MIKE KUKETZ 20. DEZEMBER 2017 7 ERGÄNZUNGEN

Google Hacking: Die dunkle Seite der Suchmaschine

1. Die dunkle Seite

Suchmaschinen wie Google nutzen die meisten für einen einfachen Zweck: Man möchte etwas im Internet finden. Doch Google hat auch eine **dunkle** Seite, mit der man Dinge im Internet aufspüren kann, die eigentlich nicht für den allgemeinen Gebrauch bestimmt sind: Geheime Dokumente, Passwörter, Gehaltsabrechnungen, urheberrechtlich geschütztes Material (Filme, Musik, etc.), ungeschützte Sicherheitskameras, Datenschutzverstöße oder Sicherheitslücken in Webanwendungen.

Bekannt unter dem Begriff »Google Hacking« oder »Google Dorks« kann jeder innerhalb weniger Minuten einen Blick in vertrauliche Regionen des Internets werfen. Das Handwerkszeug ist denkbar simpel: Mit ein paar zusätzlichen **Suchparametern** lässt sich Googles gigantischer Suchindex »anzapfen« und Informationen gewinnen, die der jeweilige Betreiber (vermutlich) nicht jedem verraten möchte.

Im vorliegenden Beitrag möchte ich euch ein paar dieser Google Suchparameter vorstellen, mit dem Konfigurationsfehler und die Unachtsamkeit der Betreiber schonungslos sichtbar gemacht werden können.

2. Vorwort

Bevor ihr anfangt mit den erweiterten Suchparametern von Google zu experimentieren, solltet ihr Folgendes berücksichtigen:

Das »Hacking« via Google ist zunächst nicht verboten oder strafbar, da die darüber gefundenen Informationen ungeschützt für jeden aus dem Internet erreichbar sind. Oder anders formuliert: Der Betreiber hat es versäumt, vertrauliche Daten in einem besonderen Maß zu schützen. Die Schwelle zur Kriminalität bzw. **Straftat** ist allerdings schnell überschritten, wenn ihr die erspähten Daten und Schwachstellen für Zwecke missbraucht, die bei den Betroffenen finanzielle, persönliche oder Schäden jedweder Art verursachen.

Grundsätzlich passiert beim Google Hacking aber zunächst nur eines: Ihr greift mit den speziellen Suchparametern auf Googles Datenbestände zu, die öffentlich und für jeden einsehbar bereitgestellt werden.

Hinweis

Mit Google Hacking ist übrigens **kein** Hacking-Angriff auf Google gemeint, sondern eine Art der Informationsgewinnung mit Hilfe der Suchmaschine von Google.

2.1 Warum Google Hacking funktioniert

Aufgrund zum Teil eklatanter Konfigurationsfehler oder fehlenden Anweisungen in der robots.txt kann es vorkommen, dass Google Daten bzw. Informationen indexiert, die im Normalfall nicht für den allgemeinen Gebrauch bestimmt sind. Die Ursache dieser Konfigurationsfehler lassen sich meist auf zwei Gründe zurückführen:

- **Mangelnde Kenntnis**: Insbesondere Privatpersonen legen oftmals sensible Daten ungeschützt auf einen FTP- oder Webserver, den sie von zu Hause aus betreiben. Es werden keinerlei Sicherheitsmaßnahmen ergriffen oder Geräte (direkt) mit dem Internet verbunden, die weder Sicherheitsupdates erhalten, noch für diesen Zweck ausgelegt sind. Wenn ihr auf solche ungeschützten Daten bzw. Informationen stoßt, dann solltet ihr den Anstand besitzen und dies dem Betreiber melden – sofern ihr diesen identifizieren könnt. Es gilt: Wer nicht weiß was er tut, der sollte eventuell darauf verzichten Daten ins Internet zu stellen oder einfach jemanden dafür bezahlen, der sich damit auskennt.
- **Unachtsamkeit:** Mit Google Hacking lassen sich allerdings nicht nur sensible Informationen von Privatanwendern aufspüren, sondern auch brisante, zum Teil auch geheime Dokumente von Unternehmen und Institutionen. Hier kann man davon ausgehen, dass die verantwortlichen Administratoren im Normalfall über ausreichend Kenntnisse verfügen sollten, um solche Fälle zu vermeiden. Ungepatchte Webserver, sensible Passwortlisten oder vollständige Backups der geschäftlichen E-Mail Korrespondenz, die sich via Google Hacking aufspüren lassen, zeichnen allerdings ein anders Bild. Hier reden wir von Unachtsamkeit, bis hin zu grober Fahrlässigkeit und menschlichem Versagen.

