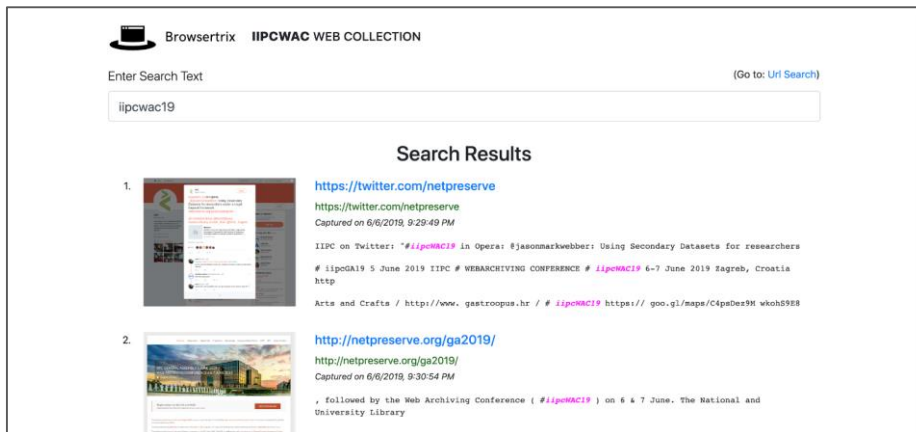
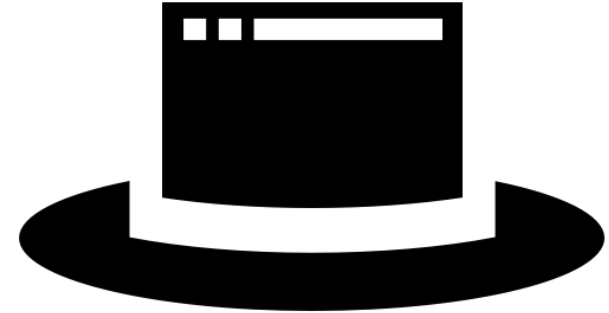


A New Browser-Based Crawling System from Webrecorder



[https://tinyurl.com/
browsertrix](https://tinyurl.com/browsertrix)



Browsertrix

- Fully Docker-Based
- Browser Crawling + In-Page Behaviors
- Uses pywb for capture and replay
- Screenshot creation (capture + replay)
- Instant text extraction while crawling
- Full-text search via Solr
- CLI and web UI interfaces
- Browser Profiles via Docker Images