Konark Modi
@konarkmodi

Josep M. Pujol
@solso
Cliqz GmbH, Munich

Local-Sheriff:

First step in users‘ data journey towards privacy

https://github.com/cliqz-oss/local-sheriff

Telltale URL : which contains sensitive information or can lead to a page which contains sentive information.

EG: http://track.emirates.email/track/click/30705682/www.emirates.com?p=eMSwicCI6IntcInVcIjozM....(REDACTED)

Telltale URLs

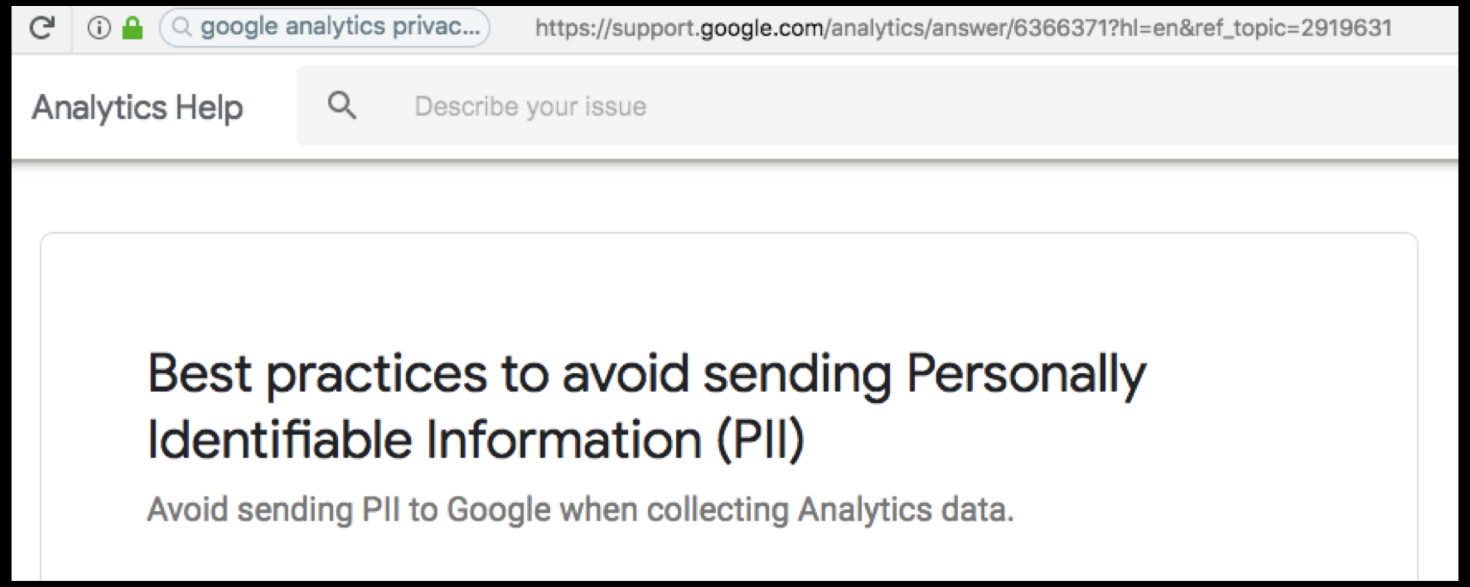
- Websites are carelessly leaking senstive PII information to pleothra of third-parties.

