A generic framework for information and consent for the IoT

Mathieu Cunche, Daniel Le Métayer, and Victor Morel

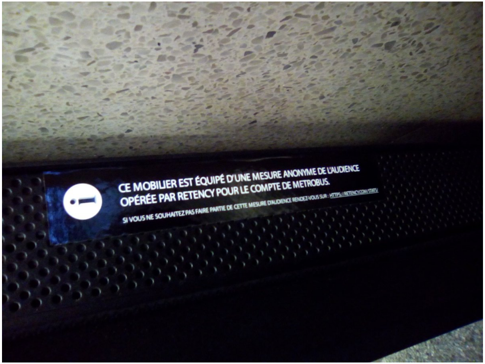
Inria

victor.morel@inria.fr

July 26, 2019 TRUSTCOM19

(Mis)Information and (Forced) Consent

[1]

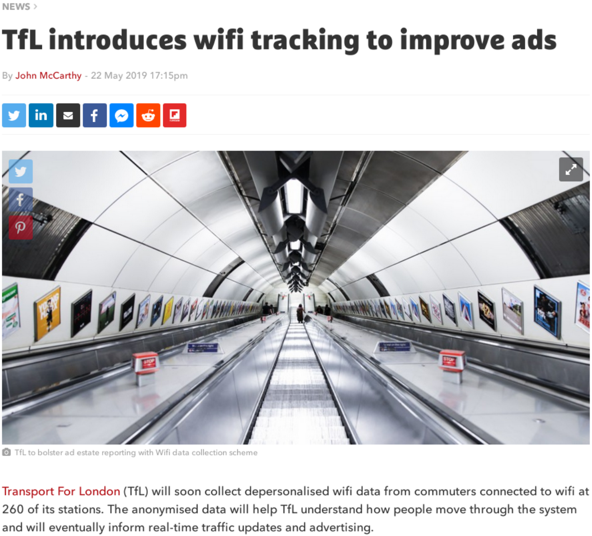

CE MOBILIER EST ÉQUIPÉ D'UNE MESURE ANONYME DE L'AUDIENCE
OPÉRÉE PAR RETENCY POUR LE COMPTE DE METROBUS.
SI VOUS NE SOUHAITEZ PAS FAIRE PARTIE DE CETTE MESURE D'AUDIENCE RENDEZ-VOUS SUR : HTTPS://RETENCY.COM/STATS/

[2]

[1] https://twitter.com/_LoboTom_/status/1109106043706109952

[2] https://www.liberation.fr/checknews/2019/03/25/les-panneaux-de-pub-du-metro-tracent-ils-les-telephones-des-usagers_1717316

No need to ask... ?



NEWS ›

TfL introduces wifi tracking to improve ads

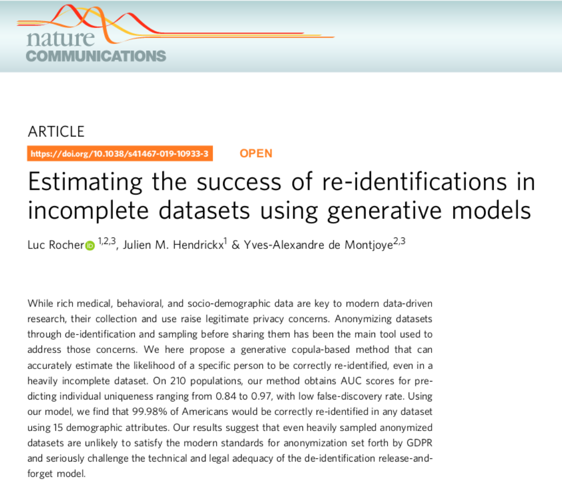
By John McCarthy - 22 May 2019 17:15pm

TfL to bolster ad estate reporting with Wifi data collection scheme

Transport For London (TfL) will soon collect depersonalised wifi data from commuters connected to wifi at 260 of its stations. The anonymised data will help TfL understand how people move through the system and will eventually inform real-time traffic updates and advertising. [3]

[3] https://www.thedrum.com/news/2019/05/22/tfl-introduces-wifi-tracking-improve-ads

“Anonymized”



nature COMMUNICATIONS

ARTICLE

https://doi.org/10.1038/s41467-019-10933-3 OPEN

Estimating the success of re-identifications in incomplete datasets using generative models

