

# TECHNISCHES SEO 2019

**ALONA DEMCHYK**  
**SEO CAMPIXX 2019 / BERLIN**

# ALONA DEMCHYK

## SEO CONSULTANT



**Alona Demchyk**

<https://alonademchyk.com>

### SEO Consultant

- » Bei Plan.Net Performance seit 2015
- » 11 Jahre Suchmaschinenoptimierung/Performance Marketing
- » Erfahrungen: BSH, MediaMarkt, ExLibris, Carglass, IKEA

### Online auf:

LinkedIn: <https://www.linkedin.com/in/alonademchyk>

Xing: [https://www.xing.com/profile/Alona\\_Demchyk](https://www.xing.com/profile/Alona_Demchyk)

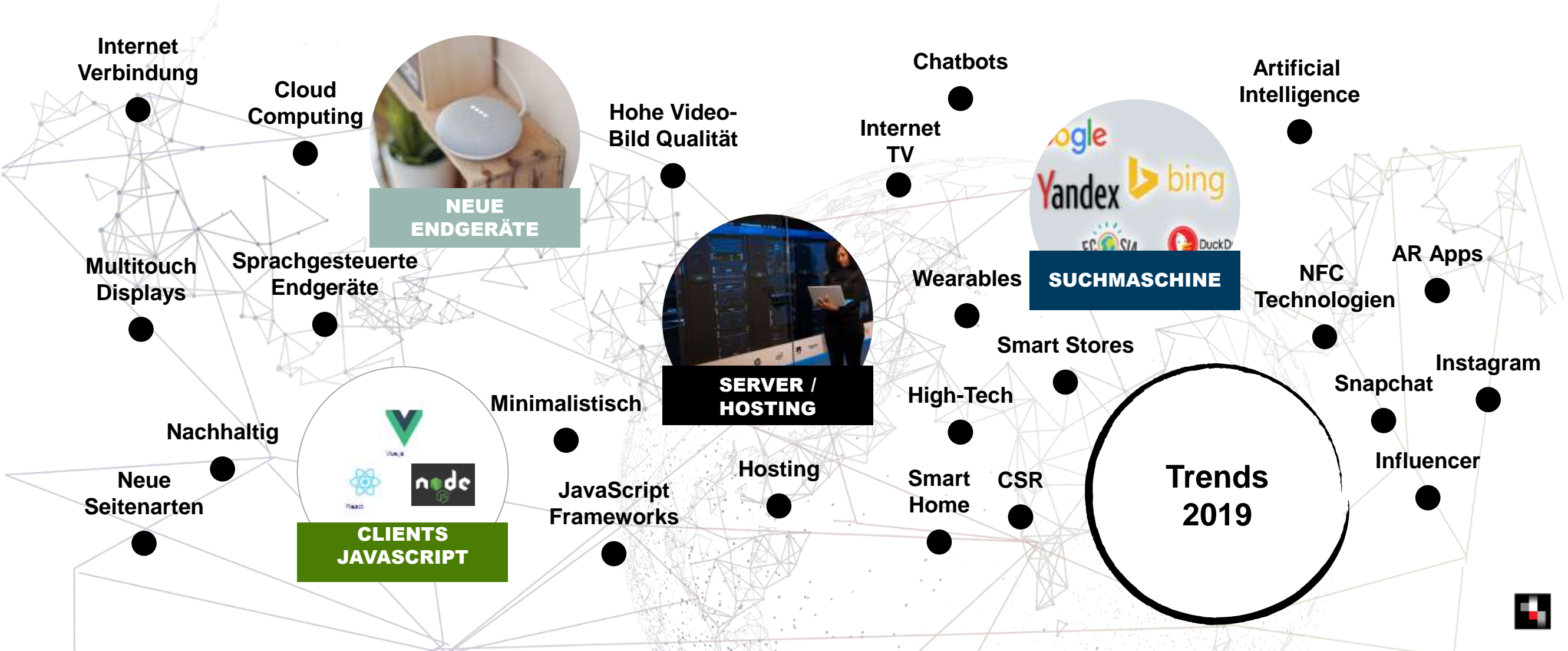
Twitter: <https://twitter.com/alonademchyk>



# INHALTE

1. Trends 2019
2. Herausforderungen
3. Serverseitige Optimierung
4. Server & Client Kommunikation
5. Clientseitige Optimierung
6. Suchmaschinen
7. Crawling- & Indexierungsprobleme
8. Identifikation von Problemen

# DIE TRENDS VON HEUTE



**DIE NUTZER WOLLEN**  
**SCHNELLER**  
**INFORMATIONEN**  
**ERHALTEN**  
**UNABHÄNGIG VOM**  
**ENDGERÄT ODER**  
**STANDORT**



# **DIE SCHNELLSTE INTERNET- VERBINDUNG GIBT ES IN SÜDKOREA MIT Ø 28.6 MBPS**



**IN  
DEUTSCHLAND  
Ø 15.3 MBPS ...**

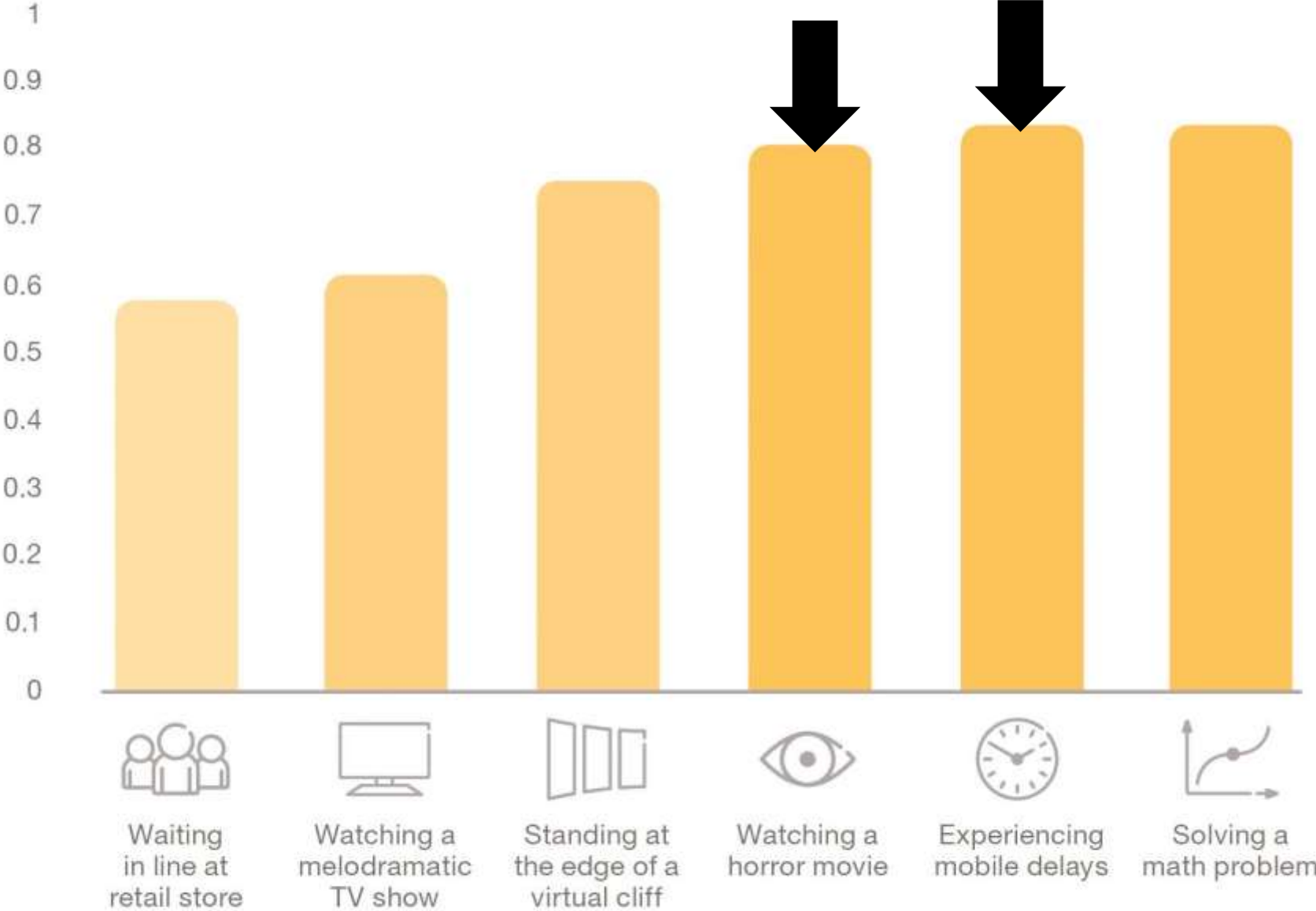


# IN DEUTSCHLAND MOBILE Ø 24.1 MBPS





# Cognitive load associated with stressful situations



Source: Ericsson ConsumerLab, Neurons Inc., 2015



As page load time goes from:

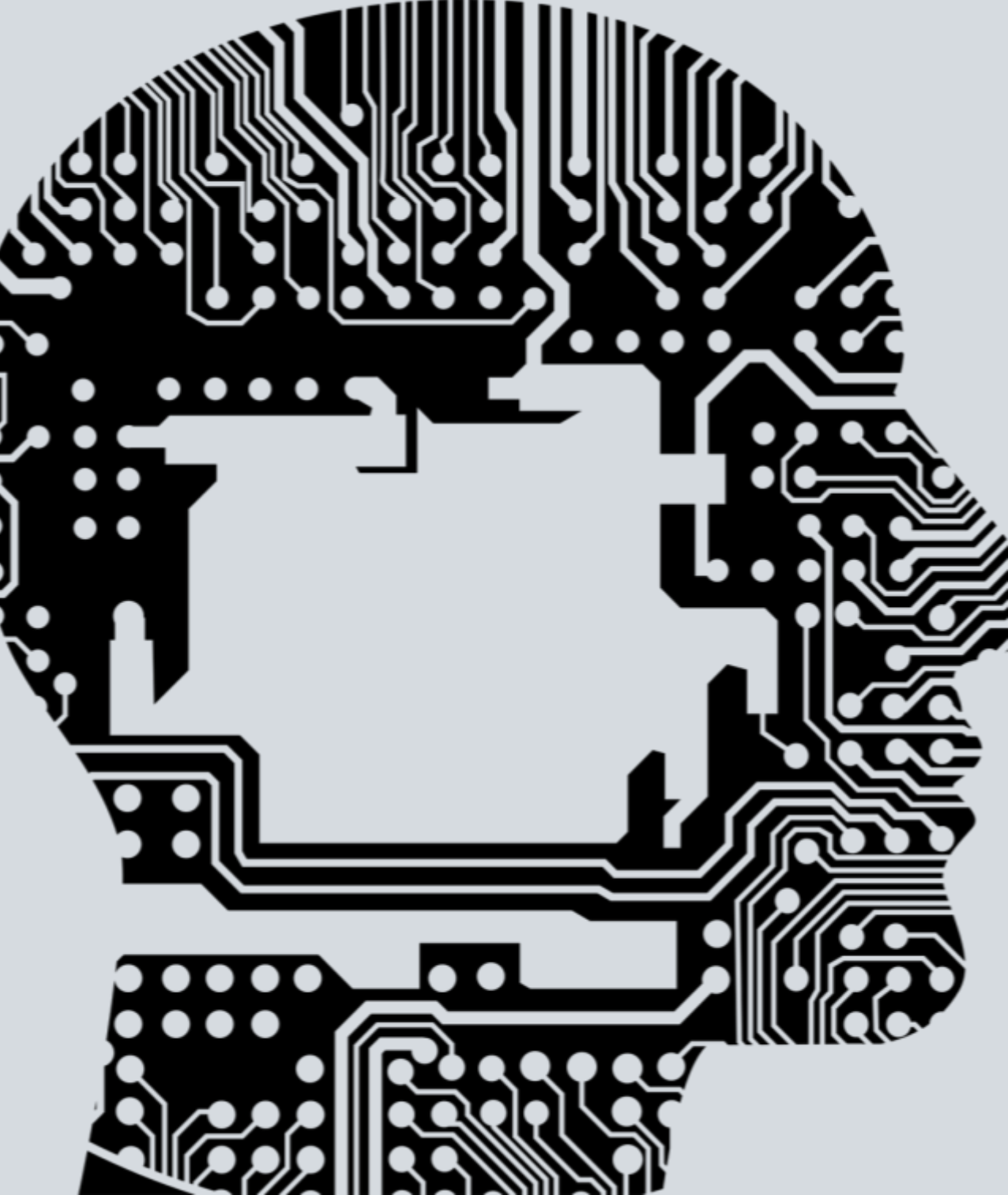
**1s to 3s** the probability of bounce **increases 32%**

