What you and I keep doing wrong when searching the web

EIJC 2019, Mechelen, 17th of May

Marcus Lindemann,
TV journalist and investigation specialist
lindemann@autorenwerk.de
www.autorenwerk.de www.recherche-info.de

Marcus Lindemann

- TV journalist with my own production house – documentaries, magazine stories for ZDF and other, mainly public, broadcasters

- Journalism lecturer and trainer since 1998, specialised in computer based investigations

- Seminars and trainings for journalists at all career levels
  - lecturer at universities and journalism schools
  - inhouse for newspapers, agencies, stations
  - workshops and trainings abroad

- since 2011 presentations and trainings for the insurance industry and the police in Germany, Austria and Switzerland

- Previously researcher, editor and reporter at ZDF
- Freelance journalist for regional papers, magazine, radio and TV
- MA in Political Economy (UK)

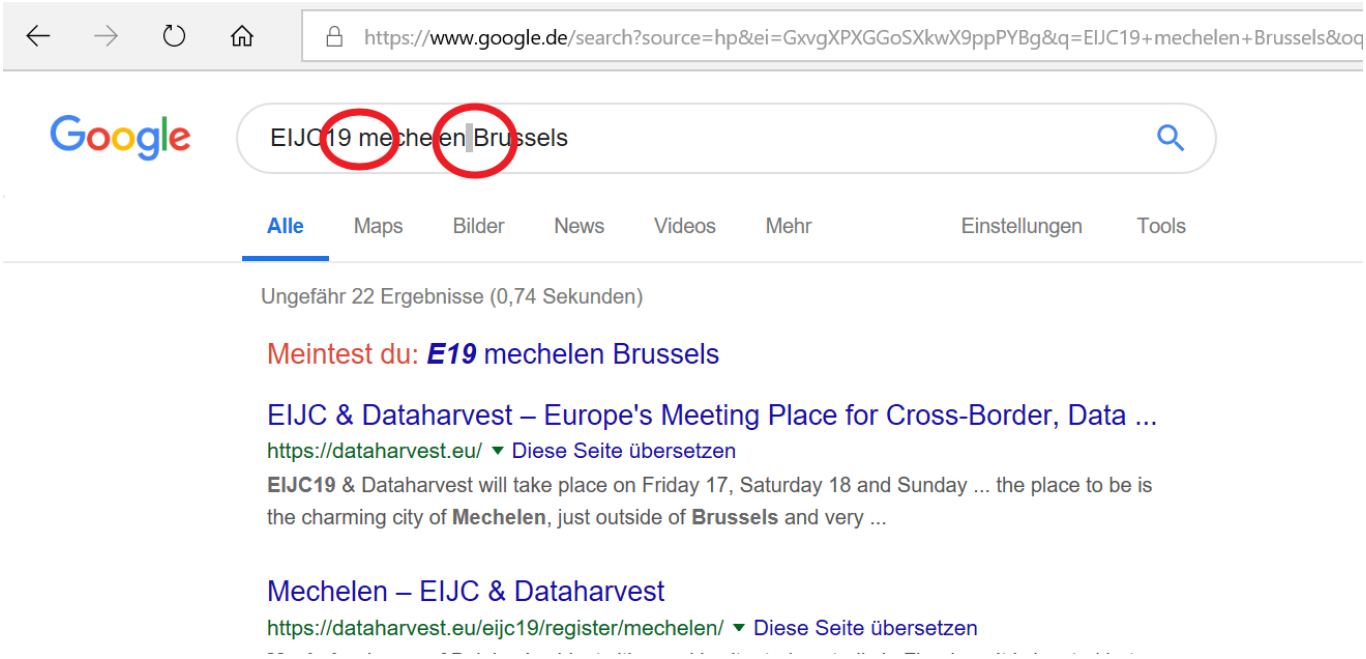
Structure of this session

I. Use Google properly
II. Think beyond Google
III. Other recommendations

I. Use Google properly

I. Use Google properly

1. There is an invisible AND.



I. Use Google properly

1. There is an invisible AND.

Take more care on choosing search terms.

I. Use Google properly

1. There is an invisible AND.

Take more care on choosing search terms.

“Search terms” now refers to full text, not keywords only

I. Use Google properly

1. There is an invisible AND.
2. You cannot search for #,§, @, ? (but for €, $ and &).

I. Use Google properly

1. There is an invisible AND.
2. You cannot search for §, @, € (but for $).
3. Use less quotations marks.

Use less quotation marks

[Curt Engelhorn]

Use less quotation marks

[“Curt Engelhorn”]

Use less quotation marks

What are you missing if you go for ["Curt Engelhorn"]?

Use less quotation marks

What are you missing if you go for ["Curt Engelhorn"]?

Engelhorn, Curt

Curt Glover Engelhorn/ Curt G. Engelhorn

Mr. Engelhorn

Kurt Engelhorn

Use less quotation marks

[“Curt Engelhorn” OR “Curt * Engelhorn” OR
“Engelhorn, Curt]

Use less quotation marks

[“Curt Engelhorn” OR “Curt * Engelhorn” OR
“Engelhorn, Curt OR “Kurt Engelhorn” OR
“Kurt * Engelhorn” OR “Engelhorn, Kurt]

Use less quotation marks
Quick & dirty

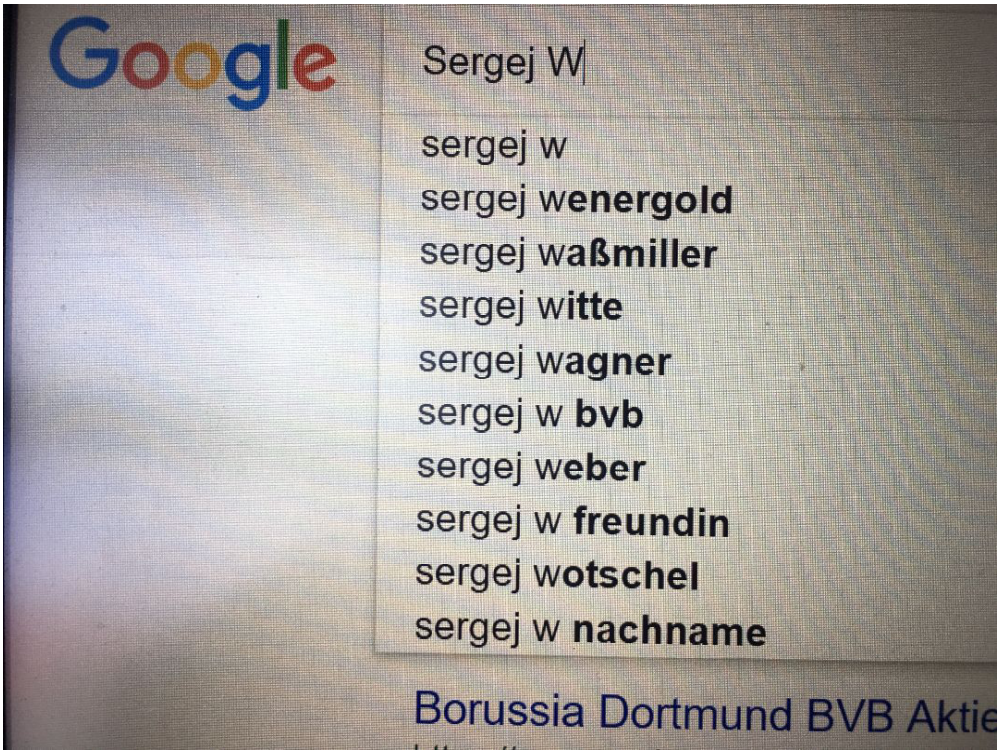
[Curt Engelhorn] is better than

[“Curt Engelhorn”]

I. Use Google properly

1. There is an invisible AND.
2. You cannot search for §, @, € (but for $).
3. Use less quotations marks.
4. Do not forget about other Google services
   (suggest, video, alerts, CSE, Earth)

Make use of google suggestions



https://www.google.de/search?biw=1280&bih=603&tbm=vid&ei=yI3eXKKWJZ-g1fAPhqmcKA&q

Google

investigativ

Alle Bilder News **Videos** Maps Mehr Einstellungen Tools

Ungefähr 330.000 Ergebnisse (0,21 Sekunden)

so geht MEDIEN | Lexikon : Investigativ-Journalismus - BR

https://www.br.de/.../so-geht-medien-lexikon-investigativ-journalismus-av:5acc840def8...

