SOFTWARE CURATION: AN ECOSYSTEM OF USERS, TOOLS, AND SERVICES

#pasig 2019

#S9B

WHAT WE'LL COVER

- HOW IS SOFTWARE DIFFERENT FROM OTHER DIGITAL OBJECTS?
- WHY SHOULD WE CARE ABOUT SOFTWARE CURATION?
- WHAT IS SOFTWARE CURATION?
- ROLES IN THE SOFTWARE CURATION ECOSYSTEM
- GROUP DISCUSSION

HOW IS SOFTWARE DIFFERENT FROM DATA?

Software is **executable**, data is not.

Data provides evidence, software **provides a tool**.

Software is a creative work, data are observations.

Software faces **dependency-level bit rot**.

Software has a **shorter projected lifespan**.

-Daniel S. Katz, Kyle E. Niemeyer, Arfon M. Smith

"Software vs. Data"

WHAT IS SOFTWARE CURATION?

"Software curation encompasses the active practices related to the creation, acquisition, appraisal and selection, description, transformation, preservation, storage, and dissemination/access/reuse of software over short- and long- periods of time."

- Alex Chassanoff
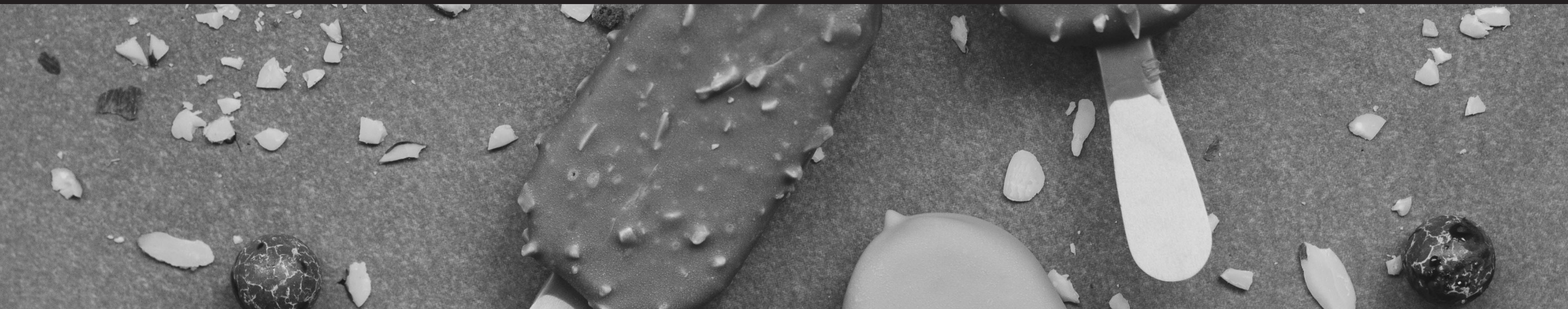"Building a Model for Software Curation"

