SNS-Mining

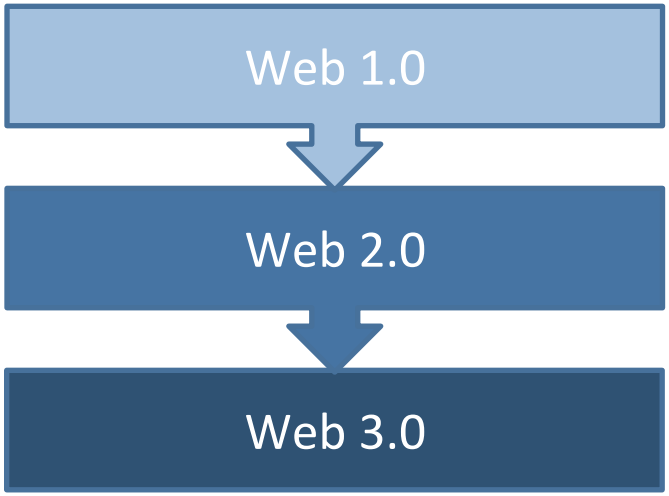
Anforderungen an einen Forensik-Crawler zur Erhebung und Sicherung beweisrelevanter Dokumente in sozialen Netzwerken

Andreas Dhein

K/15, IuK Forensik,
Techn. Ermittlungsunterstützung

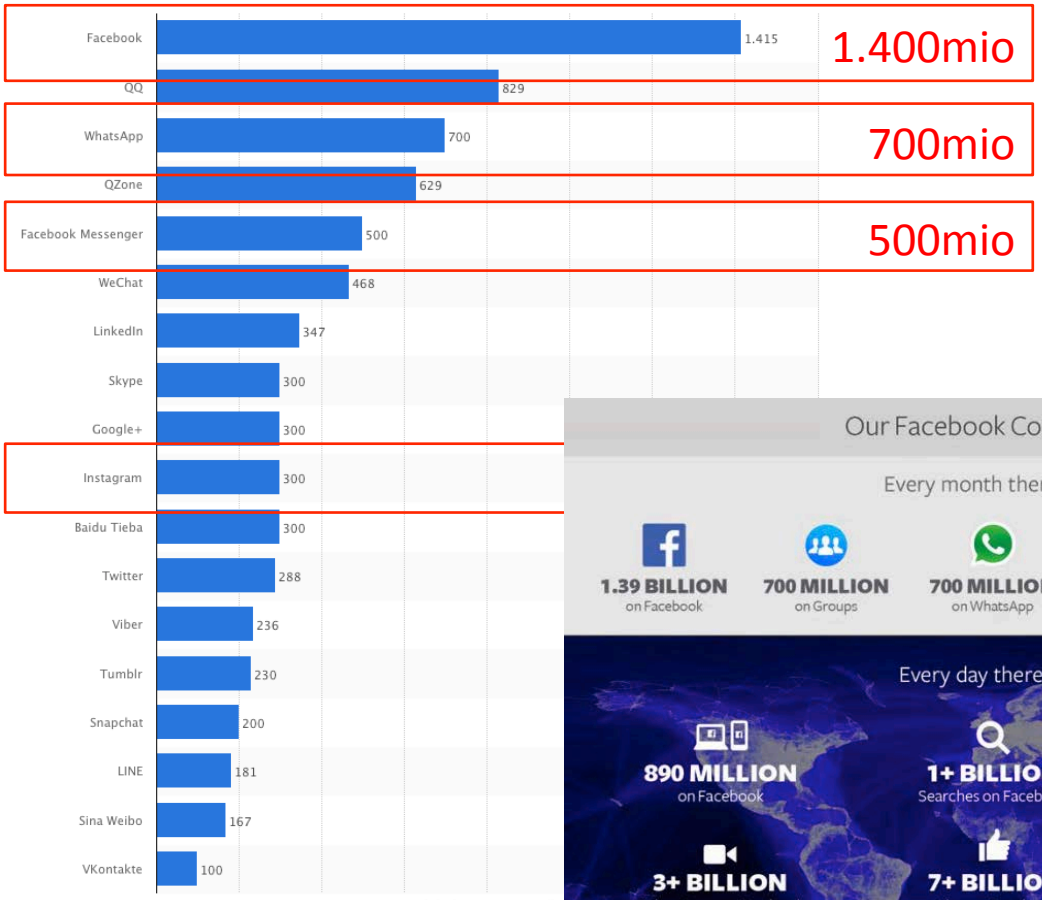
Einführung

- Motivation
  - Soziale Netzwerke sind keine „kleine Randerscheinung“ (mehr)
  - Dienstanbieter durchdringen sämtliche Bereiche des täglichen Lebens (Mobile Nutzung)
  - Mitteilungsbedürfnis der Nutzer ist enorm (in Echtzeit)



- Polizeiliche Herausforderungen:
  - Immer mehr Strafanzeigen mit Bezug zu SNS
  - Qualität der angelieferten Beweismittel mitunter mangelhaft (Ausdrucke, keine IDs)
  - Erhebung relevanter Beweismittel gestaltet sich häufig schwierig (quantitativ, technisch)

Nutzerzahlen

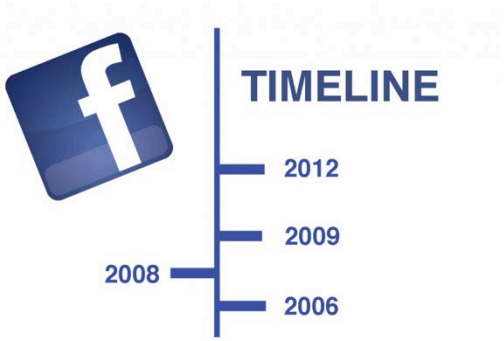


- Monthly Active Users
- And the Winner is....
  - FACEBOOK

Potential für Ermittlungen

RheinlandPfalz
POLIZEIPRÄSIDIUM KOBLENZ

- Was bieten Soziale Netze?
  - Profile (Inhaber, Urheber)
  - Beiträge (Nachrichten, Texte, Bilder)
  - Kommentare (Eigene, Fremde)
  - Chatkommunikation



- Welche Probleme können auftreten
  - Unterschiedliche Darstellungen (Desktop, Mobil)
  - Datenmenge (enorm, partielles Nachladen)
  - Sichtbarkeitseinstellungen (Privatsphäre)

K/15, IuK Forensik,
Techn. Ermittlungsunterstützung

Struktur von sozialen Netzen

RheinlandPfalz
POLIZEIPRÄSIDIUM KOBLENZ

- SNS sind Webservices, d.h.
  - Zugriff über das Internet
  - Daten liegen irgendwo zentral
  - Unterschiedliche Zugriffsmöglichkeiten
  - Gemeinsame API zum Datenabruf

- Wichtigste Erkenntnis für die Datensicherung
  - Sicherung über ein Medium genügt
  - Daten sind hochgradig volatil

