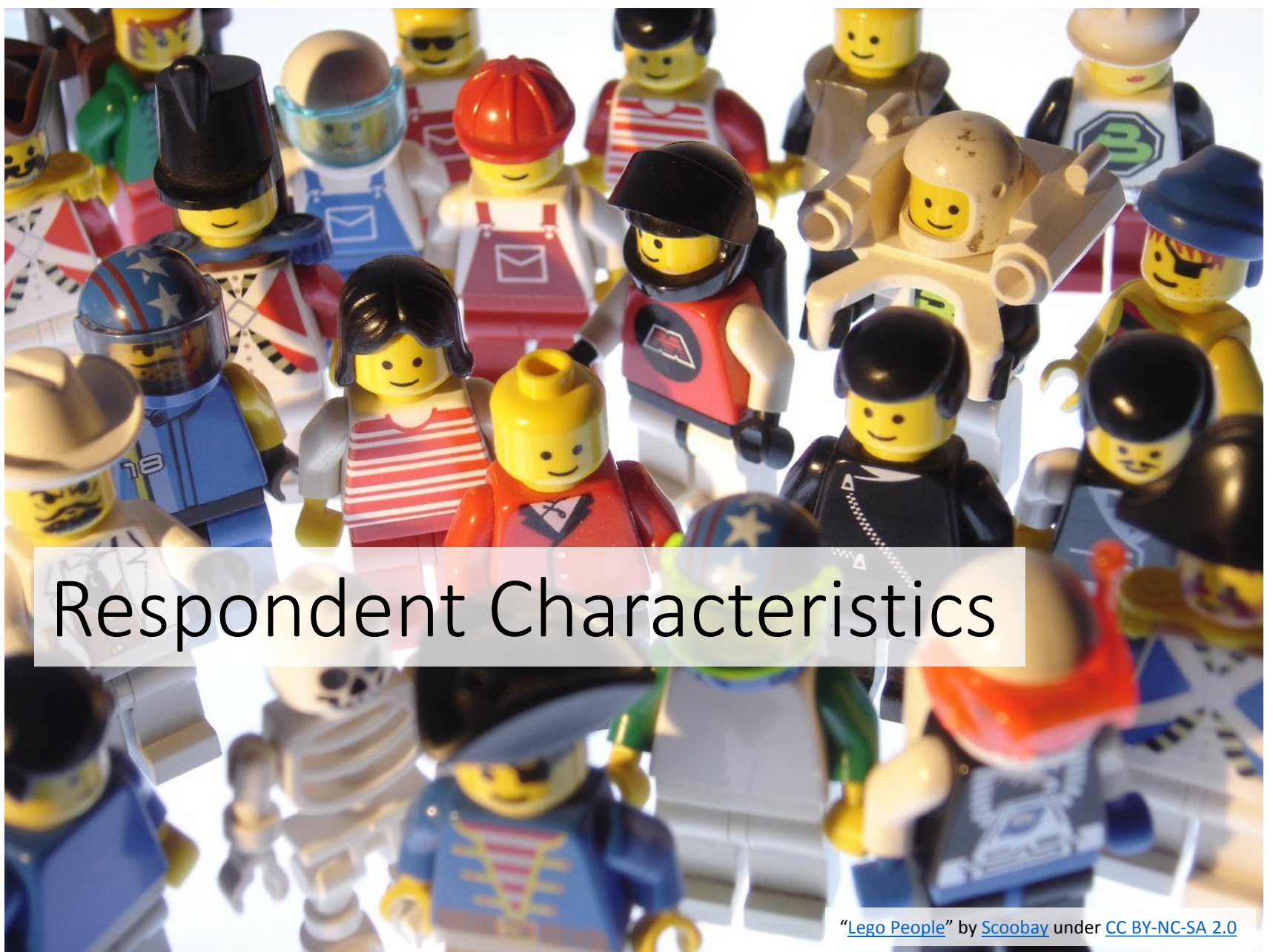
NDSA Web Archiving survey background

2011

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

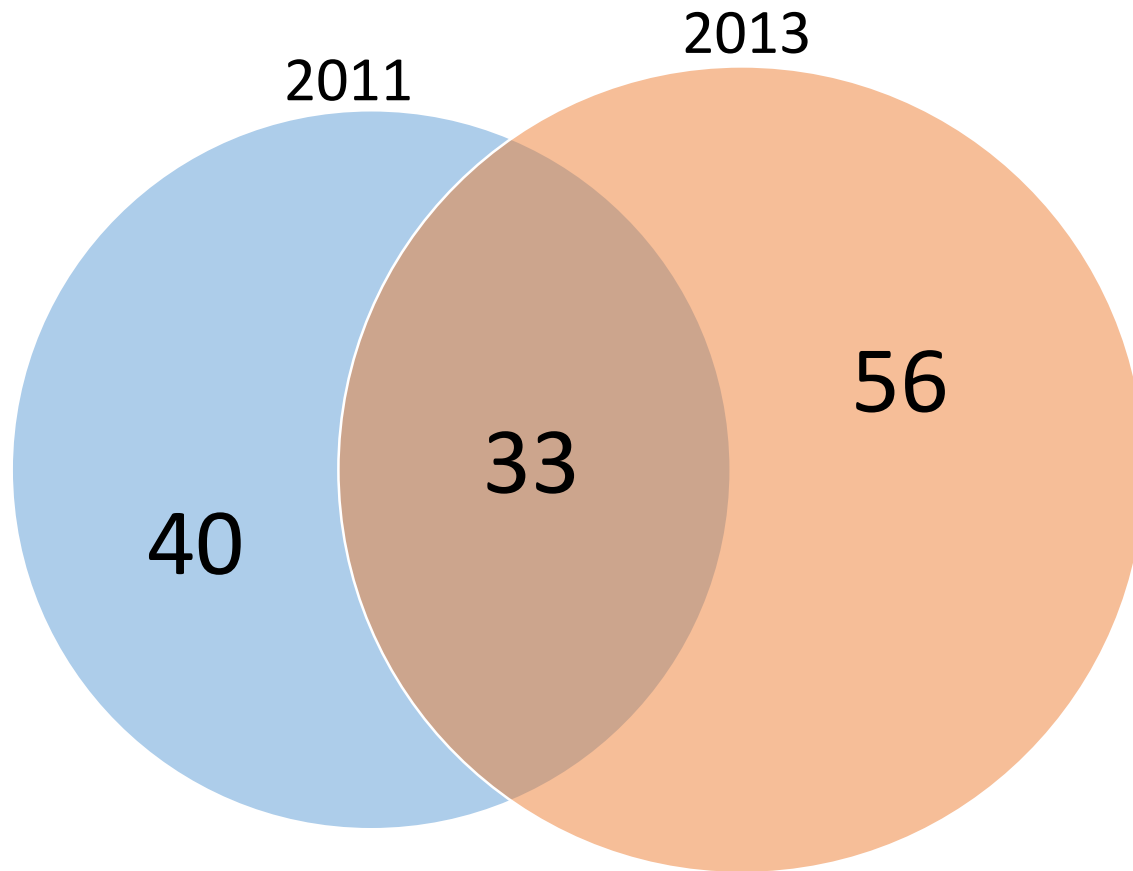
2013

- 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

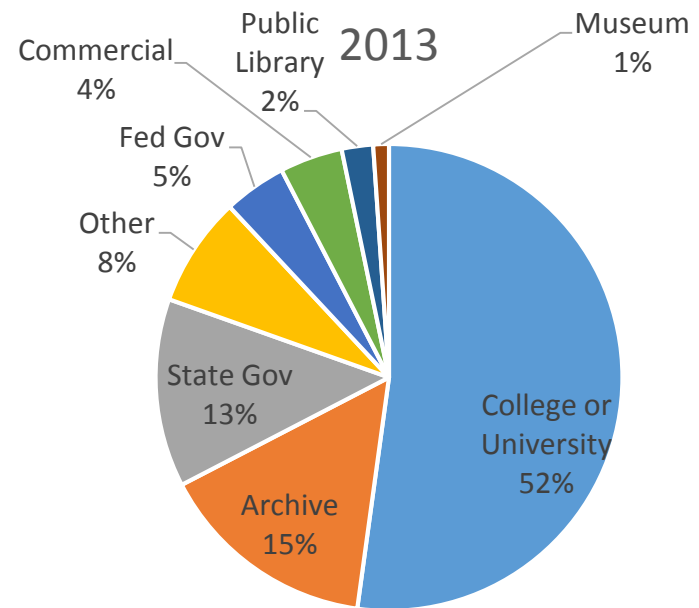
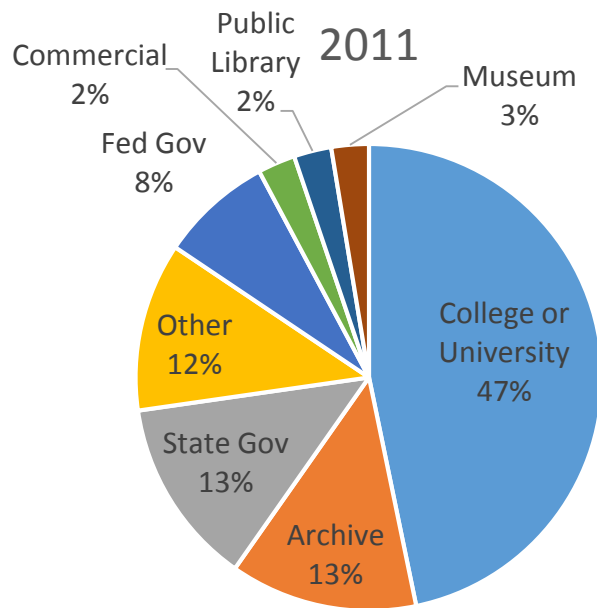


