



Web Page Clustering for More Efficient Website Accessibility Evaluations

Justyna Mucha¹, Mikael Snaprud^{1,2}, Annika Nietzio³

¹University of Agder, Norway,

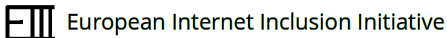
²Tingtun AS, Norway,

³Forschungsinstitut Technologie und Behinderung (FTB)
der Evangelischen Stiftung Volmarstein, Germany

15.07.2016

About the research

Research done at the University of Agder is connected to the European Internet Inclusion Initiative (EIII) and has been supervised by the company Tingtun AS.



- 1 Web Accessibility
 - EIII Accessibility Evaluation
- 2 Solution
 - Page Accessibility Profile
- 3 Results
 - Example
- 4 Future steps
- 5 Summary

Web Accessibility

- ✦ Access by everyone regardless of disability
- ✦ Removal of the accessibility barriers
- ✦ 10% to 15% of the total European population affected by a disability, in the United States, one out of every five citizens (2015)
- ✦ More and more public services go online

| | | | | |
|---|---|---|---|--|
| Blindness/ Screen Reader | Low Vision | Hearing Impaired | Mobility/ Keyboard- Only | Cognitive (Dyslexia, Non-native English) |
|  |  |  |  |  |

1

¹ source: ScienceDirect blog

- ✦ Proposed European Directive to make the websites and mobile apps of public sector bodies more accessible, especially for the blind, the deaf and the hard of hearing (3rd May, 2016)

EIII Accessibility Evaluation

- ✦ Tests are based on Web Content Accessibility Guidelines 2.0 Level A and AA, developed by the auto-wcag community group hosted by W3C
- ✦ Time-consuming process
- ✦ All pages or a random sample of 600 pages is checked



Evaluation Bottlenecks

- ❖ Crawling of websites
- ❖ High computational costs of evaluation
- ❖ Large number of websites already existing and more to be created

Research Purpose

- ❖ How to reduce the number of pages?
- ❖ Need for more focused way of selecting representative pages for assessment
- ❖ Key features of a website as clustering determinants
 - ▶ URL – Hernández et. al., 2012
 - ▶ DOM structure – Harper et. al., 2015
 - ▶ Page accessibility profile

- 1 Web Accessibility
 - EIII Accessibility Evaluation
- 2 **Solution**
 - **Page Accessibility Profile**
- 3 Results
 - Example
- 4 Future steps
- 5 Summary

A single web page is represented as a vector of n accessibility tests outcomes that have been applied on the web page:

The Rise and Rise of Smart...

by Mark Jenkins



Smart! The new buzz word for urban development in the 21st century. Smart cities, smart communities, smart society. In the 18th and 19th centuries, the industrial revolution created the wealth to build roads and railways that brought people closer together. 150 years later mobile communications

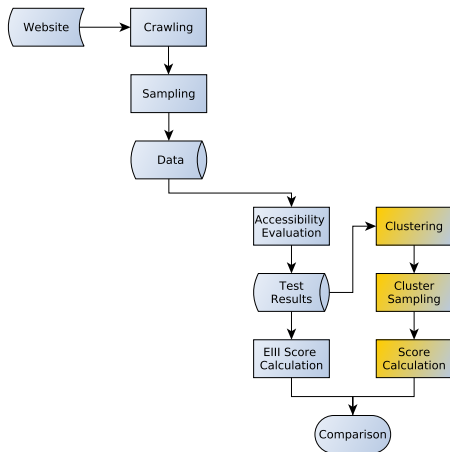
meant that people didn't need to be, well, mobile anymore. And now in 2015 we're already well into the next phase of (im)mobility with the latest generation of smart phones and tablets connecting to a myriad of electronic services allowing us to run our lives from the palm of our hand, literally. [Continue Reading](#) »



$$\mathbf{p}_i = \begin{bmatrix} 0 \\ 3 \\ 1 \\ \vdots \\ x_n \end{bmatrix}_{n \times 1}$$

Suggested enhancement

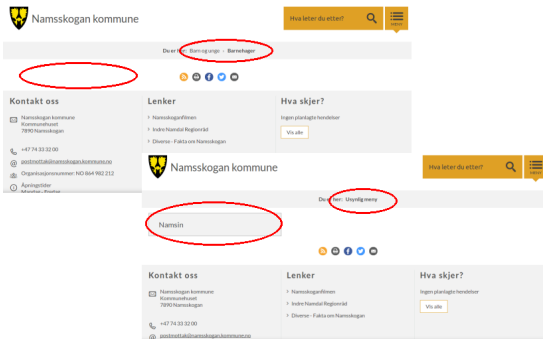
- ✚ Dynamic scripts application
- ✚ Cluster representatives



- 1 Web Accessibility
 - EIII Accessibility Evaluation
- 2 Solution
 - Page Accessibility Profile
- 3 Results
 - Example
- 4 Future steps
- 5 Summary

Results – example

Two different web pages but similar:



The image displays two screenshots of the Namsskogan kommune website, illustrating two different pages that are structurally similar. Red circles highlight specific elements in both pages.

