

# Multimedia Security and Privacy

**Slava Voloshynovskiy**

# Copyright notice

---

This course uses in part some materials from

- Book: I. Cox, M. Miller and J. Bloom, “Digital Watermarking: Principles and Practice”, *Morgan Kaufmann Publisher Inc.*, San Francisco, 2001;
- Book: M. Barni and F. Bartolini, “Watermarking Systems Engineering: Enabling Digital Assets Security and Other Applications”, *Marcel Dekker*, 2004;

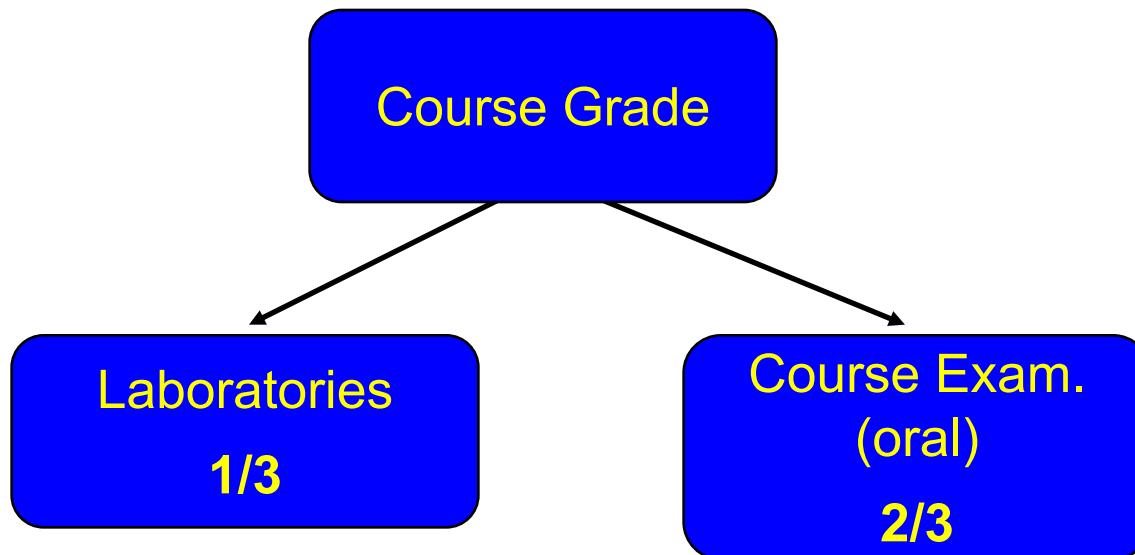
And some papers, books, presentations that will be indicated along the course.

The usage of the slides in commercial or educational purposes is prohibited without authorization of the course leaders and permission of the above document authors.

# Grading Policy

---

The final grade will be determined based on laboratories and a final examination.

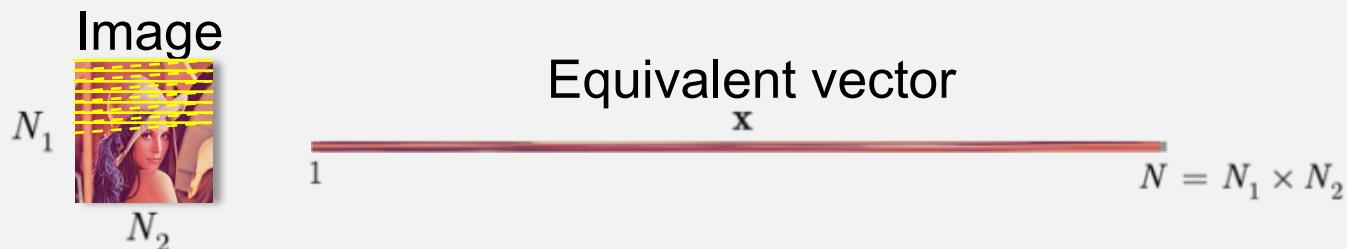


# Notations

---

## Notations

- $X$  Random variable with some distribution  $p(x)$
- $\mathbf{x}$  Random vector of length  $N$ :  $\mathbf{X} = \{X_1, X_2, \dots, X_N\}$
- $x$  Realization of random variable
- $\mathbf{x}$  Realization of random vector  $\mathbf{x} = \{x_1, x_2, \dots, x_N\}$



