UNIVERSITÄT ZU LÜBECK
INSTITUT FÜR
THEORETISCHE INFORMATIK

Täuschungsstrategien gegen Fingerprinting im Webbrowser

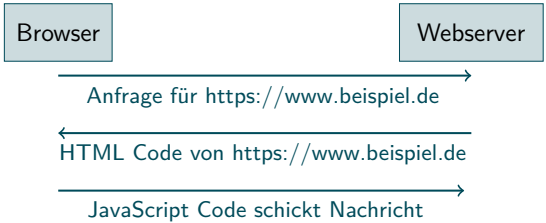
Camouflage strategies to protect against Browser Fingerprinting

Kolloqium von Malte Luttermann

28.10.2019

Worum geht es hier?

- *Fingerprinting*: Eine spezielle Web Tracking Variante



- Programmcode im Browser ermittelt die *System- und Browserkonfiguration*
- Generierung eines *Fingerprints* anhand der Konfiguration

Beispielcode Fingerprinting

```
<!doctype html>

<!-- HTML Code hier -->

<script>
  let data = JSON.stringify({
    platform: navigator.platform,
    plugins: navigator.plugins,
    ...
  });
  let http = new XMLHttpRequest();
  http.open('POST', 'https://www.beispiel.de/data.php');
  http.send(data);
</script>

</html>
```

Was tut man gegen Fingerprinting?

- Die *Herausforderung*:

> Wie können Nutzer das Internet uneingeschränkt nutzen, ohne dass sie dabei verfolgt werden?

- Mögliche *Lösungsideen*:
  - JavaScript deaktivieren?
  - Eine falsche Konfiguration vortäuschen?

$\Rightarrow$ Nicht ausreichend und/oder Einschränkung der Nutzererfahrung.

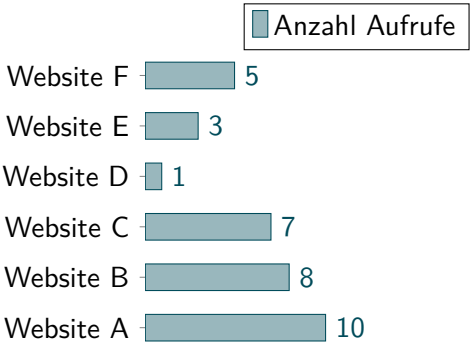
Wie können wir das besser machen?

Idee: Falsches Verhalten vortäuschen, durch

1. zusätzliche *Seitenaufrufe* und
2. zusätzliche *Suchanfragen*.

Welche Seiten sollen besucht werden?

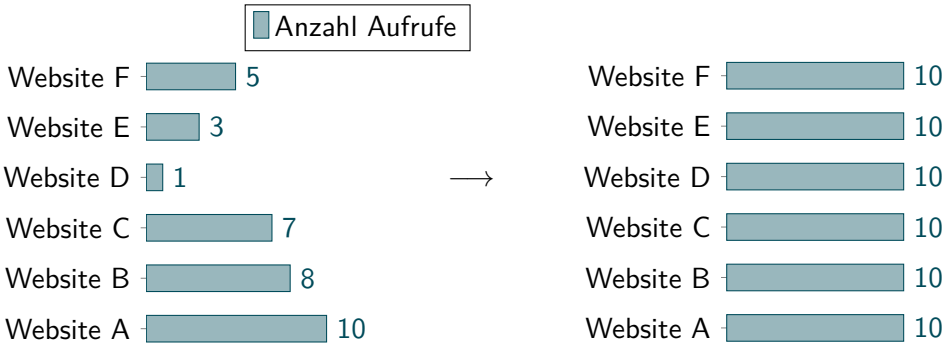
- Seiten aus dem *Browserverlauf*
- *Häufigkeit* der Aufrufe:



(Beispieldaten)

Welche Seiten sollen besucht werden?