**1s to 5s** the probability of bounce **increases 90%**

**1s to 6s** the probability of bounce **increases 106%**

**1s to 10s** the probability of bounce **increases 123%**





# WIE PASST SEO REIN?



**SNIPPET**

**TEXT, BILD, VIDEO**

**AUTHORITY**

**INHALT**

**SEO**

**TECHNIK**



**WEBSERVER /  
HOSTING**



**CLIENTS  
JAVASCRIPT**

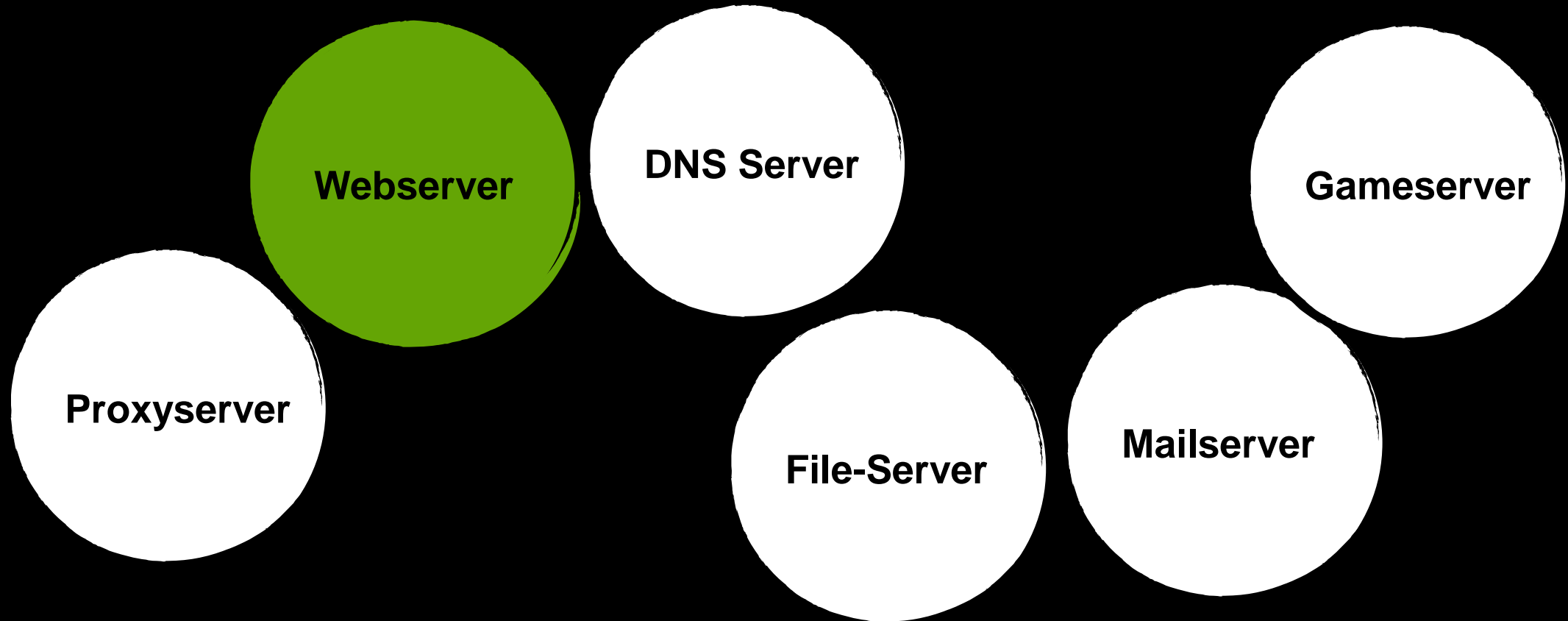


**SUCHMASCHINE**



**EIN SERVER IST EIN  
COMPUTER ODER  
PROGRAMM, DER DATEN  
ODER DIENSTE FÜR DIE  
ANDEREN PROGRAMME  
ODER COMPUTER  
BEREITSTELLT**

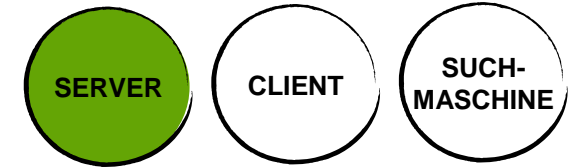
# VERSCHIEDENE SERVER-TYPEN











**EIN WEBSERVER**

**SPEICHERT WEBSEITEN,  
BEREITET DIESE AUF, UND  
LIEFERT SIE AN CLIENTS  
WIE WEBBROWSER ODER  
SUCHMASCHINEN-  
CRAWLER AUS**

# HOSTING



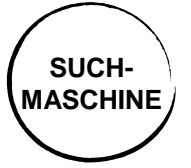
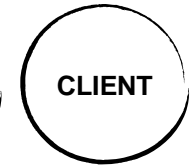
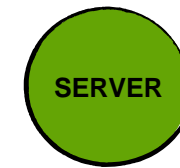
SHARED HOSTING	DEDICATED HOSTING	CLOUD	VIRTUAL PRIVATE SERVER
 	 	 	 





**BEREITTS BEIM HOSTING  
ENTSCHEIDET MAN  
ÜBER DIE LADEZEITEN  
DER SEITE**

# KLEINE WEBSEITE – KEIN PROBLEM

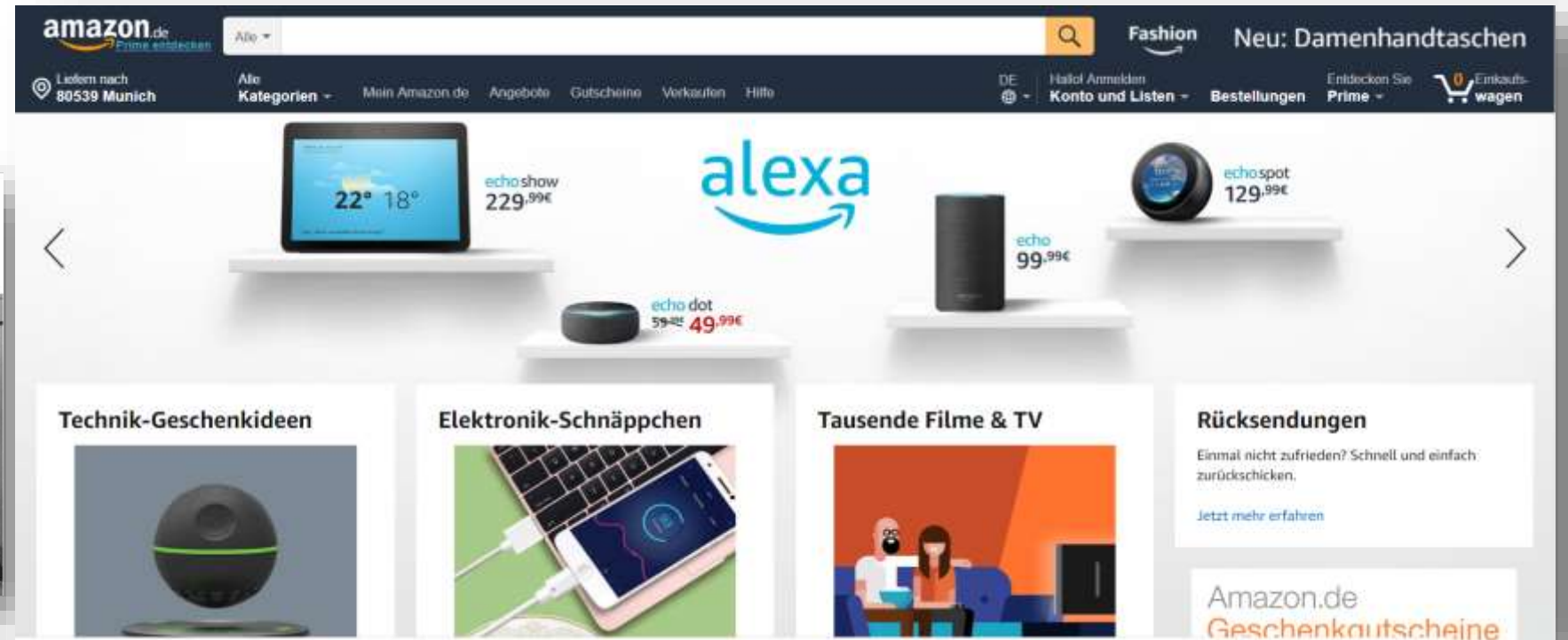
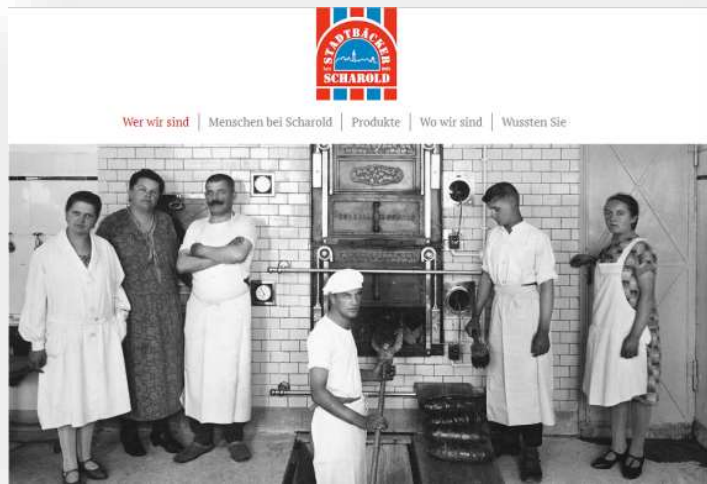
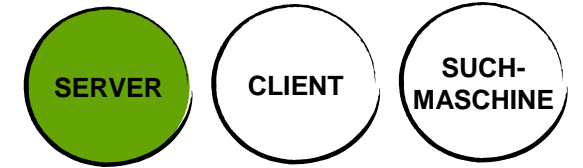


Wer wir sind | Menschen bei Scharold | Produkte | Wo wir sind | Wussten Sie



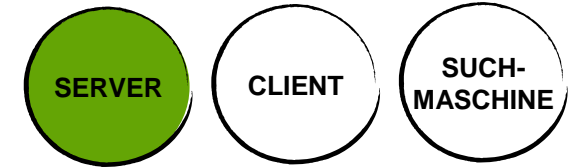
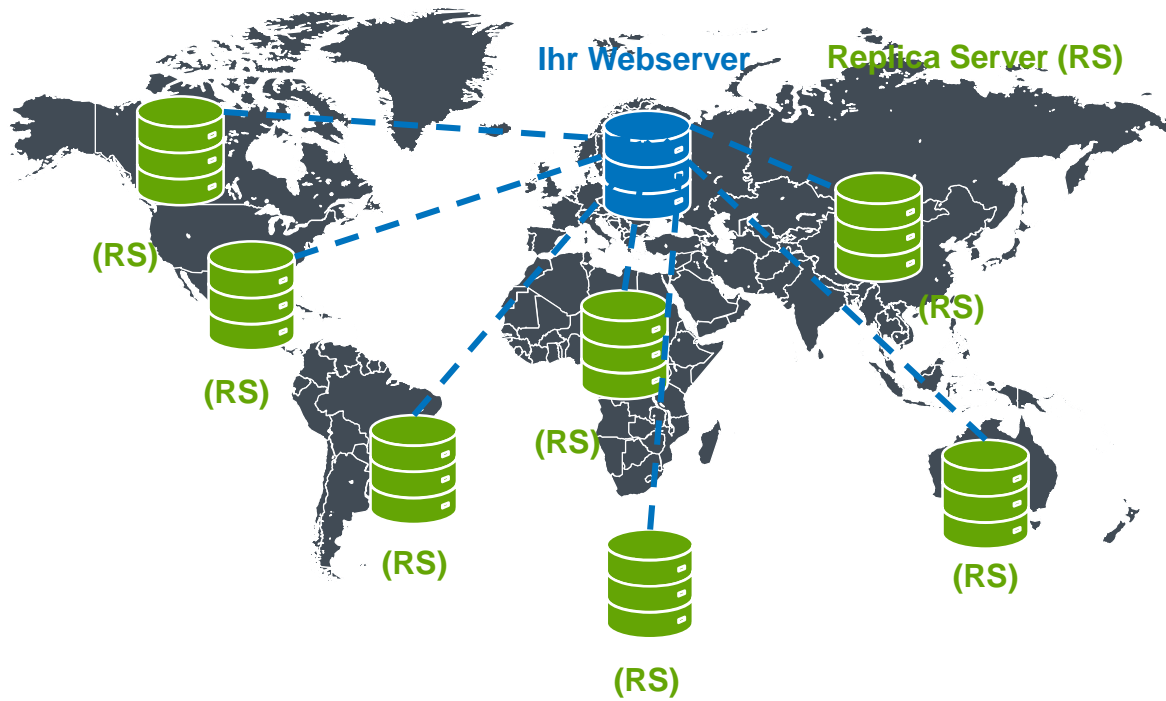
# KLEINE WEBSEITE – KEIN PROBLEM

# GROßE WEBSEITE – CDN



# **CDN – CONTENT DELIVERY NETWORK**

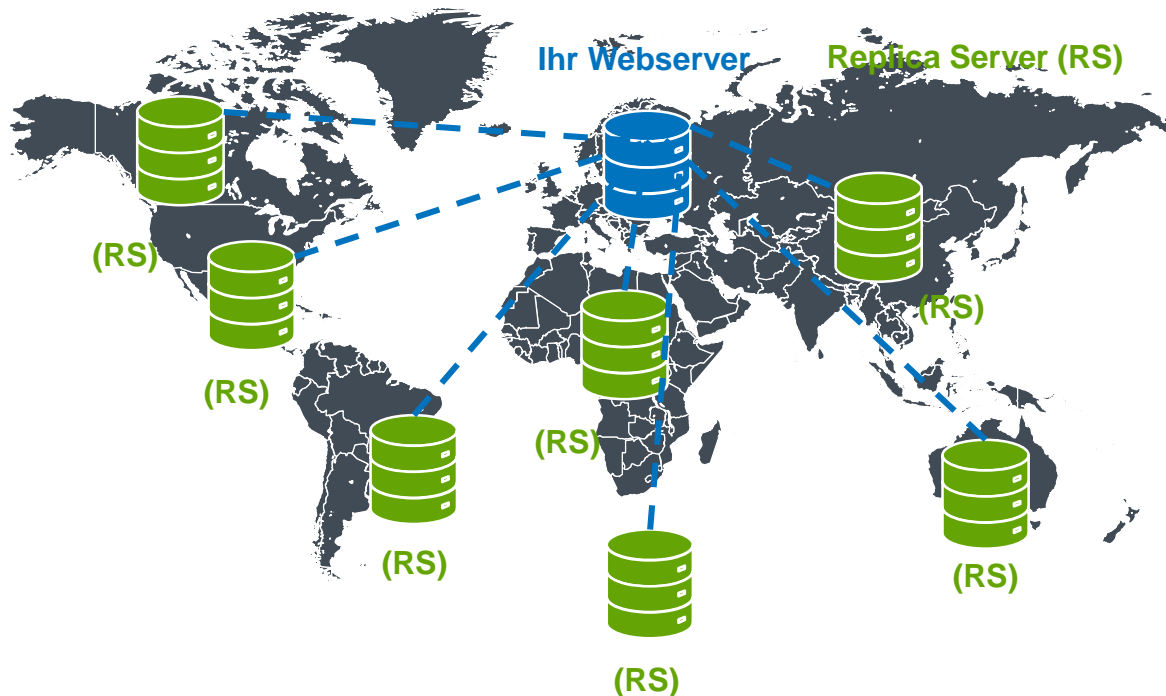
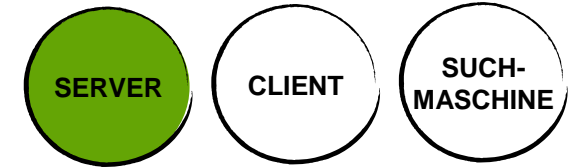
# CDN



- » CDN – Content Delivery Network
- » Statische Daten der Webserver werden auf die Replica Server übertragen (Verteilungssystem)
- » Nutzer ruft die Website auf, befindet sich dabei in Venezuela. Über Request Routing System wird der RS in Venezuela zu Hilfe gerufen. Er spielt die aktuellste Version der Website die er hat aus.