# Notations

---

	Realization	Random
Scalar	$x$	$X$
Vector	$\mathbf{x} = x^N$	$\mathbf{X} = X^N$

$$x_i \in \mathcal{X} \quad \text{or} \quad x_i \in S_X$$

- Discrete alphabets

$$\mathcal{X} = \mathbb{Z} \quad \text{- integer}$$

$$\mathcal{X} = \{0,1\} \quad \text{- binary}$$

$$\mathcal{X} = \{a,b,\dots,z\}$$

$$\mathcal{X} = \{January, February, \dots, Decemeber\}$$

- Continuous alphabets

$$\mathcal{X} = \mathbb{R}$$

- Vector  $\mathbf{x} \triangleq \begin{bmatrix} | \\ \mathbf{x} \\ | \end{bmatrix} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_N \end{bmatrix} = \begin{array}{c} 1 \\ 2 \\ \vdots \\ N \end{array} \begin{array}{c} \textcircled{1} \\ \textcircled{2} \\ \vdots \\ \textcircled{N} \end{array}$

$$\mathbf{X} \triangleq \begin{bmatrix} | \\ \mathbf{X} \\ | \end{bmatrix} = \begin{bmatrix} X_1 \\ X_2 \\ \vdots \\ X_N \end{bmatrix}$$

# Notations

---

$\Pr[X = x]$  Probability of event

- Discrete random variables

$X \sim p_X(x)$  Probability mass function (**pmf**) of discrete random variable  $X$   
(or  $X \sim p(x)$ )

$p_X(x | Y = y)$  Conditional pmf of discrete random variable  $X$  given  $Y$   
(or  $p_X(x | y)$ )

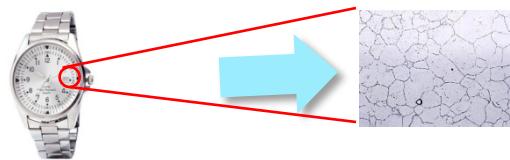
- Continuous random variables

$X \sim f_X(x)$  Probability density function (**pdf**) of continuous random variable  $X$   
(or  $X \sim f(x)$ )

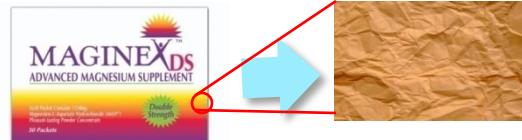
$f_X(x | Y = y)$  Conditional pdf of continuous random variable  $X$  given  $Y$   
(or  $f_X(x | y)$ )

# Applications and main concerns

## Physical Objects



Luxury products = “Annoying”

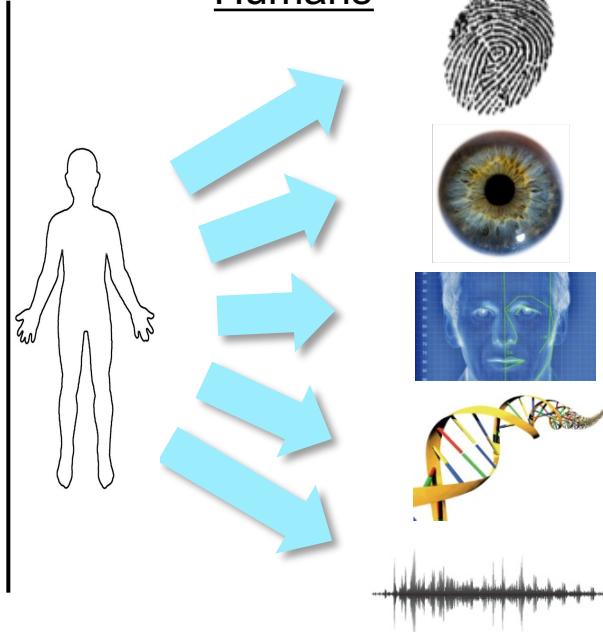


## Medicine = “Dangerous”



Electronics = “Security as whole”

## Humans



## Digital Content

## Images

## Videos

## Audios

# Text docs



In July 1845 Texas formally accepted an American president. In the United States, Andrew Jackson remained neutral, but his son-in-law, General Zachary Taylor, and his troops were sent to the Rio Grande. On April 21 at the Alamo, Texan volunteers surrendered to Mexican troops under Santa Anna. On May 8 the Mexicans massacred Taylor's men at Resaca de la Palma. In June Taylor began a march toward the Rio Grande. On September 8 he defeated the Americans near Sabine and with them often fled Santa Anna.

Santa Anna now faced the field against the American forces. He had to retreat from the Battle of Victoria to Matamoros, where he was joined by General Vicente Filisola. From San Antonio he sent Taylor's forces to the coast. On October 10, 1846, Taylor finally fighting Santa Anna pushed to the river, leaving Taylor's forces to defend the fortifications of the city.

When Taylor received the news, General W. Fawcett said "Arrr! the *Devil's* here." New Mexico, Captain Charles Loring, and the 1st Dragoons were sent to California and sending the remainder under a command of General J. W. Geary to the Pacific Coast. In September 1847 the two countries finally reached a truce.

