

Spurensuche in der virtuellen Welt „Das Leben des Norbert P.“

Logdaten-Konferenz – September 2009

Christian J. Dietrich
dietrich [at] internet-sicherheit . de

Marian Jungbauer
jungbauer [at] internet-sicherheit . de

Institut für Internet-Sicherheit
<https://www.internet-sicherheit.de>
FH Gelsenkirchen



Wie viele Daten-Skandale der letzten 2 Jahre basierten auf Logdaten?

•A: 1-9

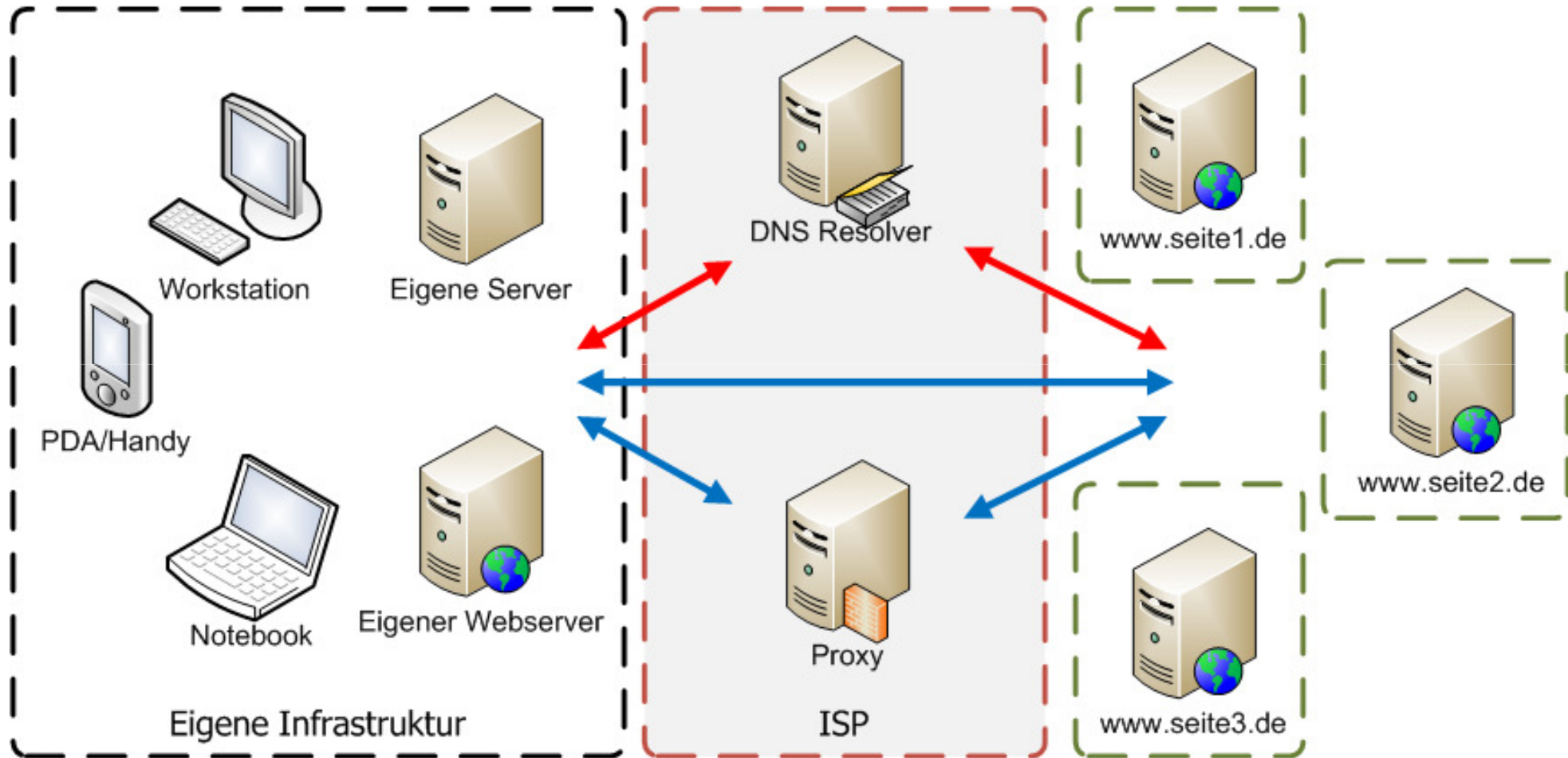
•B: 10-99

•C: 100-999

•D: mehr als 1000

- Einleitung
 - Modell zur Logdaten-Gewinnung
 - Log-Auswertung im Web
 - Abgrenzung Datenpanne vs. Logdaten-Affäre
- Eigene Datenspuren vermeiden
 - Private Browsing
 - Proxy
 - Anonymisierungsdienste
 - Eigener DNS-Resolver
- Fazit

