Introduction to privacy protection
Part 1 - Introduction and legal context

Guillaume Piolle
guillaume.piolle@supelec.fr
http://guillaume.piolle.fr/

Master "Machine Learning and Data Mining", Saint-Étienne

December 16th 2013

Introduction

Is private stuff shameful?

> *If you haven't done anything wrong, then you don't have anything to hide.*

- But then, why use an envelope when you send a letter?
- It is not because you have "nothing to hide" that you will not be blamed for anything or that nothing can hurt you.

This kind of statement is usually made by a member of the "high caste": a man, white, heterosexual, usually more than 45 year old.

An instance of risk: the privacy breach

Main kinds of privacy breaches

- Intrusion of someone in your "private" business;
- Disclosure, to the public or to a third party, of a "private" information which you were not ready to reveal, and/or which is false;
- Identity theft.

Possible consequences

- More or less severe impact on social relationships;
- Risk of discrimination ;
- Risk of legal/criminal proceedings ;
- ...

What about "public" people?

Risks: identity theft

Symptoms (identitytheft.org.uk)

- Loss of identity documents;
- Some mails (in particular from banks) don't reach you anymore;
- Unusual banking transactions;
- You are told to have requested a loan or some kind of social/government aid;
- You receive invoices, orders to pay, formal demands for goods or services that you don't know about;
- You are refused a credit although you have a good credit record;
- A mobile phone contract has been subscribed in your name;
- You are being contacted by banking institutions which you are not related to;
- ...

What is the connection with computing?

The concept of privacy, of right to privacy, of privacy protection. . . can be (and is) considered independently from computing

Computing (and Internet) brings **new sources of risk** but also **new protection tools**.

Web services, cloud computing and privacy: Mat Honan

The Mat Honan case

Senior journalist at *Wired*.

In August 2012, all his online accounts are compromised and a large part of his personal data are destroyed by a few attackers.

The damage

- Google account compromised, then deleted;
- Twitter account compromised, then used to publish racist and homophobic declarations;
- Amazon account compromised;
- AppleID account compromised, allowing remote data suppression on his iPhone, iPad and MacBook (one year of pictures, 8 years of Gmail correspondence).

Web services, cloud computing and privacy: Mat Honan

The hacker's motivation

Control of his Twitter account, `@mat`, coveted because in three characters only...

The root causes of the disaster

- Systematic links between the different identities;
- No two-factor authentication;
- No efficient backup strategy;
- Apple applications designed to control user data from a central service point (design seen as "aggressive" towards privacy);
- Critical flaws in the security procedures at Amazon and Apple (a partial credit card number displayed by Amazon, and then used as authentication factor by Apple).

Web services, cloud computing and privacy: Mat Honan

What could be saved

Mat Honan took back control over his online accounts and got most of his data back. A large part of his cloud data (Gmail & co) was restored, the remote deletion of his MacBook was halted.

Consequences: Mat Honan

Setup of an actual backup strategy, systematic activation of two-factor authentication, deactivation of Apple's *Find my* services.

Consequences: at Amazon

Correction of organizational flaws

Consequences: at Apple

No information on any procedure update... Hackers still seem to consider AppleID accounts as easy entry points...

General and non-technical definitions

Privacy (Oxford dictionary)

- *The state or condition of being free from being observed or disturbed by other people;*
- *The state of being free from public attention.*

Private sphere (Crépin 2008)

The set of the pieces of information a person considers private.

The private sphere is **personal**, **customizable** and **context-dependent**.

Privacy protection

The set of measures aimed at preserving the **control** one may have over one's private sphere (perimeter and content).

Privacy: a culturally situated concept

In France

- A fundamental right, even a "fundamental fundamental" one, necessary condition to the exercise of other fundamental rights;
- In the constitutional block since 1971;
- Central role of the State as a garantor of this right;
- Privacy traditionally seen as having a social value which ought to be protected.

In the United States

- Privacy cannot conflict with liberty of speech, legally superior (first amendment);
- Traditional mistrust towards the (federal) state;
- Central role of the market, of free enterprise;
- Privacy traditionally seen as an individual notion, upon which every ~~individual~~ citizen can decide freely.

(French-centred) History

- Warren & Brandeis 1890: *The Right to Privacy*. First elaborations on the topic, triggered by the emergence of photography and mass media;
- Progressive construction of a right to privacy in legal doctrine, in the form of an **immaterial property right**, related to **personality rights**;
- 1970: Introduction of the right to privacy in the French civil code;
- End of the 1970s: French scandal of the Safari file, creation of the *Informatique et Libertés* law;
- 90s and 2000s: too dense to be summarized here!