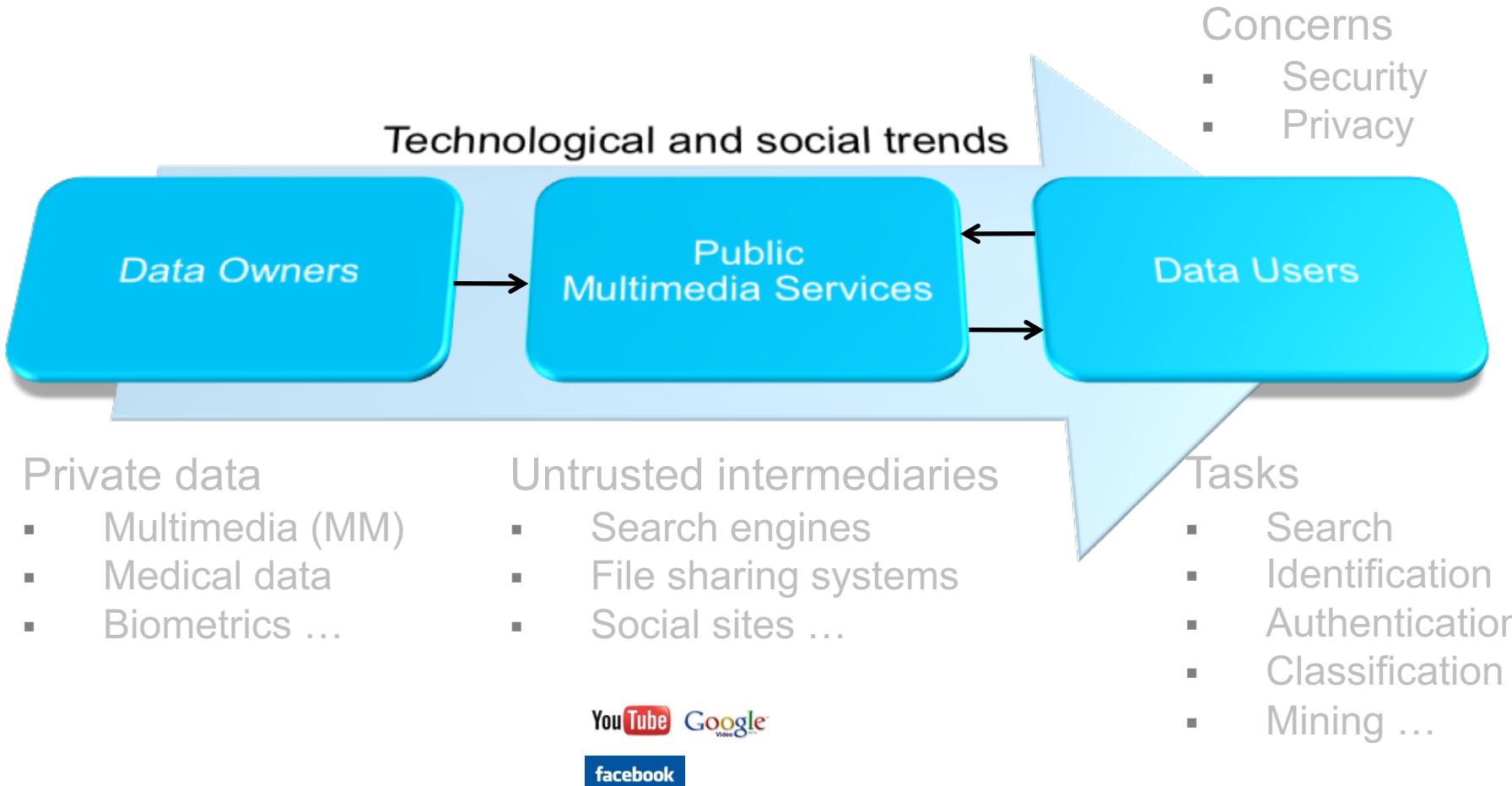
## Online sharing services



## Main concerns

- Identification: identity, authenticity, origin, ownership, ...
  - Tracking and tracing
  - Automatic tagging