"Lego People" by Scoobay under [CC BY-NC-SA 2.0](https://creativecommons.org/licenses/by-nc-sa/2.0/)





less than half repeat respondents



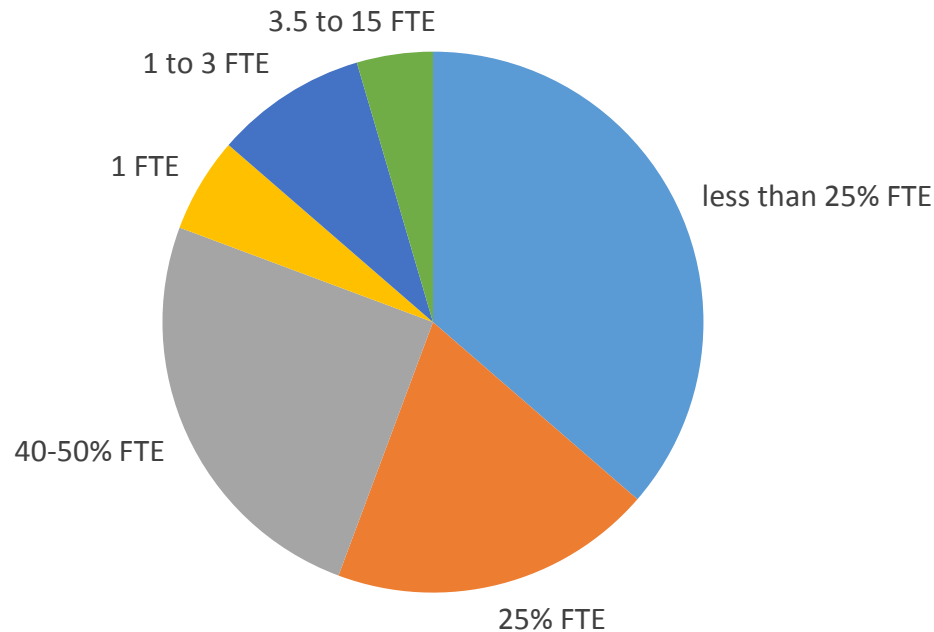
universities still make up most programs



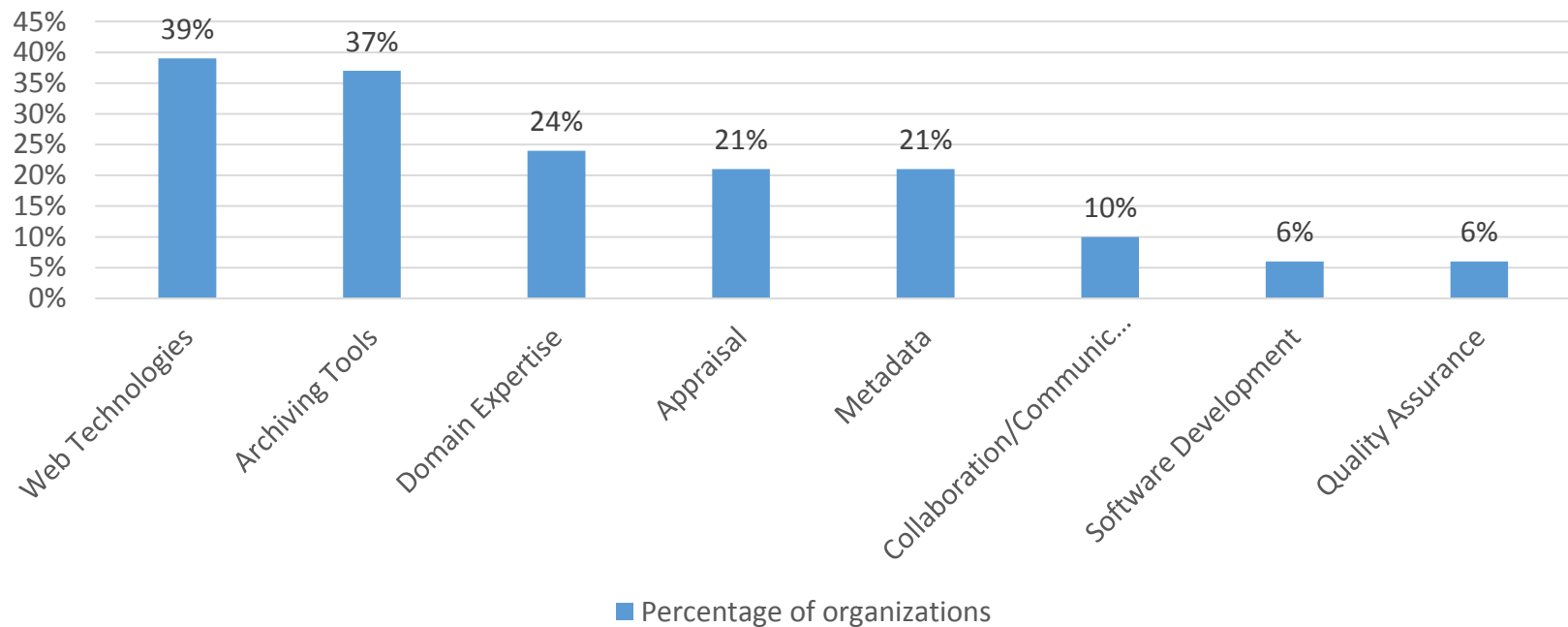
Archive-It and SAA top group affiliations

group	2011	2013
 INTERNATIONAL INTERNET PRESERVATION CONSORTIUM	8%	7%
	31%	33%
Web Archiving  Roundtable		45%
	72%	71%

