

Beurteilung Spiegelungsfähigkeit

Beschreibung Abspielumgebung

Werktitel	SearchSongs
Werkversion	20.11.2014
Autor	Johannes Auer, Beat Suter, René Bauer
Hauptseite-URL	http://searchsonata.netzliteratur.net/fiwi/searchsongs_info.html
Datum der Sichtung	20.11.2014
Bearbeiter	S. Kuch
Einrichtung	Deutsches Literaturarchiv Marbach
Betriebssystem	Windows XP
Browser	Firefox 17.0.7
Plugins	Shockwave Flash 14.0.0.145

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? - Nein.
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? - Ja.

Standardkonformität

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

Performanz

Werden Anfragen ausreichend schnell beantwortet? - Nicht bestanden.

Kohäsion

Werden Dateien auf verschiedenen Systemen gehostet? - Ja..

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

Metadaten

Sind Metadaten aus HTTP Headern vorhanden? - Nein.

Bemerkungen

Zur Überprüfung des Werks mussten zwei URLs auf zwei verschiedenen Systemen getestet werden. Die Archivierbarkeit bezieht sich auf statische Inhalte, die allerdings dynamisch generiert werden.

Beurteilung

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

Archivierbarkeit 69%

Anhang A: Liste aller zum Werk gehörenden und getesteten URLs

http://searchsonata.netzliteratur.net/fiwi/searchsongs_info_en.html
http://searchsonata.netzliteratur.net/fiwi/searchsongs_info.html
<http://searchsongs.cyberfiction.ch/>

Anhang B

URL: http://searchsonata.netzliteratur.net/fiwi/searchsongs_info.html
Testdatum: 2014-11-11 12:31:31.862362
Standard Compliance: 29
Archivierbarkeit: 69
Accessibility: 45
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 using proprietary external objects (Flash, Quicktime).

Level: 0

Gewichtung: HIGH

Beschreibung:

Proprietary formats are not recommended for , long term preservation.
Please consider changing , your content to open formats.

Titel: Invalid CSS <http://searchsonata.netzliteratur.net/fiwi/style1.css>.
Located 9 errors, 0 warnings.
Level: 0
Gewichtung: MEDIUM
Beschreibung:
Web pages which do not conform with Web Standards have a lower possibility
to be preserved correctly

Titel: Invalid HTML
http://searchsonata.netzliteratur.net/fiwi/searchsongs_info.html. Located
12 errors, 3 warnings.
Level: 0
Gewichtung: HIGH
Beschreibung:
Web pages which do not conform with Web Standards have a lower possibility
to be preserved correctly

Titel: 5 valid links identified.
Level: 100
Gewichtung: HIGH
Beschreibung:
No invalid links found.

Titel: Local CSS found: 1, remote CSS found: 0
Level: 100
Gewichtung: MEDIUM
Beschreibung:
CSS 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.
 1 local CSS:

<http://searchsonata.netzliteratur.net/fiwi/style1.css>

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: Content encoding was clearly defined in HTTP Headers.
Level: 100
Gewichtung: MEDIUM

Beschreibung:

This header defines type of encoding used on the data. (e.g. compression).

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 enabled.

Level: 100

Gewichtung: MEDIUM

Beschreibung:

HTTP Expires: was not defined.
HTTP Last-Modified: was clearly defined in HTTP headers: Tue, 11 Jan 2011 09:08:56 GMT
HTTP Etag was clearly defined in HTTP headers: "b6326f-18f2-4998e6f25da00"
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.

URL: http://searchsonata.netzliteratur.net/fiwi/searchsongs_info_en.html

Testdatum: 2014-11-11 12:31:25.448633

Standard Compliance: 29

Archivierbarkeit: 69

Accessibility: 45

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 using proprietary external objects (Flash, Quicktime).

Level: 0

Gewichtung: HIGH

Beschreibung:

Proprietary formats are not recommended for , long term preservation.

Please consider changing , your content to open formats.

Titel: Invalid CSS <http://searchsonata.netzliteratur.net/fiwi/style1.css>.
Located 9 errors, 0 warnings.
Level: 0
Gewichtung: MEDIUM
Beschreibung:
Web pages which do not conform with Web Standards have a lower possibility
to be preserved correctly

Titel: Invalid HTML
http://searchsonata.netzliteratur.net/fiwi/searchsongs_info_en.html.
Located 11 errors, 3 warnings.
Level: 0
Gewichtung: HIGH
Beschreibung:
Web pages which do not conform with Web Standards have a lower possibility
to be preserved correctly

Titel: 5 valid links identified.
Level: 100
Gewichtung: HIGH
Beschreibung:
No invalid links found.

Titel: Local CSS found: 1, remote CSS found: 0
Level: 100
Gewichtung: MEDIUM
Beschreibung:
CSS 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.
 1 local CSS:

<http://searchsonata.netzliteratur.net/fiwi/style1.css>

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: Content encoding was clearly defined in HTTP Headers.
Level: 100

Gewichtung: MEDIUM

Beschreibung:

This header defines type of encoding used on the data. (e.g. compression).

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 enabled.

Level: 100

Gewichtung: MEDIUM

Beschreibung:

HTTP Expires: was not defined.
HTTP Last-Modified: was clearly defined in HTTP headers: Tue, 11 Jan 2011 09:08:56 GMT
HTTP Etag was clearly defined in HTTP headers: "b6326e-161c-4998e6f25da00"
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.

URL: <http://searchsongs.cyberfiction.ch/>

Testdatum: 2014-11-20 08:49:03.774728

Standard Compliance: 44

Archivierbarkeit: 69

Accessibility: 33

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 using proprietary external objects (Flash, Quicktime).

Level: 0

Gewichtung: HIGH

Beschreibung:

Proprietary formats are not recommended for , long term preservation.
Please consider changing , your content to open formats.

Titel: Invalid HTML http://searchsongs.cyberfiction.ch/. Located 20 errors,
2 warnings.
Level: 0
Gewichtung: HIGH
Beschreibung:
Web pages which do not conform with Web Standards have a lower possibility
to be preserved correctly

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: No Content-Encoding or Transfer-Encoding defined in HTTP Headers.
Level: 50
Gewichtung: MEDIUM
Beschreibung:
The encoding used on the data is not defined. Web bots , would try assume
the possible encoding, leading to potential , mistakes. 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 enabled.
Level: 100
Gewichtung: MEDIUM
Beschreibung:

HTTP Expires: was not defined.
HTTP Last-Modified: was clearly defined
in HTTP headers: Thu, 10 Apr 2008 08:04:38 GMT
HTTP Etag was clearly
defined in HTTP headers: "58c04f-5ec-44d63180"
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.