Luc Rocher [1,2,3], Julien M. Hendrickx [1] & Yves-Alexandre de Montjoye [2,3]

While rich medical, behavioral, and socio-demographic data are key to modern data-driven research, their collection and use raise legitimate privacy concerns. Anonymizing datasets through de-identification and sampling before sharing them has been the main tool used to address those concerns. We here propose a generative copula-based method that can accurately estimate the likelihood of a specific person to be correctly re-identified, even in a heavily incomplete dataset. On 210 populations, our method obtains AUC scores for predicting individual uniqueness ranging from 0.84 to 0.97, with low false-discovery rate. Using our model, we find that 99.98% of Americans would be correctly re-identified in any dataset using 15 demographic attributes. Our results suggest that even heavily sampled anonymized datasets are unlikely to satisfy the modern standards for anonymization set forth by GDPR and seriously challenge the technical and legal adequacy of the de-identification release-and-forget model.

[4]

[4] Rocher, Hendrickx, and de Montjoye, 2019.

GDPR

- General Data Protection Regulation in May 2018
- Relevant guidelines for privacy protection
  - → Transparency about data collection and processing
  - → Information required for consent
- DC: Data controller (legally responsible)
- DS: Data subject (in other words: user)

Challenges for the GDPR in the IoT

- Toothless if not complemented with proper technologies
- The Internet of Things presents difficulties
  - → Numerous devices, various uses
  - → Limited capacities, inappropriate/non-existent interfaces

How do we inform and manage consent in the IoT?

- Intelligible and systematic information of DS?
- Privacy-preserving and demonstrable consent?

A generic framework

Provides facilities for the following requirements:

- With respect to information
  - → To declare DC devices
  - → To receive information by DS
  - → To facilitate understanding
- With respect to consent
  - → To minimize fatigue
  - → To ensure data is collected iff consent is provided
  - → To facilitate demonstration of obtention of consent

Unambiguity of formal semantics

Generic in the sense that...

Small number of technical requirements

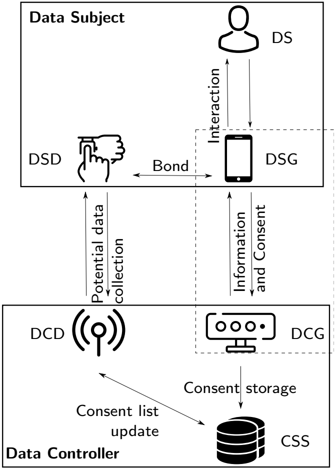
- Agnostic of collection protocol
- Agnostic of types of devices
- Agnostic of fielding configurations

Actors are represented by devices

- DC by DCG
- DS by DSG

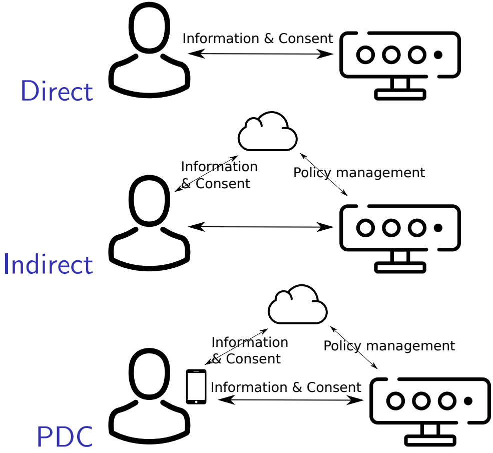
Different manners to implement the framework

Visual representation

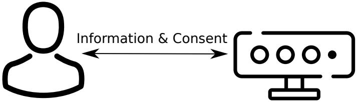


Technical options

Three complementary components:

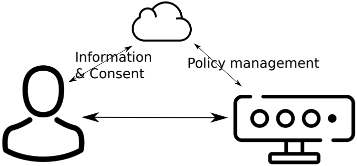


Direct communications



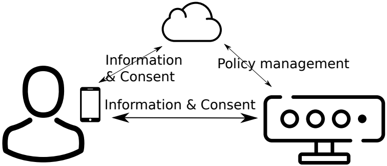
- DC policy broadcasted via beacons
- Communication is local and P2P
- Can use Bluetooth Low Energy (BLE)
- Consent can be sent using Attribute Protocol (ATT)

Indirect communications



- DC registries for information
- DS registries for consent
- Information can be *a priori*
- Especially appropriate when interaction not needed

Personal Data Custodian



- Enable interactions with DS
- Retrieve information
- Manage consent
- Definition of DS policy
- Can use PILOT privacy language[5]

[5] "Analysis of Privacy Policies to Enhance Informed Consent (Extended Version)".