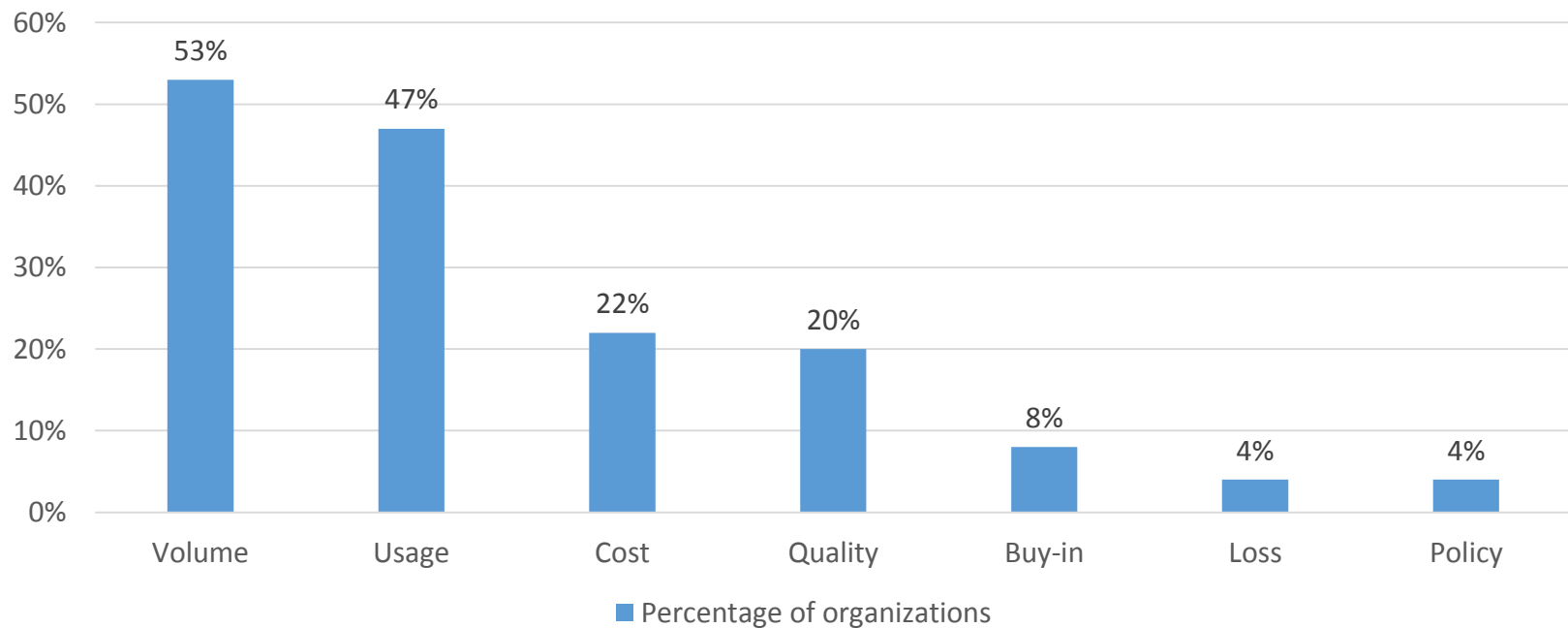
most programs are fractionally staffed



web/archiving tech savviness are key skills



data volume and archive use are key metrics

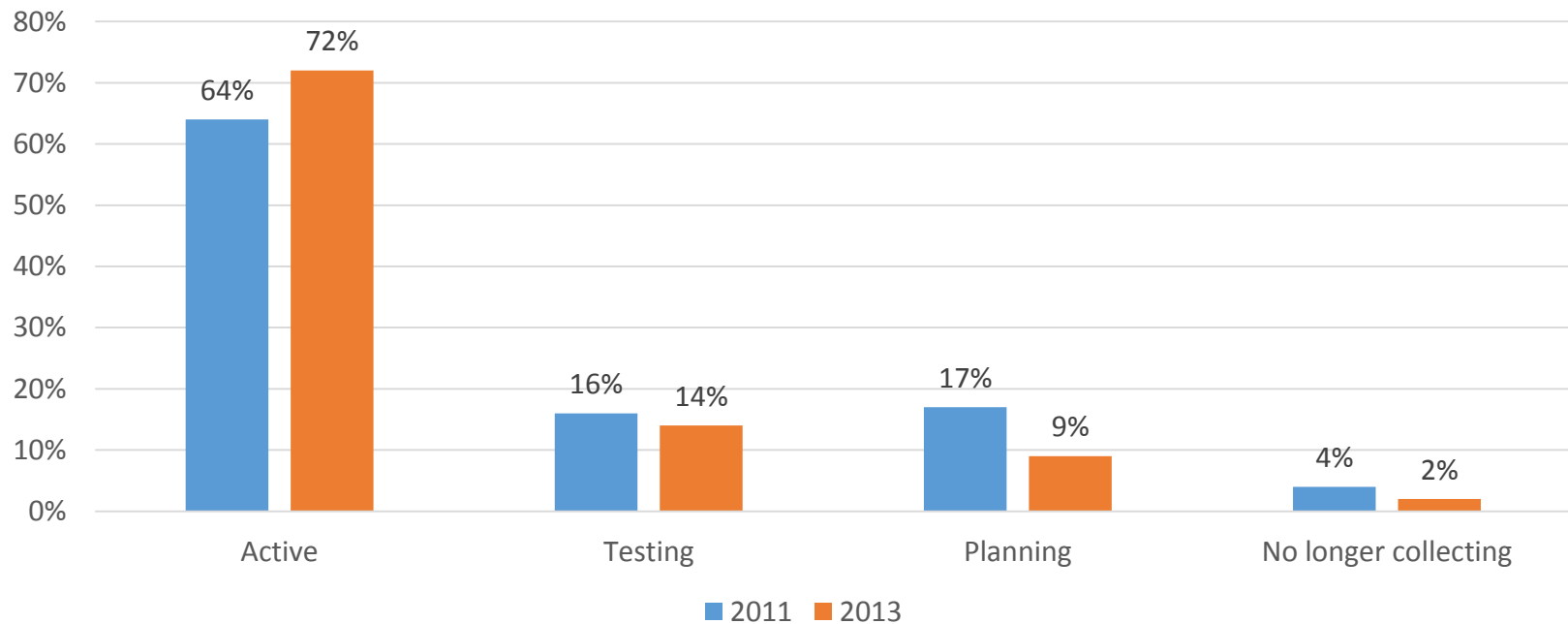




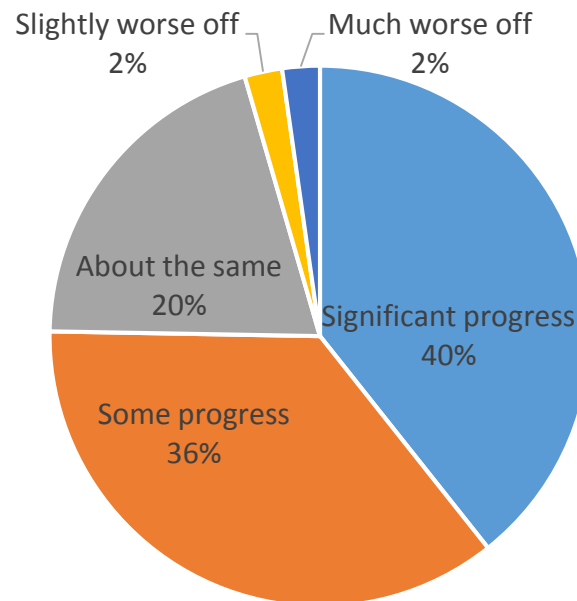
Maturity and Progress

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

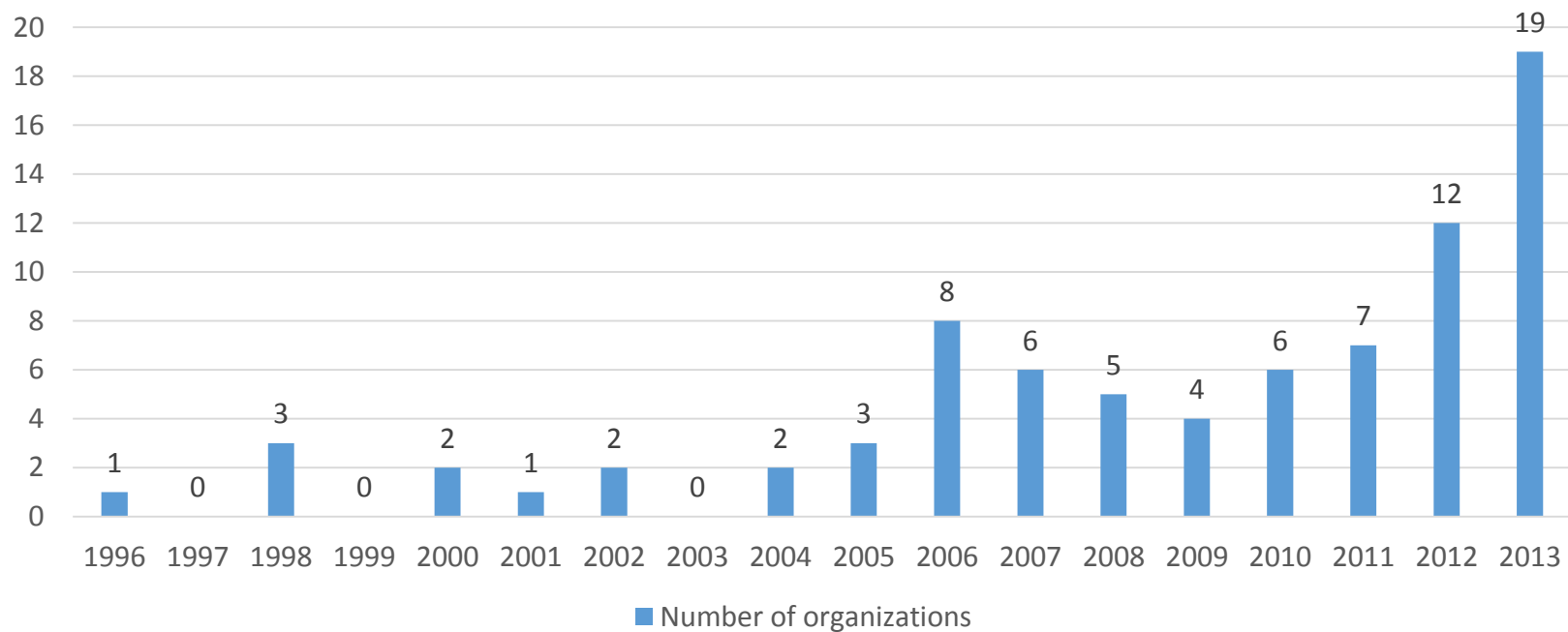
programs have matured slightly since 2011



