Institut Mines-Télécom

INFORMATION LEAKS-IDENTIFICATION OF THE SOURCE BASED ON DATA WATERMARKING

G. COATRIEUX
IMT ATLANTIQUE
LATIM INSERM UMR 1101

SUMMARY

Institut Mines-Télécom

CONTEXT

Institut Mines-Télécom

Cyber-attack = Economic losses, risks for human life …

Economic loss of 400000 Millions US$ in 2016
Estimated to **2100 Billions US$** in 2019 (+ 500 % !!!)[1]

~ 55% of the security issues come from inside [2]…
… but the most protections focus on external threats

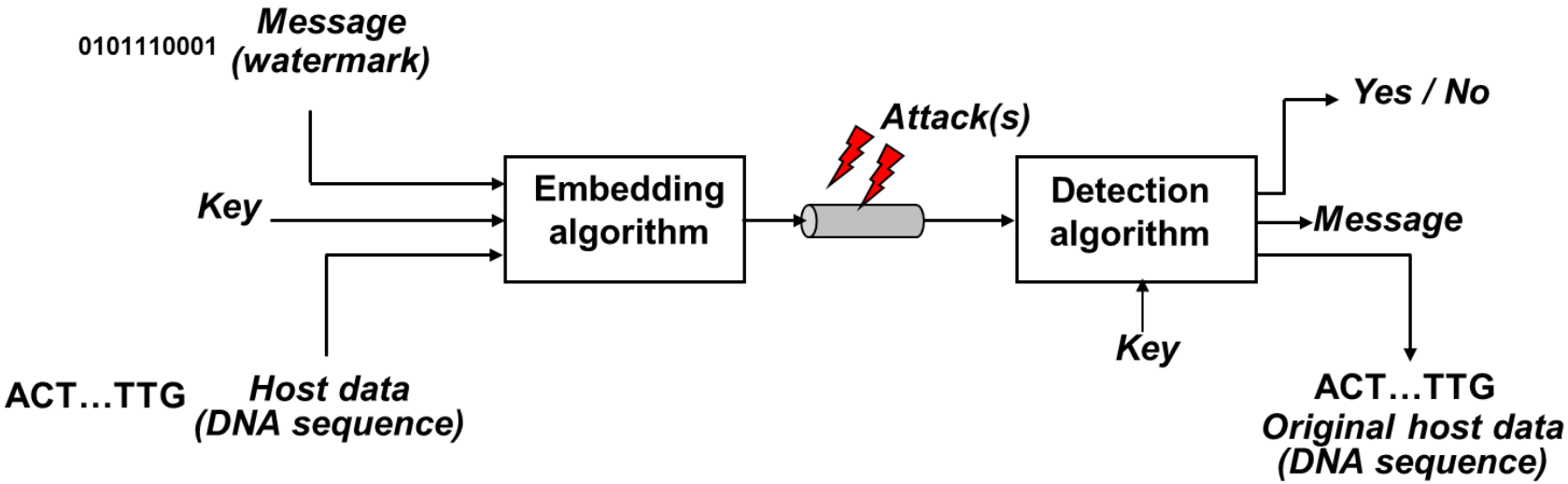
EXAMPLES ISSUED FROM PRESS

- Data leaks about DCNS submarines [3]
- Personal data of 112000 police officers leaked on the web [4]
- One physician from the AP-HM convicted for unlawful processing of health data[5]

1 Juniper Research, "Cybercrime & the Internet of Threats", Whitepaper 2015 ;
2 IBM 2015 Cyber Security Intelligence Index
3 http://www.silicon.fr/fuite-donnees-torpille-dcns-155740.html
4 http://www.lefigaro.fr/actualite-france/2016/06/27/01016-20160627ARTFIG00154-les-donnees-personnelles-de-112000-policiers-ont-fuite-sur-le-web.ph
5 http://www.ticsante.com/Une-pediatre-de-l-AP-HM-condamnee-pour-traitement-illicite-de-donnees-de-sante-NS_3701.html

