



TECHNICAL REPORT

**User Centric approach in Digital Ecosystem;
The Smart Interface;
Part 2: Smart Identity: A Proof of Concept**

ReferenceDTR/USER-0053

Keywordscybersecurity, interface, QoS, user

ETSI

650 Route des Lucioles
F-06921 Sophia Antipolis Cedex - FRANCE

Tel.: +33 4 92 94 42 00 Fax: +33 4 93 65 47 16

Siret N° 348 623 562 00017 - APE 7112B
Association à but non lucratif enregistrée à la
Sous-Préfecture de Grasse (06) N° w061004871

Important notice

The present document can be downloaded from:

<http://www.etsi.org/standards-search>

The present document may be made available in electronic versions and/or in print. The content of any electronic and/or print versions of the present document shall not be modified without the prior written authorization of ETSI. In case of any existing or perceived difference in contents between such versions and/or in print, the prevailing version of an ETSI deliverable is the one made publicly available in PDF format at www.etsi.org/deliver.

Users of the present document should be aware that the document may be subject to revision or change of status.

Information on the current status of this and other ETSI documents is available at

<https://portal.etsi.org/TB/ETSIDeliverableStatus.aspx>

If you find errors in the present document, please send your comment to one of the following services:

<https://portal.etsi.org/People/CommitteeSupportStaff.aspx>

If you find a security vulnerability in the present document, please report it through our
Coordinated Vulnerability Disclosure Program:

<https://www.etsi.org/standards/coordinated-vulnerability-disclosure>

Notice of disclaimer & limitation of liability

The information provided in the present deliverable is directed solely to professionals who have the appropriate degree of experience to understand and interpret its content in accordance with generally accepted engineering or other professional standard and applicable regulations.

No recommendation as to products and services or vendors is made or should be implied.

No representation or warranty is made that this deliverable is technically accurate or sufficient or conforms to any law and/or governmental rule and/or regulation and further, no representation or warranty is made of merchantability or fitness for any particular purpose or against infringement of intellectual property rights.

In no event shall ETSI be held liable for loss of profits or any other incidental or consequential damages.

Any software contained in this deliverable is provided "AS IS" with no warranties, express or implied, including but not limited to, the warranties of merchantability, fitness for a particular purpose and non-infringement of intellectual property rights and ETSI shall not be held liable in any event for any damages whatsoever (including, without limitation, damages for loss of profits, business interruption, loss of information, or any other pecuniary loss) arising out of or related to the use of or inability to use the software.

Copyright Notification

No part may be reproduced or utilized in any form or by any means, electronic or mechanical, including photocopying and microfilm except as authorized by written permission of ETSI.

The content of the PDF version shall not be modified without the written authorization of ETSI.

The copyright and the foregoing restriction extend to reproduction in all media.

© ETSI 2023.
All rights reserved.

Contents

Intellectual Property Rights	4
Foreword.....	4
Modal verbs terminology.....	4
Introduction	4
1 Scope	5
2 References	5
2.1 Normative references	5
2.2 Informative references.....	5
3 Definition of terms, symbols and abbreviations.....	5
3.1 Terms.....	5
3.2 Symbols.....	5
3.3 Abbreviations	5
4 Smart Identity: from definition to PoC design	6
4.1 Identity definition	6
4.2 From Identity to User Profile	7
4.3 Knowledge base for Smart Identity: Potential profile	9
4.4 PoC Design.....	12
4.4.1 Data Collection	12
4.4.2 SmartID: active profile	12
5 Data collection.....	13
5.1 User Profile service: data collection process	13
5.2 Knowledge base tools.....	13
5.3 Knowledge base	13
6 Data processing	14
6.1 Shaping the problem.....	14
6.2 Resolution tools.....	14
7 PoC: use case presentation	14
7.1 User Story.....	14
7.2 Potential Profile data	15
7.3 Active profile.....	15
8 PoC: realization (results).....	16
8.1 Platform description	16
8.1.1 Activity diagram	16
8.1.2 UML diagram of components.....	17
8.2 Models building.....	18
8.2.1 Operating principle of Transformers	18
8.2.2 Model used for Smart-ID	20
8.2.3 Similarity calculation.....	20
8.3 Choice of tools and technologies.....	21
8.4 Results	21
8.4.1 First Model Use Scenario	21
8.4.2 Second Model Use Scenario	22
8.4.2.1 Similarity between sentences	22
8.4.2.2 Resource Prediction	23
Annex A: Code file for Proof of Concept	24
Annex B: Change History	25
History	26

Intellectual Property Rights

Essential patents

IPRs essential or potentially essential to normative deliverables may have been declared to ETSI. The declarations pertaining to these essential IPRs, if any, are publicly available for **ETSI members and non-members**, and can be found in ETSI SR 000 314: *"Intellectual Property Rights (IPRs); Essential, or potentially Essential, IPRs notified to ETSI in respect of ETSI standards"*, which is available from the ETSI Secretariat. Latest updates are available on the ETSI Web server (<https://ipr.etsi.org/>).

Pursuant to the ETSI Directives including the ETSI IPR Policy, no investigation regarding the essentiality of IPRs, including IPR searches, has been carried out by ETSI. No guarantee can be given as to the existence of other IPRs not referenced in ETSI SR 000 314 (or the updates on the ETSI Web server) which are, or may be, or may become, essential to the present document.

Trademarks

The present document may include trademarks and/or tradenames which are asserted and/or registered by their owners. ETSI claims no ownership of these except for any which are indicated as being the property of ETSI, and conveys no right to use or reproduce any trademark and/or tradename. Mention of those trademarks in the present document does not constitute an endorsement by ETSI of products, services or organizations associated with those trademarks.

DECT™, **PLUGTESTS™**, **UMTS™** and the ETSI logo are trademarks of ETSI registered for the benefit of its Members. **3GPP™** and **LTE™** are trademarks of ETSI registered for the benefit of its Members and of the 3GPP Organizational Partners. **oneM2M™** logo is a trademark of ETSI registered for the benefit of its Members and of the oneM2M Partners. **GSM®** and the GSM logo are trademarks registered and owned by the GSM Association.

Foreword

This Technical Report (TR) has been produced by ETSI Special Committee User Group (USER).

The present document is part 2 of a multi-part deliverable. Full details of the entire series can be found in part 1 [i.1].

Modal verbs terminology

In the present document "**should**", "**should not**", "**may**", "**need not**", "**will**", "**will not**", "**can**" and "**cannot**" are to be interpreted as described in clause 3.2 of the [ETSI Drafting Rules](#) (Verbal forms for the expression of provisions).

"**must**" and "**must not**" are **NOT** allowed in ETSI deliverables except when used in direct citation.

Introduction

The present document is associated with a demo of the Smart Identity Proof of Concept (see annex A).

1 Scope

The present document demonstrates the feasibility of the Smart Identity as it is defined in ETSI TR 103 875-1 [i.1].