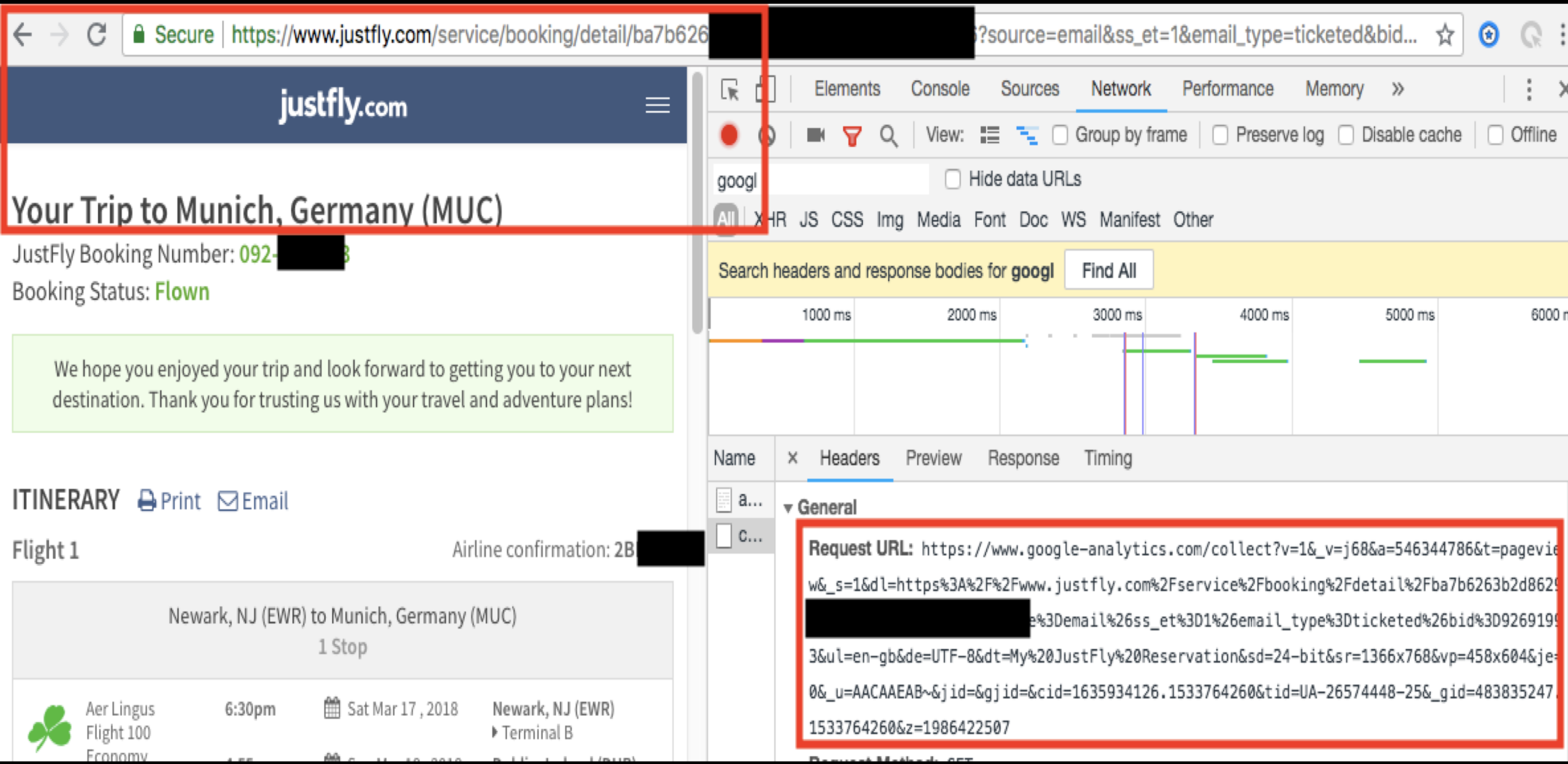
Telltale URLs

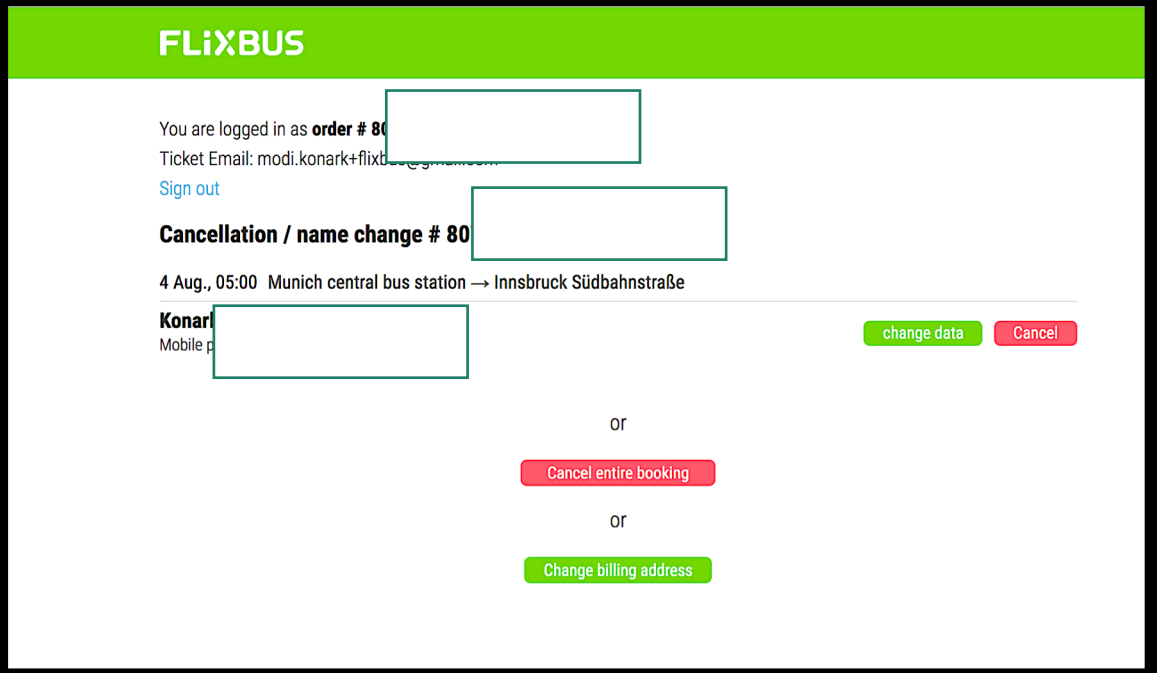
- Websites are carelessly leaking senstive PII information to pleothra of third-parties.
- Most often without users‘ consent

