Hash me, als je kan...

Het gebruik van PET's

Johan Loos

Healthcare Information Security en Privacy Consultant

Wie?

- Information en Cybersecurity Professional
- Privacy Professional
- Healthcare security en privacy practitioner (medical devices, healthcare informatics)
- Cryptography (PKI - Certificate Lifecycle Management)
- Artificiële Intelligentie

Over welke data hebben we het

```
MariaDB [twin_peaks_db]> select * from actors_tbl;
+----+------------------+---------------+------+-----------+---------+-----------------+
| ID | name             | phone         | zip  | city      | country | rrn             |
+----+------------------+---------------+------+-----------+---------+-----------------+
| 1  | SHERILYN FENN    | 0032478658791 | 1930 | Zaventem  | Belgium | 69.10.06-124.16 |
| 2  | KYLE MACLACHLAN  | 0032475635741 | 2000 | Antwerpen | Belgium | 75.12.05-137.17 |
| 3  | MADCHEN AMICK    | 0032474635802 | 2200 | Herentals | Belgium | 82.07.02-167.12 |
| 4  | LARA FLYNN BOYLE | 0032471360197 | 3500 | Hasselt   | Belgium | 72.09.21-117.18 |
| 5  | JAMES MARSHALL   | 0032479681932 | 9000 | Gent      | Belgium | 92.08.31-107.11 |
| 6  | JOHAN LOOS       | 0032474896742 | 2500 | Lier      | Belgium | 98.03.13-102.14 |
+----+------------------+---------------+------+-----------+---------+-----------------+
6 rows in set (0.001 sec)
```

Over welke data hebben we het

```
MariaDB [twin_peaks_Db]> select name,rrn from actors_tbl;
+------------------+------------------+
| name             | rrn              |
+------------------+------------------+
| SHERILYN FENN    | 69.10.06-124.16  |
| KYLE MACLACHLAN  | 75.12.05-137.17  |
| MADCHEN AMICK    | 82.07.02-167.12  |
| LARA FLYNN BOYLE | 72.09.21-117.18  |
| JAMES MARSHALL   | 92.08.31-107.11  |
| JOHAN LOOS       | 98.03.13-102.14  |
+------------------+------------------+
6 rows in set (0.001 sec)
```