Teil I

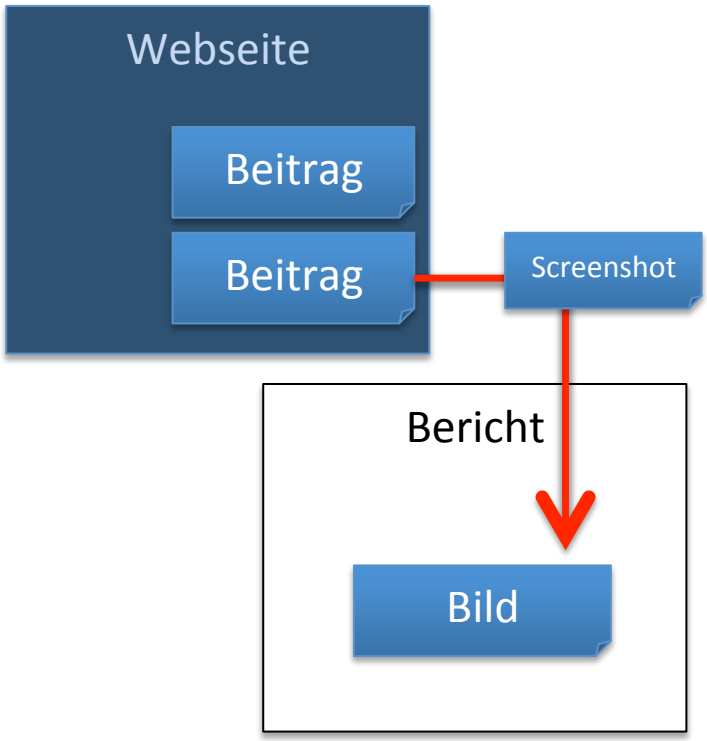
RheinlandPfalz
POLIZEIPRÄSIDIUM KOBLENZ

Grundlegende Möglichkeiten zur Datensicherung in Sozialen Netzen

K/15, IuK Forensik,
Techn. Ermittlungsunterstützung

Möglichkeit I - Screenshots

- Vorteile
  - Einfach umsetzbar
  - Gut nachvollziehbar
- Nachteile
  - Nur einzelne Beiträge
  - Sehr zeitaufwendig
- Nicht recherchierbar!



K/15, IuK Forensik,
Techn. Ermittlungsunterstützung

Weitere Möglichkeiten

- Profilarchiv erstellen
  - Vorteile
    - Recherchierbar
    - Übersichtlich aufbereitet vom Anbieter
  - Nachteile
    - Nicht immer vollständig (externe Inhalte)
    - Nicht immer durchführbar (Passwort/Kooperation?)

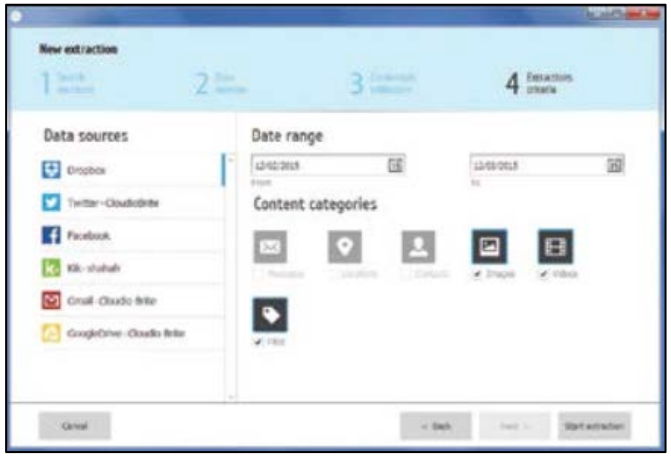
- Browser-Plugins
  - Awesome Screenshot, Fireshot (Firefox) -> Ganze Seite
  - Facepaste reloaded (Firefox) -> Bilder- Videoarchive
  - VideoDownloadhelper -> ext. Videos herunterladen

- Screencapture Tools
  - PicPick (kostenlos) -> Scrollende Webseiten
  - FastStoneCapture (komm.) -> Scrollen, Videocapture

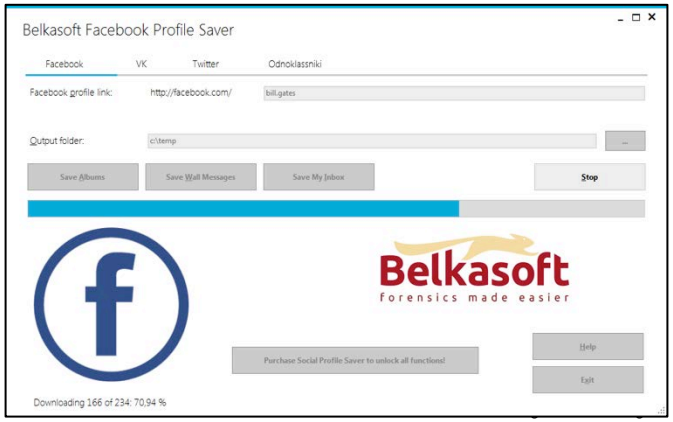
All-in-One Lösungen

... arbeiten alle nach dem gleichen Schema (API ausnutzen)