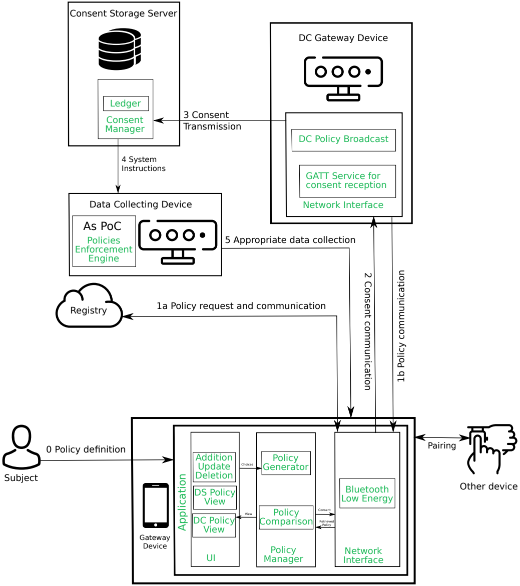
A design space

Provides guidelines for implementations

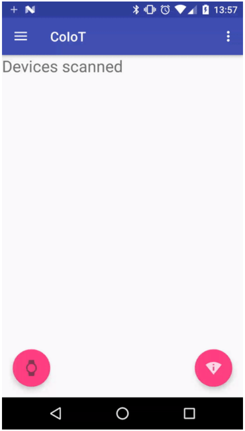
Examples

- Informing about passive sensors?
  - → Use an additional device or indirect communications
- Informing about moving sensors?
  - → Prefer direct communications
- Device with scarce resources?
  - → Direct communications without pairing are not possible

CoIoT


Consent Storage Server
Ledger
Consent Manager
3 Consent Transmission
4 System Instructions
DC Gateway Device
DC Policy Broadcast
GATT Service for consent reception
Network Interface
Data Collecting Device
As PoC
Policies Enforcement Engine
5 Appropriate data collection
Registry
1a Policy request and communication
2 Consent communication
1b Policy communication
Subject
0 Policy definition
Gateway Device
Application
Addition Update Deletion
DS Policy View
DC Policy View
UI
Choices
View
Policy Generator
Policy Comparison
Policy Manager
Consent
Retrieved Policy
Bluetooth Low Energy
Network Interface
Pairing
Other device

