



TECHNICAL REPORT

**User Centric Approach in Digital Ecosystem;  
The Smart Interface;  
Part 1: Smart Identity: user digital clone**

---

**Reference**

DTR/USER-0052

---

**Keywords**digital clone, smart identity, smart interface,  
user-centric**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](http://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 2022.  
All rights reserved.

# Contents

Intellectual Property Rights .....	5
Foreword.....	5
Modal verbs terminology.....	5
Introduction .....	5
1 Scope .....	6
2 References .....	6
2.1 Normative references .....	6
2.2 Informative references.....	6
3 Definition of terms, symbols and abbreviations.....	7
3.1 Terms.....	7
3.2 Symbols.....	8
3.3 Abbreviations .....	8
4 Analysis of the user centric context in digital ecosystem.....	9
4.1 Overview .....	9
4.2 Sociological and psychological context .....	9
4.3 Non-functional requirements .....	10
4.3.1 Overview .....	10
4.3.2 Security .....	10
4.3.3 Privacy .....	11
4.3.4 Usability.....	11
4.4 Digital maturity .....	12
4.5 Usage Evolution and digital transformation of society .....	13
4.6 User and Provider interaction for service delivery .....	15
4.6.1 Overview .....	15
4.6.2 Life cycle of digital services in provider side and user journey.....	16
4.6.3 User-centric service (Health use case).....	18
5 From User Digital Clone to Smart Identity .....	21
5.1 The Modelling .....	21
5.2 User Digital Clone representation .....	23
5.3 Profiles: user knowledge base .....	24
5.3.1 Overview .....	24
5.3.2 User Profile.....	25
5.3.2.1 User Profile definition.....	25
5.3.2.2 User Profile Model.....	25
5.3.3 Potential Profile .....	27
5.3.3.1 Potential Profile definition .....	27
5.3.3.2 Potential Profile Model .....	27
5.3.4 Active Profile .....	28
5.3.4.1 Active Profile definition.....	28
5.3.4.2 Active Profile Model.....	28
5.4 Smart ID .....	29
5.4.1 Smart ID Model .....	29
5.4.2 Smart ID: Informational Use Cases .....	29
5.4.2.1 Overview .....	29
5.4.2.2 User profile architecture.....	30
5.4.2.3 Enrichment of the User Profile by the Potential Profile .....	31
5.4.2.4 Active profile for the first time slot.....	33
5.4.2.5 Active Profile for the second time slot.....	36
5.4.2.6 Smart ID for the first time slot .....	37
5.4.2.7 Smart ID for the second time slot.....	37
6 Data processing .....	37
6.1 Data categorization.....	37

6.2	Data processing steps .....	38
7	New technologies for smart identity .....	39
7.1	Overview .....	39
7.2	Artificial Intelligence (AI).....	39
7.3	Data storage and security .....	41
7.4	Data localization (Cloud, Edge, Dew).....	42
8	Conclusion and next steps .....	42
	History .....	43

---

# 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 1 of a multi-part deliverable covering the smart identity, as identified below:

**Part 1:** "Smart Identity: user digital clone";

Part 2: "Smart Identity: A Proof of Concept".

---

# 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 the first part of a set of two Technical Reports describing the Smart Identity (as a User Clone) for the User-Centric Approach and providing a PoC (Proof of Concept) demonstrating the feasibility of this User Clone in different chosen Use cases.

According to this approach, the Smart Identity needs to have a sufficient knowledge for the user interface to anticipate and respond to the user's needs and expectations, with a more in-depth analysis of the digital ecosystem.

The present document firstly presents an analysis of the user centric context in digital ecosystem, secondly the model of smart ID as a user digital clone, and what that means about the constitution of the knowledge base, thirdly the elements relating to data processing in order to achieve a smart identity. The present document also explores the different possibilities of using new technologies, in particular artificial intelligence, to achieve a smart identity.

---

# 1 Scope

The present document contains the result of studies relating to the analysis of the user requirements, the new technologies contribution for smart identity, and a digital clone definition.

The present document examines:

- The user profiles in digital ecosystem including sociological and psychological context, non-functional requirements, digital maturity, usage evolution facing digital transformation of society and new interactions between user/provider for service delivery.
- The definition and Smart ID Model based on the information model described in the ETSI TR 103 604 [i.4].
- The user digital clone and his knowledge base needs (ACIFO Model - Informational model).
- The different profiles for the most exhaustive description possible whatever the use case and associated definitions.
- Data processing for data enrichment.
- New technologies for smart identity (AI, Data storage and security, Data localization, etc.).

