Executing private data operations using Homomorphic Encryption

Dr. Gianpiero Costantino

Consiglio
Nazionale delle
Ricerche

Outline

- **Cloud**
  - Introduction on Cloud Providers.
  - Privacy concern.
  - Marketing solutions.
- **HC@WORKS Project**
  - What is.
  - Homomorphic Encryption.
  - Tweet Analysis case study.
- **Conclusion**

Cloud Providers

- The **Cloud** is a convenient place to store our files and get them back from any places and any devices;

Dropbox

Free up to 2GB

OneDrive

Free up to 15GB

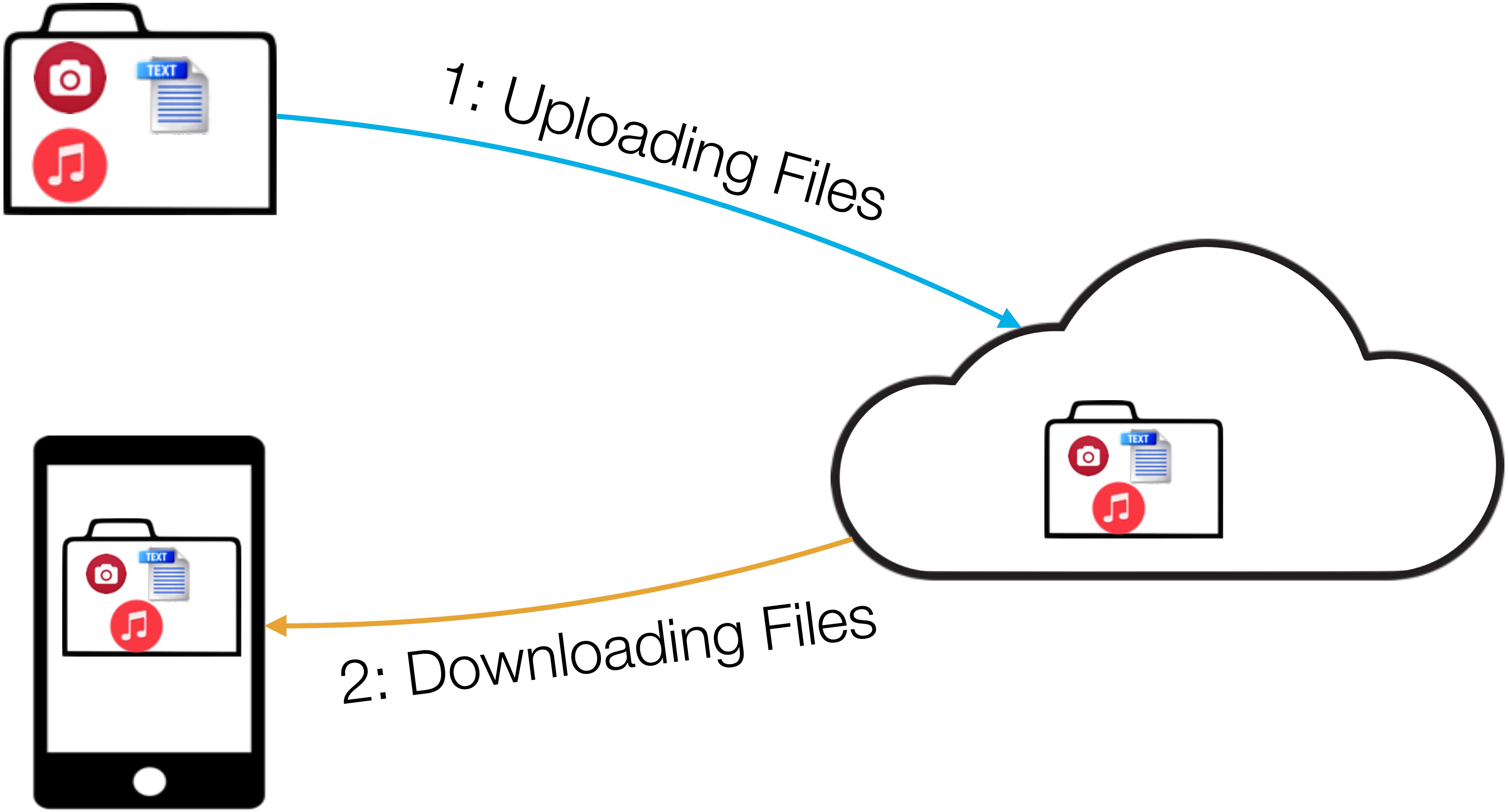
Google Drive

Free up to 15GB

SugarSync

100GB for 7.49$/month

Cloud Providers \ How do they work?



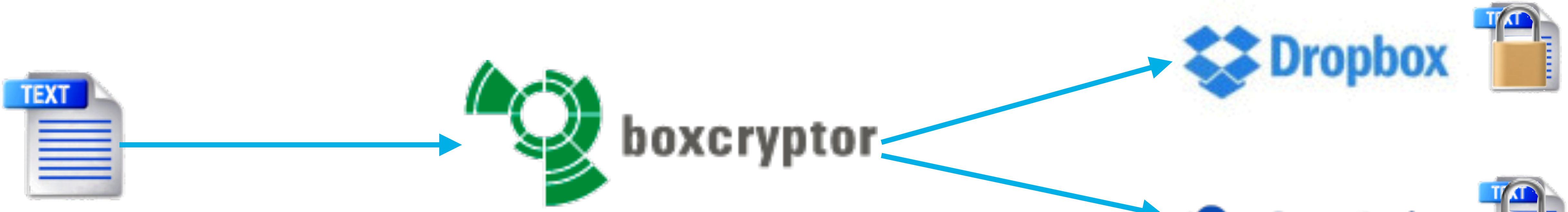
Cloud Providers \ Privacy

- Cloud Providers like Dropbox, Google Drive, etc upload our file into the cloud “in clear”.
- Sensitive files can be easily accessed by the cloud provider that we use to store our files.