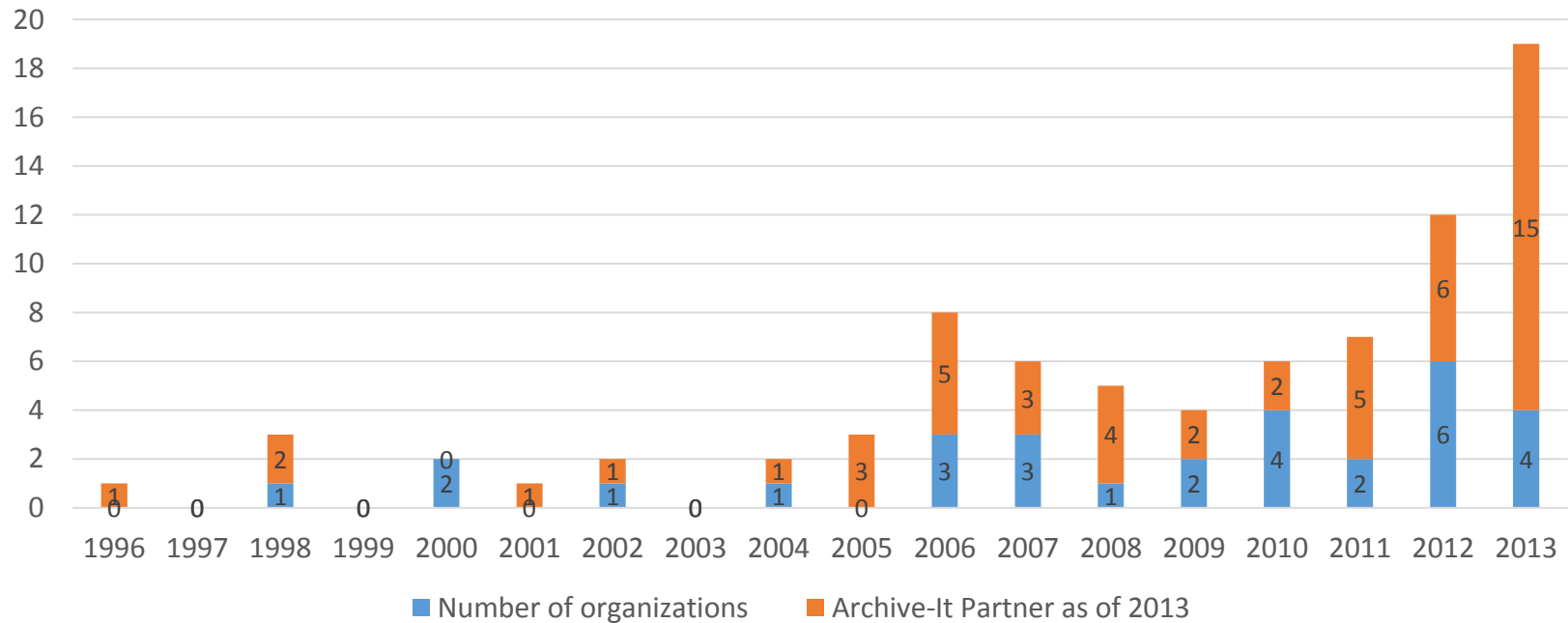
strong perceptions of progress since 2011



