Problem Statement
Cultural heritage organizations have web archives that are not being leveraged very well at this time.

Many web archives harvest and store web-published materials in a manner geared towards preservation rather than discovery.

Identifying content-rich documents among the wealth of materials available via web archives offers an opportunity to expand existing collection development policies to include a new class of materials.

Research Questions
- What role do web-published documents and publications play in developing collections in the broad categories of institutional repositories, state government documents, and publications from the federal government?
- What are the characteristics of web-published documents and publications that help content selectors identify them for inclusion in their local collections?

Methodology
Individuals (n=11) with experience in collecting publications from the web participated in semi-structured interviews.

Interviews were transcribed using Express Scribe transcription software.

After completion, transcripts were coded and analyzed using NVivo 12 Plus qualitative analysis software to identify themes and trends related to the research questions that have been posed.

Findings
The selection of publications from web archives is primarily determined by the governing collection development policies.

Content selectors can typically determine whether or not a publication fits the scope of a collection after viewing the first few pages of the document.

Distinguishable features of publications are indicative of inclusion for collection (e.g. title page, graphs & charts, and/or unique report numbers).

Acknowledgements
This research was supported by the Institute of Museum and Library Services under the award LG-71-17-0202-17.