Cloud Provider \ Privacy \ Market Solutions

boxcryptor is not a cloud provider, but it allows you to encrypt your files before uploading them;

48$/year



1: Clear Text File

2: It encrypts the Text File

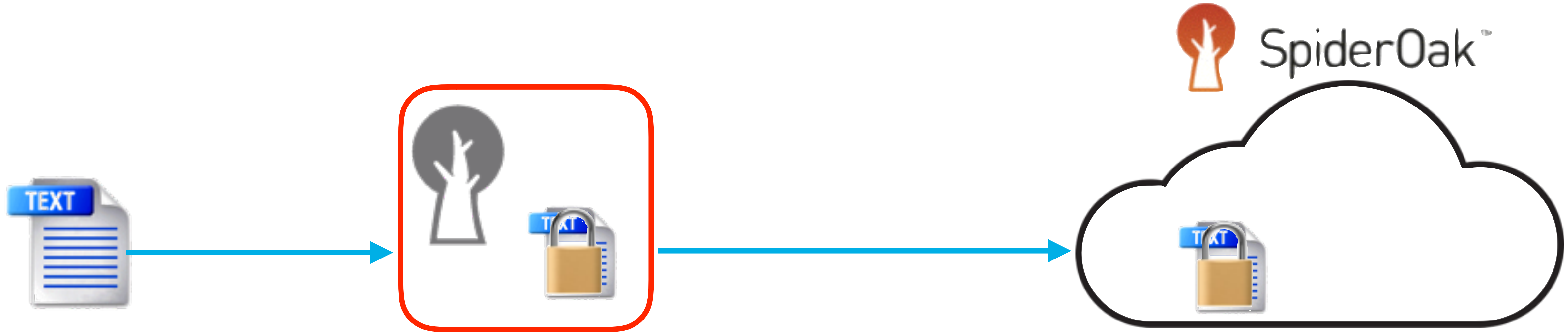
3: Encrypted Text File into your cloud providers

Cloud Provider \ Privacy \ Market Solutions

SpiderOak™

is a cloud provider that offers *“Zero Knowledge”* feature;

