Finding digital evidence in mobile devices

Research project by B.Sc. student Vince Noort
Supervised and sponsored by Dutch National Police

Dr. Hans Henseler
Professor Digital Forensics & E-Discovery
Faculty Science & Technology
University of Applied Sciences, Leiden, The Netherlands

August 2017, DFRWS USA 2017, Austin, TX

Contents

- Introduction
- Problem statement
- Contribution
- Designing and implementing the script
- Experiment results
- Conclusion
- Recommendations

Introduction

- Mobile phone extraction tools have:
  - International focus
  - Large number of apps
  - Limited app support
- Not all data analysed
- Manual search is required

Problem statement

"How can we automatically extract data from Dutch mobile apps in such a way that the data becomes accessible to investigators?"

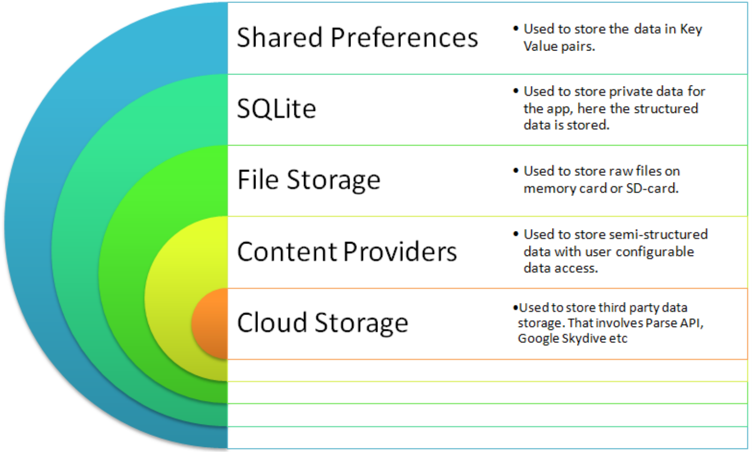
Why? Solve more cases, save time and effort

Contribution

- UFED PA plugin python script that crawls mobile phone dumps for SQLite databases and identifies interesting table headers and identity related data using regular expressions
- Validated idea and script during project with Dutch Law enforcement
- Recommendations for future work

Mobile OS storage types

- Different storage types in Android and IOS:
  - Shared preferences
  - File storage
  - Content providers
  - Cloud storage
  - Databases:
    - Sqlite,
    - IOS also CoreData and REALM
  - Text files (xml, json etc.)



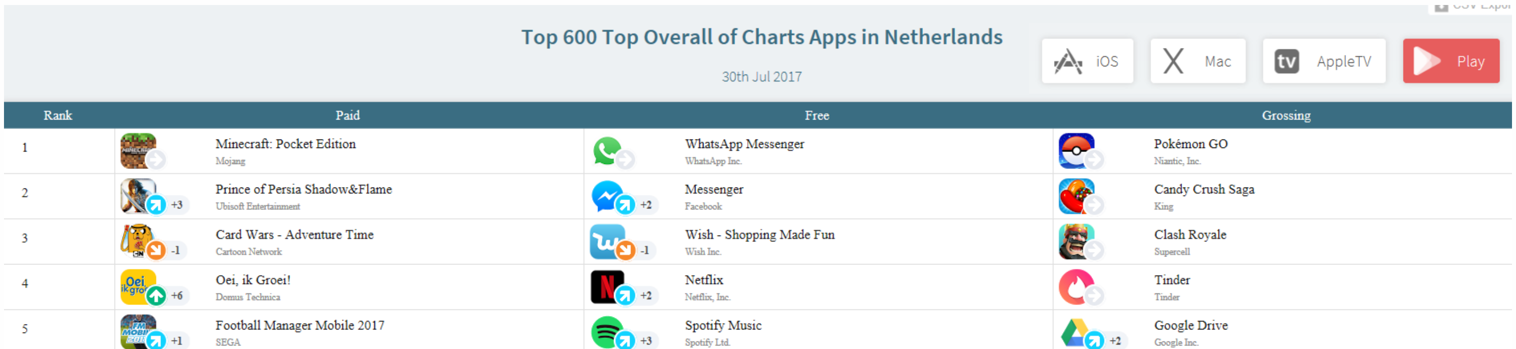
Shared Preferences	• Used to store the data in Key Value pairs.
SQLite	• Used to store private data for the app, here the structured data is stored.
File Storage	• Used to store raw files on memory card or SD-card.
Content Providers	• Used to store semi-structured data with user configurable data access.
Cloud Storage	•Used to store third party data storage. That involves Parse API, Google Skydive etc

https://stackoverflow.com/questions/9986734/which-android-data-storage-technique-to-use/9986948

Interesting entities

Question	Information Type
Who	○ User and device data ○ Contacts ○ Call history
What	○ SMS/Chat messages ○ Multi-media messages ○ E-mail ○ Social-media related data
Where	○ Application related data ○ Location data
When	○ Date, time, language and other settings ○ Calendar
Why	○ Chat messages ○ Documents
With what	○ Photos ○ Audio and video
How	○ Internet history

Popular apps in NL (or your own country)

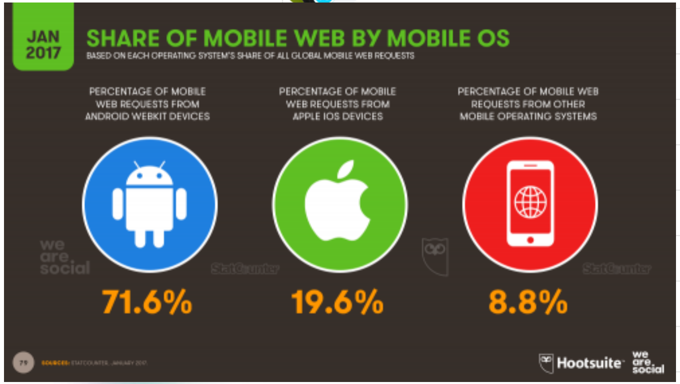


CSV Expor

Top 600 Top Overall of Charts Apps in Netherlands

30th Jul 2017

Rank	Paid	Free	Grossing
1	Minecraft: Pocket Edition Mojang	WhatsApp Messenger WhatsApp Inc.	Pokémon GO Niantic, Inc.
2	+3 Prince of Persia Shadow&Flame Ubisoft Entertainment	+2 Messenger Facebook	Candy Crush Saga King
3	-1 Card Wars - Adventure Time Cartoon Network	-1 Wish - Shopping Made Fun Wish Inc.	Clash Royale Supercell
4	+6 Oei, ik Groei! Domus Technica	+2 Netflix Netflix, Inc.	Tinder Tinder
5	+1 Football Manager Mobile 2017 SEGA	+3 Spotify Music Spotify Ltd.	+2 Google Drive Google Inc.
		-1 [nl.rijksoverheid.digid.pub]	-1 Mobile Strike Epic War
		+2 Snapchat Snapchat Inc	-1 Gardenscapes - New Acres Playrix Games
		+3 Instagram Instagram	+2 Clash of Clans Supercell
		-6 Buienalarm Franklin van Velthuizen	+6 DRAGON BALL Z DOKKAN BATTLE BANDAI NAMCO Entertainment Inc.
		-3 Buienradar Buienradar® B.V.	+2 Summoners War Com2uS
		-1 Last Day on Earth: Survival K-MOBILE	-3 Candy Crush Soda Saga King
		+1 Marktplaats Marktplaats BV	-3 Lords Mobile IGG.COM
		+3 Facebook Facebook	Hay Day Supercell



hogeschool Leiden Denk Doe Voel t

https://www.applyzer.com/?mmenu=worldcharts

Approach

- App selection
- Restrict to 2 platforms:
- Generate user test data
- Analyse application data
- Investigate with commercial tools

iOS

cellebrite
delivering mobile expertise

MAGNET
FORENSICS

XRY

App selection

- Survey NL Police (17 responses). Questions:
  1. Which apps from the NL popular app list occur on mobile phones from suspects?
  2. Which NL apps would be in your top 5 Dutch apps most valuable for investigations?
- Check which permissions apps are using to identify potential use of entity information:
  – i.e. access to contacts, location etc.