Telltale URLs

- Websites are carelessly leaking senstive PII information to pleothra of third-parties.
- Most often without users‘ consent
- More dangerously: Without the websites realizing themselves.




Secure
https://www.justfly.com/service/booking/detail/ba7b626
?source=email&ss_et=1&email_type=ticketed&bid...
justfly.com
Your Trip to Munich, Germany (MUC)
JustFly Booking Number: 092-
Booking Status: Flown
We hope you enjoyed your trip and look forward to getting you to your next destination. Thank you for trusting us with your travel and adventure plans!
ITINERARY
Print
Email
Flight 1
Airline confirmation: 2B
Newark, NJ (EWR) to Munich, Germany (MUC)
1 Stop
Aer Lingus
Flight 100
Economy
6:30pm
Sat Mar 17 , 2018
Newark, NJ (EWR)
Terminal B
Elements
Console
Sources
Network
Performance
Memory
View:
Group by frame
Preserve log
Disable cache
Offline
googl
Hide data URLs
All
XHR
JS
CSS
Img
Media
Font
Doc
WS
Manifest
Other
Search headers and response bodies for googl
Find All
1000 ms
2000 ms
3000 ms
4000 ms
5000 ms
Name
Headers
Preview
Response
Timing
a...
c...
General
Request URL: https://www.google-analytics.com/collect?v=1&_v=j68&a=546344786&t=pagevie
w&_s=1&dl=https%3A%2F%2Fwww.justfly.com%2Fservice%2Fbooking%2Fdetail%2Fba7b6263b2d8629
e%3Demail%26ss_et%3D1%26email_type%3Dticketed%26bid%3D9269199
3&ul=en-gb&de=UTF-8&dt=My%20JustFly%20Reservation&sd=24-bit&sr=1366x768&vp=458x604&je=
0&_u=AACAAEAB~&jid=&gjid=&cid=1635934126.1533764260&tid=UA-26574448-25&_gid=483835247.
1533764260&z=1986422507