many new programs since 2011



two-thirds of them now use Archive-It

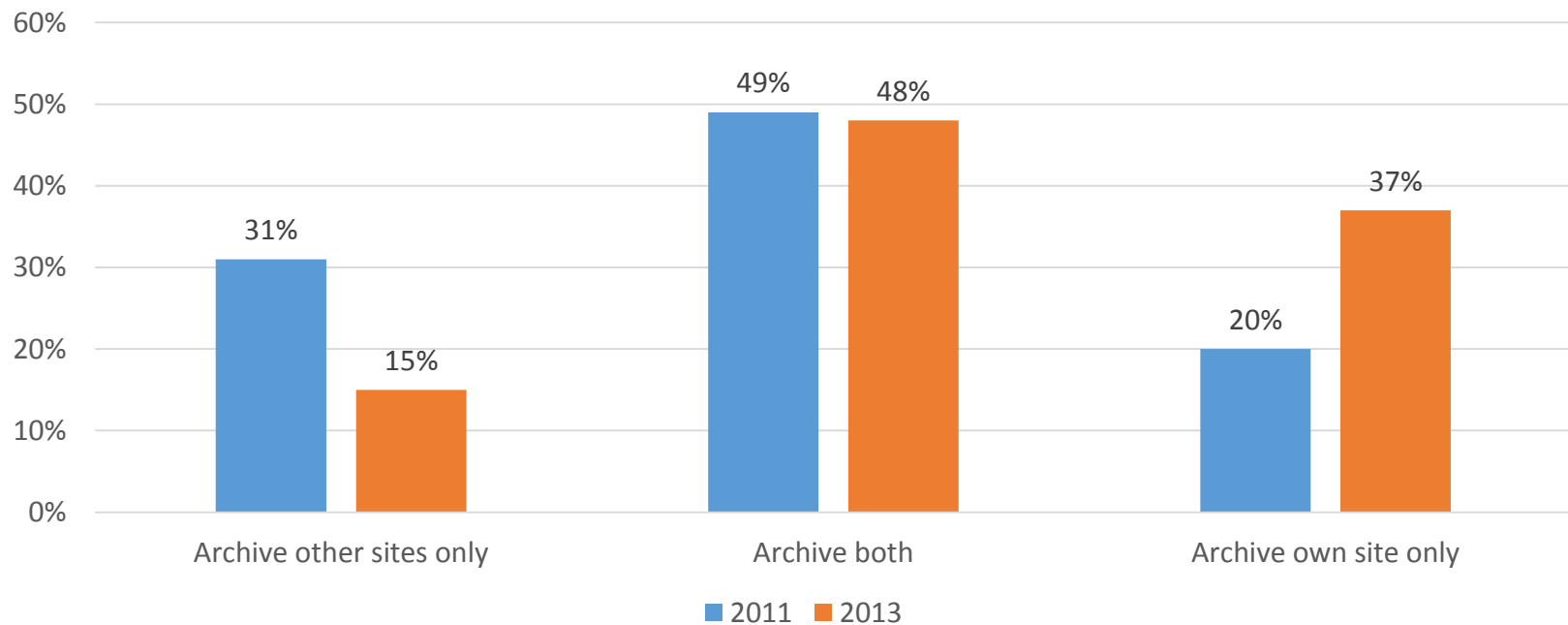




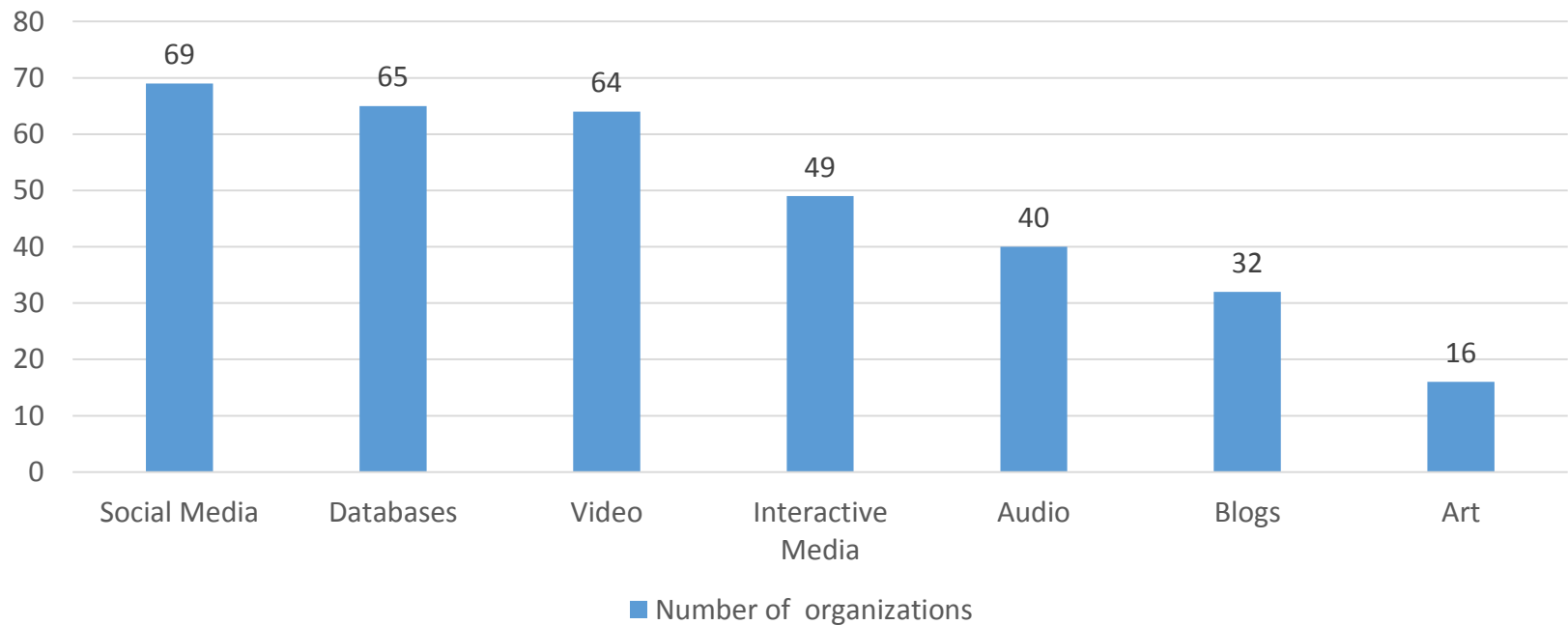
Archiving Focus

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

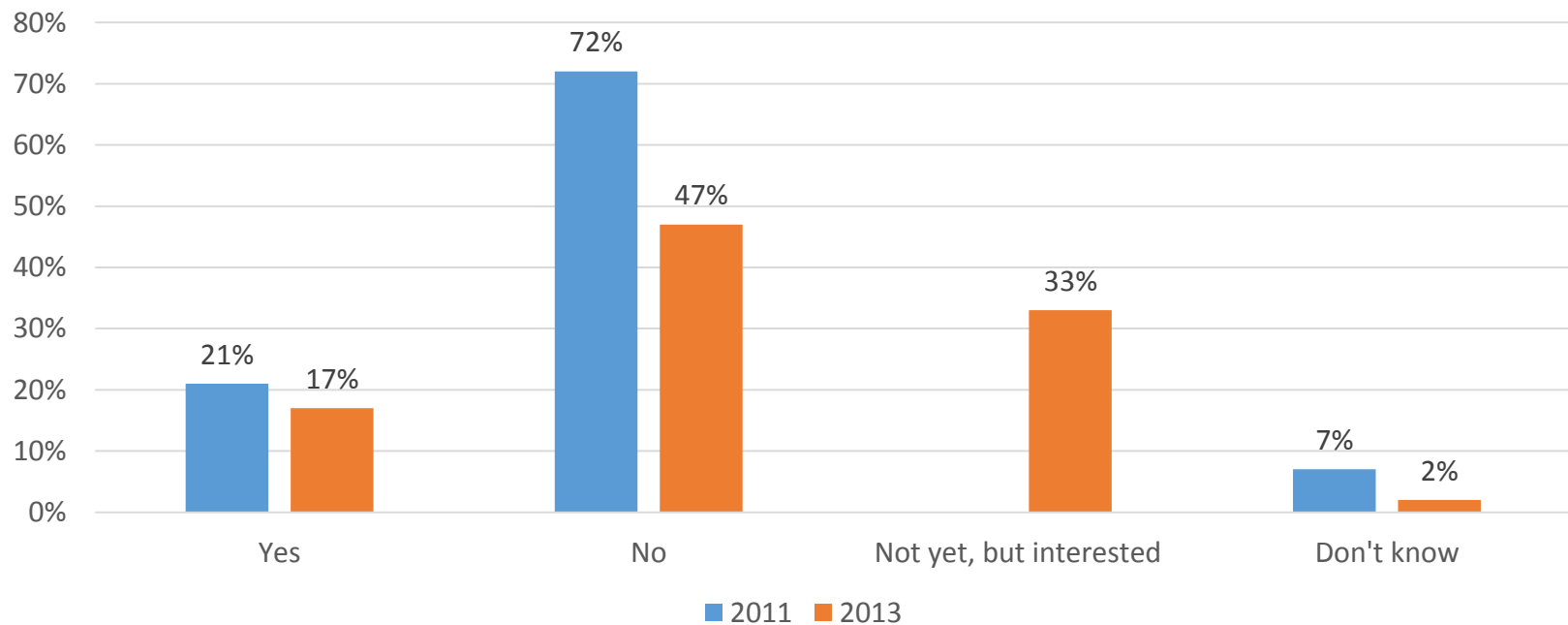
more programs are only self-archiving

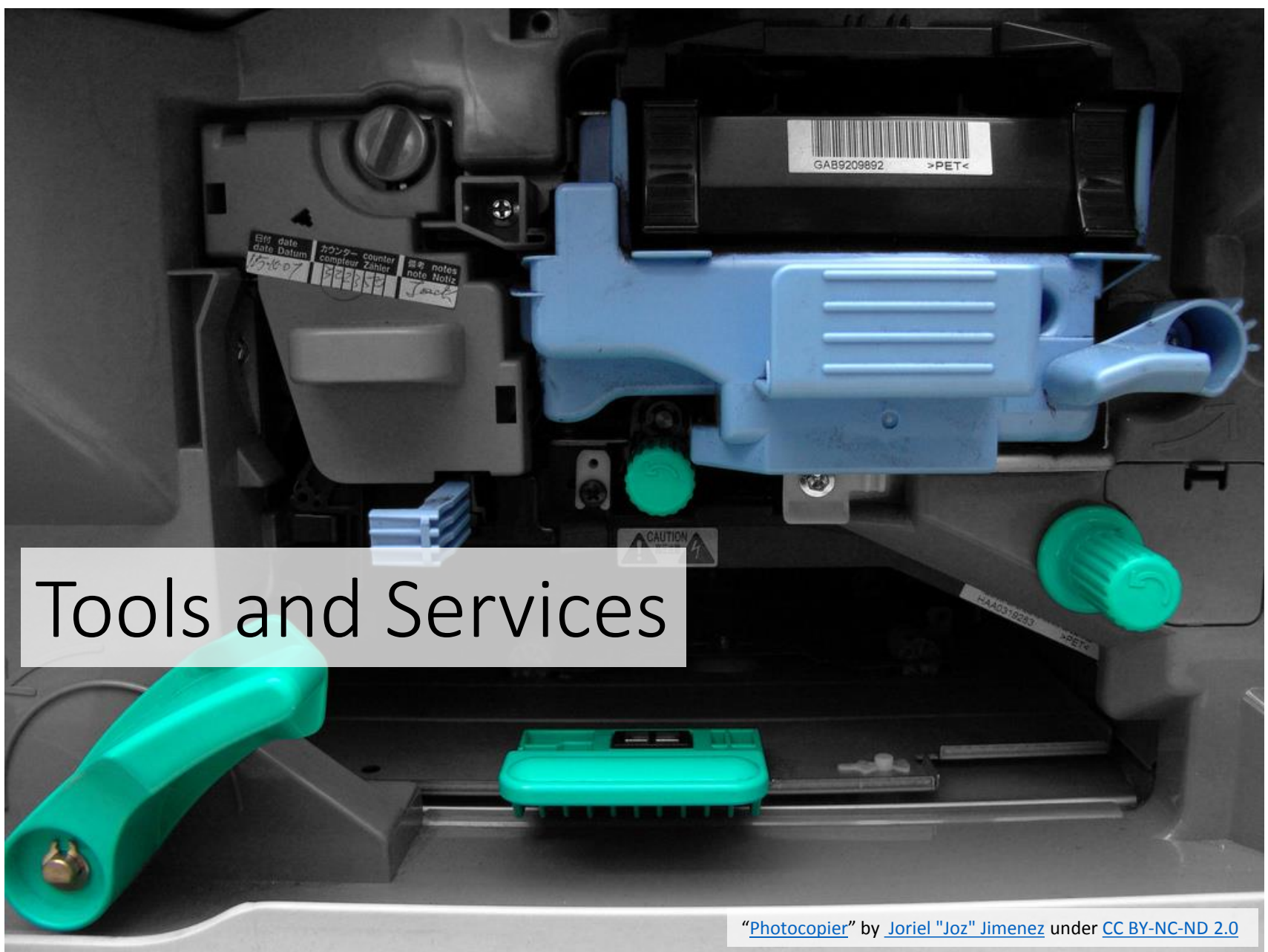