Google Hacking geht übrigens auf das Jahr 2002 zurück, als Johnny Long (j0hnny) mittels speziellen Suchparametern verwundbare bzw. ungepatchte Systeme und sensible Informationen gefunden hat – er

selbst bezeichnet das als **Google Dorks**. Mit Dork (Trottel / Idiot) sind die Betreiber der Webseite gemeint, die sich nicht ausreichend Gedanken zur Sicherheit gemacht haben und ihre Server nicht unter Kontrolle haben.

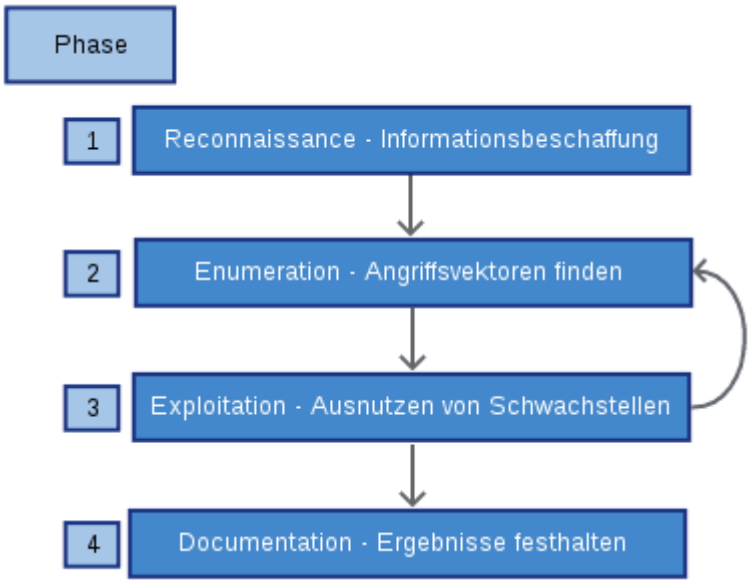
Wer nun glaubt, dass sich die Situation bis zum Jahr 2017 geändert hat, der irrt gewaltig. Dank dem Internet of Things und der damit steigenden Zahl von Geräten, die mit dem Internet vernetzt sind, lassen sich heute weitaus häufiger verwundbare Systeme aufspüren und sensible Informationen abgreifen, als jemals zuvor.

Hinweis

Nicht nur Google Hacking eignet sich zum Aufspüren von sensiblen Informationen und ungeschützten Geräten, sondern auch andere Tools und Suchmaschinen, wie bspw. Shodan, die explizit dafür entworfen wurden.

2.2 Informationsgewinnung beim Pentest

Bei der Durchführung eines Pentetrantionstests ist das Google Hacking für mich zu einem unverzichtbaren Werkzeug geworden. Ein Penetrationstest lässt sich in unterschiedlichen Phasen unterteilen, von denen ein Teil sequentiell wiederholt wird:



In der ersten Phase (Informationsbeschaffung) geht es zunächst darum, möglichst viele Informationen über das Ziel zu sammeln, die für den weiteren Verlauf von Interesse sein könnten. Zu diesem Zweck werden verschiedene öffentlich verfügbare Informationsquellen durchsucht und im Anschluss ausgewertet. Eine dieser Informationsquellen ist Google bzw. das »Hacking« via Google.

3. Einfache Beispiele

Starten wir zunächst mit einfachen Beispielen und was sich mit den speziellen Suchparametern alles im Internet aufspüren lässt.

