

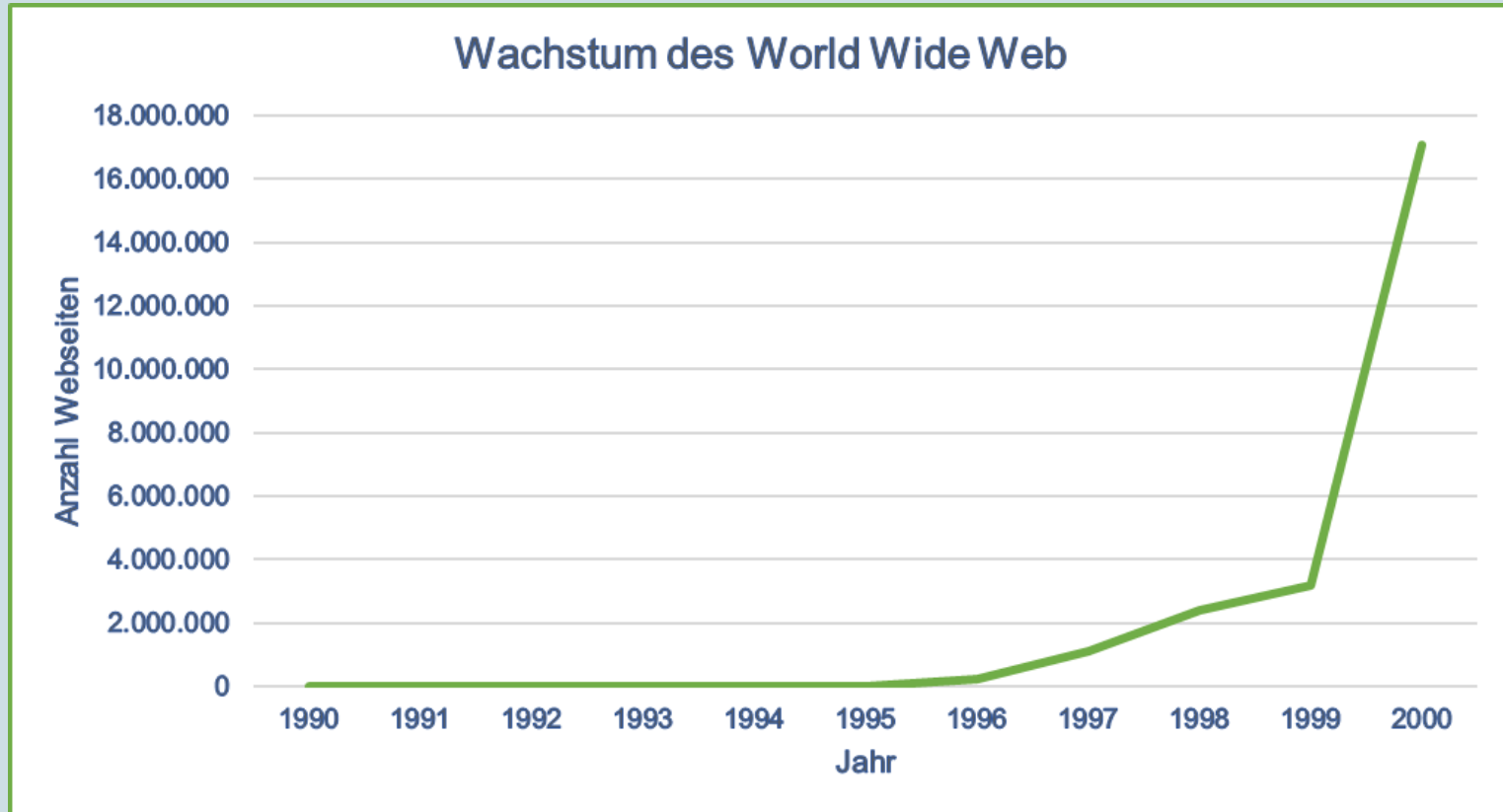
# Archivierungsansätze für (statische) Webseiten

