

WADL Submission 2017

Author Name: Ross Spencer

Job Title: Digital Preservation Analyst, Archives New Zealand

Submission Type: Poster/Demo + Lightning talk

Poster Title: HTTPreserve – Web Preservation in Documentary Heritage

Summary:

The paper paradigm is still prevalent in the government archive and the government agency. It will be for a good number of years yet as they wrestle with appropriate time-scales for the disposal of records (destruction, or transfer to an archive). Most records in most jurisdictions will be transferred between 20 and 30 years.

The web, much older than these timescales and will exist in one form or another in the records either end.

This raises questions as to whether a government agency can preserve the public record in its completeness. A hyperlink that provides evidence on how a decision was made may not be transferred through normal practices and will instead need to be archived and monitored through mechanisms developed by the web-archiving community. Some links may already be lost. Some may only exist in places like The Internet Archive. We do not have adequate tools or workflows to understand the extent of bit-rot that may exist in agencies records.

This poster and demo will demonstrate a newly developed suite of tools called HTTPreserve (<https://github.com/httpreserve>) that may fill the gap that helps government agencies, and other archives to develop the appropriate processes for preservation of links in documentary heritage.

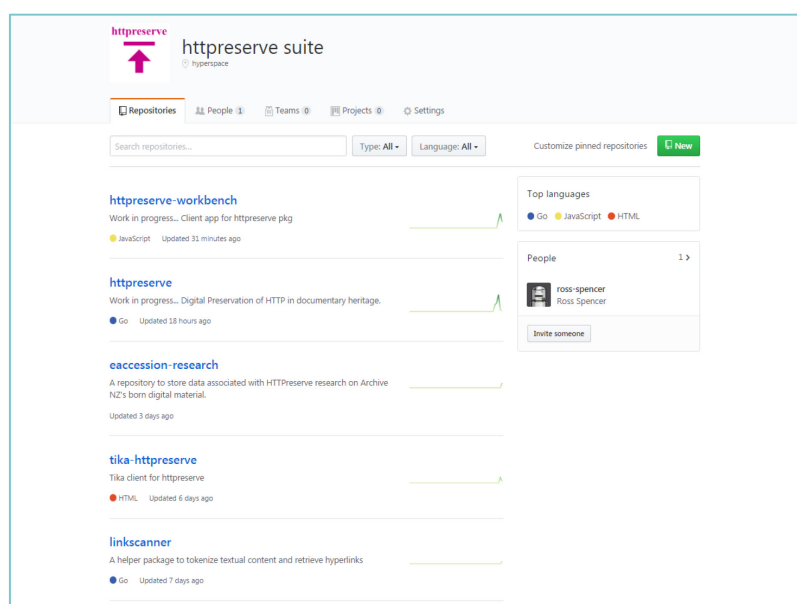


Figure 1: HTTPreserve Suite on GitHub.com