It defines, for a specific use case (e-health), the Smart Identity (ID) and provides an associated Proof of Concept (PoC).

2 References

2.1 Normative references

Normative references are not applicable in the present document.

2.2 Informative references

References are either specific (identified by date of publication and/or edition number or version number) or non-specific. For specific references, only the cited version applies. For non-specific references, the latest version of the referenced document (including any amendments) applies.

NOTE: While any hyperlinks included in this clause were valid at the time of publication ETSI cannot guarantee their long term validity.

The following referenced documents are not necessary for the application of the present document but they assist the user with regard to a particular subject area.

- [i.1] ETSI TR 103 875-1: "User Centric Approach in Digital Ecosystem; The Smart Interface; Part 1: Smart Identity: user digital clone".
- [i.2] [Workbook N°1](#): "Digital identities". Personal Information Values and Policies Chair. Mines Telecom Institut.
- [i.3] IEEE 802.11™: "IEEE Standard for Information Technology -- Telecommunications and Information Exchange between Systems -- Local and Metropolitan Area Networks -- Specific Requirements -- Part 11: Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specifications".

3 Definition of terms, symbols and abbreviations

3.1 Terms

For the purposes of the present document, the following terms apply:

deep learning: type of artificial intelligence where the machine is able to learn by itself (unlike execute rules predetermined)

3.2 Symbols

Void.

3.3 Abbreviations

For the purposes of the present document, the following abbreviations apply:

AI	Artificial Intelligence
API	Application Programming Interface

DB	Database
DNA	DeoxyriboNucleic Acid
IBAN	International Bank Account Number
ID	Identity
ML	Model Language
PoC	Proof of Concept
UML	Unified Modeling Language
XNLI	Cross-lingual Natural Language Inference

4 Smart Identity: from definition to PoC design

4.1 Identity definition

The notion of identity is complex and polysemous. A definition is available in [i.2].

It depends on the point of view to consider what an individual can be, and on the use that the Smart ID want to make of this identity.

It is possible to identify 3 domains of identification which are cumulative:

1) Civil and legal identity

This identity traditionally and mainly consists of the following elements:

- surname, first name, gender, nationality, filiation (relationships), date and place of birth.

This identity is deemed to be stable throughout life.

2) Biological identity: Height, weight, eye color, retina, fingerprint, DNA

Some attributes of this identity can be used to formally identify a person, in addition to the civil identity.

3) Social and personal identity

It is made up of many sociological and psychological elements: place of residence, profession, standard of living, hobbies, tastes, friends, beliefs, commitments, etc.

It is built, it evolves and is enriched during life, it is never fixed.

These are identities that can be described as objective or suffered (civil, legal, biological). But there are also subjective and desired identities, corresponding to the way an individual decides for themselves how they intend to present themselves to others. It is a kind of narrative identity. Social networks and the use of pseudonyms and avatars in cyberspace are a tangible manifestation of this. It can be noted in this regard that in the digital world it is possible to have several identities.

The Smart ID is created with:

- The attributes of the objective identity of the person.
- The available resources (equipment, services).

It considers the sequence of the user roles in space-time: objectives, activities, tasks, schedule.

It takes care of the information used to make awareness choices.

More simply, the Smart ID is thus the representation of a person:

- What the user is
- What the user has
- What the user is doing

- What is the user's knowledge

4.2 From Identity to User Profile

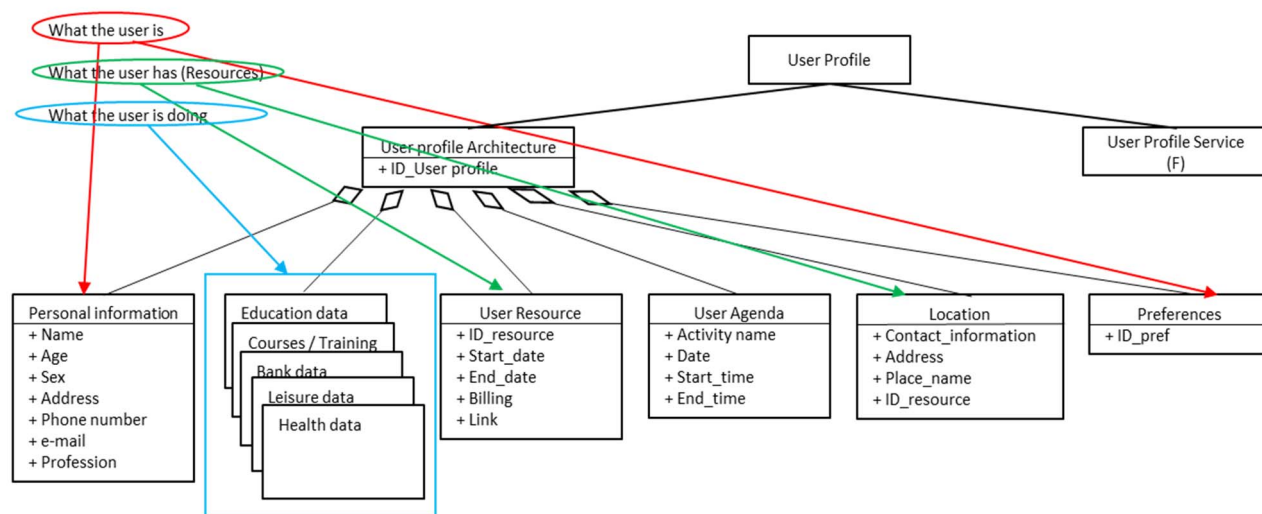


Figure 1: Relationship between Identity definition and user profile model

The "user profile" is therefore the informational representation of the user's identity (figure 1) in the digital ecosystem, including:

- Personal information:
 - Legal identity
 - Identifications
 - Roles (worker, patient, parent, benevolent, etc.)
- User centric characteristics that impact configurations:

Actions according to:

 - Preferences
 - Space-time (agenda)
 - Location
- The resource description according to the location:
 - Internal resources (equipment, network, services)
 - External resources (equipment, network, services)

More precisely the personal sheet/template (figure 2) identity can be as following:

PERSONAL INFORMATION
Social ID
Social ID
Last Name
First name(s)
Gender
Date of birth
Country of birth
City of birth
Country of birth
Nationality (s)
Profession
Personal contact information
Address
Postcode
Town or city
Country
Mobile phone (s)/landline (s)
E-mail address (es)
Professional contact information
Business address
Postcode
Town or city
Country
Mobile phone(s)/landline(s)
E-mail address(es)
Other
Residence permit
Visa
Passport
Identity Card
Licences (e.g. Driving licence)

Figure 2: Personal information template

Moreover, with the evolution of paradigms, the **relationship** between the user and the system is now an **N to N** relationship meaning that the user has N profiles according to the role (figure 3).