App selection results

1. Installed NL apps
   - Nu.nl (news)
   - Buienradar (weather)
   - Buienalarm (weather)
   - Nos (news)

2. Respondents favorites:
   - 9292 (transport)
   - Anwb onderweg (transport)
   - Marktplaats (ebay)
   - Flitsmeister (speeding)
   - Pokemon go

Research experiment

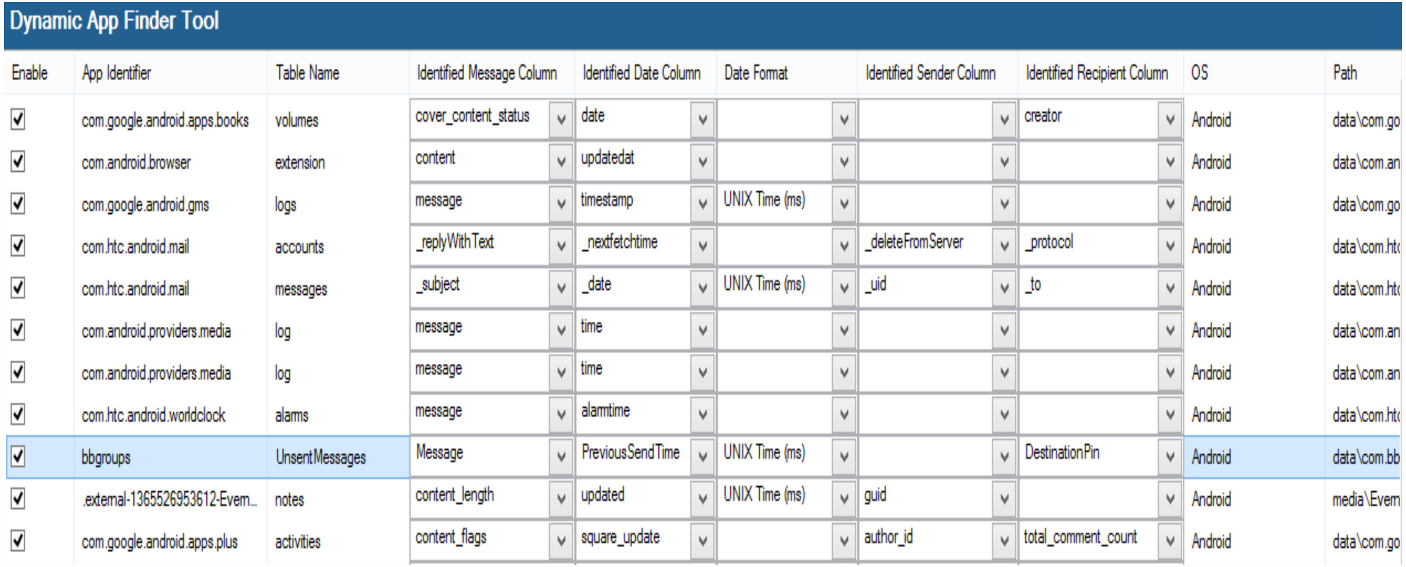
- Generate test data using test phones:
  - Android Samsung Galaxy S5
  - iOS IPhone 5s
- Protocol:
  - Factory reset
  - Fixed test protocol per app, e.g.:
    - Weather app: search current location, search given location, set favorite location, request 14 day weather forecast.

Capabilities in existing tools

	Access to unsupported apps and file formats	Identity or account information	Location information
Cellebrite UFED PA (v 5.2.5.24)	Listed but not accessible	Yes, but not for selected apps	Yes, but not for selected apps
MSAB XRY (v 7.0)	Listed but not accessible	Yes, but not for selected apps	Yes, but not for selected apps
Magnet Forensics IEF (v 6.8.2.3062)	Dynamic app finder for unknown chat apps	Yes, but not for selected apps	Yes. Also discovered for Markplaats

IEF: Dynamic App Finder

- Tries to identify unknown chat databases

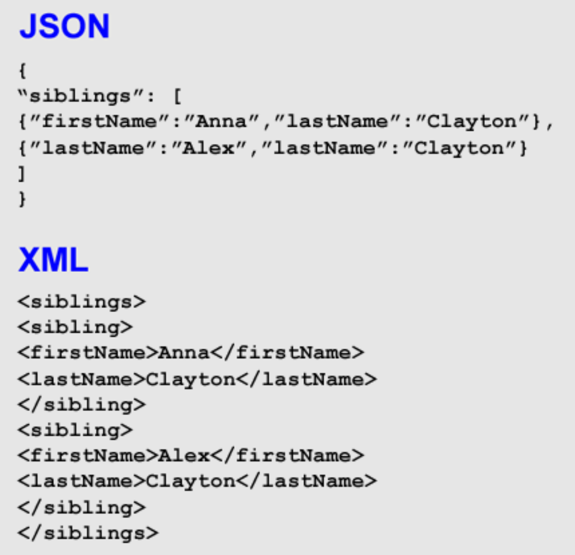


Dynamic App Finder Tool

Enable	App Identifier	Table Name	Identified Message Column	Identified Date Column	Date Format	Identified Sender Column	Identified Recipient Column	OS	Path
☑	com.google.android.apps.books	volumes	cover_content_status	date			creator	Android	data\com.go
☑	com.android.browser	extension	content	updatedat				Android	data\com.an
☑	com.google.android.gms	logs	message	timestamp	UNIX Time (ms)			Android	data\com.go
☑	com.htc.android.mail	accounts	_replyWithText	_nextfetchtime		_deleteFromServer	_protocol	Android	data\com.ht
☑	com.htc.android.mail	messages	_subject	_date	UNIX Time (ms)	_uid	_to	Android	data\com.ht
☑	com.android.providers.media	log	message	time				Android	data\com.an
☑	com.android.providers.media	log	message	time				Android	data\com.an
☑	com.htc.android.worldclock	alarms	message	alarmtime				Android	data\com.ht
☑	bbgroups	UnsentMessages	Message	PreviousSendTime	UNIX Time (ms)		DestinationPin	Android	data\com.bb
☑	.external-1365526953612-Evern...	notes	content_length	updated	UNIX Time (ms)	guid		Android	media\Evern
☑	com.google.android.apps.plus	activities	content_flags	square_update		author_id	total_comment_count	Android	data\com.go

