JENS NEUHALFEN

COMPARE DATA

GOOD CRYPTOGRAPHY


40

Problem: Securely compare two data items

Solution: Normalise & hash data, compare hashes


LOREM IPSUM . . .
==
LOREM IPSUM . . .


=> sha256(LOREM IPSUM ...) == sha256(LOREM IPSUM ...)

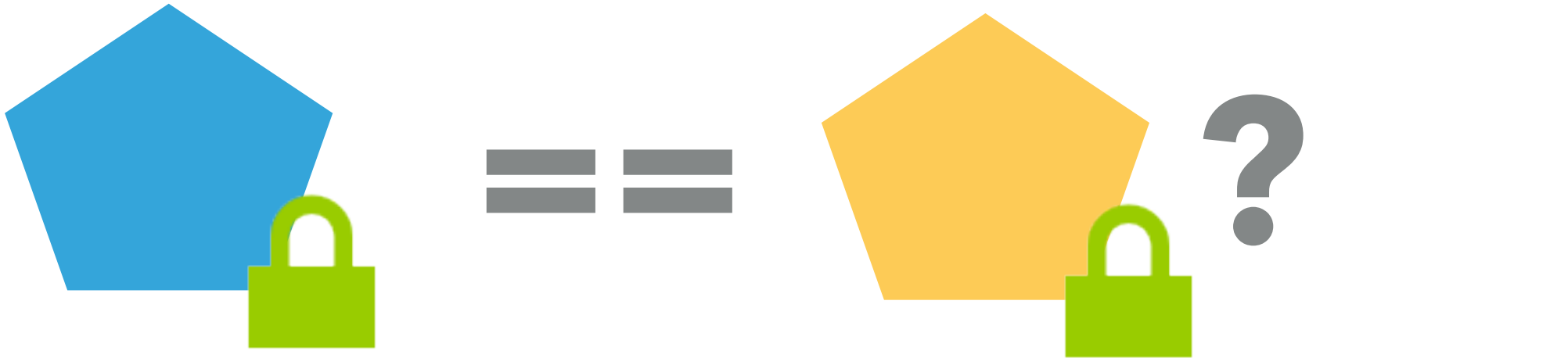
<=>

4C53E9C9...

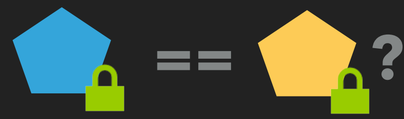
==

4C53E9C9...

- Collisions [**A != B** but **sha256(A) == sha256(B)**] are mathematically possible, but practically not relevant



COMPARE DATA



Problem: Securely compare two data items

Solution: Normalise & hash data, compare hashes

LOREM IPSUM … == LOREM IPSUM …

=>★ sha256(LOREM IPSUM …) == sha256(LOREM IPSUM …)

<=> 4C53E9C9… == 4C53E9C9…

★ Collisions [**A != B** but **sha256(A) == sha256(B)**] are mathematically possible, but practically not relevant

COMPARE DATA

Problem: Securely compare two data items

Solution: Normalise & hash data, compare hashes

1 - Normalize

E.g. J. EDGAR HOOVER → HOOVER, JOHN EDGAR →* H160, J500 E326

2 - Hash

Use *hash(salt + data)* to prevent precomputing attacks. Use multiple iterations of hashing.

- *public salt* => treat hash as *pseudonymised*
- *secret salt*=> treat hash as *anonymised*

* *Soundex - but choose whatever normalisation works for you*

HASHING DOES NOT INCREASE THE ENTROPY! (THINK OF HASHED PHONE NUMBERS)