(French-centred) History

Reminder:

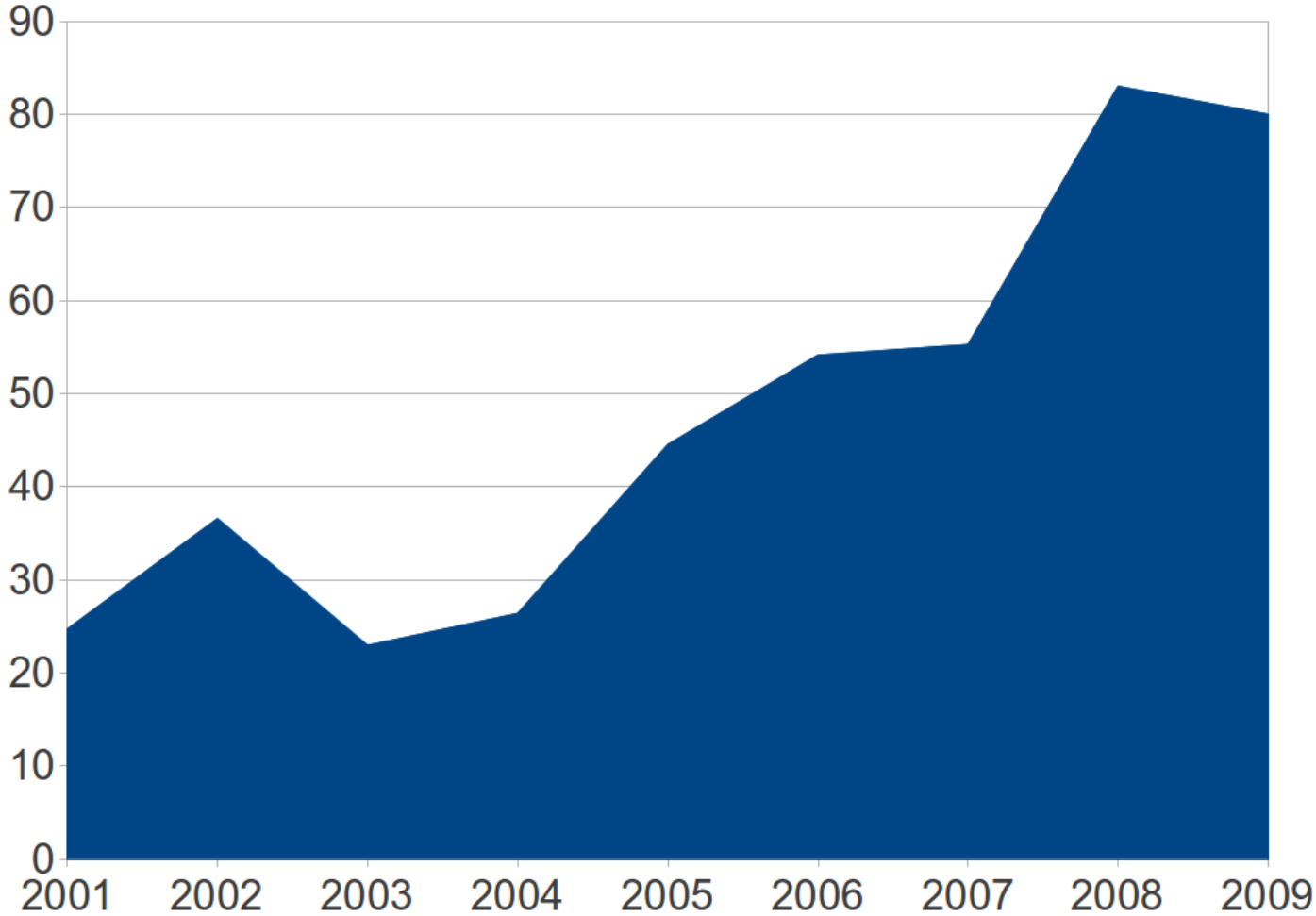
Originally, the "enemy" is not Google or Facebook, it's the State!

Since the 2000s, privacy-related concerns have turned to other actors.

Is it because it is not necessary to worry about the State anymore?