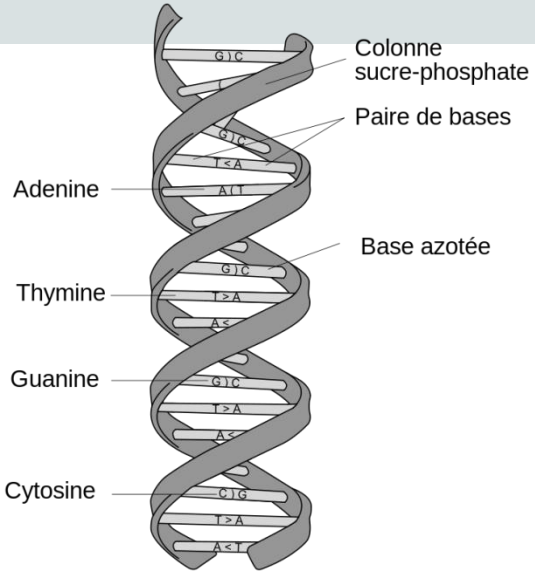
WATERMARKING PRINCIPLES

Institut Mines-Télécom



Main principles of watermarking :

- Data can be accessed while being protected by the watermark (imperceptibility of the watermark)
- Different security services that depend on the link in-between the message and the host data
- *A posteriori* protection independent of the data storage format



- Message encoding based on dictionaries

(4 possible complementary bases (A,T), (C,G)):

D_0={A,C} ➔ A, C encode ‘0’;

D_1={T,G} ➔ T, G encode ‘1’;

- Encoding : replace or not secretly selected bases in a DNA sequence so as to encode the message

Original DNA sequence: **AGCTTGCTATGCAAGTTCGCGATC**
(secret position in blue)

Message : ‘001101’

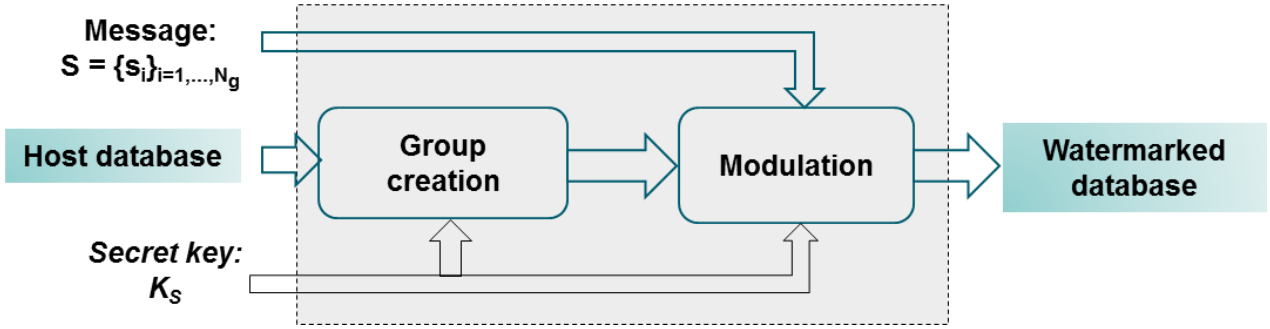
Watermarked DNA sequence: **ACCTTGCAATGCTAGTTGGCGATCAT**