---

## 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 103 438: "User Group; User centric approach in Digital Ecosystem".
- [i.2] ETSI EG 203 602: "User Group; User Centric Approach: Guidance for users; Best practices to interact in the Digital Ecosystem".
- [i.3] ETSI TR 103 603: "User Group; User Centric Approach; Guidance for providers and standardization makers".
- [i.4] ETSI TR 103 604: "User Group; User centric approach; Qualification of the interaction with the digital ecosystem".
- [i.5] ETSI TR 103 437: "USER; Quality of ICT services; New QoS approach in a digital ecosystem".
- [i.6] ISO/IEC 27001: "Information and data security".
- [i.7] ISO/IEC 27040: "Storage security".
- [i.8] Byron Reeves and Clifford Nash: "The media equation".

NOTE: Available at <https://web.stanford.edu/group/cslipublications/cslipublications/site/1575860538.shtml>.

- [i.9] Eija Kaasinen, Tiina Kymäläinen, Marketta Niemelä, Thomas Olsson, Minni Kanerva and Veikko Ikonen: "A User-Centric View of Intelligent Environments: User Expectations, User Experience and User Role in Building Intelligent Environments", *Computers*, 2(1), 1-33. 2013.
- [i.10] Zeltser: "Importance of feeling secure".
- NOTE: Available at <https://zeltser.com/importance-of-feeling-secure/>.
- [i.11] Zeltser: "Digital strategy: eprivacy regulation".
- NOTE: <https://digital-strategy.ec.europa.eu/en/policies/eprivacy-regulation>.
- [i.12] Directive (EU) 2018/1972 of the European Parliament and of the Council of 11 December 2018 establishing the European parliament and the council establishing the European Electronic Communications Code.
- [i.13] Telecom Paris values and policies of personal information chair: "Data, identities and trust in the digital age".
- [i.14] ETSI TR 103 875-2: "User Centric approach in Digital Ecosystem; The Smart Interface; Part 2: Smart Identity, A Proof of Concept".
- [i.15] 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:

**ACIFO:** The 5-dimension model ACIFO described in the different publications (ETSI TR 103 438 [i.1], ETSI EG 203 602 [i.2], ETSI TR 103 603 [i.3], ETSI TR 103 604 [i.4]) is based on 5 sub-models:

- **Architectural Model** "Acifo" describes the global structure, including semantics and is optimized for the stated objectives.
- **Communication (Relational) Model** aCifo describes the exchange protocols, including HMIs (User) and APIs (provider) exchange and management protocols over three planes: (1) Management (Monitoring), (2) Control, and (3) Usage.
- **Information Model** acIfO describes the different Profiles (User, device, service). The information covers the whole ecosystem (equipment, network, applications, services, HMIs, User, etc.) from the offer to the resource's availability for Users, Providers, and any other partners.
- **Functional Model** aciFo: describes services and service composition. The functionalities (the process) to compose any service based on "micro-service".
- **Organization Model** acifO: describes the role of any actor and which actor is responsible of each action. ("Who is doing what?").

**cloud:** network of remote servers hosted on the Internet and used to store, manage, and process data in place of local servers or personal computers

**dew:** programming model for enabling ubiquitous, pervasive, and convenient ready-to-go, plug-in facility empowered personal network

**edge:** distributed computing paradigm in which computation is largely or completely performed on distributed device nodes

**fog:** provides close computation, data storage and application services

**profile:** information template (model) to provide the data characterizing and qualifying the user and the resources of the digital ecosystem

## 3.2 Symbols

Void.

## 3.3 Abbreviations

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

ACIFO	Architectural, Communication, Information, Functional and Organizational model
AI	Artificial Intelligence
API	Application Programming Interface
CISO	Chief Information Security Officer
CPU	Central Process Unit
DAP	Digital Adoption Platform
DaaS	Device as a Service
DDoS	Distributed Denial of Service
EC	European Commission
EU	European Union
GDPR	General Data Protection Regulation
GPS	Global Positioning System
GSM	Global System for Mobile communication
GWI	Global Web Index
HMI	Human-Machine Interface
IAM	Identity and Access Management
ID	Identity
IoT	Internet of Things
IT	Information Technology
KNN	K-Nearest Neighbour
LAN	Local Area Network
NGN	New Generation Networks
NPaaS	Network Platform as a Service
PaaS	Platform as a Service
PAN	Personal Area Network
PCA	Principal Component Analysis
PCI-DSS	Payment Card Industry - Data Security Standards
PoC	Proof of Concept
QoS	Quality of Service
SMS	Short Message Service
SOA	Service-Oriented Architecture
VoLTE	Voice over Long-Term Evolution
VoWiFi	Voice over WiFi
VPCN	Virtual Private Connectivity Network
VPSN	Virtual Private Service Network
WAN	Wide Area Networks
WiFi	Wireless Fidelity



---

## 4 Analysis of the user centric context in digital ecosystem

### 4.1 Overview

Smart products and services that adapt to aspects of the users' activity, context or personality are starting to appear in the market. Now, users often expect the emerging of products or services which act intelligently with them more like they would among themselves, as humans. In the first decades of the 21<sup>st</sup> century, technical limitations impeded the creation of smart interfaces that fully live up to those expectations. But now the progress and partial trivialization of artificial intelligence gives new possibilities, and consequently, it is essential to manage user expectations throughout the design process of an intelligent interface.