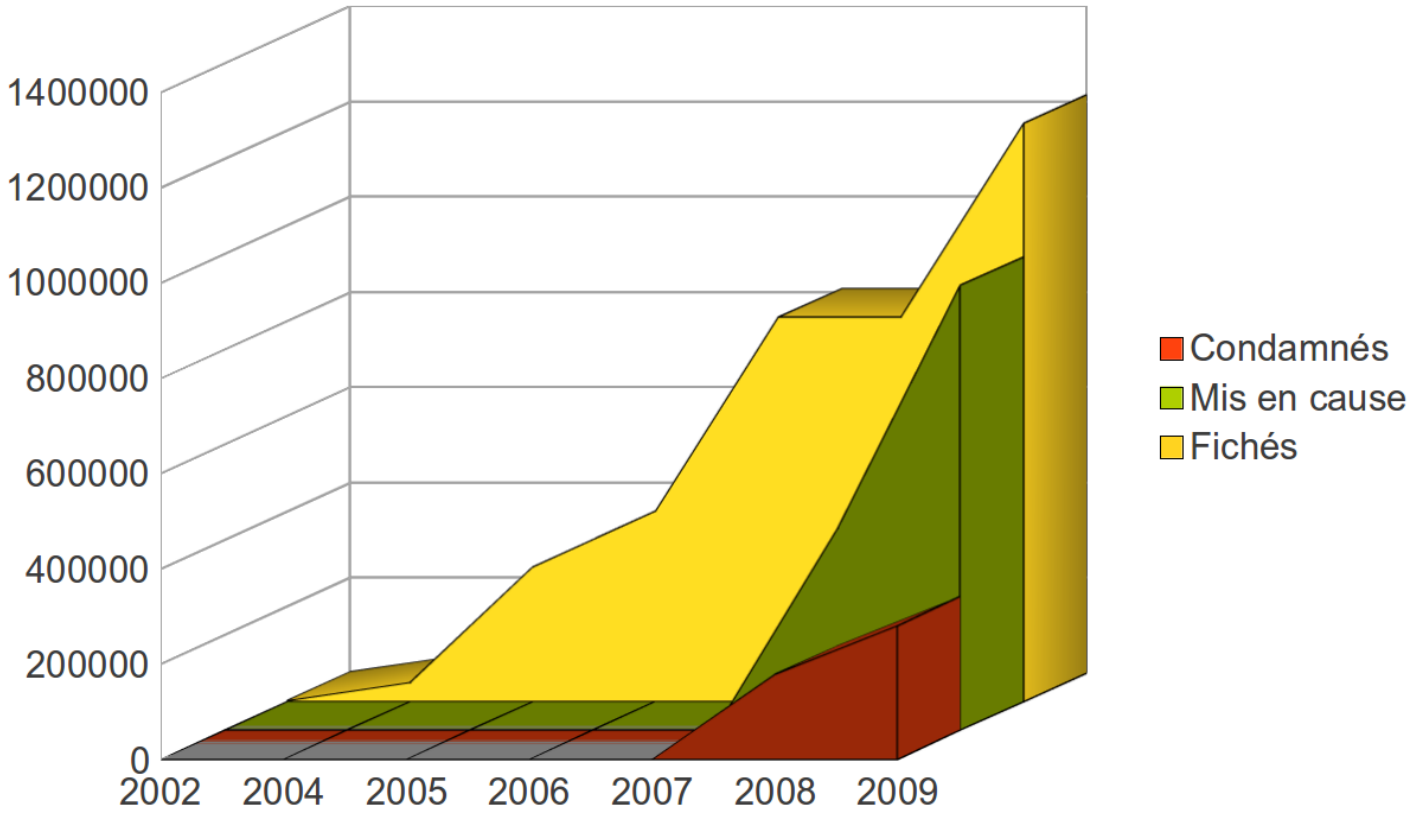
- The Safari file has been reintroduced (as the Edvirsp file in law Loppsi 2), and it has many friends (44 police and defence files created in France between 2002 and 2009);
- As of 2009, many files stay undeclared or non-compliant at the French Home and Defence ministries (cf 5-year grace period after 2004);
- Abusive access to police files;
- ...

(French-centred) History



Error level in the STIC file (source: J.-M. Manach)

(French-centred) History



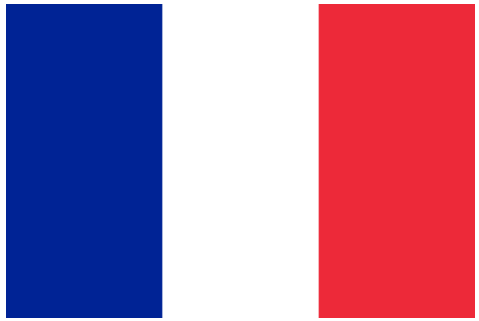
Fichier National Automatisé des Empreintes Génétiques

(national automated genetic print file)

International context

- **United Nations:**
  - 1948 - Universal Declaration of Human Rights (art. 12);
  - 1966-1980 : International Covenant on Civil and Political Rights, International Covenant on Economic, Social and Cultural Rights.
- **European Council:**
  - 1950 - Convention for the Protection of Human Rights and Fundamental Freedoms ("European Convention on Human Rights", ECHR) (art. 8);
  - 1981 - **Convention for the Protection of Individuals with regard to Automatic Processing of Personal Data** (in particular preamble, art. 5).
- **European Union:**
  - 1992-2007 - Treaty on European Union (includes the ECHR) ;
  - 2000-2010 - Charter of Fundamental Rights of the European Union;
  - 1995 : **European Directive 95/46/CE**
  - 2002 : Directive 2002/58/CE
  - ? - Regulation project, foreseen as a replacement of 95/46.