Modell



↔ Namensauflösung (www.seite1.de -> 194.94.127.40)
↔ Web/HTTP-Datenverkehr

- 1. **DNS-Resolver** (Auflösung des Hostnames `www.seite1.de`, den der Benutzer in seinen Browser tippt)
- 2. Aufruf des entfernten Webservers `www.seite1.de` (Spuren im **Webserver-Log**)
- 3. `www.seite1.de` bindet **Elemente von `www.seite2.de` und `www.seite3.de`** ein, daher auch dort Spuren im Webserver-Log
- 4. Falls ein **Proxy-Server** benutzt wird, auch dort Spuren
- 5. Nicht zuletzt auch auf dem Client selbst Spuren (**im Cache und in Cookies**)
- 6. Zusätzlich können **aktiv weitere Daten** vom Client an den Server übertragen werden (JavaScript/Flash, Google Analytics)

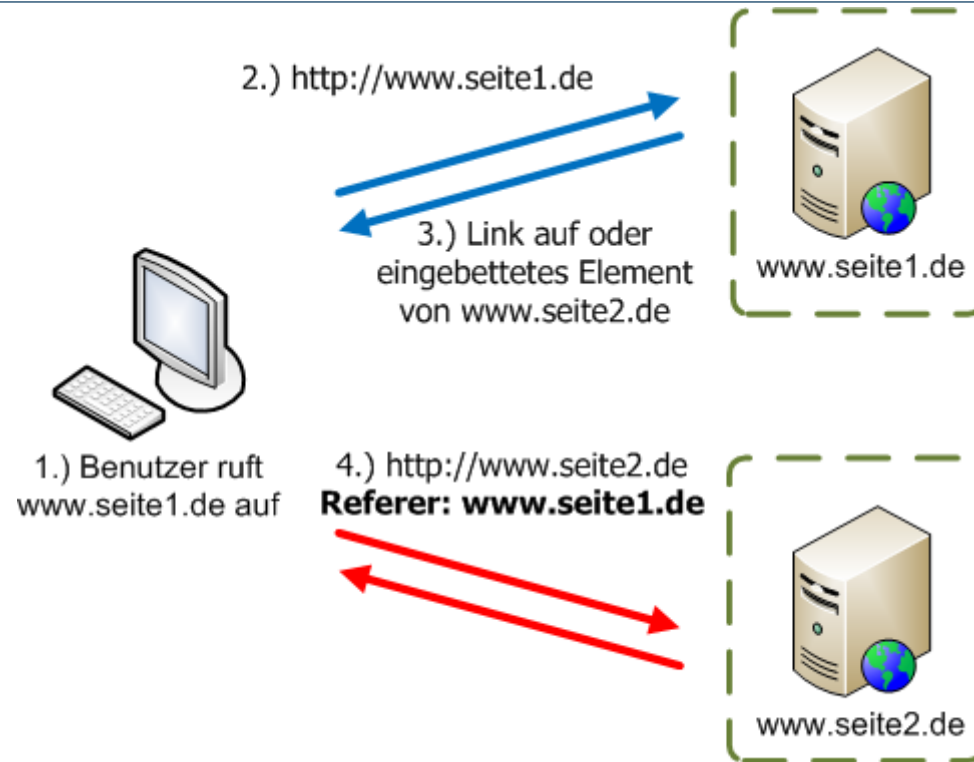
Typisches Webserver-Log (combined)

```
ntoo Firefox/3.0.13"
91.51.162.241 - - [12/Sep/2009:11:41:32 +0200] "GET /fileadmin/template/images/partner/ifis-live-ha
king-logo.jpg HTTP/1.1" 200 5899 "http://www.internet-sicherheit.de/forschung/aktuelle-forschungspr
jekte/internet-fruehwarnsysteme/" "Mozilla/5.0 (X11; U; Linux i686; de-DE; rv:1.9.0.13) Gecko/20090
2610 Gentoo Firefox/3.0.13"
91.51.162.241 - - [12/Sep/2009:11:41:32 +0200] "GET /fileadmin/template/images/partner/ifis-live-ha
```

- 9 Felder

- IP-Adresse **91.51.162.241** (p5B33A2F1.dip.t-dialin.net)
- RFC 1403 Identität (identd) - (unbekannt)
- Benutzername (HTTP Auth) - (unbekannt, keine Auth)
- Zeitstempel **12.Sep 2009 11:41:32 Uhr +2 MESZ**
- Anfrage des Client **“GET /fileadmin/.../ifis-live-hacking-logo.jpg HTTP/1.1“**
- Status Code **200** (bedeutet „Alles OK“)
- Dateigröße / Transfervolumen **5899** Bytes
- Referrer URL **http://www.internet-s...t.de/.../internet-fruehwarnsysteme/**
- User Agent **Mozilla/5.0 (X11;U;Linux i686; de-DE; rv:1.9.0.13) ...
Gentoo Firefox/3.0.13**