3.1 Video- und Audiomaterial

Anzeige von Webseiten, die mp3-Dateien öffentlich bereitstellen. Sei es aufgrund mangelnden Kenntnissen, Unachtsamkeit oder sogar absichtlich, um urheberrechtlich geschütztes Material anzubieten:

```
intitle:index of inurl:mp3
```

Ähnlich funktioniert es auch bei Videomaterial, wenn ihr avi, mp4 oder mkv als Parameter übergebt:

```
intitle:index of inurl:avi
```

Wenn ihr eure Suche anschließend noch verfeinert und einen Interpreten oder Filmtitel ergänzt, dann könnt ihr die Ergebnisse weiter verdichten.

3.2 Schwachstellen

Die Ausgabe von SQL-Fehlermeldungen sind ein erstes Indiz für die Verwundbarkeit gegenüber einer SQL-Injection:

```
ysql_connect()" | intext:"Warning: mysql_query()" | intext:"Warning: pg_connect()"
```

Aufspüren von sensiblen Konfigurationsdateien, die MySQL Login-Daten wie Benutzernamen und Passwörter beinhalten:

```
mysqli_connect ext:inc
```

Darstellung von verwundbaren Apache Webservern, mit einer bestimmten Versionsnummer und System:

```
intitle:index of "Apache/2.4.7 (Ubuntu) Server"
```

Aufspüren von Seiten die Outlook Web Application (OWA) anbieten und damit auf Microsoft Exchange aufbauen:

```
inurl:https://owa
```

3.3 Sensible Informationen / Dokumente

Aufspüren von (Cisco) VPN-Zugangsdaten, mit denen man sich Zugriff auf das lokale Netz von Unternehmen und Institutionen verschaffen kann:

```
!Host=*.* intext:enc_UserPassword=* ext:pcf
```

Anzeige von Suchergebnissen zu FTP-Servern mit Passwortdateien im Microsoft XLS-Format:

```
inurl:ftp "passwort" filetype:xls
```

Oder ihr sucht gezielt nach Seiten, die den Begriff »passwort«, »secret«, »members« oder andere Ausdrücke direkt in der URL haben sollen:

```
intitle:index of inurl:passwort
```

Suche nach Text-Dateien, die das Wort »passwort« beinhalten:

```
ext:txt intext:passwort
```

Anzeige der (Linux) Bash-History, die häufig Benutzernamen zu einem System beinhaltet:

```
intitle:index of .bash_history
```

Anzeige von möglicherweise sensiblen Dropbox Ordnern:

```
intitle:index.of.dropbox
```

Hinweis

Eine riesige Auswahl an Google Dorks findet ihr unter anderem auf der Exploit Database unter Google Hacking Database (GHDB). Dort könnt ihr dann gezielt nach einem Dork suchen oder in den unterschiedlichen Kategorien wie »Footholds« oder »Vulnerable Servers« stöbern.

4. Suchparameter

Einige der wichtigsten Suchparameter, die beim Google Hacking eingesetzt werden, habt ihr nun schon bei den Beispielen kennengelernt. Die am häufigsten verwendeten Suchparameter werde ich euch kurz erläutern – diese lassen sich oftmals auch wunderbar miteinander kombinieren.

4.1 Basisparameter

Die doppelten **Anführungsstriche** kennt wohl bereits jeder, der sich näher mit dem Prinzip einer Suchmaschine beschäftigt hat. Sie sorgen dafür, dass genau nach dem angegebenen SUCHWORT

gesucht wird. Wenn ihr mehrere SUCHWÖRTER umschlossen von den doppelten Anführungsstrichen eingebt, dann wird die Reihenfolge ebenso berücksichtigt:

```
"SUCHWORT (bzw. mehrere) SUCHWÖRTER"
```

Mit **site:** wird die Suche auf eine bestimmte Seite fokussiert. Das ist praktisch, wenn ihr die Suche nach Inhalten, sensiblen Dokumenten oder Schwachstellen (SUCHWORT) auf eine bestimmte DOMAIN (bspw. kuketz-blog.de) beschränken wollt:

```
site:DOMAIN SUCHWORT
```