Watermark: **CATGCT**

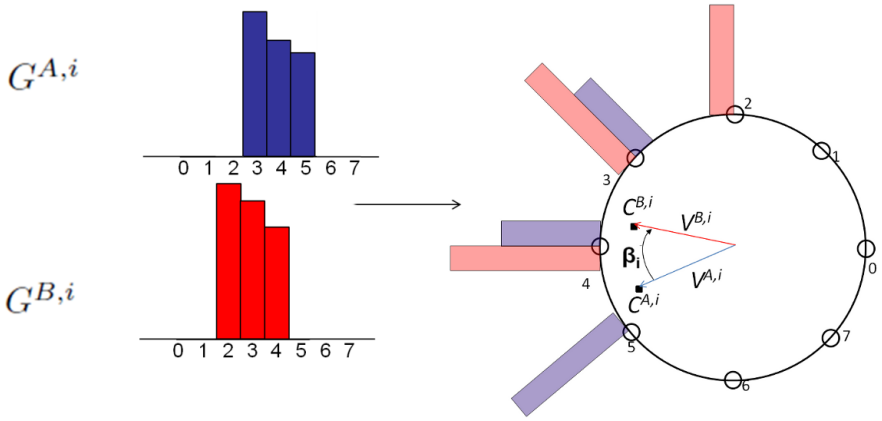
id_stay	id_patient	age	gender	drg	p_diag
4350986	75484	92,23	0	06M03W	A048
4290235	45587	42,34	0	24M11Z	A050
4372568	43567	25,39	0	24M11Z	A058
4562065	35255	54,02	1	06M03V	A058
4607357	68781	43,65	0	06M03V	A058
4546036	34885	65,87	1	06M03T	A058

Sample view of the original table

One tuple includes the attributes: stay identifier ('id_stay'), patient identifier ('id_patient'), patient age and gender, ICD-10-encoded principal diagnosis ('p_diag') ….



- **Symbol embedding in a group *G* – Modulation of circular statistics**
  - Each group of tuples is divided into two sub-groups $G^{A,i}$, $G^{B,i}$
  - Histograms of a numerical attribute in each sub-group are calculated and mapped onto a circle.



$$\beta_i = \widehat{V^{A,i}, V^{B,i}}$$

É ?"

