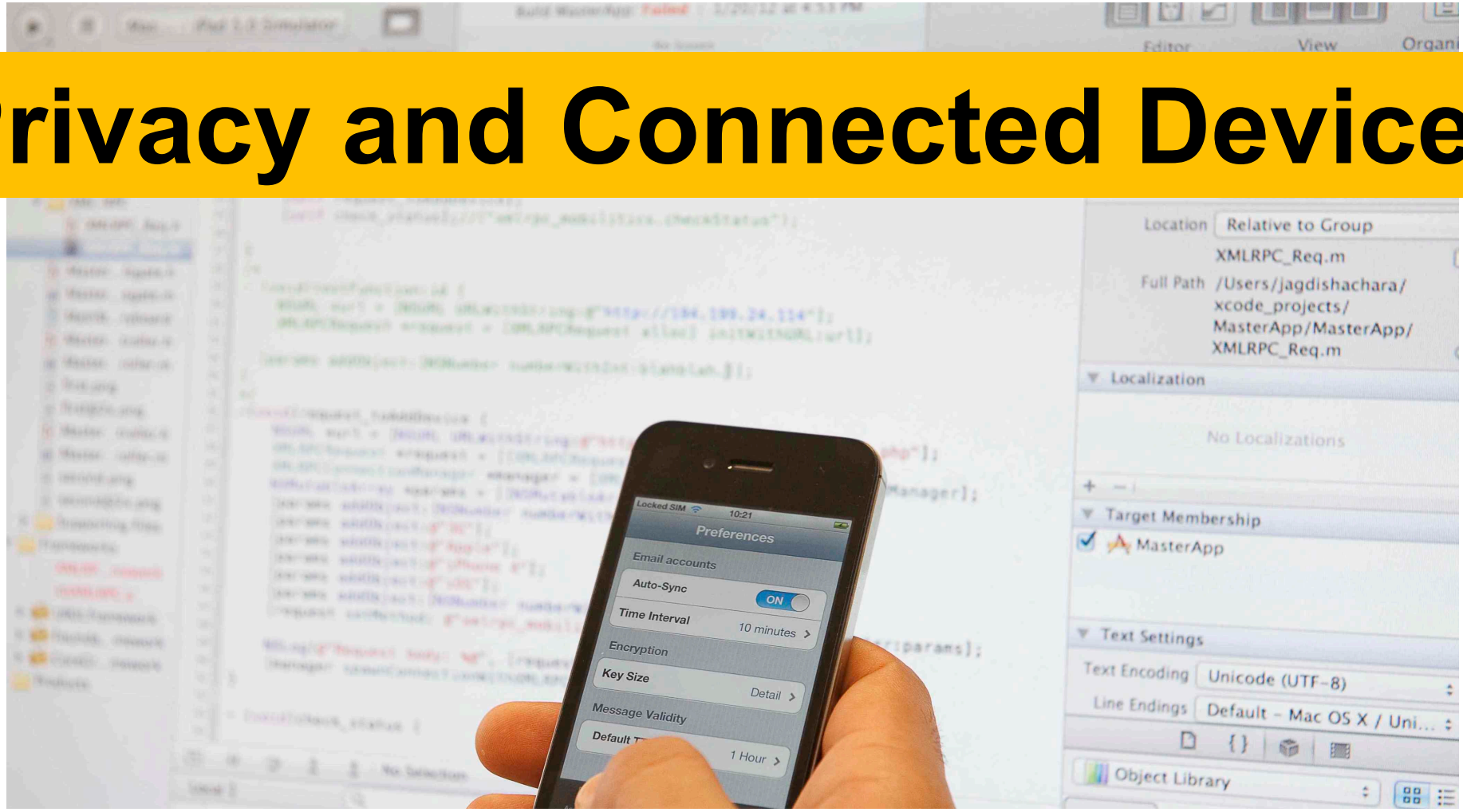
# Privacy and Connected Devices




© Inria / Photo H. Raguet

**Vincent Roca, Inria PRIVATICS, vincent.roca@inria.fr**

Eclipse IoT Days – Grenoble, January 19th, 2018

informatiques mathématiques

contact : vincent.roca@inria.fr

# Outline

- **Context: the IOTics ANR project**
- Personal information and the French/EU law
- A focus on smart-homes
- Conclusions