1TB per 12$/month



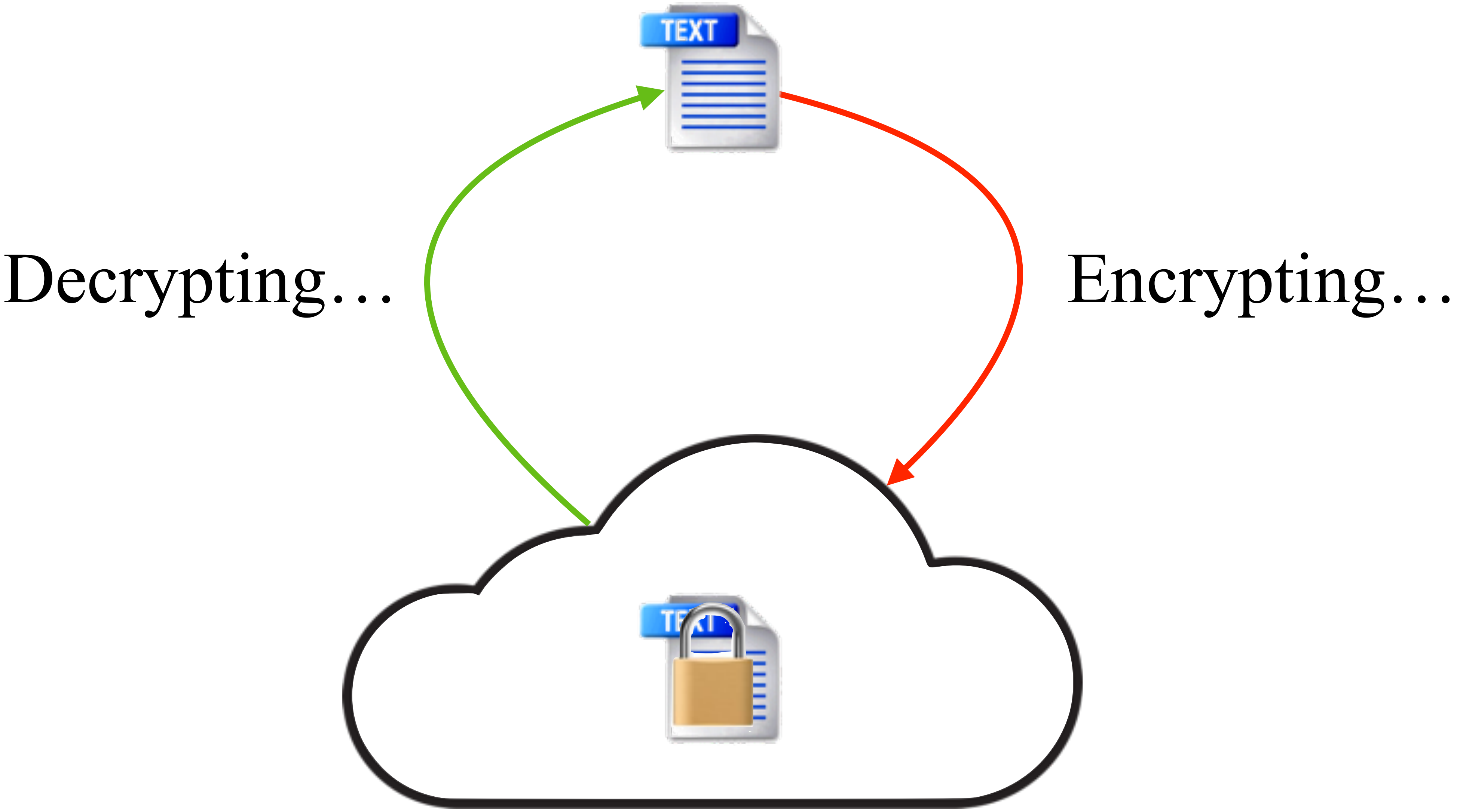
1: Clear Text File

2: The SpiderOak client encrypts the file

3: Encrypted Text File in the cloud

Cloud Provider \ Privacy \ Market Solutions

- The Loop:



Cloud Provider \ Privacy \ Other Solutions

- The question is:

Can I use a different solution to protect my data stored in the cloud?

Working o[illegible]pted data?!?!

HC@WORKS Project

- HC@WORKS is an EIT DIGITAL Project that lasts 1year;
- It aims at showing the feasibility of the **Homomorphic Encryption** for three use cases:
  - eHealth;
  - **Tweets analysis**;
  - Packet Inspection;

HC@WORKS Project \ Partners

DE LA RECHERCHE À L'INDUSTRIE
cea

Consiglio
Nazionale delle
Ricerche

AtoS

ENGINEERING

THALES

HC@WORKS Project \ Homomorphic Encryption

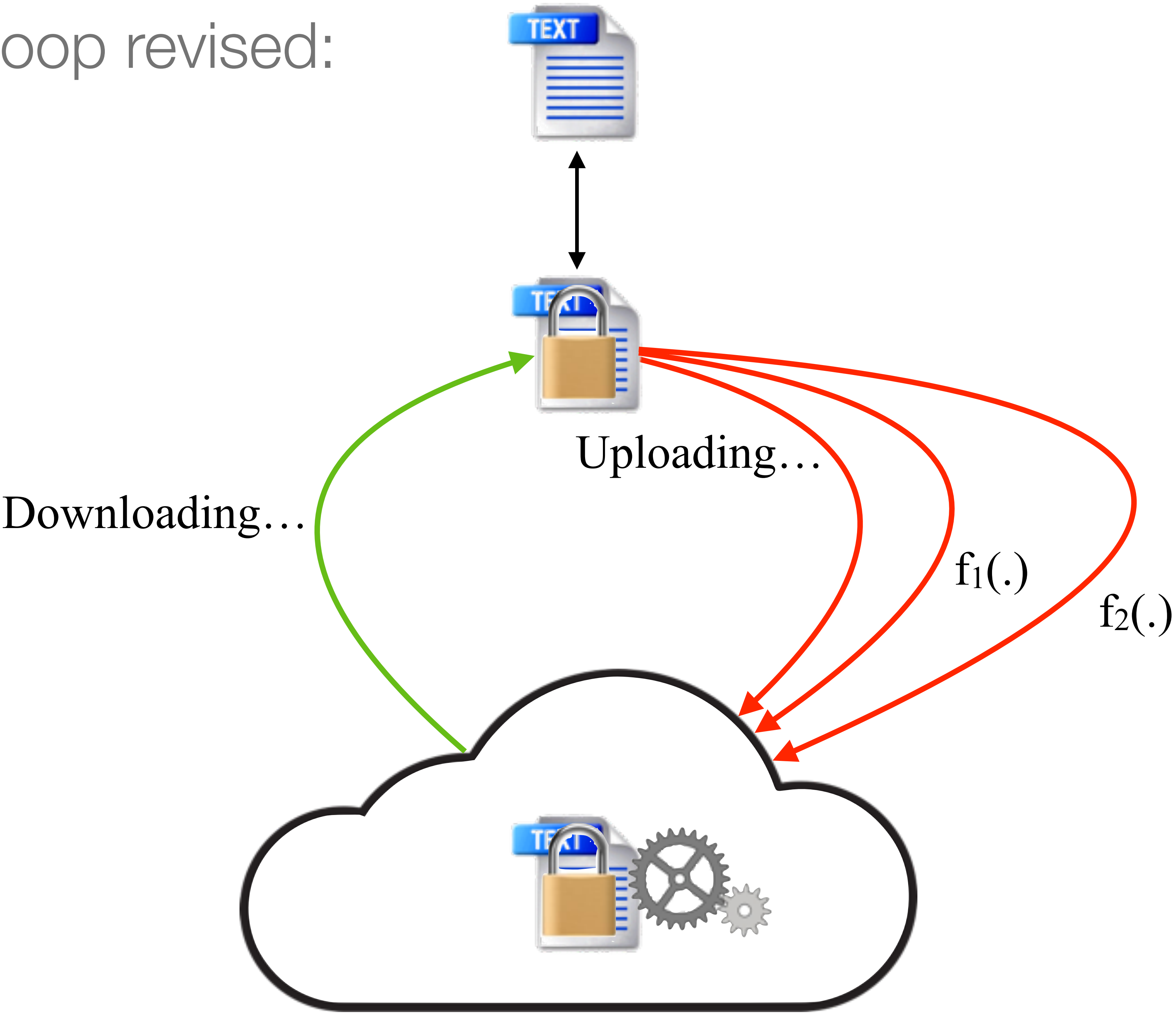
- Generically speaking, **Homomorphic Encryption** (HE) is an encryption schema to perform computations on cipher-texts;

- Basically, you work on encrypted data;

- The result of the function that you execute on encrypted data is the same of the result with the same function using clear-texts;

HC@WORKS Project \ Homomorphic Encryption

- The Loop revised:



HC@WORKS Project \ Tweets analysis

- The case study aims at analysing tweets from Twitter to find messages that belong to a specific context:

t_1

The workshop on security frameworks is great!

t_2

I am ready to shoot with my gun!