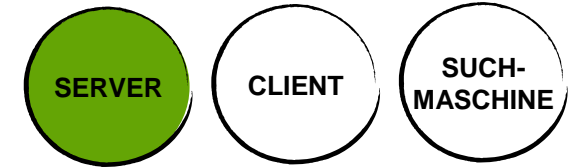
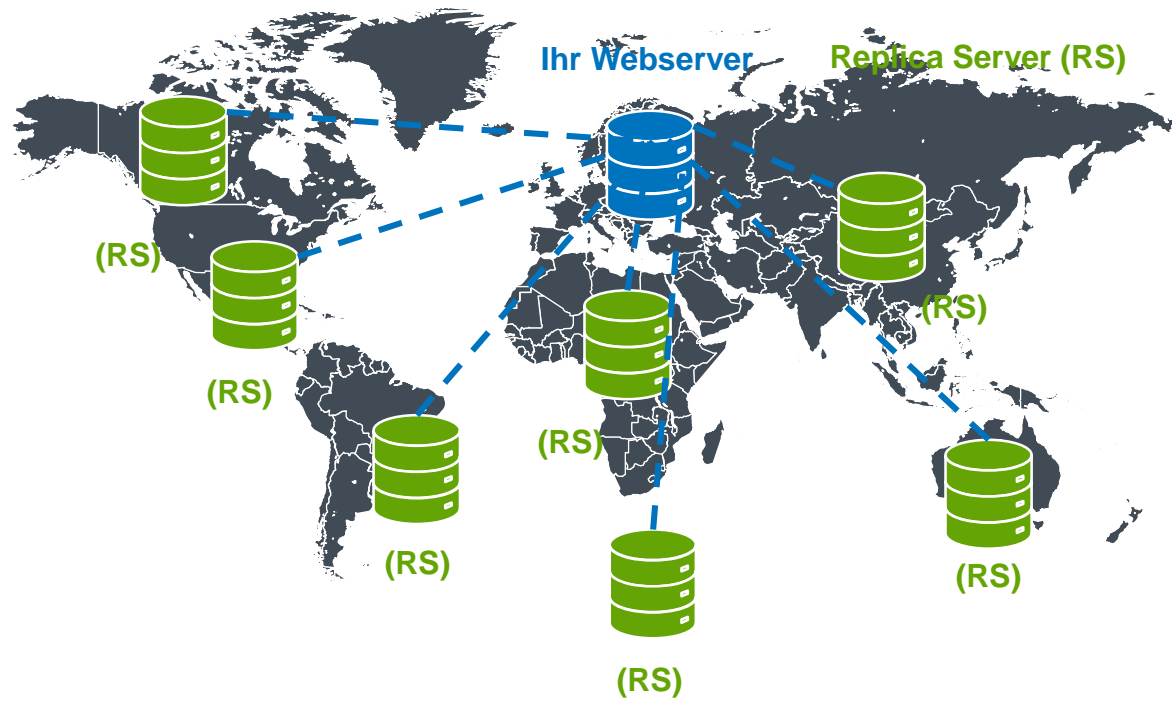
# CDN BESONDERHEITEN



- » Die CDN Anbieter unterscheiden sich nach **Server-Anzahl** und **Länderabdeckung**
- » Besonders **geeignet** sind **Videostreams**
- » Nicht CDN geeignet sind **personalisierte, öffentliche** und **häufig wechselnde** Informationen
- » **Logfile** Analyse ist **erschwert**
- » **China** ist ein Sonderfall:  
<https://www.cdnplanet.com/geo/china-cdn/>  
<https://www.kernpunkt.de/blog/digitales-marketing/webseiten-performance-in-china.html>



# WER NUTZT CDN



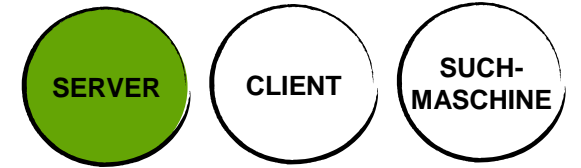
amazon

NETFLIX

facebook



# CDN ANBIETER WELTWEIT



Akamai

Amazon Web Services (CloudFront)

CDNetworks

Cloudflare

Google (Cloud CDN)

Host Europe (PlusServer)

Microsoft (Azure)





**DER CLIENT IST EIN  
PROGRAMM, DASS AUF  
DEM ENDGERÄT  
AUSGEFÜHRT WIRD UND  
MIT DEM SERVER  
KOMMUNIZIERT.**



**SERVER & CLIENT  
KOMMUNIZIEREN MIT-  
EINANDER, OBWOHL  
WIR DAS NICHT  
SEHEN**

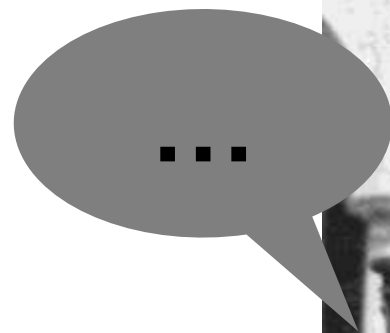






**DAS PASSIERT NACH  
DEM FRAGE-ANTWORT  
PRINZIP...**







**UND NACH  
PROTOKOLL...**

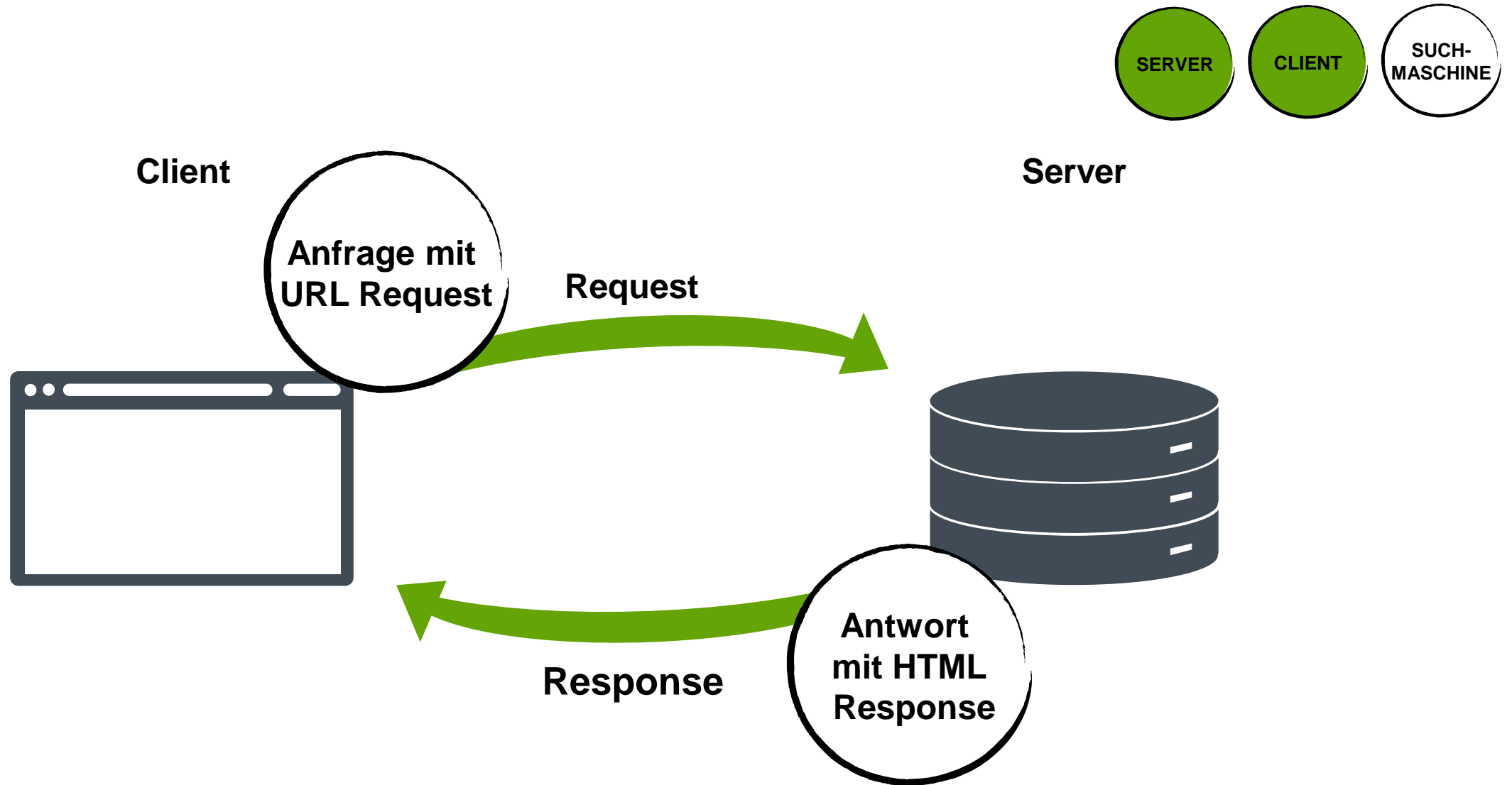




**...HTTP PROTOKOLL**



# CLIENT-SERVER-KOMMUNIKATION





# WEITERE NÜTZLICHE INFORMATIONEN



## Request

### ▼ Request Headers

```
:authority: www.campixx.de
:method: GET
:path: /seo-campixx-19/
:scheme: https
accept: text/html,application/xhtml+xml,application/xml;q=0.9,
accept-encoding: gzip, deflate, br
accept-language: de-DE,de;q=0.9,en-US;q=0.8,en;q=0.7
```

### Request-Zeilen

Methode

Request URI

HTTP-Version

## Response

### ▼ Response Headers

```
cache-control: max-age=172800
content-encoding: gzip
content-type: text/html; charset=UTF-8
date: Tue, 05 Mar 2019 21:51:49 GMT
expires: Thu, 07 Mar 2019 21:51:49 GMT
server: Apache/2.4.35
status: 200
vary: Accept-Encoding
x-powered-by: PHP/7.0.28
```

### Response-Zeilen

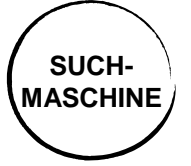
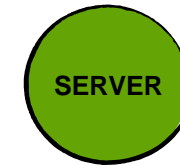
HTTP-Version

Status Code

Beschreibung



# WEITERE NÜTZLICHE INFORMATIONEN



## Request

## Response

### ▼ Request Headers

```
:authority: www.campixx.de
:method: GET
:path: /seo-campixx-19/
:scheme: https
accept: text/html,application/xhtml+xml,application/javascript;q=0.9,*/*;q=0.8
accept-encoding: gzip, deflate, br
accept-language: de-DE,de;q=0.9,en-US;q=0.8,en;q=0.7
```

### ▼ Response Headers

```
cache-control: max-age=172800
content-encoding: gzip
content-type: text/html; charset=UTF-8
date: Tue, 05 Mar 2019 21:51:49 GMT
expires: Thu, 07 Mar 2019 21:51:49 GMT
server: Apache/2.4.35
status: 200
vary: Accept-Encoding
x-powered-by: PHP/7.0.28
```

### ▼ General

```
Request URL: https://www.campixx.de/seo-campixx-19/
Request Method: GET
Status Code: 200
Remote Address: 134.119.101.129:443
Referrer Policy: origin
```