French national context



- **Civil code**, article 9 ;
- **Law n° 78-17 du 6 janvier 1978** *relative à l'informatique, aux fichiers et aux libertés* ("*Informatique et Libertés*", I&L) ;
- **Law n° 2004-801 du 6 août 2004** *relative à la protection des personnes physiques à l'égard des traitements de données à caractère personnel.*

`http://www.legifrance.gouv.fr/`

Right to privacy *vs* Personal data protection

Right to privacy

"Corrective" right

Notion of damage and redress

The damage has to be proven

Personal data protection

"Preventive" right

Operational rules aiming at avoiding privacy breaches

Violating the rules is considered a damage in itself, by principle

Informatique et Libertés: Perimeter of the law

Article 2 (excerpt – non-official translation)

The present law applies to the **automated processing of personal data**, as well as to the non-automated processing of personal data contained or to be contained in files, with the exception of processings performed for the exercise of exclusively personal activities [...].

The I&L law creates the CNIL (french data protection authority) and (in 2004) the CILs (*Correspondants Informatique et Libertés*, future "data protection officers").

Violations of the I&L law

Offence punished by 5 years' imprisonment and a 300 000 € fine ($\times$ 5 for organizations).

Informatique et Libertés: Perimeter of the law

Notion of personal data

Following of article 2 (still a non-official translation):

Personal data is **any piece of information relating to an identified or identifiable physical person**, directly or indirectly, by reference to an identification number or to one or several elements characteristic to them. In order to ascertain whether a person is identifiable, one must consider **the set of all identification means** to which **the data processor or any other person** has or may have access.

Before 2004 (and the application of directive 95/46), they spoke of **nominative** or **indirectly nominative information**.

The term **Personally identifying information** (PII) is sometimes used, but it refers to the US legal context. Plus, it is misleading (see the reste of the lecture...)

Informatique et Libertés: Principles

- **Legality principle :** It is forbidden to collect or process some kinds of data (see next slide for details);
- **Finality principle :** Data are collected with respect to a clear, declared finality, which must be respected;
- **Legitimity principle :** It must be legitimate for the data processor to pursue such a finality;
- **Proportionality principle :** Data collection must be proportional (nature, quantity, retention time) to the finality.

Informatique et Libertés: Principles

Sensitive data

It is forbidden to collect or process data about:

- Racial or ethnic origins;
- Political, philosophical, religious opinions;
- Union memberships;
- Health and sexual life.

Exceptions: express consent, human life protection, member list management, health services, official statistics, medical research, legal proceedings, "public interest" (restrictive notion).

Informatique et Libertés: Prior formalities

Declaration to CNIL

Default case (no sensitive data). No need for a declaration if organization has a CIL (data protection officer).

Authorization from CNIL

Collection or processing of sensitive data, genetic data, biometric data, data relating to offences or convictions, susceptible to deprive from a right, using the social security number...

Informatique et Libertés: Rights and obligations

Rights of data subjects

- The data processor is obligated to inform them about data collection, processing, retention, forwarding to third parties;
- Right to access and rectify the data;
- Right of opposition.

Obligations of the data processor

They must guarantee the legality, finality, legitimity and proportionality principles.
They must guarantee the rights of the data subject (notably via the "mandatory statements") and are responsible for the confidentiality of the data they keep.

The European regulation project

A few foreseen evolutions

- Reinforcement of the importance of consent (distinct and revokable consent for each processing);
- Reinforcement of a "right to be forgotten" (sometimes "right to oblivion") explicitely mentionned;
- More thorough obligations of logging and auditability for data processors (cf burden of proof);
- Obligation to notify (to data agencies and data subjects) any "privacy breach", possibly under 24h;
- Obligation to perform a "privacy impact assessment" before any request for authorization;
- Data protection officers become mandatory for public organizations and large companies. Their independence must be guaranteed;
- Graduated administrative sanctions (in addition to the national criminal sanctions) going up to 1 M€ or 2% of global turnover.