

## **Beurteilung Spiegelungsfähigkeit**

### **Beschreibung Abspielumgebung**

Werktitel	Google Snippet Picture Ad(d) Sense
Werkversion	18.11.2014
Autor	Johannes Auer
Hauptseite-URL	<a href="http://auer.netzliteratur.net/comma/english/snippet_picture.php">http://auer.netzliteratur.net/comma/english/snippet_picture.php</a>
Datum der Sichtung	18.11.2014
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“<sup>1</sup> 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 vollständig und funktionsfähig
4. Verfügt die Seite über eine Sitemap? - Nein.
5. Verfügt die Seite über eine robots.txt? Schließt diese Crawler oder bestimmten Content aus? - Ja. Nein.
6. Verfügt die Seite über Feeds? - Nein.
7. Werden proprietäre Datenformate (Flash/Quicktime) verwendet? - Nein.

### **Standardkonformität**

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

### **Performanz**

Werden Anfragen ausreichend schnell beantwortet? - Bestanden.

### **Kohäsion**

Werden Dateien auf verschiedenen Systemen gehostet? - Nein.

---

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

## Metadaten

Sind Metadaten aus HTTP Headern vorhanden? - Ja.

## Bemerkungen

## Beurteilung

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

Archivierbarkeit 88%

### Anhang A: Liste aller zum Werk gehörenden und getesteten URLs

[http://auer.netzliteratur.net/comma/english/snippet\\_picture.php](http://auer.netzliteratur.net/comma/english/snippet_picture.php)

### Anhang B

URL: [http://auer.netzliteratur.net/comma/english/snippet\\_picture.php](http://auer.netzliteratur.net/comma/english/snippet_picture.php)

Testdatum: 2014-11-11 08:56:05.183930

Standard Compliance: 75

Archivierbarkeit: 88

Accessibility: 77

Kohäsion: 100

+++ Meldungen +++

-----  
Titel: There are no "Disallow:" commands in your robots.txt file

Level: 100

Gewichtung: MEDIUM

Beschreibung:

Disallow: commands are used to block web bots from specific website areas. They may cause web archiving problems.

-----  
Titel: There is no sitemap.xml reference in your robots.txt file.

Level: 0

Gewichtung: MEDIUM

Beschreibung:

This can be used to tell search engines or other robots where your sitemap is located.

-----  
Titel: Sitemap file exists <http://auer.netzliteratur.net/sitemap.xml>

Level: 100

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.<br /> 1 local images:<br />

[http://auer.netzliteratur.net/comma/english/Logo\\_60wht.gif](http://auer.netzliteratur.net/comma/english/Logo_60wht.gif) <br />

-----  
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://auer.netzliteratur.net/comma/english/Logo\\_60wht.gif](http://auer.netzliteratur.net/comma/english/Logo_60wht.gif)

Level: 100

Gewichtung: MEDIUM

Beschreibung:

<pre>Jhove (Rel. 1.6, 2011-01-04) Date: 2014-11-11 08:55:56 CET RepresentationInformation: /tmp/57787628  
ReportingModule: GIF-hul, Rel. 1.3 (2006-09-05) LastModified: 2014-11-11 08:55:55 CET Size: 5668 Format:  
GIF Version: 89a Status: Well-Formed and valid SignatureMatches: GIF-hul MIMEType: image/gif Profile: GIF  
89a GIFMetadata: GraphicRenderingBlocks: 1 Blocks: LogicalScreenDescriptor: LogicalScreenWidth: 176  
LogicalScreenHeight: 77 ColorResolution: 8 BackgroundColorIndex: 0 PixelAspectRatio: 0  
GlobalColorTableFlag: Global color table follows; background color index meaningful  
GlobalColorTableSortFlag: Not ordered GlobalColorTableSize: 7 GlobalColorTable: 213, 213, 215, 192, 192,  
192, 249, 249, 250, 247, 247, 248, 230, 230, 231, 203, 203, 204, 201, 201, 202, 195, 196, 199, 177, 178, 181, 0,  
47, 161, 0, 39, 132, 0, 29, 98, 2, 33, 107, 179, 181, 186, 0, 62, 197, 0, 54, 176, 0, 52, 170, 0, 44, 146, 1, 60, 187,  
1, 58, 182, 1, 51, 163, 1, 49, 155, 5, 39, 116, 9, 61, 173, 9, 49, 136, 18, 64, 164, 17, 57, 147, 27, 78, 187, 26, 58,  
130, 39, 78, 164, 34, 60, 117, 55, 78, 127, 0, 66, 206, 4, 73, 213, 7, 70, 198, 11, 80, 213, 17, 70, 182, 70, 104,  
173, 98, 125, 179, 73, 92, 132, 123, 137, 164, 144, 152, 168, 171, 176, 186, 163, 167, 175, 209, 211, 215, 245,  
246, 248, 203, 204, 206, 15, 89, 222, 22, 90, 216, 27, 89, 207, 31, 100, 219, 44, 107, 217, 64, 115, 205, 90, 109,  
145, 135, 158, 198, 105, 120, 148, 154, 164, 183, 189, 193, 200, 210, 214, 221, 33, 104, 224, 44, 114, 226, 59,  
125, 229, 87, 134, 214, 113, 148, 208, 167, 183, 209, 159, 171, 191, 177, 186, 201, 193, 202, 217, 219, 226, 237,  
156, 161, 169, 212, 217, 225, 182, 185, 190, 74, 137, 232, 97, 153, 236, 125, 162, 219, 152, 175, 210, 180, 197,  
222, 202, 214, 232, 187, 194, 204, 203, 209, 218, 237, 239, 242, 233, 235, 238, 224, 226, 229, 216, 218, 221,  
118, 169, 241, 142, 179, 230, 161, 187, 223, 195, 200, 207, 225, 228, 232, 229, 236, 245, 230, 234, 239, 227,  
231, 236, 217, 222, 227, 249, 250, 251, 247, 248, 249, 221, 222, 223, 216, 217, 218, 238, 239, 239, 21, 81, 21,  
44, 159, 44, 89, 189, 89, 43, 89, 43, 102, 116, 102, 190, 204, 190, 202, 213, 202, 209, 220, 209, 227, 238, 227,  
170, 178, 170, 233, 243, 233, 218, 227, 218, 243, 248, 243, 234, 238, 234, 237, 240, 237, 236, 237, 236, 214,  
215, 214, 186, 187, 186, 21, 103, 20, 38, 123, 37, 113, 193, 112, 139, 200, 138, 173, 214, 172, 223, 231, 222,  
156, 158, 155, 162, 169, 155, 141, 144, 136, 249, 249, 247, 218, 218, 216, 250, 250, 249, 207, 207, 206, 205,  
205, 204, 199, 199, 198, 193, 193, 192, 240, 232, 194, 248, 208, 45, 240, 212, 92, 222, 206, 135, 221, 212, 174,  
203, 200, 188, 246, 243, 230, 239, 238, 234, 200, 155, 0, 218, 170, 3, 234, 184, 6, 244, 197, 19, 147, 117, 13,  
188, 156, 45, 217, 184, 60, 194, 172, 91, 146, 134, 93, 187, 174, 129, 192, 185, 160, 236, 234, 227, 172, 134, 5,  
123, 96, 10, 228, 225, 215, 247, 246, 244, 182, 179, 174, 173, 169, 164, 246, 245, 244, 215, 202, 192, 186, 182,  
179, 176, 174, 173, 171, 150, 141, 203, 198, 196, 140, 99, 85, 203, 191, 187, 214, 208, 206, 199, 196, 195, 247,  
244, 243, 245, 242, 241, 216, 213, 212, 196, 186, 183, 234, 224, 221, 226, 214, 211, 233, 144, 126, 232, 194,  
186, 232, 207, 202, 192, 182, 180, 175, 170, 169, 248, 243, 242, 223, 218, 217, 226, 68, 41, 223, 83, 59, 226,  
107, 86, 203, 135, 123, 212, 178, 172, 183, 155, 150, 216, 36, 12, 214, 43, 20, 221, 50, 24, 222, 57, 31, 207, 67,  
46, 202, 165, 159, 240, 233, 232, 203, 20, 1, 212, 24, 2, 201, 30, 12, 190, 37, 21, 204, 53, 34, 166, 112, 105, 190,  
173, 171, 247, 238, 237, 179, 16, 0, 195, 18, 1, 185, 17, 1, 171, 16, 1, 160, 13, 1, 137, 12, 2, 167, 43, 33, 136, 40,  
31, 114, 39, 32, 174, 67, 57, 130, 68, 62, 107, 10, 3, 175, 172, 172, 232, 229, 229, 212, 210, 210, 226, 225, 225,  
255, 255, 255, 254, 254, 254, 253, 253, 253, 252, 252, 252, 251, 251, 251, 249, 249, 249, 247, 247, 247, 244,  
244, 244, 242, 242, 242, 241, 241, 241, 240, 240, 240, 234, 234, 234, 232, 232, 232, 228, 228, 228, 226, 226,  
226, 224, 224, 224, 223, 223, 223, 220, 220, 220, 218, 218, 218, 217, 217, 217, 212, 212, 212, 211, 211, 211,

```
208, 208, 208, 203, 203, 203, 200, 200, 200, 196, 196, 196, 195, 195, 195, 191, 191, 191, 188, 188, 188, 184,
184, 184, 183, 183, 183, 182, 182, 182, 181, 181, 181, 178, 178, 178, 175, 175, 175, 173, 173, 173, 170, 170,
170, 165, 165, 165 GraphicControlExtension: DisposalMethod: No disposal specified UserInputFlag: User input
not expected TransparencyFlag: Transparent index given DelayTime: 0 TransparentColorIndex: 1
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:
176 ImageHeight: 77 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://auer.netzliteratur.net/comma/english/style.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://auer.netzliteratur.net/comma/english/snippet\\_picture.php](http://auer.netzliteratur.net/comma/english/snippet_picture.php). Located errors, 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: 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.<br /> 1 local CSS:<br />

<http://auer.netzliteratur.net/comma/english/style.css> <br />

-----  
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.<br />

-----  
Titel: 2 inline script elements found  
Level: 0  
Gewichtung: HIGH  
Beschreibung:  
Inline Javascript code inside HTML is considered a bad practice.

-----  
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 clearly defined in HTTP headers: Mon, 26 Jul 1997 05:00:00 GMT<br />HTTP Last-Modified: was clearly defined in HTTP headers: Tue, 11 Nov 2014 07:58:35 GMT<br />HTTP Etag: was not defined.<br />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.