**Request-Zeilen**  
Methode  
Request URI  
HTTP-Version

**Response-Zeilen**  
HTTP-Version  
Status Code  
Beschreibung

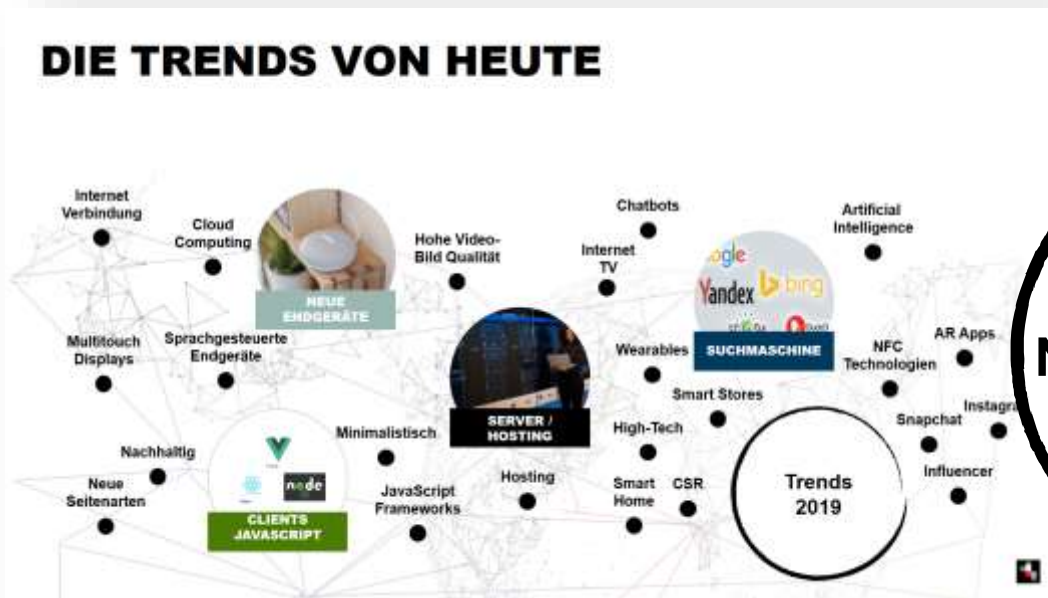


# FRONTEND



**FÜR EINE BESSERE UND  
SCHNELLERE  
KOMMUNIKATION**

# HEUTIGE TRENDS ERFORDERN NEUE CLIENTSEITIGE LÖSUNGEN



Neue Lösungen



# **ACCELERATED MOBILE PAGES (AMP)**

# AMP



- » Das AMP Projekt ist ein gemeinsames Projekt von Google und Twitter
- » Primär entwickelt für Verlage
- » Veröffentlicht von Google im Oktober 2015
- » Apache Lizenz – Open Source Code bei GitHub



# AMP BESONDERHEITEN



## Nackter HTML-Code

```
1 <img>
2 <audio>
3 <video>
4 <form>
5 <input>
6 <video>
7 <iframe>
8 <embed>
9
10 Javascript
11 Externen CSS-Dateien
12 ADS werden entfernt
```

```
1 <amp-img>
2 <amp-audio>
3 <amp-video>
4 <amp-anim>
5 <amp-iframe>
6 <amp-instagram>
7 <amp-carousel>
8 <amp-analytics>
9
10 Inline CSS
```

- » Langsame HTML Tags werden durch neue HTML Tags ersetzt
- » Websites werden gecached und bei Google gehostet (gstatic.com)
- » Das Canonical Tag, das auf die ursprüngliche Webseite verweist, ist daher sehr wichtig





# WER NUTZT AMP



Google

The Washington Post

LinkedIn

  
WORDPRESS

  
twitter

The  
Guardian



# AMP TESTEN (I)



https://amp.theguardian.com/membership/2016/feb/24/todays-release-of-accelerated-mobile-pages-amp#development=1

Support us →

# The Guardian

News Opinion Sport Culture Lifestyle

## Inside the Guardian blog

### The Guardian - now available as Accelerated Mobile Pages (AMP)

**Anthony Sullivan**  
Wed 24 Feb 2016 15:05 GMT

Today you'll notice a difference in how you see Guardian content on your mobile when you come to us from Google.

**Accelerated Mobile Pages (AMP)** is a Google-led initiative which aims to significantly speed up the mobile web. We'll make our regular content available in parallel in the AMP format, which can be served very quickly in specific areas of Google products and other non-Google apps.

### Browser Developer Console

The AMP Validator comes bundled with the AMP JS library, so it is available on every AMP page out of the box. To validate:

1. Open your AMP page in your browser.
2. Append `#development=1` to the URL, for example, `http://localhost:8000/released.amp.html#development=1`.
3. Open the **Chrome DevTools console** and check for validation errors.

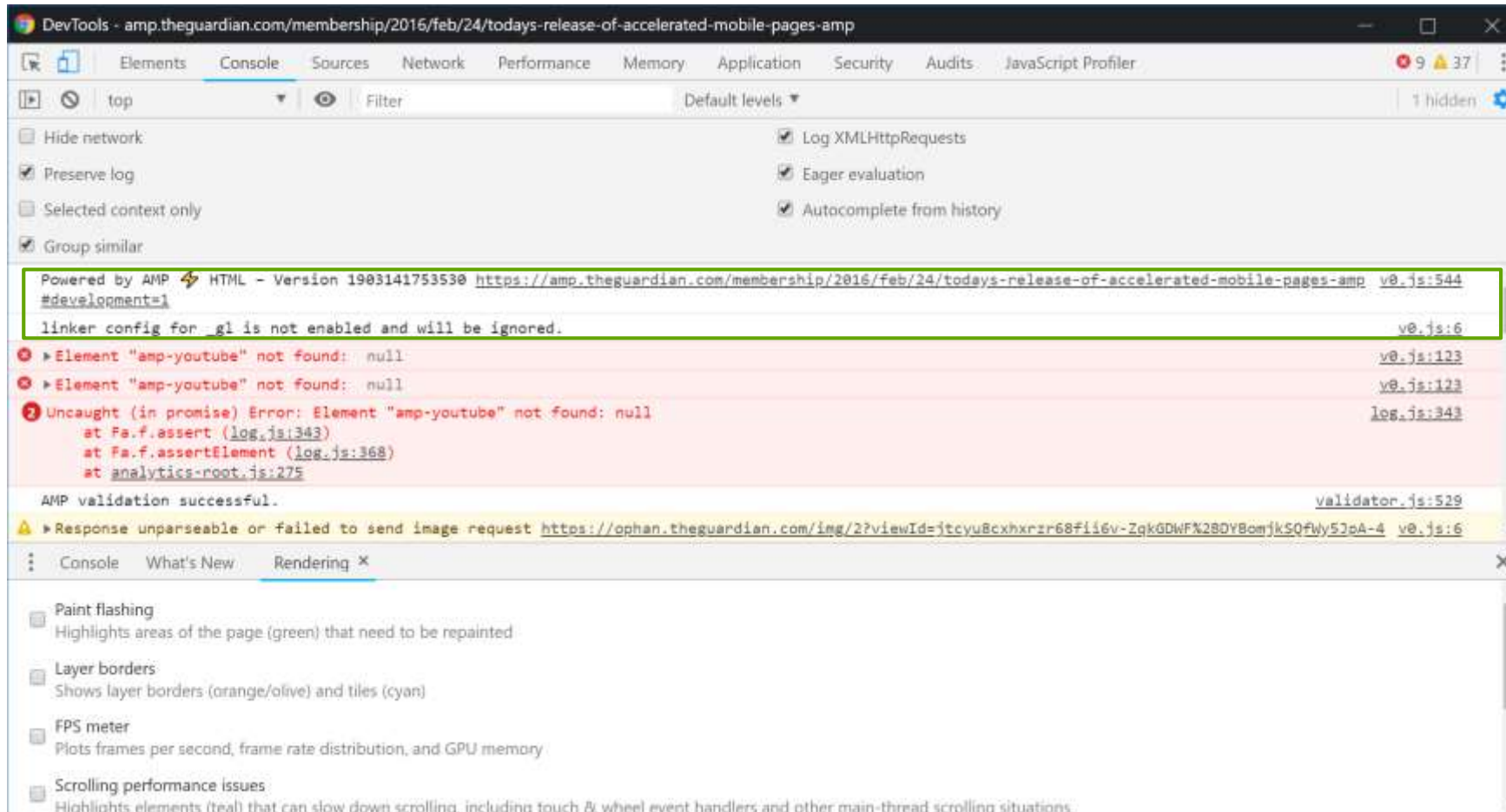
The screenshot shows the Chrome DevTools console with the following content:

- Powered by AMP HTML - Version 1903141753530 https://amp.theguardian.com/membership/2016/feb/24/todays-release-of-accelerated-mobile-pages-amp v0.js:544
- #development=1
- linker config for \_gl is not enabled and will be ignored. v0.js:16
- Element "amp-youtube" not found: null v0.js:123
- Element "amp-youtube" not found: null v0.js:123
- Uncaught (in promise) Error: Element "amp-youtube" not found: null log.js:343
- AMP validation successful. validator.js:529

Below the console, the 'Rendering' tab is active, showing options for Paint flashing, Layer borders, FPS meter, and Scrolling performance issues.



# AMP MIT CHROME DEVELOPER TOOLS TESTEN



The screenshot shows the Chrome DevTools interface for an AMP page. The top panel is the Console, which displays several log messages:

- A green message: "Powered by AMP HTML - Version 1903141753530 <https://amp.theguardian.com/membership/2016/feb/24/todays-release-of-accelerated-mobile-pages-amp> v0.js:544 #development=1 linker config for \_gl is not enabled and will be ignored. v0.js:6"
- Two red messages: "Element 'amp-youtube' not found: null" (v0.js:123)
- A red message: "Uncaught (in promise) Error: Element 'amp-youtube' not found: null" (log.js:343) with a stack trace:
  - at Fa.f.assert (log.js:343)
  - at Fa.f.assertElement (log.js:368)
  - at analytics-root.js:275
- A green message: "AMP validation successful." (validator.js:529)
- A yellow warning message: "Response unparseable or failed to send image request <https://ophan.theguardian.com/img/2?viewId=itcyu8cxhxrzr68fi6v-ZqkGDWFX2BDY8omjK5QfWY57pA-4> v0.js:6"

The bottom panel is the Rendering tab, which shows various performance and layout options:

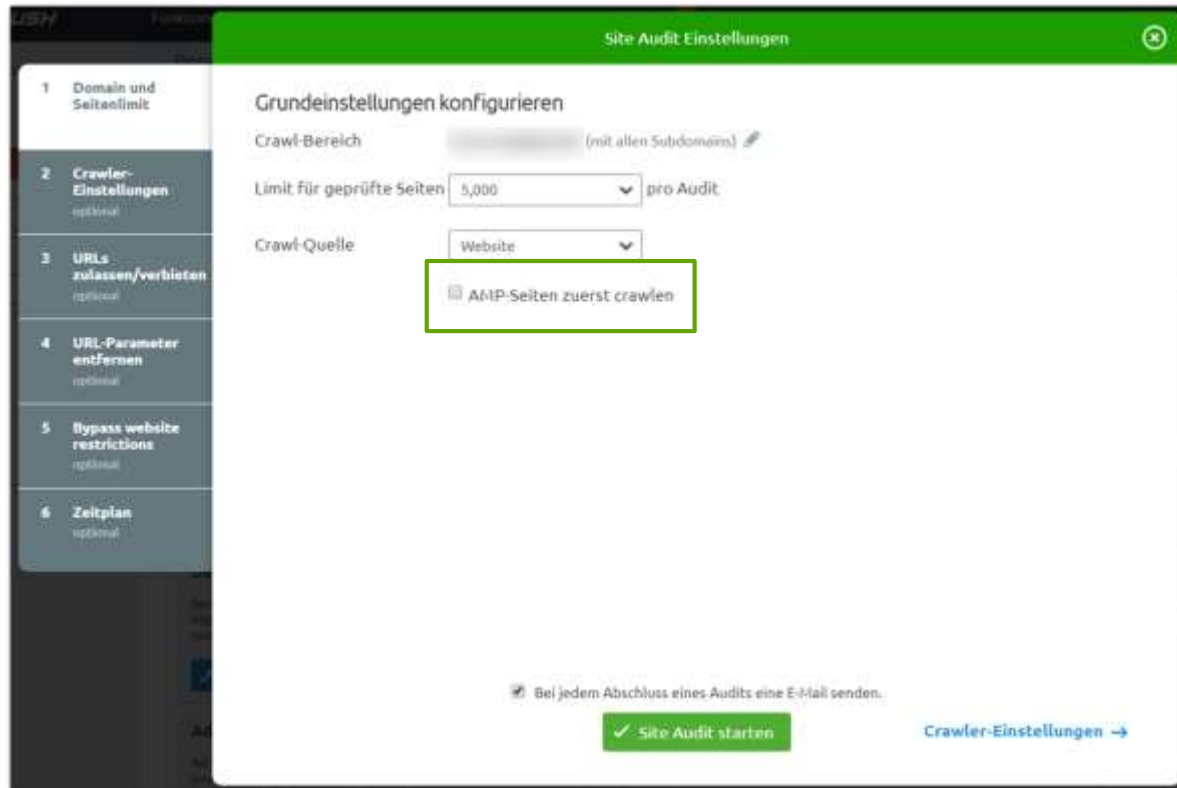
- Paint flashing: Highlights areas of the page (green) that need to be repainted.
- Layer borders: Shows layer borders (orange/olive) and tiles (cyan).
- FPS meter: Plots frames per second, frame rate distribution, and GPU memory.
- Scrolling performance issues: Highlights elements (teal) that can slow down scrolling, including touch & wheel event handlers and other main-thread scrolling situations.