Analysing applicationdata

- Database/storage file types:
  - SQLIte
  - Json
  - Xml



JSON

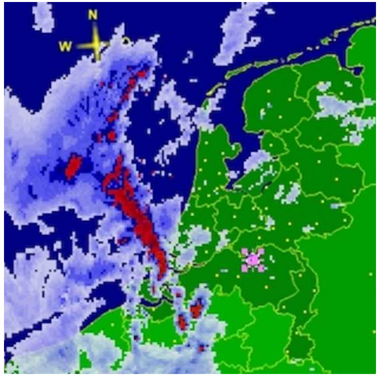
```
{
“siblings”: [
{″firstName″:″Anna″,″lastName″:″Clayton″},
{″lastName″:″Alex″,″lastName″:″Clayton″}
]
}
```

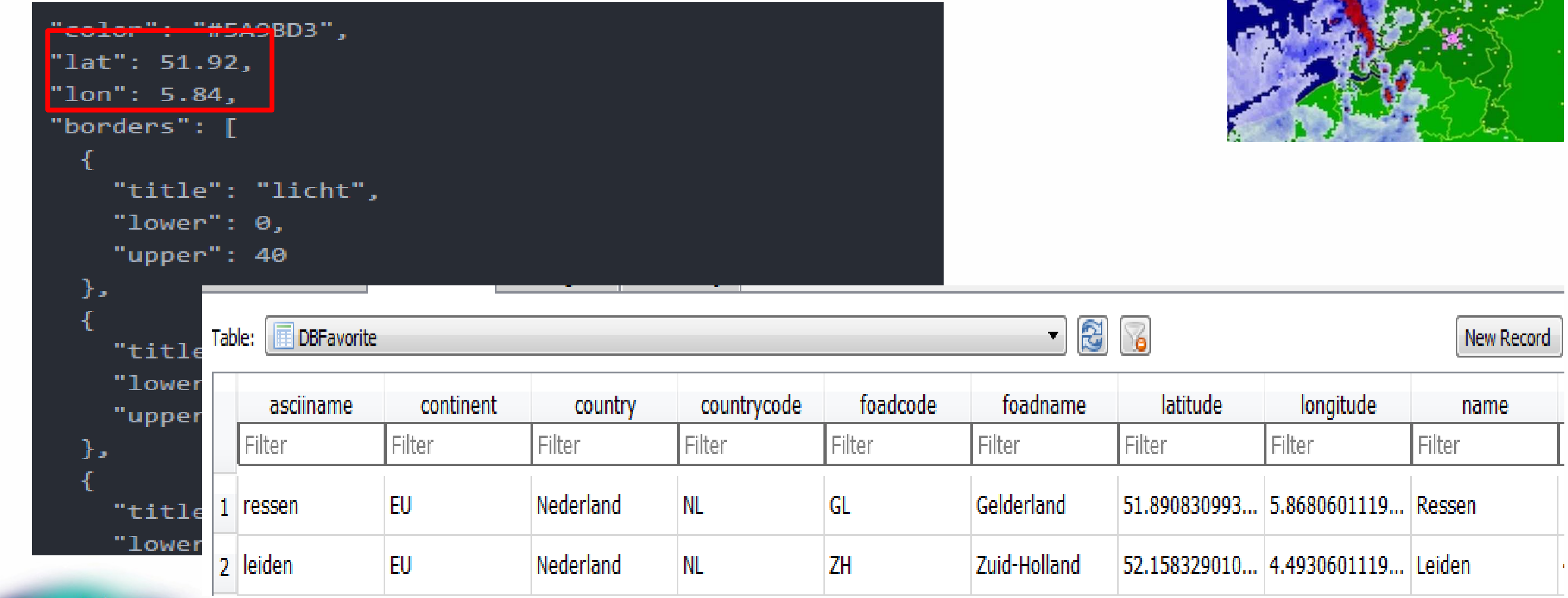
XML

```
<siblings>
<sibling>
<firstName>Anna</firstName>
<lastName>Clayton</lastName>
</sibling>
<sibling>
<firstName>Alex</firstName>
<lastName>Clayton</lastName>
</sibling>
</siblings>
```

- Digital evidence
  - Locations
  - Accounts
  - Searches
  - Timestamps

Digital traces in weather app 1





```
"color": "#5A9BD3",
"lat": 51.92,
"lon": 5.84,
"borders": [
  {
    "title": "licht",
    "lower": 0,
    "upper": 40
  },
  {
    "title
    "lower
    "upper
  },
  {
    "title
    "lower
```

Table: DBFavorite

	asciiname	continent	country	countrycode	foadcode	foadname	latitude	longitude	name
	Filter	Filter	Filter	Filter	Filter	Filter	Filter	Filter	Filter
1	ressen	EU	Nederland	NL	GL	Gelderland	51.890830993...	5.8680601119...	Ressen
2	leiden	EU	Nederland	NL	ZH	Zuid-Holland	52.158329010...	4.4930601119...	Leiden

