

Beurteilung Spiegelungsfähigkeit

Beschreibung Abspielumgebung

Werktitel	Die Aaleskorte der Ölig
Werkversion	27.04.2015
Autor	Frank Klötgen, Dirk Günther
Hauptseite-URL	http://www.aaleskorte.de/
Datum der Sichtung	27.04.2015
Bearbeiter	S. Kuch
Einrichtung	Deutsches Literaturarchiv Marbach
Betriebssystem	Windows XP
Browser	Firefox 37.0.2
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, ...) - Konnte nicht geprüft werden.
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? - Eingeschränkt.

Bemerkungen

Die Startseite <http://www.aaleskorte.de/> wurde nicht überprüft, da der Crawler dort auf andere, dort verlinkte Werke geraten wäre. Die Seite enthält html-Code und ein Java-Applet (manuell überprüft).

Beurteilung

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

Archivierbarkeit 64%

Anhang A: Liste aller zum Werk gehörenden und getesteten URLs

www.aaleskorte.de/anfang.htm
www.aaleskorte.de/drehbuch.htm
www.aaleskorte.de/index.htm
www.aaleskorte.de/start.htm

Anhang B

URL: <http://www.aaleskorte.de/anfang.htm>
Testdatum: 2015-04-22 21:23:40
Standard Compliance: 96
Archivierbarkeit: 94
Accessibility: 82
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: httpheaders.check_content_transfer_encoding.WARNING.title
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: True
HTTP Etag was clearly defined in HTTP headers:
True
HTTP caching headers are critical to improve web archiving bots
performance because they help avoid downloading unmodified content multiple
times.

Titel: Valid HTML <http://www.aaleskorte.de/anfang.htm>.
Level: 100
Gewichtung: HIGH
Beschreibung:
Web pages which conform with Web Standards , have a higher possibility to
be preserved correctly.

Titel: 2 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: Local images found: 5, 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.
5 local images:
http://www.aaleskorte.de/bilder/clipkinder.jpg
http://www.aaleskorte.de/bilder/clipaal.jpg
http://www.aaleskorte.de/bilder/cliphohmann.jpg
http://www.aaleskorte.de/bilder/cliperzaehler.jpg
http://www.aaleskorte.de/bilder/clipoelig.jpg

Titel: Checking 5 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.aaleskorte.de/bilder/clipkinder.jpg

Level: 100

Gewichtung: MEDIUM

Beschreibung:

```
<pre>Jhove (Rel. 1.6, 2011-01-04) Date: 2015-04-27 09:24:53 CEST
RepresentationInformation: /tmp/34177118 ReportingModule: JPEG-hul, Rel.
1.2 (2007-02-13) LastModified: 2015-04-27 09:24:53 CEST Size: 735 Format:
JPEG Version: 1.01 Status: Well-Formed and valid SignatureMatches: JPEG-hul
MIMEtype: image/jpeg Profile: JFIF JPEGMetadata: CompressionType: Huffman
coding, Progressive DCT Images: Number: 1 Image: NisoImageMetadata:
ByteOrder: big_endian CompressionScheme: JPEG ImageWidth: 20 ImageHeight:
29 ColorSpace: YCbCr SamplingFrequencyUnit: inch XSamplingFrequency: 150
YSamplingFrequency: 150 BitsPerSample: 8, 8, 8 BitsPerSampleUnit: integer
SamplesPerPixel: 3 Scans: 10 QuantizationTables: QuantizationTable:
Precision: 8-bit DestinationIdentifier: 0 QuantizationTable: Precision: 8-
bit DestinationIdentifier: 1 ApplicationSegments: APP0
</pre>
```