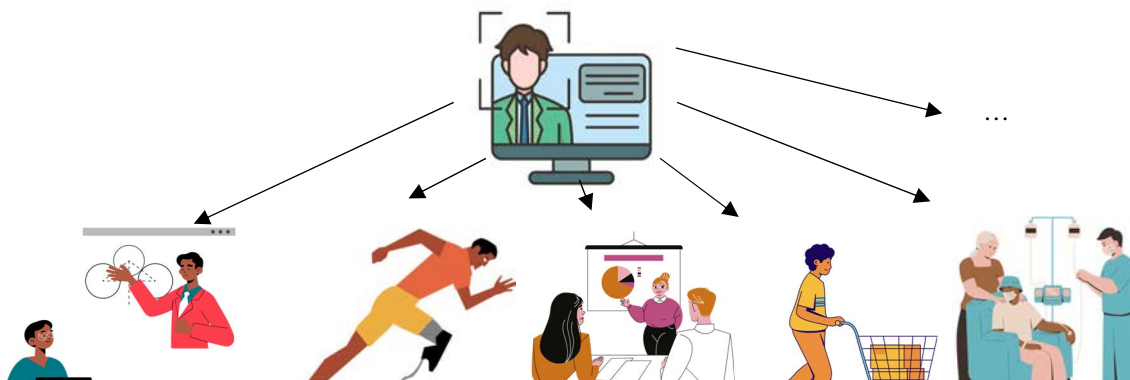


Figure 3: N user potentials roles

The instantiation of the User Profile model will give the complete picture of the user according to their roles in space-time, preferences, and location.

4.3 Knowledge base for Smart Identity: Potential profile

The user identity also takes care of "what the user knows".

This is why the knowledge base also includes the characteristics of the different fields of activity known to the user. This is why, it will be said that he acts in all awareness.

What user knows: potential profile (information allowing choice to be made awareness).

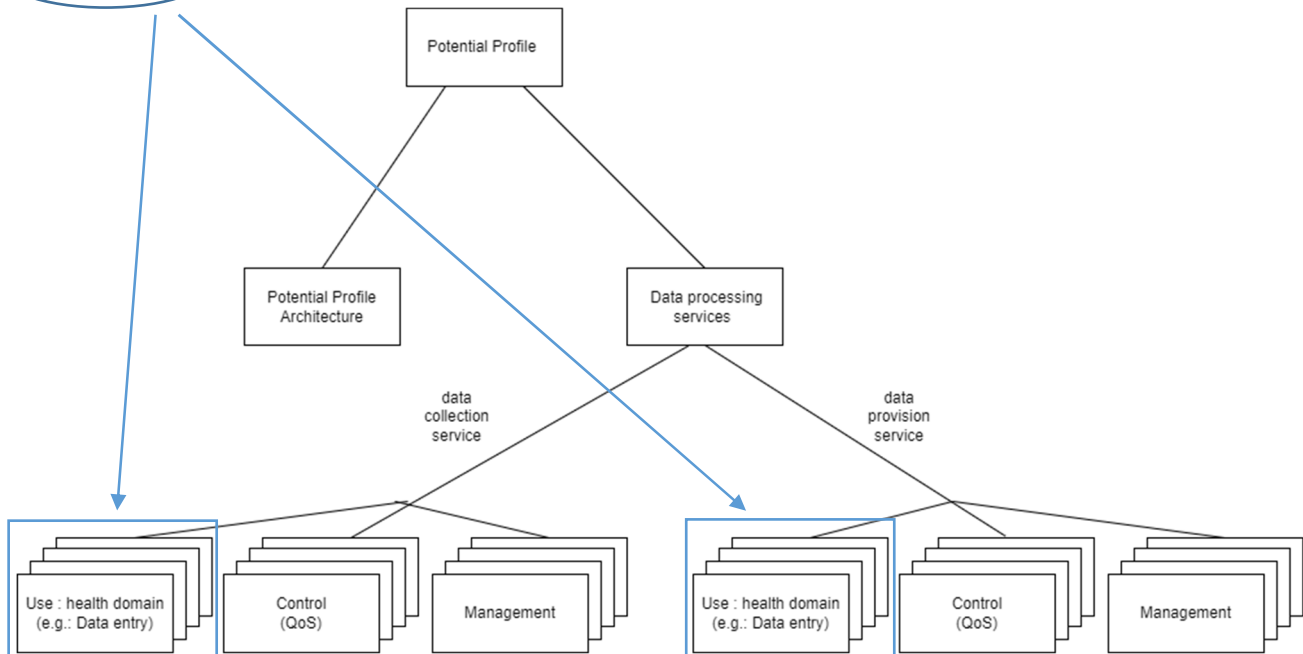


Figure 4: Knowledge base

Like any entity, the Potential Profile has an architecture and a service interface.

The first service consists of data collection (instantiation of the Potential Profile model) to store architecture information. The architecture of the Potential Profile enriches the architecture of the User Profile by informing all the areas relating to a temporal action.

The second service will be all the processing on data analysis in order to have additional decision-making information.

Each role played by the user requires data from each of the domains. That is, for each action the Smart ID needs cross-sectional data (figure 4). For example, when the user has a medical appointment, the profile will need the health domain and the finance domain, as with all actions that require payment.

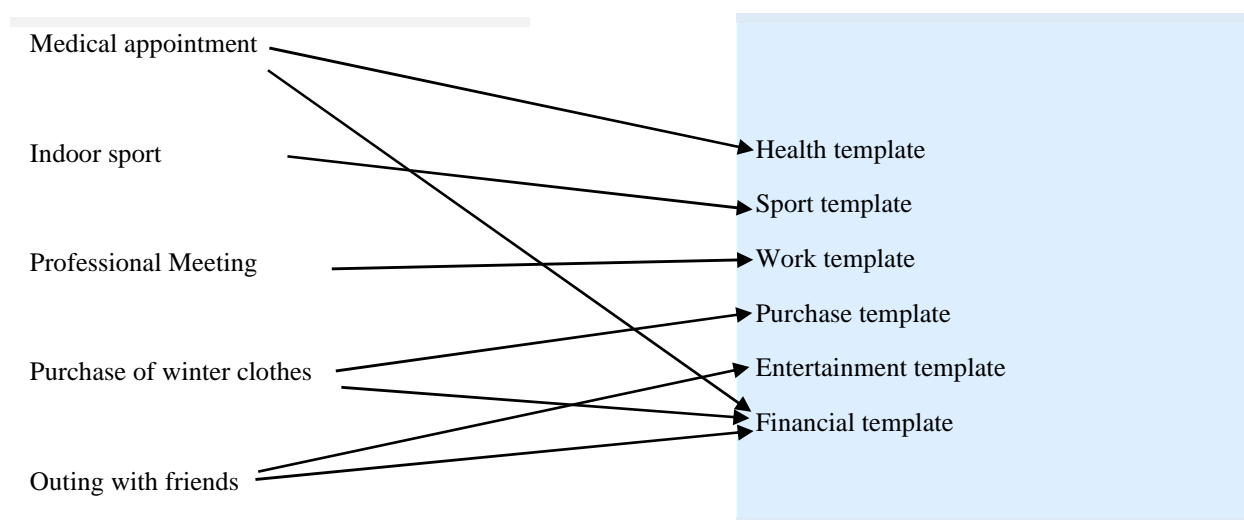


Figure 5: Example of transversal data

For the PoC it will need templates corresponding to the most common roles, that is to say the basic roles (figure 5) involved in areas such as health, work, finance, travel, etc.