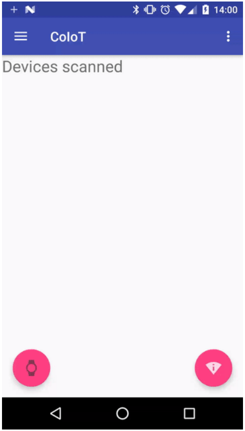
Retrieving policies



ColoT

Devices scanned

Figure: Scan

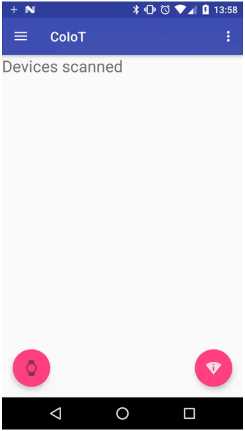


ColoT

Devices scanned

Figure: Registry

Managing Data Subject Policy

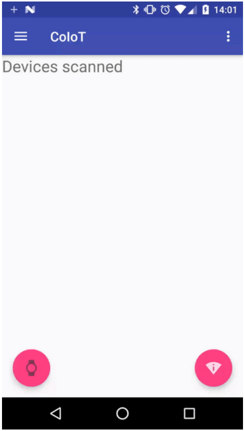


13:58

ColoT

Devices scanned

Figure: Rules



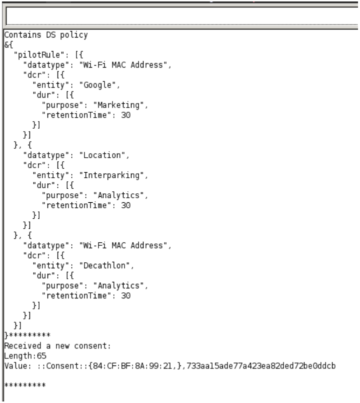
14:01

ColoT

Devices scanned

Figure: My DSP

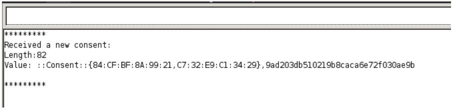
Consent and negotiation



```
Contains DS policy
&{
  "pilotRule": [{
    "datatype": "Wi-Fi MAC Address",
    "dcr": [{
      "entity": "Google",
      "dur": [{
        "purpose": "Marketing",
        "retentionTime": 30
      }]
    }]
  }, {
    "datatype": "Location",
    "dcr": [{
      "entity": "Interparking",
      "dur": [{
        "purpose": "Analytics",
        "retentionTime": 30
      }]
    }]
  }, {
    "datatype": "Wi-Fi MAC Address",
    "dcr": [{
      "entity": "Decathlon",
      "dur": [{
        "purpose": "Analytics",
        "retentionTime": 30
      }]
    }]
  }]
}*********
Received a new consent:
Length:65
Value: ::Consent::{84:CF:BF:8A:99:21,},733aa15ade77a423ea82ded72be0ddcb

*********
```

Figure: Negotiation



```
*********
Received a new consent:
Length:82
Value: ::Consent::{84:CF:BF:8A:99:21,C7:32:E9:C1:34:29},9ad203db510219b8caca6e72f030ae9b

*********
```

Figure: Consent for two devices

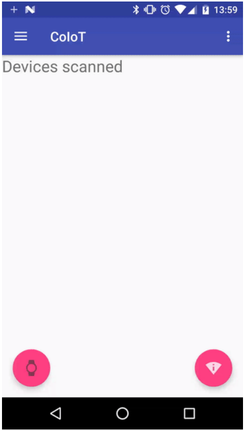
Presentation past/future work

To conclude

- A generic framework for information and consent
- Feasible options
- Prototype: CoIoT
- Work to do on consent signature & ledger
- Impact on the European ePrivacy directive?

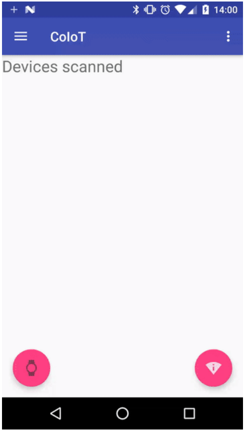
- Thank you for your attention
- Check me out: `http://perso.citi-lab.fr/vmorel/`

