

■ WEBTRACKING

A stack of old, worn books with visible spines and titles. The books are arranged in a slightly overlapping manner, showing their aged and textured covers. The lighting is warm and focused on the books, creating a sense of history and knowledge.

■ WEB TRACKING

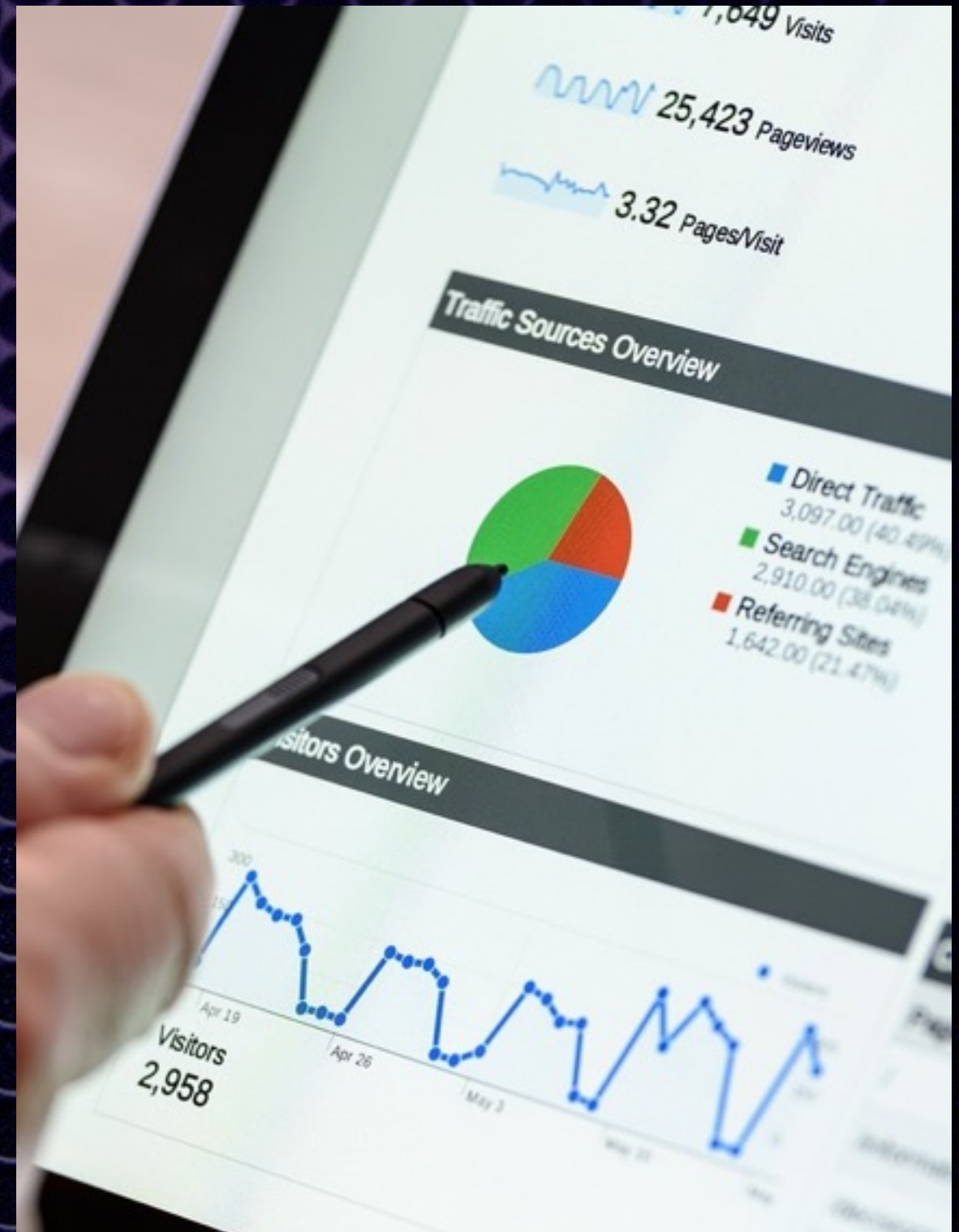
BEGRIFFSERKLÄRUNG

BEGRIFF/SYNONYME

- ▶ User-Tracking
- ▶ Website Visitor Tracking
- ▶ Ad-Tracking
- ▶ Web-Analytics
- ▶ Datenverkehrsanalyse
- ▶ Web-Controlling

ZIEL UND NUTZEN

- ▶ Marketing
 - ▶ Analyse von Interessen
 - ▶ Angebotsoptimierung
 - ▶ Zielgerichtete Werbung
 - ▶ Preisgestaltung
- ▶ Daten als Produkt





**"DER BISHER
GÜLTIGE
GRUNDSATZ, DASS
DATENSPARSAMKEIT
DAS ÜBERMAß DER
DINGE IST, DER HAT
SICH ÜBERHOLT, DER
MUSS WEG"**

Alexander Dobrindt

Bundesminister für Verkehr und digitale Infrastruktur

- ▶ Telemediengesetz § 15 Nutzungsdaten
- ▶ (3) Der Diensteanbieter darf für Zwecke der Werbung, der Marktforschung oder zur bedarfsgerechten Gestaltung der Telemedien Nutzungsprofile bei Verwendung von Pseudonymen erstellen, sofern der Nutzer dem nicht widerspricht. Der Diensteanbieter hat den Nutzer auf sein Widerspruchsrecht im Rahmen der Unterrichtung nach § 13 Abs. 1 hinzuweisen. Diese Nutzungsprofile dürfen nicht mit Daten über den Träger des Pseudonyms zusammengeführt werden.