14.04.2018

Man könnte sie als "Detektiv-Journalisten" bezeichnen, die Reporterinnen und Reporter, die ...

Journalismus 2.0 – Investigativ, unabhängig und unbequem – Wenn ...

https://www.youtube.com/watch?v=I6co7CV5Ufw

18.05.2018 - Hochgeladen von DAI Heidelberg

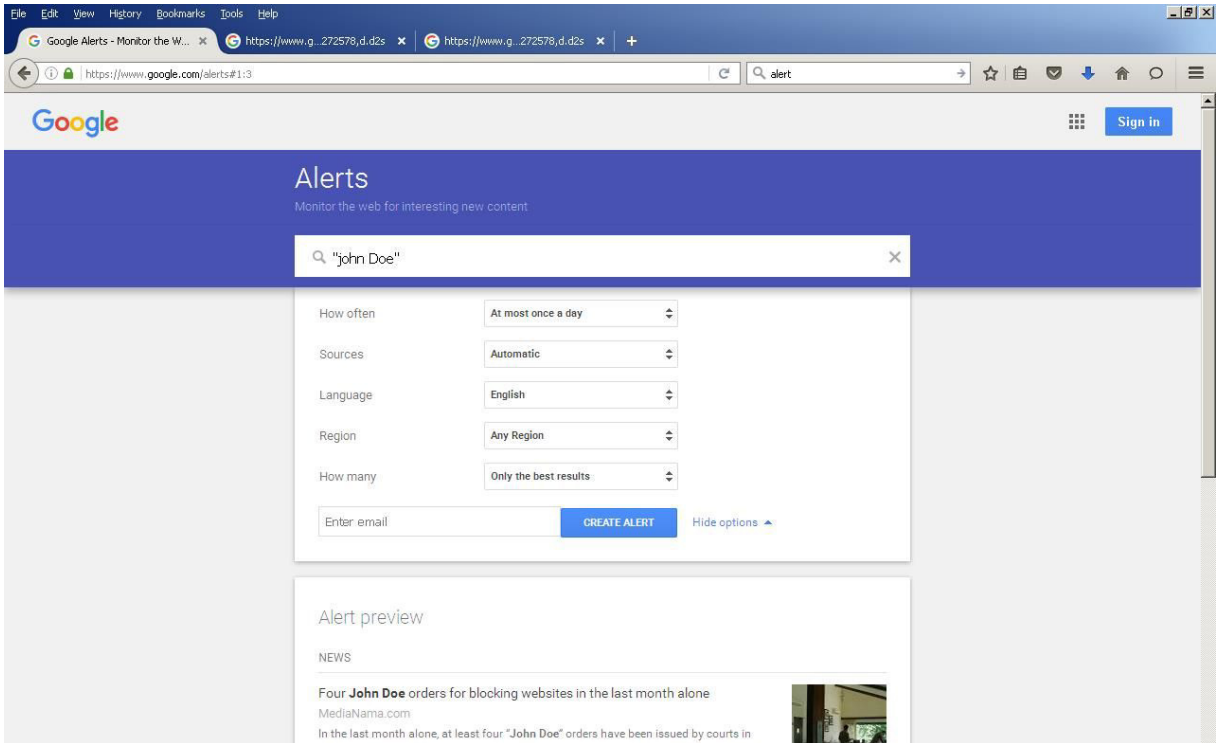
Watergate, Panama Papers, NSA-Skandal – alle Fälle wurden vom **investigativen** Journalismus ...

Die Nachwuchsjournalistin : Was ist Investigativer Journalismus? - BR

https://www.br.de/.../die-nachwuchsjournalistin-was-ist-investigativer-journalismus-av:5...

Was macht **investigativen** Journalismus aus? Worauf kommt es bei der Recherche an? Wir fragen eine ...

Control
for upcoming results with alerts



II. Think beyond Google

II. Think beyond Google

1. Do not forget about deep web sources.

II. Think beyond Google

1. Do not forget about deep web sources.

    - libraries
    - databases in general
    - official databases
    - closed forums

II. Think beyond Google

1. Do not forget about deep web sources.
2. Do not forget offline sources – it can be so much faster.

II. Think beyond Google

1. Do not forget about deep web sources.
2. Do not forget offline sources – it can be so much faster.
    - colleagues
    - telephone
    - people
    - books

II. Think beyond Google

1. Do not forget about deep web sources.
2. Do not forget offline sources – it can be so much faster.
3. For twitter and Facebook there are better options.

II. Think beyond Google

1. Do not forget about deep web sources.
2. Do not forget offline sources – it can be so much faster.
3. For twitter and Facebook there are better options.
   - twitter.com/search-home
   - facebook graph search

II. Think beyond Google

1. Do not forget about deep web sources.
2. Do not forget offline sources – it can be so much faster.
3. For twitter and Facebook there are better options.
4. There are three tasks where Google is no longer the best choice.