Titel: Image checked: http://www.aaleskorte.de/bilder/cliphohmann.jpg

Level: 100

Gewichtung: MEDIUM

Beschreibung:

```
<pre>Jhove (Rel. 1.6, 2011-01-04) Date: 2015-04-27 09:24:54 CEST
RepresentationInformation: /tmp/7680308 ReportingModule: JPEG-hul, Rel. 1.2
(2007-02-13) LastModified: 2015-04-27 09:24:53 CEST Size: 894 Format: JPEG
Version: 1.01 Status: Well-Formed and valid SignatureMatches: JPEG-hul
MIMEtype: image/jpeg Profile: JFIF JPEGMetadata: CompressionType: Huffman
coding, Progressive DCT Images: Number: 1 Image: NisoImageMetadata:
ByteOrder: big_endian CompressionScheme: JPEG ImageWidth: 20 ImageHeight:
30 ColorSpace: YCbCr SamplingFrequencyUnit: inch XSamplingFrequency: 150
YSamplingFrequency: 150 BitsPerSample: 8, 8, 8 BitsPerSampleUnit: integer
SamplesPerPixel: 3 Scans: 10 QuantizationTables: QuantizationTable:
Precision: 8-bit DestinationIdentifier: 0 QuantizationTable: Precision: 8-
bit DestinationIdentifier: 1 ApplicationSegments: APP0
</pre>
```

Titel: Image checked: <http://www.aaleskorte.de/bilder/clipaal.jpg>
Level: 100

Gewichtung: MEDIUM

Beschreibung:

```
Jhove (Rel. 1.6, 2011-01-04) Date: 2015-04-27 09:24:54 CEST
RepresentationInformation: /tmp/76648096 ReportingModule: JPEG-hul, Rel.
1.2 (2007-02-13) LastModified: 2015-04-27 09:24:53 CEST Size: 762 Format:
JPEG Version: 1.01 Status: Well-Formed and valid SignatureMatches: JPEG-hul
MIMEtype: image/jpeg Profile: JFIF JPEGMetadata: CompressionType: Huffman
coding, Progressive DCT Images: Number: 1 Image: NisoImageMetadata:
ByteOrder: big_endian CompressionScheme: JPEG ImageWidth: 20 ImageHeight:
30 ColorSpace: YCbCr SamplingFrequencyUnit: inch XSamplingFrequency: 75
YSamplingFrequency: 75 BitsPerSample: 8, 8, 8 BitsPerSampleUnit: integer
SamplesPerPixel: 3 Scans: 10 QuantizationTables: QuantizationTable:
Precision: 8-bit DestinationIdentifier: 0 QuantizationTable: Precision: 8-
bit DestinationIdentifier: 1 ApplicationSegments: APP0
```

Titel: Image checked: <http://www.aaleskorte.de/bilder/clipoelig.jpg>
Level: 100

Gewichtung: MEDIUM

Beschreibung:

```
Jhove (Rel. 1.6, 2011-01-04) Date: 2015-04-27 09:24:54 CEST
RepresentationInformation: /tmp/81404236 ReportingModule: JPEG-hul, Rel.
1.2 (2007-02-13) LastModified: 2015-04-27 09:24:54 CEST Size: 813 Format:
JPEG Version: 1.01 Status: Well-Formed and valid SignatureMatches: JPEG-hul
MIMEtype: image/jpeg Profile: JFIF JPEGMetadata: CompressionType: Huffman
coding, Progressive DCT Images: Number: 1 Image: NisoImageMetadata:
ByteOrder: big_endian CompressionScheme: JPEG ImageWidth: 20 ImageHeight:
30 ColorSpace: YCbCr SamplingFrequencyUnit: inch XSamplingFrequency: 300
YSamplingFrequency: 300 BitsPerSample: 8, 8, 8 BitsPerSampleUnit: integer
SamplesPerPixel: 3 Scans: 10 QuantizationTables: QuantizationTable:
Precision: 8-bit DestinationIdentifier: 0 QuantizationTable: Precision: 8-
bit DestinationIdentifier: 1 ApplicationSegments: APP0
```

Titel: Image checked: <http://www.aaleskorte.de/bilder/cliperzaehler.jpg>
Level: 100