There are five approaches to mix to completely cope with the whole users' needs (including special needs), requirements and expectations:

- The first is to understand the basic psychology of how people interact with smart products and services, clause 4.2.
- The second is to have an overlook of the non-functional requirements, clause 4.3.
- The third to take into consideration the user experience and his level of maturity in the use of digital, clause 4.4.
- The fourth is to have an overview of digital usage evolution, and a view on the future, both for personal and professional matter, clause 4.5.
- The fifth is an overview of user and provider interactions for service delivery. As well as a view of the service from the user's perspective to understand the composition of services and highlight the data that are under the responsibility of the user, clause 4.6.

### 4.2 Sociological and psychological context

When people meet something that seems to be intelligent, such as another person or a dog, because of the experience, it is possible to know how much intelligence to expect. On the contrary, when people meet a smart product that behaves intelligently and does things by itself, it is difficult, to anticipate exactly how much to expect. This causes different challenges when designing interaction with smart services.

Research in psychology, from scholars such as Byron Reeves and Clifford Nash [i.8], tells that users tend to treat smart products or services as though they are intelligent and intentional.

In a series of experiments, Stanford professors Byron Reeves and Clifford Nash found that even though people consciously think of computers as objects rather than persons, their immediate behaviour towards those computers sometimes resembles their behaviour towards another person.

For thousands of years, differentiating between inanimate objects and social intelligent beings was relatively straightforward. But now with technological development, the line has become more blurred.

One amazing consequence is that as products or services become more intelligent and behave as though they were intentional, they cannot always live up to the users' expectations. If the personal voice-controlled assistant on my iPhone can answer questions about the weather and which restaurant I should go tonight, why cannot it have personal opinions or feelings? Rationally, people know that it cannot because it is not conscious, but it is difficult to know what the upper limit of its ability is.

Not understanding what a product or service is capable or not capable to do, creates a negative user experience. In a review of research on how people perceive and interact with intelligent environments, researcher Eija Kaassinen and her co-authors [i.9] stated that users lose trust and satisfaction with intelligent products or services if they do not understand them.

It is necessary also to consider the social rules that apply to a smart product or service use case to anticipate deviating or unsuitable uses and avoid user disappointment. Just think of the failure of connected glasses for the general public, or a useful feature like autocorrect function that automatically corrects words in users' text messages (SMS or email) but can turn the sentence into very embarrassing or offensive words if the user trusts it too much.

## 4.3 Non-functional requirements

### 4.3.1 Overview

Non-functional requirements cover cross-cutting needs such as: QoS in ETSI TR 103 437 [i.5], security, privacy, usability, portability, customization, etc.).

The surveys conducted and attached to ETSI TR 103 438 [i.1] and ETSI EG 203 602 [i.2], sought to know the expectations of users for non-functional but essential services for a good digital experience.

First, there is a clear willing to use new services if their benefits are clear. For example, on the side of households, it is still difficult to evaluate the real value of some domestic connected object.

Secondly, the survey asked a set of questions about some new possible services able to improve the digital experience on a smart and flexible network, i.e. a New Generation Network (NGN).

People would like:

- To be informed when they are risking entering an area with low or without coverage and lose continuity of service. For 90 % of people losing the continuity of mobile service on move is a problem. This shows that communication everywhere and even on move is a standard request today for users.
- More control on the battery life.
- More control on the location data of their device.
- To find their professional configuration on different devices (for those how are in employment).
- 44 % of people would appreciate a service of bandwidth on demand fixe line and 38 % on mobile line. The level of people interesting and not interesting are quite le same on this question.

Of course, security and privacy issues are in the heart of the confidence in the future on the digital ecosystem. If the data are the new "fuel" for the digital society, trust is the "money".

The survey shows there is a very large majority of people how would like to be able to challenge their provider on what can be called "essential characteristics" of the contract, i.e. privacy, security, quality and price.

### 4.3.2 Security

Security is both a feeling and a reality, and they are different. An individual or a company can feel secure even though they are not, and they can be secure even though they do not feel it. The feeling of security matters because humans sometimes make seemingly irrational decisions that have reasonable explanations, and because sometimes emotions play a more significant role than logic. That is what makes human.

In the context of IT, even if an individual take actions that make their organization more secure, that might not be enough. They need to pay attention to making sure their actions also allow the relevant constituents (employee or customers) to feel secure.

Examples:

- EXAMPLE 1: A user of an overly quiet antimalware tool might assume that the tool is ineffective and switch to a product that makes the person feel more secure. Even if a company have a great security tool, they need to find a way to ensure that its users recognize its benefits.