- Seiten aus dem *Browserverlauf*
- *Häufigkeit* der Aufrufe:



(Beispieldaten)

Generierung neuer Suchbegriffe

- *Idee*: Neue Suchbegriffe anhand von bereits gesuchten Begriffen generieren
- *Ziel*: Gleiche Anzahl an Wörtern und alle Wörter syntaktisch verschieden

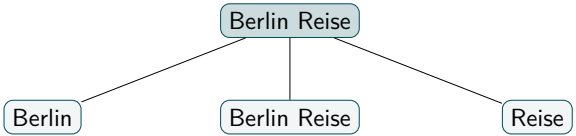
Generierung neuer Suchbegriffe

- *Idee*: Neue Suchbegriffe anhand von bereits gesuchten Begriffen generieren
- *Ziel*: Gleiche Anzahl an Wörtern und alle Wörter syntaktisch verschieden
- *Beispiel* für den Begriff „Berlin Reise“:

Berlin Reise

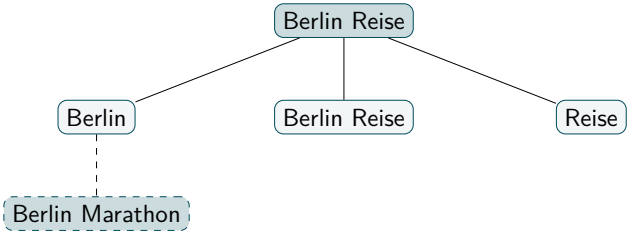
Generierung neuer Suchbegriffe

- *Idee*: Neue Suchbegriffe anhand von bereits gesuchten Begriffen generieren
- *Ziel*: Gleiche Anzahl an Wörtern und alle Wörter syntaktisch verschieden
- *Beispiel* für den Begriff „Berlin Reise“:



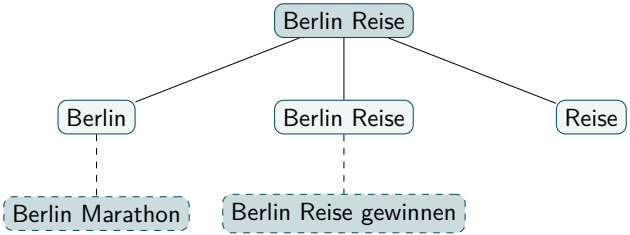
Generierung neuer Suchbegriffe

- *Idee*: Neue Suchbegriffe anhand von bereits gesuchten Begriffen generieren
- *Ziel*: Gleiche Anzahl an Wörtern und alle Wörter syntaktisch verschieden
- *Beispiel* für den Begriff „Berlin Reise“:



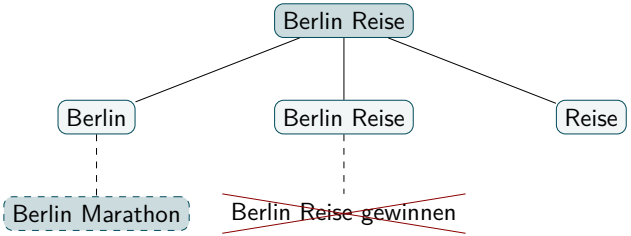
Generierung neuer Suchbegriffe

- *Idee*: Neue Suchbegriffe anhand von bereits gesuchten Begriffen generieren
- *Ziel*: Gleiche Anzahl an Wörtern und alle Wörter syntaktisch verschieden
- *Beispiel* für den Begriff „Berlin Reise“:



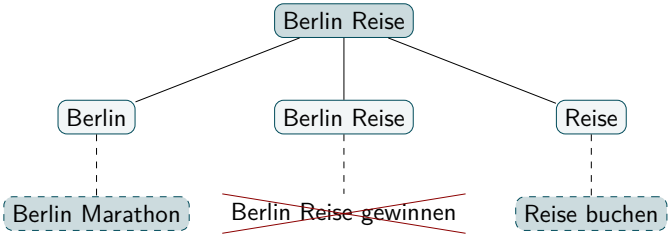
Generierung neuer Suchbegriffe

- *Idee*: Neue Suchbegriffe anhand von bereits gesuchten Begriffen generieren
- *Ziel*: Gleiche Anzahl an Wörtern und alle Wörter syntaktisch verschieden
- *Beispiel* für den Begriff „Berlin Reise“:



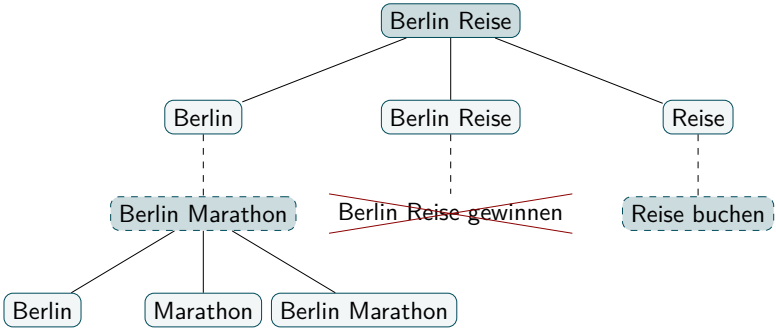
Generierung neuer Suchbegriffe

- *Idee*: Neue Suchbegriffe anhand von bereits gesuchten Begriffen generieren
- *Ziel*: Gleiche Anzahl an Wörtern und alle Wörter syntaktisch verschieden
- *Beispiel* für den Begriff „Berlin Reise“:



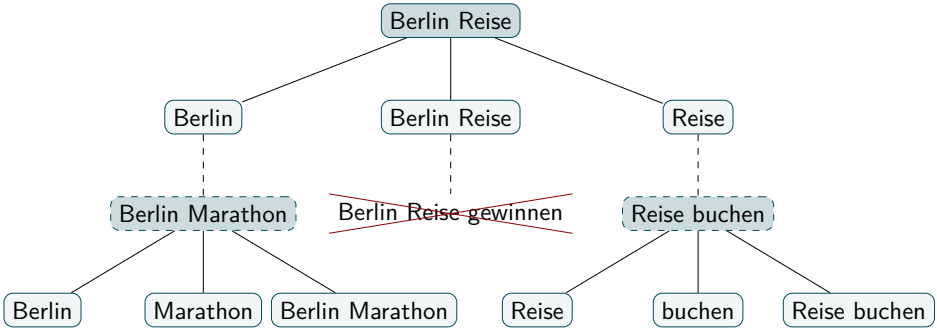
Generierung neuer Suchbegriffe

- *Idee*: Neue Suchbegriffe anhand von bereits gesuchten Begriffen generieren
- *Ziel*: Gleiche Anzahl an Wörtern und alle Wörter syntaktisch verschieden
- *Beispiel* für den Begriff „Berlin Reise“:



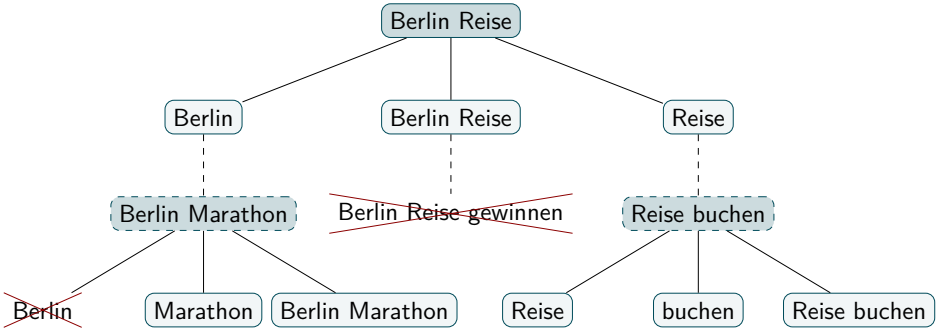
Generierung neuer Suchbegriffe

- *Idee*: Neue Suchbegriffe anhand von bereits gesuchten Begriffen generieren
- *Ziel*: Gleiche Anzahl an Wörtern und alle Wörter syntaktisch verschieden
- *Beispiel* für den Begriff „Berlin Reise“:



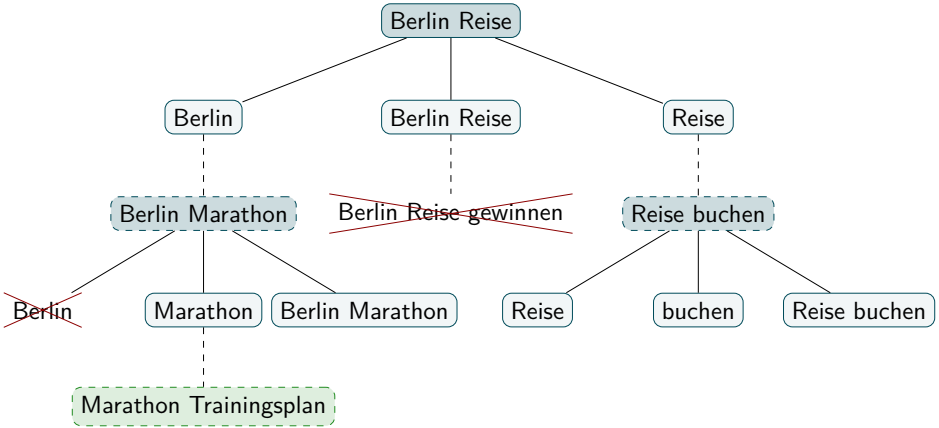
Generierung neuer Suchbegriffe

- *Idee*: Neue Suchbegriffe anhand von bereits gesuchten Begriffen generieren
- *Ziel*: Gleiche Anzahl an Wörtern und alle Wörter syntaktisch verschieden
- *Beispiel* für den Begriff „Berlin Reise“:



Generierung neuer Suchbegriffe

- *Idee*: Neue Suchbegriffe anhand von bereits gesuchten Begriffen generieren
- *Ziel*: Gleiche Anzahl an Wörtern und alle Wörter syntaktisch verschieden
- *Beispiel* für den Begriff „Berlin Reise“:



Analyse von URLs, um Suchbegriffe zu erhalten

- *Problem*: Wie finden wir bereits gesuchte Begriffe heraus?
- *Eingabe*: Eine URL aus dem Browserverlauf, zum Beispiel:
  `https://www.beispiel.de?a=2313&s=berlin+reise&bb=d`
  Parameter
- *Ausgabe*: Gesuchter Suchbegriff während des Besuchs auf der URL

Analyse von URLs, um Suchbegriffe zu erhalten

Vorgehen anhand beispielhafter URL:

`https://www.beispiel.de?a=2313&s=berlin+reise&bb=d`

1. URL ohne Parameter aufrufen:
   - `https://www.beispiel.de`

Analyse von URLs, um Suchbegriffe zu erhalten

Vorgehen anhand beispielhafter URL:

```
https://www.beispiel.de?a=2313&s=berlin+reise&bb=d
```

1. URL ohne Parameter aufrufen:
   - `https://www.beispiel.de`
2. Nach einem zufällig gewählten Suchbegriff, zum Beispiel „abc123“, suchen:
   - `https://www.beispiel.de?a=2314&s=abc123&bb=d`

Analyse von URLs, um Suchbegriffe zu erhalten

Vorgehen anhand beispielhafter URL:

`https://www.beispiel.de?a=2313&s=berlin+reise&bb=d`

1. URL ohne Parameter aufrufen:
   - `https://www.beispiel.de`
2. Nach einem zufällig gewählten Suchbegriff, zum Beispiel „abc123“, suchen:
   - `https://www.beispiel.de?a=2314&s=abc123&bb=d`
3. Den zugehörigen Parameter in der URL finden (nach „abc123“ suchen):
   - `https://www.beispiel.de?a=2314&s=abc123&bb=d`

Parameter: *s*

Analyse von URLs, um Suchbegriffe zu erhalten

Vorgehen anhand beispielhafter URL:

`https://www.beispiel.de?a=2313&s=berlin+reise&bb=d`

1. URL ohne Parameter aufrufen:
   - `https://www.beispiel.de`
2. Nach einem zufällig gewählten Suchbegriff, zum Beispiel „abc123“, suchen:
   - `https://www.beispiel.de?a=2314&s=abc123&bb=d`
3. Den zugehörigen Parameter in der URL finden (nach „abc123“ suchen):
   - `https://www.beispiel.de?a=2314&s=abc123&bb=d`