Terrorist Template = {bomb, killer, gun, shoot, …}

HC@WORKS Project \ Tweets analysis

t_2

I am ready to shoot with my gun!

- Then, we calculate the **Risk Factor (RF),** which is a simple function that takes a tweet, a template and gives the risk for that tweet;

$$RF_{t1} = 0;$$

$$RF_{t2} = \mathbf{2};$$

- Finally, an investigator queries a DB, to retrieve all RF higher than a threshold;

HC@WORKS Project \ Tweets analysis

- **Very easy so far**, but we want calculate the risk factor by preserving the users' tweets privacy!!!

t_1

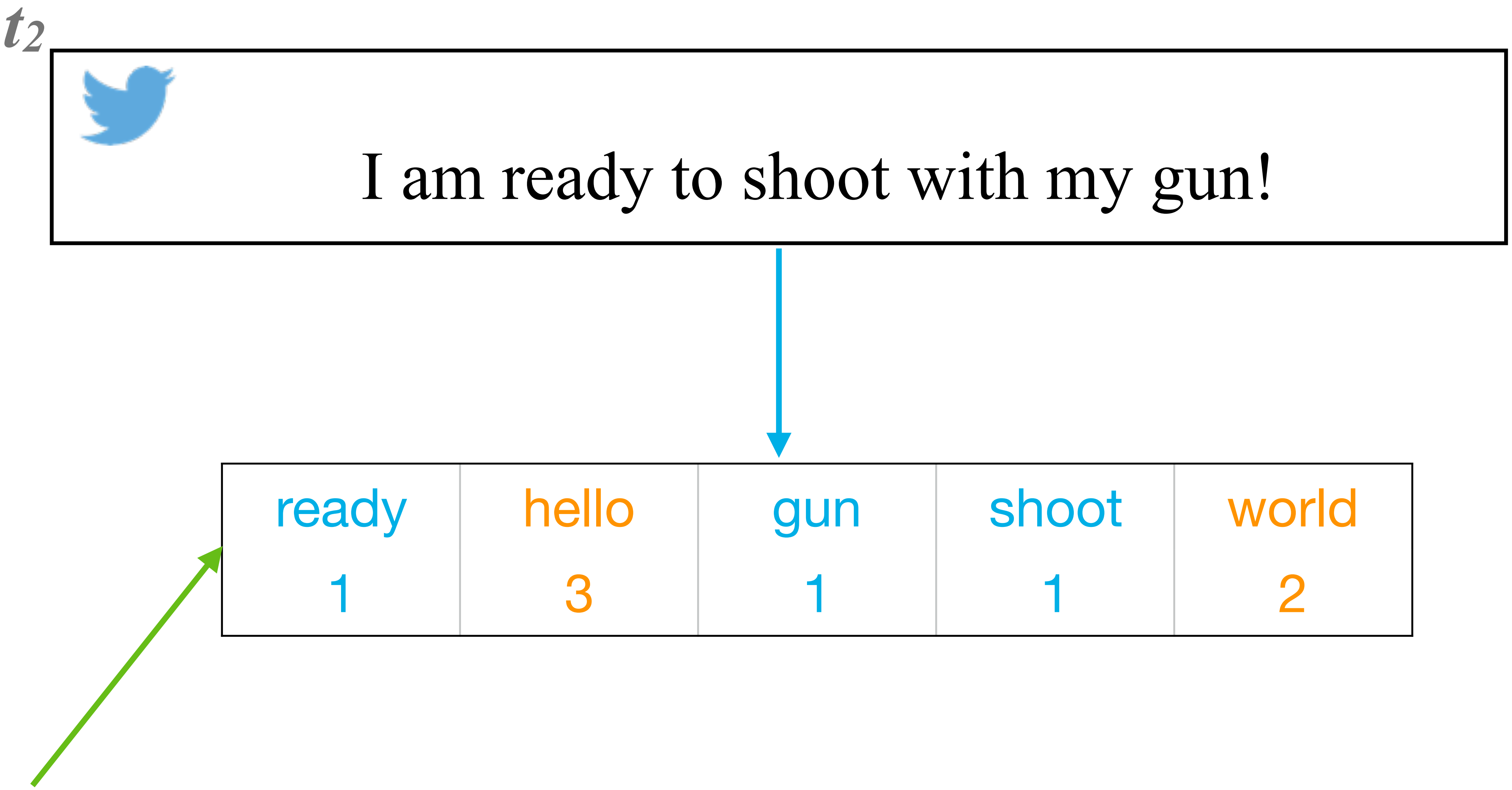
cqwe3è45à13r32189qàwecàwe22!

