Web Not For All
A Large Scale Study of Web Accessibility

Rui Lopes¹, Daniel Gomes², Luís Carriço¹

¹ LaSIGE, University of Lisbon
² FCCN
The Web is the biggest information source for Mankind. Decentralised architecture made it blossom.

Humans (and computers!) contribute to information production and consumption, leading to ~45B Web pages.
Context

- Growth of users contributing and interacting with the Web leads to significant diversity of users, including people with disabilities.

- The openness and decentralisation of the Web leads to an uncontrolled quality check of Websites’ usability (and accessibility).
What is the state of accessibility on the Web?
• It is known that Web accessibility adequacy is often **far worse** than desired.

• Studies tend to focus on a *restricted* (small) set of Web sites.

• Do *macroscopic properties* of Web accessibility emerge from analysing at a large scale?
Experiment

background

- The *Portuguese Web Archive* initiative periodically crawls contents from the Portuguese Web (.pt and others) for future preservation.

- Services are built on top of crawled collections: search (end users) & analysis framework (researchers).
Methodology

data acquisition - obtaining the document collection

- Collect a sufficiently large portion of the Web, yet representative (e.g., national Webs)
- Spider traps handled gracefully
- Boostraped with 200,000 Website addresses from the .pt TLD
- Collected March/May 2008
Methodology

data acquisition - evaluation process

• Implementation of 39 WCAG 1.0 checkpoints yield pass, fail, warn. (collection previous to WCAG 2.0 TR)

• Overcome computational effort with Hadoop cluster, streams, caching, etc.
Methodology

data analysis

- Failure rate, 3 criteria:

\[
rate_{\text{conservative}} = \frac{\text{passed}}{\text{applicable}}
\]

\[
rate_{\text{optimistic}} = \frac{\text{passed} + \text{warned}}{\text{applicable}}
\]

\[
rate_{\text{strict}} = \frac{\text{passed}}{\text{applicable} - \text{warned}}
\]
Results

• 28M Web pages were evaluated. (58%)
• 21GB evaluation data collected for analysis.
• 40B HTML elements evaluated. (~1500/page)
  • 1.5B elements passed. (56/page, 3.89%)
  • 2.9B elements failed. (103/page, 7.15%)
  • 36B elements warned. (1291/page, 89%)
Results

rates versus page count distribution

conservative

optimistic

strict
Results

rates versus page complexity (# HTML elements)

conservative

optimistic

strict
Discussion on the results

- Large scale confirms predictions of small scale studies - *the Web is still not for all.*
- Smaller Web pages tend to have greater accessibility quality.
- Nature of *warnings* is more striking than expected, completely different interpretations.
- Automated evaluation is just the beginning.
Discussion

on the limitations of the experiment

- HTML structure vs. content rhetorics.
  (CSS & Javascript can change it all)

- Collecting the Web is hard.
  (deep Web - AJAX & forms -, infinite generation, robots.txt, etc.)

- Scaling evaluation & analysis processes is hard.
  (evaluation streamability, resource inter-dependencies, billion node graphs, etc.)
Conclusions

- Large scale accessibility evaluation of the Portuguese Web.
- Re-confirmed studies at the large.
- Educating developers & designers about warnings is crucial for accessibility success!
- Automated evaluation is just the start. Always need for expert & users evaluations.
Ongoing Work

we're still at the tip of the iceberg

• Linking properties (ranking vs. accessibility)

• Evolution of accessibility compliance in time (different document collections)

• Cross-cuts: gov, e-com, personalisation, etc.

• Developing countries (Portuguese speaking African countries)
Ongoing Work

help wanted from community!

• Making available evaluation datasets (e.g., Linked Data). Ours and yours!

• Larger document collections.

• Transforming warnings into failures with machine learning.
Thank you!

rlopes@di.fc.ul.pt