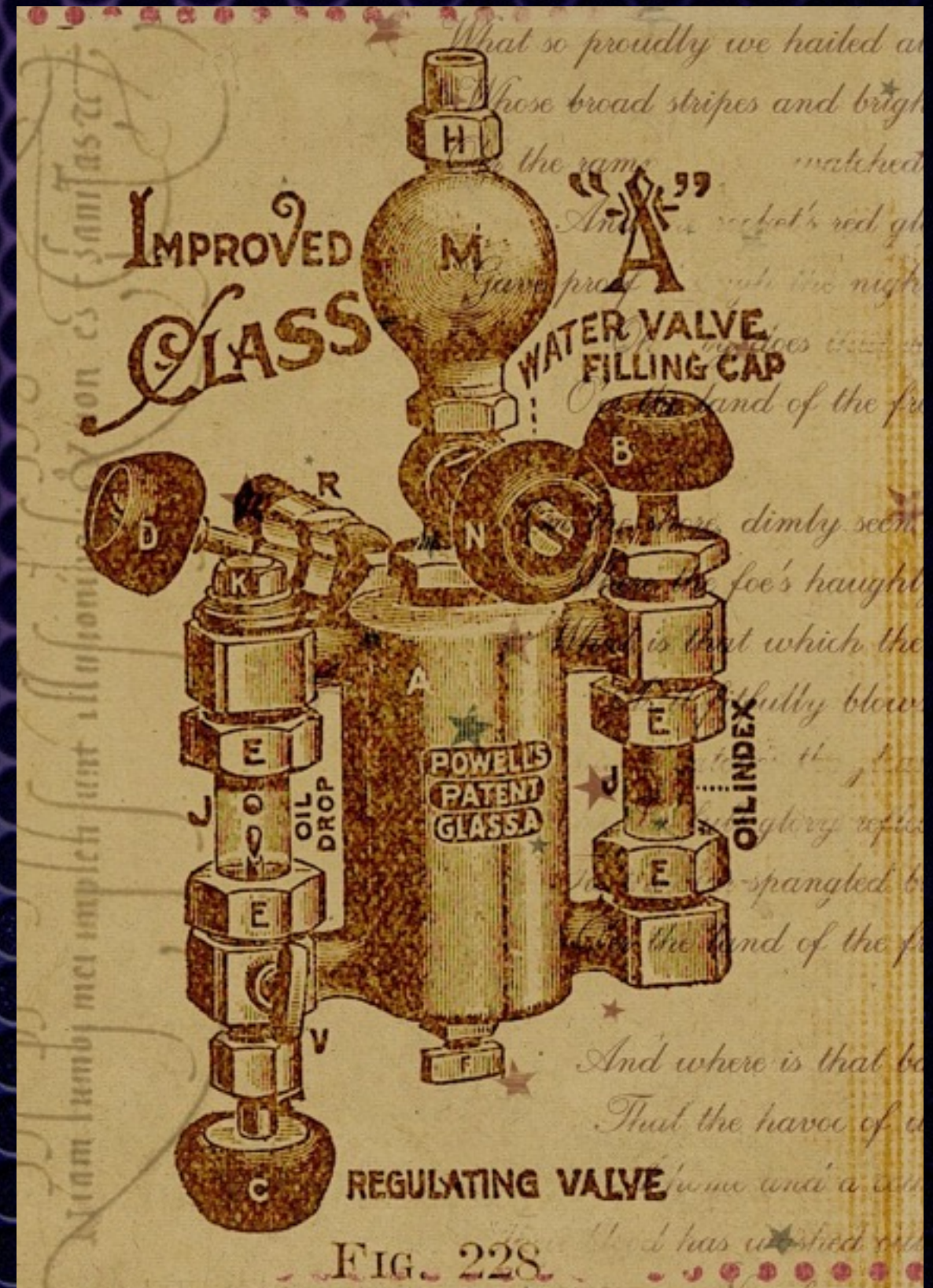


■ WEB TRACKING

TECHNIK HINTER DEN KULISSEN

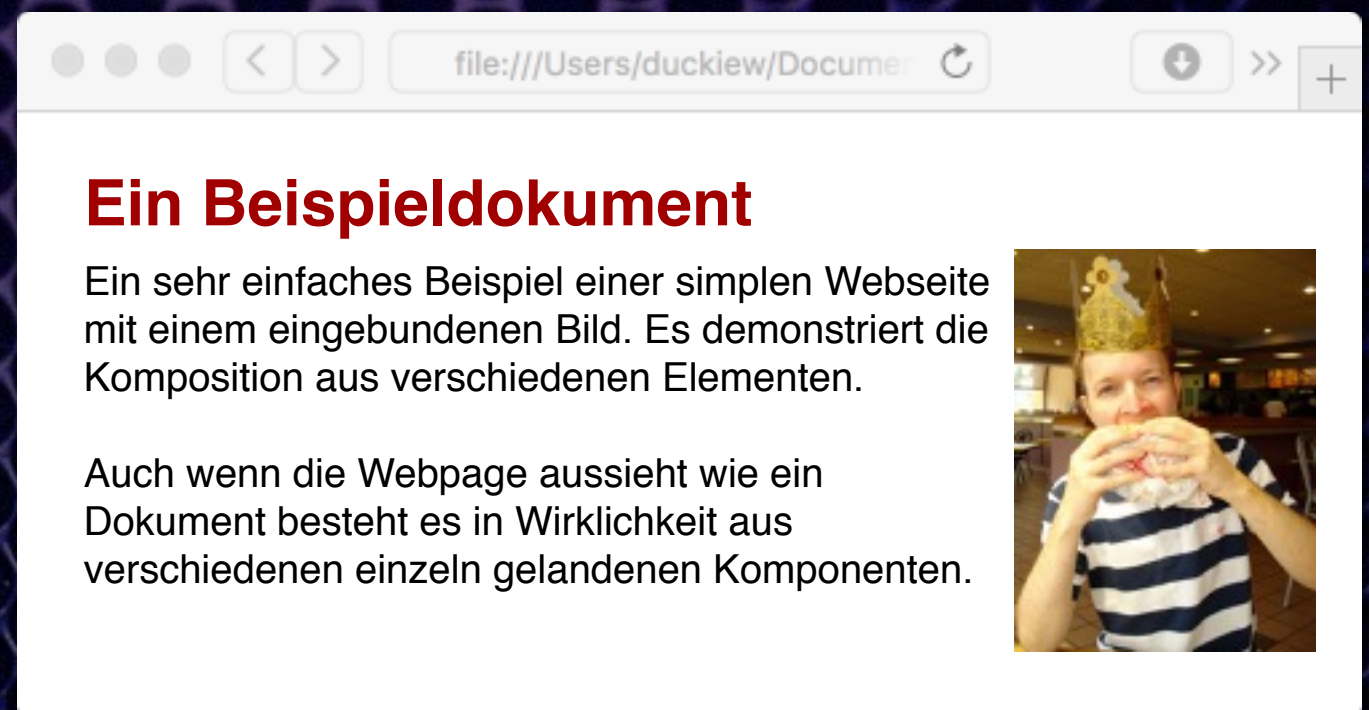
METHODEN

- ▶ Log-Analyse
- ▶ Cookies
- ▶ Pixel Tracking
- ▶ Browser-Tracking
- ▶ Social Media
- ▶ In-Browser Data
- ▶ Browser-Forensics
- ▶ Canvas-Fingerprinting



WEB BASICS

- ▶ Eine Webseite besteht aus verschiedensten Einzel-Elementen
 - ▶ Hauptdokument
 - ▶ Style-Sheet
 - ▶ Bilder
 - ▶ Skripte
 - ▶ ...




The screenshot shows a web browser window with a file:// URL. The page content includes a title, a paragraph of text, another paragraph, and a photograph of a person wearing a crown and eating a burger.

Ein Beispieldokument

Ein sehr einfaches Beispiel einer simplen Webseite mit einem eingebundenen Bild. Es demonstriert die Komposition aus verschiedenen Elementen.

Auch wenn die Webpage aussieht wie ein Dokument besteht es in Wirklichkeit aus verschiedenen einzeln geladenen Komponenten.



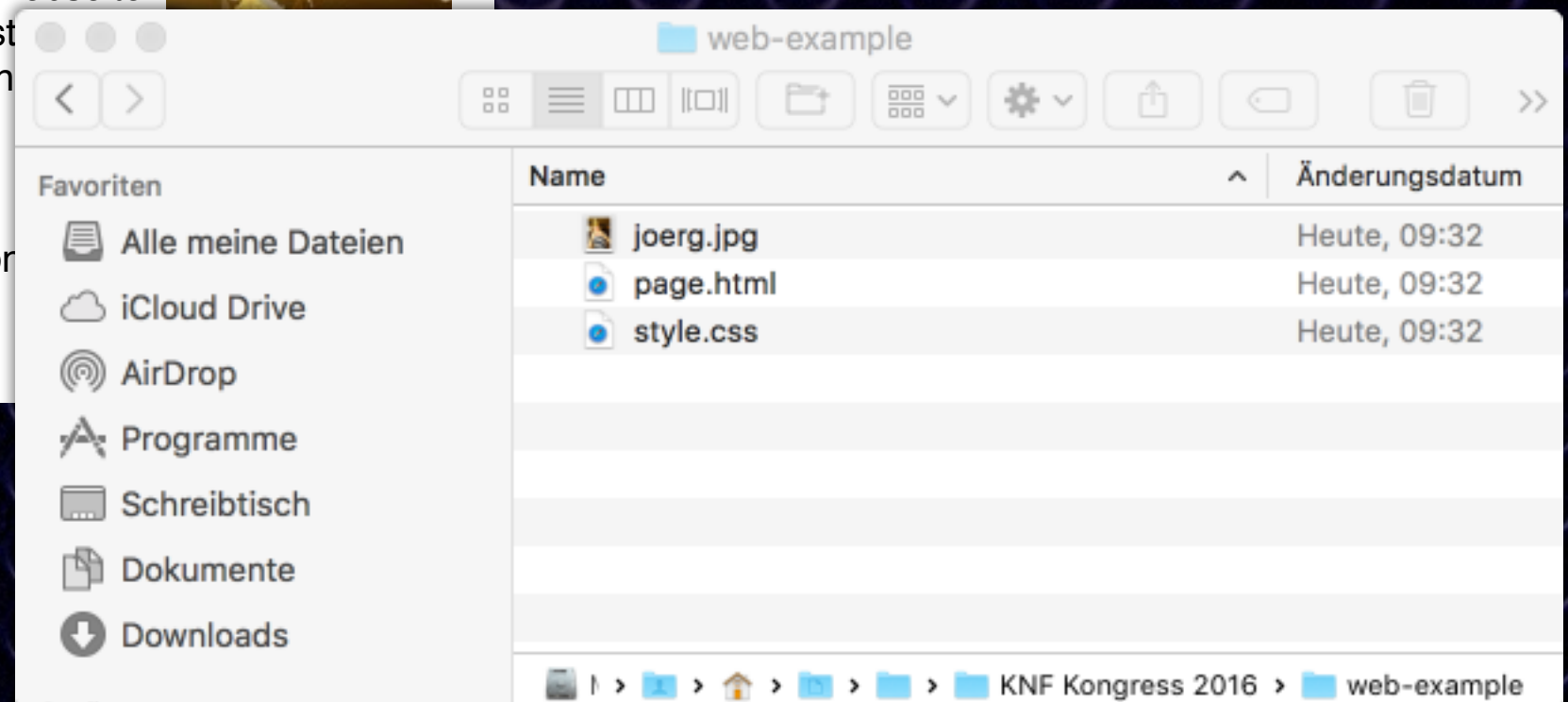
WEB BASICS

- ▶ Die einzelnen Elemente werden auch einzeln geladen...

Ein Beispieldokument

Ein sehr einfaches Beispiel einer simplen Webseite mit einem eingebundenen Bild. Es demonstriert die Komposition aus verschiedenen Elementen

Auch wenn die Webpage aussieht wie ein Dokument besteht es in Wirklichkeit aus verschiedenen einzeln geladenen Kompo



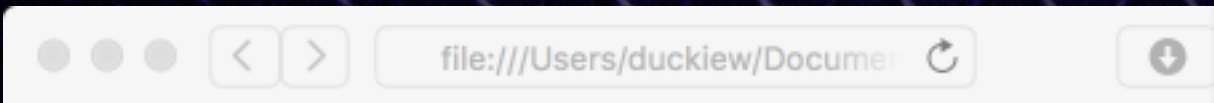
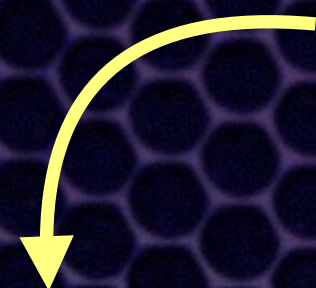
WEB BASICS

- ▶ Die einzelnen Elemente werden auch einzeln geladen...

```
page.html
!DOCTYPE html>
<html>
<head>
  <link rel="stylesheet" href="style.css">
</head>
<body>

<h1>Ein Beispieldokument</h1>
<div></div>
<p>Ein sehr einfaches Beispiel einer simplen Webseite mit einem
eingebundenen Bild. Es demonstriert die Komposition aus
verschiedenen Elementen.</p>
<p>Auch wenn die Webpage aussieht wie ein Dokument besteht es
in Wirklichkeit aus verschiedenen einzeln geladenen Komponenten.</p>

</body>
</html>
```



Ein Beispieldokument

Ein sehr einfaches Beispiel einer simplen Webseite mit einem eingebundenen Bild. Es demonstriert die Komposition aus verschiedenen Elementen.

Auch wenn die Webpage aussieht wie ein Dokument besteht es in Wirklichkeit aus verschiedenen einzeln geladenen Komponenten.



WEB BASICS

- ▶ Die einzelnen Elemente werden auch einzeln geladen...

```
page.html
<!DOCTYPE html>
<html>
<head>
<link rel="stylesheet" href="style.css">
</head>
<body>

<h1>Ein Beispieldokument</h1>
<div></div>
<p>Ein sehr einfaches Beispiel einer simplen Webseite mit einem
eingebundenen Bild. Es demonstriert die Komposition aus
verschiedenen Elementen.</p>
<p>Auch wenn die Webpage
in Wirklichkeit aus vers

</body>
</html>
```



```
file:///Users/duckiew/Documen
```

Ein Beispieldokument

Ein sehr einfaches Beispiel einer simplen Webseite mit einem eingebundenen Bild. Es demonstriert die Komposition aus verschiedenen Elementen.