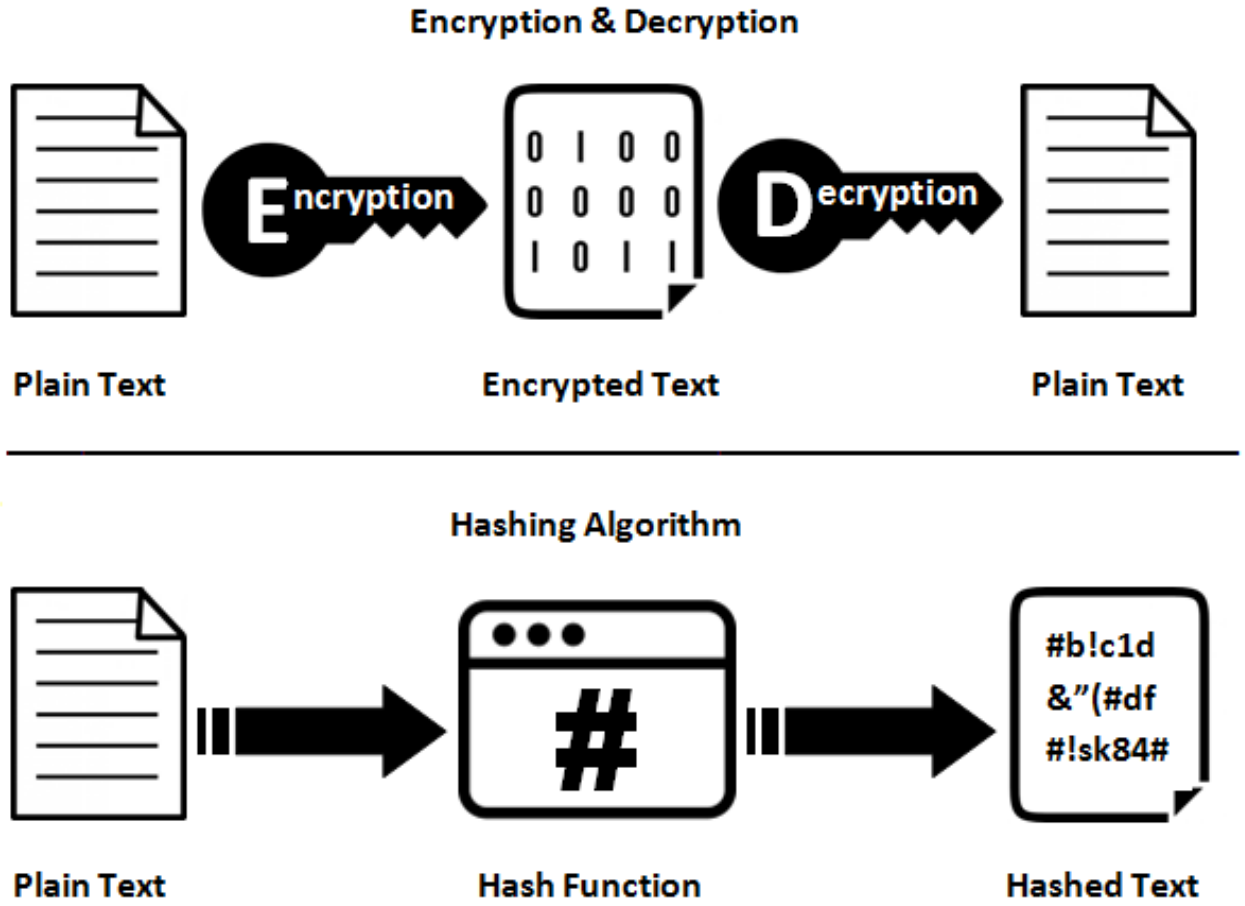
Over welke data hebben we het

```
MariaDB [twin_peaks_db]> select name,rrn from actors_sha3_224_tbl;
+----------------------------------------------------------+----------------------------------------------------------+
| name                                                     | rrn                                                      |
+----------------------------------------------------------+----------------------------------------------------------+
| cce3f3b80fcc7470b632b07f7f2dd1d5e6808650b002ada259627f60 | 16391800422d53e1f6f38ab0bffa00a83d353fcc7a8e6e882e73ae1d |
| 4af88f8eec2364af4ba7a53b8dd9bc8819840a53deb5e81829bb356c | e894f13653c6065fb61652780301921cfc7fb21138b4f26368f795d3 |
| fb2fc9f3fc3323a9600a54e4b2ddbea2406fde194a97e8a67b390643 | d212fca2ea522711b5e7b902e8c7751d88b2d5fd94ecc4a945cff90c |
| 5806c087268a2fe186503366880e2974c61f6b711238daf34e8dc642 | 670cfb6c8c9c4320a746116c05567ee5a412e0ac650a74b4aa2f51ff |
| 908adb707d761e069bdd4583fd43de2b974276611a6f8c503fc5d202 | 00ab7854ac6573c73b51d0a36d78b5d05c67c1dbfe61c2e8e234e42c |
| 5b1b9910742a58966a92f0128db58056bfeee21beeb4164052681058 | d3589f9b5f070bc0e256340f4e19fbd993a75ad2c4c83984d8422571 |
+----------------------------------------------------------+----------------------------------------------------------+
6 rows in set (0.000 sec)
```

Wat is hashing?

- Berekent een **vaste waarde (digest)** onafhankelijk van de grootte van de dataset (bit-stream),
- Grootte van digest hangt af van het gebruikte hash algoritme,
- Deterministisch,
- Doet niets met de originele data,
- Geen sleutel nodig (soms toch niet),
- Originele dataset berekenen vanuit een digest is moeilijk (onmogelijk?),
- Bij verandering van één-bit, andere digest,
- Niet hetzelfde als encryptie.

Verschil tussen hashing en encryptie



Source: https://medium.com/omarelgabrys-blog/diving-into-data-structures-6bc71b2e8f92

Gebruik van hashing?

- Integriteit,
- Bescherming tegen ongeauthoriseerde wijzigingen,
- Paswoorden.

Hash algoritmes

- SHA-2: Gebaseerd op SHA-1 en beschikbaar in verschillende uitvoeringen: SHA2-224,
  SHA2-256,
  SHA2-384,
  SHA2-512
- SHA-3: Verbeterde vorm en beschikbaar in verrschillende uitvoeringen:
  SHA3-224,
  SHA3-256,
  SHA3-384,
  SHA3-512

Hash als unieke identifier

Input voor het berekenen van de digest: annacon

Hash	Output
SHA1	27526e8fdbb10ab517673dbe478ab0730ede9b7b
SHA2-224	24526111ddbf53549e7b987ccd038174454076e2849583933a8df6c7f
SHA2-512	6a076a08146d64f099eee3129cfb0ac67e5452544e797c7f231da903225fb3c44b21e6b5f2637b6bc551b8faa22fda6d5db6dfafd4896e2ae02eee9b91d1d277
SHA3-224	57132796af508f984e50e84c55e1c09a3f2d4499fbbf50a3002170f0
SHA3-512	d8741a45f5c3e03a6d3e58024b6584afbc088836ee06496ecbf285602f205c672e2b18d0ab4ab4a12bbb7095a434771d8fee75aa1131b1a9cbca5d4887250cf4
Blake2b-512	$BLAKE2$5db6293a294e6aa2b340483e3ee901cb03769cc255e32934763834d03eb9eab5fad99bb4306efd18c177436f7d6f77ee3149c3e85080d87dc30dfcd195952ac4

Hash als pseudonimisatie of anonimisatie techniek

Pseudonymization is a method that allows you to switch the original data set (for example, e-mail or a name) with an alias or pseudonym. It is a **reversible process** that de-identifies data but **allows the re-identification** later on if necessary.