- Cellebrite Cloud Analyzer
  - Zugangsdaten aus mobilen Endgeräten können genutzt werden
  - Ergebnisse ???



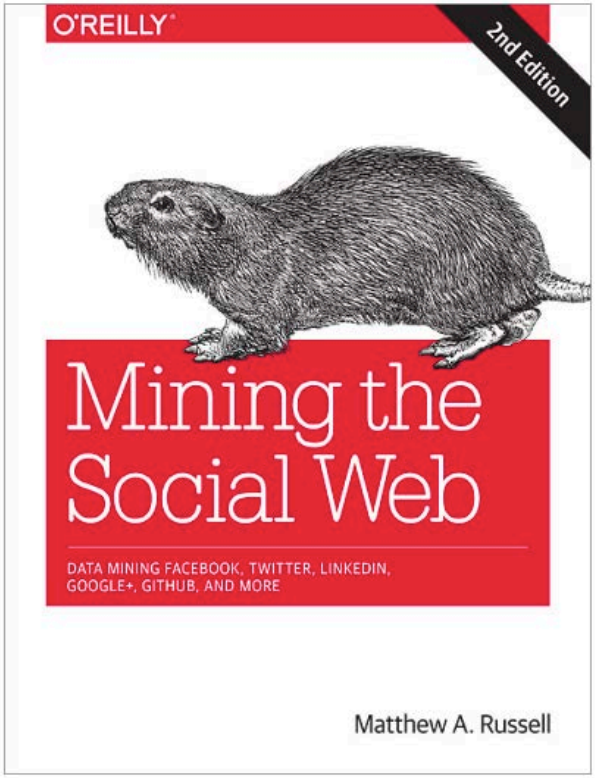
- Belkasoft Facebook Profile Saver (kostenlos)
  - Entwicklung eingestellt



Alternativ: Selbst Hand anlegen

- https://github.com/ptwobrussell/Mining-the-Social-Web-2nd-Edition

- [...]
- Chapter 1 - Mining **Twitter**: Exploring Trending Topics, Discovering What People Are Talking About, and More
- Chapter 2 - Mining **Facebook**: Analyzing Fan Pages, Examining Friendships, and More
- Chapter 3 - Mining **LinkedIn**: Faceting Job Titles, Clustering Colleagues, and More
- Chapter 4 - Mining **Google+**: Computing Document Similarity, Extracting Collocations, and More
- [...]



- Python Scripte zur Anregung als interaktive Lernplattform (iPython Notebook in VM)

- Vorteile
  - Einfache, verständliche Skripte (mit Erläuterungen)
  - Individuell anpassbar