t_2

achfn<6&6w6364£°FDSAF3afdfswwjerj

HC@WORKS Project \ Tweets analysis with HE

Preprocessing Phase:



t_2

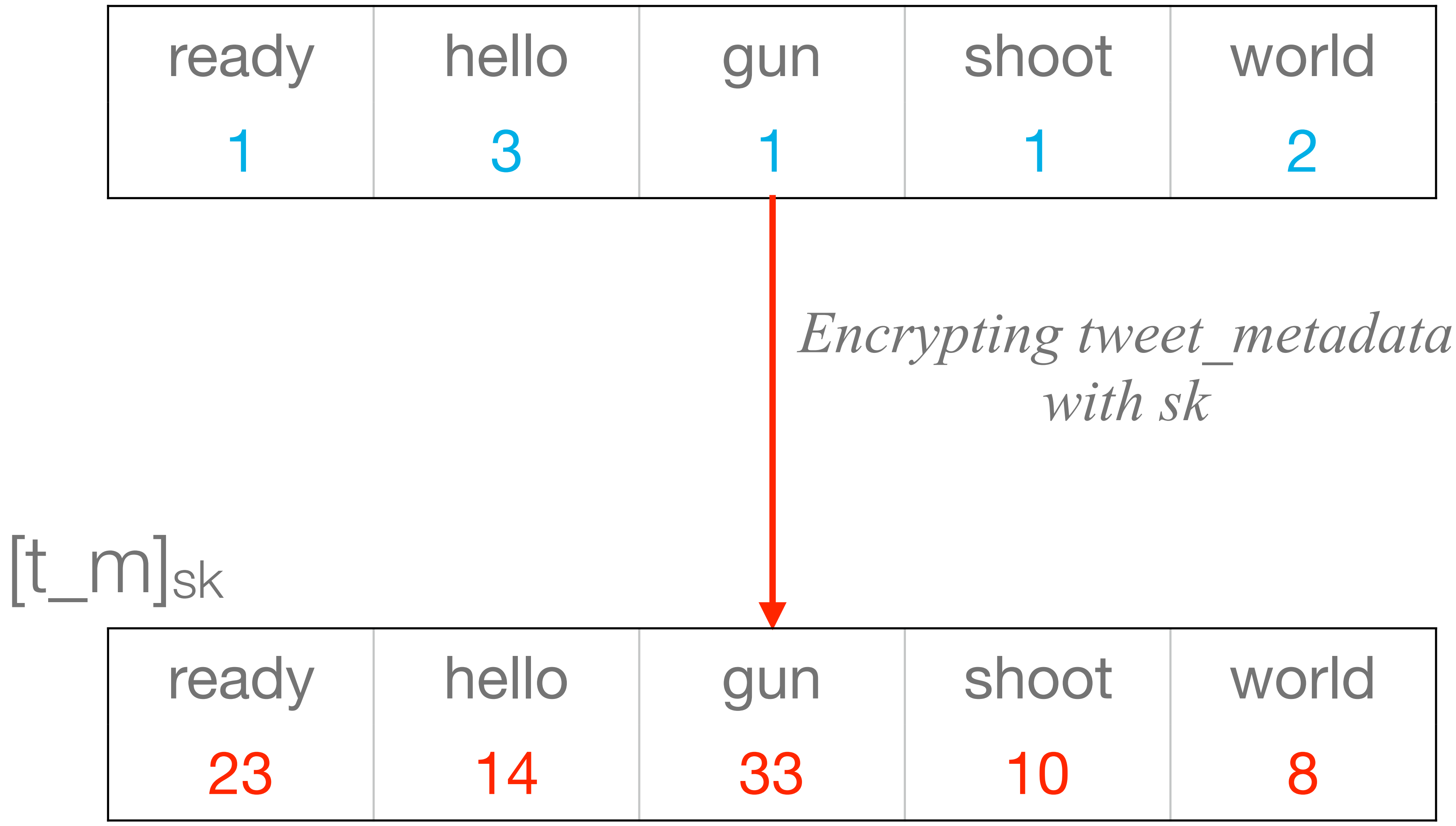
I am ready to shoot with my gun!

ready	hello	gun	shoot	world
1	3	1	1	2

Call it **tweet_metadata**

HC@WORKS Project \ Tweets analysis with HE

Preprocessing Phase:



ready	hello	gun	shoot	world
1	3	1	1	2

Encrypting tweet_metadata with sk

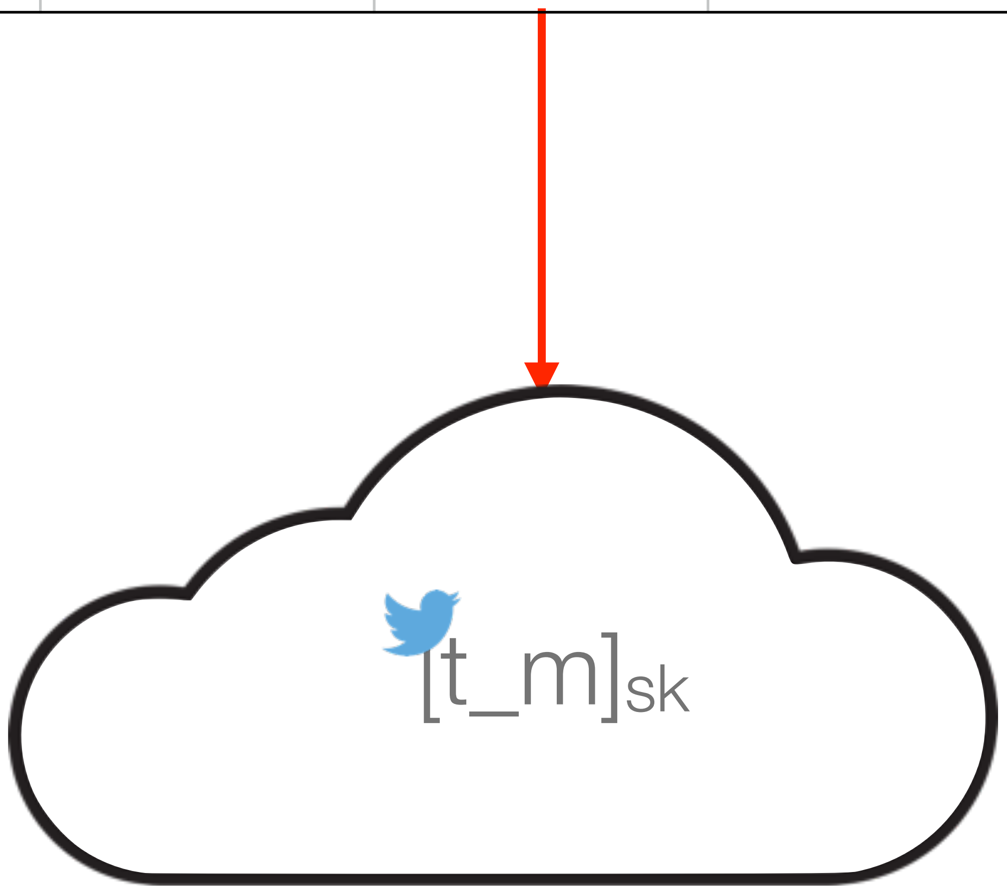
$[t\_m]_{sk}$

ready	hello	gun	shoot	world
23	14	33	10	8

HC@WORKS Project \ Tweets analysis with HE

Uploading Phase:

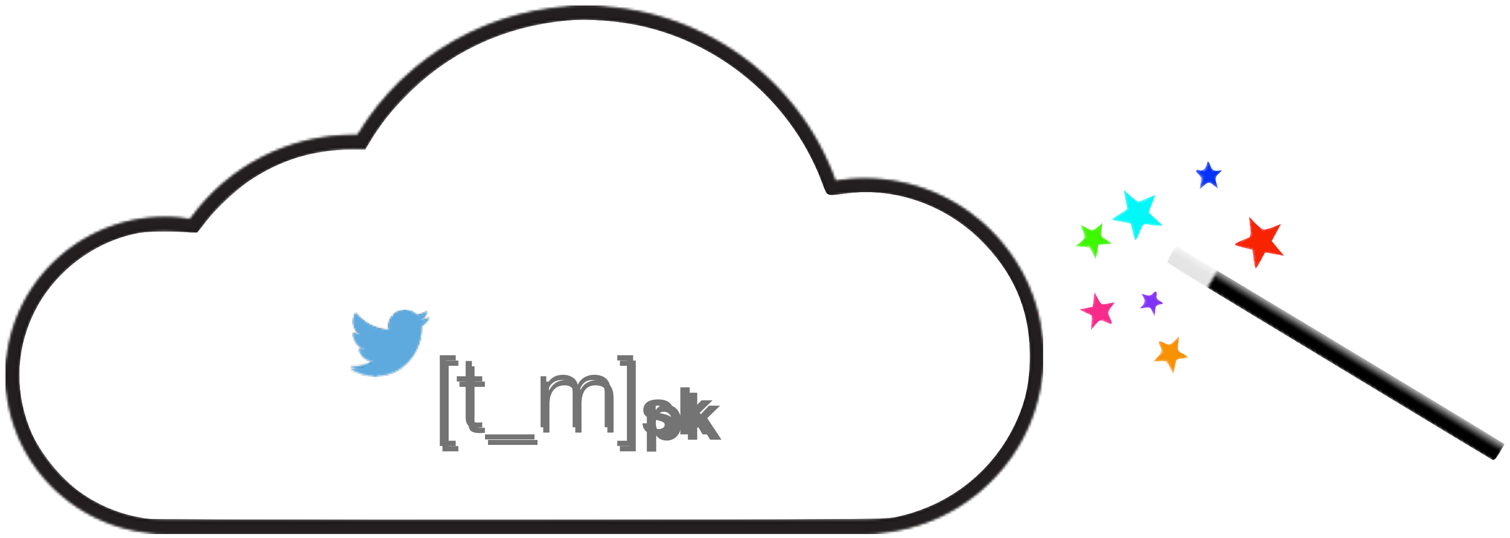
ready	hello	gun	shoot	world
23	14	33	10	8



$[t\_m]_{sk}$

HC@WORKS Project \ Tweets analysis with HE

Transencrypting Phase:



pk is the public key of the investigator

HC@WORKS Project \ Tweets analysis with HE

Analysis Phase:

$[t\_m]_{pk}$

ready	hello	gun	shoot	world
23	14	33	10	8

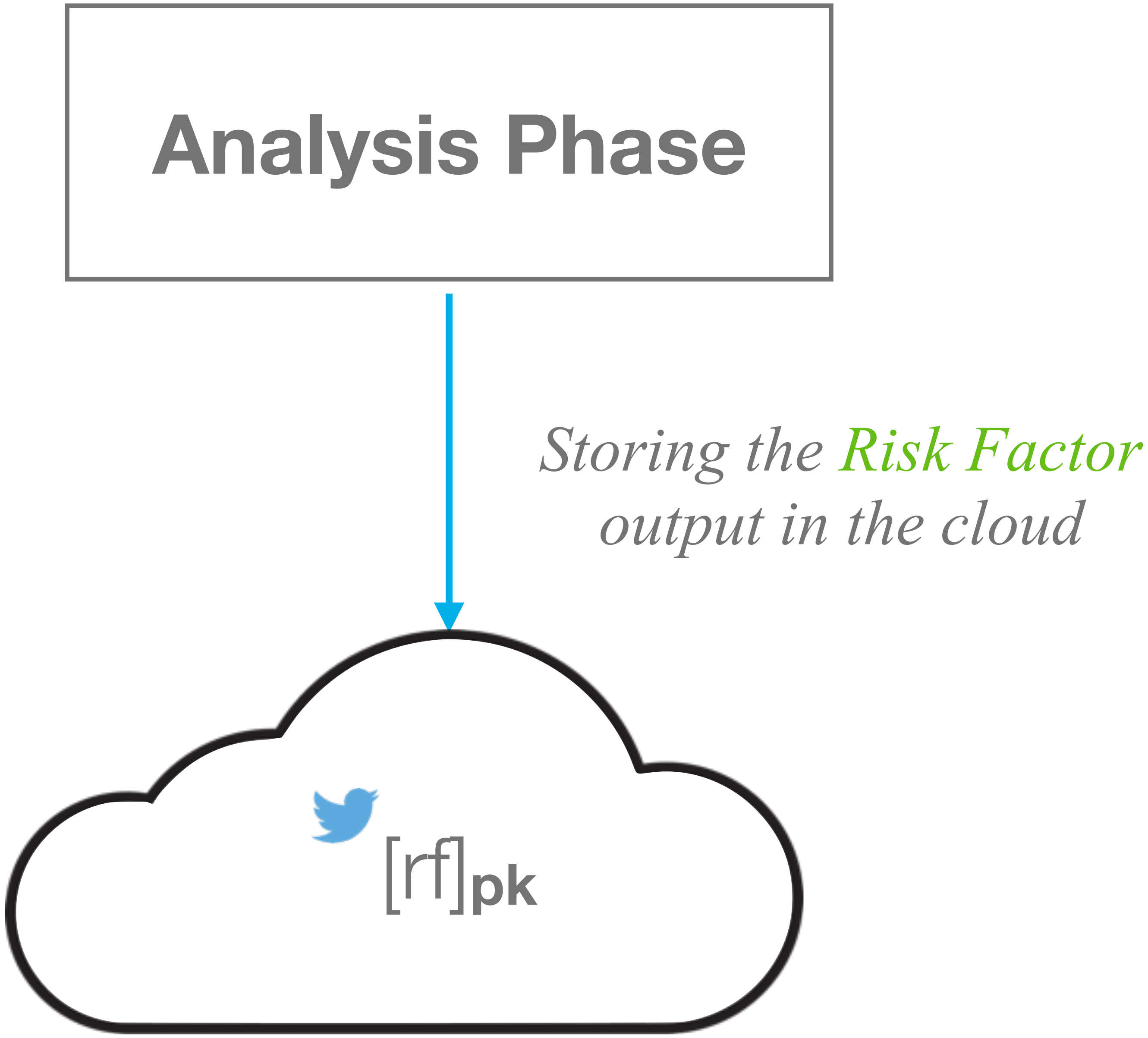
Calculating the Risk Factor

$$RF = (33*7)+(10*5) = [281]_{pk}$$

Template

bomb	killer	gun	shoot	…
46	11	7	5	…

HC@WORKS Project \ Tweets analysis with HE



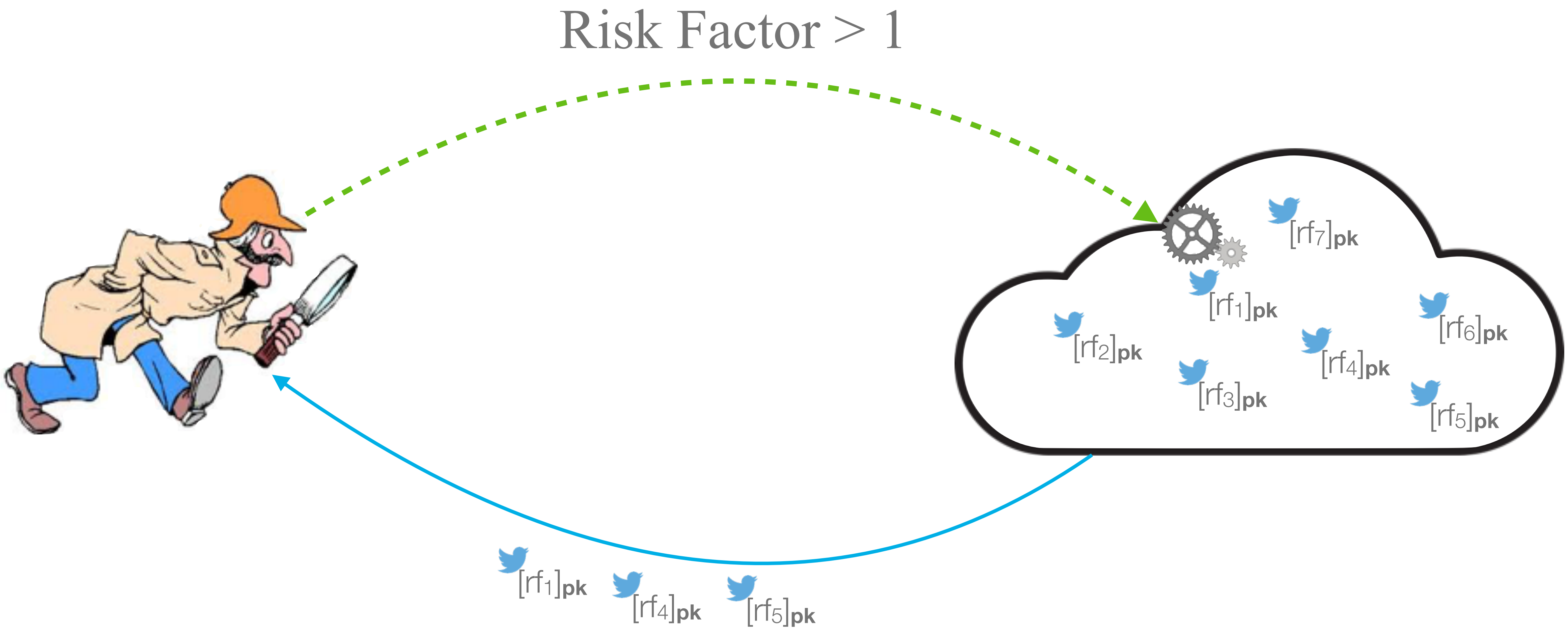
Analysis Phase

Storing the Risk Factor output in the cloud

$[rf]_{pk}$

HC@WORKS Project \ Tweets analysis with HE

Retrieving Phase:



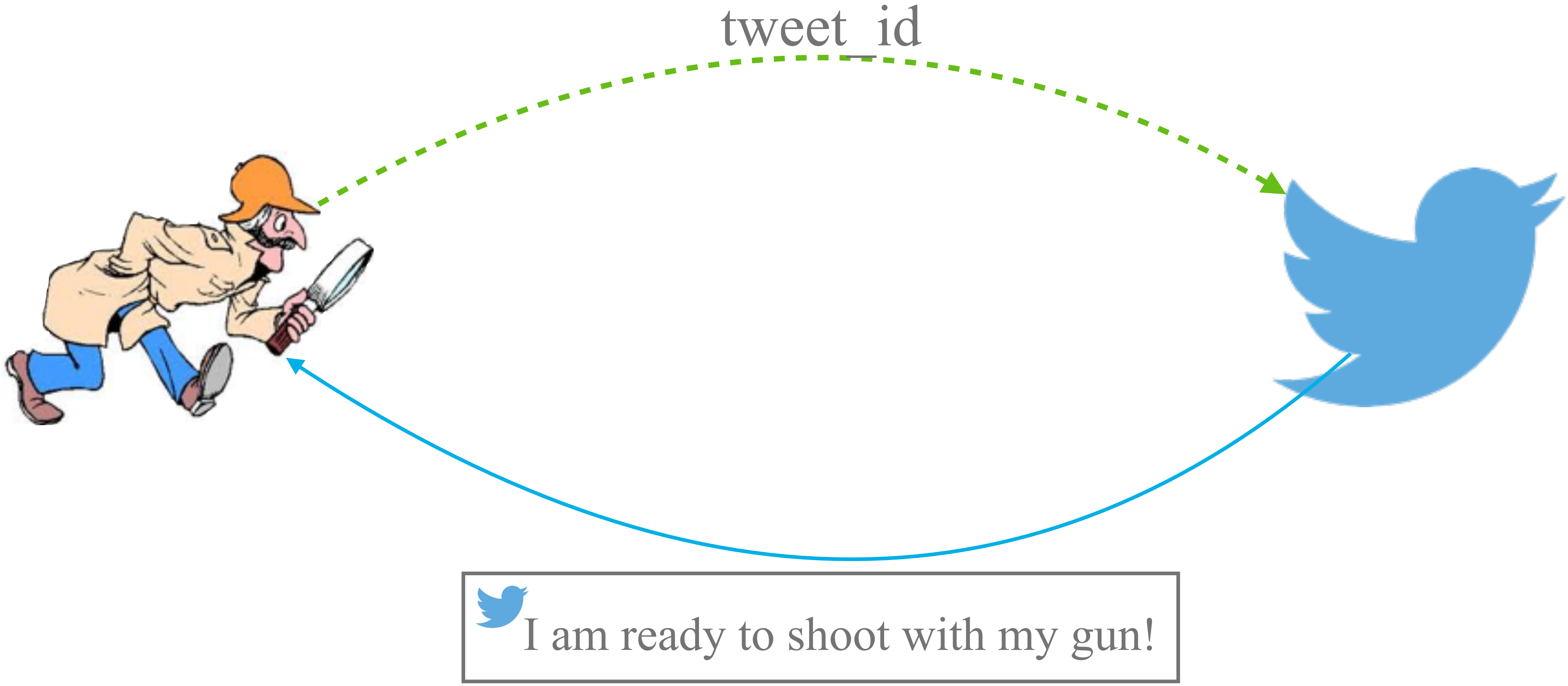
Risk Factor > 1

$[rf_7]_{pk}$ $[rf_1]_{pk}$ $[rf_2]_{pk}$ $[rf_6]_{pk}$ $[rf_4]_{pk}$ $[rf_3]_{pk}$ $[rf_5]_{pk}$

$[rf_1]_{pk}$ $[rf_4]_{pk}$ $[rf_5]_{pk}$

HC@WORKS Project \ Tweets analysis with HE

Retrieving Phase:



tweet_id

I am ready to shoot with my gun!

Conclusion

- We have seen that Cloud storage is very useful but not very privacy oriented;
- However, there are market solutions that want to bridge this gap;
- The HC@WORKS project aims at build the privacy layer that people may need;
- But it is a new technology that needs improvements and time:
    - Storage and CPU requirements;
    - Functions completeness;

Istituto di
Informatica e
Telematica del CNR

PISA

We’re hiring software engineerings

- Very good programming skills
- Security and Networking knowledge
- Mobile development
- Restfull APIs

- Multitasking and
Master Degree is desired…

CVs to
fabio.martinelli@iit.cnr.it
or
gianpiero.costantino@iit.cnr.it