Is pseudonymized data still personal data according to the GDPR?

A **pseudonym is still considered to be personal data** according to the GDPR since the ***process is reversible***, and with a ***proper key***, you can identify the individual.

Heridentificatie - Rijksregisternummer

Bestaat uit 11 cijfers:

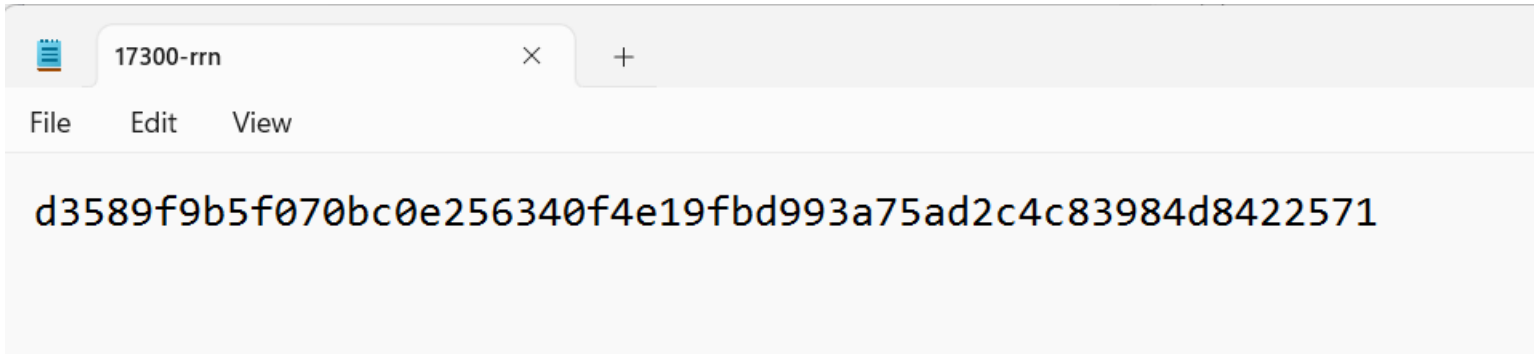
jj.mm.dd-xxx.cc

6 cijfers: geboortedatum

3 cijfers: onderscheidt tussen personen die dezelfde dag geboren zijn (dagteller)

2 cijfers: controlegetal

Heridentificatie - Rijksregisternummer



hashcat -m 17300 -a 3 “d:\tools\pet\hashcat\rrn\17300-rrn.txt” ?d?d.?d?d.?d?d-?d?d?d.?d?d