Request URL: https://fonts.googleapis.com/css?family=PT+Sans+Narrow:400,700

Request Method: GET

Status Code: 200

Remote Address: [2a00:1450:4001:821::200a]:443

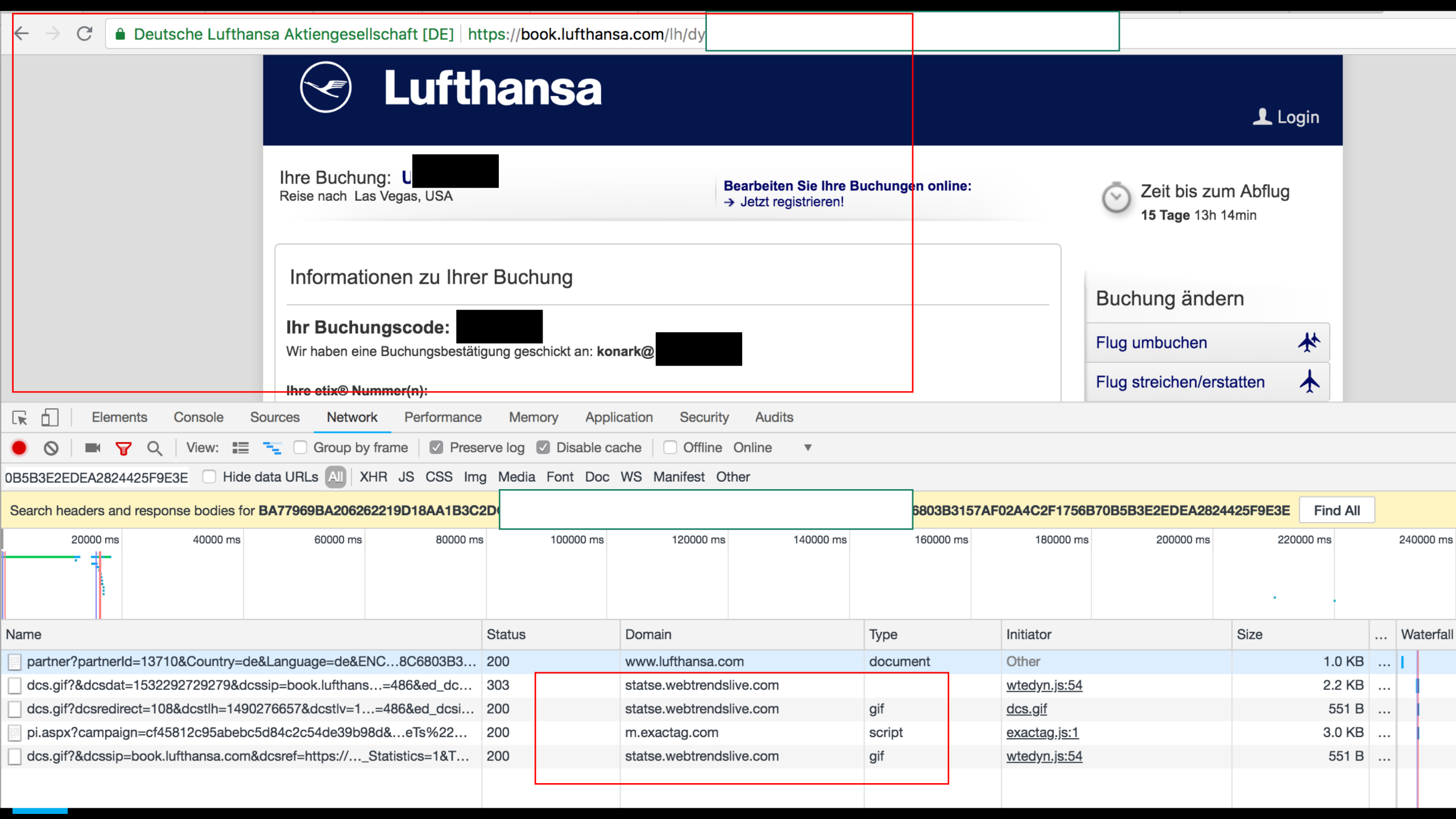
Referrer Policy: no-referrer-when-downgrade

▶ Response Headers (15)

▼ Request Headers

⚠ Provisional headers are shown

Referer: https://shop.flixbus.de/rebooking/mobile/auth?orderId=807511[redacted]oadHash=54x[redacted]fy6dhecns52gw7igtono52yszg2[redacted]


Deutsche Lufthansa Aktiengesellschaft [DE] | https://book.lufthansa.com/lh/dy
Lufthansa
Login
Ihre Buchung: U[redacted]
Reise nach Las Vegas, USA
Bearbeiten Sie Ihre Buchungen online:
→ Jetzt registrieren!
Zeit bis zum Abflug
15 Tage 13h 14min
Informationen zu Ihrer Buchung
Ihr Buchungscode: [redacted]
Wir haben eine Buchungsbestätigung geschickt an: konark@[redacted]
Ihre etix® Nummer(n):
Buchung ändern
Flug umbuchen
Flug streichen/erstatten
Elements
Console
Sources
Network
Performance
Memory
Application
Security
Audits
View:
Group by frame
Preserve log
Disable cache
Offline
Online
0B5B3E2EDEA2824425F9E3E
Hide data URLs
All
XHR
JS
CSS
Img
Media
Font
Doc
WS
Manifest
Other
Search headers and response bodies for BA77969BA206262219D18AA1B3C2D[redacted]6803B3157AF02A4C2F1756B70B5B3E2EDEA2824425F9E3E
Find All
20000 ms
40000 ms
60000 ms
80000 ms
100000 ms
120000 ms
140000 ms
160000 ms
180000 ms
200000 ms
220000 ms
240000 ms
Name
Status
Domain
Type
Initiator
Size
Waterfall
partner?partnerId=13710&Country=de&Language=de&ENC…8C6803B3…
200
www.lufthansa.com
document
Other
1.0 KB
dcs.gif?&dcsdat=1532292729279&dcssip=book.lufthans…=486&ed_dc…
303
statse.webtrendslive.com
wtedyn.js:54
2.2 KB
dcs.gif?dcsredirect=108&dcstlh=1490276657&dcstlv=1…=486&ed_dcsi…
200
statse.webtrendslive.com
gif
dcs.gif
551 B
pi.aspx?campaign=cf45812c95abebc5d84c2c54de39b98d&…eTs%22…
200
m.exactag.com
script
exactag.js:1
3.0 KB
dcs.gif?&dcssip=book.lufthansa.com&dcsref=https://…_Statistics=1&T…
200
statse.webtrendslive.com
gif
wtedyn.js:54
551 B