Digital traces in weather app 2

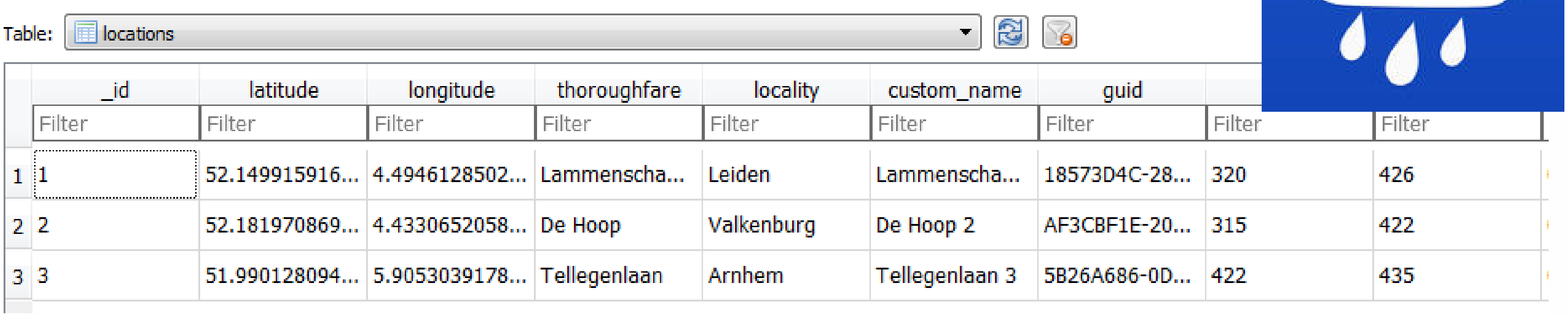


Table: locations

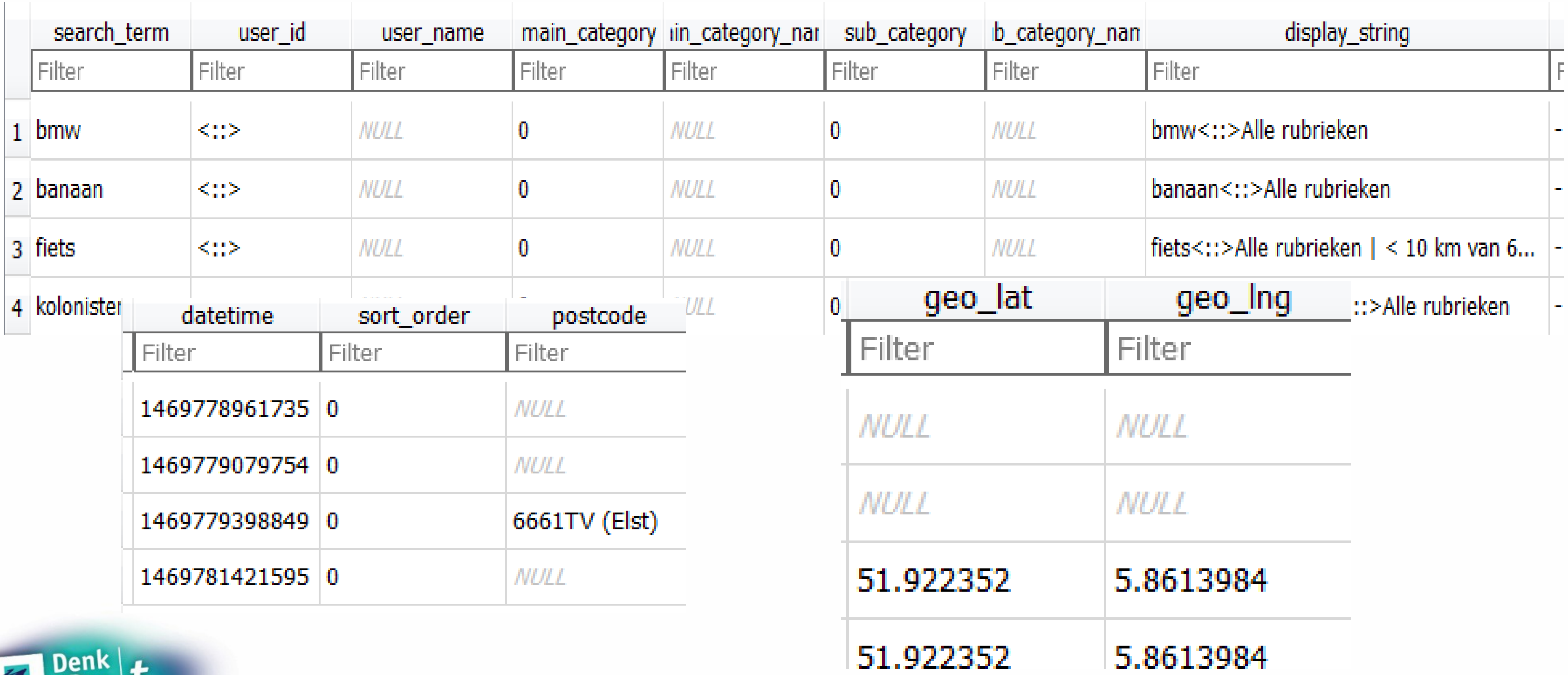
	_id	latitude	longitude	thoroughfare	locality	custom_name	guid		
	Filter	Filter	Filter	Filter	Filter	Filter	Filter	Filter	Filter
1	1	52.149915916...	4.4946128502...	Lammenscha...	Leiden	Lammenscha...	18573D4C-28...	320	426
2	2	52.181970869...	4.4330652058...	De Hoop	Valkenburg	De Hoop 2	AF3CBF1E-20...	315	422
3	3	51.990128094...	5.9053039178...	Tellegenlaan	Arnhem	Tellegenlaan 3	5B26A686-0D...	422	435

Buienalarm_preferences.xml:

```
<string name="last_location_name">Elst</string>
<long name="last_update" value="1469797603271" />
<string name="last_longitude">5.8654683</string>
<string name="last_latitude">51.90449923</string>
```

Digital traces in market place

MARKTPLAATS.NL

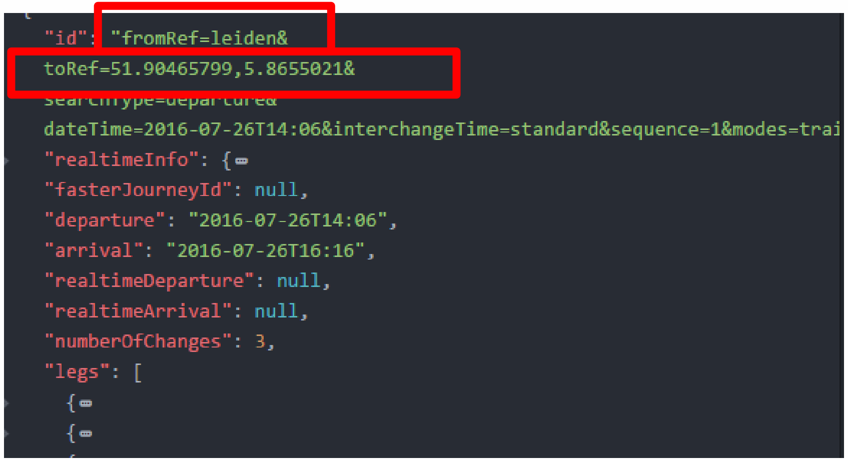


	search_term	user_id	user_name	main_category	ıin_category_nar	sub_category	b_category_nan	display_string	
	Filter	Filter	Filter	Filter	Filter	Filter	Filter	Filter	F
1	bmw	<::>	NULL	0	NULL	0	NULL	bmw<::>Alle rubrieken	-
2	banaan	<::>	NULL	0	NULL	0	NULL	banaan<::>Alle rubrieken	-
3	fiets	<::>	NULL	0	NULL	0	NULL	fiets<::>Alle rubrieken | < 10 km van 6...	-
4	kolonister				ULL	0		::>Alle rubrieken	-

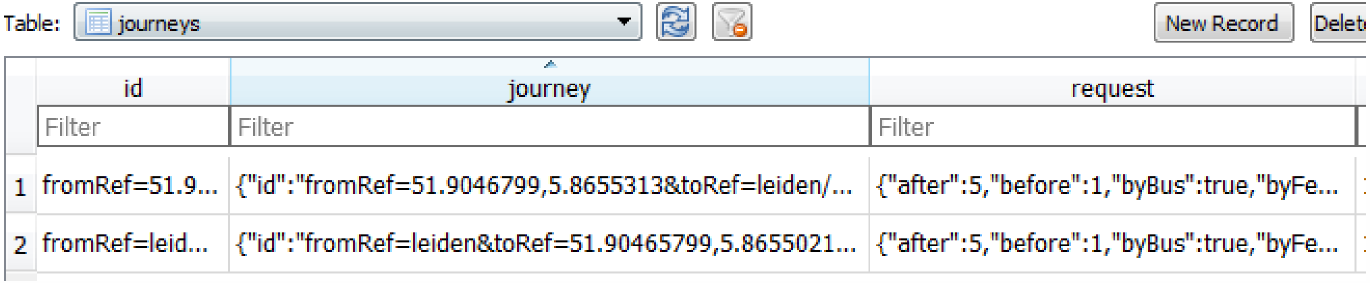
datetime	sort_order	postcode
Filter	Filter	Filter
1469778961735	0	NULL
1469779079754	0	NULL
1469779398849	0	6661TV (Elst)
1469781421595	0	NULL

geo_lat	geo_lng
Filter	Filter
NULL	NULL
NULL	NULL
51.922352	5.8613984
51.922352	5.8613984