- Nachteile
  - Programmierkenntnisse notwendig

K/15, IuK Forensik, Techn. Ermittlungsunterstützung

Teil II

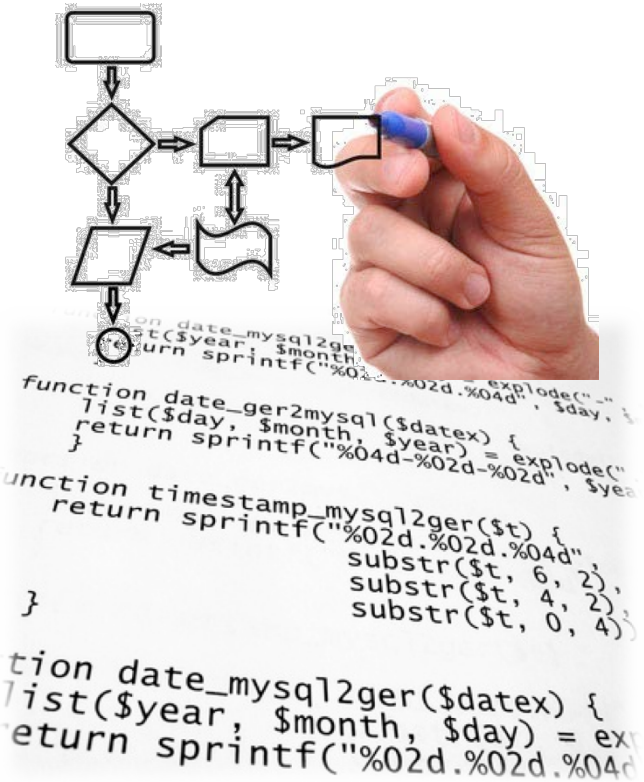
RheinlandPfalz
POLIZEIPRÄSIDIUM KOBLENZ

Anforderungen an einen Forensic SNS Crawler

K/15, IuK Forensik,
Techn. Ermittlungsunterstützung

Grundlegende Anforderungen

- Automatisierung
  - Einfache Bedienung
  - Geringe Interaktion
- Vollständigkeit
  - Rekursive Sicherung aller Daten
- Skalierung
  - Große Datenmengen
  - Unterschiedliche Problemstellungen
- Präsentation
  - Originalgetreu?
  - Bestmöglich?
  - Nachvollziehbar!



K/15, IuK Forensik,
Techn. Ermittlungsunterstützung

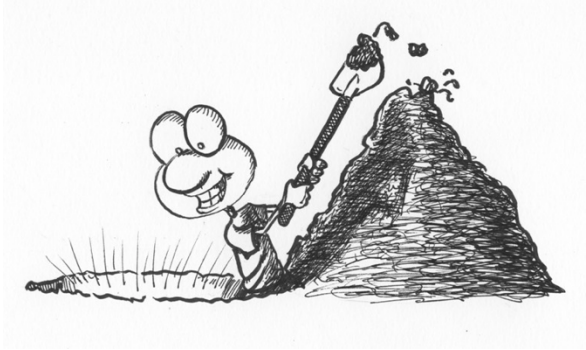
Probleme bisher (Screenshots)

- Automatisierung
  - Viele Einzelschritte notwendig
    - Facebook Anmeldung
    - Nach unten scrollen
- Vollständigkeit
  - Einsatz von AJAX („See more“)
    - Nachladen von Texten
    - Nachladen von Beiträgen
  - Kommentare von Kommentaren...
  - Fehlende Informationen (Uhrzeit, Datum)
- Skalierung
  - Fragestellung zu Beginn nicht immer klar
- Präsentation
  - Nicht recherchierbar

K/15, IuK Forensik,
Techn. Ermittlungsunterstützung

Der Königsweg – Ideenskizze

- Sicherung der Daten
  - Über die Facebook API
  - Gefiltert (Zeit, Inhalt)
  - In eine Datenbank
  - Multimediainhalte separat



- Aufbereiten der Daten
  - Anschaulich, Interaktiv
  - Wenn möglich druckbar

- Nach dem Konzept
  - Gesamtsicherung *versus*
  - Fallbezogene Darstellung

K/15, IuK Forensik,
Techn. Ermittlungsunterstützung

Teil III

RheinlandPfalz
POLIZEIPRÄSIDIUM KOBLENZ

Programmiertechnische Umsetzung (Lösung einzelner Problemstellungen)

K/15, IuK Forensik,
Techn. Ermittlungsunterstützung

Facebook Graph API

» Allgemeine Programmierschnittstelle für Facebook

» Aufbau einer Anfrage

https://graph.facebook.com/	*baseURL*
{user-id}	***FB_ID!!***
/feed | posts | photos | videos [...]	*Edge*
?fields | limit | summary | type	*Parameter*
&access_token=CAH3[..]ZD	***Access_Token!!***