Erfahrungen aus dem Stadtarchiv Münster

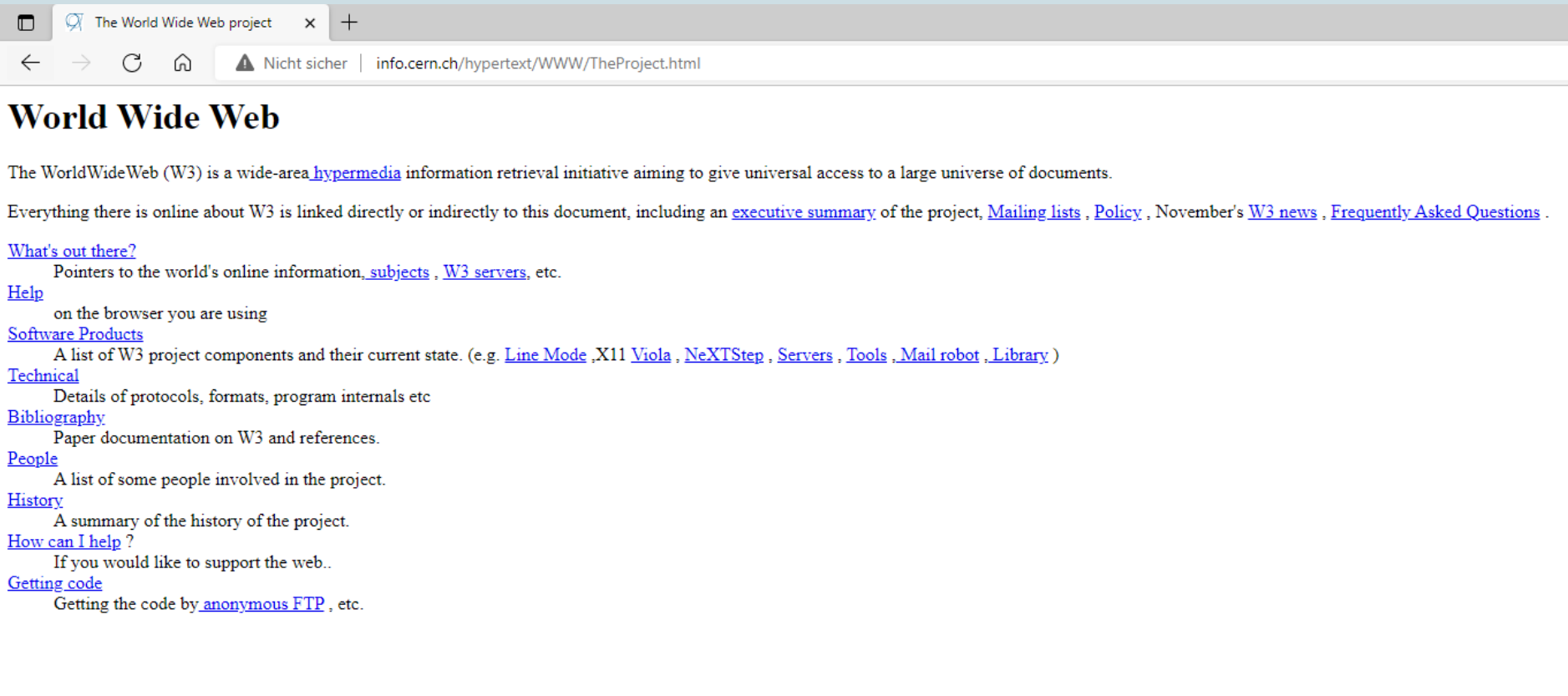
# Statische vs. Dynamische Webseiten

# „Das Internet ist für uns alle Neuland“

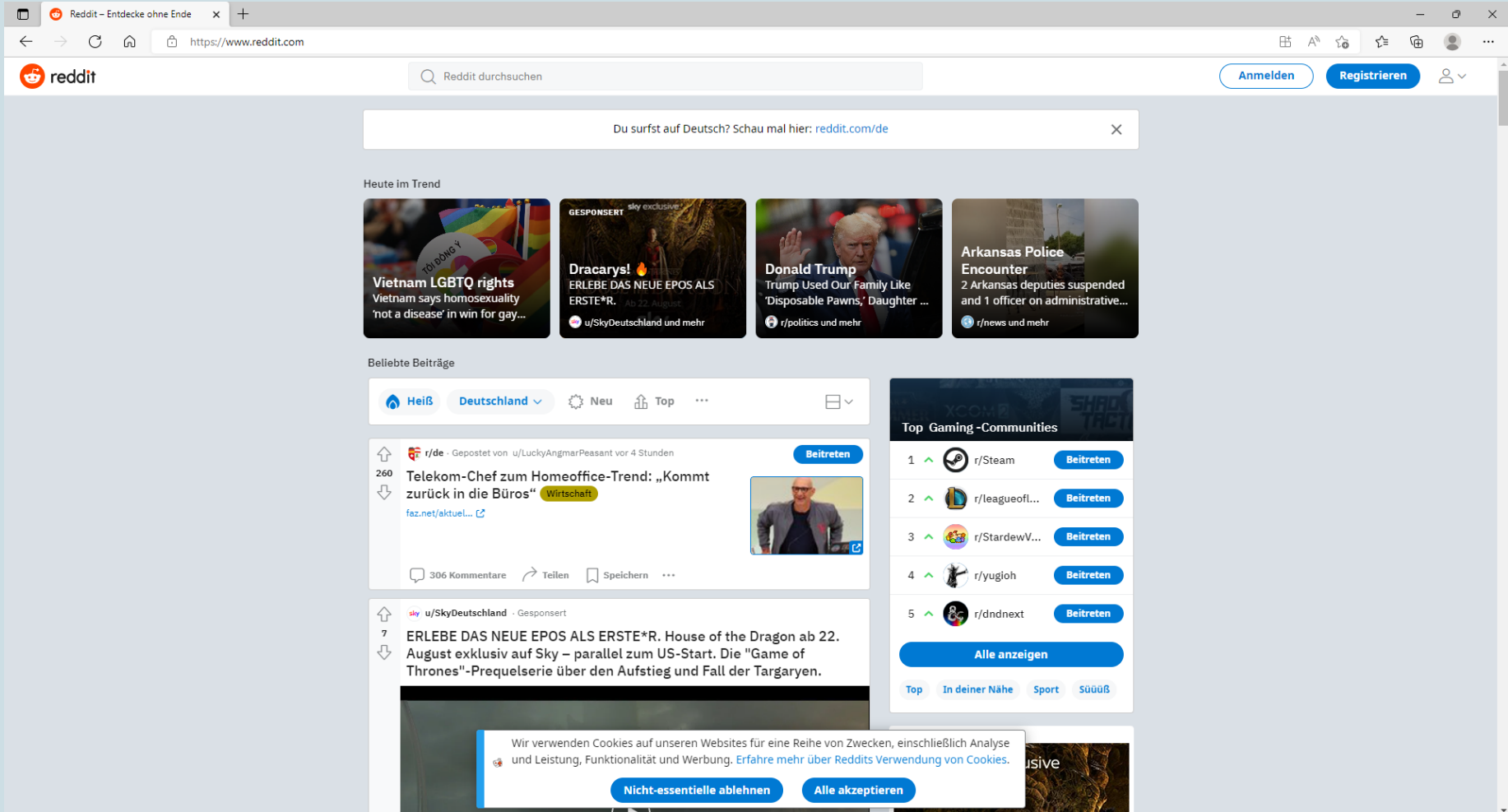
Angela Merkel, Juni 2013



Heute: 2 Milliarden



*Die erste Website der Welt vom 6. August 1991, Tim Berners-Lee*



Das größte Forum der Welt: reddit.com

# WARC - Web ARChive file format



# Alternativen zu WARC?





# Beispiel: [www.wiedertaeufer.ms](http://www.wiedertaeufer.ms)



Home Münster Klartext Rathaus Stadtgeschichte(n) Redaktionsblog Über uns Unterstützen Kontakt



**Willkommen in der virtuellen Geisterstadt!**



Die Wiedertäufer kneifen den Arsch zu



Der Lazarettbunker und das Geheimnis der kyrillischen Schriftzeichen



Die Zuckervilla: Geschichte hinter hohen Hecken



Das Haus, in dem Adolf Hitler schlief

Highlights

Unser brandneuer Ne