HEALTH TEMPLATE
Social security number
Mutual insurance (personal, professional, etc.)
Health information
General information
Weight
Height
Health parameters
Temperature
Pulse/heart rate
Oxygen saturation
Respiratory rate
Blood pressure
Medical record
Blood type
Rhesus
Vaccination booklet
Vaccination certificates (covid-19)
Current treatment(s)
Allergies
Last tests
Latest prescriptions
Medical appointments (past and future)
Doctor notes
Emergency contacts

TEMPLATE SPORT (well-being)
Membership card
Information
User behaviours
Time spent sleeping/in bed
Inactivity time
Number of steps
Distance (walking and running)
Stages climbed
Stability of walking
Other
Calories
Information sheet
Type of subscription (classic, premium)
Duration of subscription
Wellness goal(s)

TEMPLATE TRAVEL
<p>Information</p> <p>Tickets/tickets for transport</p> <p>Ticketing</p> <p>Flight/train/bus tickets</p> <p>Advantage/discount card</p> <p>Carpooling information</p> <p>Name and contact of the driver</p> <p>Departure time</p> <p>Meeting point (departure)</p> <p>Arrival point</p> <p>Stay</p> <p>Hotel/accommodation reservation</p>

TEMPLATE WORK
<p>Registration number</p> <p>Access badge</p> <p>Tickets/restaurant card</p> <p>Professional file</p> <p>Job function</p> <p>Office</p> <p>Department/service</p> <p>Other</p> <p>Employment contract/internship agreement</p> <p>Pay slips</p>

TEMPLATE EDUCATION
<p>Personnel number</p> <p>Student number</p> <p>Education card</p> <p>Student information</p> <p>Education</p> <p>Timetable</p> <p>Teaching units</p> <p>Documents</p> <p>Graduations</p>

TEMPLATE SHOPPING/PURCHASES
<p>Information</p> <p>Item preferences</p> <p>Product size</p> <p>Customer record by supplier</p> <p>Customer number</p> <p>Membership/loyalty card</p> <p>Username</p> <p>Gift card</p> <p>Discount coupons</p> <p>Ordering</p> <p>Points</p> <p>My addresses (delivery and billing)</p> <p>My payment methods (credit card, multiple payments, etc.)</p>

TEMPLATE DOMOTIC
<p>Access card/key</p> <p>Charges</p> <p>Energy supplier</p> <p>Hot and cold water</p> <p>Heating</p> <p>Maintenance costs</p> <p>Green space</p> <p>Common areas</p> <p>Household waste collection tax</p> <p>Connected house</p> <p>Management of: security, access, lighting, temperature, fans, objects, etc.</p> <p>Other</p> <p>Lease agreement/certificate of residence</p>

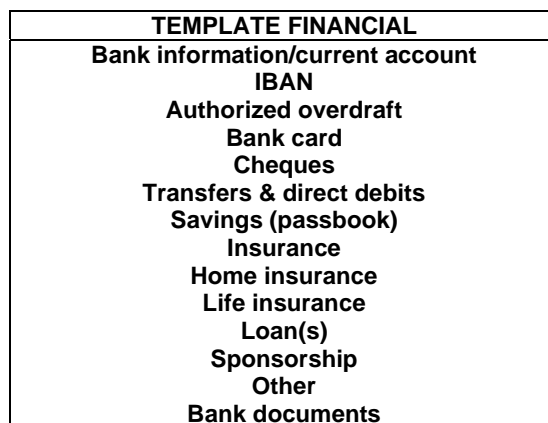


Figure 6: Common knowledge templates

4.4 PoC Design

4.4.1 Data Collection

Data collection consists of filling in the Potential Profile, which includes User Profile data, data on the various user domain interest, and data from the services requested (e.g. weather, etc.), as well as data processing contributions.

4.4.2 SmartID: active profile

The intelligent user profiling process, after the collection of the user data, begins with the consideration of their user centric parameters (space-time, space-place and preferences), which constitute the system input data allowing the obtaining the active profile.

Then an identification of the activity and a classification in relation to the different domain will be made using artificial intelligence algorithms during the data processing phase. There will then follow a selection of templates that may correspond to the areas of user action during this time slot. For example, a medical appointment might require the health, finance, and travel templates.

The preferences will make it possible to finalize the personalization of the active profile (figure 7) and to select the user's preferred resources (equipment, networks, services) according to user location.

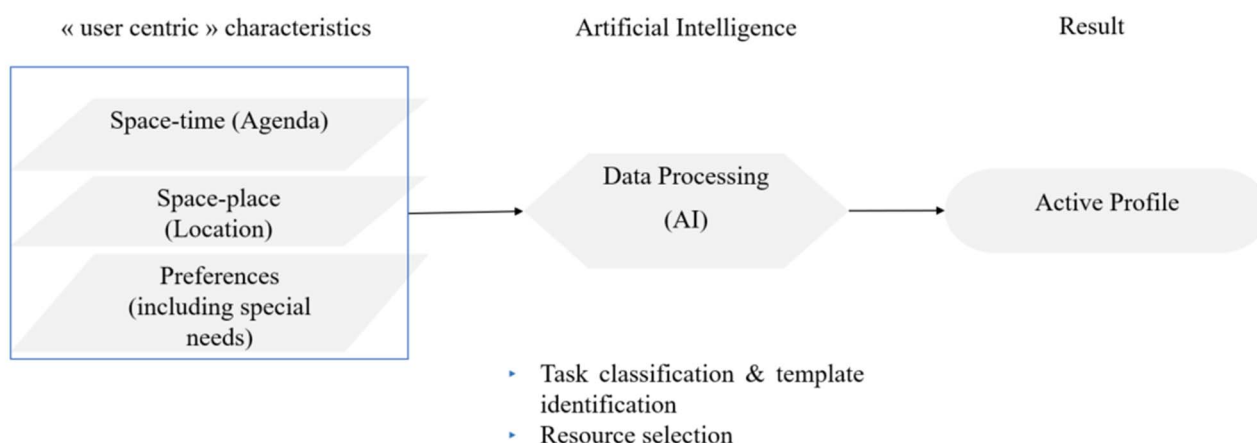


Figure 7: Smart user profiling process and active profile procurement

5 Data collection

5.1 User Profile service: data collection process

The dataset included a set of information allowing AI (deep learning) models to understand the challenge of Smart ID and to become familiar with user-centric characteristics.