concern about social media, databases, video



untapped interest in collaboration

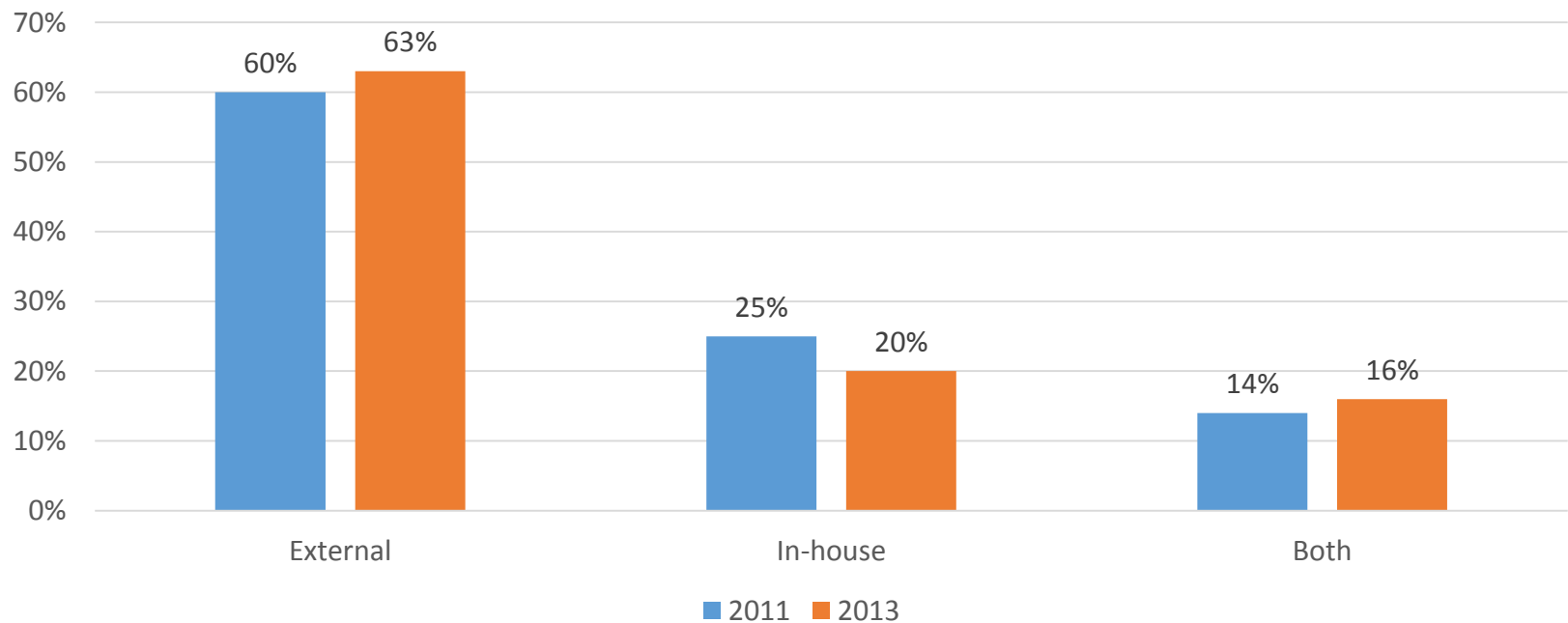




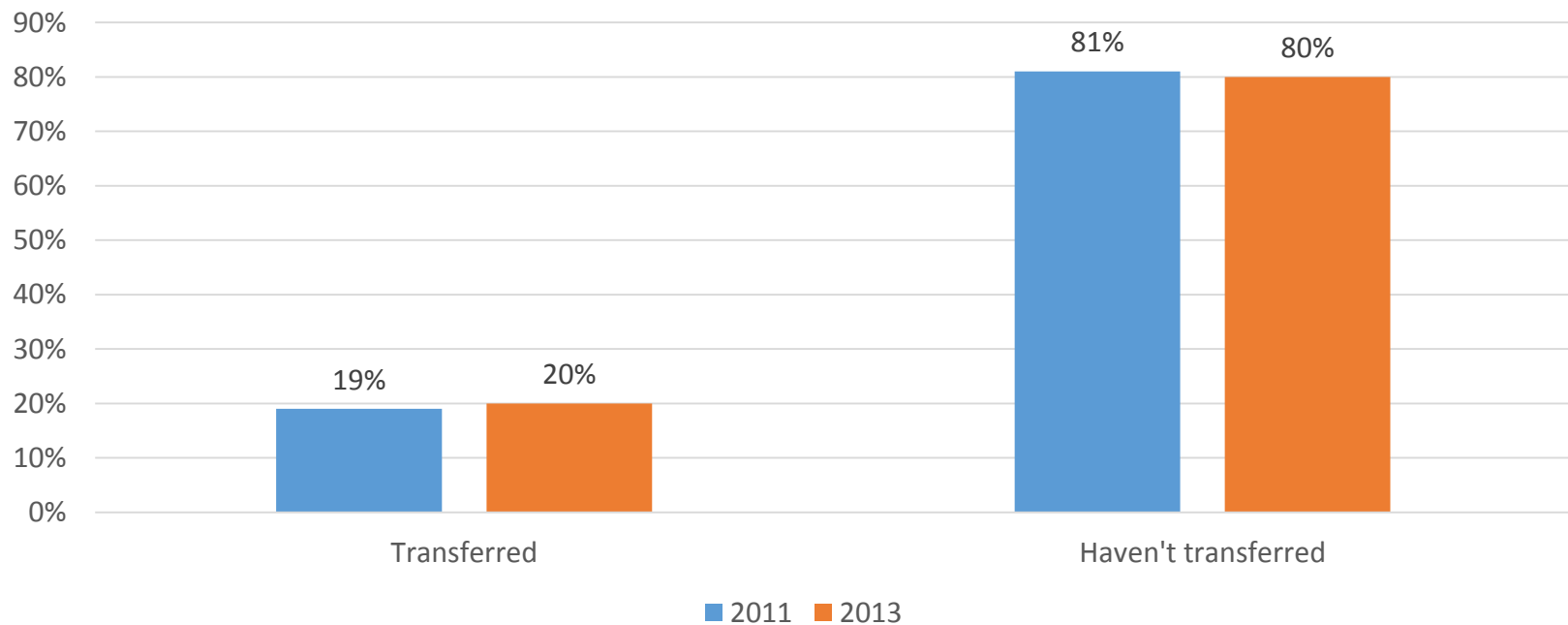
Tools and Services

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

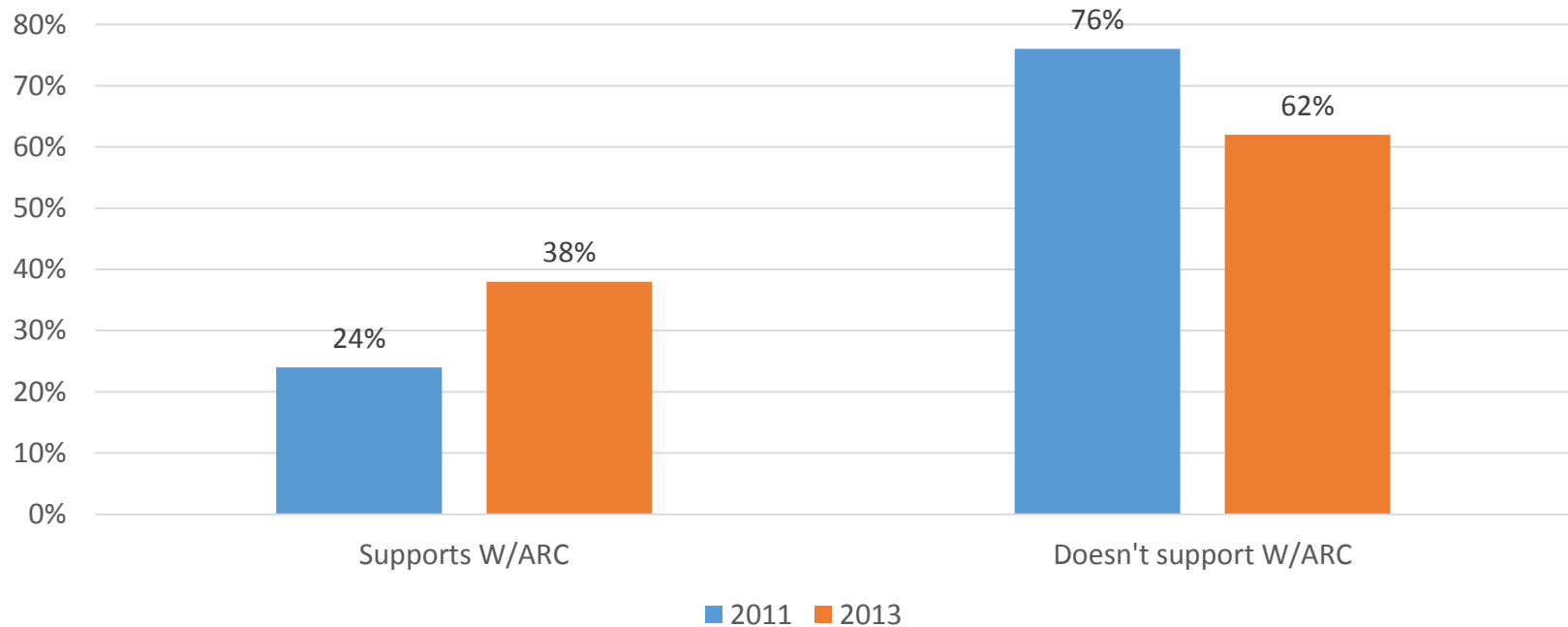
web archiving as a service still most popular