Auch wenn die Webpage aussieht wie ein Dokument besteht es in Wirklichkeit aus verschiedenen einzeln geladenen Komponenten.



```
style.css
body {
background-color: #ffffff;
}
h1 {
color: #a00000;
font-family: Helvetica, Arial, sans-serif;
font-size: 24px;
}
p {
color: #000000;
font-family: Helvetica, Arial, sans-serif;
font-size: 15px;
}
div {
float: right;
width: 118px;
margin: 0 0 0 0;
padding: 1px;
text-align: center;
}
```


WEB BASICS

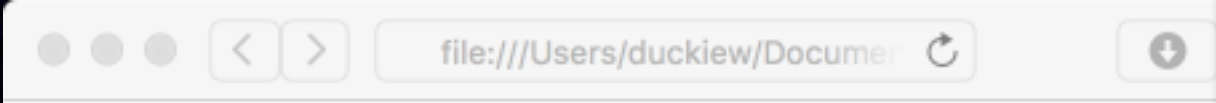
- ▶ Die einzelnen Elemente werden auch einzeln geladen...

```
page.html
!DOCTYPE html>
<html>
<head>
<link rel="stylesheet" href="style.css">
</head>
<body>

<h1>Ein Beispieldokument</h1>
<div></div>
<p>Ein sehr einfaches Beispiel einer simplen Webseite mit einem
eingebundenen Bild. Es demonstriert die Komposition aus
verschiedenen Elementen.</p>
<p>Auch wenn die Webpage
in Wirklichkeit aus vers

</body>
</html>
```

```
style.css
body {
  background-color: #ffffff;
}
h1 {
  color: #a00000;
  font-family: Helvetica, Arial, sans-serif;
  font-size: 24px;
}
p {
  color: #000000;
  font-family: Helvetica, Arial, sans-serif;
  font-size: 15px;
}
div {
  float: right;
  width: 118px;
  margin: 0 0 0 0;
  padding: 1px;
  text-align: center;
}
```



Ein Beispieldokument

Ein sehr einfaches Beispiel einer simplen Webseite mit einem eingebundenen Bild. Es demonstriert die Komposition aus verschiedenen Elementen.

Auch wenn die Webpage aussieht wie ein Dokument besteht es in Wirklichkeit aus verschiedenen einzeln geladenen Komponenten.



WEB BASICS

- ▶ Die einzelnen Elemente werden auch einzeln geladen...

```
page.html
!DOCTYPE html>
<html>
<head>
  <link rel="stylesheet" href="style.css">
</head>
<body>
  <h1>Ein Beispieldokument</h1>
  <div></div>
  <p>Ein sehr einfaches Beispiel einer simplen Webseite mit einem eingebundenen Bild. Es demonstriert die Komposition aus verschiedenen Elementen.</p>
  <p>Auch wenn die Webpage aussieht wie ein Dokument besteht es in Wirklichkeit aus verschiedenen einzeln geladenen Komponenten.</p>
</body>
</html>
```



```
file:///Users/duckiew/Doc
```

Ein Beispieldokument

Ein sehr einfaches Beispiel einer simplen Webseite mit einem eingebundenen Bild. Es demonstriert die Komposition aus verschiedenen Elementen.

Auch wenn die Webpage aussieht wie ein Dokument besteht es in Wirklichkeit aus verschiedenen einzeln geladenen Komponenten.



```
style.css
body {
  background-color: #ffffff;
}
h1 {
  color: #a00000;
  font-family: Helvetica, Arial, sans-serif;
  font-size: 24px;
}
p {
  color: #000000;
  font-family: Helvetica, Arial, sans-serif;
  font-size: 15px;
}
div {
  float: right;
  width: 118px;
  margin: 0 0 0 0;
  padding: 1px;
  text-align: center;
}
```




■ WEB TRACKING

TECHNIK HINTER DEN KULISSEN

LOGFILE ANALYSE

- ▶ Logfile des Webservers protokolliert meist:
 - ▶ Zeitpunkt eines Zugriffs
 - ▶ Object (URL)
 - ▶ Quelle (IP)
 - ▶ Informationen über Zugriff
 - ▶ Browser inkl. Version
 - ▶ „Referer“

```
8.208.150.117 - - [04/Nov/2016:07:24:47 +0100] "GET /roundmail/ HTTP/1.1" 200 140 "-" "Mozilla/4.0 (compatible; MSIE 8.0; Windows NT 6.1; Trident/4.0; .NET CLR 2.0.50727; .NET CLR 3.5.30729; .NET CLR 3.0.30729; Media Center EUIA.5.0.60020.05000; InfoPath.3; .NET CLR 1.1.4322.5722)"
8.208.150.117 - - [04/Nov/2016:07:24:48 +0100] "GET /roundmail/ HTTP/1.1" 200 140 "-" "Mozilla/4.0 (compatible; MSIE 8.0; Windows NT 6.1; Trident/4.0; .NET CLR 2.0.50727; .NET CLR 3.5.30729; .NET CLR 3.0.30729; Media Center EUIA.5.0.60020.05000; InfoPath.3; .NET CLR 1.1.4322.5722)"
82.47.3 - - [04/Nov/2016:07:46:11 +0100] "GET / HTTP/1.1" 200 140 "-" "-"
82.47.3 - - [04/Nov/2016:07:49:38 +0100] "-" 400 0 "-" "-"
82.47.3 - - [04/Nov/2016:07:51:09 +0100] "-" 400 0 "-" "-"
211.99.1 - - [04/Nov/2016:08:16:01 +0100] "GET /robots.txt HTTP/1.1" 404 0 "-" "Needle/1.0.0 (Node.js v5.11.0; linux x64)"
211.99.1 - - [04/Nov/2016:08:16:16 +0100] "GET / HTTP/1.1" 302 5 "-" "Hare/1.3 (Googlebot/2.1; +https://hypefactors.com/webcrawler)"
8.210.216.140 - - [04/Nov/2016:08:19:11 +0100] "GET /about/duckiew/ HTTP/1.1" 200 160 "-" "Mozilla/4.0 (compatible; MSIE 6.0; Windows NT 5.1)"
1.212.122.64 - - [04/Nov/2016:09:20:11 +0100] "-" 400 0 "-" "-"
249.73.183 - - [04/Nov/2016:09:46:59 +0100] "GET /robots.txt HTTP/1.1" 404 0 "-" "Mozilla/5.0 (compatible; Googlebot/2.1; +http://www.google.com/bot.html)"
249.73.183 - - [04/Nov/2016:09:46:59 +0100] "GET / HTTP/1.1" 200 129 "-" "Mozilla/5.0 (compatible; Googlebot/2.1; +http://www.google.com/bot.html)"
249.65.106 - - [04/Nov/2016:10:17:46 +0100] "GET /robots.txt HTTP/1.1" 404 0 "-" "Mozilla/5.0 (compatible; Googlebot/2.1; +http://www.google.com/bot.html)"
249.65.106 - - [04/Nov/2016:10:17:46 +0100] "GET /blog/index.php?archiv=knowing-what-to-say-is-not-good-enough-reason-not-to-start...html HTTP/1.1" 200 3808 "-" "Mozilla/5.0 (compatible; Googlebot/2.1; +http://www.google.com/bot.html)"
6.39.60.180 - - [04/Nov/2016:10:20:24 +0100] "-" 400 0 "-" "-"
6.39.60.180 - - [04/Nov/2016:10:20:34 +0100] "GET / HTTP/1.0" 200 140 "-" "-"
6.39.60.180 - - [04/Nov/2016:10:20:41 +0100] "GET / HTTP/1.0" 400 270 "-" "-"
6.39.60.180 - - [04/Nov/2016:10:20:41 +0100] "-" 400 0 "-" "-"
6.39.60.180 - - [04/Nov/2016:10:20:41 +0100] "-" 400 0 "-" "-"
6.39.60.180 - - [04/Nov/2016:10:20:41 +0100] "GET / HTTP/1.1" 400 270 "-" "-"
6.39.60.180 - - [04/Nov/2016:10:20:45 +0100] "GET / HTTP/1.1" 200 140 "-" "Mozilla/4.0 (compatible; MSIE 7.0; Windows NT 5.1; 360SE)"
6.39.60.180 - - [04/Nov/2016:10:20:46 +0100] "GET / HTTP/1.1" 200 140 "-" "Mozilla/4.0 (compatible; MSIE 6.0; ) Opera/UCWEB7.0.2.37/28/999"
212.203.188 - - [04/Nov/2016:10:27:25 +0100] "GET /robots.txt HTTP/1.1" 404 0 "-" "Needle/1.0.0 (Node.js v5.11.0; linux x64)"
249.65.98 - - [04/Nov/2016:11:27:43 +0100] "GET /blog/index.php?archiv=knowing-what-to-say-is-not-good-enough-reason-not-to-start...html HTTP/1.1" 200 1661 "-" "Mozilla/5.0 (compatible; Googlebot/2.1; +http://www.google.com/bot.html)"
212.203.188 - - [04/Nov/2016:12:14:05 +0100] "GET /robots.txt HTTP/1.1" 404 0 "-" "Needle/1.0.0 (Node.js v5.11.0; linux x64)"
212.203.188 - - [04/Nov/2016:12:14:27 +0100] "GET / HTTP/1.1" 302 5 "-" "Hare/1.3 (Googlebot/2.1; +https://hypefactors.com/webcrawler)"
# cat /var/log/nginx#
```