Referrer



- Referrer informiert die neu besuchte Seite über die vorher besuchte Seite
- www.seite2.de erfährt dadurch, dass www.seite1.de besucht wurde

```
1.2.3.4 - - [29/Sep/2009:11:41:32 +0200]  
„GET /seite2/ HTTP/1.1“ 200 2069 „http://www.seite1.de/“
```

- Typischerweise werden die Logdaten des Webserver ausgewertet
- Clients sind menschliche Besucher, daher besonders interessant
 - Wie viele Besucher hat die Web-Präsenz?
 - Wie lange verweilen die Benutzer?
 - Wie ist die Browser-Verteilung?
 - Welche Bildschirmauflösungen haben die Benutzer?
- Typische Werkzeuge
 - awstats/webalizer
 - Google Analytics

- Demo

- Demo

Abgrenzung Datenpanne vs. Logdaten-Affäre

Erneute Datenpanne bei eBay

17.05.2009 12:53

Bahn überprüft Gewerkschaftsbeiträge

Meldung vom 13.12.2008 10:08

Bericht: Datenleck bei Kreditkarten-Dienstleister

News-Meldung vom 04.05.2009 - 13:46

Strafanzeige gegen private Bahn-Ermittler

07.05.2008 09:51

Über 1,4 GByte an gestohlenen Daten auf Server gefunden

23.06.2008 12:52

 [« Vorige](#) | [Nächste »](#)

Fernsehmagazin: Datenpanne bei Einwohnermeldeämtern

26.11.2008 16:55


 [« Vorige](#) | [Nächste »](#)

Bericht: Telekom-Kundendaten mit Konto-Informationen im Umlauf

Meldung vom 27.05.2008 15:53


Daten von tausenden Studenten der Uni Magdeburg im Netz

17.09.2008 16:17

 [« Vorige](#) | [Nächste »](#)


Massive Datenpanne in norwegischer Steuerverwaltung

26.09.2008 10:37

 [« Vorige](#) | [Nächste »](#)

Britische Lehrerorganisation vermisst CD mit Daten von 11.000 Lehrern

06.10.2008 18:50

 [« Vorige](#) | [Nächste »](#)

Berlin: Gewerkschaft der Polizei räumt Datenpanne ein

News-Meldung vom 30.03.2009 - 14:13

Zeitung: Datenskandal bei Kabel Deutschland

News

Meldung vom 19.01.2008 12:01

Datenpannen beim britischen Militär und US-Kreditkartenunternehmen

 [« Vorige](#) | [Nächste »](#)

08.09.2008 11:19

 [« Vorige](#) | [Nächste »](#)

Erneut Sorgen um den Datenschutz im britischen Gesundheitssystem

Meldung vom 16.10.2008 08:53

Datenpanne auf Internetseite des Kinderkanals

22.01.2008 00:49


TELEPOLIS [« Vorige](#) | [Nächste »](#)

Die nächste Datenpanne beim britischen Militär

Meldung vom 01.09.2008 14:45

Beate Uhse: Tausende E-Mail-Adressen veröffentlicht

07.05.2008 09:51

 « [Vorige](#) | [Nächste](#) »

Über 1,4 GByte an gestohlenen Daten auf Server gefunden

Meldung vom 01.09.2008 14:45

Beate Uhse: Tausende E-Mail-Adressen veröffentlicht

- Über 1,4 GByte an gestohlenen Daten auf Server gefunden

Daten von normalen Websurfern ebenso wie von Firmen, namhaften Organisationen und Dienstleistern aus dem Gesundheitssektor. [...] Die Dateien enthalten [...] kompromittierte Patientendaten, Daten von Bankkunden, geschäftliche E-Mails sowie Outlook-Konten mitsamt der E-Mail-Kommunikation. [...]

<http://www.heise.de/newsticker/Ueber-1-4-GByte-an-gestohlenen-Daten-auf-Server-gefunden--/meldung/107517>