3 tasks where Google is no longer best

Similar images

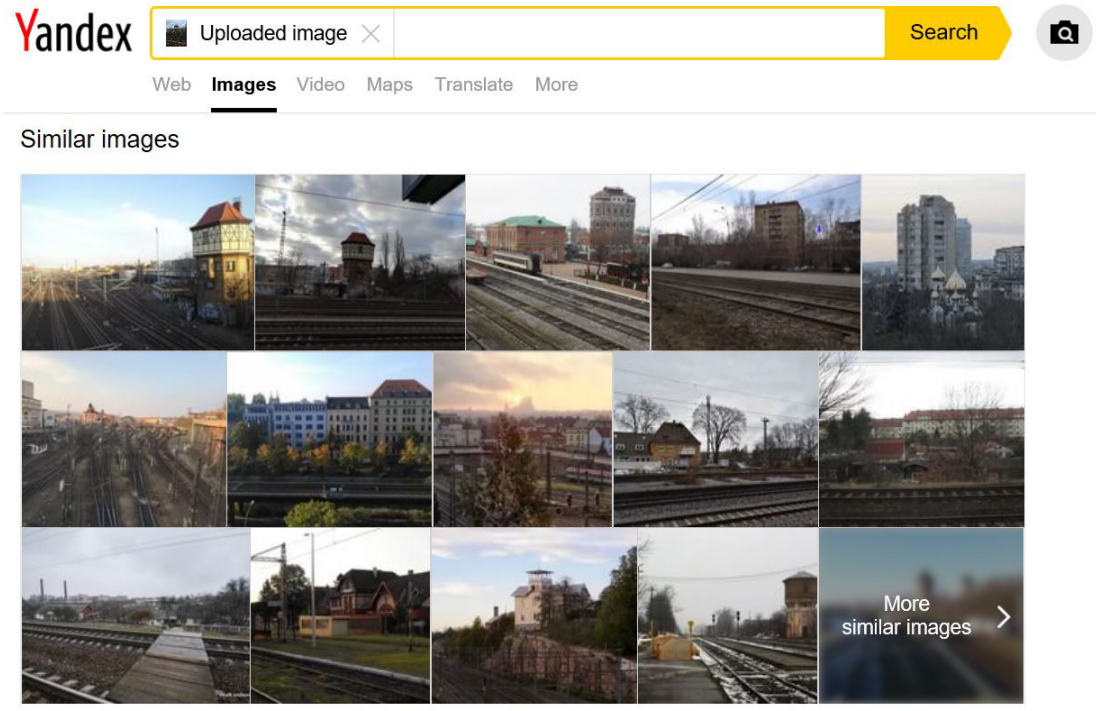
3 tasks where Google is no longer best

Similar images

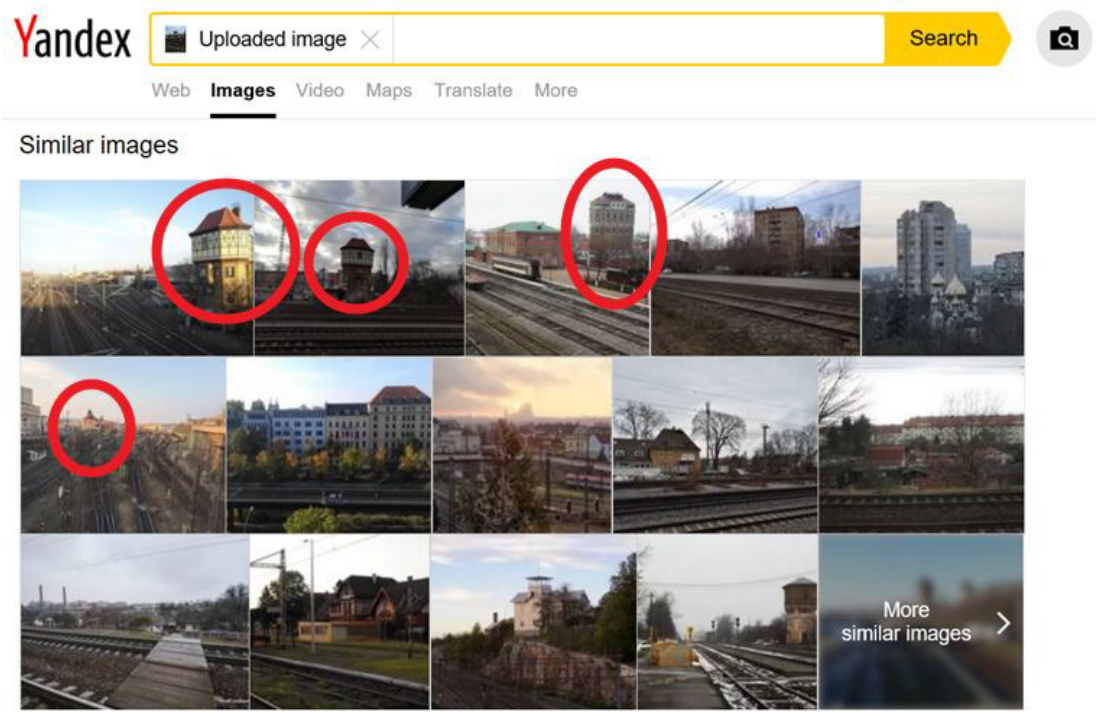
Optisch ähnliche Bilder

Unangemessene Bilder melden

3 tasks where Google is no longer best

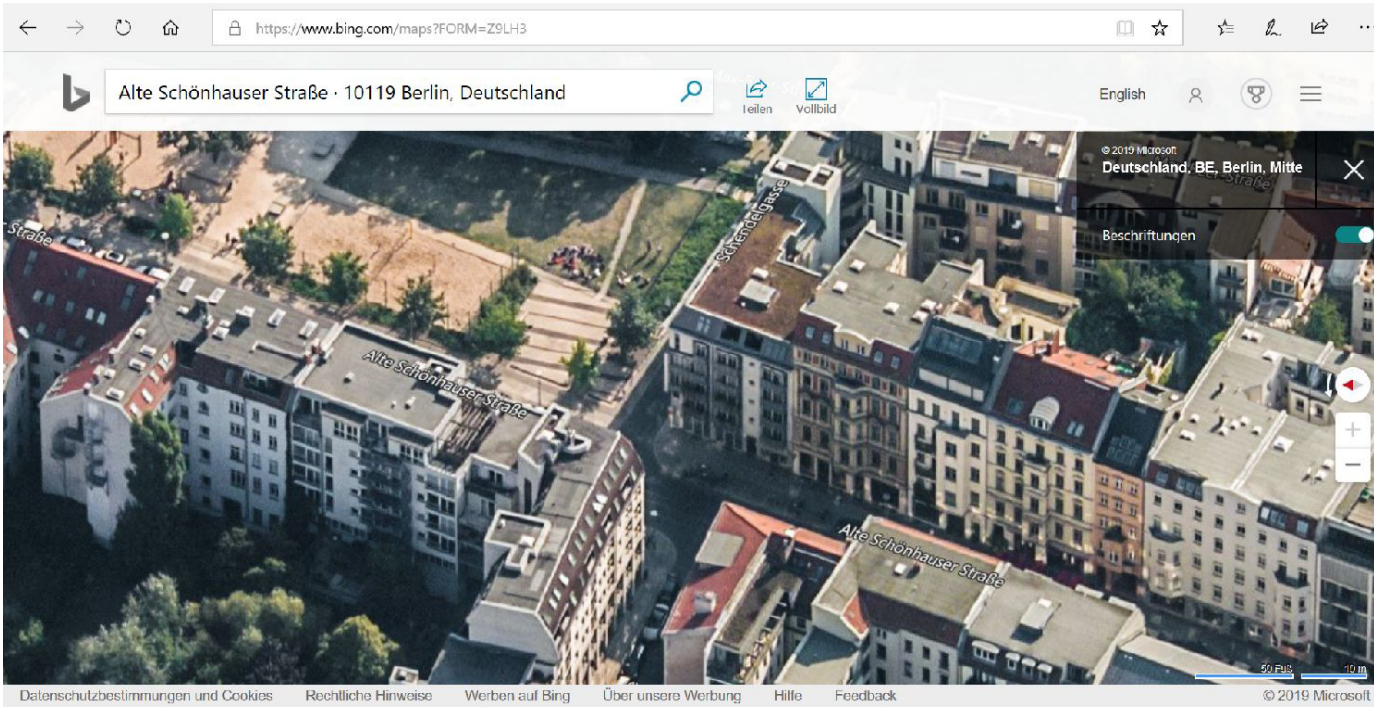


3 tasks where Google is no longer best

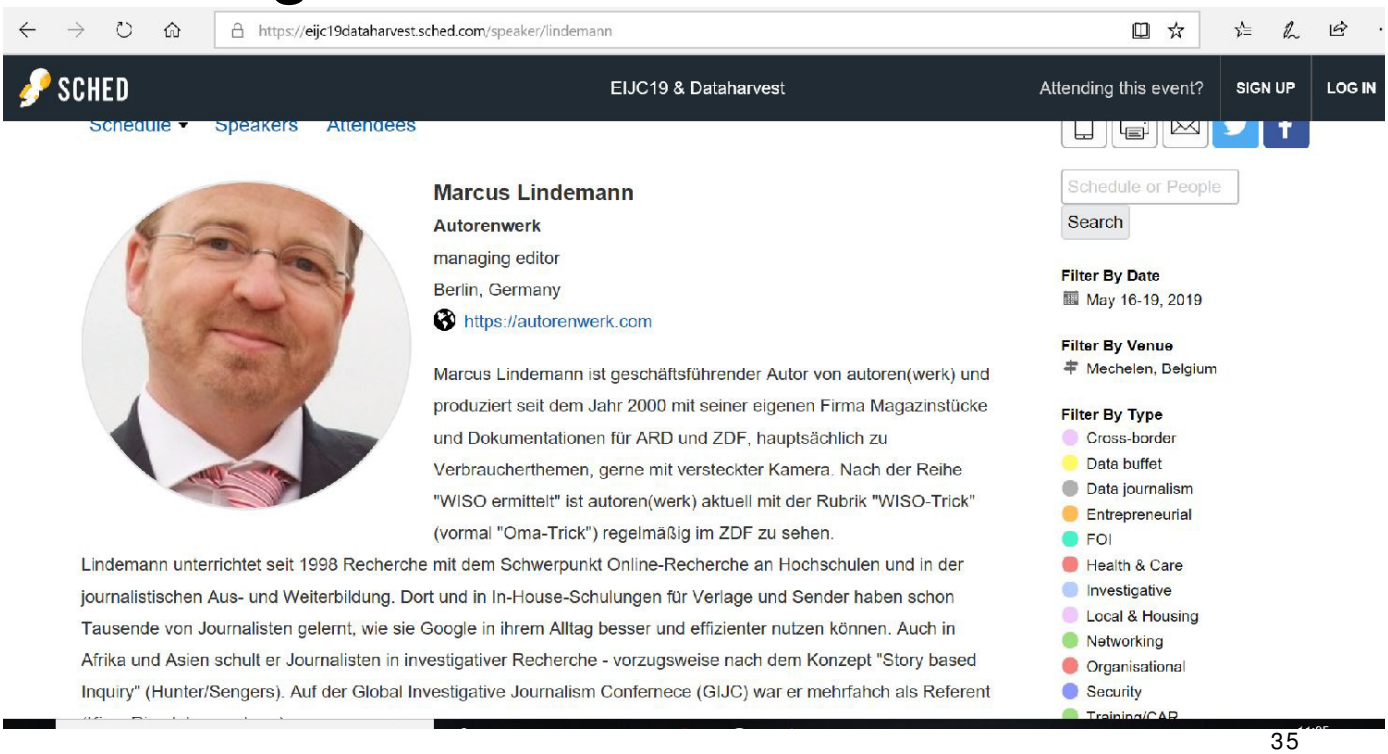


3 tasks where Google is no longer best

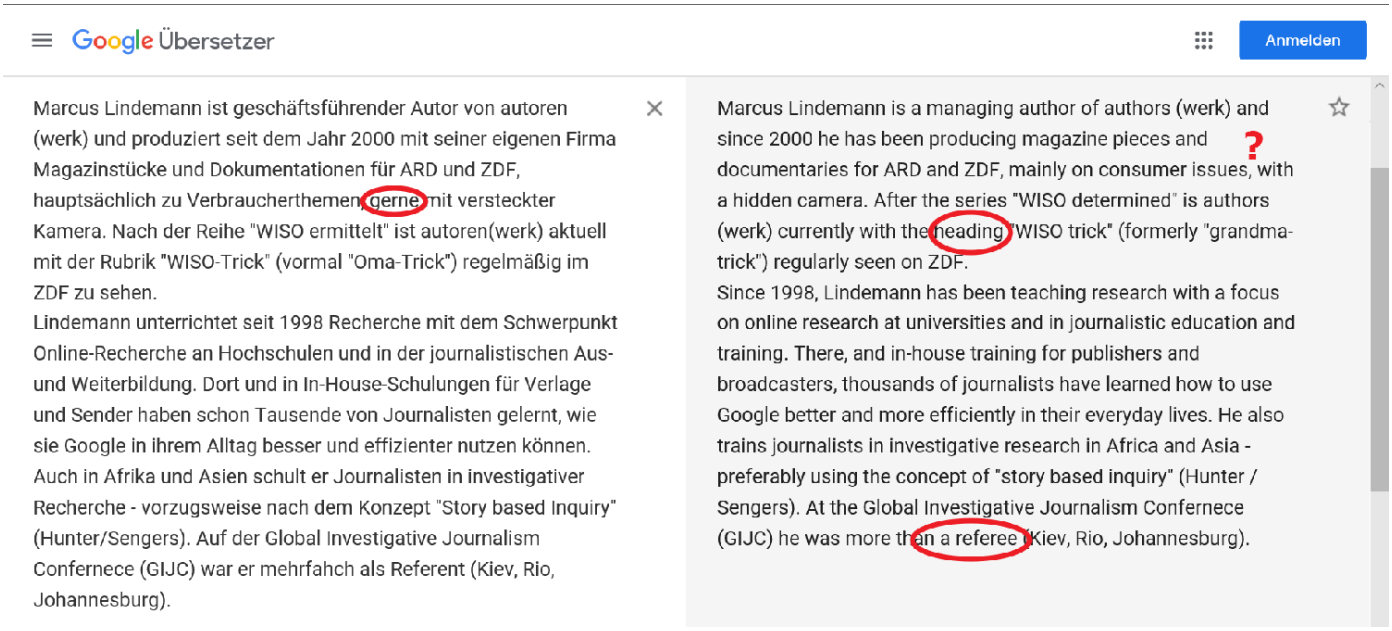
3 tasks where Google is no longer best



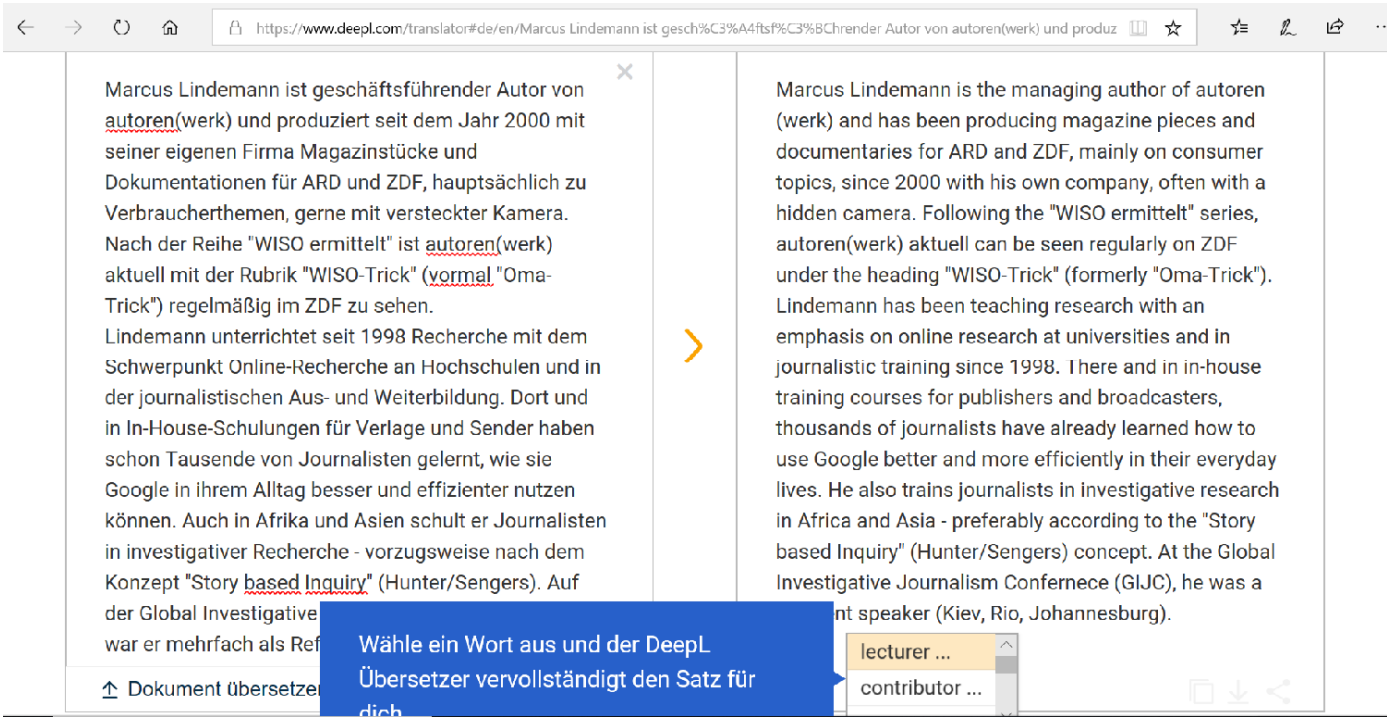