# Modern networked multimedia applications

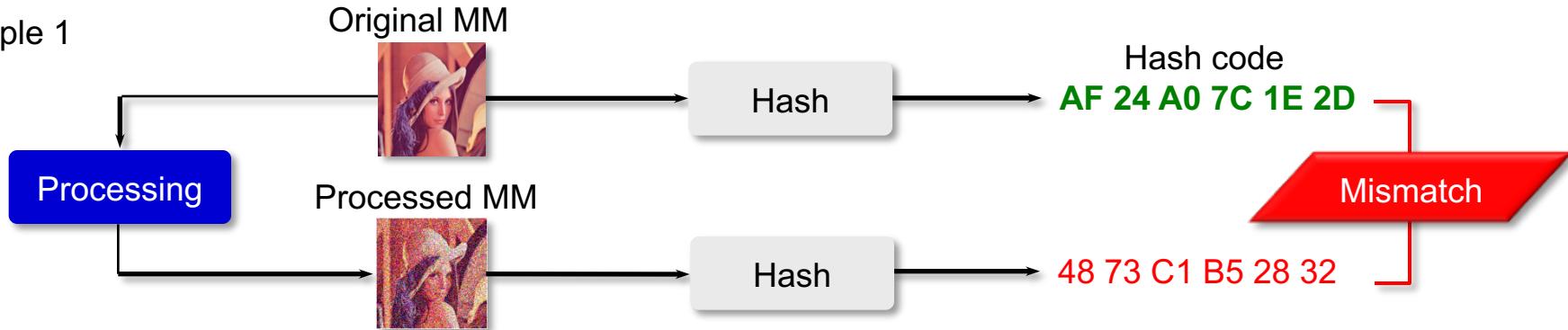


# Multimedia security

## Why are traditional security tools not suitable for multimedia?

- Traditional security:
  - Cryptographic encryption (confidentiality of information)
  - Cryptographic hashes (authentication, trust, access control)
- Main concerns of classical crypto-based algorithms:
  - Sensitivity to noise and unintentional distortions in input data
  - Data handling in the encrypted domain

Example 1

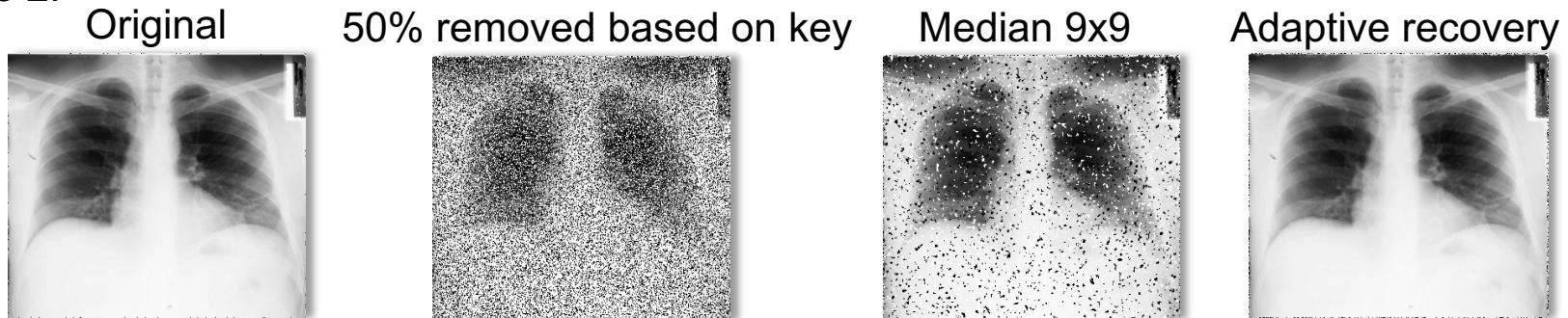


# Multimedia privacy

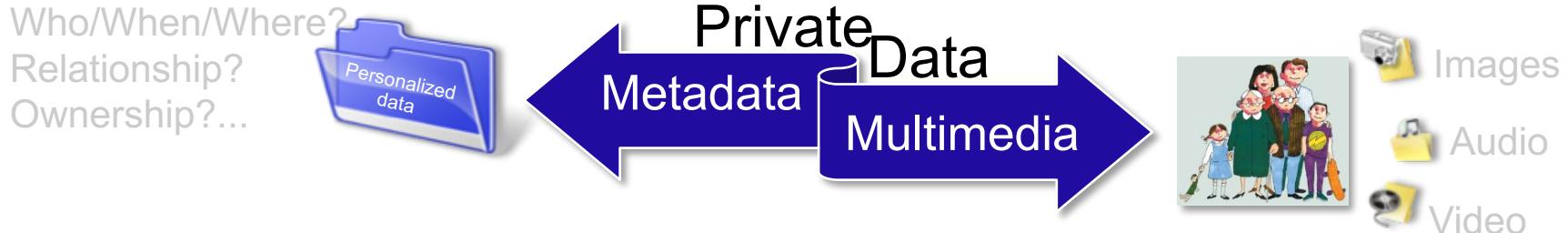
## Why are traditional privacy tools not suitable for multimedia?