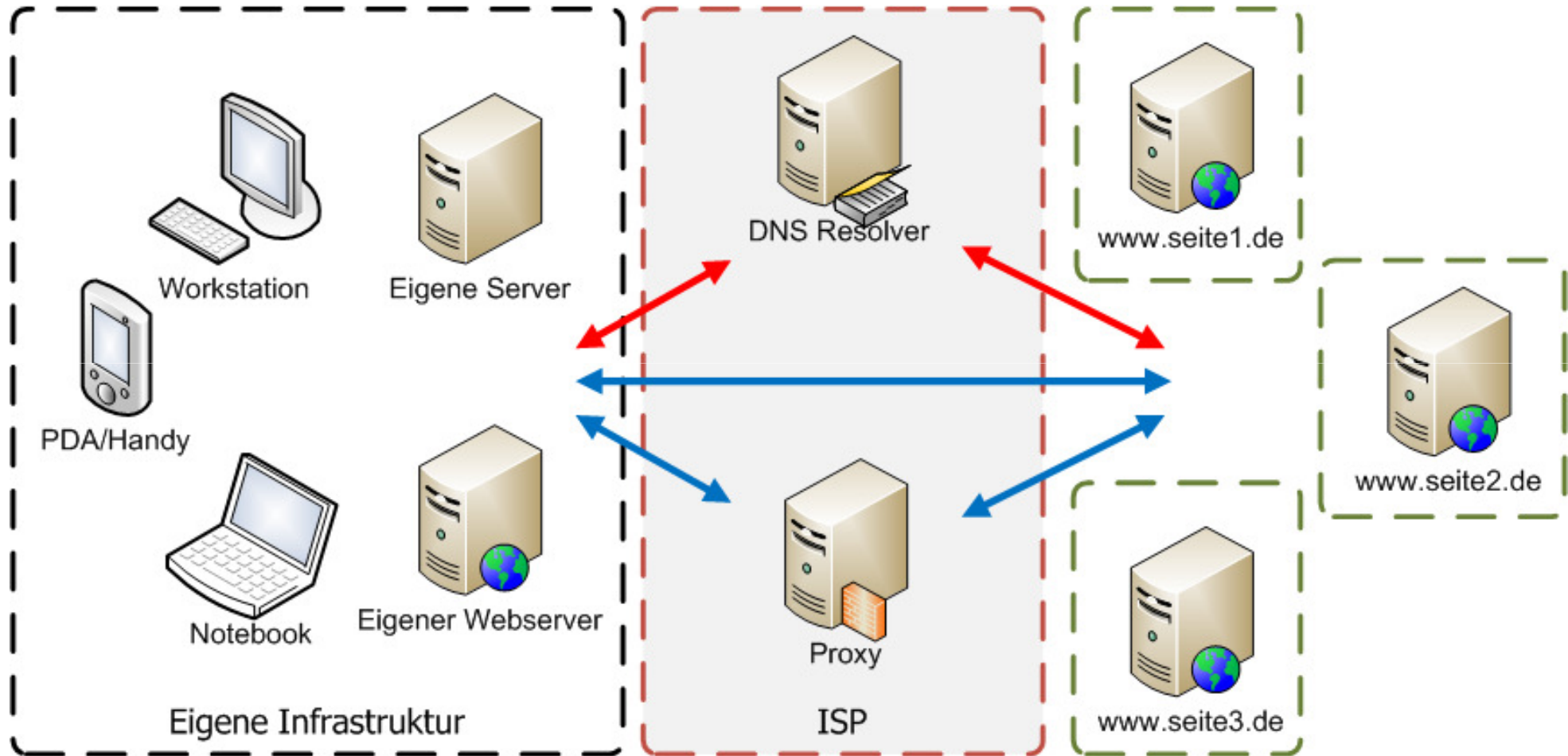
- Beate Uhse: Tausende E-Mail-Adressen veröffentlicht

[...] Online-Adventskalender. [...] Bei jedem Zugriff wird das Datum, die Zeit sowie eine E-Mail-Adresse protokolliert, letzteres vermutlich im Kontext des Newsletter-Versandes. Diese Logdateien waren ungeschützt und konnten über das Internet aufgerufen werden [...] Eins dieser Verzeichnisse ist sogar in den Index der Suchmaschine Google gelandet

<http://www.datenleck.net/?inputtype=art&ftext=logdateien>

- Abgrenzung
 - Passive Messungen / Logdatenanalyse
 - Aktive Messungen / keine klassische Logdatenanalyse
- **Die typische Datenpanne basiert nicht auf Logdaten!**

Eigene Datenspuren vermeiden



↔ Namensauflösung (www.seite1.de -> 194.94.127.40)
↔ Web/HTTP-Datenverkehr

Eigene Datenspuren vermeiden

→ „Private Browsing“

■ Firefox Privacy Mode (Firefox >= 3.5)



Privater Modus

Firefox wird keinerlei Chronik für diese Sitzung anlegen.

Im privaten Modus wird Firefox keinerlei Chronik anlegen. Das beinhaltet besuchte Seiten, Sucheinträge, Download-Chronik, Formulardaten, Cookies und temporäre Internetdateien. Allerdings werden alle Lesezeichen, die Sie anlegen, oder Dateien, die Sie herunterladen, beibehalten.

Um den privaten Modus zu beenden, wählen Sie Extras > Privaten Modus beenden, oder beenden Sie Firefox.

i Während auf diesem Computer keine Spuren Ihrer Browserchronik gespeichert werden, kann Ihr Internetprovider oder Ihr Arbeitgeber trotzdem nachverfolgen, welche Seiten Sie besuchen.