3 tasks where Google is no longer best



3 tasks where Google is no longer best



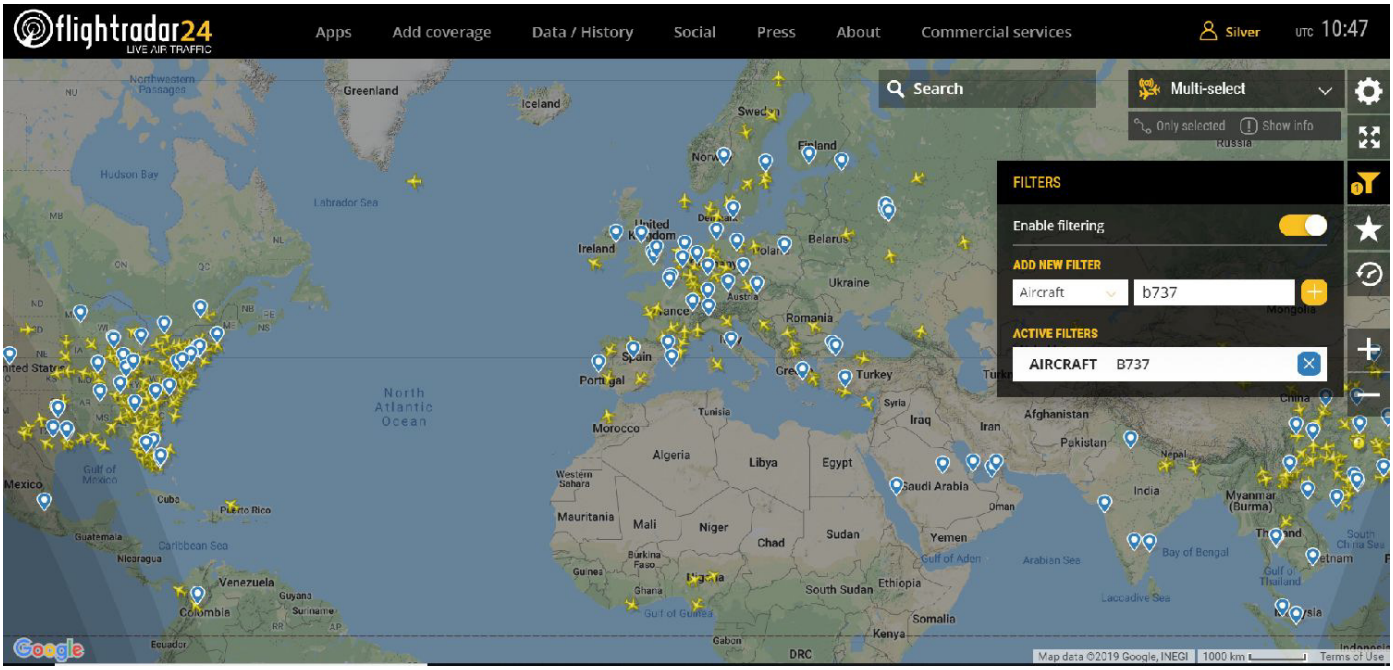
3 tasks where Google is no longer best

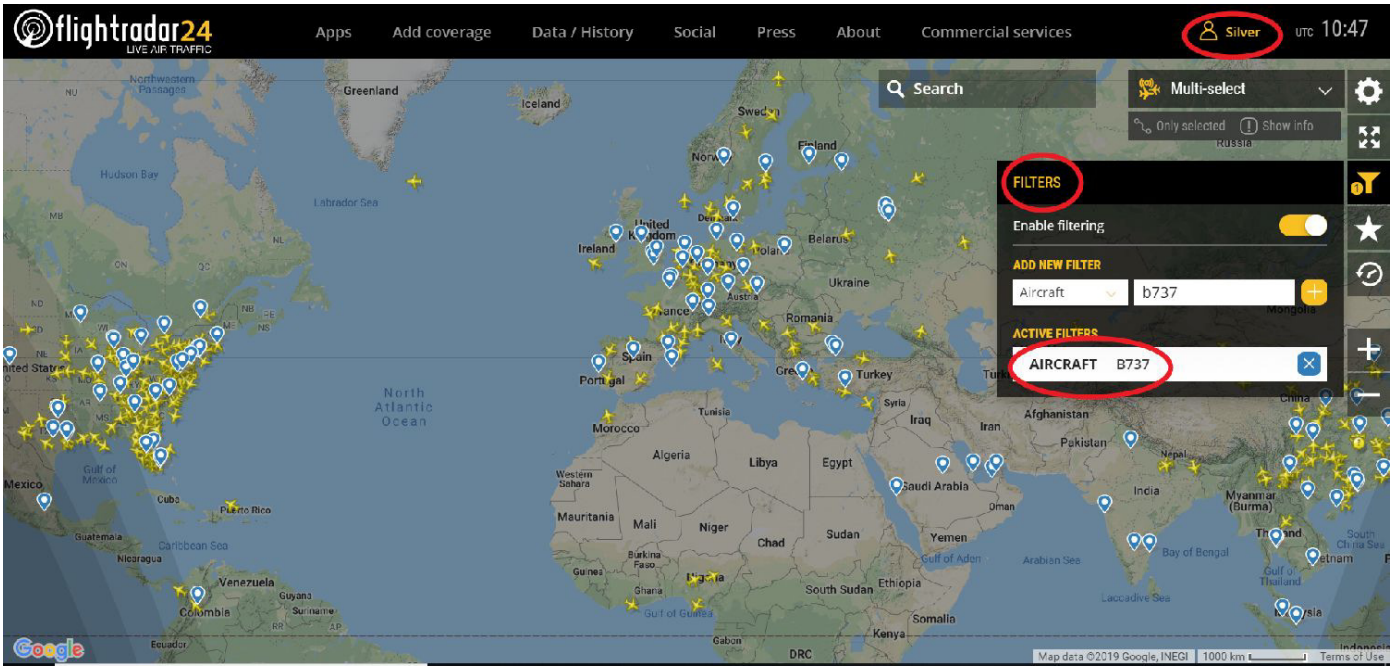


II. Other recommendations

II. Other recommendations

1. If necessary, register for services.

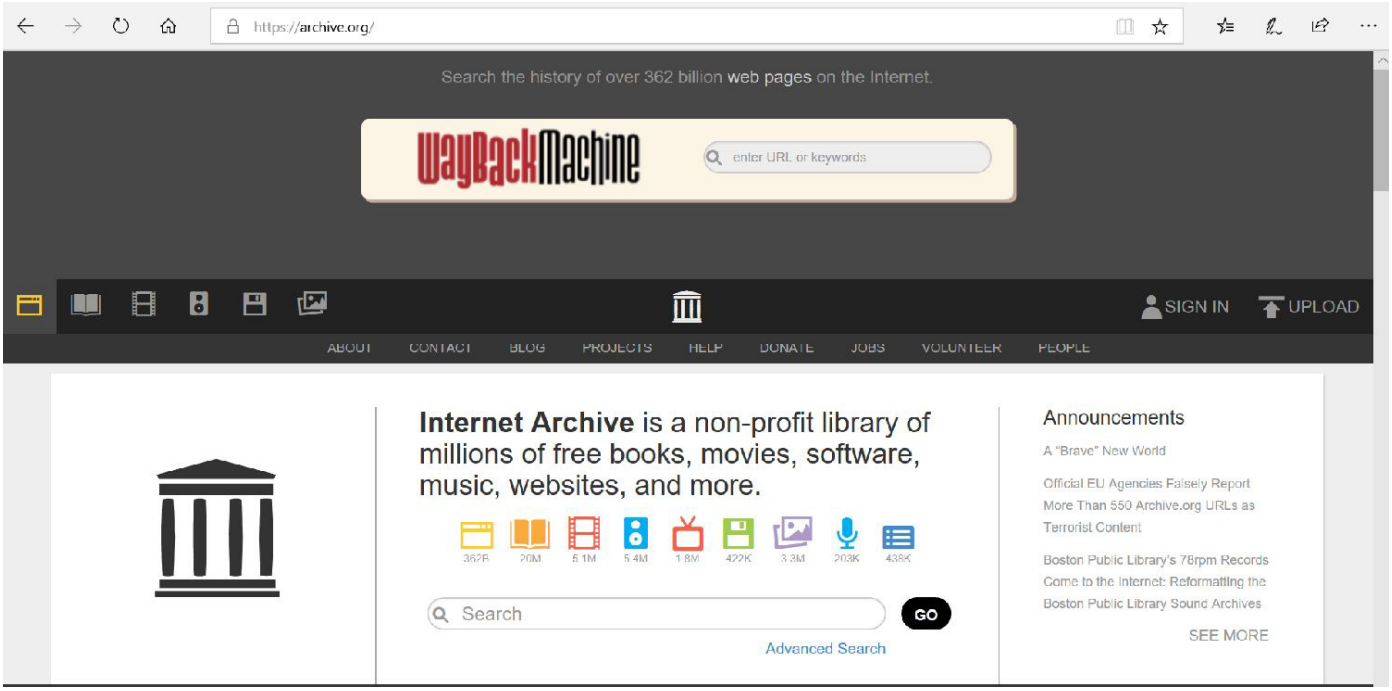

flightradar24
LIVE AIR TRAFFIC
Apps
Add coverage
Data / History
Social
Press
About
Commercial services
Silver
UTC 10:47
Search
Multi-select
Only selected
Show info
FILTERS
Enable filtering
ADD NEW FILTER
Aircraft
b737
ACTIVE FILTERS
AIRCRAFT B737
Greenland
Iceland