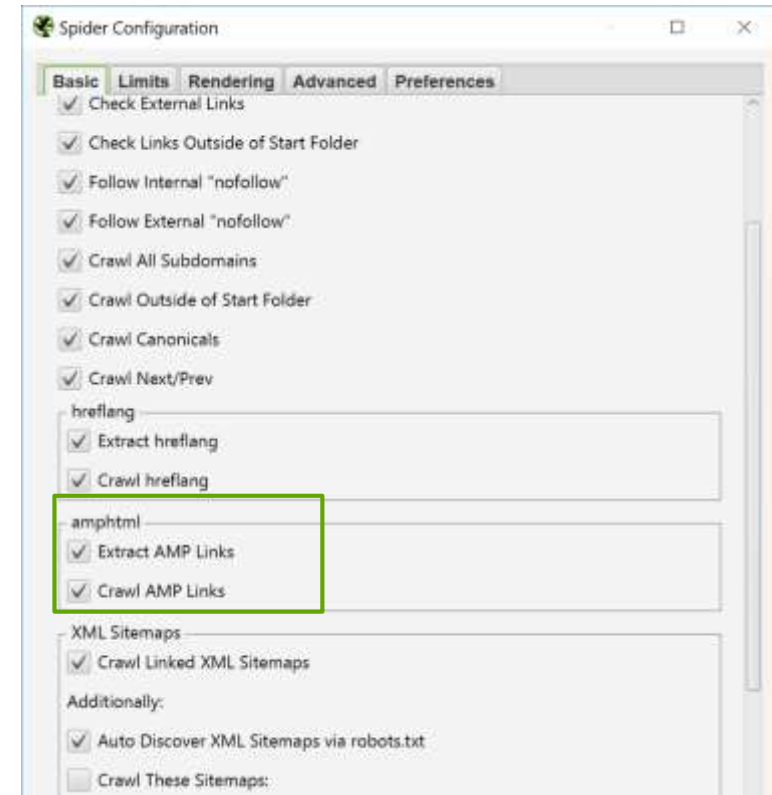
# AMP TESTEN (II)



## SEMrush Site Audit



## Screaming Frog



# AMP PROS UND KONTRAS



## Vorteile

- » AMP sind vor allem für News, Verlage und Content Websites vorteilhaft
- » Wesentlich schnellere Ladezeiten durch redundanten Code, Lazy Loading und Rendering

*The Washington Post*

**23% mehr mobile**

**88% bessere Ladezeiten**

**1000+ AMP Artikel-Seiten täglich**

## Nachteile

- » Schlecht für eigene Ads, da die AMP Seiten auf dem Google Server gecached und gehostet werden
- » Weniger Kontrolle über Style und Inhalte
- » Limitierte Anzahl an HTML Tags
- » Zweite Version einer Website:
  - » Google Analytics – aufwändige Einstellung
  - » Aufwand doppelte Programmierung

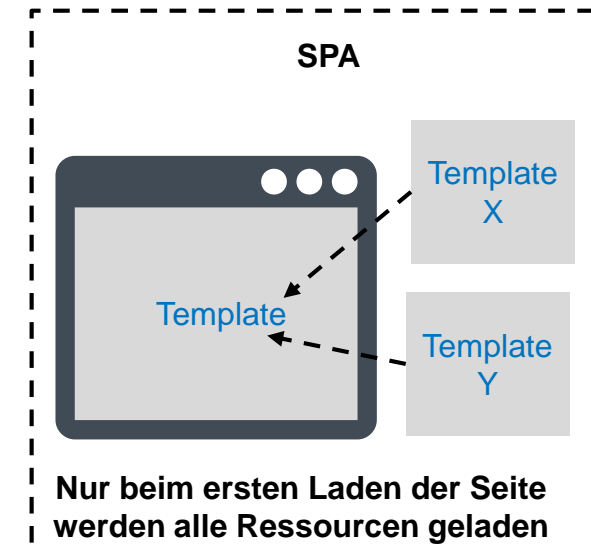


# **SINGLE PAGE APPLICATION (SPA)**

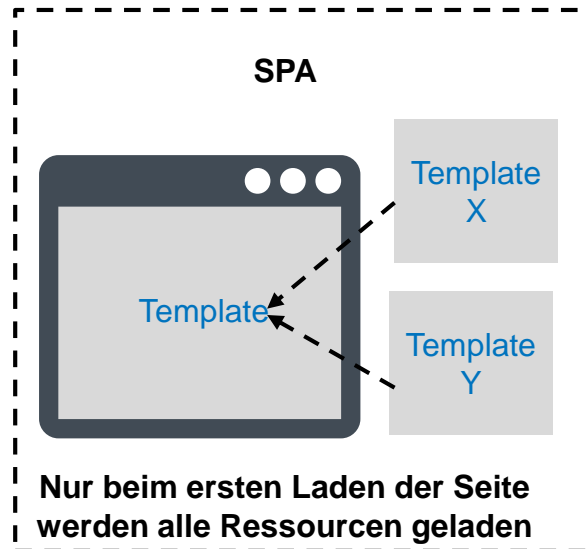
# SPA



» Single Page Application ist eine Webanwendung, die aus einem HTML-Dokument besteht, deren Inhalte dynamisch nachgeladen werden.



# SPA BESONDERHEITEN



- » Es werden **clientseitige JavaScript** Frameworks eingesetzt (AngularJS, Ember.js, ExtJS, Knockout.js, Meteor.js, ReactJS, Vue.js)
- » **React** verbreitet sich dank Facebook. Es bietet eine Syntax-Lösung (JSX) an, die suchmaschinen-freundlichen Inhalt bereitstellt.





# WER NUTZT SPA



# SPA PROS UND KONTRAS



## Vorteile

- » Mehrere Ressourcen müssen nur einmal geladen werden: (HTML, CSS, Scripts)
- » Sie können mit Chrome debugged werden
- » Sie können lokales Cache effektiv nutzen

## Nachteile

- » JavaScript muss vorhanden und aktiviert sein
- » Unvollständiges Rendering ist möglich
- » SPA-Seiten können längere Ladezeiten verursachen, umso grösser die Websites sind (Serverside Rendering sollte ermöglicht werden)



# **PROGRESSIVE WEB APPLICATION (PWA)**

# PWA



	Native App	Responsive Website	Progressive Web App
Functions Offline	✓	✗	✓
Push Notifications	✓	✗	✓
Installable on home screen	✓	✗	✓
Full screen experience	✓	✗	✓
Indexable by search engines	✗	✓	✓
One place to enter content	✗	✓	✓
Works across all devices	✗	✓	✓
No download required	✗	✓	✓
Doesn't require updates	✗	✓	✓

- » Ein Nutzer muss mindestens 30 Sekunden mit der Domain interagieren
- » Weitere Infos:  
<https://developers.google.com/web/fundamentals/web-app-manifest/>

- » Eine Progressive Web Application ist eine Website, die sich wie eine App verhält.
- » Vier technologische Voraussetzungen:
  - (1) manifest.json File
  - (2) Service Workers
  - (3) HTTPs
  - (4) Application Shell Architecture



	Native App	Responsive Website	Progressive Web App
Functions Offline	✓	✗	✓
Push Notifications	✓	✗	✓
Installable on home screen	✓	✗	✓
Full screen experience	✓	✗	✓
Indexable by search engines	✗	✓	✓
One place to enter content	✗	✓	✓
Works across all devices	✗	✓	✓
No download required	✗	✓	✓
Doesn't require updates	✗	✓	✓