## Archivierungsauftrag gem. §1, Abs. 2 Nr. 5 Satzung

*„Das Stadtarchiv kann Unterlagen von anderen Stellen oder von natürlichen oder juristischen Personen als Archivgut übernehmen, soweit eine dauernde Verwahrung, Erschließung, Bereitstellung und Nutzung im öffentlichen Interesse liegen und das Archivgut Bezüge zur Stadt Münster oder ihren Rechtsvorgängern hat. Für dieses Archivgut gelten vorbehaltlich abweichender Vereinbarungen mit den Eigentümerinnen/Eigentümern des Archivguts bzw. Rechteinhaberinnen/Rechteinhabern die nachfolgenden Regelungen entsprechend.“*

## Signifikante Eigenschaften

- Die Gesamtwahrnehmung der Seite ist gewährleistet
- Das Look & Feel der Seite bleibt wahrnehmbar
- Die menschliche Lesbarkeit der Inhalte ist gewährleistet
- Die maschinelle Recherchierbarkeit der Inhalte ist gewährleistet
- Die Weiterverwendbarkeit der Inhalte ist gewährleistet
- Die Lesbarkeit von Bildern und Videos ist gewährleistet
- Die maschinelle Verarbeitbarkeit von Metadaten ist gewährleistet

# HTML to PDF - sejda

Die Webseiten bleiben noch bis zum Jahr 2025 online (ist ja schon bezahlt), danach werden die Webseiten auch in der digitalen Welt verschwinden.