   Parameter: *s*
4. In der ursprünglichen URL den korrespondierenden Wert für den Parameter auslesen:
   - `https://www.beispiel.de?a=2313&s=berlin+reise&bb=d`

   Wert für *s*

Automatisiert nach Suchbegriffen suchen

- *Problem*: Wie finden wir das Suchfeld auf einer beliebigen Website?
- *Idee*: CSS-Selektoren nutzen:

```
// Suchfeld finden
var inputField = $(':input[type=search]').first();

if (inputField != null) {
  // Zuvor ermittelten Suchbegriff in das Suchfeld eingeben
  $(inputField).val(searchTerm);

  // Suche starten
  $(inputField).closest('form').submit();
}
```

- *Sonderfälle* müssen berücksichtigt werden

Evaluation

- Ablauf:
  - *Zwei Systeme* für *eine Woche* lang im Internet unterwegs
  - Suche nach „*Herren Schuhe*“ auf beiden Systemen
  - *System 1* nutzt die vorgestellten Täuschungsstrategien, *System 2* nicht
  - *Frage*: Gibt es auffällige Unterschiede zwischen den Werbeanzeigen, die wir erhalten?

Evaluation

- Ablauf:
  - *Zwei Systeme* für *eine Woche* lang im Internet unterwegs
  - Suche nach „*Herren Schuhe*“ auf beiden Systemen
  - *System 1* nutzt die vorgestellten Täuschungsstrategien, *System 2* nicht
  - *Frage*: Gibt es auffällige Unterschiede zwischen den Werbeanzeigen, die wir erhalten?
- Auszug der Ergebnisse (vom ersten Tag):

Seite	Werbung für System 1	Werbung für System 2
wetter.com	TV-Streaming	Fotobuch und Schuhe auf amazon.de
tz.de	Outdoorbekleidung	Dell Laptop
focus.de	ADAC (Pannenhilfe)	ADAC (Pannenhilfe)
web.de	web.de Mobilfunk	Lotto (Glücksspiel)

- Automatisch gesuchte Suchbegriffe (vom ersten Tag) scheinen noch nicht mit Werbeanzeigen zusammen zu hängen

Evaluation

- Ablauf:
  - *Zwei Systeme* für *eine Woche* lang im Internet unterwegs
  - Suche nach „*Herren Schuhe*“ auf beiden Systemen
  - *System 1* nutzt die vorgestellten Täuschungsstrategien, *System 2* nicht
  - *Frage*: Gibt es auffällige Unterschiede zwischen den Werbeanzeigen, die wir erhalten?
- Auszug der Ergebnisse (vom letzten Tag):

Seite	Werbung für System 1	Werbung für System 2
wetter.com	Mazda und Eurowings	Kaminausstellung und Immobilien
tz.de	Eurowings	Snipes Schuhe
focus.de	Treppenlift	Treppenlift
web.de	Aldi	Aldi

- Automatisch gesuchte Suchbegriffe umfassen neben anderen:
  „*skyscanner*“ (mehrfach) und „*kielius*“

  „*skyscanner*“: Flüge vergleichen und buchen

  „*kielius*“: Bus zum Flughafen

Zusammenfassung und Ausblick

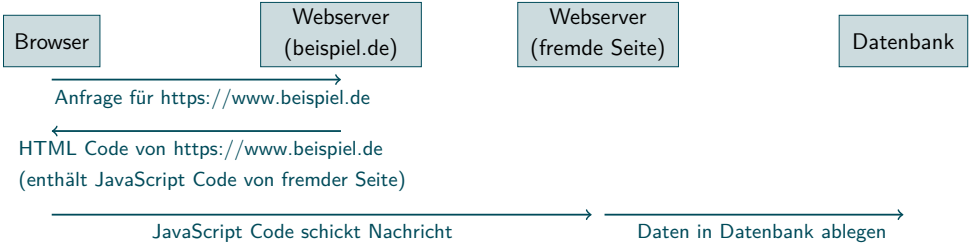
- Zusammenfassung der Ergebnisse:
  - *Extraktion* von Suchbegriffen aus URLs
  - *Generierung* neuer Suchbegriffe mit Hilfe von Google
  - Automatisiert nach Suchbegriffen *suchen*