# WER NUTZT PWA



The Washington Post



tinder



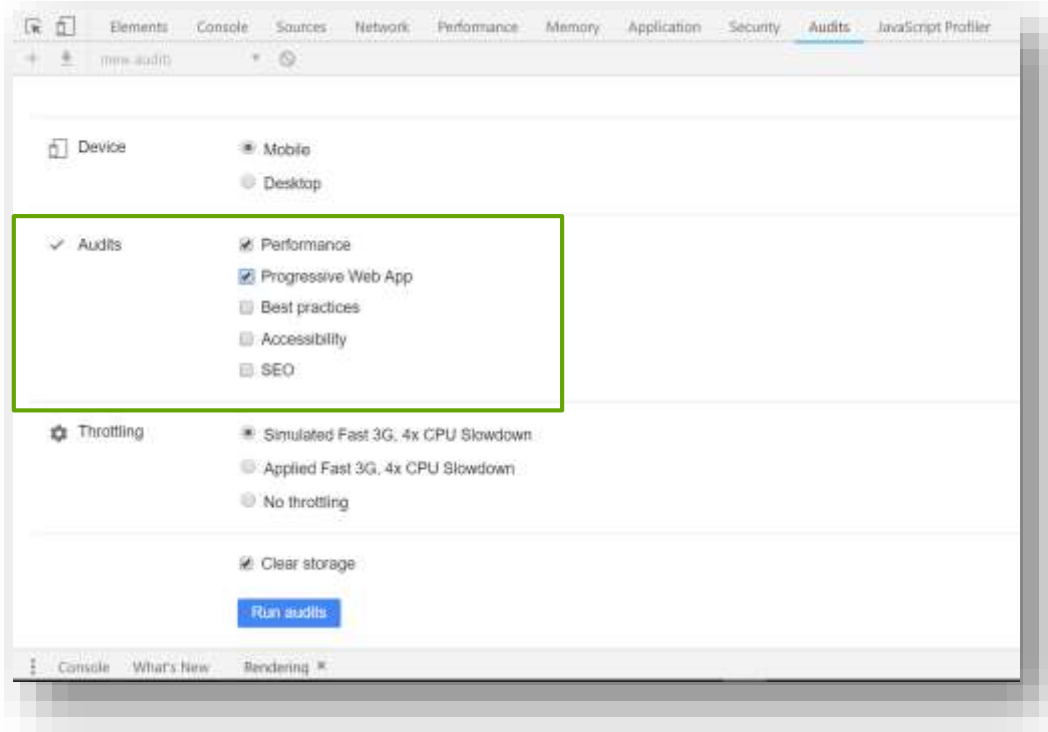
Forbes



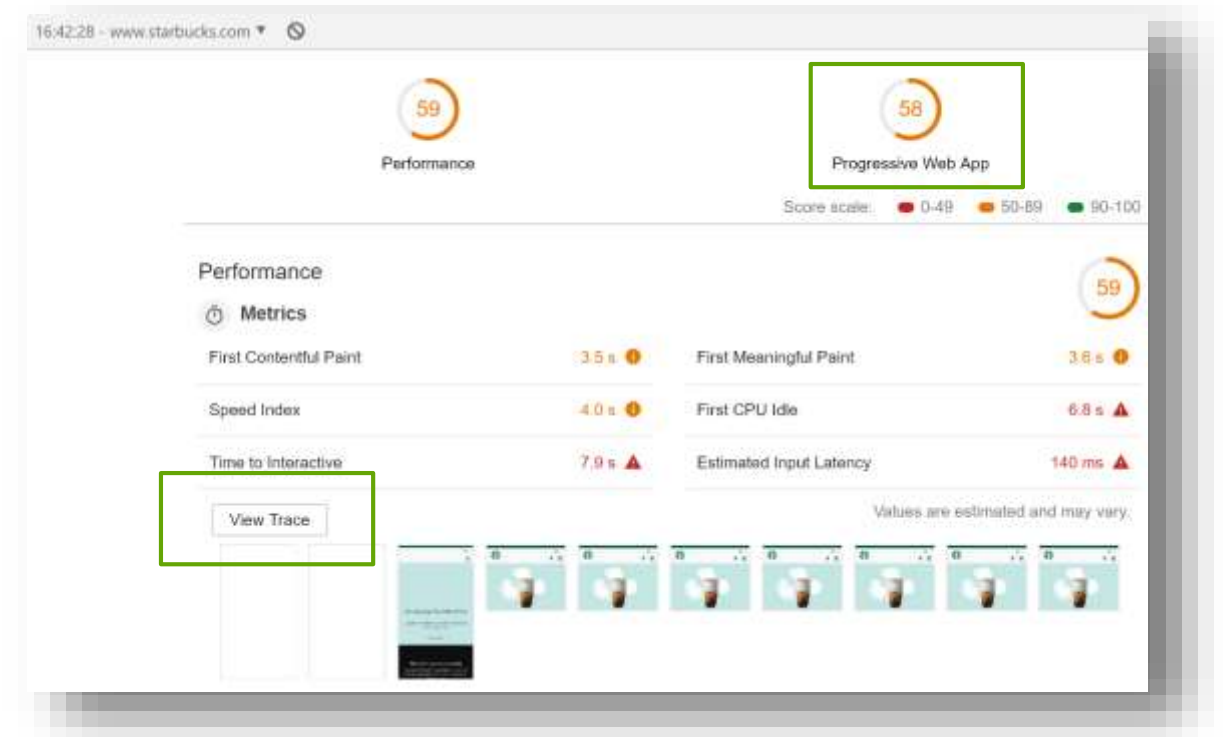
# PWA TESTEN



## Google Developer Tools



## PWA Testen mit Google Lighthouse



**DIE SUCHMASCHINE IST  
EIN PROGRAMM,  
WELCHES AUTOMATISCH  
WEBSEITEN SAMMELN  
UND AUSWERTEN KANN**





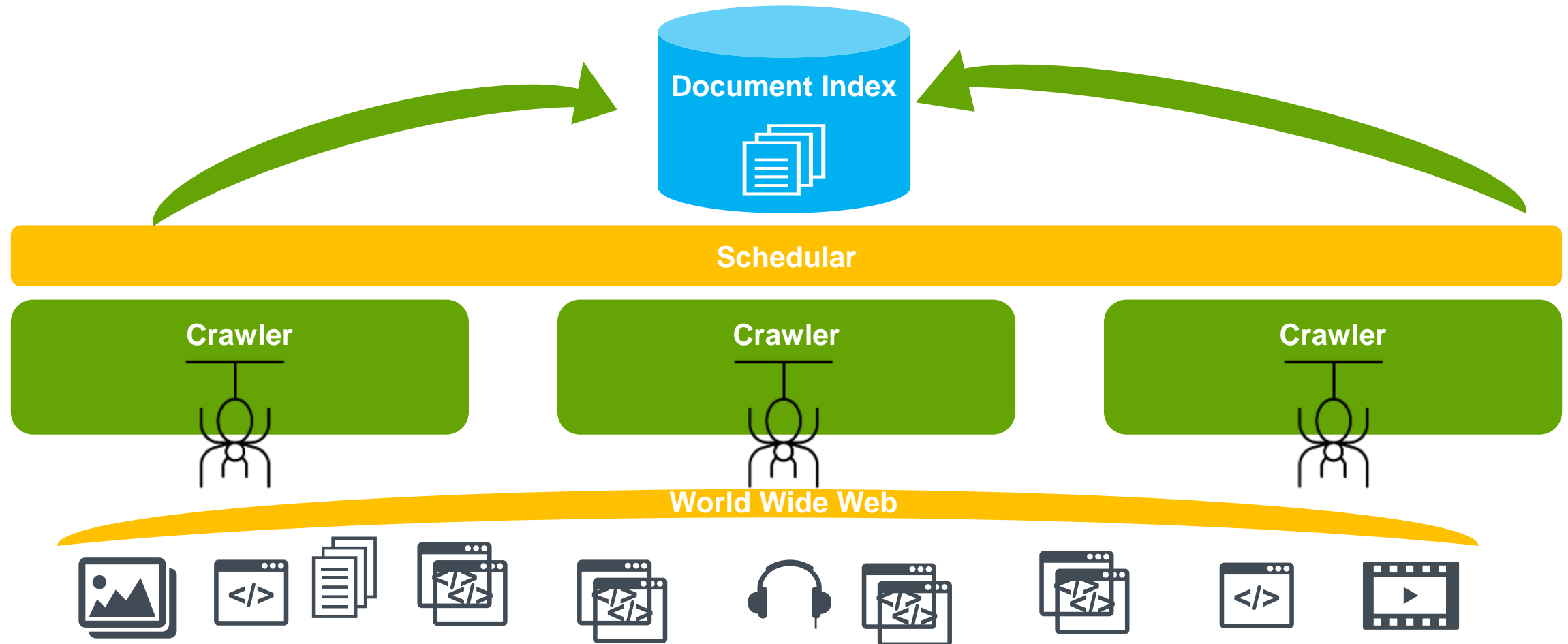
CAUTION

ROT

**GOOGLE IST WIE EIN  
KOPIERER, DER SEITEN  
AUS WEBSERVERN  
KOPIERT**

**CRAWLING**

# CRAWLING



# INDEXIERUNG

# INFORMATION RETRIEVAL SYSTEM



# BLIND FIVE YEAR OLD

HOME ABOUT CLIENTS CONTACT US DESKTOP SEARCH VOLUME TREND SUBS

## CRAWL OPTIMIZATION

// AJ KOHN // JULY 29TH 2013 // SEO

Crawl optimization should be a priority for any large site looking to improve their SEO efforts. By tracking, monitoring and focusing Googlebot you can gain an advantage over your competition.

## Crawl Budget



# CRAWL BUDGET





Erforschen Sie mehr als zehn Jahre Innovation in unseren Rechenzentren.



2012

## Wir zeigen, wer wir sind. Innen wie außen.

Erstmals gestatten wir jedem, einen Blick auf die Anlagen zu werfen, auf denen unsere Produkte ausgeführt werden. Sie können nun in [Street View](#) an einer virtuellen Tour durch eines unserer Rechenzentren teilnehmen. Oder Sie stöbern in unseren [Fotoalben](#) mit Bildern unserer Systeme, unserer Mitarbeiter und der Standorte, an denen Google-Produkte ausgeführt werden. [Learn more.](#)

2009

## Zusammenarbeit in der Branche

2010

## Erneuerbare Energien

2011

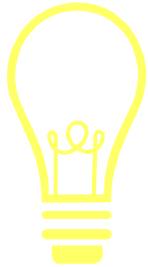
## Externe Zertifizierung

## 2012 Transparenz



# LESEN UND AUSWERTEN KOSTET GOOGLE BARES GELD

» Jeder Google-Nutzer könnte mit seinen monatlichen Suchanfragen eine **60-Watt-Glühbirne für drei Stunden** mit Strom versorgen.



» Oder 5,7 Terawattstunden in 2015 = der jährliche Energiekonsum von San Francisco



# CRAWL BUDGET

Crawler



Website



**Crawl budget** is the number of pages Google will **crawl** on your site on any given day. This number varies slightly from day to day, but overall it's relatively stable. ... At that point, it becomes time to optimize your **crawl budget**. 05.07.2016

Crawl budget optimization • Yoast  
<https://yoast.com/crawl-budget-optimization/>



# CRAWL BUDGET

Crawler



Website



**Crawl budget** is the number of pages Google will **crawl** on your site on any given day. This number varies slightly from day to day, but overall it's relatively stable. ... At that point, it becomes time to optimize your **crawl budget**. 05.07.2016

Crawl budget optimization • Yoast  
<https://yoast.com/crawl-budget-optimization/>



*Dieses Budget kann aufgebraucht werden.*



**CRAWLING UND  
INDEXIERUNG KANN  
EINZELN GESTEUERT  
WERDEN**

# ROBOTS.TXT

```
← → ↻ 🔒 https://www.google.de/robots.txt
User-agent: *
Disallow: /search
Allow: /search/about
Allow: /search/static
Allow: /search/howsearchworks
Disallow: /sdch
Disallow: /groups
Disallow: /index.html?
Disallow: /?
Allow: /?hl=
Disallow: /?hl=*
Allow: /?hl=*&gws_rd=ssl$
Disallow: /?hl=*&*&gws_rd=ssl
Allow: /?gws_rd=ssl$
Allow: /?ptl=true$
Disallow: /imgres
Disallow: /u/
Disallow: /preferences
Disallow: /setprefs
Disallow: /default
Disallow: /m?
Disallow: /m/
Allow: /m/finance
Disallow: /wml?
Disallow: /wml/?
Disallow: /wml/search?
Disallow: /xhtml?
Disallow: /xhtml/?
Disallow: /xhtml/search?
```



An die Angaben in der robots.txt Datei halten sich Google, Bing und Baidu.

- » Die robots.txt Datei wird von der Suchmaschine öfter „besucht“
- » Die Steuerung erlaubt oder verbietet bestimmte Seiten zu crawlen
- » Die robots.txt verhindert nicht die Indexierung
- » **Hauptaufgabe: Crawling Steuerung**
- » **Testen: GSC-> Crawling-> robots.txt Tester**



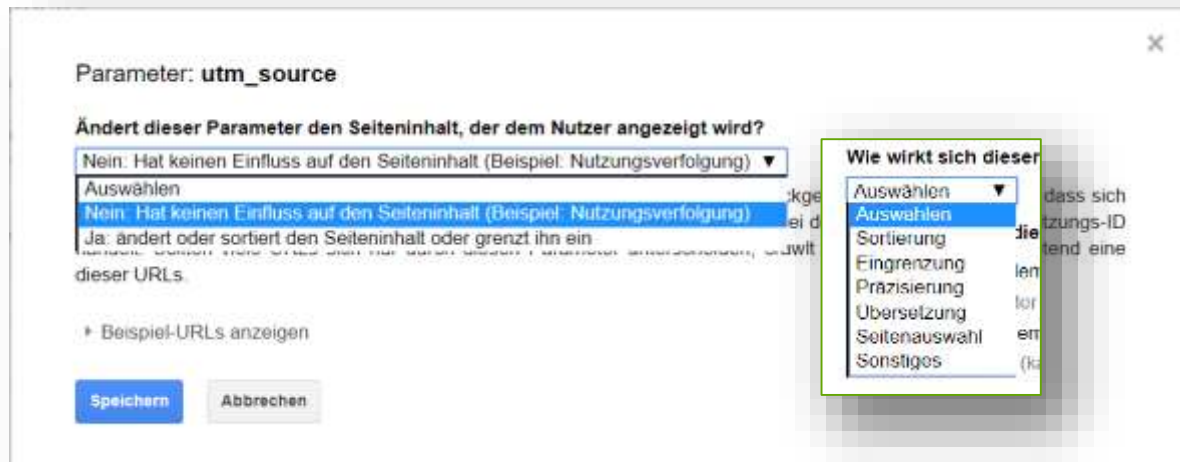
# XML-SITEMAP

```
<?xml version="1.0" encoding="UTF-8" ?>
<sitemapindex xmlns="http://www.google.com/schemas/sitemap/0.84">
  <sitemap>
    <loc>https://www.google.com/gmail/sitemap.xml</loc>
  </sitemap>
  <sitemap>
    <loc>https://www.google.com/forms/sitemaps.xml</loc>
  </sitemap>
  <sitemap>
    <loc>https://www.google.com/slides/sitemaps.xml</loc>
  </sitemap>
  <sitemap>
    <loc>https://www.google.com/sheets/sitemaps.xml</loc>
  </sitemap>
  <sitemap>
    <loc>https://www.google.com/drive/sitemap.xml</loc>
  </sitemap>
  <sitemap>
    <loc>https://www.google.com/docs/sitemaps.xml</loc>
  </sitemap>
  <sitemap>
    <loc>https://www.google.com/get/sitemap.xml</loc>
  </sitemap>
</sitemapindex>
```

- » Die XML-Sitemap Datei ermöglicht es, eine **indexrelevante Liste mit URLs** direkt an die Suchmaschine **zu kommunizieren**
- » Hier können **hreflang Tags** direkt vergeben werden
- » Größere Websites sollten mehrere, nach Themen erstellte XML-Sitemaps in einer **Index XML-Sitemap bereitstellen**
- » XML-Sitemaps können für **Medien** (Bilder, Videos) erstellt werden
- » **Hauptaufgabe: Indexierung Steuerung**
- » **Testen: GSC-> Crawling-> Sitemap testen**



# URL-PARAMETER TOOL IN DER GOOGLE SEARCH CONSOLE



- » Ein sehr wirksames Tool, das bei einer falschen Einstellung die gültige Steuerung (Canonical Tag, Meta Tag Robots) verhindern kann
- » Einzelne Parameter können hinzugefügt werden
- » Nicht konfigurierte Einstellungen haben kein Datum

Konfiguriert	Effekt	Crawling	
05.03.2019	Keine	Stellvertretende URL	<a href="#">Bearbeiten / Löschen</a>

- » **Hauptaufgabe: Crawling und Indexierung Steuerung**
- » **Testen: Logfile Analyse, Google Index**



# ANDERE URL-PARAMETER TOOLS

**Diese URL-Parameter ignorieren** ⓘ

Welche Parameter sollen von Bing ignoriert werden?

**Zu ignorierende Parameter**

Sie haben keine zu ignorierende Parameter.

**Diese URL-Parameter ignorieren**

Mit dieser Funktion können Sie Abfrageparameter festlegen, die der Bing Crawler ignorieren soll. Wenn Sie z. B. die Website `http://example.com/` besitzen und legen fest, dass für diese Seite der Parameter "abc" zu ignorieren ist, so wird die URL `http://example.com/home?abc=123` als `http://example.com/home` gesehen. D. h., dass bei URLs mit dem Abfrageparameterstring "abc" dieser Parameter vor der Indexierung entfernt wird. Vorteile sind: doppelten Inhalt im Bing Index vermeiden, der Indexwert einer Seite wird nicht über mehrere URL-Varianten verteilt, und keine unnötige Nutzung von Website-Bandbreite durch den Crawler.