# Context: the IOTics ANR project (2017 – 2020)


PRIVATICS team **(leader)**
DIANA team

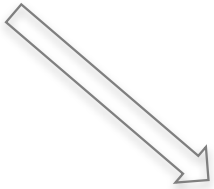
Computer scientists (privacy, network)


RITM team

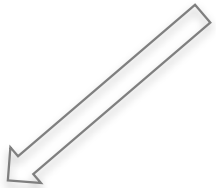
Economists


Digital Security department

Computer scientists
(embedded systems security)



IOTics


Lab. Innovation et
Prospective (LINC)

Regulator

Lawyers

Plus two close associates: **Université Versailles St-Quentin**, DANTE
**Université Paris-Sud**, CERDI

# Context: the IOTics ANR project (2)

- Focuses on **smart-homes**, **quantified self**, and **childhood** connected devices



- Goals
  - focus on personal information leaks
  - give transparent information of hidden behaviors
  - highlight good and bad practices
  - reduce information asymmetry
  - give back control to end-users
  - hopefully convince stakeholders to change practices

# Context: the IOTics ANR project (3)

- Three complementary directions
  - ✓ analysis of the **hidden personal information leaks** of a set of devices
    - within the **device** (breaking firmware security if needed to analyze it)
    - within the interconnection **network** (monitor/analyze data flows)
    - within the **smartphone** (monitor/analyze leaks through the app)
  - ✓ analysis of the **privacy policies** provided (or not) by the companies
  - ✓ analysis of the underlying **ecosystem**

EURECOM

Inria informatiques mathématiques

UNIVERSITÉ PARIS SUD

université PARIS-SACLAY

# Outline

- Context: the IOTics ANR project
- **Personal information and the French/EU law**
- A focus on smart-homes
- Conclusions

# Some vocabulary…

**Private company, administration**

**“data controller”**

**(responsable de traitements)**

*has the responsibility of*

Data Base containing

**“Personal Information”**

**(données à caractère personnel)**

*relative to*

**Physical persons**

# Personal information according to the “Loi informatique et libertés” of 1978 (1)

identity is not required as long as a path to a physical person can be found

**Article 2 :** […] Constitue une donnée à caractère personnel toute information relative à une **personne physique identifiée ou qui peut être identifiée, directement ou indirectement**, par référence à un numéro d’identification ou à un ou plusieurs éléments qui lui sont propres. Pour déterminer si une personne est identifiable, il convient de **considérer l’ensemble des moyens** en vue de permettre son identification dont dispose ou auxquels peut avoir accès le **responsable du traitement ou toute autre personne.** […]

http://www.cnil.fr/documentation/textes-fondateurs/loi78-17/

no limit on the technical means

no limit: anybody in the world

# Personal information according to the “Loi informatique et libertés” of 1978 (2)

- the **nature** of the information does not matter…
  - ✓ can be anything (e.g., temperature in a home)
- …if there is a **link to a person**, it’s a Personal Info (PI)

- this link can be **direct**…
  - ✓ e.g., we record temperature + name
- or **indirect**
  - ✓ e.g., we record temperature + EDF client ID

# Personal information according to the “Loi informatique et libertés” of 1978 (3)

- a person is considered **identifiable** if the data controller has the information to identify him
  - ✓ e.g., EDF collects your home temperature + EDF client ID
- or **anybody else** in the world
  - ✓ e.g., EDF collects your home temperature + IP address of the sensor.
  - ✓ Here the ISP can link the IP to the ADSL user

# The particular case of “sensitive information” (1)

it’s forbidden!

direct or indirect information

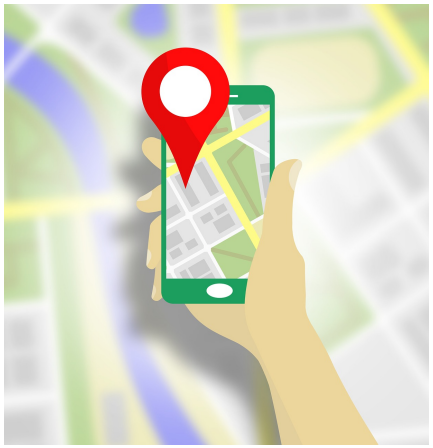
**Article 8 :** Il est **interdit** de collecter ou de traiter des données à caractère personnel qui font apparaître, **directement ou indirectement**, les origines raciales ou ethniques, les opinions politiques, philosophiques ou religieuses ou l'appartenance syndicale des personnes, ou qui sont relatives à la santé ou à la vie sexuelle de celles-ci. […]