- Traditional *privacy protection*:
  - Data owner (protect data): based on data degradation and randomization (noise addition, lossy compression, data removal, dimensionality reduction...)
  - Data user (protect requests): based on anonymization, randomized rules
- Main concerns of classical privacy preserving algorithms:
  - Reduction of accuracy (data utility)
  - Not very efficient against experienced attackers (sensitivity analysis)

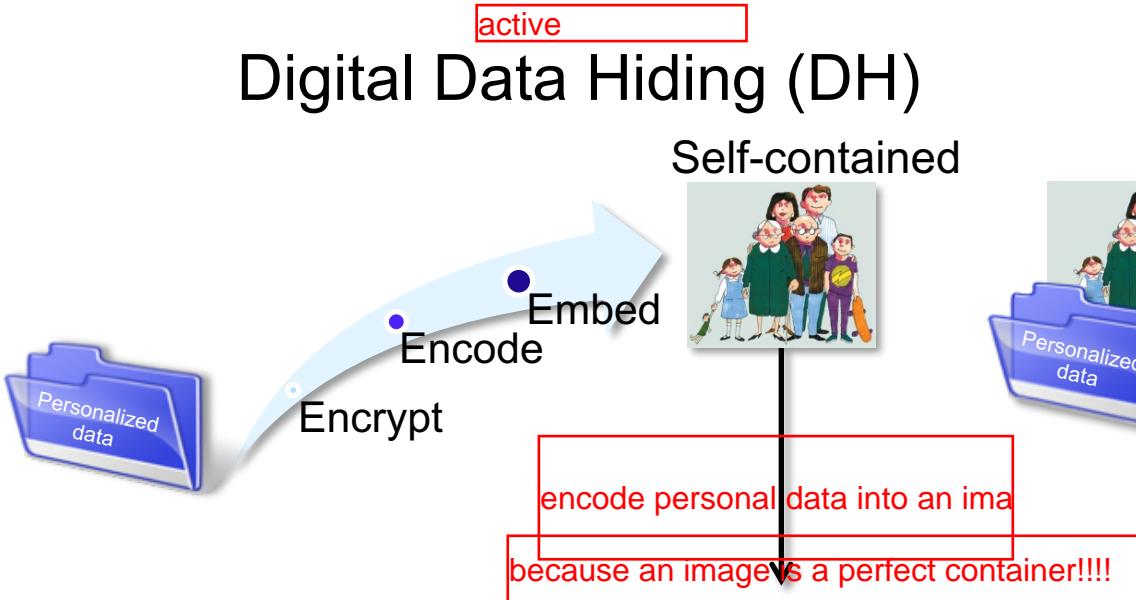
Example 2:



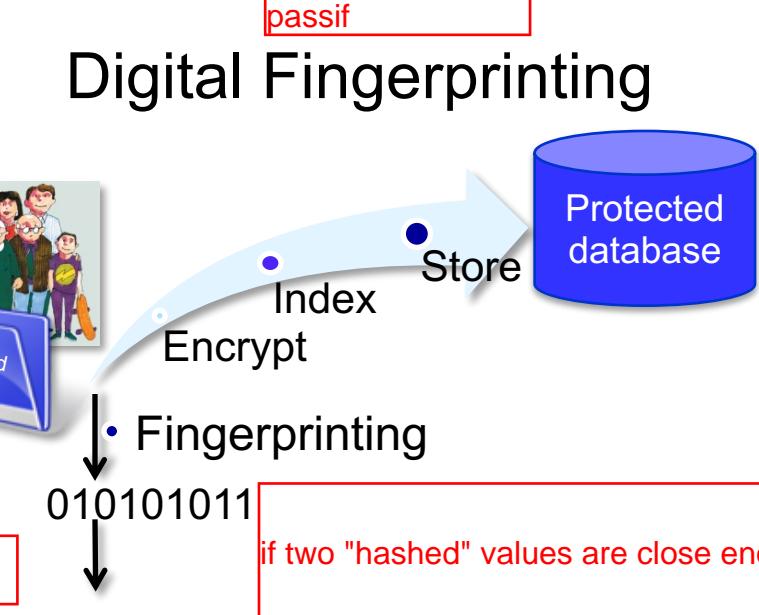
# New tools for multimedia privacy and security