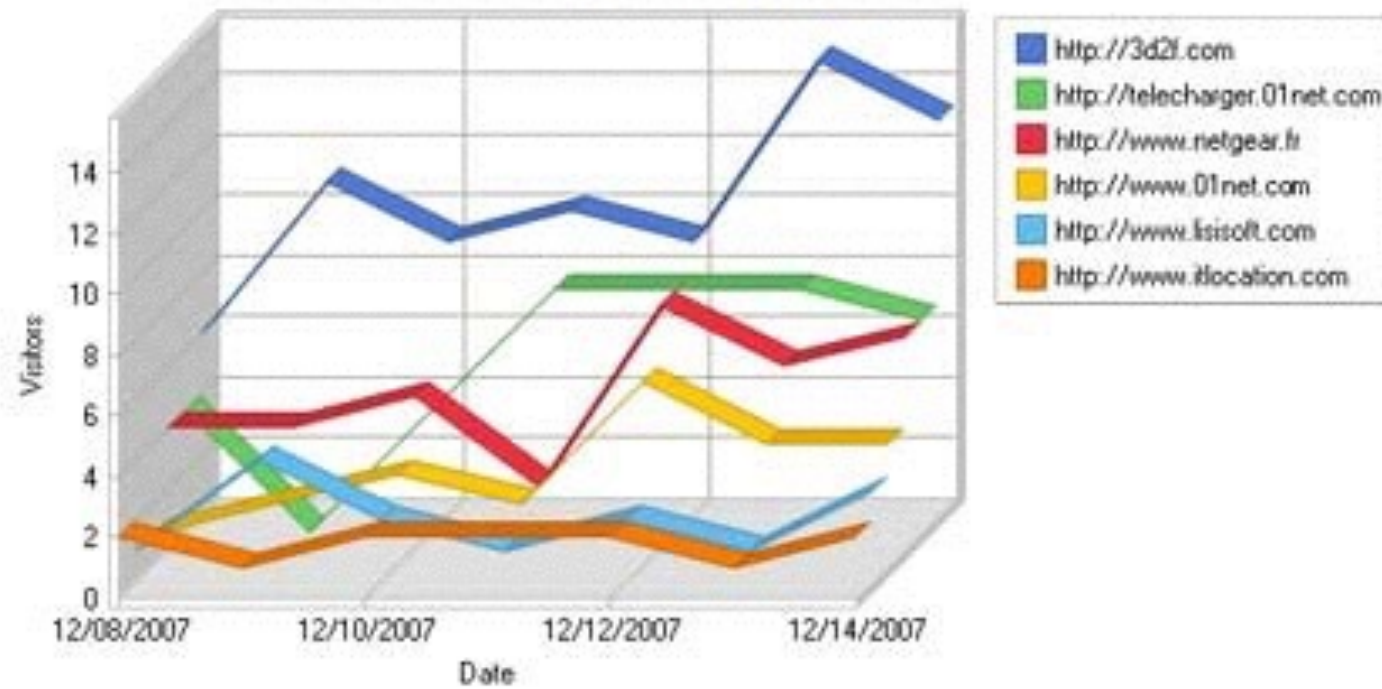

- [-] General Statistics
- [-] Activity Statistics
- [-] Access Statistics
- [-] Visitors
- [-] Referrers
 - [-] Referring Sites
 - [-] Referring URLs
 - [-] Search Engines
 - [-] Engines and Phrases
 - [-] Search Phrases
 - [-] Search Keywords
 - [-] Entry Pages / Engines
 - [-] Entry Pages / Phrases
- [-] Browsers
- [-] Errors
- [-] Tracked Files

Time range: 12/08/2007 00:00:41 - 12/14/2007 23:59:56

Generated on Fri Mar 14, 2008 - 17:36:33

Referring Sites

Daily Referring Sites



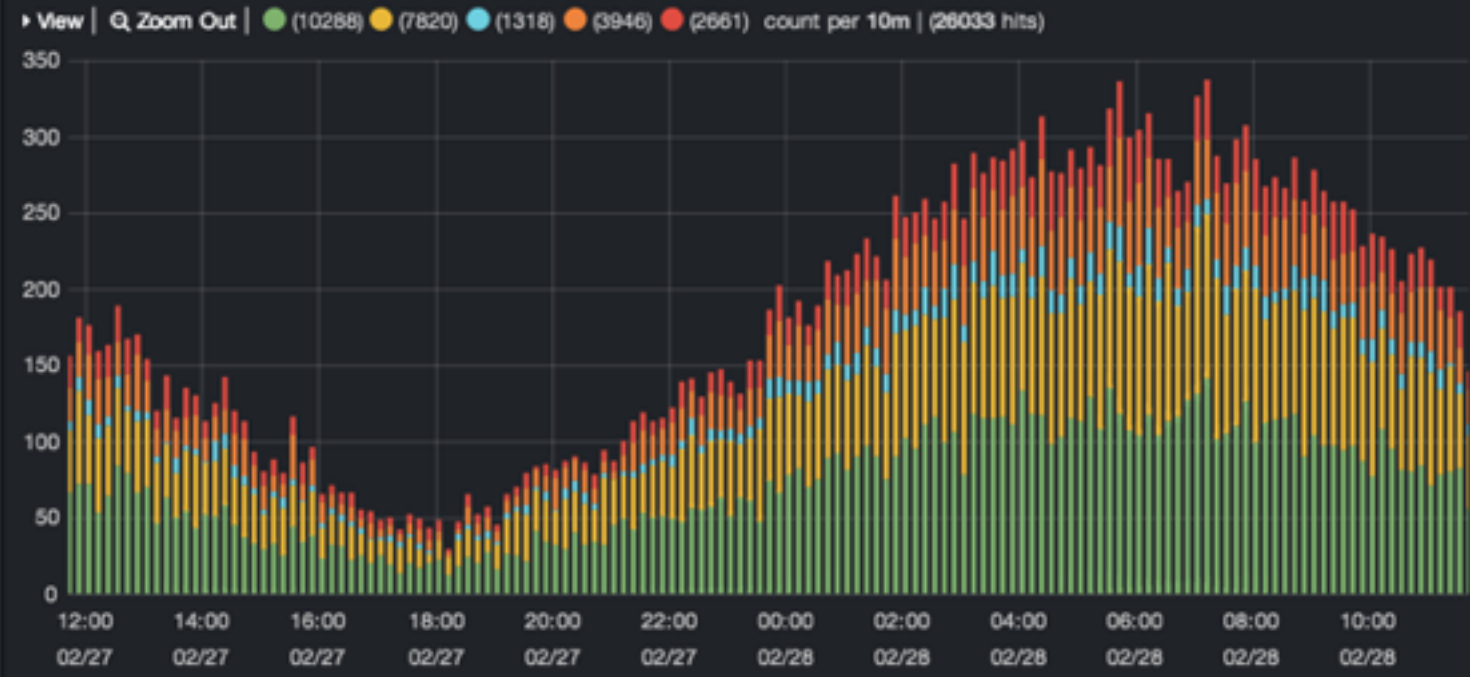
Top Referring Sites



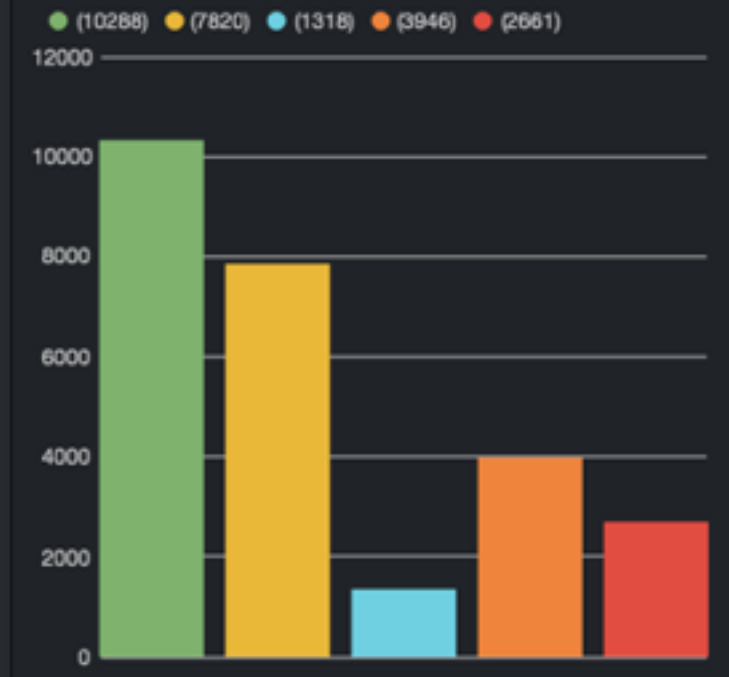
html php css png gif

QUERY FILTERING

EVENTS OVER TIME



QUERY HITS



Fields

- _id
- _index
- _type
- @message
- @timestamp
- agent
- bytes
- clientip
- coordinates
- country
- extension
- id

0 to 100 of 500 available for paging

@timestamp	@message
2014-02-28T17:48:30Z	131.82.81.46 - - [28/Feb/2014:17:48:30 +0000] "GET /HEHK.gif HTTP/1.1" 200 1077 "-" Mozilla/5.0 (X11; Linux x86_64; rv:6.0a1) Gecko/20110421 Firefox/...
2014-02-28T17:48:20Z	96.175.158.187 - - [28/Feb/2014:17:48:20 +0000] "GET /KEHl.php HTTP/1.1" 200 1111 "-" Mozilla/5.0 (X11; Linux x86_64; rv:6.0a1) Gecko/20110421 Firefo...
2014-02-28T17:48:20Z	170.59.70.27 - - [28/Feb/2014:17:48:20 +0000] "GET /KEHK.css HTTP/1.1" 200 9104 "-" Mozilla/5.0 (X11; Linux x86_64; rv:6.0a1) Gecko/20110421 Firefox/...
2014-02-28T17:48:18Z	4.129.143.176 - - [28/Feb/2014:17:48:18 +0000] "GET /HHEE.html HTTP/1.1" 404 0 "-" Mozilla/5.0 (X11; Linux i686) AppleWebKit/534.24 (KHTML, like Geck...
2014-02-28T17:48:18Z	1.204.129.229 - - [28/Feb/2014:17:48:18 +0000] "GET /EBKH.php HTTP/1.1" 200 9739 "-" Mozilla/5.0 (X11; Linux i686) AppleWebKit/534.24 (KHTML, like Ge...
2014-02-28T17:48:18Z	112.216.136.213 - - [28/Feb/2014:17:48:18 +0000] "GET /HHEE.css HTTP/1.1" 200 7657 "-" Mozilla/5.0 (X11; Linux i686) AppleWebKit/534.24 (KHTML, like Ge...

LOGFILE ANALYSE

AUF DEM SERVER

- ▶ Zugriffsverteilung nach Zeit
- ▶ Beliebteste Zeiten
- ▶ Zugriffe nach Quellort
- ▶ Statistik über Browser/Betriebssystem
- ▶ „Referer“ um verweisende Dienste zu finden
- ▶ Ermittlung von Aufenthaltsdauer von Nutzern

- ▶ Problem: keine wiederkehrenden Nutzer erkennbar



■ WEB TRACKING
MIT COOKIES

COOKIES IM BROWSER

- ▶ Cookies werden von jedem marktüblichen Browser unterstützt
- ▶ Server fordert den Browser auf, ab sofort bei jeder Anfrage ein Key-Value-Paar mitzuschicken
- ▶ Client wird dieses Key-Value-Paar (Cookie) von nun an mitschicken
 - ▶ Session-Cookies
 - ▶ Persistente Cookies
 - ▶ Cookies mit Ablaufdatum



Webinformationen - www.sueddeutsche.de

170 | 12 | 32 | 15

Suchen

Elemente | Netzwerk | Ressourcen | Timelines | Debugger | Speicher | Konsole

Kompletter Speicher | Cookies

Cookies

- dmp.theadex.com
- googleads.g.doubleclick.net
- platform.twitter.com
- staticxx.facebook.com
- tpc.googlesyndication.com
- www.facebook.com
- www.sueddeutsche.de**