http://www.cnil.fr/documentation/textes-fondateurs/loi78-17/

list of sensitive domains

# The particular case of “sensitive information” (2)

- Sensitive information cannot be collected and processed (except in a few particular cases).
  - ✓ The “Loi Informatique et Libertés” lists a few exceptions
  - ✓ ex. Health professionals and medical urgencies



- What about “**inferences**”?
  - ✓ in practice it’s pretty complex because of inference
  - ✓ if Google knows I’m at a church every Sunday morning (e.g., thanks to geolocation data), he knows something whose collection is prohibited

# Other viewpoints on personal information (1)

- The FR and European definitions of PI are in line and protective ☺

- In certain countries, the link between data and physical person is only considered for the **data controller**
  - ✓ Changes everything!
  - ✓ *e.g., if X collects "temperature + IP address of the sensor », data is not considered as PI unless X is an ISP…*

# Other viewpoints on personal information (2)

- Question 1: what about the following claim?

  **“we don’t collect your name, age or address, only non personal information”**

  ✓ wrong if linkability to a person remains possible

- Question 2: is an IP address a PI?

  ✓ yes in France and in EU

  ✓ no in the US, apart from the ISP

# Obligations for the data controller

**Article 6 :** […]

1° Les données sont collectées et traitées de **manière loyale et licite ;**

2° Elles sont collectées pour des **finalités déterminées, explicites et légitimes** et ne sont pas traitées ultérieurement de manière incompatible avec ces finalités. […] ;

3° Elles sont adéquates, pertinentes et **non excessives** au regard des finalités pour lesquelles elles sont collectées et de leurs traitements ultérieurs ; […]

5° Elles sont conservées sous une forme permettant l'identification des personnes concernées **pendant une durée qui n'excède pas** la durée nécessaire aux finalités […]

fair collection

well defined goal

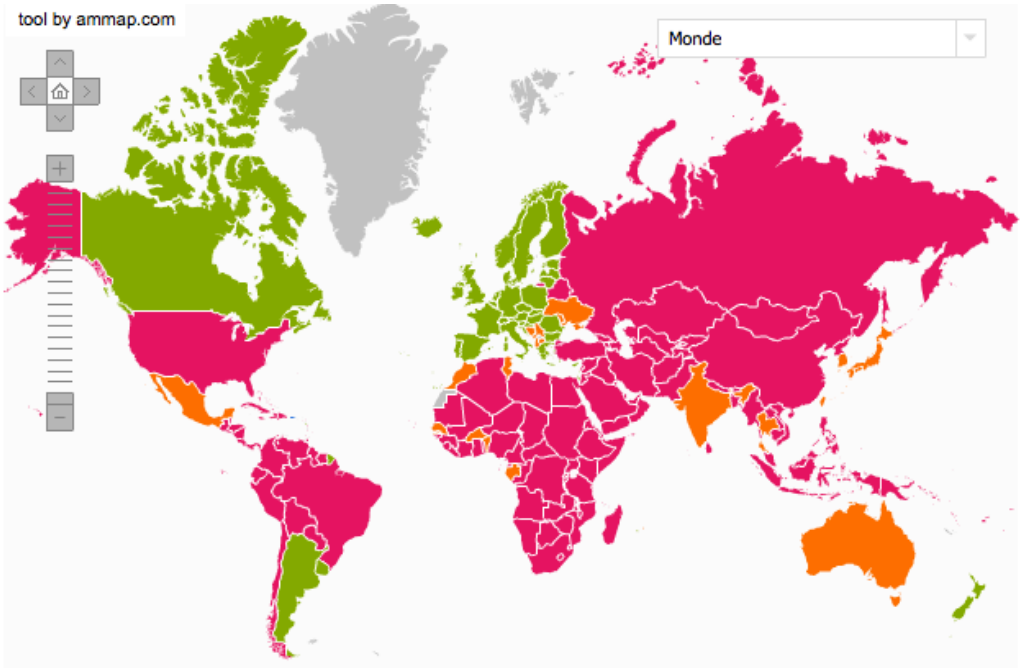
collect the bare minimum

limited duration

# PI transmission beyond EU (1)

- Personal information of EU citizens **cannot** be sent beyond EU borders.

