

Beurteilung Spiegelungsfähigkeit

Beschreibung Abspelunggebung

Werktitel	Nord
Werkversion	19.11.2014
Autor	Esther Hunziker, Felix Zbinden
Standort / URL	http://www.ref17.net/nord/
Datum der Sichtung	19.11.2014
Bearbeiter	S. Kuch
Einrichtung	Deutsches Literaturarchiv Marbach
Betriebssystem	Windows XP
Browser	Firefox 17.0.7
Plugins	Java Plugin 10.65.2.20, 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? - Nein.
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? - Ja. (manuell)

Standardkonformität

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

Performanz

Werden Anfragen ausreichend schnell beantwortet? - Eingeschränkt.

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? - Nein.

Bemerkungen

Da es nicht möglich ist, automatisiert eine Liste aller referenzierten Dokumente zu erhalten, musste dies manuell durchgeführt werden.

Beurteilung

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

Archivierbarkeit

Anhang A: Liste aller zum Werk gehörenden und getesteten URLs

<http://www.ref17.net/hunziker/nord.html>
<http://www.ref17.net/nord/>

Anhang B

URL: <http://www.ref17.net/hunziker/nord.html>
Testdatum: 2014-11-11 12:34:56.528304
Standard Compliance: 69
Archivierbarkeit: 81
Accessibility: 56
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://www.ref17.net/hunziker/nord.html>. Located 19 errors, 0 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: Wed, 30 Oct 2013 03:01:08 GMT
HTTP Etag was clearly defined in HTTP headers: "2cee-4e9ec8b81a3e7"
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://www.ref17.net/nord/>
Testdatum: 2014-11-11 11:50:05.167603
Standard Compliance: 69
Archivierbarkeit: 86
Accessibility: 75
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: Local images found: 1, remote images found: 0
Level: 100
Gewichtung: MEDIUM
Beschreibung:
images 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 images:
 http://www.ref17.net/nord/de/_images/index.gif

Titel: Checking 1 images with JHOVE Object Validation Environment
Level: -1
Gewichtung: MEDIUM
Beschreibung:
Format validation is the process of determining the level of compliance of a digital object to the specification for its purported format, e.g.: I have an object purportedly of format F; is it?

Titel: Image checked: http://www.ref17.net/nord/de/_images/index.gif
Level: 100
Gewichtung: MEDIUM
Beschreibung:
<pre>Jhove (Rel. 1.6, 2011-01-04) Date: 2014-11-11 11:49:56 CET
RepresentationInformation: /tmp/41332060 ReportingModule: GIF-hul, Rel. 1.3
(2006-09-05) LastModified: 2014-11-11 11:49:56 CET Size: 1457 Format: GIF

```
Version: 89a Status: Well-Formed and valid SignatureMatches: GIF-hul
MIMEtype: image/gif Profile: GIF 89a GIFMetadata: GraphicRenderingBlocks: 1
Blocks: LogicalScreenDescriptor: LogicalScreenWidth: 600
LogicalScreenHeight: 90 ColorResolution: 4 BackgroundColorIndex: 0
PixelAspectRatio: 0 GlobalColorTableFlag: Global color table follows;
background color index meaningful GlobalColorTableSortFlag: Not ordered
GlobalColorTableSize: 3 GlobalColorTable: 10, 10, 10, 1, 1, 1, 245, 245,
245, 236, 236, 236, 20, 20, 20, 19, 19, 19, 246, 246, 246, 29, 29, 29, 28,
28, 28, 237, 237, 237, 38, 38, 38, 254, 254, 254, 235, 235, 235, 227, 227,
227, 222, 222, 222, 255, 255, 255 GraphicControlExtension: DisposalMethod:
No disposal specified UserInputFlag: User input not expected
TransparencyFlag: Transparent index is not given DelayTime: 0
TransparentColorIndex: 0 ImageDescriptor: ImageLeftPosition: 0
ImageTopPosition: 0 InterlaceFlag: Image is not interlaced
LocalColorTableFlag: No local color table; use global table if available
LocalColorTableSortFlag: Not ordered LocalColorTableSize: 0
NisoImageMetadata: ByteOrder: little_endian CompressionScheme: LZW
ImageWidth: 600 ImageHeight: 90 ColorSpace: palette color Orientation:
normal BitsPerSample: 8 BitsPerSampleUnit: integer
</pre>
```

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: Valid CSS http://www.ref17.net/nord/de/_styles/index.css.
Level: 100
Gewichtung: MEDIUM
Beschreibung:
Web pages which conform with Web Standards , have a higher possibility to be preserved correctly

Titel: Invalid HTML <http://www.ref17.net/nord/>. Located 2 errors, 0 warnings.
Level: 0
Gewichtung: HIGH
Beschreibung:
Web pages which do not conform with Web Standards have a lower possibility to be preserved correctly

Titel: There are 3 valid and 2 invalid links.
Level: 60
Gewichtung: HIGH
Beschreibung:
Invalid links prevent web bots from accessing the full content , in order to preserve it. Also check out http://en.wikipedia.org/wiki/Link_rot ,
Invalid links:
 , http://www.ref17.net/nord/de/index_de.html
http://www.ref17.net/nord/en/index_en.html
 , Valid links:
 , <http://www.adobe.com/shockwave/download>
<http://www.ref17.net>
http://www.ref17.net/nord/de/_styles/index.css

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://www.ref17.net/nord/de/_styles/index.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: 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: Fri, 06 Jun 2008 01:09:52 GMT
HTTP Etag was clearly defined in HTTP headers: "b05-44ef51e7ca892"
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 10 ms
Level: 100

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.