» Beispiel

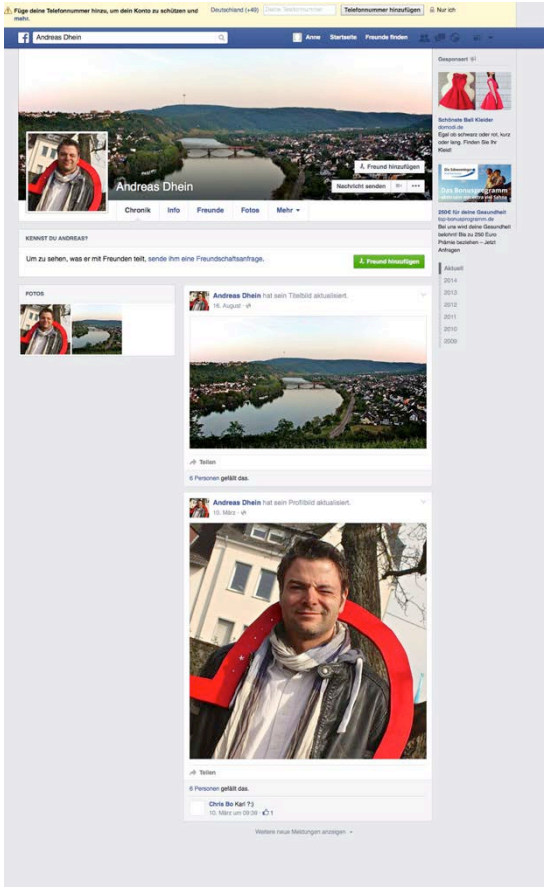
» https://graph.facebook.com/**{user-id}**/picture?width=9999&height=9999**&access_token=C[..]ZD**

K/15, IuK Forensik,
Techn. Ermittlungsunterstützung

Facebook Graph API

RheinlandPfalz
POLIZEIPRÄSIDIUM KOBLENZ

what you see



what you get ;)


```
{
        "picture": "https://scontent.xx.fbcdn.net/[...]s130x130/10731094_[...]06868_n.jpg?",
        "story": "Andreas Dhein updated his cover photo.",
        "likes": {
            "paging": {
                "cursors": {
                    "after": "NjY5NTYzNDk2NDUwMzY4",
                    "before": "MTUwNjAwNTU4MzAxNTQ0OA=="
                }
            },
            "data": [
                {
                    "id": "15060[...]5448",
                    "name": "Ein Freund"
                },
               [...]
            ]
        },
        "from": {
            "name": "Andreas Dhein",
            "id": "10202845130204269"
        },
        "name": "Andreas's cover photo",
        "is_expired": false,
        "updated_time": "2015-08-16T17:37:24+0000",
        "link": "https://www.facebook.com/photo.php?fbid=10201996458308002&type=3",
        "object_id": "10201996458308002",
        "created_time": "2015-08-16T17:37:24+0000",
        "is_hidden": false,
        "type": "photo",
        "id": "10202845130204269_10203491177155039",
        "status_type": "added_photos",
}
```

Webautomatisierung

- Selenium
  - http://www.seleniumhq.org
  - „Selenium automates browsers. That‘s it!“
  - Verfügbar für Java, C#, Ruby, python, JS
  - Webdriver für Firefox, Ie, Safari (manuell), etc.
- Was kann man damit tun?
  - Browser starten, beenden
  - Seiten aufrufen
  - Bestimmte Seitenelements detektieren
  - Eingaben automatisieren (Bsp. Einloggen)
  - Bildschirmfotos ganzer Seiten speichern