Given that the objective is to have independent data of application services, a dataset using a relational database could not meet these linkage requirements.

The approach was to create a dataset written in Python[®] and containing key information, including the task, internal resources (equipment, networks, and services) by location, external resources and templates.

5.2 Knowledge base tools

The tools used to create the dataset to build and train Artificial Intelligence models are as follows:

Jupyter[®] Notebooks implementation is used, allowing to write programs in different languages.

"Python[®]" is a Programming language used for the dataset.

"Pandas[™]" is the Python[®] library to create datasets as dataframes.

5.3 Knowledge base

Figure 8 shows the draft of our dataset used to build and train the built deep learning models.

	task	location	internal equipment	internal networks	internal services	external equipment	external networks	external services	information	templates
0	medical appointment booking	office	smartphone or computer	802.11		equipment participating in the session	internet	web, health platform services	health insurance	health
1	book medical appointment	house	smartphone or computer	802.11		equipment participating in the session	internet	web, health platform services	health insurance	health
2	medical appointment booking	outside	smartphone	3GPP		equipment participating in the session	internet	web, health platform services	health insurance	health
3	Go to doctor	outside	smartphone	3GPP	set location	equipment participating in the session	internet	weather, get location	travel ticket	displacement
4	Consultation by doctor	at destination	smartphone						health insurance, mutuals	health
5	Teleconsultation	house	smartphone or computer	802.11	Consultation reminder notification	equipment participating in the session	internet	health platform services	health insurance, mutuals	health
6	Consultation fee payment	house	smartphone	802.11	notification for payment validation	equipment participating in the session	internet	banking services	credit card	finance
7	Consultation fee payment	at destination	smartphone					banking services	credit card, health insurance, mutual funds	finance

Figure 8: Draft dataset used for Smart ID

6 Data processing

6.1 Shaping the problem

Before using Artificial Intelligence algorithms, it is necessary to transform the input textual data (user-centric characteristics) into number vectors of the same size (embeddings) so that they can be understood by neural networks.

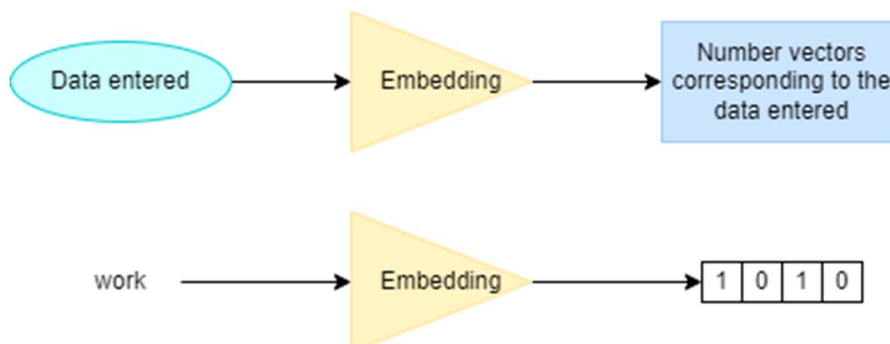


Figure 9: Embedding generation

The steps generally performed for this processing of textual data are as follows:

- 1) "Lowercase", to put all the characters of the sentences in lower case.
- 2) Remove punctuation, numbers, and special characters.
- 3) Remove phrases in English (or French, depending on the problem to be solved).
- 4) "Tokenization" to reduce the size of vocabulary needed to encode a message.

6.2 Resolution tools

For the Smart ID, the **Transformers** model and its **AutoTokenizer** library were used to pre-process the data presented in clause 6.1.

```
from transformers import AutoTokenizer,
```

Figure 10: Line of code used to pre-process data

7 PoC: use case presentation

7.1 User Story

As an example, the use case for the PoC is to consider the Provisional Agenda and Location of September 19, 2022, as presented in figure 11.

	SCHEDULE	LOCATION
15h - 17h	Team Meeting	Office
17h15 - 18h	(Travel)	
18h - 18h30	Preparation for the gym	House
18h45 - 20h	Sports session	Nearest gym
20h15 - 21h	Dinner	House
21h - 22h	Purchase of autumn / winter clothes	Online
22h	Rest	

Figure 11: Day Program of September 19, 2022

7.2 Potential Profile data

In the following example, it is 6 p.m. and the user has just returned home. Users would like to go for a workout and need to know if the weather allows them to go out, but also to find a gym close to their home.

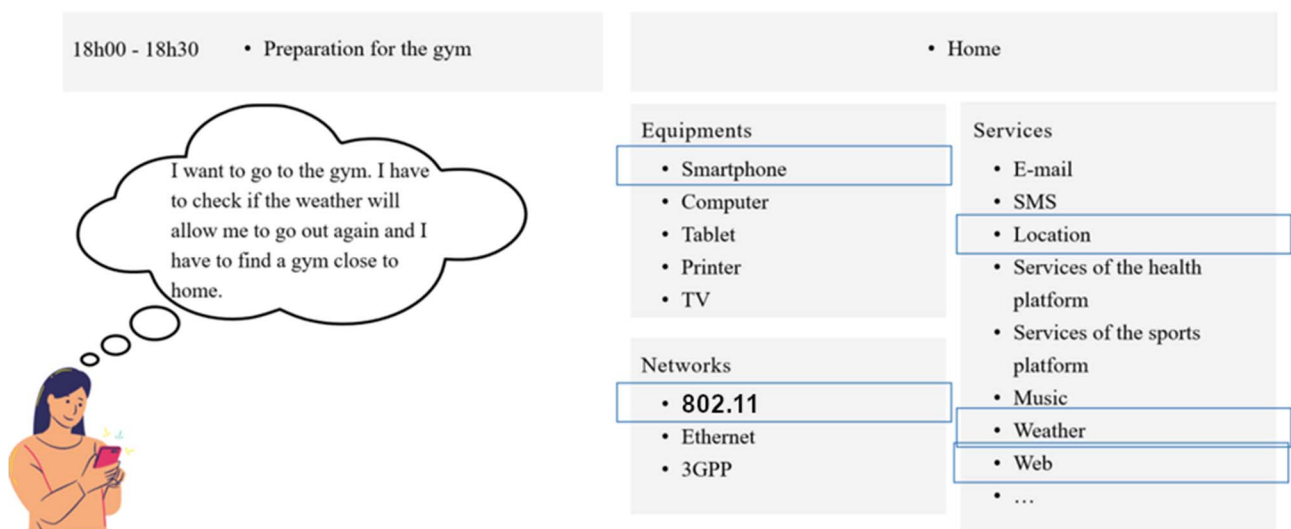


Figure 12: Available Resources for the specific use case

The resources available in the user's environment are shown in figure 12.

7.3 Active profile

To perform the tasks presented in clause 7.2, the user needs a terminal, network access and services.