- **Solution 1:** there are exceptions for countries whose data protection law is compliant with that of EU




https://www.cnil.fr/fr/la-protection-des-donnees-dans-le-monde

# PI transmission beyond EU (2)

- US is not concerned by this exception
  - ✓ US is not recognized as trustworthy W.R.T. PI protection

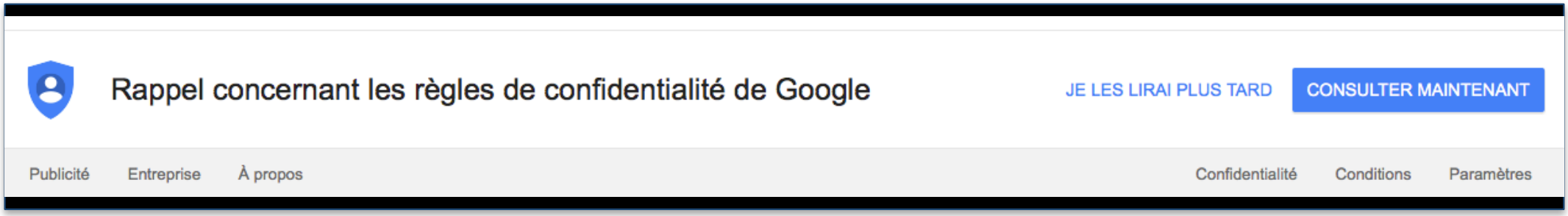
- **Solution 2:** join the **Privacy Shield** program that rules PI transfers to the US
  - ✓ The company commits to respect the contractual obligations
  - ✓ A previous program, “Safe Harbor”, has been canceled in 2015 by the EU Court of Justice: see the EUJC judgment (Max Schrems)

https://www.cnil.fr/fr/le-privacy-shield
http://www.cnil.fr/linstitution/actualite/article/article/invalidation-du-safe-harbor-par-la-cour-de-justice-de-lunion-europeenne-une-decision-cl/

# Ways to escape the PI rules (1)

The data collector can do much more if…

- **Solution 1:** he obtains the **"free and informed consent"** of the user
  - ✓ "consentement libre et éclairé"
  - ✓ explains why Google urges the user to read their confidentiality rules




- is it sufficient?
  - ✓ no if the user is not free to use the service (no alternative)
  - ✓ no if the privacy rules are not compliant with French / EU law (ex. Facebook)

# Ways to escape the PI rules (2)

- **Solution 2:** data is **anonymized**
  - if linkability to a person is impossible it is no longer PI
  - but secure anonymization can be pretty hard to achieve and not necessarily sufficient
    - ✓ because of **inference attacks with side information**
    - ✓ Example: *if a group of people is known to have a certain property, and if somebody knows I belong to this group then he knows I have this property too, even if my individual record cannot be identified in the database*

# Last but not least…


- **GDPR**, the EU **General Data Protection Regulation**, enters in application on May 25th, 2018

  - ✓ immediately and uniformly applicable throughout the European Union

  - ✓ additional rights and requirement to obtain an **explicit and positive consent** from users