K/15, IuK Forensik,
Techn. Ermittlungsunterstützung

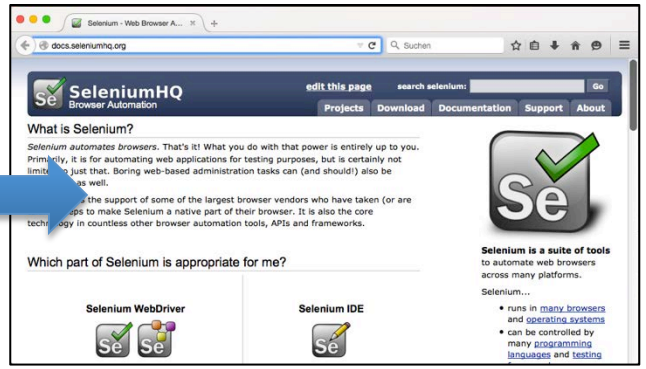
Selenium – kurze Einführung

» Selenium Installation unter python

```
# pip install -U selenium
```

» Beispiel 1 (Browser öffnen und Webseite laden)

```
from selenium import webdriver

browser = webdriver.Firefox()
browser.get('http://seleniumhq.org/')
```



» API online abrufbar unter:

» http://seleniumhq.github.io/selenium/docs/api/py/api.html

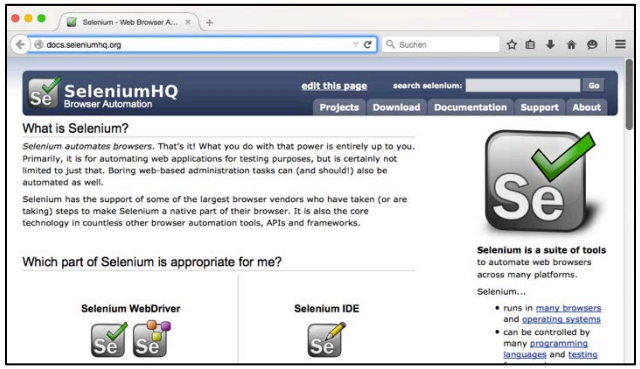
K/15, IuK Forensik,
Techn. Ermittlungsunterstützung

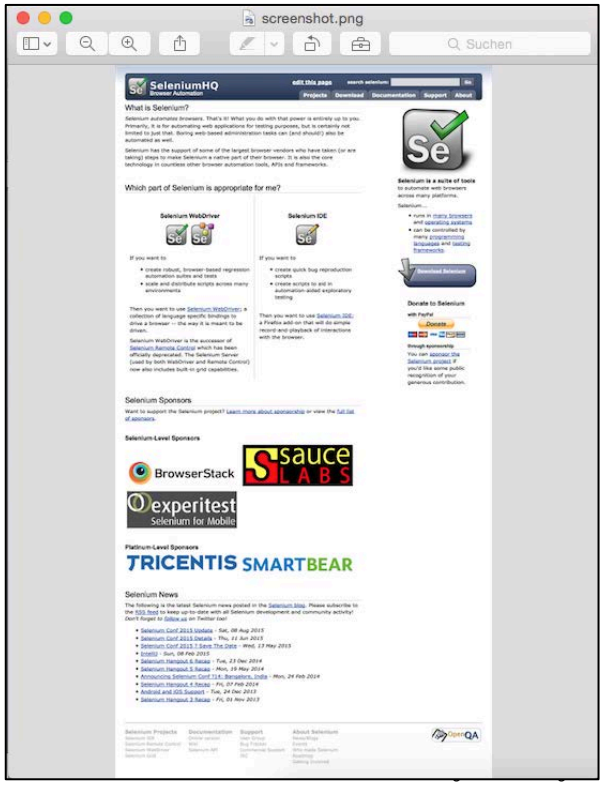
Selenium – Screenshots

RheinlandPfalz
POLIZEIPRÄSIDIUM KOBLENZ

» Beispiel 2

```
from selenium import webdriver

browser = webdriver.Firefox()
browser.get('http://seleniumhq.org/')

browser.save_screenshot('screenshot.png')
```





