

# Sustainable Web Archiving at Scale: An Introduction

June 21-22, 2021



## Worksheet: Defining Scope and Developing a Collecting Statement

Use the prompts on the following pages to identify the specifics of your collecting mission and design a new web archive for your institution. This workshop exercise is designed to envision a new collection of materials published to the web: how those materials are acquired, managed, and made accessible.

### Scope considerations:

- What will this collection(s) consist of?
- Review the mission of the collecting organization and close constituents, and consider how a proposed collection further supports that mission.
- Identify your target audience for use of the collection and consider potential secondary uses/users.
- If you are also collecting print and non-digital materials, consider how this born-digital collection might complement those resources.
- What does not fit within this collection? Cite specific exclusions.

### Policy considerations:

- Why do we need this program/project?
- What is the end-goal objective for initiating web archiving as a collection development priority within my institution?
- Rights issues are rarely clear-cut and risk assessment is often necessary.
- Will we seek permission from site owners to archive their site?
- Will we provide open/public access to our web archive collections?
- What is the timeframe for initiating a program of web archiving?

# Sustainable Web Archiving at Scale: An Introduction

June 21-22, 2021



## Resources/Stakeholder considerations:

- Which staff will be responsible for this work and what is their level of expertise?
- What is the budget for the implementation of this collection?
- What frequency of capture will be sustainable for these collections (or will the objective be solely one-time capture)?
- Does the organization presently use a web harvesting technology in some capacity?
- What is our capacity to curate and manage the collection during the entire lifecycle?



## Complete the following prompts to compile your collecting statement:

The goal of this collecting effort is to preserve born-digital materials that pertain to:

The institutions and/or stakeholders of this web archiving program/project are:

This program aligns with the institutional/stakeholder mission in this way:

The target audience for use of collected material is:

# Sustainable Web Archiving at Scale: An Introduction

June 21-22, 2021



## DPOE-N

Digital Preservation Outreach  
& Education Network

Web material will be selected by (designated staff, faculty, public, etc.):

This collecting effort will respect web material creator/owners' intellectual property rights by (select one):

- Obtaining written consent prior to acquisition
- Providing "opt-out" recourse prior to and/or after acquisition
- Other

Will these materials be made accessible (publicly/on-site only)?:

Describe what successful use of these web archives might look like:

What is specifically not in scope within this collecting mission (an example being social media)?: