Security and PrIvacy foR the Internet of Things

SPIRIT

Gareth Howells

University of Kent

La Rochelle Université

Université de Genève

University of Essex

Consortium

University of Kent

Lead partners, ICMetric technology

University of La Rochelle

Content based signatures

Semantic Firewall

University of Geneva

Data Analysis / Semantic Extraction

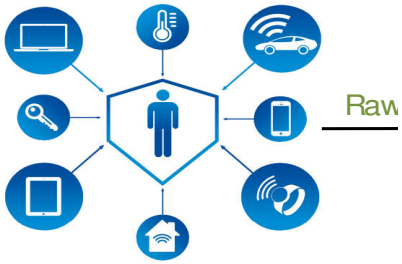
University of Essex

ICMetric technology, Demonstrator integration

Background

- Enhance trust and integrity of IoT technology
- Address lack of user confidence in the technology
- Specifically
  - How to ensure data originates from claimed device?
  - How to ensure it has not been altered?
  - How to ensure users comply with security requirements?
  - How to ensure the solution is easy and cheap to deploy?
  - How to ensure system may evolve to changing requirement?

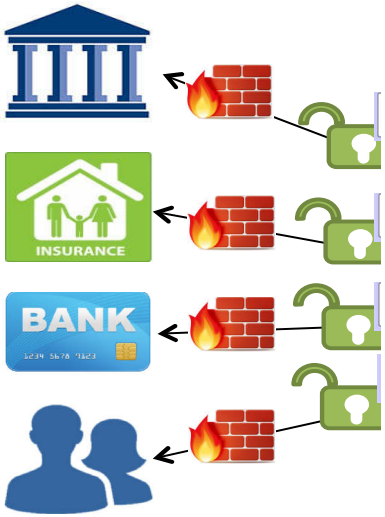
System Overview



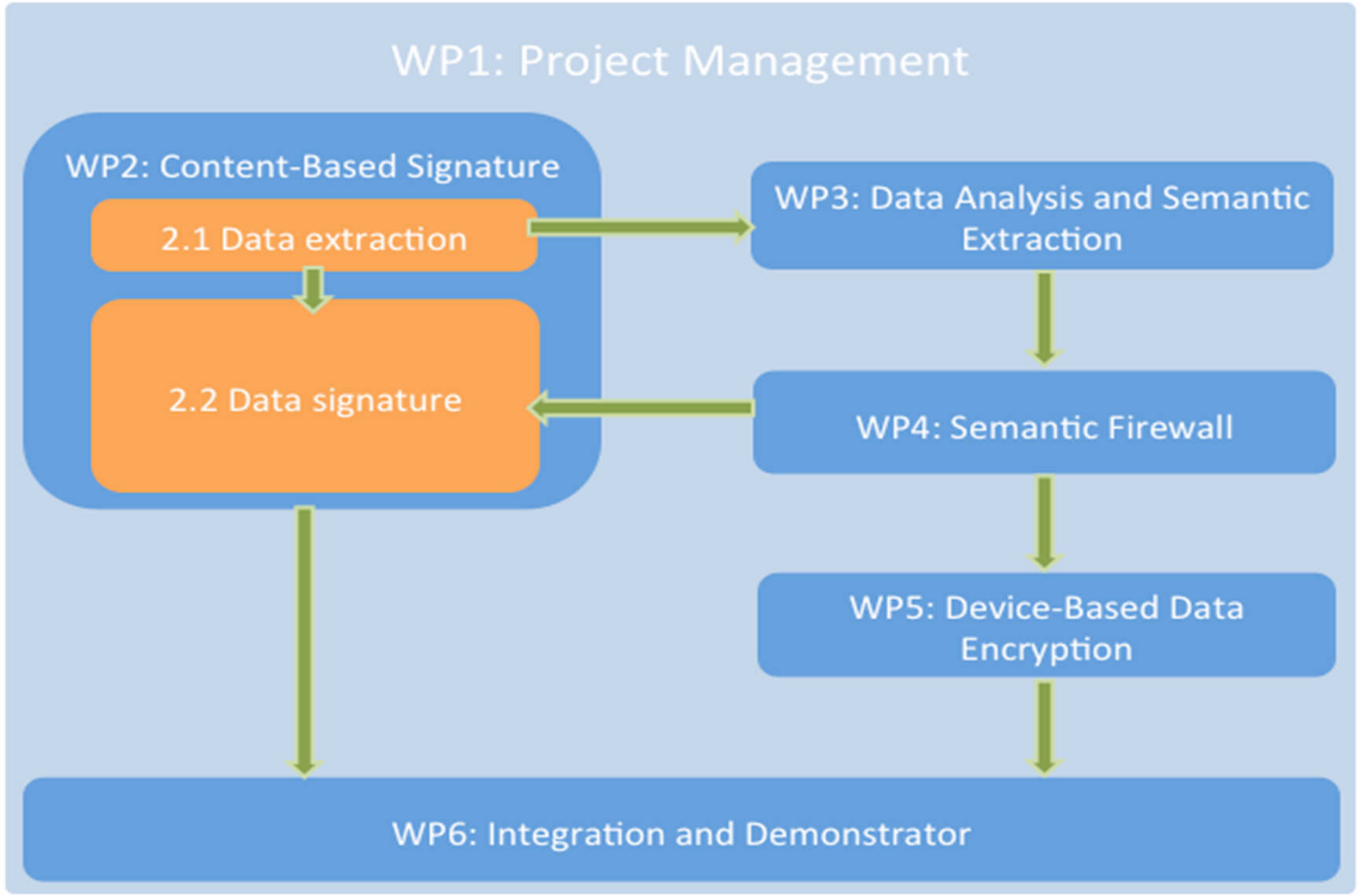
Raw

User's IoT devices
Data producer

IoT Hub



Data consumers


WP1: Project Management
WP2: Content-Based Signature
2.1 Data extraction
2.2 Data signature
WP3: Data Analysis and Semantic Extraction
WP4: Semantic Firewall
WP5: Device-Based Data Encryption
WP6: Integration and Demonstrator

WP2 – Content-based Signature

Partners: Univ. La Rochelle + Univ. Geneva

Main goal:

- Computing a unique content-based signature using the content of IoT device data

- First year
    - Decomposing the layout of a document in order to separate the different content composing a page
    - Proposing a layout descriptor based on this decomposition
    - Hiring a post-doc

WP2 – Content-based Signature

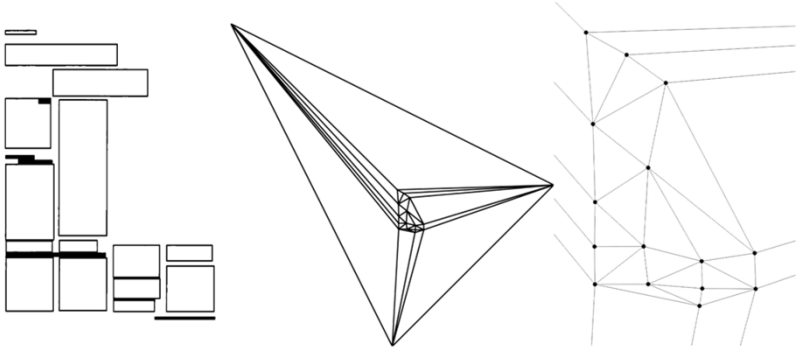
Task 2.1: Data Document analysis: layer decomposition

- First step: Identifying the regions of a document
- A survey has been done to highlight the variety of the approaches since 2008
- We identified that the physical layout of a document refers to the physical location of various regions in the document image

Task 2.2: Classification and extraction of content

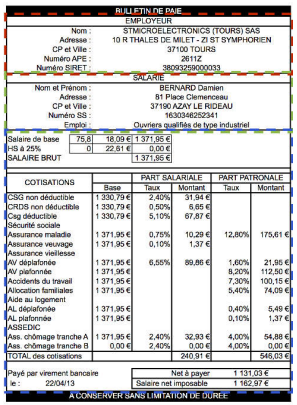
Task 2.3: Testing and optimization of data extraction

- This work has started lately
- Design of a layout descriptor without any threshold
- Creation of a new public dataset for forgery detection and spotting for the validation process



Dataset

Payslip-like documents.
5 fonts, 4 text sizes
20 pairs generated
Fixed layout, different data:
Company information
Employee information
Wage info information



BULLETIN DE PAIE

EMPLOYEUR

Nom : STMICROELECTRONICS (TOURS) SAS
Adresse : 10 R THALES DE MILET - ZI ST SYMPHORIEN
CP et Ville : 37100 TOURS
Numéro APE : 2611Z
Numéro SIRET : 38093259000033

SALARIE

Nom et Prénom : BERNARD Damien
Adresse : 91 Place Clemenceau
CP et Ville : 37190 AZAY LE RIDEAU
Numéro SS : 1630346252341
Emploi : Ouvriers qualifiés de type industriel