[Weitere Informationen](#)

■ Distrust/Stealthier Erweiterung (auch Firefox < 3.5)

Eigene Datenspuren vermeiden

→ „Private Browsing“ - Einschränkungen

- *„Während auf diesem Computer keine Spuren Ihrer Browserchronik gespeichert werden, kann Ihr Internetprovider oder Ihr Arbeitgeber trotzdem nachverfolgen, welche Seiten Sie besuchen.“*
- Kein Einfluss auf die gesendeten Daten

```
172.16.52.98 - - [13/Sep/2009:12:08:26 +0200] "GET  
/favicon.ico HTTP/1.1" 404 287 "-" "Mozilla/5.0  
(Windows; U; Windows NT 6.0; de; rv:1.9.1.3)  
Gecko/20090824 Firefox/3.5.3 (.NET CLR 3.5.30729)"
```

- Kein Schutz vor „aktiver“ Datensammlung
 - z.B. per JavaScript (Google Analytics)

Eigene Datenspuren vermeiden

→ „Private Browsing“ - Erweiterungen

- **BetterPrivacy:** „Super Cookie Safeguard vor LSO Flash Objekte, DOM Storage Objekte, eBay Langzeitverfolgung“
- **Targeted Advertising Cookie Opt-Out (TACO):** „Stops behavioral advertising by 90 different companies who quietly track you as you surf the Web “
- **NoScript:** Blockiert per default alle aktiven Inhalte
- **Adblock Plus:** Blockierung von Werbeeinblendungen
- **UserAgentSwitcher:** Veränderung des UserAgent-Strings
- **Dienste:** Plugins für Fakeaccounts (BugMeNot), Wegwerf-Email-Adressen (spamavert.com, mailinator, 10minutemail.com), ...



Eigene Datenspuren vermeiden

→ Proxy

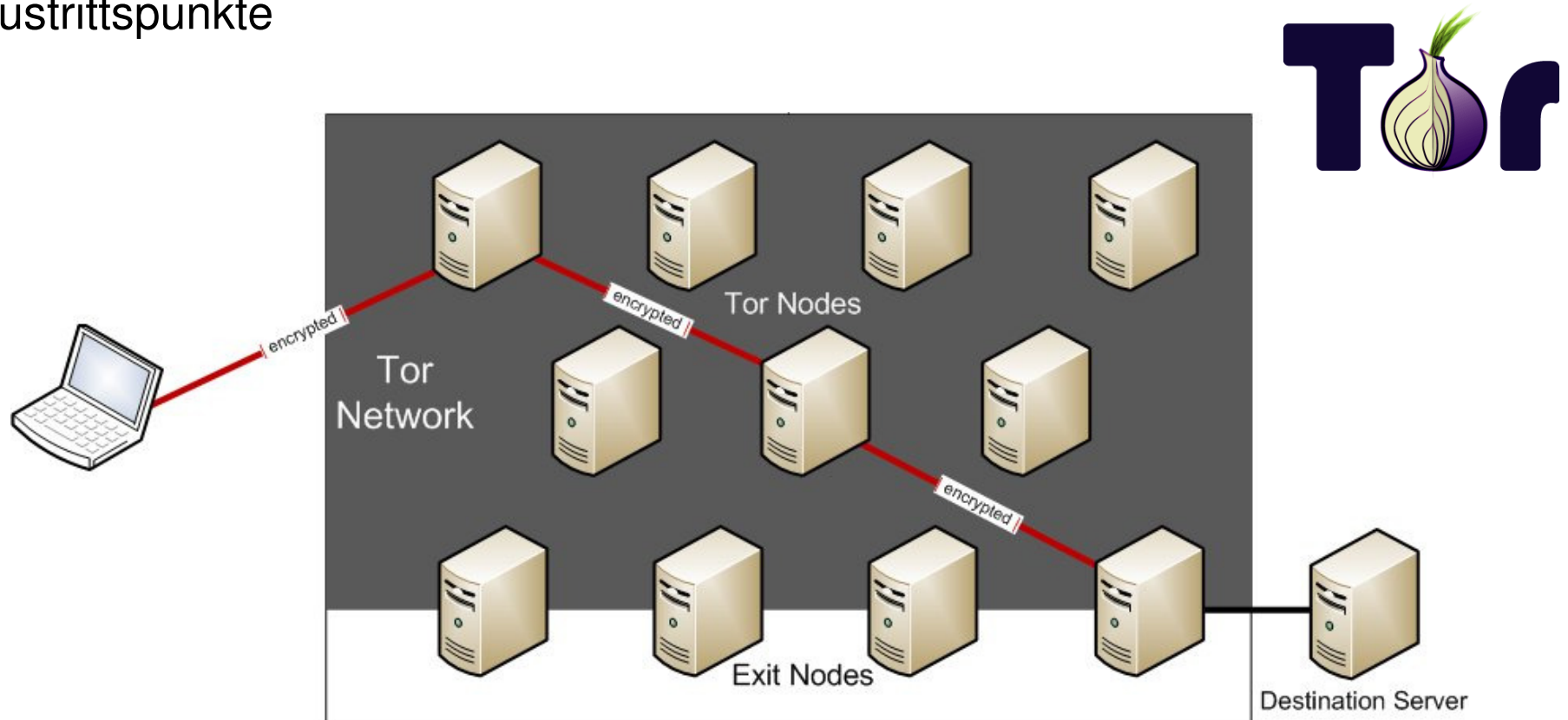
- **IP-Adressen lassen sich nicht alleine über lokale Maßnahmen verschleiern**
- **Einsatz von Proxy-Servern**
 - z.B. Firefox: FoxyProxy, QuickProxy, gladder, ...
 - Verwendung für alle Anfragen, oder
 - nur für definierte Seiten
- **Achtung: Proxy-Betreiber können**
 - Loggen
 - (unverschlüsselte) Daten mitschneiden
 - ➔ Zusammenhang zur tatsächlichen IP herstellen