Top Screenshot (Barn-og-unge - Barnehager):

- Header: Namsskogan kommune logo and search bar.
- Breadcrumb: **Barn-og-unge - Barnehager** (circled in red).
- Social media icons: Facebook, Twitter, YouTube, Instagram.
- Left Column (Kontakt oss):
 - Namsskogan kommune
 - Kommuneskiltet: 7890 Namsskogan
 - +47 74 33 32 00
 - post@stallinnamsskogan.kommune.no
 - Organisasjonsnummer: NO 864 902 212
 - Apsingelider
 - Møteland - Postboks
- Middle Column (Lenker):
 - Namsskogefilmen
 - Indre Namdal Regionråd
 - Diverse - Fakta om Namsskogan
- Right Column (Hva skjer?):
 - Ingen planlagte hendelser
 - Vis alle
- Bottom Right: Search bar and menu icon.

Bottom Screenshot (Usynlig-meny):

- Header: Namsskogan kommune logo and search bar.
- Breadcrumb: **Usynlig-meny** (circled in red).
- Search Input: **Namsin** (circled in red).
- Social media icons: Facebook, Twitter, YouTube, Instagram.
- Left Column (Kontakt oss):
 - Namsskogan kommune
 - Kommuneskiltet: 7890 Namsskogan
 - +47 74 33 32 00
 - post@stallinnamsskogan.kommune.no
- Middle Column (Lenker):
 - Namsskogefilmen
 - Indre Namdal Regionråd
 - Diverse - Fakta om Namsskogan
- Right Column (Hva skjer?):
 - Ingen planlagte hendelser
 - Vis alle
- Bottom Right: Search bar and menu icon.

www.namsskogan.kommune.no/barn-og-unge/barnehager
www.namsskogan.kommune.no/usynlig-meny

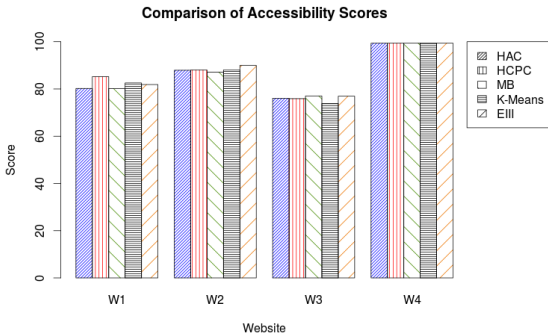
Accessibility page profiles example

| | |
|---|----------|
| Provide descriptive titles for web pages [2.4.2] | ✓1 |
| Provide links to navigate to related Web pages [2.4.5] | ✓1 |
| Provide descriptive headings [2.4.6] | ✓7 |
| Primary language of page [3.1.1] | ✓1 |
| Use identical xml:lang and lang attributes [3.1.1] | ✓1 |
| Languages within the body [3.1.2] | ✓1 |
| Provide a submit button to initiate a change of context [3.2.2] | ✓1 |
| Provide descriptive labels [3.3.2] | ✗1 |
| Use a button to identify the purpose of a form control [3.3.2] | ✓1 |
| Label groups of form elements [3.3.2] | ✓1 |
| Define ids for elements [4.1.1] | ✓67 |
| Provide unique accesskeys [4.1.1] | ✓3 |
| Provide role name for div/span with event handler [4.1.2] | ✓170, ✗1 |
| Use HTML form controls and links [4.1.2] | ✓85, ✗2 |
| Accessible name for image links [4.1.2] | ✓1 |
| Title attribute to identify controls [4.1.2] | ✓1 |

| | |
|---|----------|
| Provide descriptive titles for web pages [2.4.2] | ✓1 |
| Provide links to navigate to related Web pages [2.4.5] | ✓1 |
| Provide descriptive headings [2.4.6] | ✓7 |
| Primary language of page [3.1.1] | ✓1 |
| Use identical xml:lang and lang attributes [3.1.1] | ✓1 |
| Languages within the body [3.1.2] | ✓1 |
| Provide a submit button to initiate a change of context [3.2.2] | ✓1 |
| Provide descriptive labels [3.3.2] | ✗1 |
| Use a button to identify the purpose of a form control [3.3.2] | ✓1 |
| Label groups of form elements [3.3.2] | ✓1 |
| Define ids for elements [4.1.1] | ✓67 |
| Provide unique accesskeys [4.1.1] | ✓3 |
| Provide role name for div/span with event handler [4.1.2] | ✓173, ✗1 |
| Use HTML form controls and links [4.1.2] | ✓85, ✗2 |
| Accessible name for image links [4.1.2] | ✓1 |
| Title attribute to identify controls [4.1.2] | ✓1 |

- ✗ Similar number of tests applied
- ✗ Common accessibility barriers detected

Accessibility scores compared



New accessibility scores comparable to the previous EIII scores

- 1 Web Accessibility
 - EIII Accessibility Evaluation
- 2 Solution
 - Page Accessibility Profile
- 3 Results
 - Example
- 4 Future steps
- 5 Summary

- ✦ Clustering of web pages *on the fly*, i.e., while performing accessibility analysis of a website
- ✦ Evaluation of the mobile content on websites by configuring the crawler
- ✦ Addition of further tests in EIII checking tools

- 1 Web Accessibility
 - EIII Accessibility Evaluation
- 2 Solution
 - Page Accessibility Profile
- 3 Results
 - Example
- 4 Future steps
- 5 Summary

- ✦ The approach can reduce the number of pages by 70% up to 99%
- ✦ Page accessibility profile can be used for clustering
- ✦ Page clusters can identify website templates
- ✦ The barriers found in templates can enable the developer to remove them efficiently
- ✦ Detection of error pages/invalid URLs