Other features of ColoT



Devices scanned

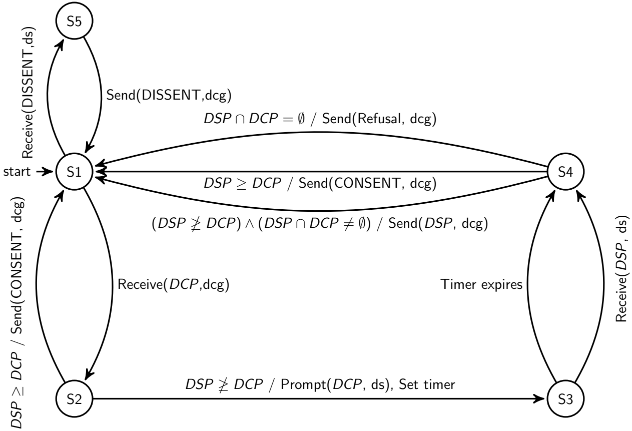
Figure: Bond



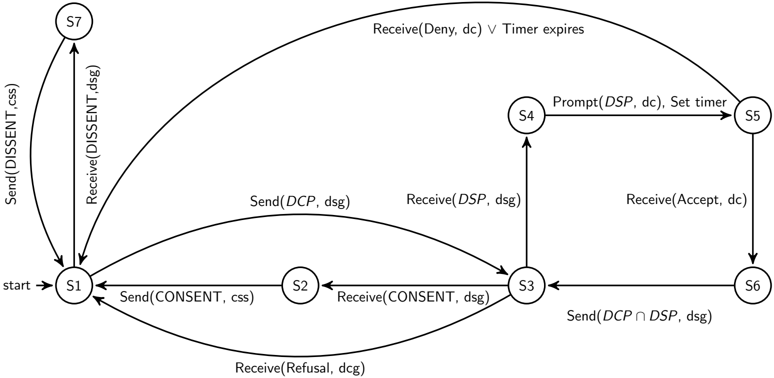
Devices scanned

Figure: Generic rules

Data Subject Gateway


S5
Receive(DISSENT,ds)
Send(DISSENT,dcg)
$DSP \cap DCP = \emptyset$ / Send(Refusal, dcg)
start
S1
S4
$DSP \geq DCP$ / Send(CONSENT, dcg)
$(DSP \ngeq DCP) \wedge (DSP \cap DCP \neq \emptyset)$ / Send(DSP, dcg)
$DSP \geq DCP$ / Send(CONSENT, dcg)
Receive(DCP,dcg)
Timer expires
Receive(DSP, ds)
S2
$DSP \ngeq DCP$ / Prompt(DCP, ds), Set timer
S3

Data Controller Gateway



Policies match

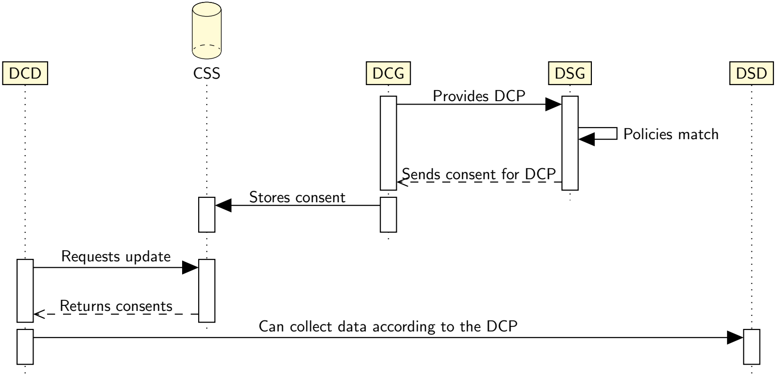


Figure: Policies match

Request interaction from DS

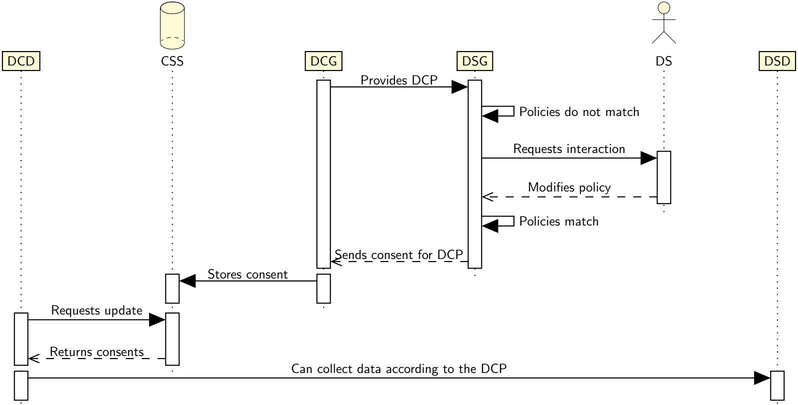


Figure: The policies do not match at first, the *DSG* requests an interaction from the data subject. The modification results in a match.