Digital traces in route planner



```
{
  "id": "fromRef=leiden&
  toRef=51.90465799,5.8655021&
  searchType=departure&
  dateTime=2016-07-26T14:06&interchangeTime=standard&sequence=1&modes=trai
  "realtimeInfo": {…
  "fasterJourneyId": null,
  "departure": "2016-07-26T14:06",
  "arrival": "2016-07-26T16:16",
  "realtimeDeparture": null,
  "realtimeArrival": null,
  "numberOfChanges": 3,
  "legs": [
    {…
    {…
```

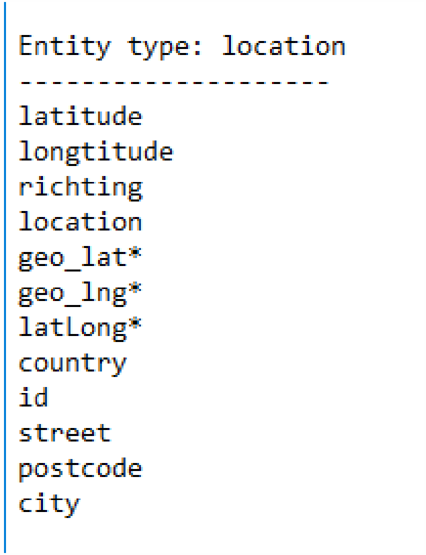


	id	journey	request
	Filter	Filter	Filter
1	fromRef=51.9...	{"id":"fromRef=51.9046799,5.8655313&toRef=leiden/...	{"after":5,"before":1,"byBus":true,"byFe...
2	fromRef=leid...	{"id":"fromRef=leiden&toRef=51.90465799,5.8655021...	{"after":5,"before":1,"byBus":true,"byFe...

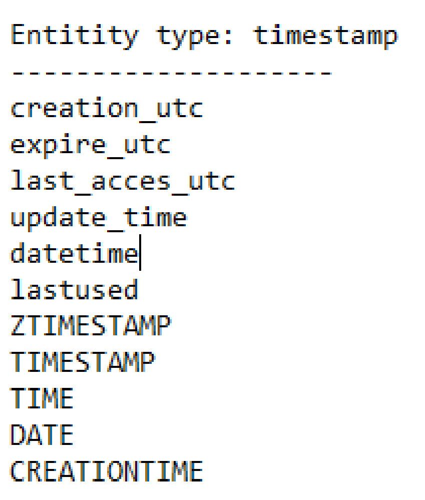
Most relevant entities and their structure

Type	Example
Time stamps	1476829788, 1997-07-16T19:20
Email addresses	blabla@gmail.com
Phone numbers	06 1234 5678
GPS location	51.123455, 4.800928

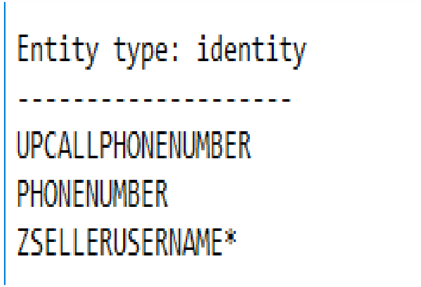
Typical db column headers



```
Entity type: location
--------------------
latitude
longtitude
richting
location
geo_lat*
geo_lng*
latLong*
country
id
street
postcode
city
```



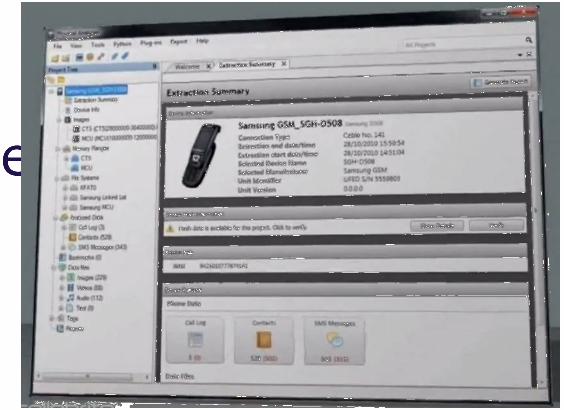
```
Entitity type: timestamp
--------------------
creation_utc
expire_utc
last_acces_utc
update_time
datetime
lastused
ZTIMESTAMP
TIMESTAMP
TIME
DATE
CREATIONTIME
```



```
Entity type: identity
--------------------
UPCALLPHONENUMBER
PHONENUMBER
ZSELLERUSERNAME*
```

Research led to the following plan:

1. Perform data analysis on:
   - Structured file types (for now only SQLite)
   - Look for entities in record fields
   - Look for familiar tag names or column headers
2. Develop a proof of concept:
   - Python script
   - Make use of the UFED PA plugin feature
   - Regular expressions and Keywords
3. Do the experiment



Extraction Summary

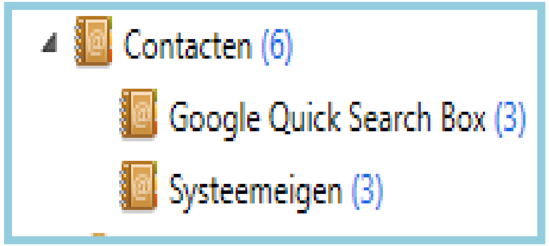
Samsung GSM_SGH-D508

Connection Type	Cable No. 141
Extraction end date/time	28/10/2010 15:59:54
Extraction start date/time	28/10/2010 14:51:04
Selected Device Name	SGH-D508
Selected Manufacturer	Samsung GSM
Unit Identifier	UFED S/N 5559803
Unit Version	0.0.0.0

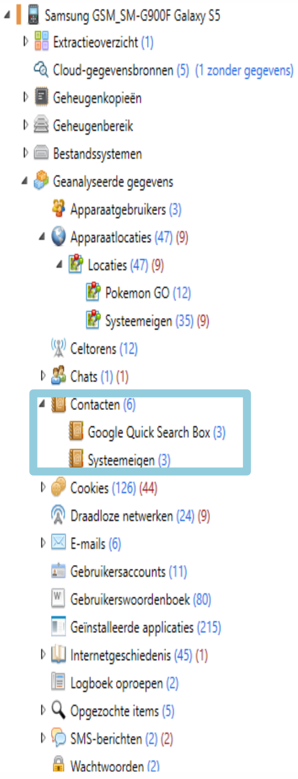
Hash data is available for this project. Click to verify.