## Digital Data Hiding (DH)



## Digital Fingerprinting

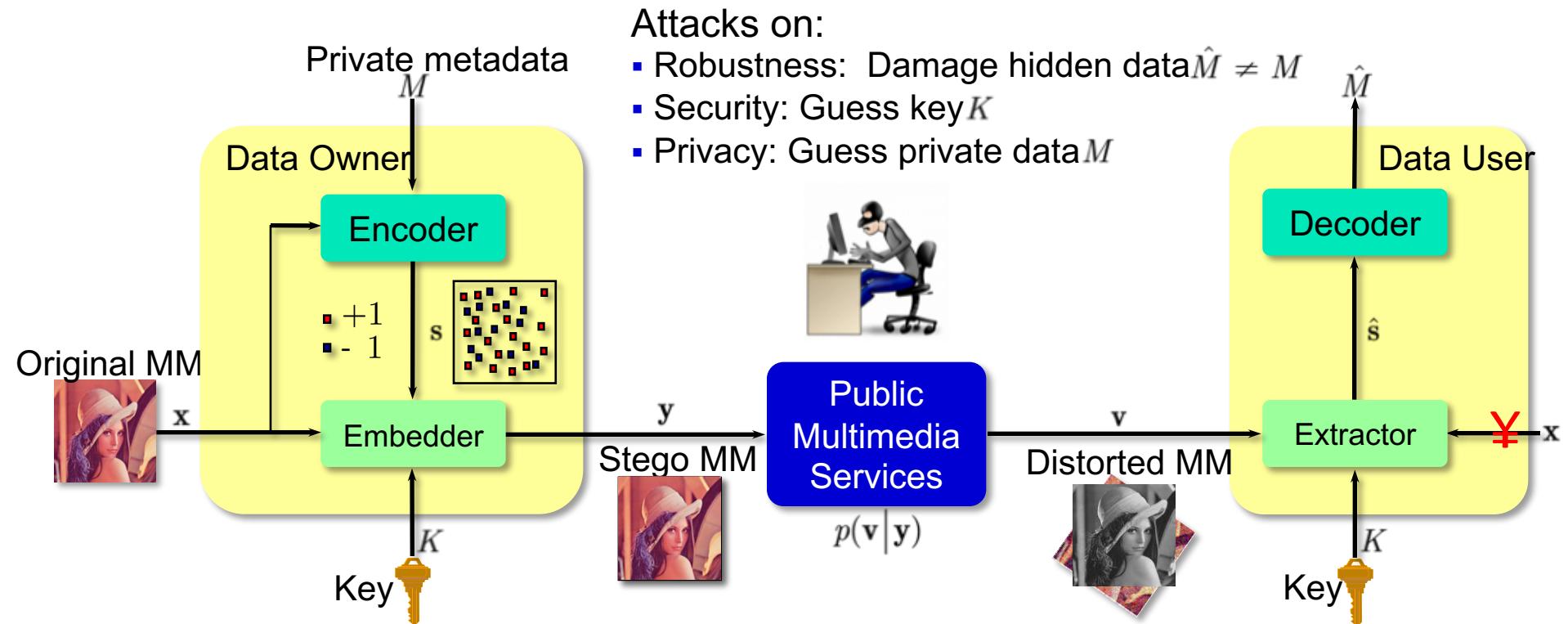


## Public Domain Multimedia Management

# Digital data hiding: definition and concept

## Definition (Digital data hiding)

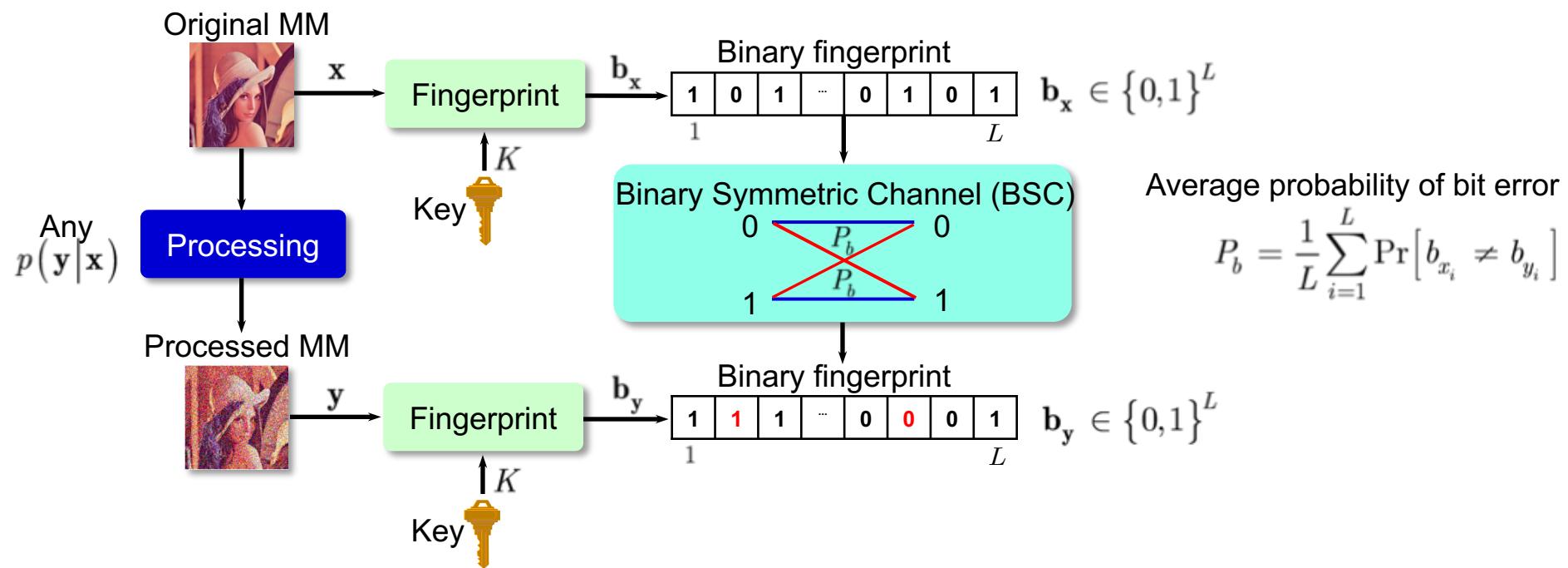
Digital data hiding (*a.k.a. steganography*) is the art of perceptually and statistically undetectable robust information embedding in multimedia content.