No collection

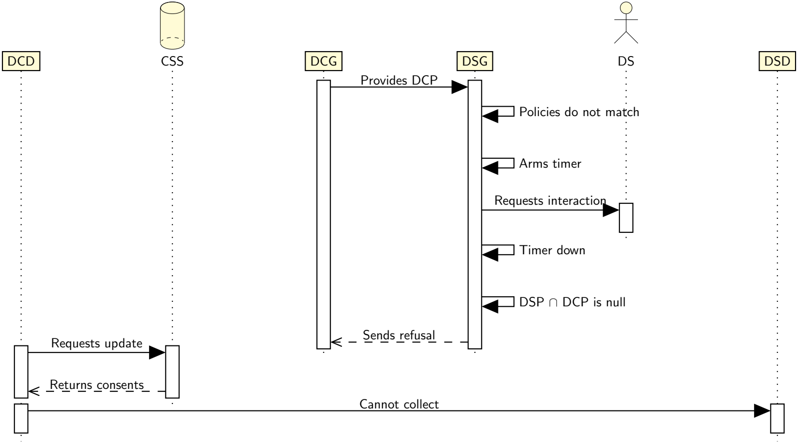


Figure: The policies do not match, and the data subject does not interact

Intersection

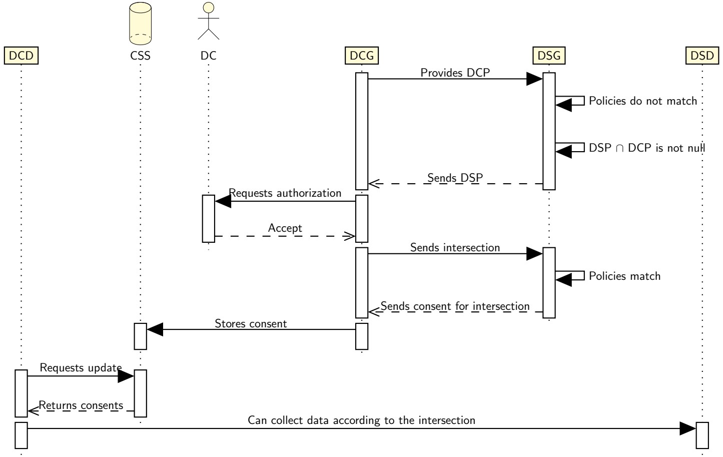


Figure: The policies do not match, but an agreement is made on the intersection of policies

Similar projects

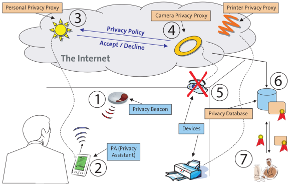
Smart places

THIS VENUE USES LOCATION INFORMATION FROM MOBILE DEVICES

For more information and your choices, visit

WWW.SMART-PLACES.ORG

PawS



Wombat

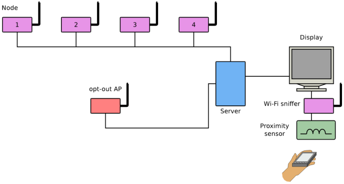
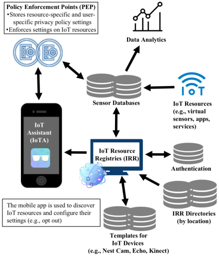


Figure VII.1 – Architecture of the Wombat system in a demonstration configuration.

PPA for IoT


Policy Enforcement Points (PEP)
•Stores resource-specific and user-specific privacy policy settings
•Enforces settings on IoT resources
Data Analytics
Sensor Databases
IoT Resources (e.g., virtual sensors, apps, services)
IoT Assistant (IoTA)
IoT Resource Registries (IRR)
Authentication
IRR Directories (by location)
The mobile app is used to discover IoT resources and configure their settings (e.g., opt out)
Templates for IoT Devices (e.g., Nest Cam, Echo, Kinect)

Limitations of the SotA

Flaws of the related work

- Cost → heavy infrastructure
- Scalability → heavy infrastructure
- User interaction → no negotiation
- Flexibility → lack of granularity
- GDPR compliance → framework devised to this end

PawS

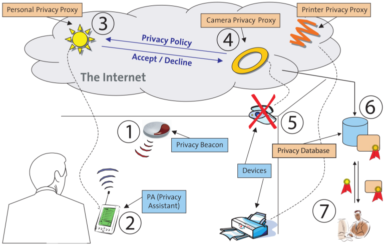


Figure: PawS architecture

- Early work[6]
- Distinction between assistant and proxy
- P3P for privacy language
  - → No negotiation
  - → Requires prior knowledge

[6] Langheinrich, 2002.

Mobile location analytics opt-out

Wombat[7]

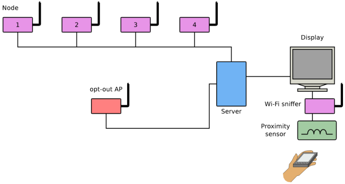
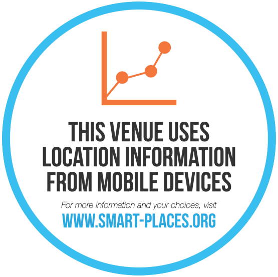


Figure VII.1 – Architecture of the Wombat system in a demonstration configuration.

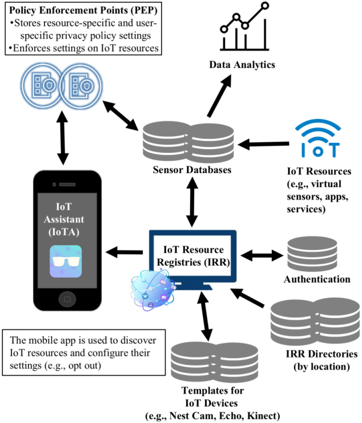
Smart places [8]


THIS VENUE USES LOCATION INFORMATION FROM MOBILE DEVICES
For more information and your choices, visit
WWW.SMART-PLACES.ORG

[7] "Wombat: An Experimental Wi-Fi Tracking System".