Mode 17300 : SHA3-224

Heridentificatie - Rijksregisternummer

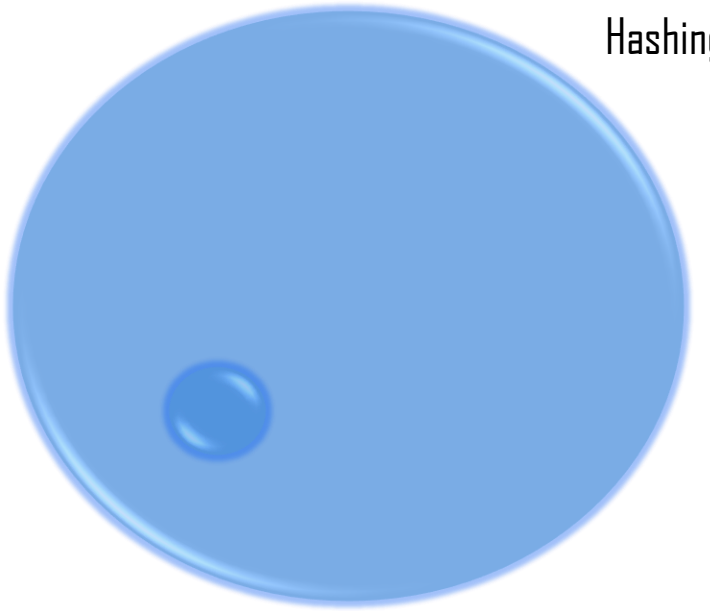
```
Session..........: hashcat
Status...........: Running
Hash.Mode........: 17300 (SHA3-224)
Hash.Target......: d3589f9b5f070bc0e256340f4e19fbd993a75ad2c4c83984d8422571
Time.Started.....: Wed Oct 11 12:49:16 2023 (1 sec)
Time.Estimated...: Wed Oct 11 12:53:20 2023 (4 mins, 3 secs)
Kernel.Feature...: Optimized Kernel
Guess.Mask.......: ?d?d.?d?d.?d?d-?d?d?d.?d?d [15]
Guess.Queue......: 1/1 (100.00%)
Speed.#1.........:   386.2 MH/s (12.16ms) @ Accel:128 Loops:62 Thr:32 Vec:1
Speed.#3.........: 23231.4 kH/s (7.48ms) @ Accel:4 Loops:3 Thr:512 Vec:1
Speed.#*.........:   409.6 MH/s
Recovered........: 0/1 (0.00%) Digests (total), 0/1 (0.00%) Digests (new)
Progress.........: 314441728/100000000000 (0.31%)
Rejected.........: 0/314441728 (0.00%)
Restore.Point....: 0/100000000 (0.00%)
Restore.Sub.#1...: Salt:0 Amplifier:620-682 Iteration:0-62
Restore.Sub.#3...: Salt:0 Amplifier:273-276 Iteration:0-3
Candidate.Engine.: Device Generator
Candidates.#1....: 11.43.67-206.12 -> 03.48.21-960.12
Candidates.#3....: 38.21.09-278.12 -> 46.28.33-420.12
Hardware.Mon.#1..: Temp: 45c Util: 97% Core:1447MHz Mem:5000MHz Bus:8
Hardware.Mon.#3..: N/A
```

Heridentificatie - Rijksregisternummer

```
d3589f9b5f070bc0e256340f4e19fbd993a75ad2c4c83984d8422571:98.03.13-102.14

Session..........: hashcat
Status...........: Cracked
Hash.Mode........: 17300 (SHA3-224)
Hash.Target......: d3589f9b5f070bc0e256340f4e19fbd993a75ad2c4c83984d8422571
Time.Started.....: Wed Oct 11 12:49:16 2023 (2 mins, 14 secs)
Time.Estimated...: Wed Oct 11 12:51:30 2023 (0 secs)
Kernel.Feature...: Optimized Kernel
Guess.Mask.......: ?d?d.?d?d.?d?d-?d?d?d.?d?d [15]
Guess.Queue......: 1/1 (100.00%)
Speed.#1.........:   284.8 MH/s (17.11ms) @ Accel:128 Loops:62 Thr:32 Vec:1
Speed.#3.........: 31450.9 kH/s (5.44ms) @ Accel:4 Loops:3 Thr:512 Vec:1
Speed.#*.........:   316.2 MH/s
Recovered........: 1/1 (100.00%) Digests (total), 1/1 (100.00%) Digests (new)
Progress.........: 50555486208/100000000000 (50.56%)
Rejected.........: 0/50555486208 (0.00%)
Restore.Point....: 49692672/100000000 (49.69%)
Restore.Sub.#1...: Salt:0 Amplifier:0-62 Iteration:0-62
Restore.Sub.#3...: Salt:0 Amplifier:588-591 Iteration:0-3
Candidate.Engine.: Device Generator
Candidates.#1....: 12.07.13-465.14 -> 04.04.06-628.14
Candidates.#3....: 73.95.20-367.14 -> 16.94.80-222.14
Hardware.Mon.#1..: Temp: 74c Util: 97% Core: 990MHz Mem:5000MHz Bus:8
Hardware.Mon.#3..: N/A
```

Probleem – Brute Force



1/ Hash algoritme is ontwikkeld om snel te zijn. Betere hardware beschikbaar en meer bewerkingen. Dit wil ook zeggen dat brute force ook snel zal zijn.

2/ In welke mate zijn de huidige hash algoritmes minder onderhevig aan snellere hardware.

Hardware – Brute Force

RRN	RTX A2000		RTX 4090	
	Hash	Time	Hash	Time
SHA1	2440 M	36 sec	29982 M	2 sec
SHA2-512	475 M	3 min 19 sec	5700 M	15 sec
SHA3-512	367 M	4 min 10 sec	4324 M	18 sec
Blake2b	377 M	3 min 51 sec	2704 M	30 sec
PBKDF2-HMAC_256	668 K	1d 14 uur	8290 K	3 uur 21 min
PBKDF2-HMAC_512	236 K	4d 15 uur	2951 K	9 uur 24 min

Wat kunnen we hieraan doen?

- Heridentificatie analyse,
- Algoritme kwetsbaar voor brute force,
- Salting (personal data + salt = hash),
- Cost factor verhogen,
- Encryptie en dan hashing.


UM...
I HAVE ADDITIONAL QUESTIONS