# Digital fingerprinting: definition and concept

## Definition (Digital fingerprinting)

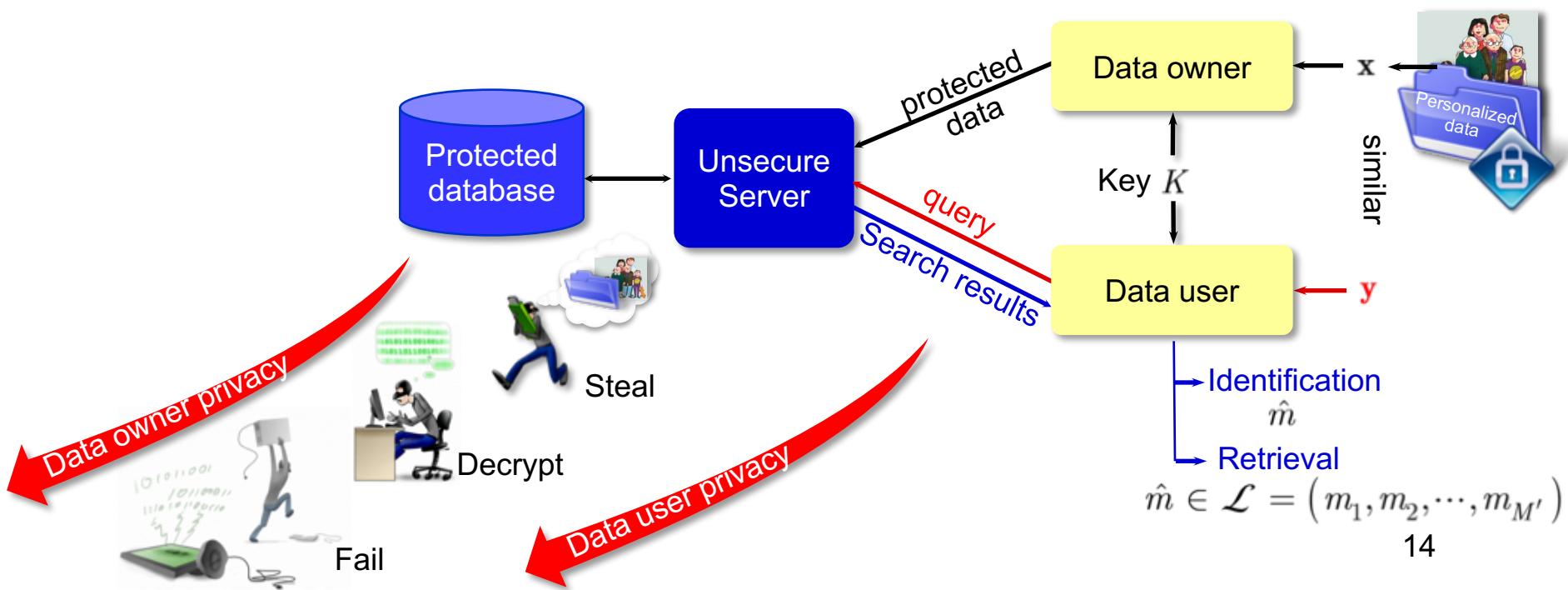
Digital fingerprinting (*a.k.a. robust perceptual hashing*) is a technique for computing a compact robust, secure and private binary representation of multimedia content.