Salaire de base	75,8	18,09 €	1 371,95 €
HS à 25%	0	22,81 €	0,00 €
SALAIRE BRUT			1 371,95 €

COTISATIONS	Base	PART SALARIALE Taux	PART SALARIALE Montant	PART PATRONALE Taux	PART PATRONALE Montant
CSG non déductible	1 330,79 €	2,40%	31,94 €		
CRDS non déductible	1 330,79 €	0,50%	6,65 €		
Csg déductible	1 330,79 €	5,10%	67,87 €		
Sécurité sociale					
Assurance maladie	1 371,95 €	0,75%	10,29 €	12,80%	175,61 €
Assurance veuvage	1 371,95 €	0,10%	1,37 €		
Assurance vieillesse					
AV déplafonée	1 371,95 €	6,55%	89,86 €	1,60%	21,95 €
AV plafonnée	1 371,95 €			8,20%	112,50 €
Accidents du travail	1 371,95 €			7,30%	100,15 €
Allocation familiales	1 371,95 €			5,40%	74,09 €
Aide au logement					
AL déplafonée	1 371,95 €			0,40%	5,49 €
AL plafonnée	1 371,95 €			0,10%	1,37 €
ASSEDIC					
Ass. chômage tranche A	1 371,95 €	2,40%	32,93 €	4,00%	54,88 €
Ass. chômage tranche B	0,00 €	2,40%	0,00 €	4,00%	0,00 €
TOTAL des cotisations			240,91 €		546,03 €

Payé par virement bancaire — Net à payer 1 131,03 €
le : 22/04/13 — Salaire net imposable 1 162,97 €