22. August 2022 Ich suche nach...

**DIE WIEDERTÄUFER**

Home Münster Klartext Rathaus Stadtgeschichte(n) Redaktionsblog Über uns Unterstützen Kontakt

Home / Redaktionsblog / Die Wiedertäufer kneifen den Arsch zu

**Redaktionsblog**

## Die Wiedertäufer kneifen den Arsch zu

Von Nils Dietrich Vor 12 Monaten 14



Vereinhalb Jahre sind dann auch genug gewesen. Die Wiedertäufer stellen das Erscheinen mit sofortiger Wirkung ein. Im März 2017 sind wir mit der ersten Geschichte online gegangen. Echt lange her. Die Welt hat sich seither weiter gedreht, auch Münster. Deswegen haben wir heute nochmal zusammengefasst, was die Wiedertäufer sind und was sie für uns tun.

**Unser brandneuer Nachrichtendienst**

Email-Adresse:  
 Trag' Dich hier ein und bleibe über neue Artikel auf dem Laufenden!  
 Ausführung!

Unser Newsletter enthält Informationen zu unseren journalistischen Produkten. Hinweise zum Datenschutz, Widerruf, Protokollierung sowie der von der Einwilligung umfassten Erfolgsmessung, erhaltet ihr in unserer Datenschutzerklärung.

**Beliebteste Artikel**



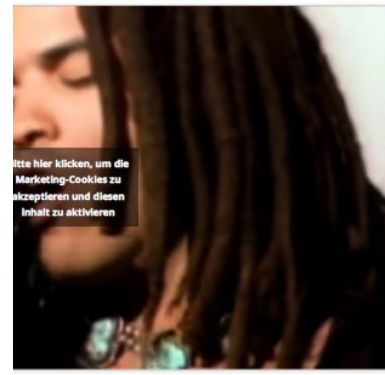
Verwaister Straßenschild: Die Not der Frauen von der Siemensstraße

Philipp Aug 27, 2021

...habe deine/eure Berichte gerne gelesen!  
 ...hoffe, dass sie trotzdem in der digitalen Welt archiviert bleiben, gibt es da Planungen über web.archive.org hinaus?

By Nils Dietrich Aug 27, 2021

Nein.



Abtreibungsgegner und Uni-Professor: Paul Callen polarisiert

Folge uns auf Facebook!

Facebook icon

Bitte hier klicken, um die Marketing-Cookies zu akzeptieren und diesen Inhalt zu aktivieren

...oder auf Twitter!

Twitter icon

Bitte hier klicken, um die Marketing-Cookies zu akzeptieren und diesen Inhalt zu aktivieren

## PDF-Ergebnisse nach der Konvertierung

## Signifikante Eigenschaften

- Die Gesamtwahrnehmung der Seite ist gewährleistet
- Das Look & Feel der Seite bleibt wahrnehmbar
- Die menschliche Lesbarkeit der Inhalte ist gewährleistet
- Die maschinelle Recherchierbarkeit der Inhalte ist gewährleistet
- Der Weiterverwendbarkeit der Inhalte ist gewährleistet
- Die Lesbarkeit von Bildern und Videos ist gewährleistet
- Die maschinelle Verarbeitbarkeit von Metadaten ist gewährleistet



# Ausblick

The screenshot shows a Twitter interface. On the left is a navigation sidebar with icons and labels: Startseite, Entdecken, Mitteilungen, Nachrichten, Lesezeichen, Listen, Profil, and Mehr. Below the sidebar is a blue button labeled 'Twittern'. The main content area shows a tweet from Markus Lewe (@LeweMarkus) dated 19. Sep. 2020. The tweet text reads: 'OB Lewe begleitete in seiner Doppelfunktion als Vizepräsident des Deutschen Städtetages und als OB von Münster den deutschen Bundespräsidenten Steinmeier bei einer Kurzreise nach Mailand. Ziel der Reise war die Vertiefung der kommunalen Zusammenarbeit innerhalb der EU, 🇪🇺🇪🇺🇪🇺'. Below the text is a photo of a meeting around a table. The tweet has 3 replies, 2 retweets, and 8 likes. Below the tweet is a reply from Lucky Look (@realeyes\_Oo) dated 19. Sep. 2020, which says: 'Antwort an @LeweMarkus: Schwätzen kann er ja. Nur handeln halt nicht. Ganz die Kanzlerin. 🙄'. The reply has 1 'Gefällt mir'-Angabe. At the bottom of the interface are icons for reply, retweet, like, and share.