UFED presentation: Identified Identies

Without script

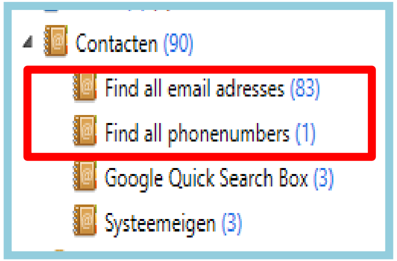


- Contacten (6)
  - Google Quick Search Box (3)
  - Systeemeigen (3)

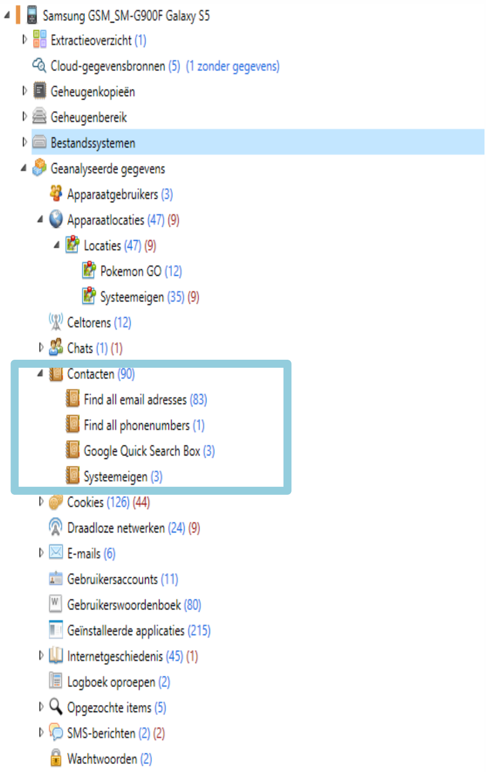


- Samsung GSM_SM-G900F Galaxy S5
  - Extractieoverzicht (1)
  - Cloud-gegevensbronnen (5) (1 zonder gegevens)
  - Geheugenkopieën
  - Geheugenbereik
  - Bestandssystemen
  - Geanalyseerde gegevens
    - Apparaatgebruikers (3)
    - Apparaatlocaties (47) (9)
      - Locaties (47) (9)
        - Pokemon GO (12)
        - Systeemeigen (35) (9)
    - Celtorens (12)
    - Chats (1) (1)
    - Contacten (6)
      - Google Quick Search Box (3)
      - Systeemeigen (3)
    - Cookies (126) (44)
    - Draadloze netwerken (24) (9)
    - E-mails (6)
    - Gebruikersaccounts (11)
    - Gebruikerswoordenboek (80)
    - Geïnstalleerde applicaties (215)
    - Internetgeschiedenis (45) (1)
    - Logboek oproepen (2)
    - Opgezochte items (5)
    - SMS-berichten (2) (2)
    - Wachtwoorden (2)

With script



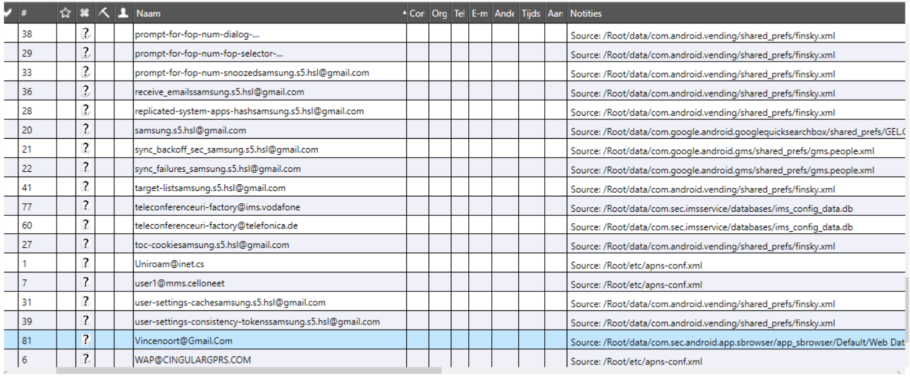
- Contacten (90)
  - Find all email adresses (83)
  - Find all phonenumbers (1)
  - Google Quick Search Box (3)
  - Systeemeigen (3)



- Samsung GSM_SM-G900F Galaxy S5
  - Extractieoverzicht (1)
  - Cloud-gegevensbronnen (5) (1 zonder gegevens)
  - Geheugenkopieën
  - Geheugenbereik
  - Bestandssystemen
  - Geanalyseerde gegevens
    - Apparaatgebruikers (3)
    - Apparaatlocaties (47) (9)
      - Locaties (47) (9)
        - Pokemon GO (12)
        - Systeemeigen (35) (9)
    - Celtorens (12)
    - Chats (1) (1)
    - Contacten (90)
      - Find all email adresses (83)
      - Find all phonenumbers (1)
      - Google Quick Search Box (3)
      - Systeemeigen (3)
    - Cookies (126) (44)
    - Draadloze netwerken (24) (9)
    - E-mails (6)
    - Gebruikersaccounts (11)
    - Gebruikerswoordenboek (80)
    - Geïnstalleerde applicaties (215)
    - Internetgeschiedenis (45) (1)
    - Logboek oproepen (2)
    - Opgezochte items (5)
    - SMS-berichten (2) (2)
    - Wachtwoorden (2)

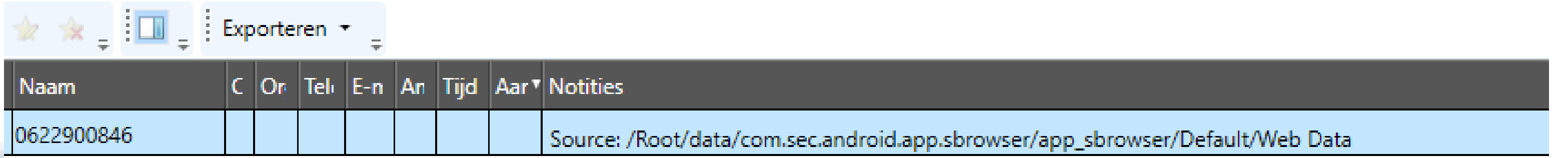
What identities were found?

- New Email addresses:



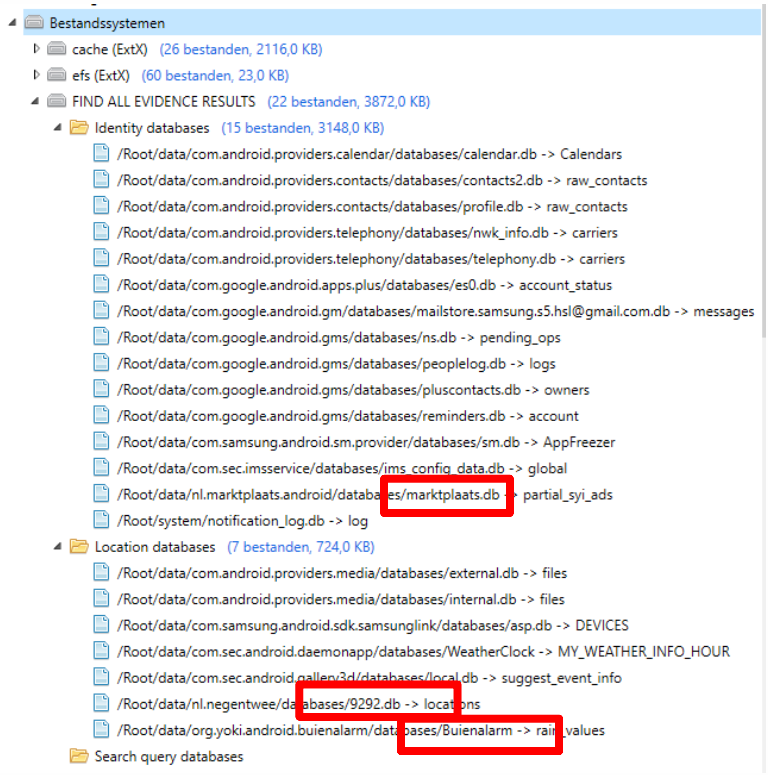
✓	#	☆	✖	⛏	👤	Naam	Cor	Org	Tel	E-m	Ande	Tijds	Aan	Notities
	38		?			prompt-for-fop-num-dialog-...								Source: /Root/data/com.android.vending/shared_prefs/finsky.xml
	29		?			prompt-for-fop-num-fop-selector-...								Source: /Root/data/com.android.vending/shared_prefs/finsky.xml
	33		?			prompt-for-fop-num-snoozedsamsung.s5.hsl@gmail.com								Source: /Root/data/com.android.vending/shared_prefs/finsky.xml
	36		?			receive_emailssamsung.s5.hsl@gmail.com								Source: /Root/data/com.android.vending/shared_prefs/finsky.xml
	28		?			replicated-system-apps-hashsamsung.s5.hsl@gmail.com								Source: /Root/data/com.android.vending/shared_prefs/finsky.xml
	20		?			samsung.s5.hsl@gmail.com								Source: /Root/data/com.google.android.googlequicksearchbox/shared_prefs/GEL.(
	21		?			sync_backoff_sec_samsung.s5.hsl@gmail.com								Source: /Root/data/com.google.android.gms/shared_prefs/gms.people.xml
	22		?			sync_failures_samsung.s5.hsl@gmail.com								Source: /Root/data/com.google.android.gms/shared_prefs/gms.people.xml
	41		?			target-listsamsung.s5.hsl@gmail.com								Source: /Root/data/com.android.vending/shared_prefs/finsky.xml
	77		?			teleconferenceuri-factory@ims.vodafone								Source: /Root/data/com.sec.imsservice/databases/ims_config_data.db
	60		?			teleconferenceuri-factory@telefonica.de								Source: /Root/data/com.sec.imsservice/databases/ims_config_data.db
	27		?			toc-cookiesamsung.s5.hsl@gmail.com								Source: /Root/data/com.android.vending/shared_prefs/finsky.xml
	1		?			Uniroam@inet.cs								Source: /Root/etc/apns-conf.xml
	7		?			user1@mms.celloneet								Source: /Root/etc/apns-conf.xml
	31		?			user-settings-cachesamsung.s5.hsl@gmail.com								Source: /Root/data/com.android.vending/shared_prefs/finsky.xml
	39		?			user-settings-consistency-tokenssamsung.s5.hsl@gmail.com								Source: /Root/data/com.android.vending/shared_prefs/finsky.xml
	81		?			Vincenoort@Gmail.Com								Source: /Root/data/com.sec.android.app.sbrowser/app_sbrowser/Default/Web Dat
	6		?			WAP@CINGULARGPRS.COM								Source: /Root/etc/apns-conf.xml

- New Phone number:



Naam	C	Or	Tel	E-n	An	Tijd	Aar	Notities
0622900846								Source: /Root/data/com.sec.android.app.sbrowser/app_sbrowser/Default/Web Data

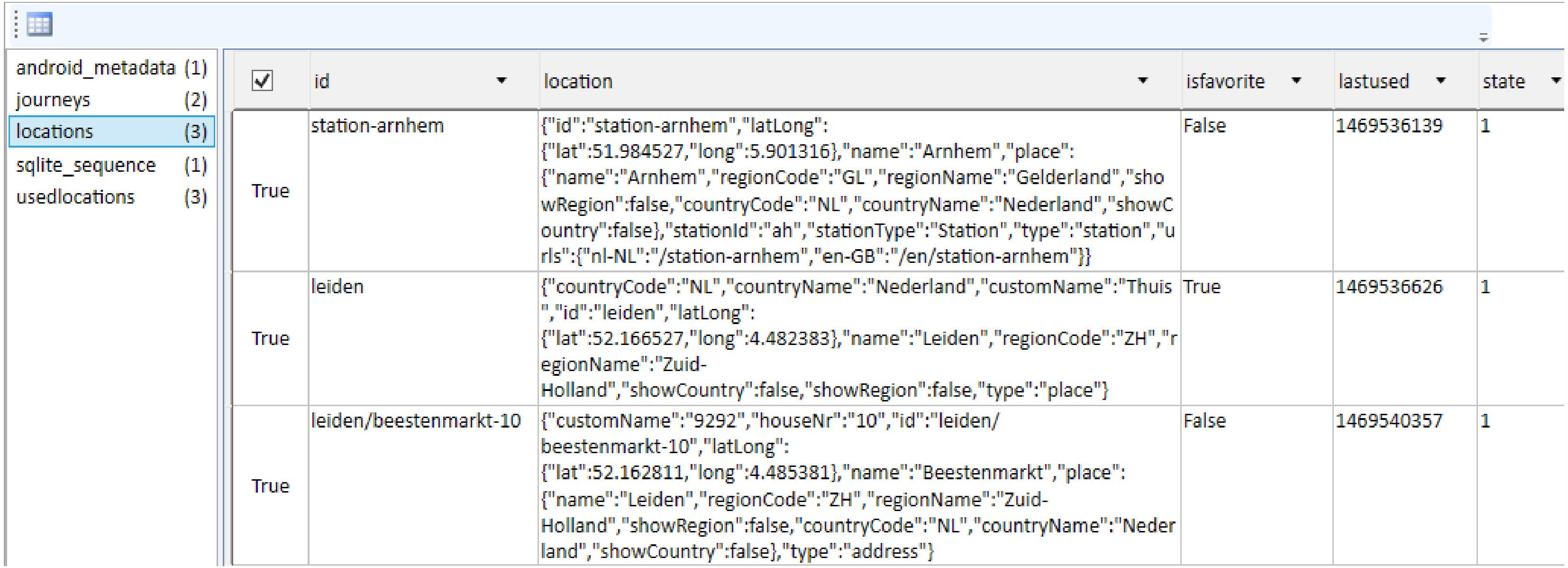
UFED presentation: Identified Databases



```
◢ Bestandssystemen
  ▷ cache (ExtX)  (26 bestanden, 2116,0 KB)
  ▷ efs (ExtX)  (60 bestanden, 23,0 KB)
  ◢ FIND ALL EVIDENCE RESULTS  (22 bestanden, 3872,0 KB)
    ◢ Identity databases  (15 bestanden, 3148,0 KB)
        /Root/data/com.android.providers.calendar/databases/calendar.db -> Calendars
        /Root/data/com.android.providers.contacts/databases/contacts2.db -> raw_contacts
        /Root/data/com.android.providers.contacts/databases/profile.db -> raw_contacts
        /Root/data/com.android.providers.telephony/databases/nwk_info.db -> carriers
        /Root/data/com.android.providers.telephony/databases/telephony.db -> carriers
        /Root/data/com.google.android.apps.plus/databases/es0.db -> account_status
        /Root/data/com.google.android.gm/databases/mailstore.samsung.s5.hsl@gmail.com.db -> messages
        /Root/data/com.google.android.gms/databases/ns.db -> pending_ops
        /Root/data/com.google.android.gms/databases/peoplelog.db -> logs
        /Root/data/com.google.android.gms/databases/pluscontacts.db -> owners
        /Root/data/com.google.android.gms/databases/reminders.db -> account
        /Root/data/com.samsung.android.sm.provider/databases/sm.db -> AppFreezer
        /Root/data/com.sec.imsservice/databases/ims_config_data.db -> global
        /Root/data/nl.marktplaats.android/databases/marktplaats.db -> partial_syi_ads
        /Root/system/notification_log.db -> log
    ◢ Location databases  (7 bestanden, 724,0 KB)
        /Root/data/com.android.providers.media/databases/external.db -> files
        /Root/data/com.android.providers.media/databases/internal.db -> files
        /Root/data/com.samsung.android.sdk.samsunglink/databases/asp.db -> DEVICES
        /Root/data/com.sec.android.daemonapp/databases/WeatherClock -> MY_WEATHER_INFO_HOUR
        /Root/data/com.sec.android.gallery3d/databases/local.db -> suggest_event_info
        /Root/data/nl.negentwee/databases/9292.db -> locations
        /Root/data/org.yoki.android.buienalarm/databases/Buienalarm -> rain_values
      Search query databases
```