A CONSERVER SANS LIMITATION DE DUREE

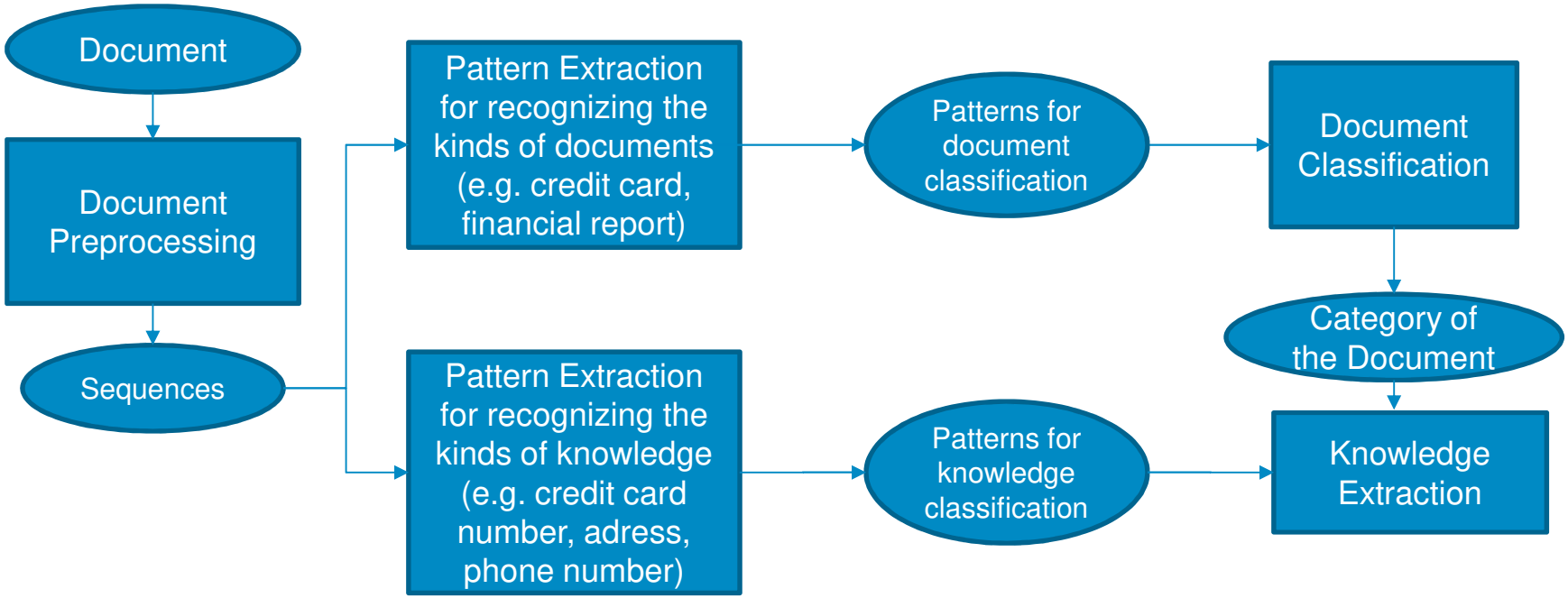
Objectives or Work-Package WP3

Partners: Univ. Geneva + Univ. La Rochelle

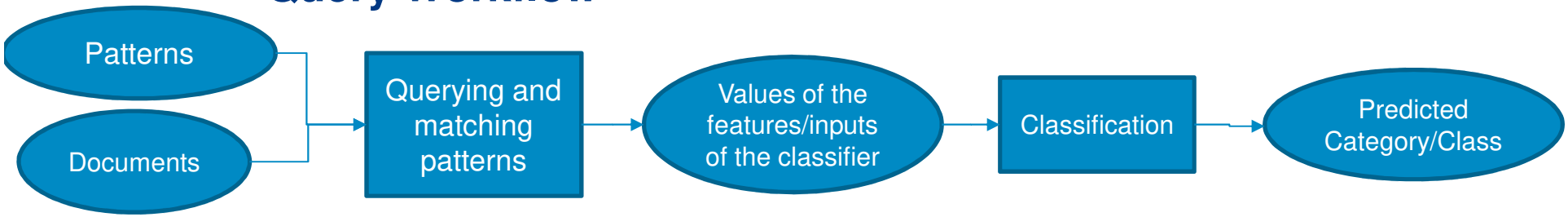
Objectives	Document Analysis (for each category of *Documents*)	Knowledge Extraction (for each category of *Knowledge*)
Preprocessing	X (only once for a given corpus)	—
Pattern Extraction	X	X
Querying and Matching patterns	X	X
Training a classifier and model selection	X	X
Testing the chosen (the selected/best model) model/classifier	X	X

Workflow WP3

Analysis Workflow



Query Workflow



Summary of work-package WP3

Task	Description
Document Preprocessing	• Cleaning the corpus; • POS tagging, deep syntactic analysis. • Transforming our data into sets of sequences; • Dividing the original dataset into (I) a training dataset and (II) a testing dataset.
Pattern mining (for Feature Extraction)	• Pattern mining algorithm for extracting frequent and maximal patterns satisfying gap constraints provided by the user.
Pattern matching for querying purposes	• Finding which patterns are included in a sequence.
Feature selection	• Extracting patterns that are representative of a class of documents; • Using FCA (Formal Concept Analysis) to identify the most discriminative sets of patterns.
Classification	• FCA-based classifier; • Neural network classifier.

Work-package 4

Partners: Univ. La Rochelle + Univ Kent + Univ Essex

Design and implementation of a semantic firewall based on a user-centric approach for personal data protection and privacy preservation within an IoT environment

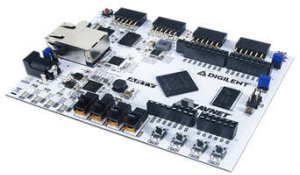
- Rules for Personal Data Protection
- Privacy Knowledge Base Description
- Privacy Policy Infringement Detection & Protection
- Reasoning mechanisms for access control

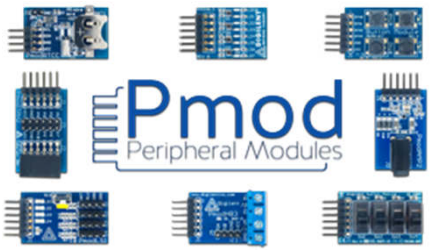
Objective of Work-Package WP5

Partners: Univ Kent + Univ Essex

To develop practical and secure technologies for IoT devices, which lock-down the communication between the semantic firewall and IoT devices on its network.

Task		Acomplishments
5.1		✓ Programmable embedded platform developed ✓ Various IoT Sensors investigated: MEMs gyros, accelerometers, ambient light sensors, etc. ✓ Signal processing techniques developed and investigated to stabilise readings
5.2		✓ Mathematical model derived to validate and model the data sets based on the Fisher transform and Probability Central Limit theorem.





Summary

WP1: Management ongoing

WP2: First results available. Improvements in progress

WP3: Document Classifiers complete

Alternative analysis underway

WP4: Rules definition commenced

WP5: Sensor analysis completed. Mathematical model developed

WP6: Not yet commenced