Zusammenfassung und Ausblick

- Zusammenfassung der Ergebnisse:
  - *Extraktion* von Suchbegriffen aus URLs
  - *Generierung* neuer Suchbegriffe mit Hilfe von Google
  - Automatisiert nach Suchbegriffen *suchen*
- Ausblick für die Zukunft:
  - Suchergebnisse *auswählen*?
    - ↪ *Problem*: Semantische Programmtextanalyse

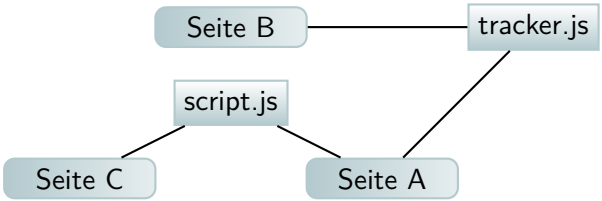
Anhang

Anhang I – Third-party Web Tracking



Anhang II – Third-party Web Tracking Graph

Besuchte Seite	Angefragte JavaScript Datei
Seite A	facebook.de/tracker.js
Seite A	google.de/script.js
Seite B	facebook.de/tracker.js
Seite C	google.de/script.js



(Beispieldaten)

Anhang III – Beispiel für einen Fingerprint

Merkmal	Wert
User Agent	Mozilla/5.0 (Windows NT 10.0; Win64; x64) AppleWebKit/537.36 (KHTML, like Gecko) Chrome/75.0.3770.100 Safari/537.36
Accept Header	text/html,application/xhtml+xml,application/xml;q=0.9,image/webp,image/apng,*/*;q=0.8,application/signed-exchange;v=b3
Zeichenkodierung	gzip, deflate, br
Bevorzugte Sprachen	de-DE,de;q=0.9,en-US;q=0.8,en;q=0.7
Host	amiunique.org
Upgrade-Insecure-Requests Header	1
Referrer	https://amiunique.org/
IP-Adresse	31.17.50.35
Protokoll	https
...	...

Anhang IV – Beispiel für einen Fingerprint

Merkmal	Wert
Betriebssystem	Win32
Cookies aktiviert?	Ja
Zeitzone	-120
Installierte Schriftarten	Arial, Arial Black, Arial Narrow, Bahnschrift, Brush Script MT, Calibri, Cambria, Cambria Math, Candara, Comic Sans MS und 56 weitere
Canvas	
Do Not Track Header	nicht angegeben
Local storage aktiv	Ja
Session storage aktiv	Ja
...	...

Anhang V – Beispiel für einen Fingerprint

Merkmal	Wert
Bildschirmbreite	1920
Bildschirmhöhe	1080
Farbtiefe	24
Erster Pixel oben	153
Erster Pixel links	-1920
Verfügbare Breite	1920
Verfügbare Höhe	1040
Pixeldistanz links	undefiniert
Pixeldistanz oben	undefiniert
Plugins	Plugin 0: Chrome PDF Plugin; Portable Document Format; internal-pdf-viewer. Plugin 1: Chrome PDF Viewer; ; mhjfbmdgcfjbbpaeojofohoefgiehjai. Plugin 2: Native Client; ; internal-nacl-plugin.
BuildID	undefiniert
...	...