https://www.spotify.com/de/account/overview/?utm_source=spotify&utm_medium=menu&utm_campaign=your_account&**oauth%255ftoken=NAph…(REDACTED)**

```
2018-03-18 03:27:23 GET https://sb.scorecardresearch.com/b?c1=2&c2=15654041&ns__t=1521340042970&ns_c=UTF-8&c8=Konto%C3%BCbersicht%20-%20Spot
                        fy&c7=https%3A%2F%2Fwww.spotify.com%2Fde%2Faccount%2Foverview%2F%3Futm_source%3Dspotify%26utm_medium%3Dmenu%26utm_ca

                        ← 204 No Content [no content] 2.25s
                  Request                              Response                              Detail
Host:              sb.scorecardresearch.com
Connection:        keep-alive
User-Agent:        Mozilla/5.0 (Macintosh; Intel Mac OS X 10_12_6) AppleWebKit/537.36 (KHTML, like Gecko) Chrome/65.0.3325.146 Safari/537.36
                   OPR/52.0.2871.20 (Edition beta)
Accept:            image/webp,image/apng,image/*,*/*;q=0.8
Referer:           https://www.spotify.com/de/account/overview/?utm
                   =NAphCkkKDVNwb3RpZnktdXNlcnMSCmtvbmFyazMwMDMaA5g
                   UL4shIvIrpBMcA
Accept-Encoding:   gzip, deflate, br
Accept-Language:   en-GB,en-US;q=0.9,en;q=0.8
Cookie:            UID=17523a194187173be228dcg1521336255; UIDR=1521336255
Query                                                                                                                  [m:auto
c1:     2
c2:     15654041
ns__t:  1521340042970
ns_c:   UTF-8
c8:     Kontoübersicht - Spotify
c7:     https://www.spotify.com/de/account/overview/?utm_source=sp
b3RpZnktdXNlcnMSCmtvbmFyazMwMDMaA5gBASUnxq1aMiIKCFqOQbXqyPevEhbh
```


OH THE HORROR

Empowering users‘

- Install browser extension. (Firefox, Chrome, Opera) .
- Does not need to know anything that happens under the hood.
- Store all the data exchanged among website and 3rd parties.
- Identify companies behind 3rd parties
- Provide a search interface on top of the data stored.
- Provide automated **“Points of interest“**.

Local-Sheriff

Input information which should not have been shared without consent with services you have not interacted with. Like Email address, Phone number, Name, Address etc.

7R[redacted]

7R[redacted]

has been leaked to **7 third-party domains** , owned by **6 different companies** courtesy **1 website.**

Companies which now have your Booking reference [redacted]

Website leaking Booking reference

Company	Website
Tealium	www.lufthansa.com
BuySellAds.com	www.lufthansa.com
Unknown	www.lufthansa.com
Google	www.lufthansa.com
Exactag	www.lufthansa.com
Refined Labs	www.lufthansa.com

Who benefits from Local Sheriff ?

- **Users'** can educate themselves about:
  - What sensitive information with is being shared this which parties?
  - What companies are behind these third parties?
  - What can they doing with this information? EG: de-anonymize users on the internet, create shadow profiles.
- **Organizations** can audit:
  - Which all the third-parties that are being used on their websites.
  - Audit privacy policies.
  - Audit implementation of 3rd parties.

Under the hood

1. Monitor network request using webRequest APIs
2. Identify first-party, third-party calls using **whotracks.me** database.
3. Store information shared with a third-party via headers like cookies, referrer, query parameters.
4. Do anonymous GET request of the URL shared with a third-party.
5. Store the HTML content received.
6. Make the data stored in steps 3, 4 available via search interface.
7. Purge data older than 10 days.*

DEMO

https://streamable.com/yl3qq

Roadmap

- POST requests are not analyzed right now.
- GET requests could be base64, we need to reverse and save them.
- Notification when information entered in a form is leaked.
- **„Points of interest“:** Automated selection of data that could be sensitive.

THANK YOU