# Tools

- ~~HTML to PDF Converter~~
- ~~Twitter API – eigener Export~~
- ~~Harvesting in WARC-Container~~

→ Harvesting mit **TWINT**

```
C:\Windows\system32\cmd.exe
ID 1201894887592988674 <Landschaftsverband Westfalen-Lippe <LWL> 03.12.2019 17:04:05) "Das neue #BTHG verunsichert [.]
ID 1199946871822856192 <Landschaftsverband Westfalen-Lippe <LWL> 28.11.2019 08:03:22) "Achtung, Zuckerschok! Die Kinder
[.]"
ID 1199609361032765440 <Landschaftsverband Westfalen-Lippe <LWL> 27.11.2019 09:42:13) "Die @fr berichtet über [.] : Tes
t auf 2 Antworten
  ID 1199747375184658432 <Volker Höring 27.11.2019 18:50:39) "@lwl_aktuell @fr @Museum247 Hingehen! [.]
  ID 1214893738230734848 <ev ?? 08.01.2020 13:56:53) "@lwl_aktuell @fr @Museum247 Tolle [.]! No more data! Scraping
will stop now.
found 0 deleted tweets in this search.
: Test auf 1 Antwort
  ID 1214963691084566529 <LWL-Museum für Kunst und Kultur 08.01.2020 18:34:51) "@InmersEdna @lwl_aktuell @fr Wir [
.]"! No more data! Scraping will stop now.
found 0 deleted tweets in this search.
: Test auf 1 Antwort
ID 1199313688949334022 <Landschaftsverband Westfalen-Lippe <LWL> 26.11.2019 14:07:20) "Wie steht es um [.]
ID 1197530143247536129 <Landschaftsverband Westfalen-Lippe <LWL> 21.11.2019 16:00:09) "Heute haben wir der [.] : Test a
uf 1 Antwort
ID 1195312295636848642 <Landschaftsverband Westfalen-Lippe <LWL> 15.11.2019 13:07:13) "Das @Museum247 in #Münster [.]
ID 1195312030837886978 <Landschaftsverband Westfalen-Lippe <LWL> 15.11.2019 13:06:10) "Die #LWL-Klinik #Münster hat [.]"
```

*twint-Ausgabe während des Harvestens von Antworten*

## Signifikante Eigenschaften

- Die Gesamtwahrnehmung der Seite ist gewährleistet
- Das Look & Feel der Seite bleibt wahrnehmbar
- Die menschliche Lesbarkeit der Inhalte ist gewährleistet
- Die maschinelle Recherchierbarkeit der Inhalte ist gewährleistet
- Der Weiterverwendbarkeit der Inhalte ist gewährleistet
- Die Lesbarkeit von Bildern **und Videos** ist gewährleistet
- Die maschinelle Verarbeitbarkeit von Metadaten ist gewährleistet



## Twitter-Export des Nutzers „lewemarkus“

September 2020

4 Tweets

**Markus Lewe** lewemarkus

1307350050012639236

19.09.2020, 18:05:15

OB Lewe begleitete in seiner Doppelfunktion als Vizepräsident des Deutschen Städtetages und als OB von Münster den deutschen Bundespräsidenten Steinmeier bei einer Kurzreise nach Mailand. Ziel der Reise war die Vertiefung der kommunalen Zusammenarbeit innerhalb der EU, <https://t.co/UzKY7Jnrlt>



1307350050012639236.EiSj8FYWkAcnKcCt.jpg

8 „Gefällt mir“-Angaben - 2 Retweets

**Lucky Look** realeyes\_00

1307350959266332673

19.09.2020, 18:08:52 - Antwort an LeweMarkus

@LeweMarkus Schwätzen kann er ja. Nur handeln halt nicht. Ganz die Kanzlerin. 🤖

1 „Gefällt mir“-Angabe

**Tephtor** tephtor

1307353415593361408

19.09.2020, 18:18:38 - Antwort an LeweMarkus

@LeweMarkus Mailand oder Madrid - Hauptsache Italien

1 „Gefällt mir“-Angabe

*Heutiges Ergebnis  
des TWINT-basierten  
Exports*

## Fazit

- Kein einheitlicher Lösungsweg
- Jede Webseite muss erneut geprüft werden
  - **Signifikante Eigenschaften**
  - **Erhaltungschance**
  - **Methodenauswahl**
  - **Protokollierung (Metadaten)**
- Komplexe Einzellösungen ergeben „Web-Archivierungsbaukasten“