Der Suchparameter **cache:** in Verbindung mit einer URL zeigt Googles Cache-Version der angegebenen Webseite an, also die Version der letzten Indexierung:

```
cache:URL
```

Eine Suche mit dem Suchparameter **intitle:** in Verbindung mit einem SUCHWORT liefert Ergebnisse von Webseiten, deren Titel diesen Suchbegriff enthält. Wird häufig dafür benutzt, um Seiten bzw. Verzeichnisse zu finden, die Directory Listing (index of / index.of) aktiviert haben:

```
intitle:SUCHWORT
```

Mit **allintitle:** werden nur Ergebnisse angezeigt, die alle angegebenen SUCHWÖRTER im Titel enthalten.

```
allintitle:SUCHWÖRTER
```

Mit dem Suchparameter **intext:** in Verbindung mit einem SUCHWORT werden Webseiten angezeigt, in denen der Begriff im Text der Seite vorkommt:

```
intext:SUCHWORT
```

Mit **allintext:** werden nur Ergebnisse angezeigt, die alle angegebenen SUCHWÖRTER im Text der Seite beinhalten:

```
allintext:SUCHWÖRTER
```

Eine Suche mit dem Suchparameter **inurl:** in Verbindung mit einem SUCHWORT liefert Ergebnisse von Webseiten, deren URL den Suchbegriff enthalten:

```
inurl:SUCHWORT
```

Eine Suche mit dem Suchparameter **filetype:** oder **ext:** beschränkt die Ergebnisse auf Dokumente eines bestimmten FORMATs (bspw. pdf oder docx Dateien):

```
filetype:FORMAT
```

Mit dem Suchoperator **related:** und der Angabe einer DOMAIN lassen sich Seiten finden, die der angegebenen Seite inhaltlich ähnlich sind. Das könnt ihr mal mit dem Kuketz-Blog ausprobieren (related:www.kuketz-blog.de):

```
related:DOMAIN
```

Hinweis

Google erkennt den Einsatz der erweiterten Suchparameter übrigens und wird euch hin und wieder Captchas einblenden, um herauszufinden ob ihr Mensch oder Bot seid.

4.2 Kombinieren von Suchparametern

Die eben kennengelernten Suchparameter könnt ihr übrigens wunderbar miteinander kombinieren. Anbei ein paar Beispiele.

Anzeige von Ergebnissen, die Directory Listing aktiviert haben und bei denen das Suchwort »backup« innerhalb der URL vorkommt. Damit findet ihr unzählige Server bzw. Webseiten, auf denen mitunter komplette Backups von Privatpersonen und Unternehmen ungeschützt abgelegt sind:

```
intitle:index of inurl:backup
```

Aufspüren von WordPress-Seiten, die das Plugin WP Security Audit Log benutzen und die Log-Files ungeschützt für jeden bereitstellen:

```
inurl:"wp-security-audit-log" ext:log
```

Anzeige von Webseiten, die einen Schwachstellen-Report, durchgeführt von IBM AppScan, ungeschützt im PDF-Format bereitstellen. Das funktioniert übrigens auch mit anderen Vulnerability Scannern – brisante Details inklusive:

```
lication Report" intext:"This report was created by IBM Security AppScan" ext:pdf
```

4.3 Eigene Webseite »hacken«

Abgesehen von der Verfeinerung eurer Suchergebnisse solltet ihr Google Hacking in erster Linie zum Aufspüren eigener Konfigurationsfehler benutzen. Anbei habe ich euch ein paar Standard Google Hacks zusammengefasst, mit denen ihr eure eigene Webseite / Internetpräsenz prüfen könnt. Ersetzt »kuketz-blog.de« einfach mit eurer Domain.