Name	Wert	Domain	Gültig bis	...	Sicher
POPUPCHECK	1478322251263	www.sueddeutsche.de	5.11.2016, 06:04:11		
creid	1549532674538139785	.sueddeutsche.de	1.1.2038, 00:55:55	✓	
BIGipServerIb-pol-web_va...	3070893740.20480.0000	www.sueddeutsche.de	Sitzung		
BIGipServerIb-play-prod	3288997548.20480.0000	www.sueddeutsche.de	Sitzung		
BIGipServerIb-pay_http	1107959468.20480.0000	www.sueddeutsche.de	Sitzung		
axd	1000511471558360000	www.sueddeutsche.de	28.11.2016, 15:58:36		
adp_segs	e0	www.sueddeutsche.de	7.11.2016, 14:26:43		
_lp4_u	3rWO5qAwu8	.sueddeutsche.de	1.11.2017, 18:12:26		
_gat_UA-19474199-5	1	.sueddeutsche.de	4.11.2016, 14:35:16		
_ga	GA1.2.1697258674.1477749...	.sueddeutsche.de	4.11.2018, 14:26:43		
_dc_gtm_UA-19474199-5	1	.sueddeutsche.de	4.11.2016, 14:35:16		
_chartbeat2	.1477749518652.14782660...	www.sueddeutsche.de	4.12.2017, 14:26:44		
_cGtmS	1308023260.2	.sueddeutsche.de	4.11.2016, 14:56:42		
_cGtmAD	1	.sueddeutsche.de	4.11.2016, 14:56:44		
_cb_js	1	www.sueddeutsche.de	28.11.2017, 14:58:35		
_cb	B-uP_bCdGVkHCgX1qx	www.sueddeutsche.de	4.12.2017, 14:26:44		
__utmz	6611437.1477749515.1.1.ut...	.sueddeutsche.de	6.5.2017, 03:26:41		
__utmt	1	.sueddeutsche.de	4.11.2016, 14:35:16		
__utmz	6611437	.sueddeutsche.de	Sitzung		
__utmb	6611437.2.10.1478265916	.sueddeutsche.de	4.11.2016, 14:56:41		
__utma	6611437.1697258674.1477...	.sueddeutsche.de	4.11.2018, 14:26:41		
__gads	ID=6a8909226e97c12f:T=14...	.sueddeutsche.de	29.10.2018, 15:03:06		

Speicherliste filtern

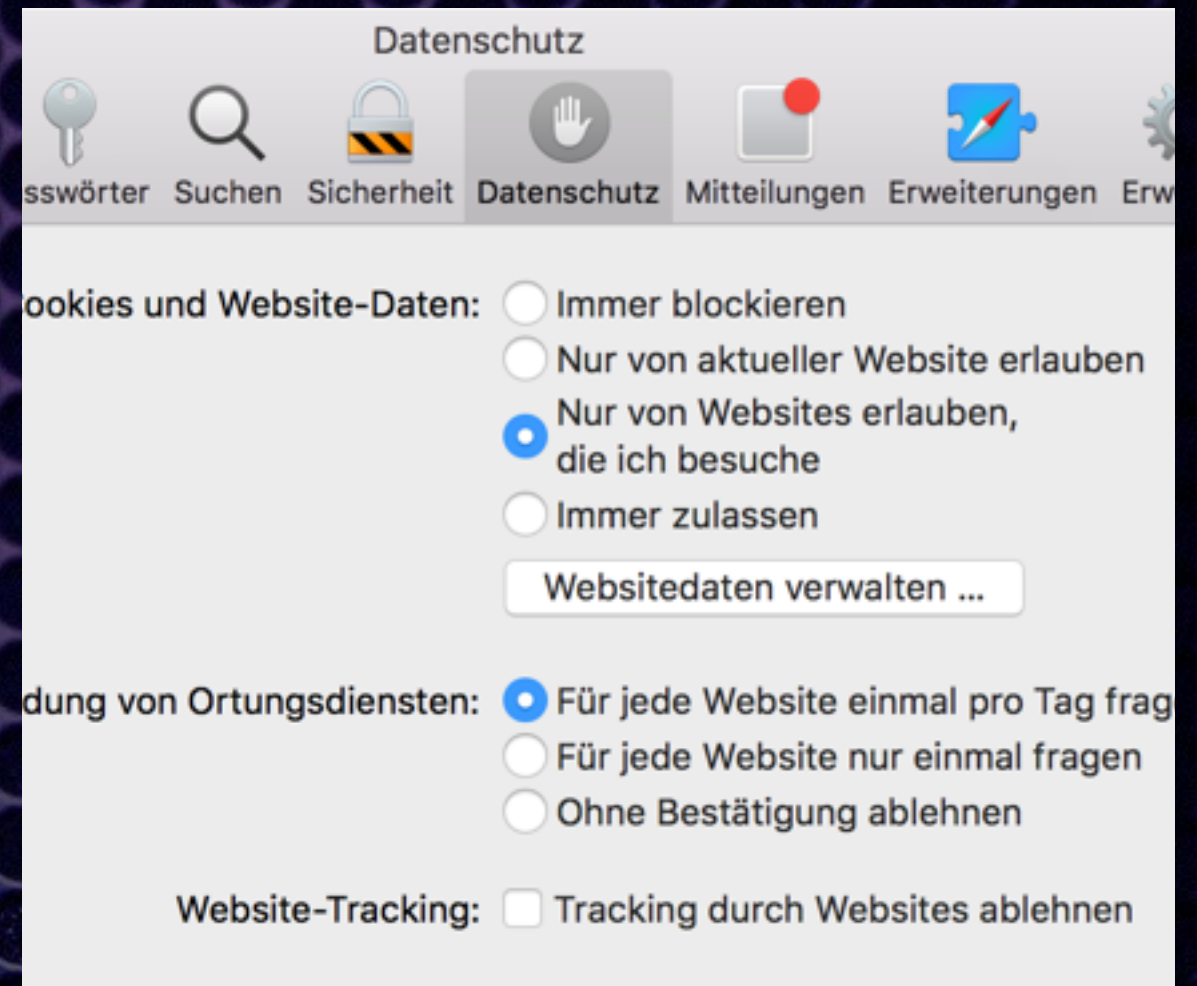
Haupt-Frame

COOKIES

NACH ZUGRIFF AUF SUEDEUTSCHE.DE

COOKIE LIMITIERUNGEN

- ▶ Nutzer können Cookies abschalten
- ▶ Nutzer können Cookies löschen
- ▶ Beschränkt auf eine Domain
 - ▶ Nutzer kann innerhalb einer Domain identifiziert werden, aber nicht auf verschiedenen Domains um ein globaleres Surfverhalten zu verfolgen



UMGEHUNG VON DOMAIN-EINSCHRÄNKUNG



toller-hopfen.de



mach-dich-blau.de



COOKIEdatabase

TrackID=u101; domain=toller-hopfen.de

TrackID=u220; domain=toller-mach-dich-blau.de

UMGEHUNG VON DOMAIN-EINSCHRÄNKUNG



toller-hopfen.de



mach-dich-blau.de



COOKIEdatabase

TrackID=u101; domain=toller-hopfen.de

TrackID=e0f12a; domain=evil-tracker

TrackID=u220; domain=mach-dich-blau.de

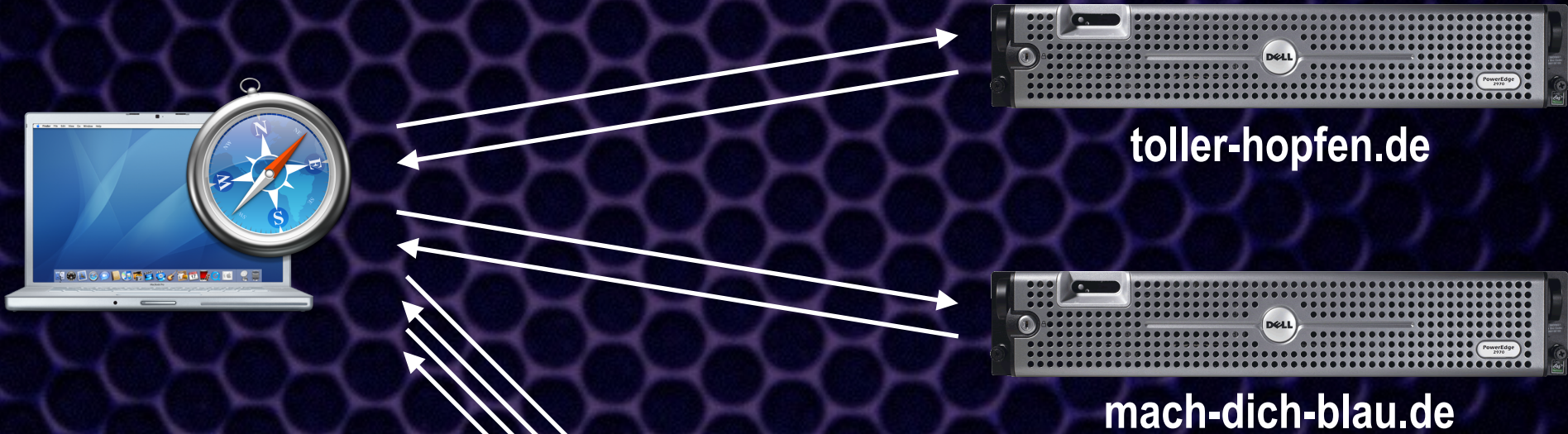
<http://evil-tracker/?ref=toller-hopfen&u=101>



evil-tracker

<http://evil-tracker/?ref=mach-blau&u=220>

UMGEHUNG VON DOMAIN-EINSCHRÄNKUNG



UID	Zeitpunkt	Affiliate	Resource
e0f12a	20.12.16 08:03	toller-hopfen	Paulaner
e0f12a	20.12.16 08:04	toller-hopfen	Spaten
e0f12a	20.12.16 08:05	mach-blau	Löwenbräu
e0f12a	20.12.16 08:06	mach-blau	Tucher

<http://evil-tracker/?ref=toller-hopfen&u=101>



evil-tracker

<http://evil-tracker/?ref=mach-blau&u=220>

OHNE COOKIES

- ▶ keine Identifizierung von wiederkehrenden Nutzern möglich (?!?)
- ▶ weiterhin werden über Referer oder in URL übergebene Parameter Daten von/über die Quelle weitergegeben
- ▶ Für Nutzer Probleme mit einigen Shops/Social-Media-Sites, die kein alternatives Session-Keeping unterstützen



WERKZEUGE

- ▶ Plugins für Browser
 - ▶ Better Security
 - ▶ Ghostery
 - ▶ RequestPolicy
 - ▶ Close'n Forget
 - ▶ Privacy Badger (erkennt auch 3rd Party Canvas Fingerprinting!)