Active Development

Publication and Active Use

Preservation and Long-term Reuse
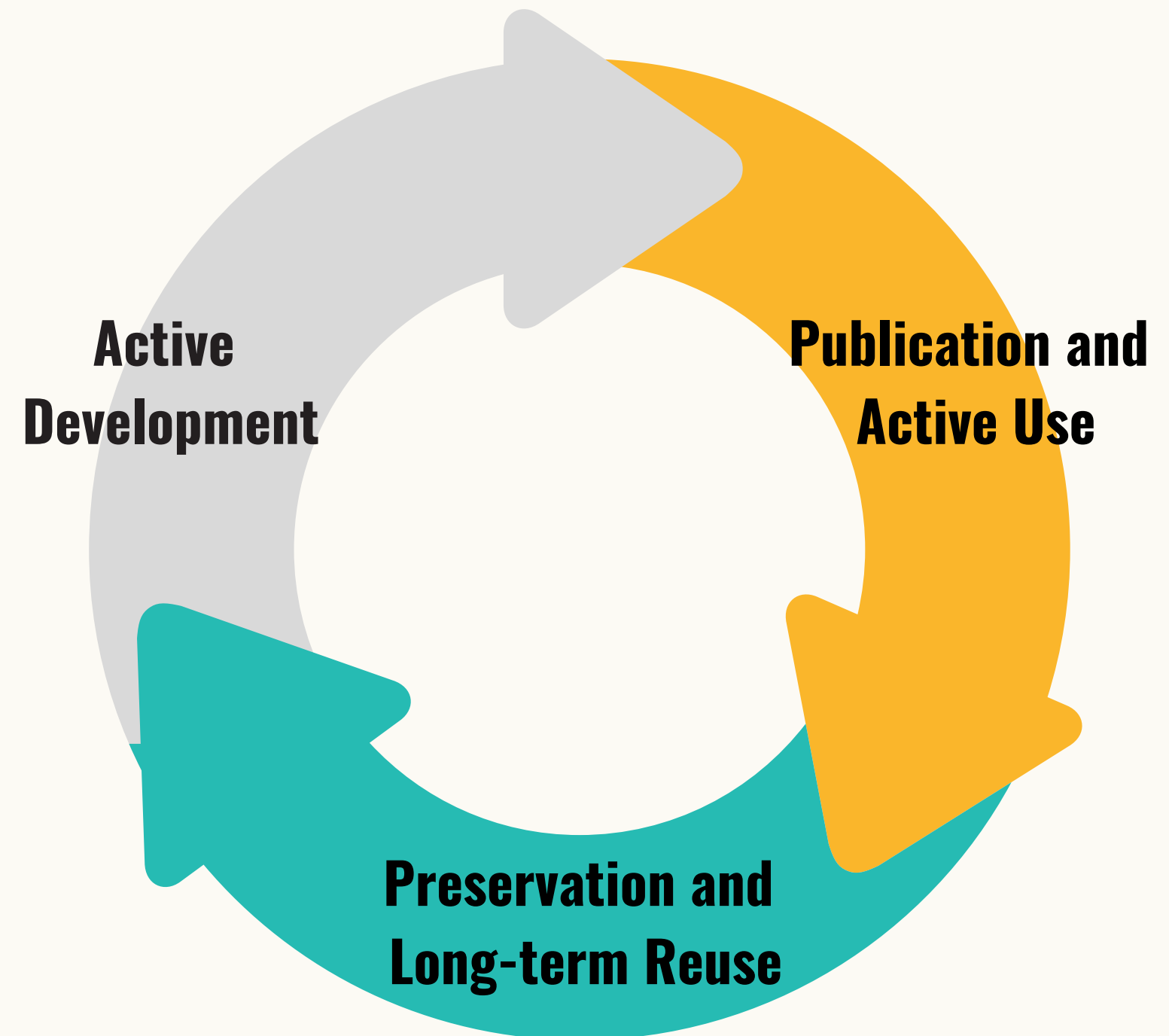
Imagine your favorite food.

If you wanted to share your favorite food with humans far, far in the future - what would you save?

a) recipe to make your favorite food **(the code)**
b) ingredients **(what's needed to run the code)**

Software is the recipe.
Software is also the ingredients.

SOURCE CODE V. BINARIES

SOURCE CODE IS HUMAN READABLE AND HAS TO BE COMPILED
BINARIES ARE COMPUTER READABLE AND ARE PRECOMPILED

- What we save or what we preserve is informed by our users are and what their reuse cases are.
- Time is another consideration.
- In the medium term, there are migration, reuse and access needs best served by binaries.
- Longer term, source code fragments may be the equivelent of parchment fragements we use to reconstruct some understanding of peoples of the past.

TIME TO SWITCH METAPHORS

LOOKING AT USER COMMUNITIES, TOOLS, AND SERVICES FOR SOFTWARE CURATION

USER COMMUNITIES

Groupings of users that share a similar context of use and reuse. Examples will focus on include: research software engineers, software curators, repository managers and digital archivists.