Since the users are at home, they will choose:

- 1) the Smartphone as a terminal for mobility concerns;
- 2) Wi-Fi® (IEEE 802.11™ [i.3]), to save mobile data and to respect the environment;
- 3) weather, location, and web services.

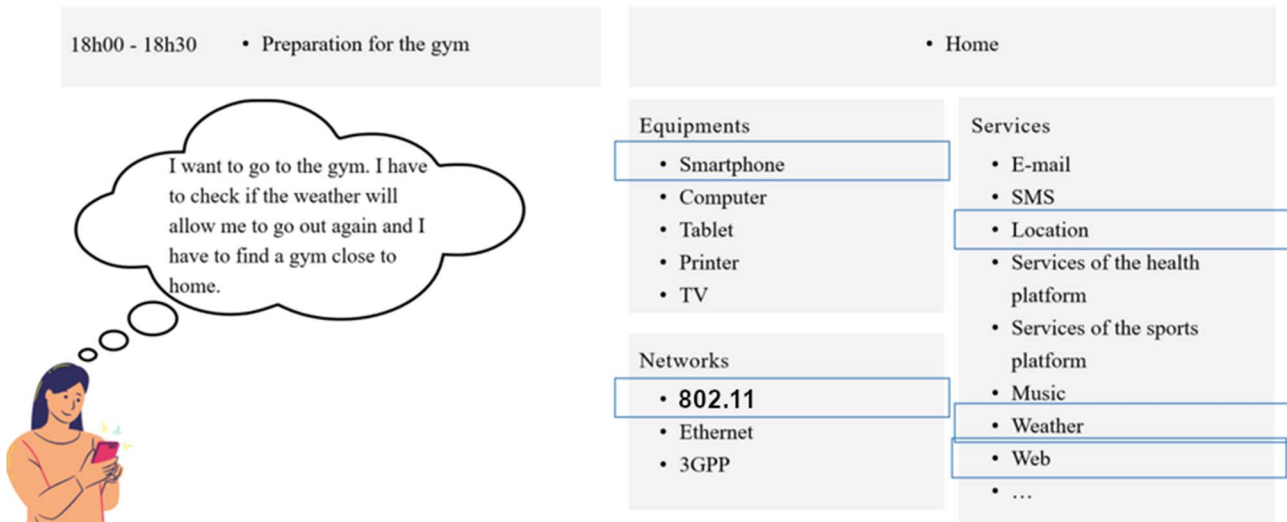


Figure 13: Active user profile from 6 p.m. to 6:30 p.m.

The active profile of the user from 6 p.m. to 6:30 p.m. at home is shown in figure 13.

8 PoC: realization (results)

8.1 Platform description

8.1.1 Activity diagram

The activity diagram is the representation of the behaviour of the system in the form of flows or sequences of activities.

Figure 14 shows the operations performed to process user data and build AI models relevant to Smart ID.

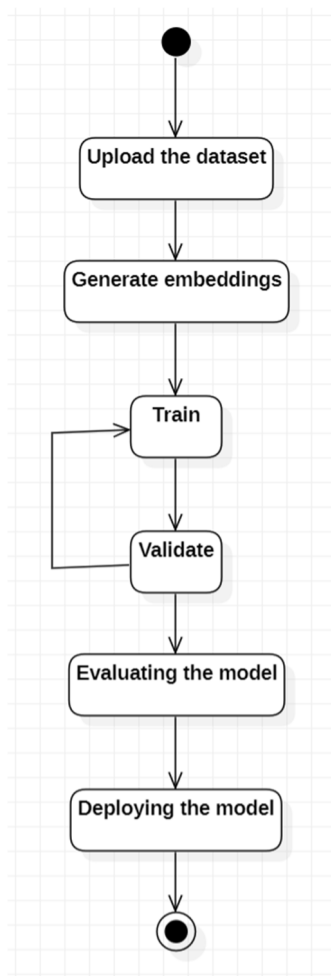


Figure 14: Choosing the right AI model for Smart ID

8.1.2 UML diagram of components

Figure 15 shows the UML diagram of the components of the Smart ID system. The data entered as text correspond to the tasks in the user's agenda. These are classified by the text classification component that will output the percentages of the different domains to which the task can belong to facilitate the choice of "templates".

Next, the semantic similarity component will browse the dataset to find an existing task that can be like the one the user wants to perform. The output of the semantic similarity component is therefore the set of textual data that are like the input data.

Finally, the prediction component will determine the resources required by the user to perform this task.

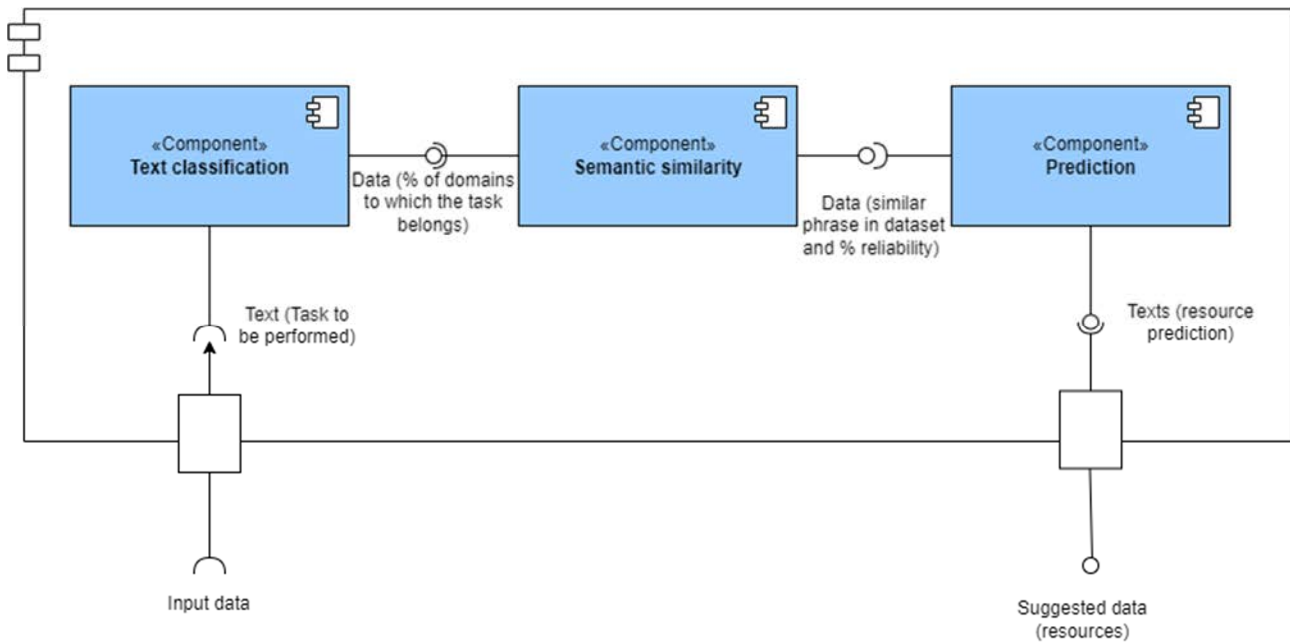


Figure 15: UML diagram of system components

8.2 Models building

8.2.1 Operating principle of Transformers

The Transformer is a deep learning model that uses attention by differentially weighting the importance of each part of the input data to increase its training speed.

The Transformer consists of two main components, namely: the encoder and the decoder and the connections between them (figure 16).

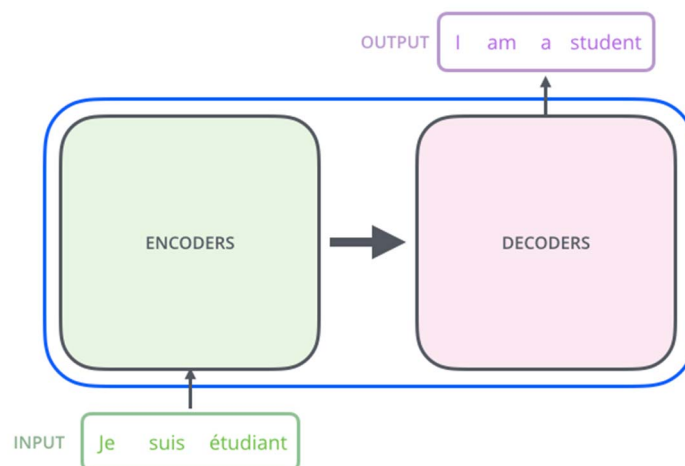


Figure 16: Components of a transformer

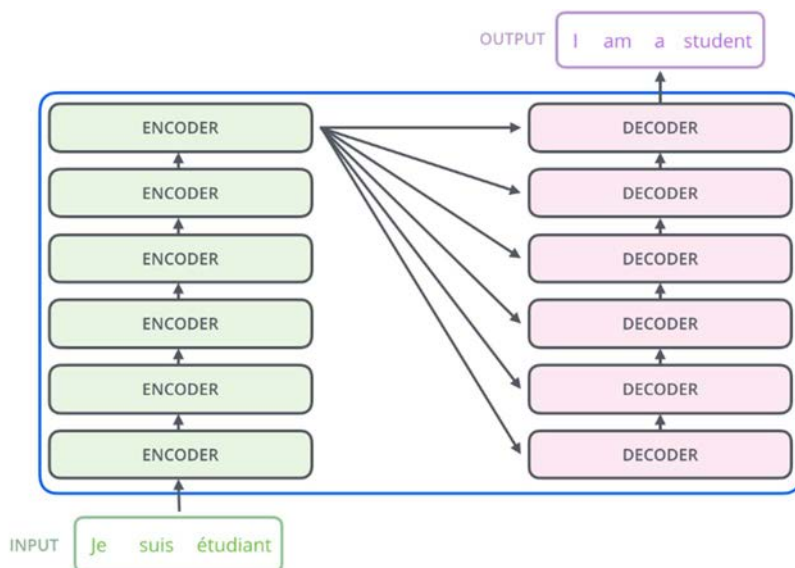


Figure 17: Encoders - Decoders

The **encoding** component is generally composed of six (6) encoders stacked on top of each other (figure 17).

The **decoding** component is composed of decoders stacked on top of each other of the same number as the encoding component.

The overall architecture of the transformer is shown in figure 18.

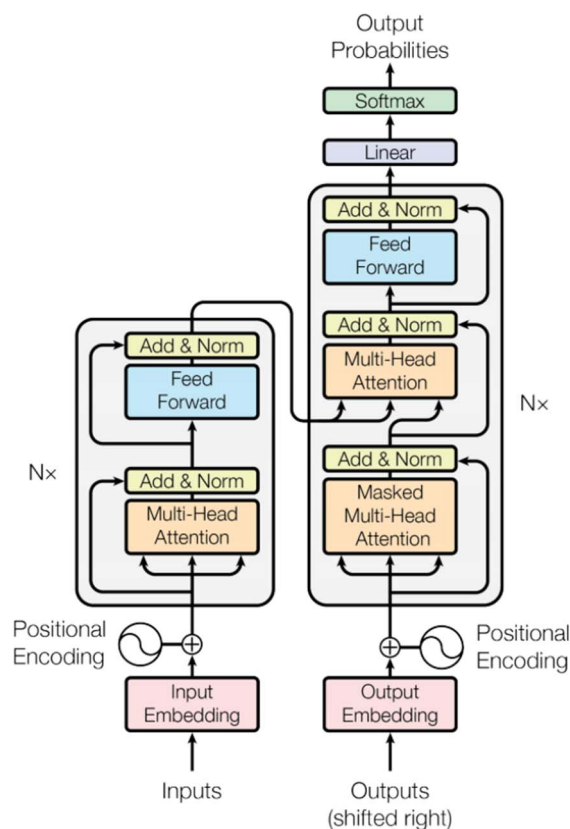


Figure 18: Global transformer architecture

In accordance with the overall architecture of the transformer (figure 18), there are three (3) attention mechanisms:

- 1) Self-Attention in the encoder
It is a layer that helps the encoder look at other words in the input phrase when encoding a specific word.

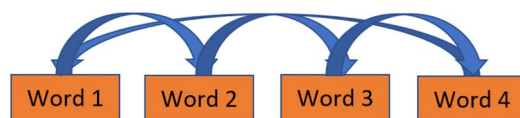


Figure 19: Self-Attention Mechanism

- 2) Self-attention with all previously generated elements in the input of the decoder (at the level of "Encoder-decoder Attention") to help the decoder focus on the relevant parts of the input sentence.

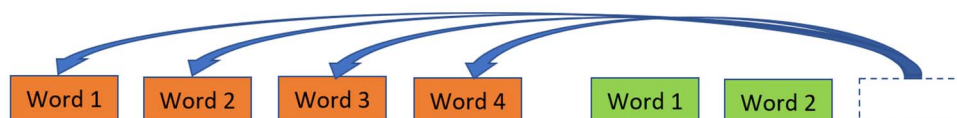


Figure 20: Mechanism of "Encoder-Decoder Attention"

- 3) "Masked" Attention between the element generated by the decoder and all the elements of the encoder. This mechanism can be used for text generation or simultaneous translation and is autoregressive.



Figure 21: "Masked Attention" mechanism

8.2.2 Model used for Smart-ID

For the creation of AI models for Smart ID, a pre-trained neural network model based on Transformers was used. It is called **CamemBERT™**.

The Camembert™-Base-XNLI zero-stroke pre-trained transfer learning algorithm was used because classical machine learning algorithms did not give accurate results during training on the dataset.

Camembert™-base-XNLI is a transformer-based natural language processing model written in Python®. It was trained on XNLI (Multilingual Natural Language Inference) which was published by Facebook™. It is mainly used to determine the probability of a corpus of text belonging to a predefined class.