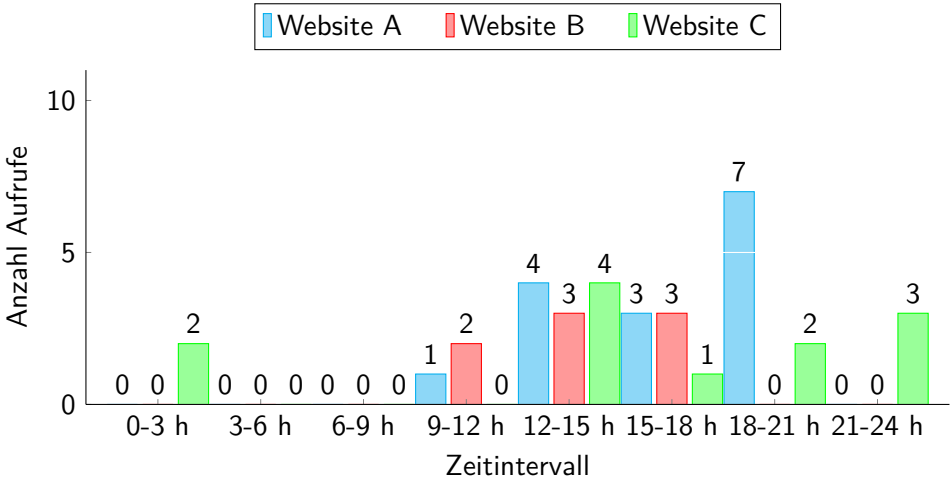
Anhang VI – Beispiel für einen Fingerprint

Merkmal	Wert
Audioformate	audio/aac, audio/flac, audio/mpeg, audio/mp4; codecs=“mp4a.40.2“, audio/ogg; codecs=“flac“, audio/ogg; codecs=“vorbis“, audio/ogg; codecs=“opus“, audio/wav; codecs=“1“, audio/webm; codecs=“vorbis“, audio/webm; codecs=“opus“
Videoformate	video/mp4; codecs=“flac“, video/ogg; codecs=“theora“, video/ogg; codecs=“opus“, video/webm; codecs=“vp9, opus“, video/webm; codecs=“vp8, vorbis“
WebGL Vendor	Google Inc.
WebGL Renderer	ANGLE (NVIDIA GeForce GTX 970 Direct3D11 vs_5_0 ps_5_0)
WebGL Daten	
...	...

Anhang VII – Beispiel für einen Fingerprint

Merkmal	Wert
WebGL Parameter	3380 verschiedene Parameter
Installierte Schriftarten (Flash)	Kein Flash vorhanden
Bildschirmauflösung (Flash)	Kein Flash vorhanden
Betriebssystem (Flash)	Kein Flash vorhanden
Adblock aktiviert?	Ja

Anhang VIII – Besuchsstatistik mit Zeiten



(Beispieldaten)

Anhang IX – Ergebnisse der Evaluation (Tag 1)

Seite	Gesuchte Begriffe durch FPFool
amazon	hotmail, h&m, speedtest, hagebaumarkt, hermes, skyscanner, sane, hvv, samsung, hvv, spotify, drangstedt, wochenendticket, denizli, speedtest, skyscanner, hagebaumarkt, bankdrücken, h&m, speedtest, saturn
google	hermes, hagebaumarkt, speedtest, haspa, skyscanner, h&m, hsv, spotify, saturn, hagebaumarkt, parkett, hotmail, pinterest

Seite	Werbung mit FPFool	Werbung ohne FPFool
wetter.com	TV-Streaming	Fotobuch und Schuhe auf amazon.de
tz.de	Outdoorbekleidung	Dell Laptop
focus.de	ADAC (Pannenhilfe)	ADAC (Pannenhilfe)
web.de	web.de Mobilfunk	Lotto (Glücksspiel)

Anhang X – Ergebnisse der Evaluation (Tag 2)

Seite	Gesuchte Begriffe durch FPFool
amazon	speedtest, skyscanner, spotify, spotify, spiegel, domino, spiegel, hotmail, moldaustausee, sheego
google	spotify, spotify, hermes, hsv, skyscanner, penny, hermes, spiegel, parkett, hvv