# Privacy-preserving search: concept

## Main concerns of privacy-preserving search in large-scale systems

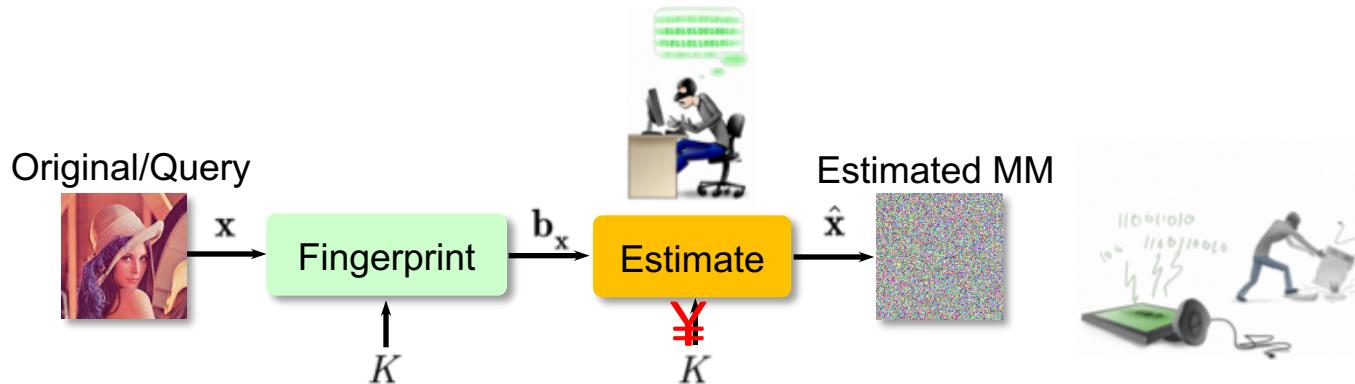
- Exact queries and binary distributed hash tables in P2P systems (incl. LSH)
- Complexity:
  - Optimal Maximum Likelihood-based search is a NP-hard problem ( $\mathcal{O}(2^L)$ )
  - Cryptographic homomorphic encryption is computationally expensive



# Privacy-preserving search: concept

---

Privacy



# Course outline

---

- Digital watermarking
- Digital content fingerprinting
- Privacy protection

# Extensions

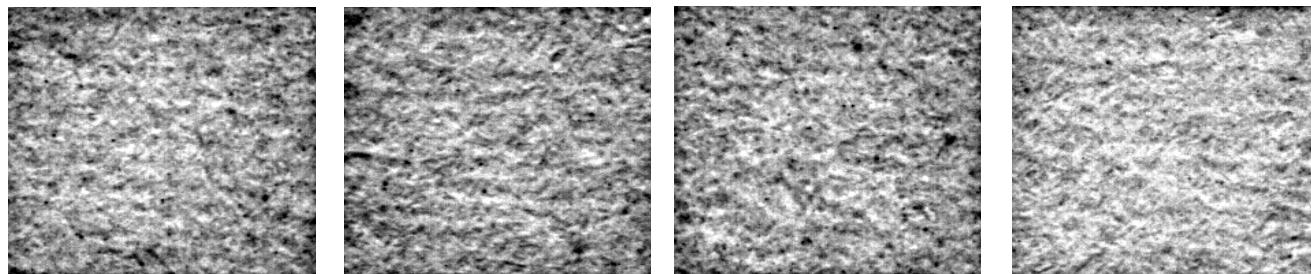
---

- Physical world security
- Security based on smart phones

# Protection of physical objects: natural randomness

## Definition (Natural randomness)

Natural randomness represents unclonable features created by nature or some uncontrollable randomized process initiated by humans.



## Features

- Easy to evaluate but is hard to characterize
- Structure is unique for each item
- Manufacturer not-reproducible

# Protection of physical objects: natural randomness

---

Sample items with random surface structure  
(cardboard)



Item 1, unique surface features on every item



Item 2, unique surface features on every item

**u-nica**<sup>®</sup>

Global  
Security  
Solutions

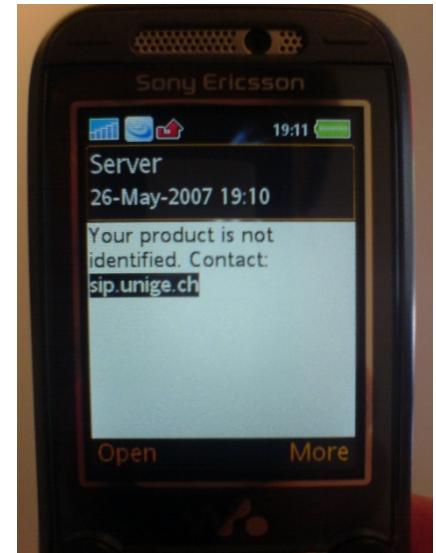
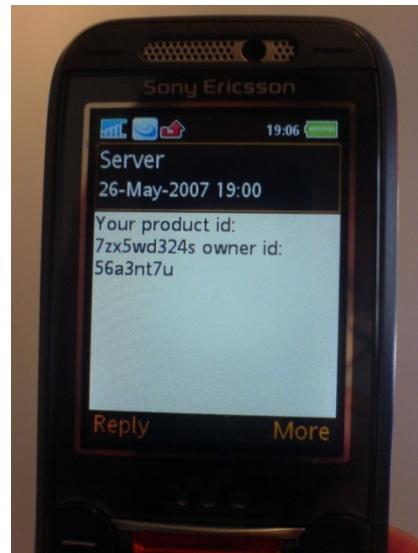


© U-NICA Security AG 2011

# Verification on mobile phones

---

## Watch identification



# Verification on mobile phones

---