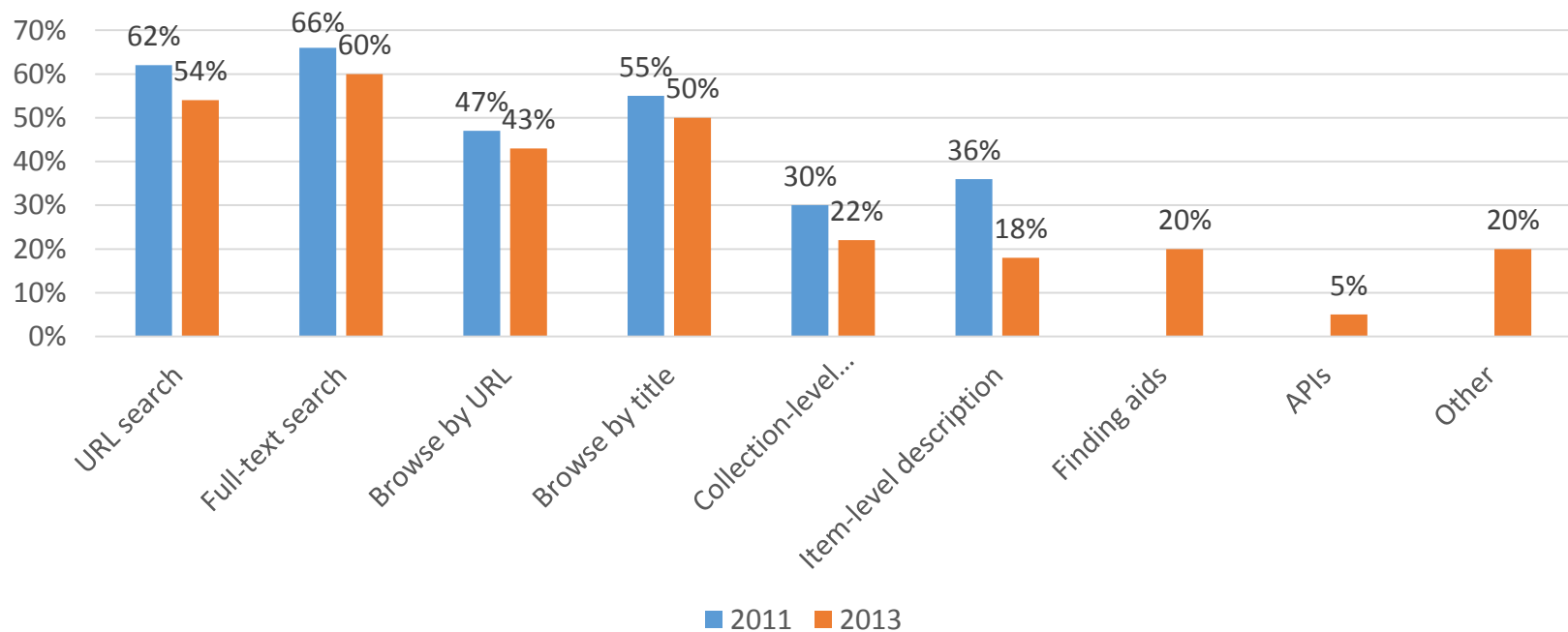
data not transferred from service provider



increased use of tools supporting W/ARC



less granular descriptive metadata

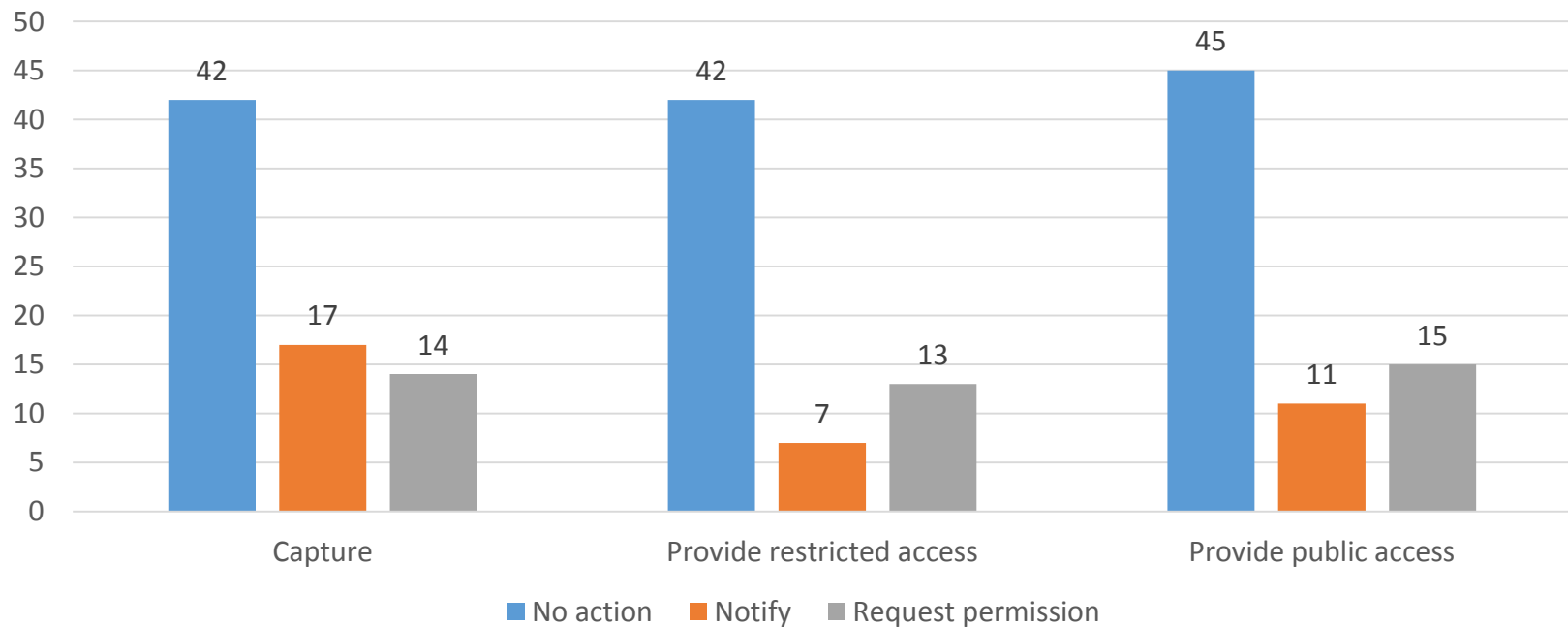


The background of the slide is a close-up photograph of three wooden crates. Each crate has a black line-art label. The first crate on the left has two upward-pointing arrows. The middle crate has a line-art of a wine glass. The third crate on the right has a line-art of an open umbrella. Below the umbrella icon, the words "Keep Dry" are partially visible. A semi-transparent white box with the text "Archiving Policies" is overlaid on the middle crate.