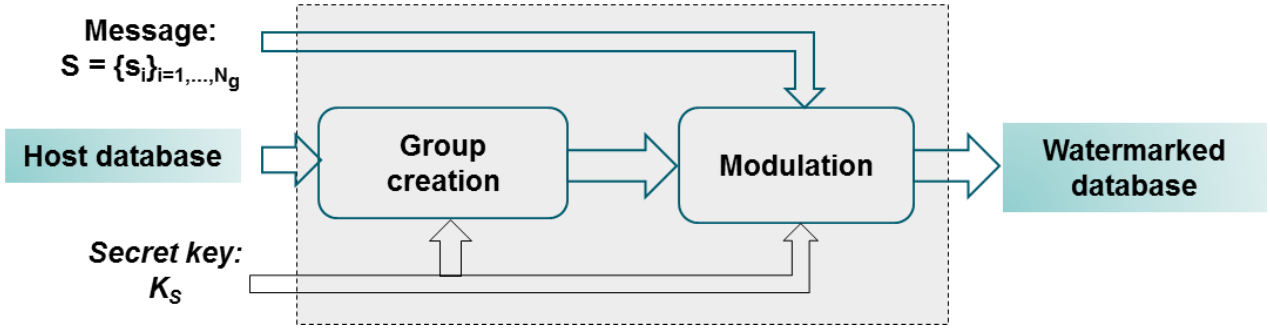
Institut Mines-Télécom

Logotype partenaire

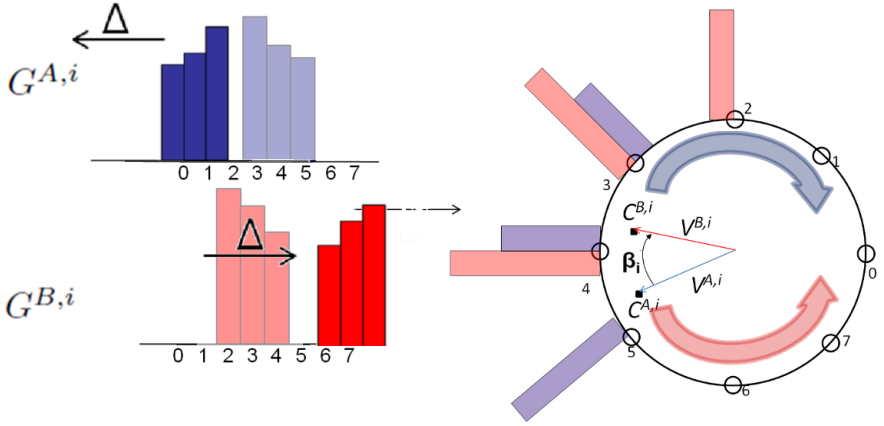
id_stay	id_patient	age	gender	drg	p_diag
4350986	75484	92,23	0	06M03W	A048
4290235	45587	42,34	0	24M11Z	A050
4372568	43567	25,39	0	24M11Z	A058
4562065	35255	54,02	1	06M03V	A058
4607357	68781	43,65	0	06M03V	A058
4546036	34885	65,87	1	06M03T	A058

Sample view of the original table

One tuple includes the attributes: stay identifier ('id_stay'), patient identifier ('id_patient'), patient age and gender, ICD-10-encoded principal diagnosis ('p_diag') ….



- **Symbol embedding in a group *G* – Modulation of circular statistics**
  - Each group of tuples is divided into two sub-groups $G^{A,i}$, $G^{B,i}$
  - Histograms of a numerical attribute in each sub-group are calculated and mapped onto a circle.