Verzeichnisse, bei denen das Directory Listing aktiv ist:

```
site:kuketz-blog.de intitle:index.of
```

Vergessene oder ungeschützte Konfigurationsdateien:

```
site:kuketz-blog.de ext:xml | ext:conf | ext:cnf | ext:reg | ext:inf | ext:rdp | e
```

Vergessene oder ungeschützte (SQL-)Datenbanken bzw. Backups:

```
site:kuketz-blog.de ext:sql | ext:dbf | ext:mdb
```

Vergessene oder ungeschützte Log-Dateien:

```
site:kuketz-blog.de ext:log
```

Backups und veraltete / vergessene (Sicherungs-)Dateien:

```
site:kuketz-blog.de ext:bkf | ext:bkp | ext:bak | ext:old | ext:backup
```

Seiten, die »login« in der URL beinhalten und auf einen zu schützenden Bereich hindeuten:

```
site:kuketz-blog.de inurl:login
```

Ausgabe von SQL-Fehlermeldungen. Ein erstes Indiz für SQL-Injections:

```
site:kuketz-blog.de intext:"sql syntax near" | intext:"syntax error has occurred"
```

Vergessene oder ungeschützte Dokumente:

```
site:kuketz-blog.de ext:doc | ext:docx | ext:odt | ext:pdf | ext:rtf | ext:sxw | e
```

Ausgabe der phpinfo():

```
site:kuketz-blog.de ext:php intitle:phpinfo "published by the PHP Group"
```

Wie ihr seht ist keine spezielle »**Hacker-Software**« notwendig, um Schwachstellen oder Konfigurationsfehler auf euren Webseiten ausfindig zu machen. Mit Google Hacking könnt ihr euch einen schnellen Überblick verschaffen und Angreifern, die es nicht wirklich gut mit euch meinen, zuvorkommen.

5. Fazit

Google Hacking ist zwar legal, allerdings ist die Schwelle zu einer Straftat rasch überschritten. Allzu schnell wird man neugierig und klickt sich durch die riesigen Datenbestände, die Google indexiert hat. Lasst euch nicht in Versuchung führen, denn zu einem Teil der Ergebnisse zählen sicherlich auch Honeypots, die euer Verhalten aufzeichnen und analysieren.

Mit Google Hacking lassen sich zwar auch Schwachstellen in der IT von (fremden) Unternehmen aufspüren, allerdings solltet ihr dies lieber professionellen Penetrationstestern überlassen, die meist **auftragsgesteuert** handeln. Nutzt das erworbene Wissen über die Google Dorks lieber dafür, um eure persönlichen Suchergebnisse zu optimieren oder eure eigenen Webseiten / Geräte auf Konfigurationsfehler abzuklopfen.

Bildquellen:

Skull: Useful Objects from www.flaticon.com is licensed by CC 3.0 BY
Google: Freepik from www.flaticon.com is licensed by CC 3.0 BY

Weitersagen | Unterstützen

Wenn dir der Beitrag gefallen hat, dann **teile** ihn mit deinen Freunden, Bekannten und Mitmenschen. Nutze dafür soziale Medien (datenschutzfreundlich via Mastodon), Foren, Messenger, E-Mails oder einfach die nächste Feier / Veranstaltung. Gerne darfst du meine Arbeit auch unterstützen!

Über den Autor | Kuketz

In meiner freiberuflichen Tätigkeit als Pentester / Sicherheitsforscher (Kuketz IT-Security) schlüpfe ich in die Rolle eines »Hackers« und suche Schwachstellen in IT-Systemen, Webanwendungen und Apps (Android, iOS). Des Weiteren bin ich **Lehrbeauftragter** für IT-Sicherheit an der dualen Hochschule Karlsruhe, schärfe durch Workshops und Schulungen das **Sicherheits-** und **Datenschutzbewusstsein** von Personen und bin unter anderem auch als Autor für die Computerzeitschrift c't tätig.

Der Kuketz-Blog bzw. meine Person ist regelmäßig in den Medien (heise online, Spiegel Online, Süddeutsche Zeitung etc.) vertreten.

Mehr Erfahren ➡

Unterstützung erhalten

Wenn du Fragen hast oder Hilfe suchst, sind das offizielle Forum oder der Chatraum geeignete Anlaufstellen, um den Sachverhalt dort zu erörtern.