- Script inserts findings as a new category under filesystems as FIND ALL EVIDENCE RESULTS

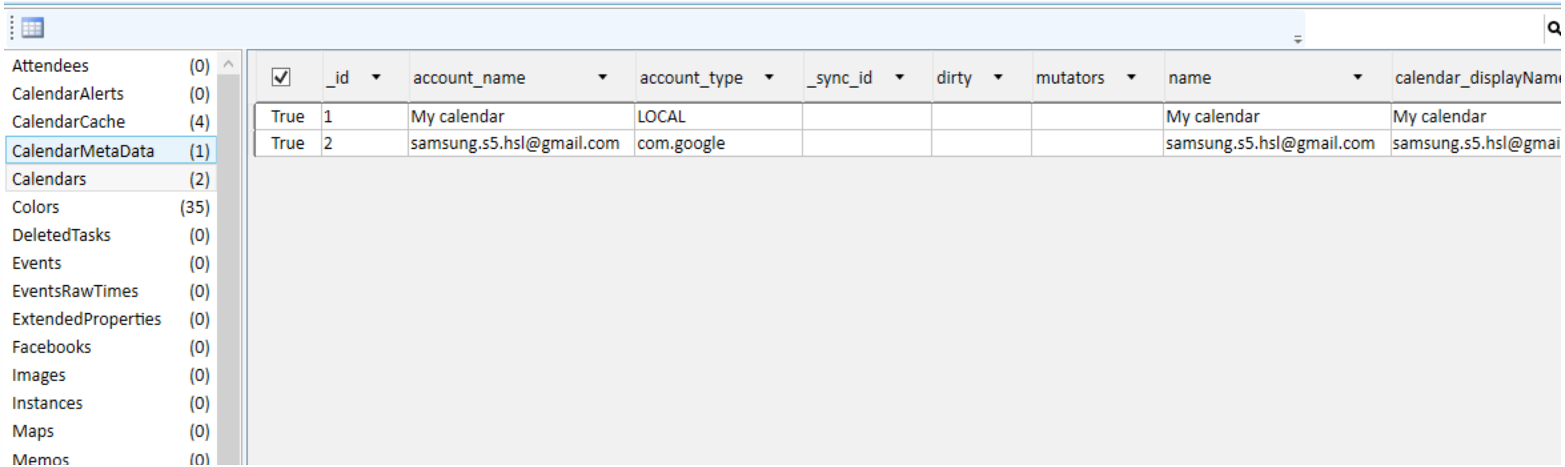
Locations found in 9292 app



android_metadata	(1)
journeys	(2)
locations	(3)
sqlite_sequence	(1)
usedlocations	(3)

☑	id	location	isfavorite	lastused	state
True	station-arnhem	{"id":"station-arnhem","latLong":{"lat":51.984527,"long":5.901316},"name":"Arnhem","place":{"name":"Arnhem","regionCode":"GL","regionName":"Gelderland","showRegion":false,"countryCode":"NL","countryName":"Nederland","showCountry":false},"stationId":"ah","stationType":"Station","type":"station","urls":{"nl-NL":"/station-arnhem","en-GB":"/en/station-arnhem"}}	False	1469536139	1
True	leiden	{"countryCode":"NL","countryName":"Nederland","customName":"Thuis","id":"leiden","latLong":{"lat":52.166527,"long":4.482383},"name":"Leiden","regionCode":"ZH","regionName":"Zuid-Holland","showCountry":false,"showRegion":false,"type":"place"}	True	1469536626	1
True	leiden/beestenmarkt-10	{"customName":"9292","houseNr":"10","id":"leiden/beestenmarkt-10","latLong":{"lat":52.162811,"long":4.485381},"name":"Beestenmarkt","place":{"name":"Leiden","regionCode":"ZH","regionName":"Zuid-Holland","showRegion":false,"countryCode":"NL","countryName":"Nederland","showCountry":false},"type":"address"}	False	1469540357	1

Identity found in Calendar



Attendees	(0)
CalendarAlerts	(0)
CalendarCache	(4)
CalendarMetaData	(1)
Calendars	(2)
Colors	(35)
DeletedTasks	(0)
Events	(0)
EventsRawTimes	(0)
ExtendedProperties	(0)
Facebooks	(0)
Images	(0)
Instances	(0)
Maps	(0)
Memos	(0)

☑	_id	account_name	account_type	_sync_id	dirty	mutators	name	calendar_displayNam
True	1	My calendar	LOCAL				My calendar	My calendar
True	2	samsung.s5.hsl@gmail.com	com.google				samsung.s5.hsl@gmail.com	samsung.s5.hsl@gmai

Note: this information is probably also presented by the standard tools as this is related to Google calendar

Conclusions

- Automated scanning for identity related patterns in mobile phone data including app databases without prior knowledge on table structure, column headers and record content
- Built a python script that serves as a plugin in UFED PA so that output is presented as part of an existing process (if using UFED).
- Approach worked for Dutch mobile phone apps

Recommendations

- Improve location entity extraction, e.g.:
  - Use predefined list of streetnames, cities, countries, continents etc
- More test data is required for extensive testing and to reduce false positives
- Extend beyond SQLITE database to JSON & XML
- Support for other tools besides UFED
  - Currently students working on commandline version with standard python SQLite support

Thank your for your attention

Questions?

Email:
henseler.h@hsleiden.nl

Linked In:
www.linkedin.com/in/henseler

hogeschool Leiden

Denk Doe Voel t