$$\beta_i = \widehat{V^{A,i}, V^{B,i}}$$

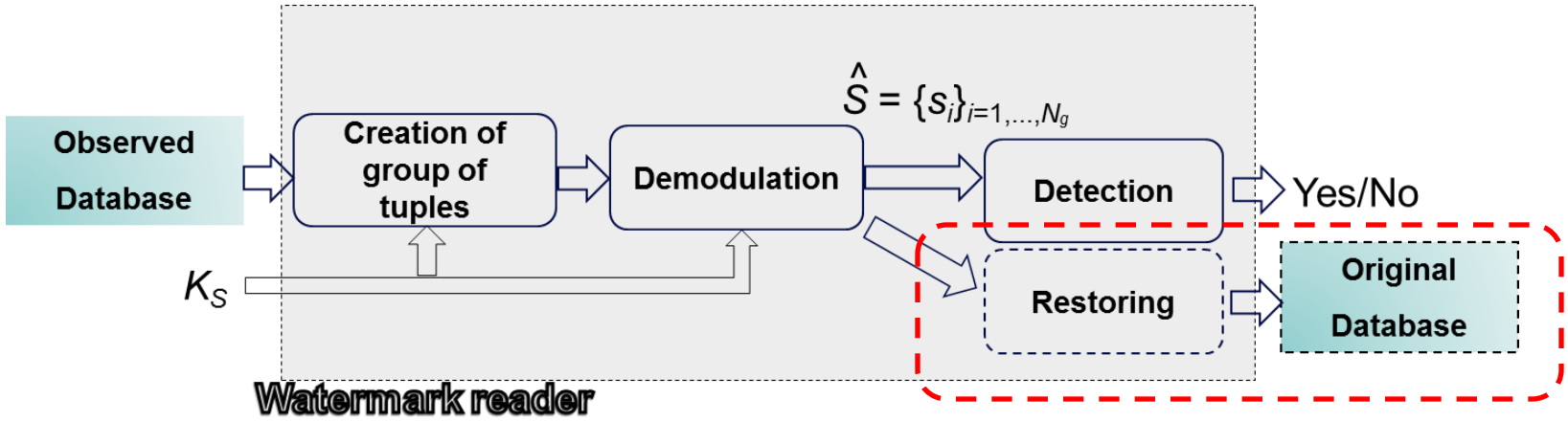
$s_i=1 \rightarrow \beta_i^W>0$

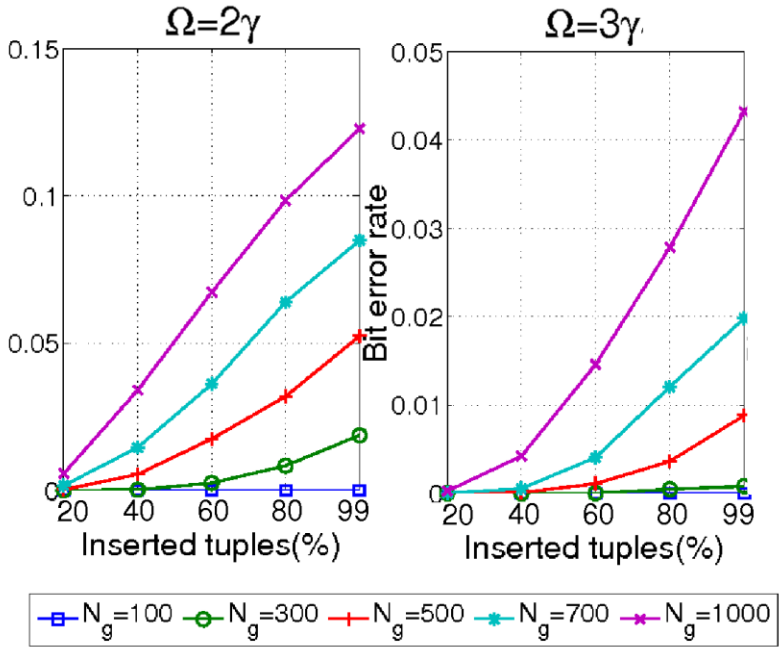
$s_i=0 \rightarrow \beta_i^W<0$

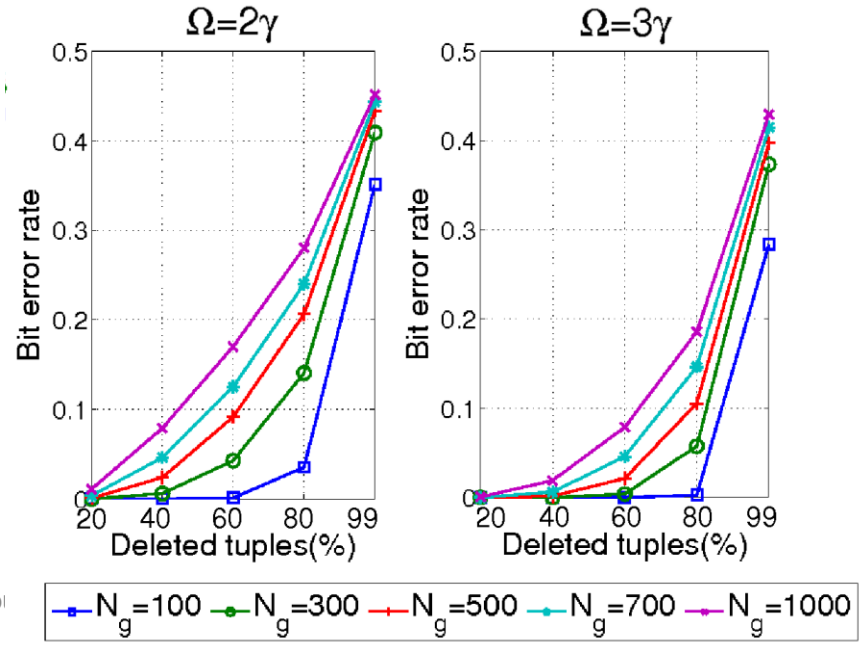
É ?"