8.2.3 Similarity calculation

To obtain the similarity score between two (2) sentences, the system takes as input:

- 1) the reference phrase from the dataset;
- 2) the candidate phrase to be challenger (entered by the user).

Embeddings (number vector) are generated from this textual data and these will be used to calculate the cosine similarity (cosine similarity per word pair + linear regression) between sentences.

8.3 Choice of tools and technologies

To implement the Camembert™-Base-XNLI algorithm for data entry and resource prediction, the tools used are listed in table 1.

Table 1

Tools/Technologies	Version	Role
Python®	3.7	The most widely used programming language for data analysis, ML and AI in general. It is platform independent.
Transformers	4.24.0	Library for downloading and training pre-trained natural language processing models.
Tensorflow®-Text	2.9.0	It is a library of TensorFlow® to perform operations on texts for pre-processing.
Pandas™	1.3.5	For managing datasets using dataframes
Google Colab®	1.0.0	A cloud service offered by Google®, based on Jupyter® Notebook and allowing to train ML models directly online, without the need to install anything.

For a better visualization of the results of the main model, web interfaces have been developed with the Gradio API version 3.12.1. Gradio allows for deploying locally to test models.

8.4 Results

8.4.1 First Model Use Scenario

Before running the model corresponding to the first component of the Smart ID, the **pipeline** library was imported to link to the camemBERT™ pre-trained model. Then a classifier based on this model was defined.

```

✓ [6] from transformers import pipeline
      classifier = pipeline("zero-shot-classification", model="BaptisteDoyen/camembert-base-xnli")

Downloading: 100% ██████████ 882/882 [00:00<00:00, 28.7kB/s]
Downloading: 100% ██████████ 443M/443M [00:08<00:00, 48.0MB/s]
Downloading: 100% ██████████ 433/433 [00:00<00:00, 14.7kB/s]
Downloading: 100% ██████████ 811k/811k [00:00<00:00, 12.9MB/s]
Downloading: 100% ██████████ 299/299 [00:00<00:00, 6.30kB/s]

```

Figure 22: Task classifier definition

To test Model 1, the Gradio interface developed for this purpose is run. Then simply fill in the task to be executed as well as the templates that constitute here the candidate classes.

The behaviour of Model 1 is shown in figure 23.

Task

Purchase of drugs

Templates (classes) candidats

"Health";"Finance";"Sport";"Travel"

Clean Submit

Classification

Report

Figure 23: Entering the task to classify in relation to the different templates

Figure 24 shows the result after clicking "submit". Model 1 is therefore able to classify a given task in relation to domains, giving the probabilities that it belongs to one class rather than another.

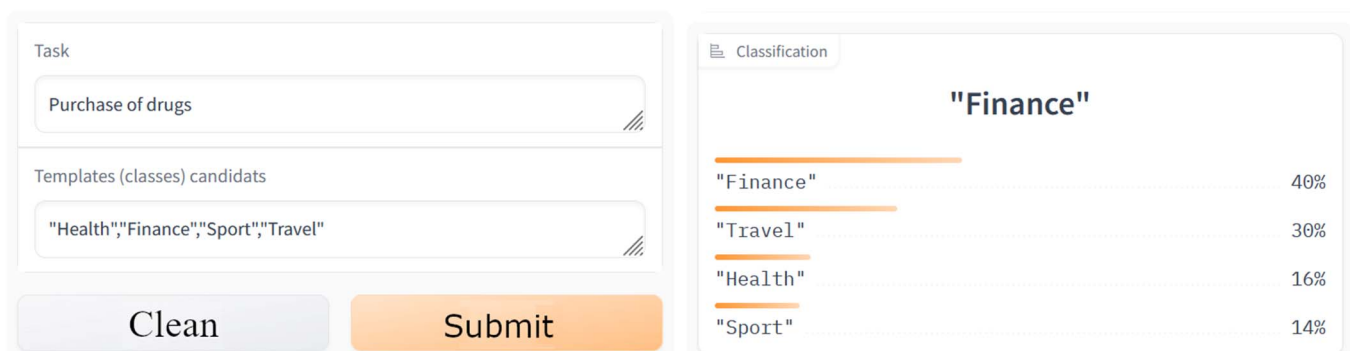


Figure 24: Result of Model 1

8.4.2 Second Model Use Scenario

8.4.2.1 Similarity between sentences

For model 2, it is first necessary do the embedding (see figure 25) with the AutoTokenizer library of the Transformers library so that the model can understand and process the candidate task that will be entered.

```

✓ [9] #Importation de la bibliothèque permettant de faire le pré-traitement des données
0s
from transformers import AutoTokenizer, AutoModelForSequenceClassification

✓ [10] # Génération des embeddings et importation du modèle pré-entraîné
4s
nli_model = AutoModelForSequenceClassification.from_pretrained("BaptisteDoyen/camembert-base-xnli")
tokenizer = AutoTokenizer.from_pretrained("BaptisteDoyen/camembert-base-xnli")

```

Figure 25: Embeddings

```
det_tache("Book a slot with the doctor", df)
(52.55739688873291, 'medical appointment booking')
```

Figure 26: Result

This component gives the result of the similarity calculation between the candidate sentence (task to be performed) and the reference phrase (task that is in the dataset), as well as the sentence that it considers to be reference.

Figure 26 shows the result of the template for the "book a slot with doctor" task.

Score: 52,56 %

Reference phrase: Medical appointment booking.

8.4.2.2 Resource Prediction

The resource prediction model considers the task at hand, the calculation of similarity between sentences, as well as the location to propose internal and external resources (equipment, networks, services) that may be useful.

Here is the behaviour of Model 2.

```
det_ressources1("Composition of the training program", "house")
['smartphone',
 '802.11',
 'set location',
 'equipment participating in the session',
 'internet',
 'web, get location, weather']
```

The screenshot shows a web interface with two input fields: 'entry' containing 'medical appointment booking' and 'loc' containing 'outside'. Below these are 'Clean' and 'Submit' buttons. To the right, an 'output' field displays the result: ['smartphone', '3GPP', '', 'equipment participating in the session', 'internet', 'web, health platform services'].

Figure 27: Proposal of internal and external resources for the entered task

Annex A: Code file for Proof of Concept

The Smart Identity Proof of Concept is run in the Google® Colaboratory notebook (<https://colab.research.google.com/>).

The code file for the present document is available on the ETSI Forge repository, at the following link:

- <https://forge.etsi.org/rep/user/tr-103-875-2/smartid-poc/-/tree/v1.1.1>

The *POC_SmartID_v3.ipynb* file is to be uploaded to Google® Colaboratory and the PoC is executed from there.

Annex B: Change History

Date	Version	Information about changes
01-2023	V0.0.5	Final draft for Approval
03-2023	V1.1.1	First published version

History

Document history		
V1.1.1	March 2023	Publication