Eigene Datenspuren vermeiden

→ Anonymisierungsdienste

- **Anonymisierungsdienste wie Tor, I2P & JAP**
effektiver als reine Proxy-Seiten
- **Kaskadierung** von mehreren Servern und wechselnde Austrittspunkte

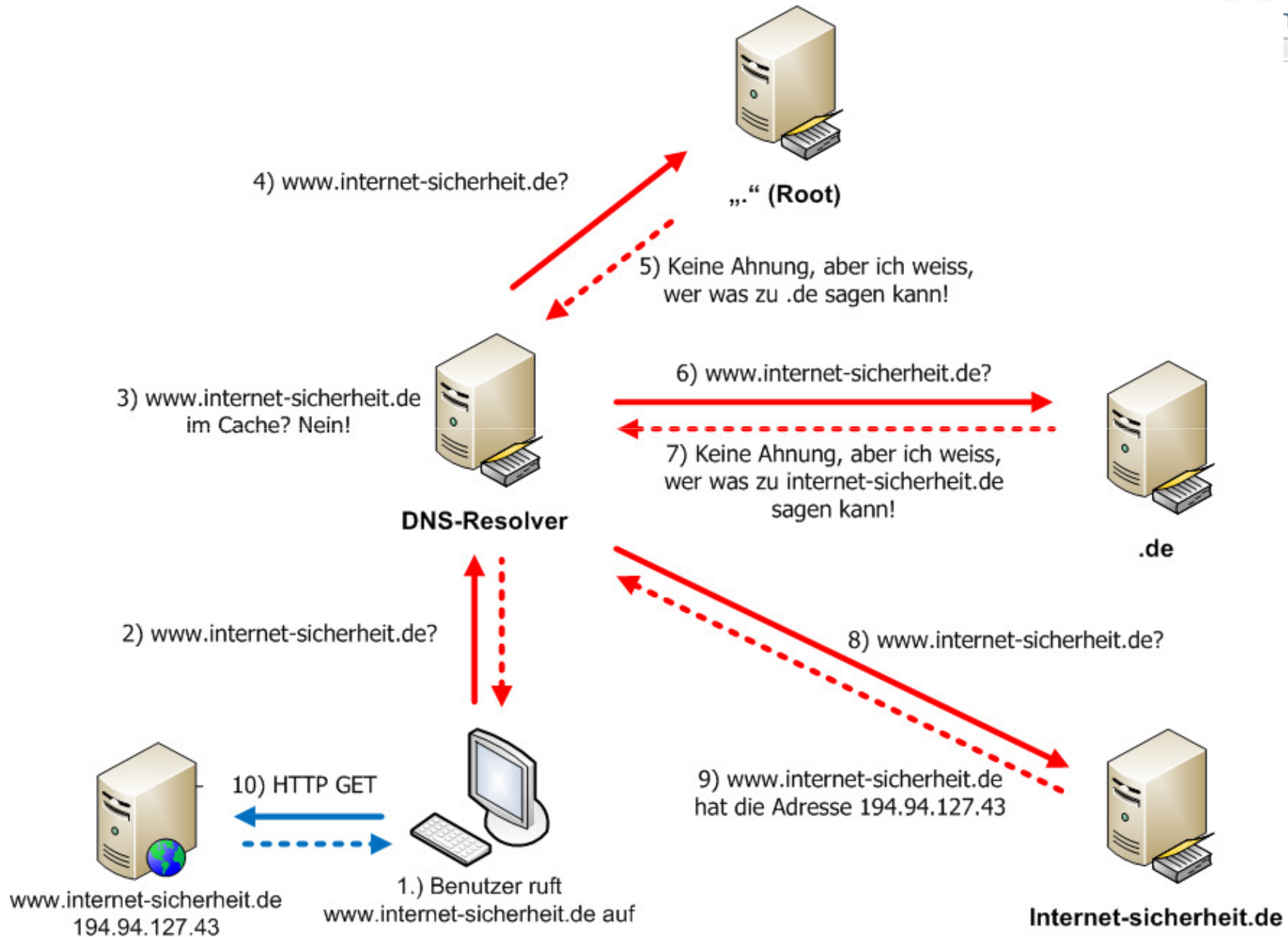


Eigene Datenspuren vermeiden

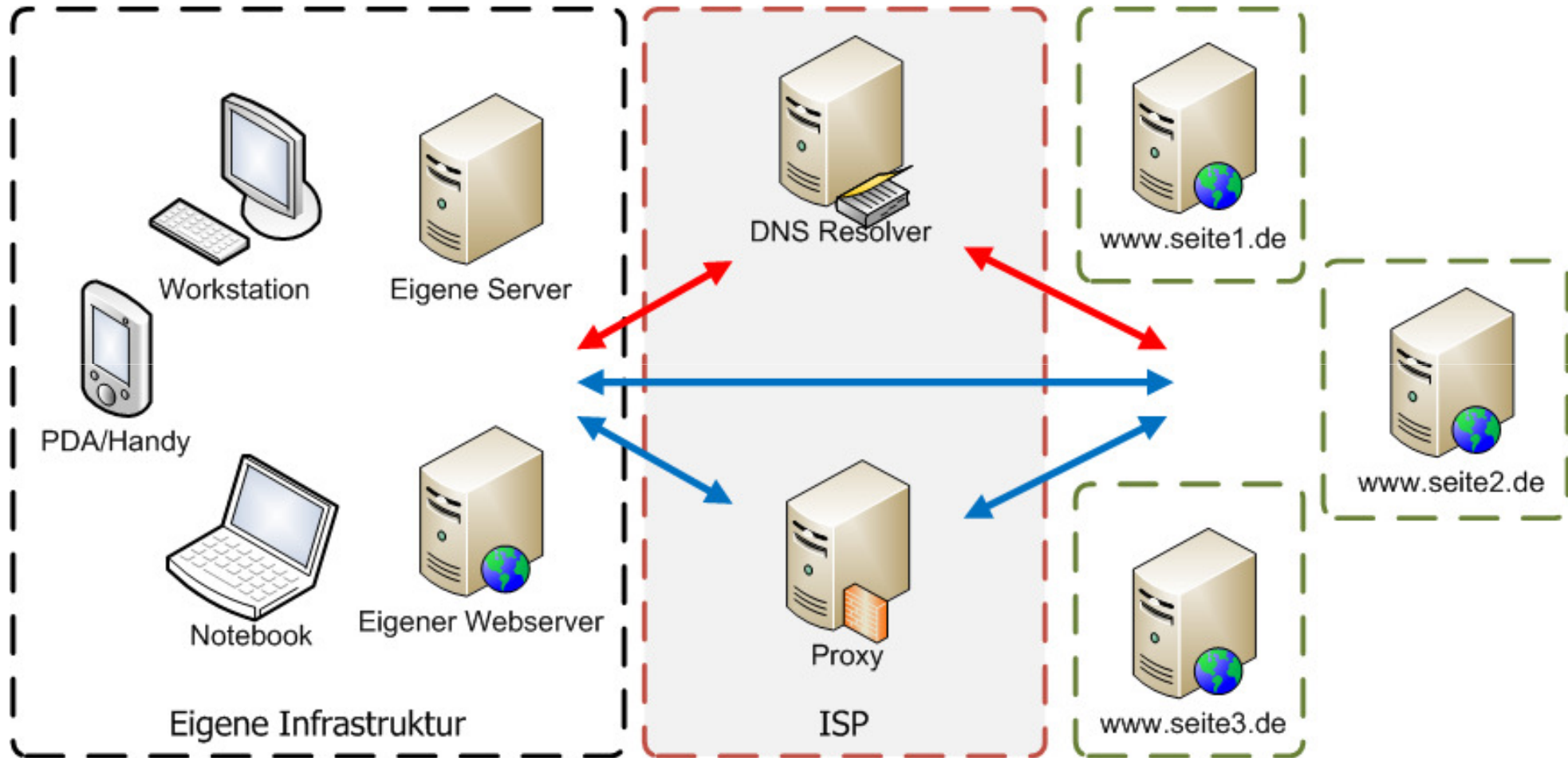
→ Anonymisierungsdienste - Einschränkungen

- **Credentials** (Logins, Cookies,...) **können** vom Webseiten-Betreiber **benutzt werden**, um den **Benutzer zu identifizieren**
 - Deshalb: Nur sinnvoll, wenn sie mit den Mitteln zur Verwischung von Datenspuren kombiniert werden!
- (Kostenlose Dienste) **nicht immer** ausreichend **performant**
- Wie bei Proxies können die Betreiber von „Exit-Nodes“ unverschlüsselte Daten mitlesen