Institut Mines-Télécom

Logotype partenaire


Observed Database
Creation of group of tuples
Demodulation
$\hat{S} = \{s_i\}_{i=1,...,N_g}$
Detection
Yes/No
K_S
Restoring
Original Database
Watermark reader

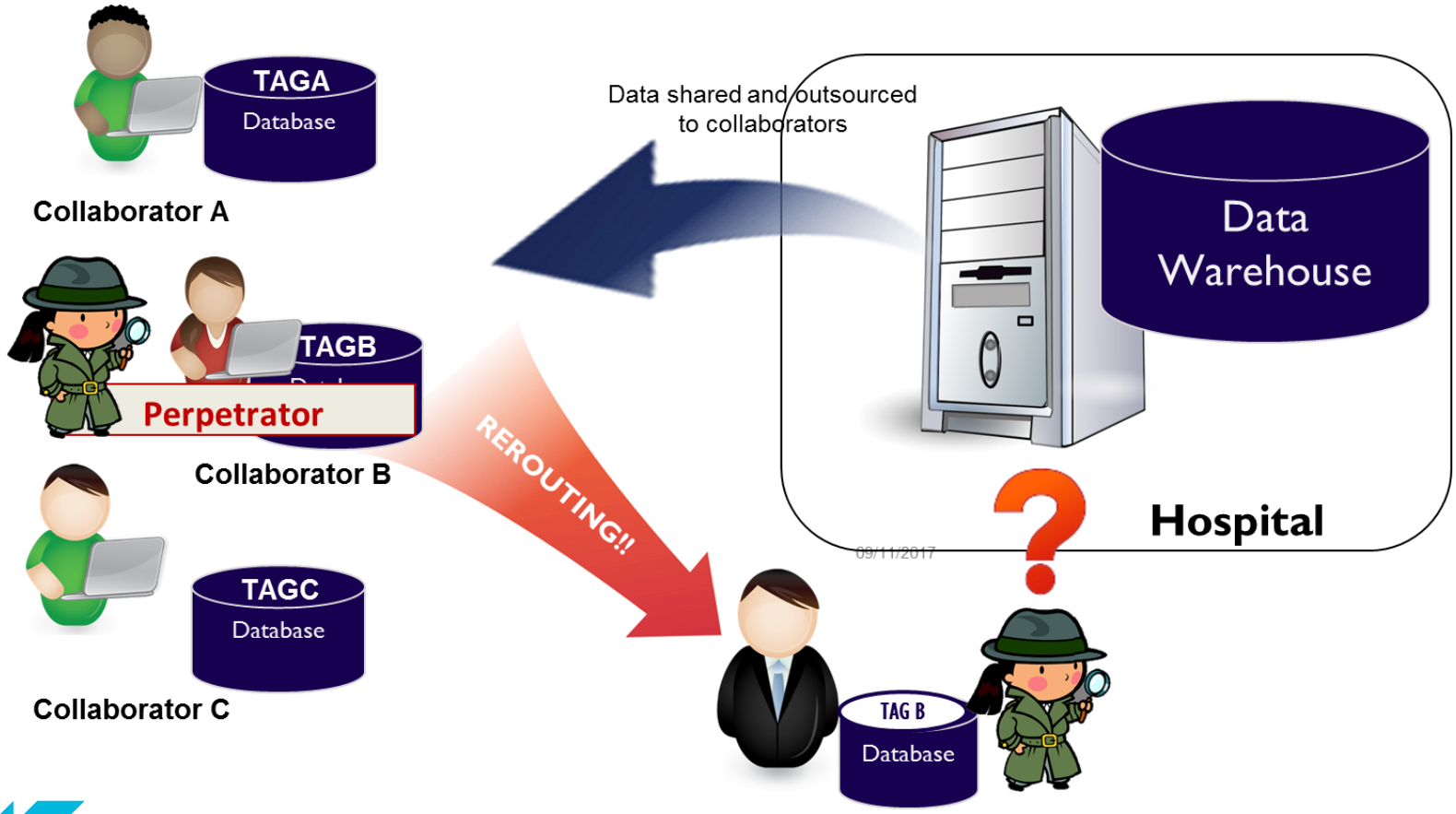

Ω=2γ
Ω=3γ
Bit error rate
Inserted tuples(%)
N_g=100 N_g=300 N_g=500 N_g=700 N_g=1000