TOOLS

Software libraries and platforms that allow user communities to achieve their software goals - whether that is to create, to share, to assign attribution, to track provenance or to reuse.

SERVICES

A constellation of tools, community governance, staff, sustainability considerations and additional layers of mediation between a user community and a set of tools or digital objects.

SOFTWARE CURATION ECOSYSTEM

ECOSYSTEMS

USER COMMUNITIES, TOOLS AND SERVICES CAN OCCUPY DIFFERENT NICHES WITHIN A SOFTWARE CURATION ECOSYSTEM

The ecosystems metaphor is often used to informally describe the connections among organizations that share common or complementary features... information and resource flows connect organizations within organizational ecosystems in spite of the presence of diverse and sometimes even competing goals and agendas.

SOFTWARE CURATION NICHES

1 TRAINER
Selina Aragon
Software Sustainability Institute

2 DISCIPLINARY LIASION
Daina Bouquin
Center for Astrophysics
Harvard | Smithsonian

3 INFRASTRUCTURE PROVIDERS
Don Brower
Notre Dame University
Seth Anderson
Yale University

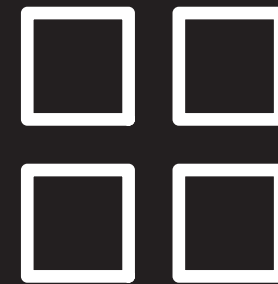**4** LEGAL ADVOCATE
Krista Cox
Association of Research Libraries

OUR DISCUSSION QUESTIONS

SHARING NICHE PERSPECTIVES

Q1: What does software curation mean to each of the panelists?

Q2: To what extent are software curation practices shaped by discipline-specific needs?

Q3: How do we present/discuss the evolving ecosystem of tools and services when the configuration of those tools and services will vary by user community?

SOFTWARE CURATION ECOSYSTEM

Q&A: LET'S TALK

POSSIBLE DISCUSSION QUESTIONS

What software curation niche or niches are you apart of?

When preserved materials and software are held in different archives/repositories and distributed to various platforms, how do we avoid building an ever-increasingly fragile network of dependencies? How is governance implicated in the scale and complexity of curating software?

How does the legal context of this work influence what we can and can't do? And what is being done to change in this area?

Basic question: What does the future of software preservation and curation look like?

SOFTWARE CURATION RESOURCES

PLEASE COME AND CHAT WITH US
WE WANT TO LEARN FROM YOU

Software Sustainability Institute

- Software Deposit Guidelines for Researchers: https://softwaresaved.github.io/software-deposit-guidance/
- Digital Preservation and Curation: https://www.software.ac.uk/resources/guides/digital-preservation-and-curation-danger-overlooking-software

Software Preservation Network

- Scaling Software Preservation and Emulation Infrastructure: http://www.softwarepreservationnetwork.org/eaasi/
- Fostering a Community of Practice: Software Preservation and Emulation in Libraries, Archives and Museums: https://www.softwarepreservationnetwork.org/fcop/
- Code of Best Practices for Fair Use in Software Preservation

SOFTWARE CURATION ECOSYSTEM

SOFTWARE CURATION RESOURCES

PLEASE COME AND CHAT WITH US
WE WANT TO LEARN FROM YOU

FORCE 11

- Software Citation Working Group: https://github.com/force11/force11-sciwg
- Software Citaiton Principles: https://www.force11.org/software-citation-principles

Research Data Alliance

- Software Source Code Interest Group: https://www.rd-alliance.org/groups/software-source-code-ig

Data and Software Preservation Quality Tool (PreQT)

- PresQT on OSF: https://osf.io/d3jx7/

Software Hertiage

- Software Heritage Archive: https://archive.softwareheritage.org/
- Software Heritage Development Documentation: https://docs.softwareheritage.org/devel/

SOFTWARE CURATION ECOSYSTEM