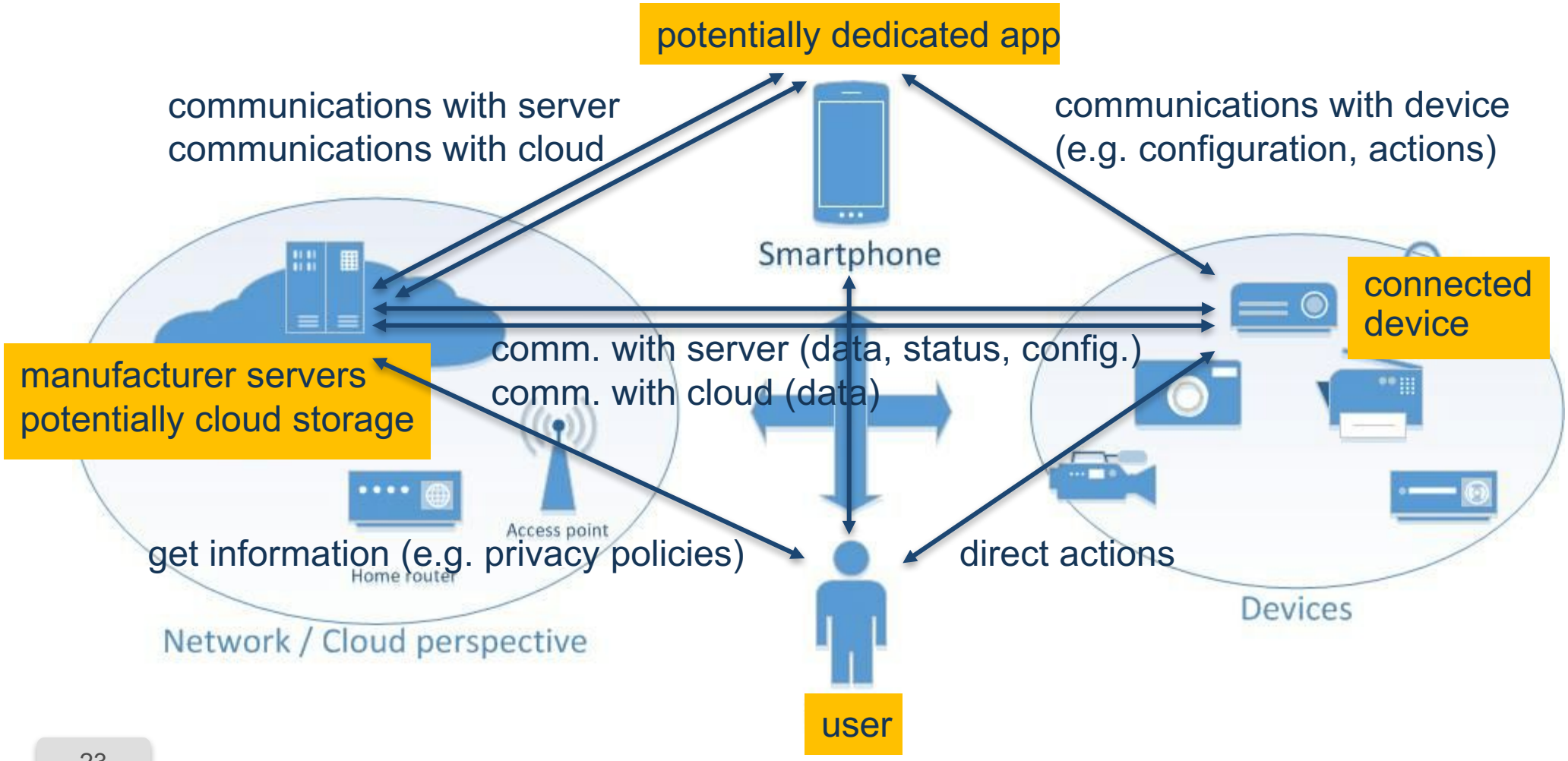
    it will make third-party ad server's work pretty complex!


  - ✓ above all a **sanction that can reach 4% of the annual worldwide turnover or 20 Million €**, whichever is greater

# Outline

- Context: the IOTics ANR project
- Personal information and the French/EU law
- **A focus on smart-homes**
- Conclusions

# Smart-homes are complex

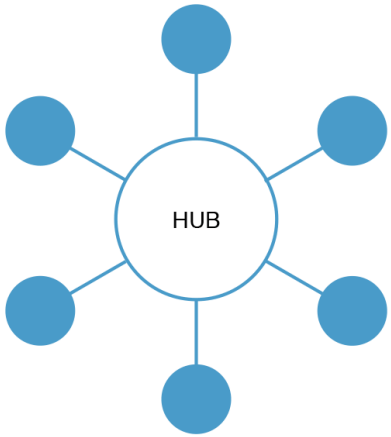


# About smart-home ecosystems

- Several ecosystems exist, setup by **device manufacturers**
  - ✓ NEST
  - ✓ TP-Link Smarthome
  - ✓ Philips Hue
- … or by voice-controlled **smart-speaker manufacturers**
  - ✓ Amazon Echo
  - ✓ Google Home
- … or **smartphone manufacturers**
  - ✓ Apple Home
  - ✓ Samsung SmartThings
- … and **service providers**
  - ✓ Hubnumérique from La Poste