Ω=2γ
Ω=3γ
Bit error rate
Deleted tuples(%)
N_g=100 N_g=300 N_g=500 N_g=700 N_g=1000

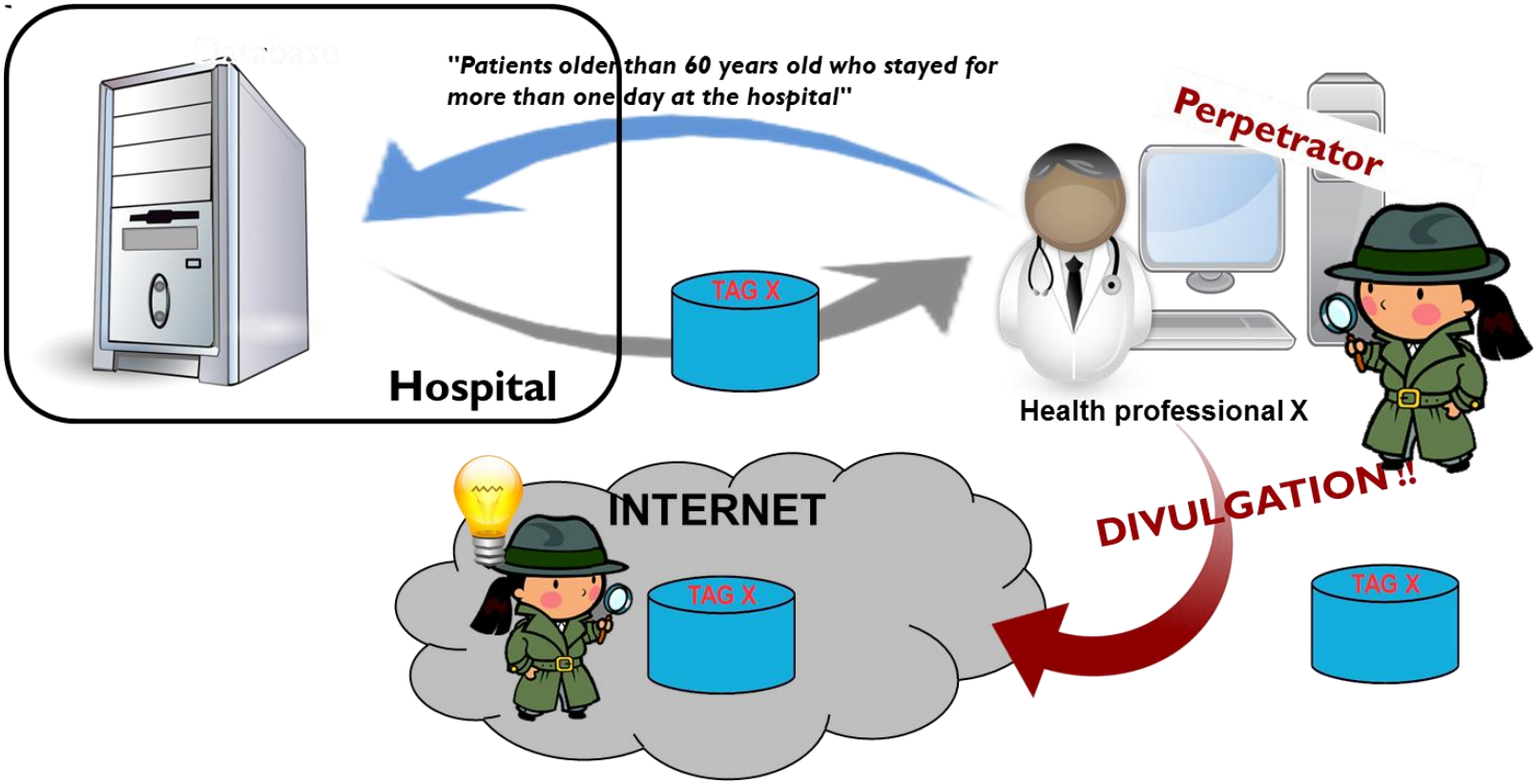
DATABASE WATERMARKING APPLICATIONS

Institut Mines-Télécom

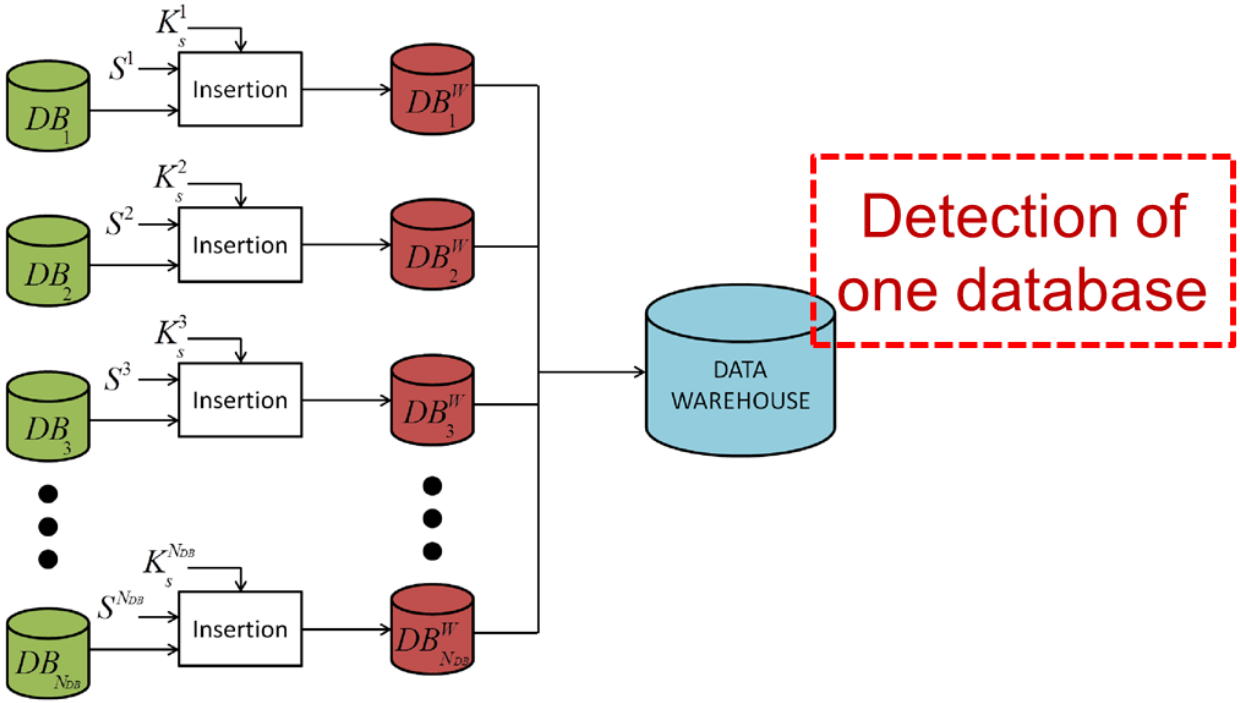
- Traitor tracing: identifying the user at the origin of an information leak, illegal data resale …



- **Traitor tracing: identifying the user at the origin of an information leak, illegal data resale …**



❑ Data Traceability: identifying the presence of protected data in a set



Data watermarking applications:

- **Traitor tracing**: identifying the user at the origin of an information leak,
- **Data Traceability**: identifying the presence of protected data in a set
- **Integrity control**: detecting a data set has been illegally altered
- **Authenticity control**: provide the proof of the data origin (acquisition) and its attachment to a person
- **Access control policy compliance and consent verification**: verifying data are used accordingly a Service Level Agreement – e.g. data should be removed after a period of time.
- **Meta-data insertion**: provide functionalities directly from data (e.g. indexing)