Hudson Bay
Labrador Sea
North Atlantic Ocean
Ireland
Belarus
Ukraine
Romania
Turkey
Morocco
Algeria
Libya
Egypt
Mauritania
Mali
Niger
Chad
Sudan
Saudi Arabia
Yemen
Iraq
Iran
Afghanistan
Pakistan
India
Mexico
Cuba
Puerto Rico
Venezuela
Colombia
Guatemala
Nicaragua
Caribbean Sea
Gulf of Mexico
Gulf of Guinea
Gulf of Aden
Arabian Sea
Bay of Bengal
Laccadive Sea
South Sudan
Ethiopia
Somalia
Kenya
Gabon
DRC
Google
Map data ©2019 Google, INEGI
1000 km
Terms of Use


flightradar24
LIVE AIR TRAFFIC
Apps
Add coverage
Data / History
Social
Press
About
Commercial services
Silver
UTC 10:47
Search
Multi-select
Only selected
Show info
FILTERS
Enable filtering
ADD NEW FILTER
Aircraft
b737
ACTIVE FILTERS
AIRCRAFT B737
Greenland
Iceland
Hudson Bay
Labrador Sea
North Atlantic Ocean
Ireland
Belarus
Ukraine
Romania
Turkey
Morocco
Algeria
Libya
Egypt
Mauritania
Mali
Niger
Chad
Sudan
Saudi Arabia
Yemen
Iraq
Iran
Afghanistan
Pakistan
India
Mexico
Cuba
Puerto Rico
Venezuela
Colombia
Guatemala
Nicaragua
Caribbean Sea
Gulf of Mexico
Gulf of Guinea
Gulf of Aden
Arabian Sea
Bay of Bengal
Laccadive Sea
South Sudan
Ethiopia
Somalia
Kenya
DRC
Google
Map data ©2019 Google, INEGI
1000 km
Terms of Use