Selenium – Facebook-Login

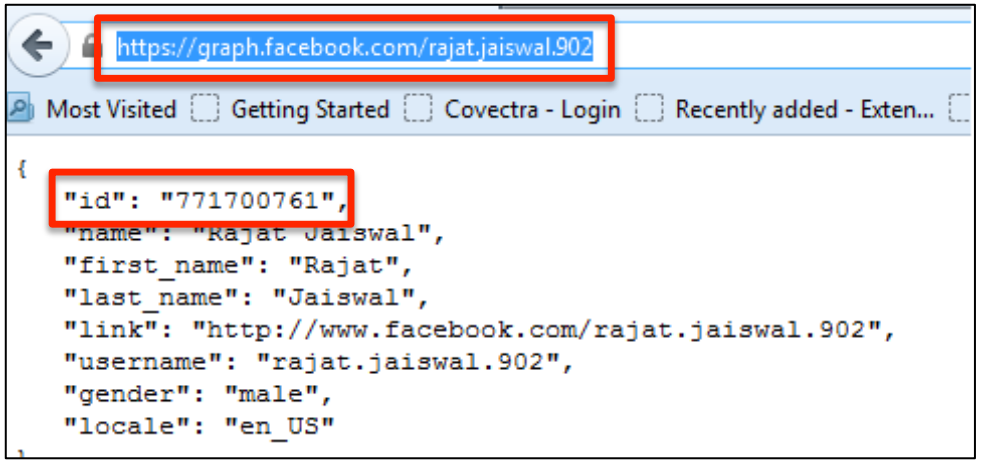
» Beispiel 3

```
import time
from selenium import webdriver

browser = webdriver.Firefox()
browser.get('http://www.facebook.com/')

inputElement = browser.find_element_by_name('email')
inputElement.send_keys('user@domain.com')
inputElement = browser.find_element_by_name('pass')
inputElement.send_keys('Test1234')

time.sleep(5) #inputElement.submit()
```

-> „submit“ Erkennung als „Robot“, besser selbst klicken...

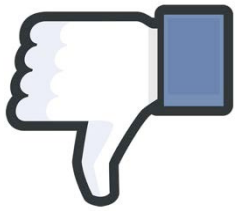
K/15, IuK Forensik,
Techn. Ermittlungsunterstützung

Ermitteln der Facebook ID



» Früher

» Heute



K/15, IuK Forensik,
Techn. Ermittlungsunterstützung

Selenium – Facebook ID

» Trick: FacebookID steckt (noch) im Link zum Profilbild

```
def getFBID(fbUserName):
      browser = webdriver.Firefox()
      browser.get("http://www.facebook.com/" + fbUserName)
      #get link to profile picture
      link = browser.find_element(By.CSS_SELECTOR,
                'a.profilePicThumb').get_attribute('href')
      #match FacebookID using regular expressions
      mg = re.search('([0-9]+\.[0-9]+\.)([0-9]+) ', link)
      fb_id = mg.group(2)
      return fb_id
```

K/15, IuK Forensik,
Techn. Ermittlungsunterstützung

Selenium – Access Token

» Beispiel 4

```
import time
from selenium import webdriver

browser = webdriver.Firefox()
browser.get('http://www.facebook.com/')

inputElement = browser.find_element_by_name('email')
inputElement.send_keys('user@domain.com')
inputElement = browser.find_element_by_name('pass')
inputElement.send_keys('Test1234')
time.sleep(5) #inputElement.submit()

browser.get("https://developers.facebook.com/tools/explorer")

access_token = browser.find_element_by_css_selector('input._58al')
       .get_attribute('value')
```

Zeit für Fragen/Anregungen

KD Koblenz, K15

IuK Forensik, Technische Ermittlungsunterstützung

Moselring 10-12
56068 Koblenz
+49 261 103 2489

@4rensiker
4rensics@gmx.de
andreas.dhein@polizei.rlp.de