[8] https://smart-places.org/

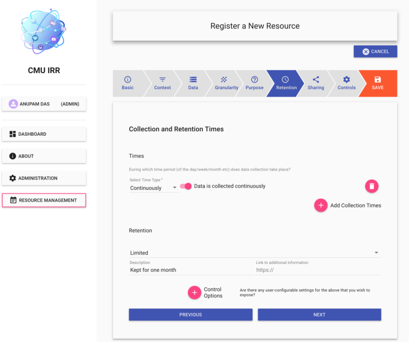
Personalized Privacy Assistant for the IoT


Policy Enforcement Points (PEP)
•Stores resource-specific and user-specific privacy policy settings
•Enforces settings on IoT resources
Data Analytics
Sensor Databases
IoT
IoT Resources (e.g., virtual sensors, apps, services)
IoT Assistant (IoTA)
IoT Resource Registries (IRR)
Authentication
IRR Directories (by location)
The mobile app is used to discover IoT resources and configure their settings (e.g., opt out)
Templates for IoT Devices (e.g., Nest Cam, Echo, Kinect)

- CMU project[9]
- IoT Resource Registries
- IoT Assistant
- Possibility to set privacy *preferences* through an assistant
  - → Registration mandatory
  - → GDPR?

[9] Das et al., 2018.

IoT registry



Register a New Resource

Collection and Retention Times

Times

During which time period (of the day/week/month etc) does data collection take place?

Select Time Type *: Continuously

Data is collected continuously

Retention

Limited

Description: Kept for one month

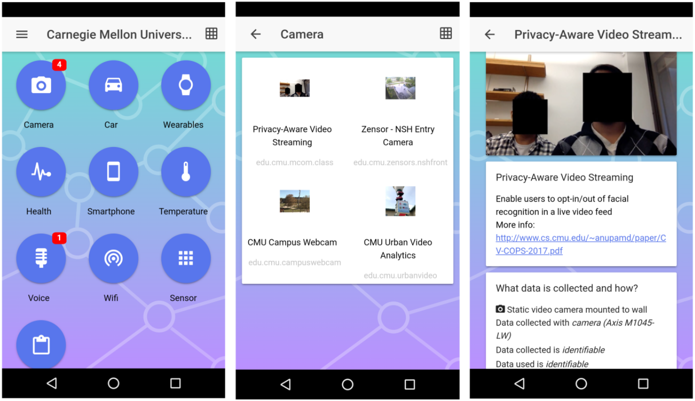
Link to additional information: https://

Are there any user-configurable settings for the above that you wish to expose?

- CMU project[10]
- Smart building
- Online registry of devices
- Information about data collection and processing
  - → Costly and specific
  - → Heavy infrastructure

[10] Pappachan et al., 2017.

IoT Assistant



Privacy-Aware Video Streaming

Enable users to opt-in/out of facial recognition in a live video feed
More info:
http://www.cs.cmu.edu/~anupamd/paper/CV-COPS-2017.pdf

What data is collected and how?

Static video camera mounted to wall
Data collected with *camera (Axis M1045-LW)*
Data collected is *identifiable*
Data used is *identifiable*

Figure: Privacy Assistant of CMU

A design space

Table: Technical options for information as a function of the DC device

Features of DC device	Direct communications without beacon	Direct communications with beacon	Indirect communications
Passive sensor	✗		
Active sensor with extensible protocol		(✗)	
Active sensor without extensible protocol	✗		
Fixed sensor			
Moving sensor			(✗)

A design space

Table: Technical options for consent as a function of the DS device

Features of the DS device	Direct communications without pairing	Direct communications with pairing	Indirect communications	A priori enforcement	A posteriori enforcement
Device with extensible protocol		(✗)	(✗)		
Device without extensible protocol	✗				
Device with substantial resources		(✗)	(✗)		
Device with scarce resources	✗				
Systematic collection process				✗	
Selective collection process					(✗)
Pre-existing relationship					
No pre-existing relationship			✗		