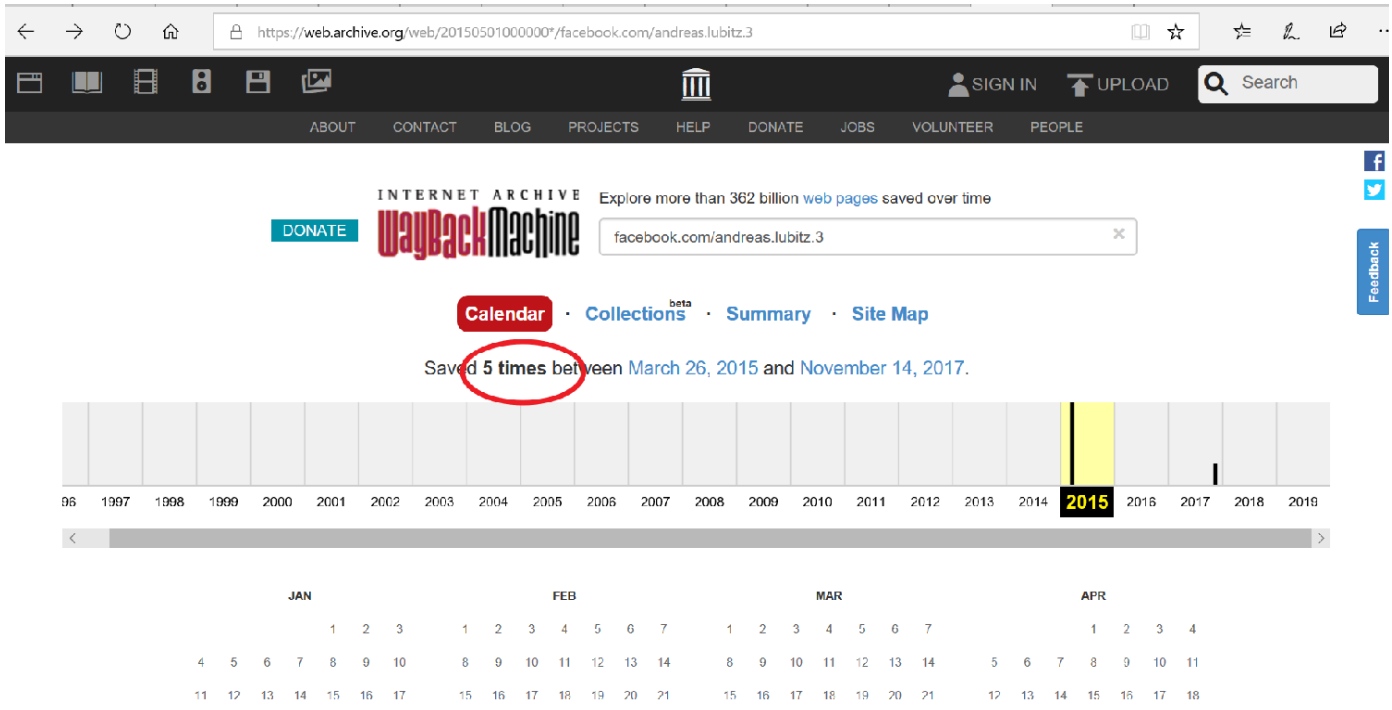
II. Other recommendations

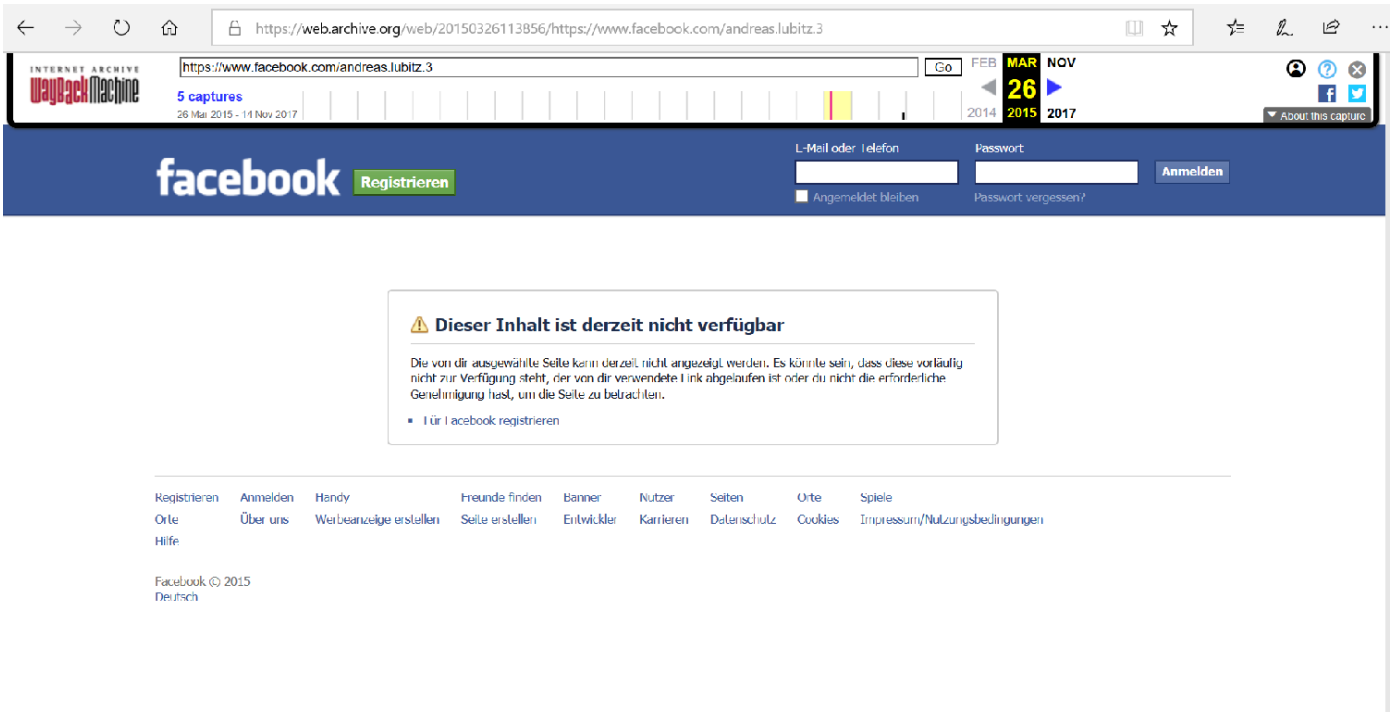
1. If necessary, register for services.
2. The wayback maschine is great, but…