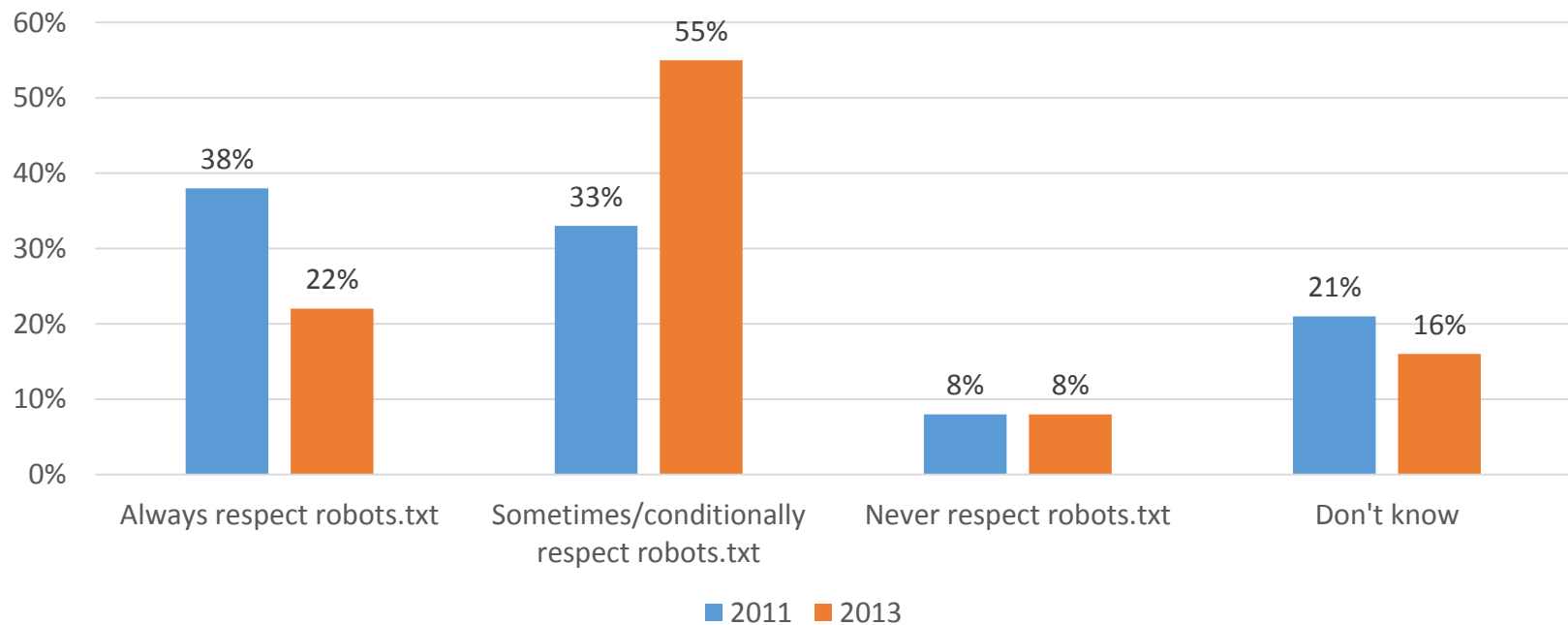
Archiving Policies

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

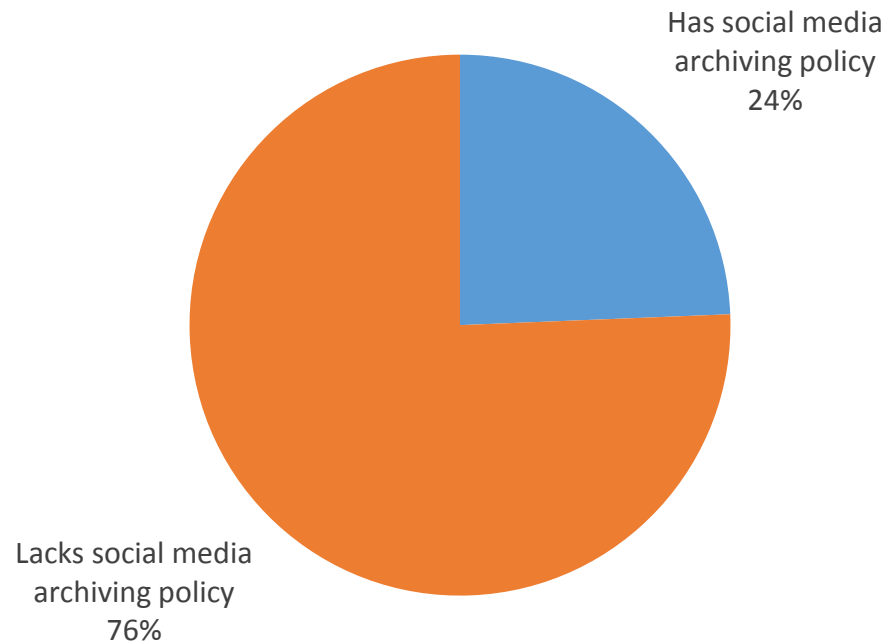
most don't notify or seek permission



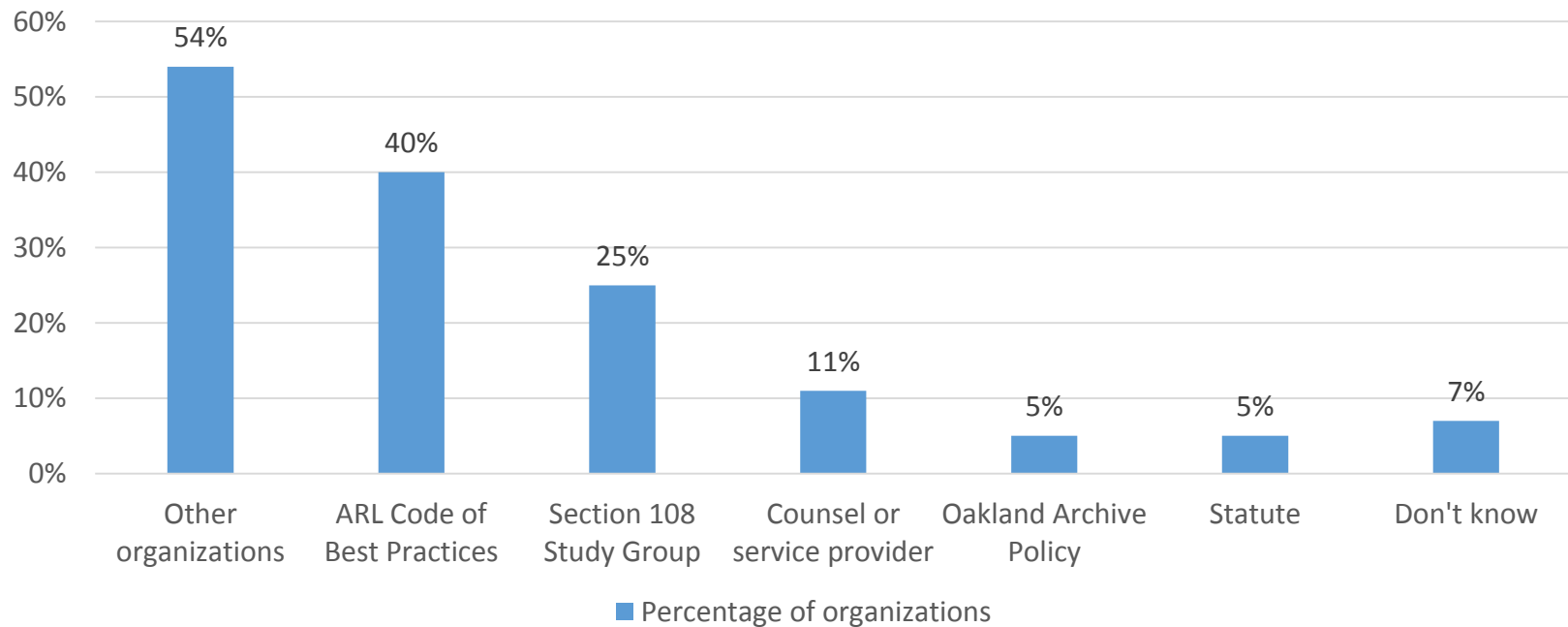
more conditional handling of robots.txt



social media archiving policies are uncommon



policies based on community practices



PREF

033.9732°N / 117.2846°W 2710ft 09:00:51

MAIL

POSITION - ALTITUDE - TIME



MAP

LOG



HORIZON

-00.8°

ANGLE

ELEVATION

+03.8°

ANGLE



Landscape Summary

AZIMUTH - BEARING



304° N56W 5404mils MAG

ZERO

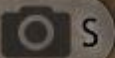
A-B

CAL



LENS

1.0X



"Mt Baldy from Box Springs Mountain wi Theodolite" by [signal mirror](#) under [CC BY 2.0](#)

profile of the average survey respondent

- university archive
- started in last three years
- Archive-It user
- ¼ FTE web-savvy archivist
- concerned w/ content capture, cost, and use
- broad level of description
- ambivalent about collaboration



"Container" by Glyn Lowe under CC BY 2.0

maturity and convergence

- maturity
 - 75% cite [some or significant progress](#) since 2011
 - 38% [started programs](#) since 2011
 - 8% more [programs in active status](#) since 2011
- convergence
 - 79% using [external service providers](#)
 - 81% devoting [½ FTE or less](#) to web archiving
 - 67% rely on [community practices](#) for policy-making
 - 13% more [using Wayback](#) since 2011

challenges and opportunities

- challenges
 - 53% concerned about data volume growth
 - 47% concerned about fostering access
 - more than 73% concerned about content capture
- opportunities
 - 33% interested but not yet involved in collaborations
 - 76% lack social media archiving policies
 - less than 23% of archived materials are described

implications and questions

- implications
 - web archiving **not (yet) a top institutional priority**
 - demand for **ongoing Archive-It technical investment**
 - U.S. web archiving **landscape is changing quickly**
- questions
 - how to **build institutional support?**
 - **collaboration** with whom and on what?
 - what's **not being archived?**
 - how well are we **curating what we do archive?**

THANK
YOU

Nicholas Taylor
[@nullhandle](#)

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