- » Im Bing Webmaster Tools können einzelne Parameter oder eine Liste mit Parametern eingereicht werden
- » Im Bing Webmaster Tools werden ausschließlich die Parameter eingegeben, die ignoriert werden sollen
- » Die Baidu Suchmaschine hat kein URL Parameter Tool
- » **Hauptaufgabe: Crawling und Indexierung Steuerung**
- » **Testen: Logfile Analyse, Bing Index**

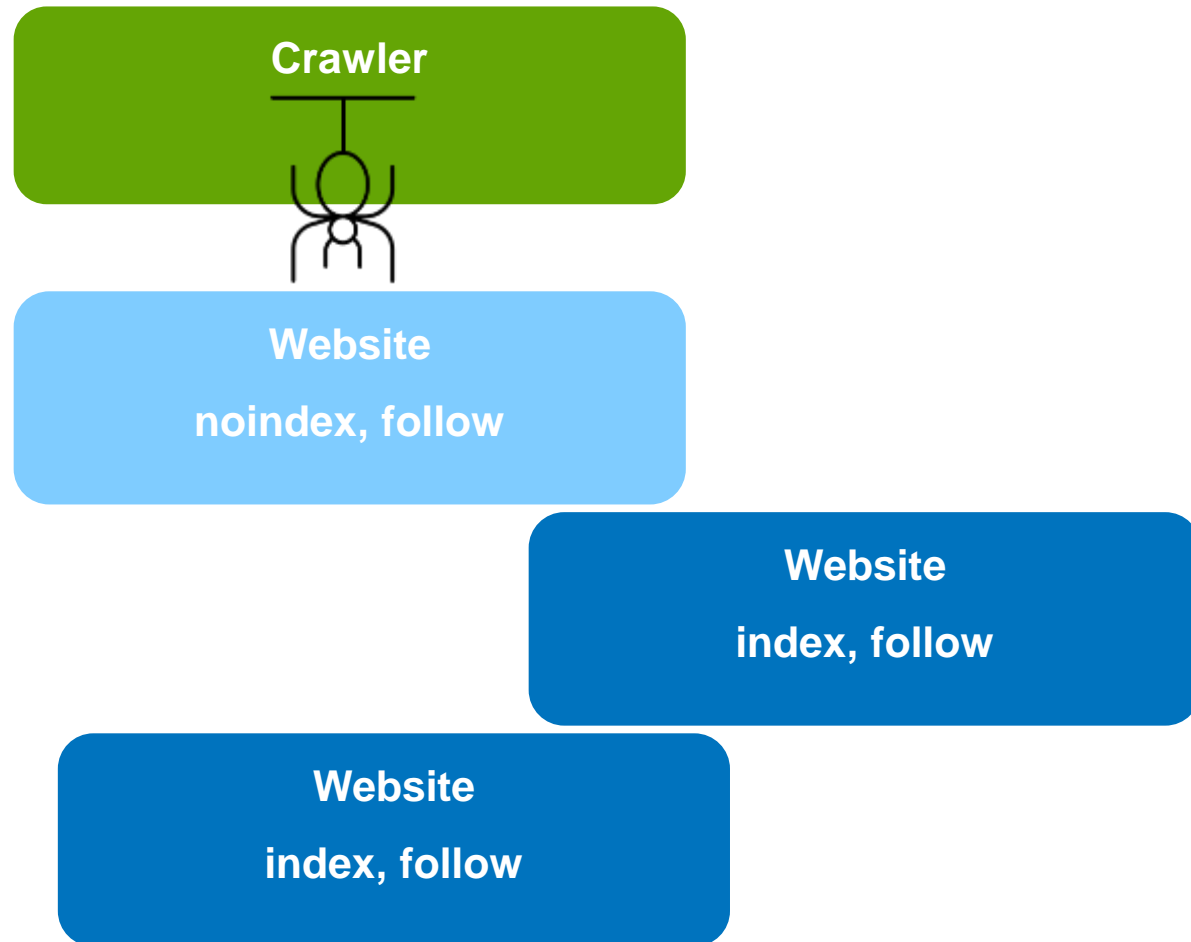
# POPULÄRSTE MAßNAHMEN UND DEREN EINSATZ

Maßnahme	Crawlen	Indexieren	Crawlen nicht Indexieren
Meta Tags Robots	V	V	V
Sitemap.xml	V	V	V
rel=prev rel=next	V	V	V
GSC Parameter Tool	V	V	V
hreflang tag	V	V	X
Canonical Tag	V	*	*
robots.txt	V	X	X

\* Ist von den anderen Signalen stark abhängig



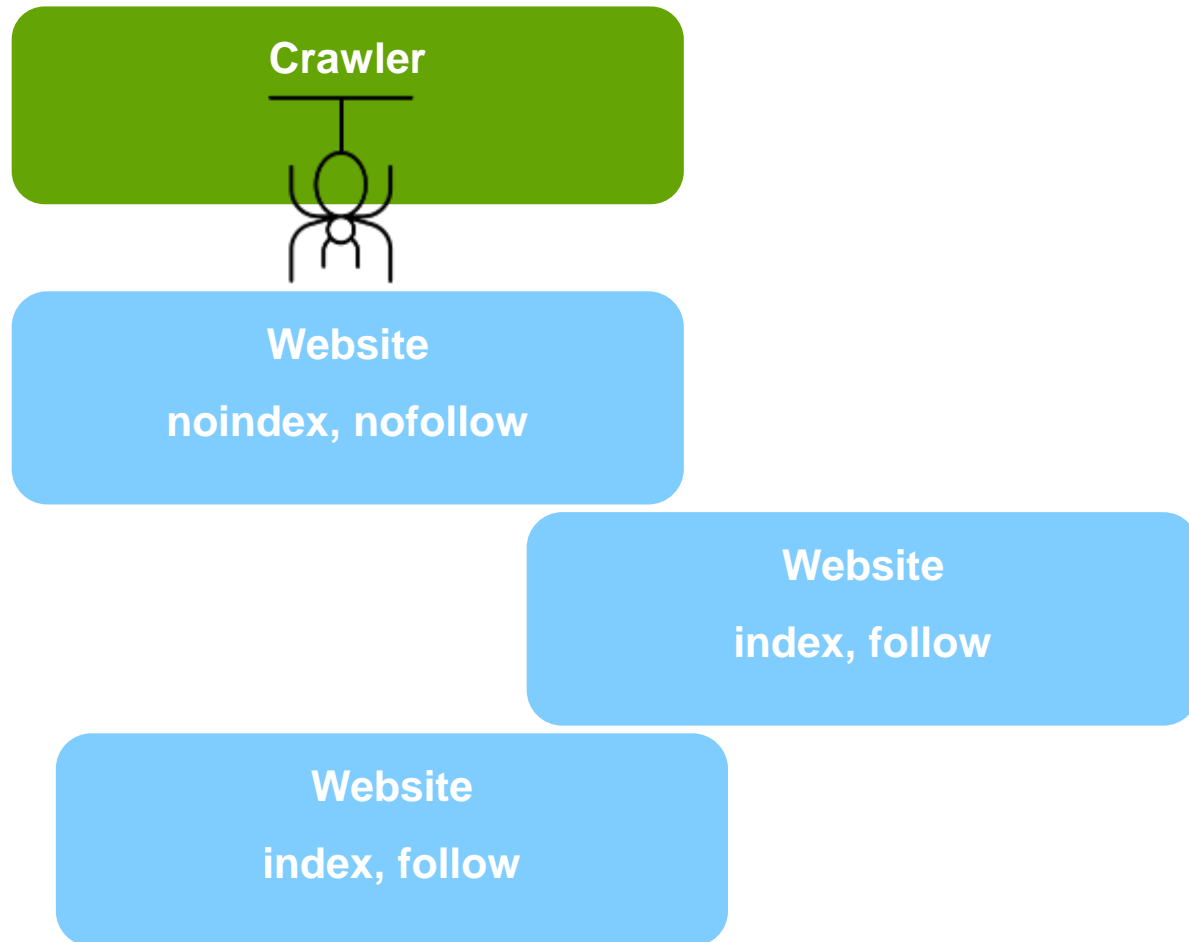
# NOINDEX WIRD ZU NOFOLLOW?



- » Je **länger** die Seiten mit **noindex, follow** versehen werden, desto weniger werden sie gecrawlt.
- » **Irgendwann** hört die Suchmaschine auf, sie und die ausgehenden URLs der Seite zu crawlen.
- » Die ausgehenden Links dieser Seiten werden zu **noindex, nofollow**.



# NOINDEX WIRD ZU NOFOLLOW?



- » Je **länger** die Seiten mit **noindex, follow** versehen werden, desto weniger werden sie gecrawlt.
- » **Irgendwann** hört die Suchmaschine auf, sie und die ausgehenden URLs der Seite zu crawlen.
- » Die ausgehenden Links dieser Seiten werden zu **noindex, nofollow**.

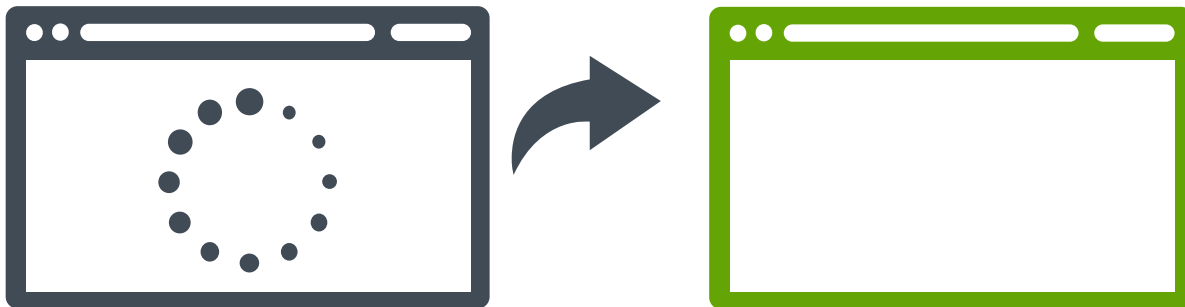


# **CRAWLING- & INDEXIERUNGS- PROBLEME**

# WEITERLEITUNGEN

## URL-Änderung

- » Technische Änderungen (.html, Trailing Slash)
- » Alte Produkt URLs werden gelöscht
- » Kategorien werden verschoben
- » Unterkategorien werden umbenannt, usw.



### Serverseitige

- 301 - permanent
- 302 - temporär
- 307 - bitte über https



### Clientseitige

Weiterleitung wie  
JavaScript und Meta  
Refresh sind nicht  
suchmaschinen-  
freundlich



# CRAWLING BUDGET UND WEITERLEITUNGEN



**Gary "鯨理" Illyes** ✓

@methode

Follow



30x redirects don't lose PageRank anymore.

8:59 AM - 26 Jul 2016 from Zurich, Switzerland

43 276 267



Tweet your reply



**Eric Enge** ✓ @stonetemple · 26 Jul 2016

Replying to @methode

not even a dampening factor?

1 1



**Gary "鯨理" Illyes** ✓ @methode · 26 Jul 2016

for PageRank, no.

1 4



**Ulf Heyden** @UlfHeyden · 7 Apr 2017

Replying to @methode

Is this still true?

1 1



**Gary "鯨理" Illyes** ✓ @methode · 7 Apr 2017

Yes

1 1

1 more reply



# CRAWLING BUDGET UND WEITERLEITUNGEN

The image shows a screenshot of a Twitter thread. The main tweet is from Gary "鯨理" Illyes (@methode) dated 26 Jul 2016, with the text "30x redirects don't lose Pa". Below it is a reply from Eric Enge (@stonetemple) dated 26 Jul 2016, replying to @methode with the text "not even a dampening factor?". A second tweet from Jenny Halasz (@jennyhalasz) dated 10 Oct 2017 is overlaid on the thread, with the text "#StateofSearch @methode 301/302 pass the same pagerank. But only pagerank. (Wish he would mention crawl budget)". The background shows other tweets from Ulf Heyden (@UlfHeyden) dated 7 Apr 2017. The interface includes profile pictures, names, handles, dates, and interaction icons (reply, retweet, like, share).

**Gary "鯨理" Illyes** @methode  
30x redirects don't lose Pa  
8:59 AM - 26 Jul 2016 from Zurich, Switzerland  
43 276 267

**Eric Enge** @stonetemple · 26 Jul 2016  
Replying to @methode  
not even a dampening factor?  
1 1

**Gary "鯨理" Illyes** @methode · 26 Jul 2016  
for PageRank, no.  
1 4

**Jenny Halasz** @jennyhalasz  
#StateofSearch @methode 301/302 pass the same pagerank. But only pagerank. (Wish he would mention crawl budget).  
2:45 PM - 10 Oct 2017  
5

**Ulf Heyden** @UlfHeyden · 7 Apr 2017





# LINKS, DIE KEINE LINKS SIND

Use proper link markup!

```
<a href="/stuff">Yes, good link! ~(^-^~</a>  
<a href="/stuff" onclick="goTo('stuff')">Okay ~\_(\ツ)\_/~</a>  
<a href="#" onclick="goTo('stuff')">Please don't. (...)</a>  
<a href="javascript:goTo('stuff')">Still wrong! ๓_๓</a>  
<a href="/stuff" onclick="changePage('loljk')">No, just don't. ๓_๓</a>  
<span onclick="goTo('stuff')">WAT? ๓益๓</span>  
<span onclick="goTo('stuff')">WAT NO (ノ°□°)ノ 一_一</span>
```


@g33konaut



# SERVERFEHLER MEIDEN



# BROKEN CODE MEIDEN

 **Kristine Schachinger** @schachin · Mar 12  
.@badams you did the presentation that **iframes in the head tag can break** Google crawler yes? Have you had an issue with Akamai tracking (which puts an iframe in the header (sigh why?)). Do you know about this? cc @JohnMu

1. Validate whether the Boomerang `<IFRAME>` in the `<HEAD>`.

After the `loader.snippet` executes, it will create an `<IFRAME>` in the `<HEAD>` of the HTML document. This `<IFRAME>` is used to load `boomerang.js` and `config.js[on]`.

This `<IFRAME>` must remain in the page for Boomerang to work.


3 1 8

 **Barry Adams**  @badams

Following

Replying to @schachin @JohnMu

Yes and yes. The Akamai iframe broke Google's processing of the site's hreflang tags. Definitely an issue.

 **John** 🙌 @JohnMu · Mar 12  
We don't have exceptions for sites that "can't make a clean head section", so you either have to fix it (maybe move it to the bottom of the head?), use a different method (headers, sitemap), or assume we won't always be able to pick it up.

1 2 13

## Testen

### Testtool für strukturierte Daten - Google

<https://search.google.com/structured-data/testing-tool?hl=de>

Testtool für strukturierte Daten. Anmelden · Google-Konto · Google Assistant · Suche · Maps · YouTube · Play · News · Gmail · Kontakte · Drive · Kalender.

#### Google Structured Data

Google apps. Main menu. Structured Data Testing Tool ...

[Weitere Ergebnisse von google.com »](#)

#### Testing Tool

Google apps. Main menu. Structured Data Testing Tool ...



Unsere Mission: Die Informationen  
dieser Welt organisieren und  
allgemein zugänglich und nutzbar  
machen.