The wayback maschine is great, but…



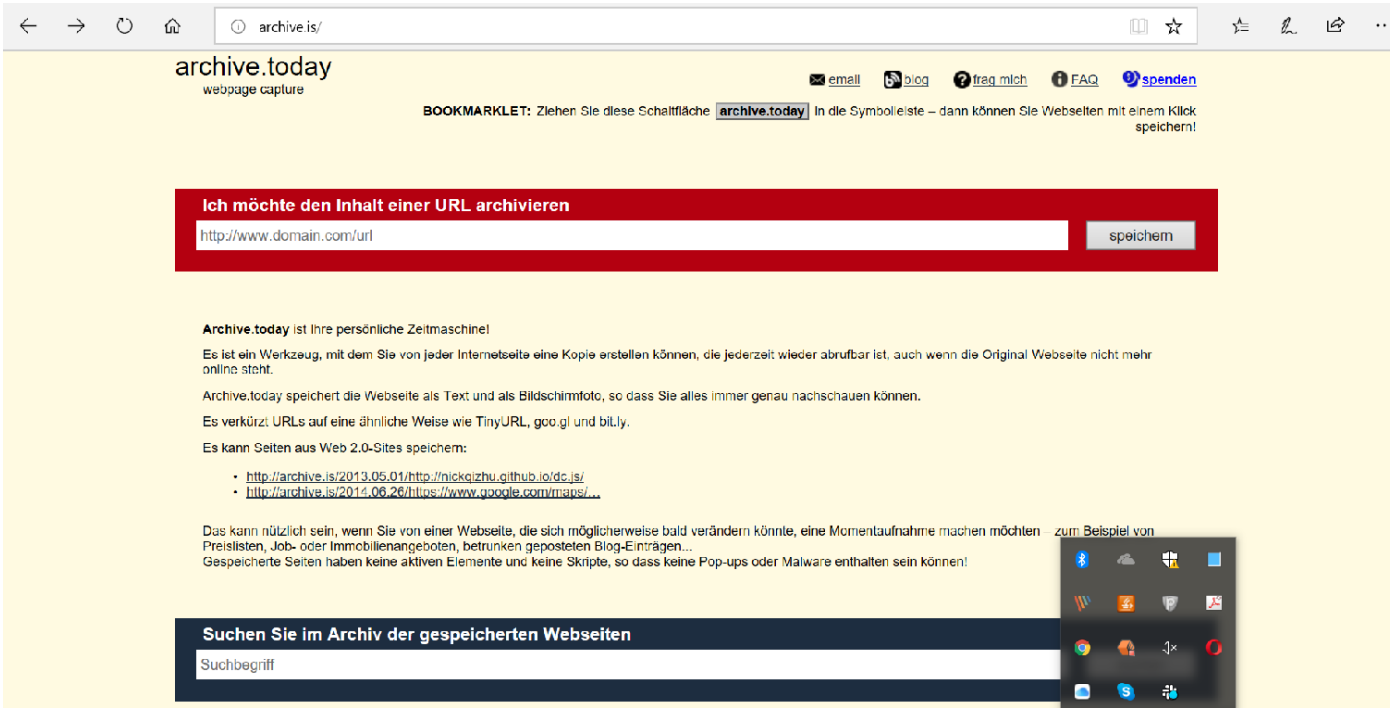
The wayback maschine is great, but…

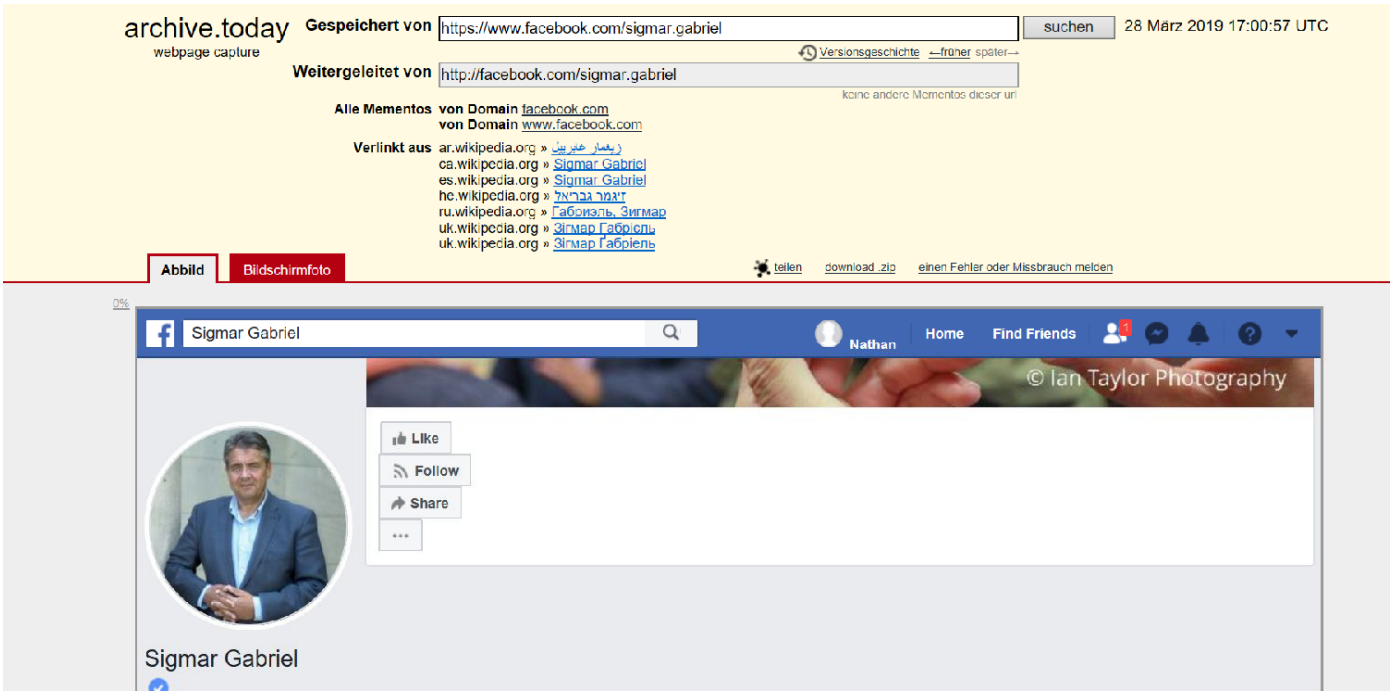



https://web.archive.org/web/20150326113856/https://www.facebook.com/andreas.lubitz.3
INTERNET ARCHIVE
WaybackMachine
https://www.facebook.com/andreas.lubitz.3
Go
5 captures
26 Mar 2015 - 14 Nov 2017
FEB
MAR
NOV
26
2014
2015
2017
About this capture
facebook
Registrieren
L-Mail oder Telefon
Passwort
Anmelden
Angemeldet bleiben
Passwort vergessen?
Dieser Inhalt ist derzeit nicht verfügbar
Die von dir ausgewählte Seite kann derzeit nicht angezeigt werden. Es könnte sein, dass diese vorläufig nicht zur Verfügung steht, der von dir verwendete Link abgelaufen ist oder du nicht die erforderliche Genehmigung hast, um die Seite zu betrachten.
Für Facebook registrieren
Registrieren
Anmelden
Handy
Freunde finden
Banner
Nutzer
Seiten
Orte
Spiele
Orte
Über uns
Werbeanzeige erstellen
Seite erstellen
Entwickler
Karrieren
Datenschutz
Cookies
Impressum/Nutzungsbedingungen
Hilfe
Facebook © 2015
Deutsch

```
User-agent: ia_archiver
Disallow: /
Disallow: /ajax/
Disallow: /album.php
Disallow: /checkpoint/
Disallow: /contact_importer/
Disallow: /dialog/
Disallow: /fbml/ajax/dialog/
Disallow: /feeds/
Disallow: /file_download.php
Disallow: /hashtag/
Disallow: /l.php
Disallow: /live/
Disallow: /moments_app/
Disallow: /p.php
Disallow: /photo.php
Disallow: /photos.php
Disallow: /share.php
Disallow: /share/
Disallow: /sharer.php
Disallow: /sharer/

User-agent: msnbot
Disallow: /ajax/
Disallow: /album.php
Disallow: /checkpoint/
Disallow: /contact_importer/
```


archive.is/
archive.today
webpage capture
email
blog
frag mich
FAQ
spenden
BOOKMARKLET: Ziehen Sie diese Schaltfläche archive.today in die Symbolleiste – dann können Sie Webseiten mit einem Klick speichern!
Ich möchte den Inhalt einer URL archivieren
http://www.domain.com/url
speichern
Archive.today ist Ihre persönliche Zeitmaschine!
Es ist ein Werkzeug, mit dem Sie von jeder Internetseite eine Kopie erstellen können, die jederzeit wieder abrufbar ist, auch wenn die Original Webseite nicht mehr online steht.
Archive.today speichert die Webseite als Text und als Bildschirmfoto, so dass Sie alles immer genau nachschauen können.
Es verkürzt URLs auf eine ähnliche Weise wie TinyURL, goo.gl und bit.ly.
Es kann Seiten aus Web 2.0-Sites speichern:
http://archive.is/2013.05.01/http://nickqizhu.github.io/dc.js/
http://archive.is/2014.06.26/https://www.google.com/maps/...
Das kann nützlich sein, wenn Sie von einer Webseite, die sich möglicherweise bald verändern könnte, eine Momentaufnahme machen möchten – zum Beispiel von Preislisten, Job- oder Immobilienangeboten, betrunken geposteten Blog-Einträgen...
Gespeicherte Seiten haben keine aktiven Elemente und keine Skripte, so dass keine Pop-ups oder Malware enthalten sein können!
Suchen Sie im Archiv der gespeicherten Webseiten
Suchbegriff


archive.today
webpage capture
Gespeichert von
https://www.facebook.com/sigmar.gabriel
suchen
28 März 2019 17:00:57 UTC
Versionsgeschichte ←früher später→
Weitergeleitet von
http://facebook.com/sigmar.gabriel
keine andere Mementos dieser url
Alle Mementos von Domain facebook.com
von Domain www.facebook.com
Verlinkt aus ar.wikipedia.org » زيغمار غبريل
ca.wikipedia.org » Sigmar Gabriel
es.wikipedia.org » Sigmar Gabriel
hc.wikipedia.org » זיגמר גבריאל
ru.wikipedia.org » Габриэль, Зигмар
uk.wikipedia.org » Зігмар Габріель
uk.wikipedia.org » Зігмар Габріель
Abbild
Bildschirmfoto
teilen
download .zip
einen Fehler oder Missbrauch melden
Sigmar Gabriel
Nathan
Home
Find Friends
© Ian Taylor Photography
Like
Follow
Share
Sigmar Gabriel

II. Further recommendations from participants

3. http://timetravel.mementoweb.org/
   Meta-Search engine for archived data
4. https://hunter.io

   identifies the e-mail-structure of organisations and companies
5. https://checkwa.online/wp/

   Gives you whatsapp pictures & status

Any questions left?
Get in contact!

lindemann@autorenwerk.de

Marcus Lindemann
mobile: +49-178-257619-0

office +49-30-257619-12

recherche-info.de