Seite	Werbung mit FPFool	Werbung ohne FPFool
wetter.com	Gelenkschmerzen, Krankenversicherung und Nike Schuhe	Gumbies (Schuhe), Tagesgeldkonto und H&M Bekleidung
tz.de	Nike Jacken, Nike Schuhe auf footlocker.de und Nike Schuhe	Regalsysteme, Schuhe auf amazon.de und Dell Laptop
focus.de	keine Werbung, keine Werbung und Gehaltsvergleich	keine Werbung, 1&1 und ADAC sowie FRITZ!Box und ADAC
web.de	web.de Energie, web.de Bonusprogramm und web.de Mobilfunk	web.de Energie, web.de Energie und web.de Mobilfunk

Anhang XI – Ergebnisse der Evaluation (Tag 3)

Seite	Gesuchte Begriffe durch FPFool
amazon	news, neumünster, smileys, speedtest, nospa, netflix, shein, spotify, netflix, sternzeichen, sport1, spiegel, nachrichten, spotify
google	hornbach, skyscanner, undertale, nachrichten, news, nike, hagebaumarkt, spotify

Seite	Werbung mit FPFool	Werbung ohne FPFool
wetter.com	Autoversicherung, Autoversicherung und Mazda (Auto) sowie Mazda	Mazda, Mazda und Gumbies (Schuhe) sowie Mazda
tz.de	zalando, Abnehmen und Auto verkaufen	Mastercard, Halbleiter und Bundeswehr IT
focus.de	Saturn, Saturn und Saturn	Saturn, Saturn und Saturn
web.de	Aldi, Aldi sowie Parship (Partnervermittlung) und Aldi	Tschibo und Aldi, Lidl und Aldi sowie Tschibo und Aldi

Anhang XII – Ergebnisse der Evaluation (Tag 4)

Seite	Gesuchte Begriffe durch FPFool
amazon	hlg, haspa, hotmail, versicherungsnummer, hagebaumarkt, haspa, haspa
google	ndr2, hsv

Seite	Werbung mit FPFool	Werbung ohne FPFool
wetter.com	zalando.de Schuhe	Kaminausstellung
tz.de	zalando.de Schuhe	Snipes Schuhe
focus.de	Kochpfanne	Lotto
web.de	Aldi	Aldi

Anhang XIII – Ergebnisse der Evaluation (Tag 5)

Seite	Gesuchte Begriffe durch FPFool
amazon	hsv, haspa, dompfaff, hvv, haspa, hotmail, hvv, haspa, hagebaumarkt, haspa, liverpool, hagebaumarkt, baubeschlagshop, hochzeitsgeschenke, hagebaumarkt, parkplatz
google	nike, h&m, hsv, ndr, nospa

Seite	Werbung mit FPFool	Werbung ohne FPFool
wetter.com	Kaminausstellung	Kaminausstellung
tz.de	Snipes Schuhe	Mercedes-Benz
focus.de	Treppenlift	Lotto
web.de	Aldi	Aldi

Anhang XIV – Ergebnisse der Evaluation (Tag 6)

Seite	Gesuchte Begriffe durch FPFool
amazon	hltv, vsco, hlae, wetter, programmiersprache, livestream, liveticker, wetter, testamentsvollstrecker, hvv, hagebaumarkt, hagebaumarkt, klumpen, hagebaumarkt, haspa, wochenendtrip
google	sport1, ndr, hagebaumarkt, ndr

Seite	Werbung mit FPFool	Werbung ohne FPFool
wetter.com	Snipes Schuhe	Kaminausstellung und Snipes Schuhe
tz.de	Snipes Schuhe	Mastercard
focus.de	Treppenlift	Treppenlift
web.de	Aldi	Aldi

Anhang XV – Ergebnisse der Evaluation (Tag 7)

Seite	Gesuchte Begriffe durch FPFool
amazon	kiellauf, lidl, haspa, indeed, kielius, kindergeburtstag, payback, derbe, h&m, hsv, postbank, medpex, hotmail, wahlomat, liveticker, ingolstadt, hansaland, wikipedia, lieferheld
google	hvv, netflix, ndr2

Seite	Werbung mit FPFool	Werbung ohne FPFool
wetter.com	Mazda und Eurowings	Kaminausstellung und Immobilien
tz.de	Eurowings	Snipes Schue
focus.de	Treppenlift	Treppenlift
web.de	Aldi	Aldi