# NEUE LÖSUNGEN MÜSSEN FÜR DIE SUCHMASCHINE VERSTÄNDLICH SEIN

## PWA



	Native App	Responsive Website	Progressive Web App
Offline-Modus	✓	✗	✓
Push-Benachrichtigungen	✓	✗	✓
App-Look & Feel	✓	✗	✓
Offline-Modus	✓	✗	✓
Push-Benachrichtigungen	✓	✗	✓
App-Look & Feel	✓	✗	✓
Offline-Modus	✓	✗	✓
Push-Benachrichtigungen	✓	✗	✓
App-Look & Feel	✓	✗	✓

- » Ein Nutzer muss mindestens 30 Sekunden mit der Domain interagieren
- » Weitere Infos:  
<https://developers.google.com/web/fundamentals/web-app-manifest/>

Quelle: [https://de.wikipedia.org/wiki/Progressive\\_Web\\_App](https://de.wikipedia.org/wiki/Progressive_Web_App)



- » Eine Progressive Web Application ist eine Website, die als eine App agiert.
- » Vier technologische Voraussetzungen:
  - (1) manifest.json File
  - (2) Service Workers
  - (3) HTTPs
  - (4) Application Shell

## SPA

- » Single Page Application ist eine Webanwendung, die aus einem HTML-Dokument besteht, dessen Inhalte dynamisch nachgeladen werden.



Quelle: [https://de.wikipedia.org/wiki/Single\\_Page\\_Application](https://de.wikipedia.org/wiki/Single_Page_Application)

## AMP



- » Das AMP Projekt ist ein gemeinsames Projekt von Google und Twitter
- » Es wurde primär für Verlage entwickelt
- » Veröffentlicht von Google im Oktober 2015
- » Apache Lizenz – Open Source Code bei GitHub



Quelle: [https://de.wikipedia.org/wiki/Accelerated\\_Mobile\\_Page](https://de.wikipedia.org/wiki/Accelerated_Mobile_Page)



2019

**PWA**

- » Eine Progressive Web Application ist eine Website, die als eine App agiert.
- » Vier technologische Voraussetzungen:
  - (1) manifest.json File
  - (2) Service Workers
  - (3) HTTPS
  - (4) Application Shell
- » Ein Nutzer muss mindestens 30 Sekunden mit der Domain interagieren
- » Weitere Info: <https://developers.google.com/web/fundamentals/web-app-manifest/>

**AMP**


- » Das AMP Projekt ist ein gemeinsames Projekt von Google und Twitter
- » Es wurde primär für Verlage entwickelt
- » Veröffentlicht von Google im Oktober 2015
- » Apache Lizenz – Open Source Code bei GitHub

**SPA**

- » Single Page Application ist eine Webanwendung, die aus einem HTML-Dokument Inhalte dynamisch nachgeladen werden.

**Google**

2015

Ilya Grigorik  @igrigorik Following

Rendering on Google Search: [bit.ly/2wffmpL](http://bit.ly/2wffmpL) - web rendering is based on Chrome 41; use feature detection, polyfills, and log errors!

Google Search

Rendering on Google Search

Googlebot uses a web rendering service (WRS) that is based on Chrome 41 (M41). Generally, WRS supports the same web platform features and capabilities that the Chrome version it uses - for a full list refer to [chromestatus.com](http://chromestatus.com), or use the [compare function on caniuse.com](http://caniuse.com).

Track errors and debug rendering problems

In addition to using feature detection and polyfills, be sure to collect and analyze client-side errors that may prevent your application from behaving correctly when accessed by users and Googlebot.

NEXT: DEBUG RENDERING PROBLEMS

10:47 PM - 4 Aug 2017

178 Retweets 315 Likes



5 178 315



Tweet your reply



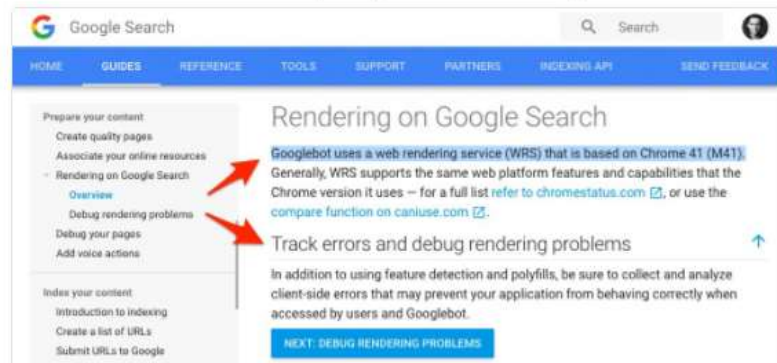
# DIE INHALTE MÜSSEN LESBAR SEIN



Ilya Grigorik   
@igrigorik

Following

Rendering on Google Search: [bit.ly/2wffmpL](https://bit.ly/2wffmpL) - web rendering is based on Chrome 41; use feature detection, polyfills, and log errors!



10:47 PM - 4 Aug 2017

178 Retweets 315 Likes



5 178 315



Tweet your reply

- » Verwendung von serverseitigem Rendering
- » Inhalte mit verfügbaren URLs, keine #
- » JavaScript und CSS-Dateien dürfen nicht blockiert werden
- » Inhalte, die nach einem ausgeführten Event erreichbar sind, sind schwer zu indexieren (History API)

Mehr dazu: <https://builtvisible.com/on-infinite-scroll-pushstate/>



# **CLIENT & SERVER PROBLEME**



# PRÜFEN



Tools

Google Search Console

- URL PRÜFEN
- BLOCKIERUNG DURCH ROBOTS.TXT PRÜFEN
- ABRUF WIE DURCH GOOGLE
- ALS SUCHERGEBNIS ANSEHEN

keycdn

- Site Speed Test
- Performance Test
- IP Location Finder

Screamingfrog

WEBPAGETEST

Server / Client



# SUCHMASCHINE PROBLEME



 **Glenn Gabe** @glenngabe - Apr 5, 2016   
And there may, or may not be, a great app coming out soon to help you easily analyze log files. Stay tuned. :) #SEO

  **John**    
@JohnMu

[@glenngabe](#) Log files are so underrated, so much good information in them.

♥ 17 9:53 PM - Apr 5, 2016 



**Endlose  
Daten**

# ANALYSE MIT FRAGEN STARTEN

**Was crawlen die Suchmaschinen?**

**Werden wichtige Seiten weniger gecrawlt als Unwichtige?**

**Gibt es Seiten, die nicht in der Sitemap sind?**

**Werden alle Sprachversionen entdeckt?**

**Mögliche Fragen**

**Gibt es veraltete Seiten (http, Dateiendungen usw.)?**

**Gibt es verwaiste Seiten?**

**Haben sich die Status Codes auf den Seiten geändert?**

**Mögliche Fragen**



# PRÜFEN



**Logfile Analyse Tools**

**Graylog, logrunner, logz.io, loggly, sumo logic, papertrail**  
Besonders logrunner.io sowie logz.io mit starkem Fokus auf SEO-basiertem Auditing (vordefinierte Reportings, nützliche Dashboards etc.)



**WAS MACHT MAN, DAMIT  
MAN TROTZ LANGSAMER  
VERBINDUNG DIE GANZE  
ACTION  
MITBEKOMMT?**



# SCHNELLER MIT...

## Non-nested Table

```
1 <table>
2 <table>
3 <table>
4 ...
5 </table>
6 </table>
7 </table>
```

-

```
1 <table>...</table>
2 <table>...</table>
3 <table>...</table>
```

+

## PHP Update auf 7.1.

Im Vergleich zu den früheren PHP Versionen, kann die 7.1. den Script um 26% schneller ausführen\*

## Version mit „/“

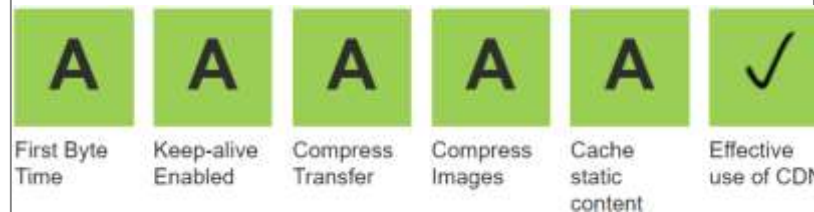
```
1 https://www.meinewebsite.de/ordner
2 https://www.meinewebsite.de/definitivordner/
```

Der Server schaut nach einer File, wenn die URL ohne “/” ist (extra Weiterleitung für den Server)

## CSS - oben, JS - unten

```
1 <!DOCTYPE html>
2 <html>
3 <head>
4 <link rel="stylesheet" href="meinstylesheet.css" />
5 </head>
6 <body>
7 Mein Hauptinhalt
8 <script type="text/javascript" src="yoursource.js"></script></body></html>
9 </body>
10 </html>
```

## Web Page Performance Test





## Hotlinking sperren

```
1 RewriteEngine onRewriteCond %{HTTP_REFERER}
2 !^$RewriteCond %{HTTP_REFERER}
3 !^http://(www\.)example.com/.*$
4 [NC]RewriteRule \.(gif|jpg|jpeg|pdf)$ - [F]
```



# DIE HAUPTENTSCHEIDUNG WIE EINE APP AUSGEFÜHRT WIRD, LIEGT AM RENDERING

	Server				Browser
					
	Server Rendering	"Static SSR"	SSR with (Re)hydration	CSR with Prerendering	Full CSR
Overview:	An application where input is navigation requests and the output is HTML in response to them.	Built as a Single Page App, but all pages prerendered to static HTML as a build step, and the JS is <b>removed</b> .	Built as a Single Page App. The server prerenders pages, but the full app is also booted on the client.	A Single Page App, where the initial shell/skeleton is prerendered to static HTML at build time.	A Single Page App. All logic, rendering and booting is done on the client. HTML is essentially just script & style tags.
Authoring:	Entirely server-side <small>(request-response, HTML)</small>	Built as if client-side <small>(components, DOM*, fetch)</small>	Built as client-side	Client-side	Client-side
Rendering:	Dynamic HTML	Static HTML	Dynamic HTML <b>and</b> JS/DOM	Partial static HTML, then JS/DOM	Entirely JS/DOM
Server role:	Controls all aspects. <small>(with client)</small>	Delivers static HTML	Renders pages <small>(navigation requests)</small>	Delivers static HTML	Delivers static HTML
Pros:	<ul style="list-style-type: none"> <li>👍 TTI = FCP</li> <li>👍 Fully streaming</li> </ul>	<ul style="list-style-type: none"> <li>👍 Fast TTFB</li> <li>👍 TTI = FCP</li> <li>👍 Fully streaming</li> </ul>	<ul style="list-style-type: none"> <li>👍 Flexible</li> </ul>	<ul style="list-style-type: none"> <li>👍 Flexible</li> <li>👍 Fast TTFB</li> </ul>	<ul style="list-style-type: none"> <li>👍 Flexible</li> <li>👍 Fast TTFB</li> </ul>
Cons:	<ul style="list-style-type: none"> <li>👎 Slow TTFB</li> <li>👎 Inflexible</li> </ul>	<ul style="list-style-type: none"> <li>👎 Inflexible</li> <li>👎 Leads to hydration</li> </ul>	<ul style="list-style-type: none"> <li>👎 Slow TTFB</li> <li>👎 TTI &gt;&gt;&gt; FCP</li> <li>👎 Usually buffered</li> </ul>	<ul style="list-style-type: none"> <li>👎 TTI &gt; FCP</li> <li>👎 Limited streaming</li> </ul>	<ul style="list-style-type: none"> <li>👎 TTI &gt;&gt;&gt; FCP</li> <li>👎 No streaming</li> </ul>
Scales via:	Infra size / cost	build/deploy size	Infra size & JS size	JS size	JS size
Examples:	Gmail HTML, Hacker News	Docusaurus, Netflix*	<a href="#">Next.js</a> , <a href="#">Razzle</a> , etc	Gatsby, Vuepress, etc	Most apps





# BEIM RENDERING SEO BEACHTEN



	Server Side Rendering	"Static SSR"	SSR with (Re)hydration	CSR with Prerendering	Full CSR
<i>Overview:</i>	An application where input is navigation request and the output is HTML in response to them,	Built as a Single Page App, but all pages prerendered to static HTML as a built step and the JS is removed.	Built as a Single Page App. The server prerenders pages, but the full app is also booted on the client	A Single Page App, where the initial shell/ skeleton is prerendered to a static HTML at build time.	A Single Page App. All logic, rendering and booting is done on the client. HTML is essentially just script & style tags.
<i>Authoring:</i>	Entirely server-side	Built as if client-side	Built as client-side	Client-side	Client-side
<i>Rendering:</i>	Dynamic HTML	Static HTML	Dynamic HTML and JS/DOM	Partial Static HTML, then JS/DOM	Entirely JS/DOM
<i>Server role:</i>	Controls all aspects	Delivers static HTML	Renders Page	Delivers static HTML	Delivers static HTML
<i>SEO Pros:</i>	All SEO required elements will be available in initial response	SEO required elements will be available in initial response	SEO required elements can be available in initial response	Initial shell can contain SEO required elements	What do you think yourself?
<i>SEO Cons:</i>	None	Make sure outdated content is not cached	Make sure outdated content is not cached	Implementations forget populating elements with SEO datapoints	Everything!



# VIELEN DANK



plan-net.com | Creating Relevance

<https://www.plan-net.com/de/>

LinkedIn: <http://www.linkedin.com/in/alonademchuk>

Xing: [https://www.xing.com/profile/Alona\\_Demchuk](https://www.xing.com/profile/Alona_Demchuk)

Twitter: <http://twitter.com/alonademchuk>



# BILDER- UND ICONS-QUELLEN

- » <https://unsplash.com>
- » <https://www.pexels.com>
- » <https://thenounproject.com>
- » <https://giphy.com/gifs/bc-the-matrix-BMuWMPmykHL6E>

