

**LOC Recommended Formats Statement
2021-2022**

i. Websites		
B. Formats	1. The Library, and other organizations involved in web archiving, are preserving web content in the Web ARChive (WARC) format	B. Internet Archive's ARC IA format, a precursor to the WARC format
C. Delivery Method	1. Capture using tools that produce non-proprietary output, to conform with standard formats and requirements	1. Transmission of WARC or ARC IA files created by web content producers or other archiving organizations
D. Metadata	1. Refer to the WARC ISO-standard specification for mandatory and recommended metadata fields 2. When displaying archived content, the following should be clearly indicated: <ol style="list-style-type: none"> a. archiving institution, b. dates and time of capture, c. statements about functionality within the archive to distinguish from the live site 	1. The ARC IA should be named in a manner that easily identifies the archiving institution (see WARC standard for recommended naming conventions)
E. Technological Measures	1. Websites should not contain measures (such as content behind logins or only accessible through search functions) that control access to or prevent capture of the digital work. 2. Robots.txt restrictions should be set so as not to block crawlers from capturing important content, such as image and style sheets, which allow for replay of the site as it looked at the time of capture.	1. Tools currently available cannot capture all web content, so certain types of web content may not be preservable through web capture at this time. These include: <ol style="list-style-type: none"> a. Multi-media rich content b. Streaming media c. Deep web content d. Databases

OLD VERSION DO NOT USE