OHNE COOKIES

- ▶ Single-Sign-On?
- ▶ „share“ - Funktionalitäten
- ▶ Vertrauen in Plugins?
 - ▶ siehe „Web of Trust“-Plugin
- ▶ Ist eine Identifizierung ohne Cookies wirklich unmöglich?

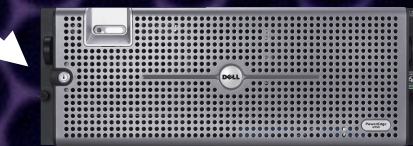




MARKETING

LIVE BIDDING PLATFORMS

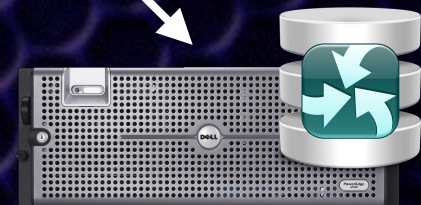
Werbeagentur



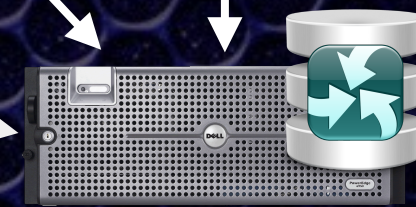
ATD
Agency Trading Desk



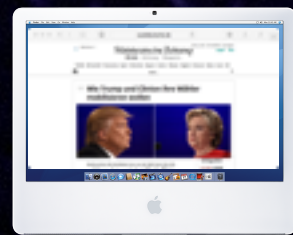
WebServer



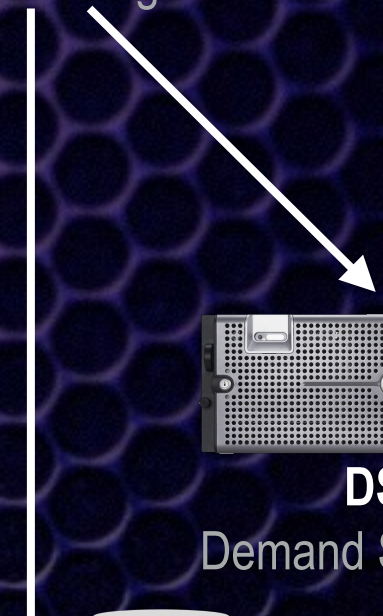
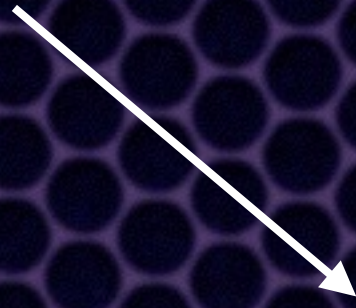
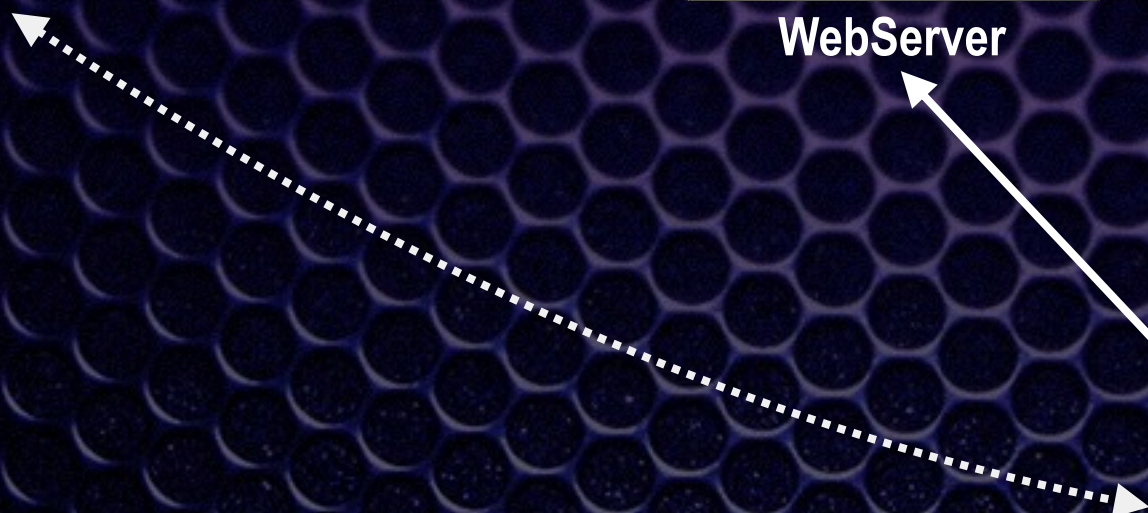
DSP
Demand Side Platform



DSP



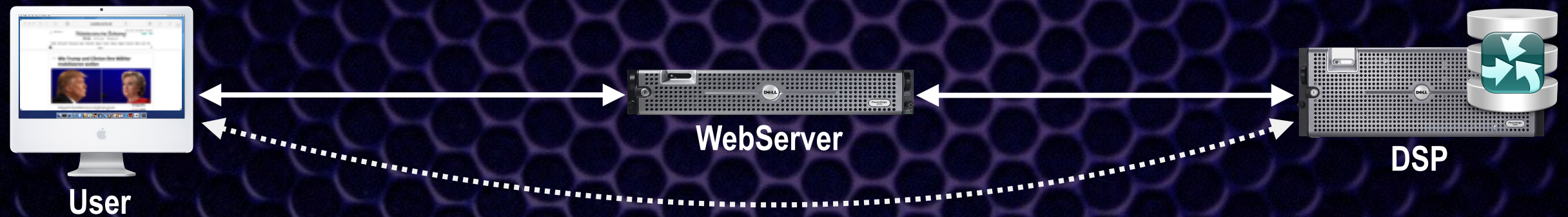
User



MARKETING

WERBUNG IM EINUNDZWANZIGSTEN JAHRHUNDERT

User surft auf Seiten über Pizza ...



**... bekommt Werbung für
Pizza-Lieferdienste angezeigt**

■ **MARKETING**

WERBUNG IM EINUNDZWANZIGSTEN JAHRHUNDERT

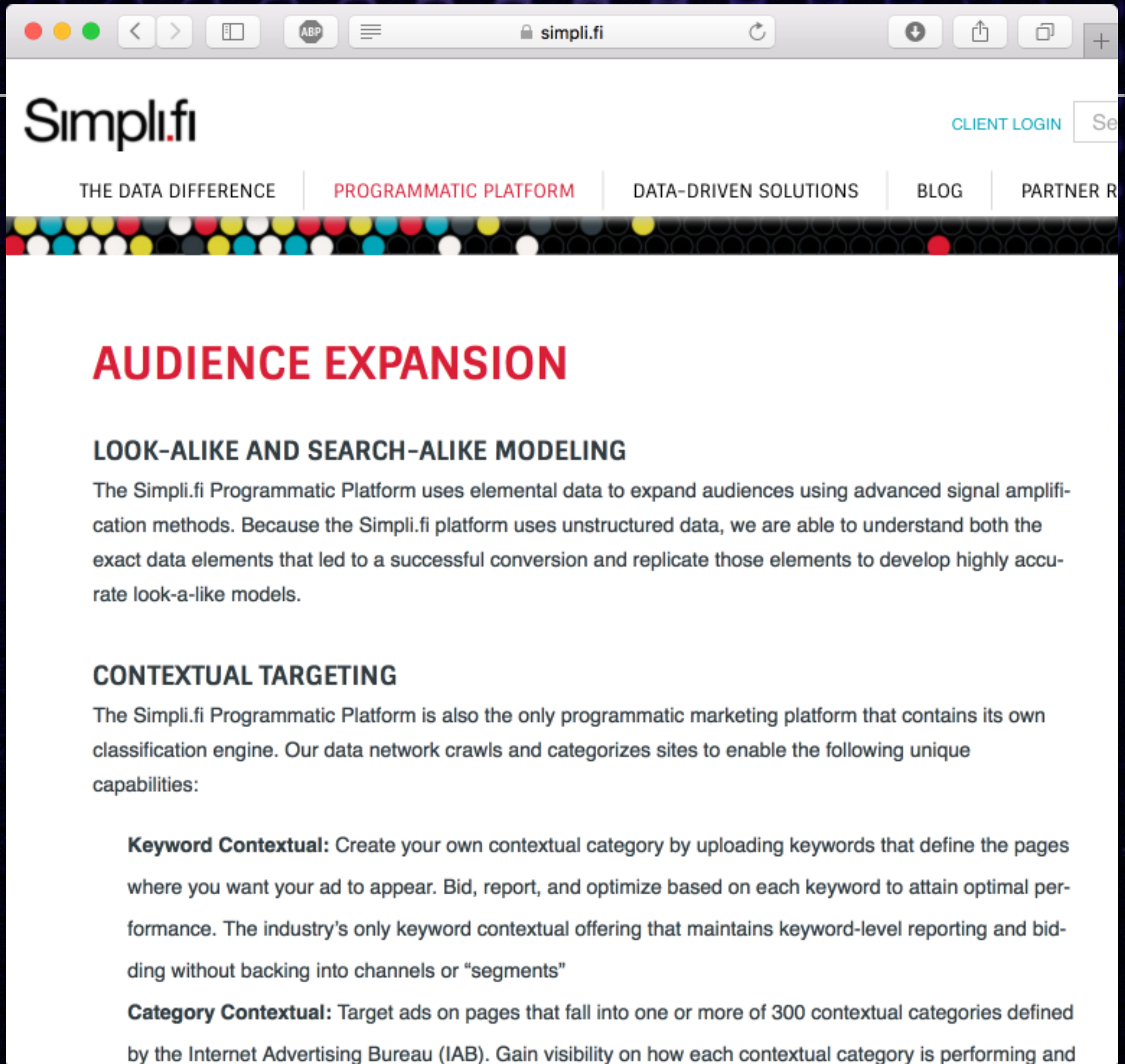
BEKANNTE DSP-PLATTFORMEN

- ▶ Facebook Exchange (FBX)
 - ▶ Facebook User-Interessen bereits bekannt über eigene Plattform
 - ▶ Live-Suchdaten über
 - ▶ simpli.fi
 - ▶ chango
 - ▶ und weitere...
 - ▶ Auch außerhalb der Plattform: Facebook's Audience Network
- ▶ Yahoo RightMedia
- ▶ Google DoubleClick



WERBENUTZUNG VON TRACKING-DATEN

- ▶ Behavioral (re)targeting
 - ▶ nutzt Surf-Verhalten der Vergangenheit (History) und bekannte Interessen für Werbung
 - ▶ weist den Nutzer oft auf bereits bekanntes hin und kann dadurch Marken-Wahrnehmung erhöhen
- ▶ Search retargeting
 - ▶ nutzt aktuelle Suchen um Werbung für neue Interessen anzuzeigen
 - ▶ zielt auf Kaufinteressen ab, die potentiell noch nicht befriedigt wurden
 - ▶ kann den Nutzer potentiell verstören



The image is a screenshot of a web browser displaying the Simpli.fi website. The browser's address bar shows 'simpli.fi'. The website's header includes the Simpli.fi logo, a 'CLIENT LOGIN' button, and a search bar. A navigation menu below the header lists 'THE DATA DIFFERENCE', 'PROGRAMMATIC PLATFORM' (highlighted in red), 'DATA-DRIVEN SOLUTIONS', 'BLOG', and 'PARTNER R'. The main content area features a decorative border of colorful dots. The primary heading is 'AUDIENCE EXPANSION' in large red letters. Below it, the sub-heading 'LOOK-ALIKE AND SEARCH-ALIKE MODELING' is followed by a paragraph explaining the platform's use of elemental data for audience expansion. The next sub-heading is 'CONTEXTUAL TARGETING', followed by a paragraph describing the platform's classification engine. Two specific capabilities are detailed: 'Keyword Contextual' and 'Category Contextual'.