Gewichtung: MEDIUM

Beschreibung:

```
Jhove (Rel. 1.6, 2011-01-04) Date: 2015-04-27 09:24:54 CEST
RepresentationInformation: /tmp/75440718 ReportingModule: JPEG-hul, Rel.
1.2 (2007-02-13) LastModified: 2015-04-27 09:24:54 CEST Size: 672 Format:
JPEG Version: 1.01 Status: Well-Formed and valid SignatureMatches: JPEG-hul
MIMEtype: image/jpeg Profile: JFIF JPEGMetadata: CompressionType: Huffman
coding, Progressive DCT Images: Number: 1 Image: NisoImageMetadata:
ByteOrder: big_endian CompressionScheme: JPEG ImageWidth: 20 ImageHeight:
29 ColorSpace: YCbCr SamplingFrequencyUnit: inch XSamplingFrequency: 150
YSamplingFrequency: 150 BitsPerSample: 8, 8, 8 BitsPerSampleUnit: integer
SamplesPerPixel: 3 Scans: 10 QuantizationTables: QuantizationTable:
Precision: 8-bit DestinationIdentifier: 0 QuantizationTable: Precision: 8-
bit DestinationIdentifier: 1 ApplicationSegments: APP0
```

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: 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.

URL: <http://www.aaleskorte.de/drehbuch.htm>

Testdatum: 2015-04-22 21:23:40

Standard Compliance: 69

Archivierbarkeit: 79

Accessibility: 45

Kohäsion: 100

+++ Meldungen +++

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: 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: httpheaders.check_content_transfer_encoding.WARNING.title

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: True
HTTP Etag was clearly defined in HTTP headers: True
HTTP caching headers are critical to improve web archiving bots performance because they help avoid downloading unmodified content multiple times.

Titel: Invalid HTML <http://www.aaleskorte.de/drehbuch.htm>. Located 1 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: There are 0 valid and 2 invalid links.

Level: 0

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:
 TypeError: 'undefined' is not a function (evaluating 'parent.neue_Seite(2)')
<http://www.aaleskorte.de/drehbuch.htm/>
<http://www.aaleskorte.de/drehbuch.htm:8> in onload
 Valid links:

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: 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.aaleskorte.de/index.htm>

Testdatum: 2015-04-22 21:23:40

Standard Compliance: 71

Archivierbarkeit: 64

Accessibility: 36

Kohäsion: 50

+++ 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: httpheaders.check_content_transfer_encoding.WARNING.title

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: True
HTTP Etag was clearly defined in HTTP headers: True
HTTP caching headers are critical to improve web archiving bots performance because they help avoid downloading unmodified content multiple times.

Titel: Local images found: 0, remote images found: 1

Level: 0

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 remote images:
<http://fastcounter.linkexchange.com/fastcounter?481790+963587>

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: Could not download file
<http://fastcounter.linkexchange.com/fastcounter?481790+963587>
Level: 0
Gewichtung: HIGH
Beschreibung:
('Connection aborted.', gaierror(-2, 'Name or service not known'))

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.aaleskorte.de/index.htm>. Located 14 errors, 5 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 6 valid and 1 invalid links.
Level: 85
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://stadtkind.tagesspiegel.de/>
 Valid links:
 <http://www.hirnpoma.de/trashical/index.html>
<http://www.hirnpoma.de>
<http://www.aaleskorte.de/anfang.htm>
<http://www.internetkrimi.de>
<http://member.linkexchange.com/cgi-bin/fc/fastcounter-login?481790>
<http://www.amazon.de/gp/redirect.html?ie=UTF8&location=http%3A%2F%2Fwww.amazon.de%2FWill-Kacheln-Buch-mit-CD%2Fdp%2F3938424192&site-redirect=de&>tag=hirnpomade-21&linkCode=ur2&camp=1638&creative=6742>

-----URL:
<http://www.aaleskorte.de/start.htm>
Testdatum: 2015-04-22 21:23:40

Standard Compliance: 69
Archivierbarkeit: 76
Accessibility: 33
Kohäsion: 100

+++ Meldungen +++

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: 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: httpheaders.check_content_transfer_encoding.WARNING.title
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: True
HTTP Etag was clearly defined in HTTP headers: True
HTTP caching headers are critical to improve web archiving bots performance because they help avoid downloading unmodified content multiple times.

Titel: Invalid HTML http://www.aaleskorte.de/start.htm. Located 441 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: 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: 1 inline script elements found
Level: 0
Gewichtung: HIGH
Beschreibung:
Inline JavaScript code inside HTML is considered a bad practice. The
absence of inline JavaScript simplifies archiving.

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.