# About smart-home ecosystems (2)

- Who will provide the **main hub** to smart homes (integration)?
  - ✓ at the **center** of all devices, no matter their manufacturer
  - ✓ **aware** of all events of interest
  - ✓ **collecting** large amounts of data
  - ✓ capable of **influencing** users




- What's the **business model/motivation?**
  - ✓ get revenue by selling hardware…
  - ✓ better **know** the user and its equipment
    - useful for other company services, cross device tracking, profiling, etc.
  - ✓ **influence** the user
    - Amazon Echo users will probably use Amazon store

# A smart building for sale

- http://www.zdnet.fr/actualites/batiment-connecte-livraison-d-un-immeuble-d-habitation-une-premiere-en-france-39859052.htm
  - "ces logements communicants permettent aux résidents de contrôler par la voix ou depuis leur smartphone l'accès à la résidence, le chauffage, les éclairages et les appareils électriques"
  - "Les logements sont contrôlables depuis l'application **« Maison » d'Apple** (**entre autre**) et compatibles avec le **hub numérique de La Poste**. **Legrand** (portier connecté) et **Netatmo** (prises, thermostat et interrupteurs connectés) participent au projet côté infrastructure et connectivité.



## Bâtiment connecté : livraison d'un immeuble d'habitation, une première en France

***Technologie :*** *La résidence Issy Préférence, composée de logements communicants et connectés, vient d'être livrée. Les logements sont contrôlables depuis une application sur smartphone ou par la voix.*

Par La rédaction de ZDNet.fr | Mardi 24 Octobre 2017

C'est à Issy-les-Moulineaux, en banlieue parisienne, que vient d'être inauguré le premier immeuble d'habitation (du studio au 5 pièces) dont la soixantaine d'appartements proposés à la vente sont équipés d'une installation électrique et d'un chauffage connectés.

# Many questions raised by those ecosystems!

- Q1(technical viewpoint): how to **interconnect** heterogeneous devices?
  - ✓ different wireless communication techs, protocol stacks (Thread?) and configuration approaches
- Q2 (security viewpoint): how is **security** managed?
  - ✓ what's the price to pay for a smooth integration of heterogeneous devices?
  - ✓ each device comes with its security model. Does their interconnection imply a lowest common denominator security?
- Q3 (technical viewpoint): **where** does data go?
  - mapping the various data flows between devices, smartphone apps, remote servers, cloud-based storage/services.

# Many questions raised by those ecosystems (2)

- Q4 (legal viewpoint): **who** is(are) the data controller(s)?
  - perhaps all the stakeholders are!
  - how to trust them all?
- Q5 (legal viewpoint): how to bring **transparency**?
  - importance of clear and readable Privacy Policies
- Q6 (legal viewpoint): how to bring **user-control**?
  - far from obvious
- Q6 (legal viewpoint): how to provide **accountability**?
  - a key requirement for trust

# Outline

- Context: the IOTics ANR project
- Personal information and the French/EU law
- A focus on smart-homes
- **Conclusions**

# About legal aspects…

- Notions of **personal information** and **sensitive information**
  - ✓ are the foundation of laws that protect privacy in EU
- A data controller that owns PI must comply with several key **obligations**
- PI transmission **beyond EU** is sometimes possible but laws exist that protect it
- In order to **escape** these obligations
  - ✓ get the “free and informed consent” of the users;
  - ✓ or anonymize the database.
- **GDPR** frightens many stakeholders
  - ✓ explicit and positive consent, potential financial sanctions in case of infringements

# About conn. devices – the case of smart homes

- Many companies compete to **be at the center** of an ecosystem
  - ✓ it's key for business
- Current situation is pretty complex and obscure
  - ✓ complexity and trust are hard to reconcile!
- Raises many questions
  - ✓ how is **security** managed when interconnecting heterogeneous devices? A lowest common denominator security model?
  - ✓ what are the **data flows** and who are the **data controllers**?
  - ✓ how to bring **transparency, user control, and accountability**?

# *Thank you… ☺*

Includes content from Mathieu Thiery (Inria, PhD)