AUDIENCE EXPANSION

LOOK-ALIKE AND SEARCH-ALIKE MODELING

The Simpli.fi Programmatic Platform uses elemental data to expand audiences using advanced signal amplification methods. Because the Simpli.fi platform uses unstructured data, we are able to understand both the exact data elements that led to a successful conversion and replicate those elements to develop highly accurate look-a-like models.

CONTEXTUAL TARGETING

The Simpli.fi Programmatic Platform is also the only programmatic marketing platform that contains its own classification engine. Our data network crawls and categorizes sites to enable the following unique capabilities:

- Keyword Contextual:** Create your own contextual category by uploading keywords that define the pages where you want your ad to appear. Bid, report, and optimize based on each keyword to attain optimal performance. The industry's only keyword contextual offering that maintains keyword-level reporting and bidding without backing into channels or "segments"
- Category Contextual:** Target ads on pages that fall into one or more of 300 contextual categories defined by the Internet Advertising Bureau (IAB). Gain visibility on how each contextual category is performing and



■ WEB TRACKING

TECHNIK HINTER DEN KULISSEN

TECHNIK

- ▶ es geht auch ohne Cookies...



SPYPLUGINS UND WERBEPLUGINS

- ▶ Awesome Screenshot (1,3 Mio Nutzer)
- ▶ Hover Zoom (1 Mio Nutzer)
- ▶ CrxMouse
- ▶ Power Zoom
- ▶ Eat my cookies
- ▶ Orbvious interes
- ▶ Autocopy original152
- ▶ Fabulous186 (500,000 Nutzer)
- ▶ Alle Plugins von wips.com (tracking is in the code, not always enabled)
- ▶ Youtube Ratings Preview (Author removed the tracking...)
- ▶ BBC News Reader (kein offizielles Plugin)
- ▶ Autocopy
- ▶ Iminent (1 Mio Nutzer)
- ▶ Bookmark Sentry
- ▶ Allow Right Click
- ▶ Clickable links
- ▶ Reload all tabs
- ▶ ruul screen ruler
- ▶ Clickonce for chrome
- ▶ Mail checker plus for google mail
- ▶ Send using Gmail™ (no button)
- ▶ Troll emoticons
- ▶ AwesomeNewTabPage
- ▶ Sexy undo Close Tab
- ▶ 40-50 small extensions by brandthunder.com
- ▶ yet another google bookmarks
- ▶ ChromeReload
- ▶ HoverZoom (1 Mio users)
- ▶ Neat Bookmarks
- ▶ ScrollToTopButton
- ▶ SmoothGestures
- ▶ Smooth Scroll
- ▶ Translate Selection
- ▶ Webpage Screenshot Capture
- ▶ Window Resizer
- ▶ GetThemAll Downloader
- ▶ Facebook Secret Emoticons
- ▶ Add to Feedly (bereits entfernt)
- ▶ Tweet This Page (bereits entfernt)

PLUGINS



- ▶ Flash „Local Shared Objects“ (LSO)
 - ▶ bis 100kb Daten
 - ▶ schwer vom Nutzer zu entfernen
 - ▶ von allen Browsern geteilt
 - ▶ LocalConnection object
 - ▶ Verbindung zwischen allen laufenden Flash Objekten
 - ▶ ermöglicht Datenaustausch zwischen Fenstern, auch Fenstern im „Private Browsing“-Modus
- ▶ Microsoft Silverlight Isolated Storage
 - ▶ bis 100kb Daten
 - ▶ von allen Instanzen geteilt

HTML5



- ▶ HTML5 Local Storage (Ersatz für Cookies)
 - ▶ Key-Value-Pairs (wie Cookies)
 - ▶ bis 5 MB Daten
 - ▶ keine Ablaufdatum
 - ▶ werden von Fenstern geteilt, aber nicht zwischen http und https, ebenso sind „anonymous browsing“-Fenster separat
 - ▶ In den meisten Browsern führt das Löschen von Cookies auch zum Entfernen des HTML5 Local Storage.
- ▶ HTML5 IndexedDB
 - ▶ Indizierte Objekt-Datenbank für große Datenmengen

INTERNET EXPLORER



- ▶ Internet Explorer userData storage
 - ▶ proprietäre Datenablage
 - ▶ seit IE5
 - ▶ seit IE7 deprecated, aber auch in IE11 nicht unterstützt
 - ▶ bis zu 64kB Daten
 - ▶ XML-Format

SSL SESSION



SSL/TLS

- ▶ SSL/TLS verwendet Session IDs um Session Settings und Keys längerfristig wiederverwenden zu können um bei erneutem Verbindungsaufbau Zeit/Resourcen zu sparen
 - ▶ Session ID wird vom Server generiert und dem Client angeboten um zu einem späteren Zeitpunkt die Session „fortzusetzen“
 - ▶ Browser ohne Session-ID werden mehr CPU für https-Surfen verbrauchen
 - ▶ Einige Browser löschen Session-IDs beim Restart

BROWSER CACHE



- ▶ Browser dazu bringen eine Nutzer-ID im Cache abzulegen und bei zukünftigen Zugriffen mitzuschicken
 - ▶ Hinterlegen eines Objekts mit eindeutiger ID des Nutzers (z.B. verstecken in `<div>`-Tag) und Caching für sehr lange Zeit erlauben
 - ▶ Request an Tracking-Server mit eindeutiger ID in URL
 - ▶ HTML-Frame im Cache referenzieren, welches ein kleines Bild mit eindeutiger URL nachlädt
 - ▶ Javascript das aus anderem (gecachtem) Dokument ID aus `<div>`-Tag auslesen

BROWSER CACHE UND ETAG



- ▶ ETag mit HTTP/1.1 eingeführt
- ▶ Dient zum Erkennen von Änderungen an vom Client bereits abgefragten Ressourcen und ist gedacht um Änderungen an Dokumenten im Cache zu erkennen
 - ▶ Server generiert eine ID für ein vom Client angefragtes Object
 - ▶ Client trägt in Zukunft bei Anfragen das ETag mit in den Request ein, damit der Server auf Änderung prüfen und gegebenenfalls mit „304 Not Modified“ antworten kann
- ▶ Tracking-Server können den ETag als Nutzer-ID missbrauchen, indem sie jeder Anfrage ohne ETag eine neue vom Status des Objekts unabhängige ID vergeben

WEITERE CACHE TRICKS



- ▶ Redirect-Caches
 - ▶ Server sendet 301 Document Moved
- ▶ Authentication-Cache
 - ▶ Java-Script erzeugt „fake“ Login, die authentication-Information wird dann vom Browser gespeichert
- ▶ HTTP Strict Transport Security
 - ▶ „Pinning“ von HTTPS-needed Information
 - ▶ fake Subdomains die mit HSTS „gesichert werden“, allerdings muss der Status mit Java-Script über brute-Force getestet werden

EVERCOOKIES

- ▶ „respawning“ von cookies
 - ▶ wenn Nutzer seine regulären Cookies löscht
 - ▶ Wiederherstellung des Cookies aus anderen Quellen
 - ▶ Flash cookies
 - ▶ localStorage
 - ▶ sessionStorage
 - ▶ ETags



■ BROWSER

FORENSICS/FINGERPRINTING

FINGERPRINTING

- ▶ Identifikation unabhängig von Cookies oder Cookie-artigen Identifizierung-Token, ohne eine unbedingt nötige zentrale Instanz
- ▶ Erkennung von einzelnen Merkmalen des Benutzers und Erstellung eines Identifiers durch Kombination aller Merkmale zu einem Gesamtbild.
 - ▶ Teilweise auch ohne aktiviertes Java-Script möglich
 - ▶ Teilweise auch durch passives Monitoring des Traffics möglich



FORENSICS/FINGERPRINTING

- ▶ Passiv:
 - ▶ IP Adresse
 - ▶ Provider
 - ▶ Standort
 - ▶ Betriebssystem
 - ▶ Browser (Typ und Version)
 - ▶ Sprache
 - ▶ HTTP-Accept Header

FORENSICS/FINGERPRINTING

- ▶ Aktiv:
 - ▶ CPU-Typ
 - ▶ Zeitzone
 - ▶ Zeit-Ungenauigkeit
 - ▶ Display-Settings
 - ▶ Installierte Schriftarten
 - ▶ Installierte Plugins
 - ▶ Unterstützte MIME-Typen
 - ▶ Einstellungen für Cookies (on/off/thirdparty)
 - ▶ Battery-Status

**[CANVAS FINGERPRINTING] IS
PRESENT ON MORE THAN 5.5 % OF
TOP 100 000 ALEXA WEBSITES**

N. Nikiforakis, A. Kapravelos, W. Joosen, C. Kruegel, F. Piessens, and G. Vigna

**[CANVAS FINGERPRINTING] IS MORE
HEAVILY USED ON THE TOP SITES,
REACHING 2.5% OF THE TOP 1000**

Steven Englehardt, Arvind Narayanan

**WE FOUND WEBRTC BEING USED TO
DISCOVER LOCAL IP ADDRESSES
WITHOUT USER INTERACTION ON 715
SITES OUT OF THE TOP 1 MILLION**

Steven Englehardt, Arvind Narayanan



■ WEB TRACKING

CANVAS FINGERPRINTING

CANVAS FINGERPRINTING

- ▶ Verwendet API um „unsichtbare“ Grafik mithilfe des Browsers und potentiell dessen Grafikkarte zu zeichnen.
- ▶ Wird nicht von üblichen Anti-Tracker-Tools erkannt.
- ▶ Beispiel 1:
 - ▶ Rufe `measureText()` für 50 verschiedene Fonts auf und berechne die Zeichenbreite.
 - ▶ unterschiedliche Browser und Font-Definitionen liefern unterschiedliche Textlängen
 - ▶ Berechne Summe/Prüfsumme aller Ergebnisse.



