

Beurteilung Spiegelungsfähigkeit

Beschreibung Abspielumgebung

Werktitel	Pleintekst.nl
Werkversion	19.11.2014
Autor	Florian Cramer
Hauptseite-URL	http://pleintekst.nl/
Datum der Sichtung	19.11.2014
Bearbeiter	S. Kuch
Einrichtung	Deutsches Literaturarchiv Marbach
Betriebssystem	Windows XP
Browser	Firefox 17.0.7
Plugins	

Eine Liste aller zum Werk gehörenden und getesteten URLs findet sich im Anhang A. Das Ergebnis der automatisierten Evaluation mittels ArchiveReady findet sich in Anhang B.

Für die Beurteilung der Archivierbarkeit ist die Anzahl der negativ gewerteten Eigenschaften unerheblich.

Teilbereiche

Dieser Beurteilungsleitfaden beruht auf dem Paper „CLEAR: a credible method to evaluate website archivability“¹ und dient als Grundlage für die Entwicklung eines Programms, das die Beurteilung semiautomatisch durchführt.

Zugänglichkeit

1. Validieren alle Hyperlinks? - Ja.
2. Werden Inhalte dynamisch generiert? - Ja.
3. Gibt es bereits eine archivierte Version, z. B. bei archive.org? Welche Qualität hat diese? - Ja, aber nicht funktionsfähig
4. Verfügt die Seite über eine Sitemap? - Nein.
5. Verfügt die Seite über eine robots.txt? Schließt diese Crawler oder bestimmten Content aus? - Nein.
6. Verfügt die Seite über Feeds? - Nein.
7. Werden proprietäre Datenformate (Flash/Quicktime) verwendet? - Nein.

Standardkonformität

1. Validierung referenzierter Medien (Bilder, Audio, Video, ...) - Keine vorhanden.
2. Compliance-Prüfung von Webelementen (HTML, CSS) - nicht bestanden (invalides HTML)

Performanz

Werden Anfragen ausreichend schnell beantwortet? - Nicht bestanden.

Kohäsion

Werden Dateien auf verschiedenen Systemen gehostet? - Nein.

¹ Banos et al.: *CLEAR: a credible method to evaluate website archivability*. Lisabon. 2013.

Metadaten

Sind Metadaten aus HTTP Headern vorhanden? - Ja.

Bemerkungen

Auch wenn der hohe Wert der Archivierbarkeit darauf hindeutet, dass das Werk archivierbar ist, so lässt sich nicht der dynamische Charakter des Werks einfangen.

Beurteilung

Der schlechteste Wert aus der Evaluierung aller Dokumente ist maßgeblich für die Archivierbarkeit der gesamten Webresource.

Archivierbarkeit 77%

Anhang A: Liste aller zum Werk gehörenden und getesteten URLs

<http://pleintekst.nl/>

Anhang B

URL: <http://pleintekst.nl/>
Testdatum: 2014-11-11 12:30:19.816722
Standard Compliance: 75
Archivierbarkeit: 77
Accessibility: 59
Kohäsion: 100

+++ Meldungen +++

Titel: No sitemap.xml found.
Level: 0
Gewichtung: HIGH
Beschreibung:
Sitemaps are an easy way for webmasters to inform web bots about available web pages so that they can more intelligently crawl the site.

Titel: No images found in this page.
Level: 100
Gewichtung: MEDIUM
Beschreibung:
Therefore, it is easier to be archived.

Titel: You are not using any external proprietary objects (Flash, Quicktime)
Level: 100
Gewichtung: HIGH
Beschreibung:
Proprietary formats are not recommended for , long term preservation.
Please consider changing , your content to open formats.

Titel: Invalid HTML http://pleintekst.nl/. Located 6 errors, 4 warnings.
Level: 0

Gewichtung: HIGH

Beschreibung:

Web pages which do not conform with Web Standards have a lower possibility to be preserved correctly

Titel: 1 valid links identified.

Level: 100

Gewichtung: HIGH

Beschreibung:

No invalid links found.

Titel: No CSS files used in this webpage.

Level: 100

Gewichtung: HIGH

Beschreibung:

The absence of CSS simplifies archiving.

Titel: Local scripts found: 0, remote scripts found: 0

Level: 100

Gewichtung: MEDIUM

Beschreibung:

scripts used in a website but hosted elsewhere may cause problems in web archiving because they may not be captured when the site is archived. What is more, the site is depended on 3rd party sites which may not be available in the future, causing incontrollable problems.

Titel: No inline script elements found

Level: 100

Gewichtung: HIGH

Beschreibung:

Inline JavaScript code inside HTML is considered a bad practice. The absence of inline JavaScript simplifies archiving.

Titel: Transfer encoding was clearly defined in HTTP Headers.

Level: 100

Gewichtung: MEDIUM

Beschreibung:

A Transfer-Encoding header field, which can be , used to specify an auxiliary encoding that was applied , to the data in order to allow it to pass through mail , transport mechanisms which may have data or , character set limitations.

Titel: Content type was clearly defined.

Level: 100

Gewichtung: MEDIUM

Beschreibung:

Correctly defined HTTP headers help web bots understand the content.

Titel: HTTP caching headers are not available.

Level: 50

Gewichtung: MEDIUM

Beschreibung:

HTTP caching headers are critical to improve web archiving bots performance because they help avoid downloading unmodified content multiple times.

Titel: Network response time is 3000 ms

Level: 0

Gewichtung: HIGH

Beschreibung:

Website performance affects greatly the process of data aggregation and web accessibility. Performance evaluation rating is 100 when response time is 200ms or less and 0 when response time is more than 2sec.