Eigener DNS-Resolver

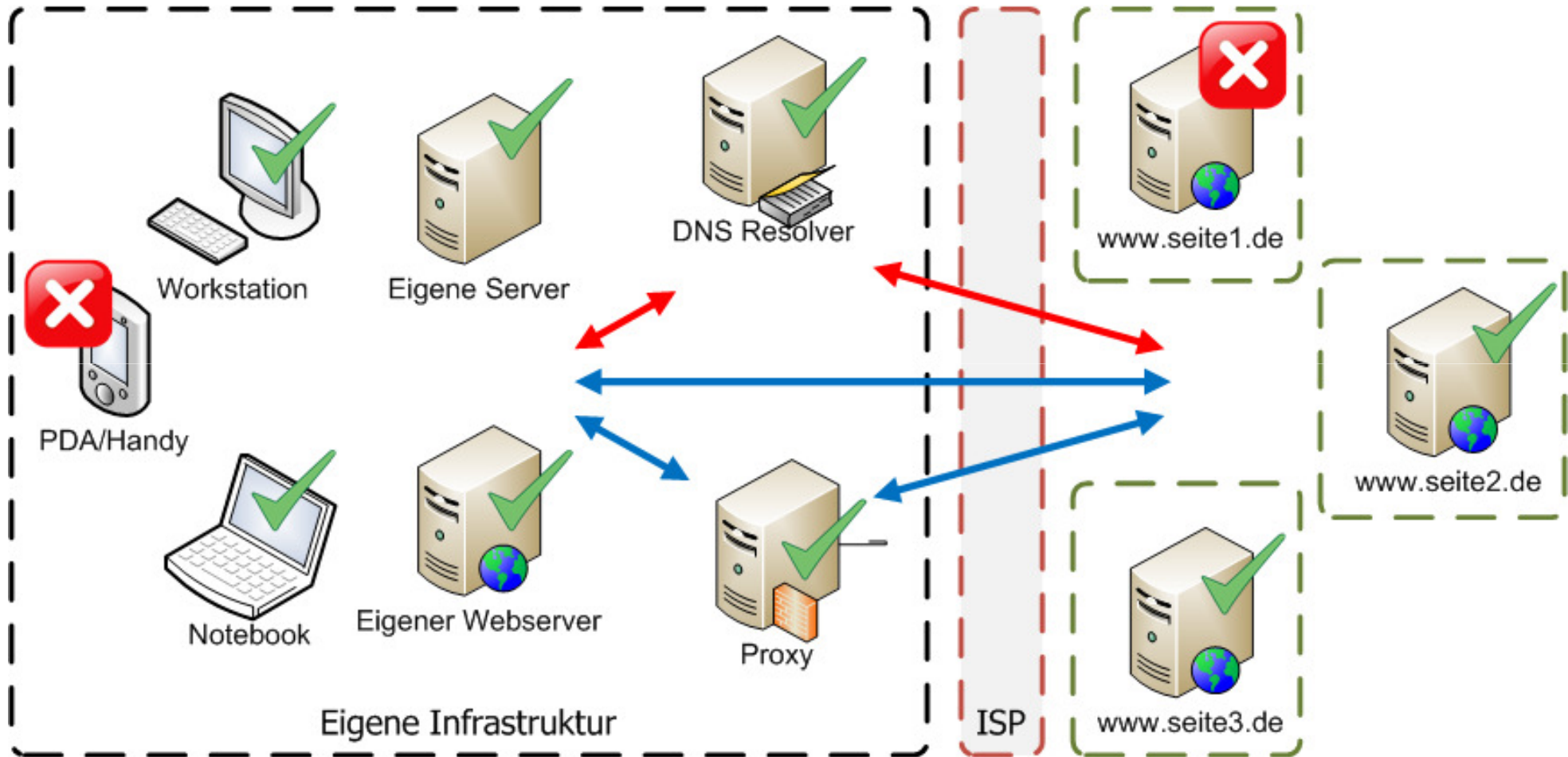


Modell



- Namensauflösung (www.seite1.de -> 194.94.127.40)
- Web/HTTP-Datenverkehr

Modell



- ↔ Namensauflösung (www.seite1.de -> 194.94.127.40)
- ↔ Web/HTTP-Datenverkehr

Logdaten – Spurensuche in der virtuellen Welt

Vielen Dank für Ihre Aufmerksamkeit!

Fragen?

Christian J. Dietrich
dietch [at] internet-sicherheit . de

Marian Jungbauer
jungbauer [at] internet-sicherheit . de

Institut für Internet-Sicherheit
<https://www.internet-sicherheit.de>
FH Gelsenkirchen