FOLGE DEM BLOG

Wenn du über aktuelle Beiträge informiert werden möchtest, hast du verschiedene Möglichkeiten, dem Blog zu folgen:

Bleib aktuell ➡

ÄHNLICHE BEITRÄGE

21. AUGUST 2018

Google Captcha auf EU-Seite: Leserbeschwerde führt zu Erfolg

15. OKTOBER 2018

Shodan: Suchmaschine für das »Internet of Shit«

21. NOVEMBER 2017

Google Hacking: Immer wieder faszinierend

5. DEZEMBER 2017

Firefox-Update ändert Suchmaschine auf Google

28. NOVEMBER 2013

Hacking WordPress – Ein Blick hinter die Kulissen

ERGÄNZUNGEN

7 Ergänzungen zu “Google Hacking: Die dunkle Seite der Suchmaschine”

Danilo sagt:

20. Dezember 2017 um 14:34 Uhr

Danke für diesen interessanten, aufschlussreichen Beitrag, Mike!

Tobias sagt:

20. Dezember 2017 um 15:22 Uhr

Vielen Dank für den Informativen Artikel! Ich würde gerne darauf Hinweisen, dass Google mittlerweile allerdings teilweise Suchergebnisse „zensiert“. Es lassen sich über andere Suchmaschinen wie Duckduckgo oder Bing teilweise deutlich bessere Ergebnisse anzeigen. Oft erscheint auch nach mehreren Suchen die nervigen Captchas. Aber da gilt wahrscheinlich das gleiche wie immer. Nicht auf eine Suchmaschine fixieren ;-)

SecInt sagt:

20. Dezember 2017 um 21:15 Uhr

Wieder mal ein kurzweiliger, interessanter Beitrag. Danke Mike! Du hast ein gutes Gespür dafür Wissen zu vermitteln.

Anonymous sagt:

21. Dezember 2017 um 07:40 Uhr

Wie immer ausgezeichnet. Nirgendwo sonst finde ich solch ansprechenden Beiträge, wie bei dir Mike. +1

Vielleicht kannst du auch mal Shodan vorstellen, wäre sicherlich interessant.

savant sagt:

21. Dezember 2017 um 21:06 Uhr

Es gibt zwei kurze Videos von SemperVideo, die sich auf Shodan beziehen.

Shodan.io: Die besondere Suchmaschine*
Metasploitable zum Austoben*

* Link führt zu YouTube

Herrmann Mann sagt:

22. Dezember 2017 um 11:11 Uhr

Ist mittlerweile ein paar Jahre her, aber ich hatte auf meiner Webseite einen Honey-Pot, d.h. einen unklickbaren link, welcher einzig dazu dar war Webcrawler die sich nicht an die robots.txt halten zu identifizieren. Der Google crawler ist ständig ins „Fettnäpfchen“ (Honigtöpfchen) getreten.

Ronin sagt:

23. Dezember 2017 um 12:47 Uhr

Super Artikel Mike! Ich habe zwar keine Ahnung in Erstellung von Websites und ihren Schutz, aber wenn ich mal eine Website betreibe, werde ich mich an diesen Artikel erinnern!

Dir und deiner Familie schöne Feiertage und einen guten Start ins neue Jahr!

Ergänzungen sind geschlossen / Aktualität

Dieser Beitrag ist älter als vier Wochen. **Hinweis**: Die Blog-Beiträge haben nicht wie Enzyklopädie-Einträge (bspw. Wikipedia) den Anspruch, dauerhaft aktuell und richtig zu sein, sondern beziehen sich wie Zeitungsartikel auf den Informationsstand zum Redaktionsschluss.

Wenn du Fragen hast oder Hilfe suchst, sind das offizielle Forum oder der Chatraum geeignete Anlaufstellen, um den Sachverhalt dort zu erörtern. Kritik, Anregungen oder Korrekturvorschläge zum Beitrag nehme ich gerne per E-Mail entgegen.

Kuketz-Forum

Artikel (324)

Hacking (15) Pentest (18)

NACH OBEN ⬆