■ CANVAS FINGERPRINTING

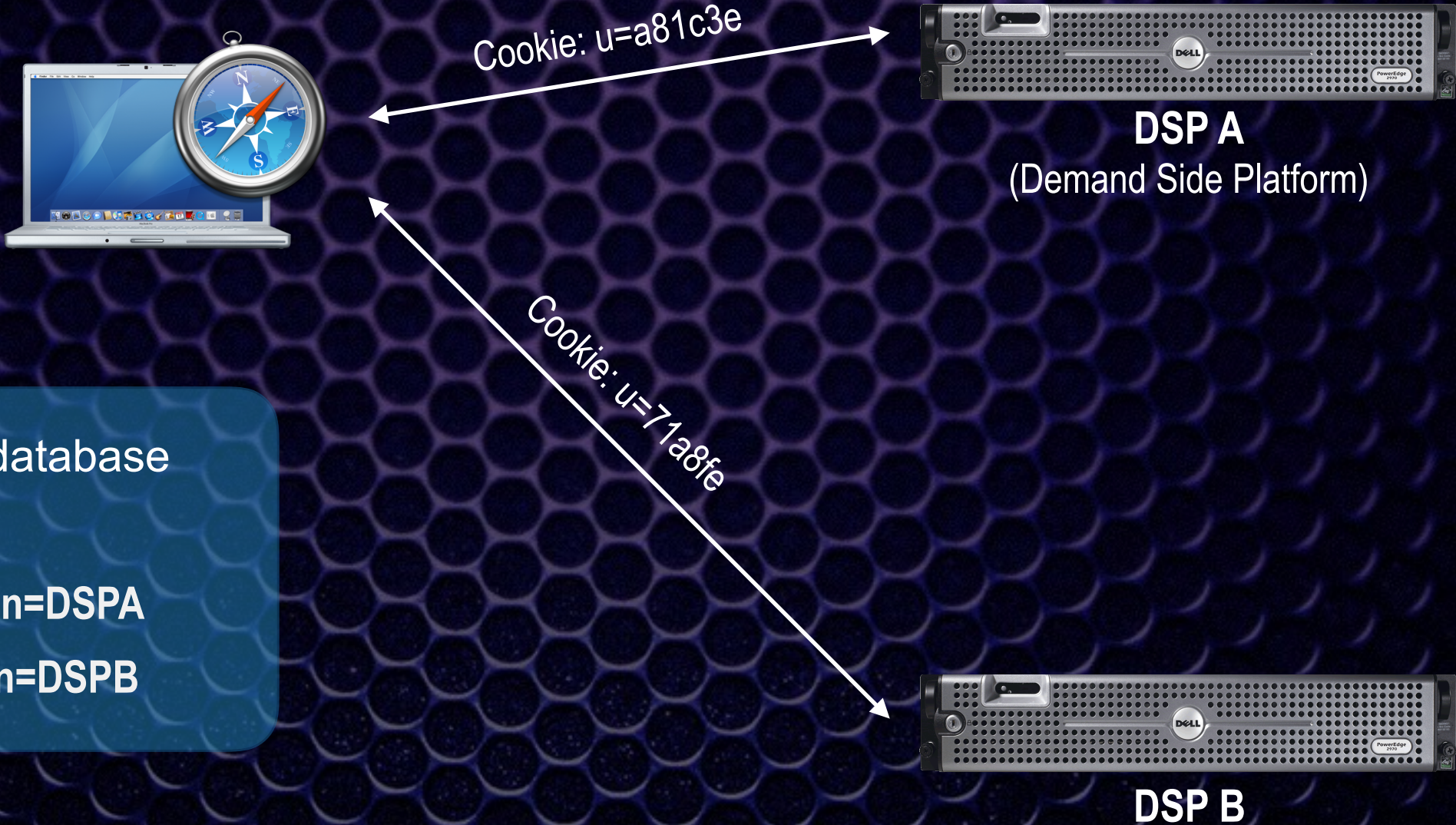
ANALYSE VON UNSICHTBAR BERECHNETEN GRAFIKEN/TEXTEN

Fingerprinting Script	Anzahl Sites	Gezeichneter Text
ct1.addthis.com/static/r07/core130.js	5282	Cwm fjordbank glyphs vext quiz, 😊
i.ligatus.com/script/fingerprint.min.js	115	http://valve.github.io
src.kitcode.net/fp2.js	68	http://valve.github.io
admicro1.vcmmedia.vn/fingerprint/figp.js	31	http://admicro.vn/
amazonaws.com/af-bdaz/bquery.js	26	Centillion
*.shorte.st/js/packed/smeadvert-intermediate-ad.js	14	http://valve.github.io
stat.ringier.cz/js/fingerprint.min.js	4	http://valve.github.io
cya2.net/js/STAT/89946.js	3	ABCDEFGHIJKLMUNROLPEQncRodSeTd UqV=W%FX0%Y9ZF%9 abcdefghijklmnopqrstuvUwTxFy8z0123f405 96f79898+3/
images.revtrax.com/RevTrax/js/fp/fp.min.jsp	3	http://valve.github.io
pof.com	2	http://www.plentyoffish.com
*.rackcdn.com/mongoose.fp.js	2	http://api.gonorthleads.com

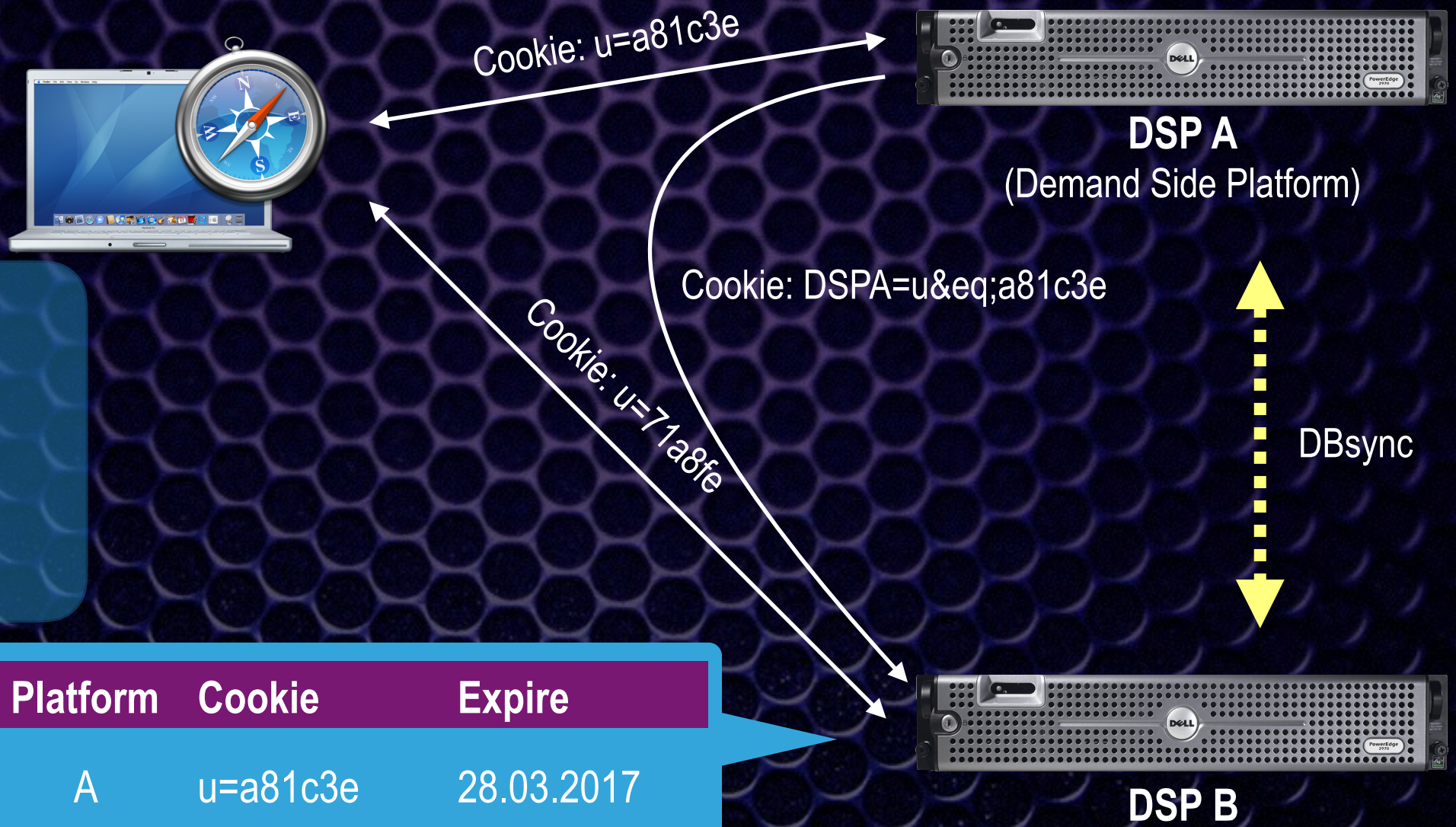
■ CANVAS FINGERPRINTING


SCRIPTS MIT TEXT UND VORKOMMEN IN ALEXA TOP 100K SITES

COOKIE TRADING/SYNCING



COOKIE TRADING/SYNCING



 **COOKIE**database

u=a81c3e; domain=DSPA
u=71a8fe; domain=DSPB

Platform	Cookie	Expire
A	u=a81c3e	28.03.2017
B	u=71a8fe	24.02.2018
C	u=6f71a2	01.12.2017



 FRAGEN?

BILDRECHTE

- ▶ „[Fingerprints](#)“ von Siri Handeland, 2011
Attribution 2.0 Generic (CC BY 2.0)