**EXAMPLE 2:** A corporation may have a Chief information security officer who is very effective at strengthening the company's security posture and managing IT risk; however, the management may feel insecure unless the CISO captures the right metrics and offers meaningful reports. Part of this includes ensuring compliance with certification e.g. ISO/IEC 27001 [i.6] or regulation e.g. UK Data Protection Act or EU GDPR (General Data Protection Regulation).

**EXAMPLE 3:** A client who commissioned a security assessment may have received competent service. However, unless the deliverable includes a comprehensive review of the findings and methodology, the client may feel unsatisfied.

**EXAMPLE 4:** A company may select a security service provider that meets the firm's requirements based purely on polished sales interactions and marketing documents. Such collateral can make the prospect feel security, regardless of the vendor's actual capabilities.

Those are just a few examples for remind not to underestimate the importance of not only being secure (like in example 3), but also feeling secure (like in example 4). These two concepts are distinct yet interrelated (see more details in [i.10]).

### 4.3.3 Privacy

Privacy is the ability to control who can access information about user's private life and activities. It is important because it gives users the power to choose their thoughts and feelings and who they share them with, for example to sites or services with a social and/or messaging function. It protects their information they do not want to be shared publicly such as health or personal finances. Or real-time location data that could even compromise physical security (see more details in [i.11]). As part of a company or service providers obligation under the EC ePrivacy regulation for example will have to ensure certain rules are met. For the user keys rules include:

- 1) Communications content and metadata: privacy should be guaranteed for communications content and metadata. Metadata - data that describes other data, such as author, date created and location etc. - has a high privacy component and should be anonymized or deleted if users did not give their consent, unless the data is needed for billing.
- 2) Simpler rules on cookies: the cookie provision, which has resulted in an overload of consent requests for internet users, will be streamlined. The new rule will be more user-friendly as browser settings will provide an easy way to accept or refuse tracking cookies and other identifiers. The proposal also clarifies that no consent is needed for non-privacy intrusive cookies that improve internet experience, such as cookies to remember shopping-cart history or to count the number of website visitors.
- 3) Protection against spam: this proposal bans unsolicited electronic communications by email, SMS, and automated calling machines. Depending on national law people will either be protected by default or be able to use a do-not-call list to stop marketing phone calls. Marketing callers will need to display their phone number or use a special prefix that indicates a marketing call.

### 4.3.4 Usability

The main reasons why it is so hard to create usable products is that there is a conflict between a high-usability level and great user-experience. This might seem to be a contradiction, but there is an important difference between the two. Usability is about the "ability to use" something. The aim for a usable product or service is to make it easy to use.

A product or service can be considered to have a high level of usability when:

- 1) It requires less mental effort to use.
- 2) The frequency of mistakes using it is less, or when the mistakes are less disastrous.
- 3) It is more powerful, where "more powerful" means that it can be used to do more or do it faster.
- 4) It is more learnable, that is, when a user can figure it out quicker.

There are many guides and tools which can provide metrics to measure the usability of a product or service. While user-experience is not like usability - it is about feelings. The aim here is to create satisfaction. The product or service provider wants the user to feel satisfied before, during and after they have used their product or service. To do that they need to take all kinds of things into consideration. These can include and take care of the user's special needs:

- 1) Environment
- 2) Colours/moods (affects interaction/navigation within a digital space)
- 3) Smell
- 4) Touch
- 5) Audio feedback
- 6) Visual feedback
- 7) Trust (confidence in actions or a system)
- 8) Branding
- 9) Show-off effect
- 10) Usefulness
- 11) Practicality
- 12) Coexistence
- 13) Emotional effect

This is much harder to achieve. None of these things can be accurately analysed for example when it comes to the user providing feedback one user might have a high satisfaction from a particular chosen design choice, but another user could have low satisfaction from the design choice. The reason why there are so few great products is because of this difference. Most developers try to find the right balance between high usability and high user-experience. This is not easy. It requires a bit of luck, intuition, a sense of humble pride, and something called "usable happiness (is a product that is simple to use, and makes the user smile every time they use it)". This means the user should never ever feel helpless or stupid when using a product or service. This equally applies to all kinds of products, including software ones - mobile or web apps, websites, etc. When it comes to the usability, there is a big difference between user's assumptions and reality therefore testing, reviewing, and applying feedback is important in ensuring the balance between usability and the user experience.

## 4.4 Digital maturity

In the Digital world, users can have the ability to access large numbers of digital services applications and contents covering almost every time and everywhere a big part of their daily life activities, personal or professional.

So, the question is: are they able to master all these services by themselves, in other words what is their skills level?

To assess this, the European Commission has been surveying each year since 2015 on the level of skills of the European population.

Persons that have been using internet during last 3 months are attributed a score on four digital competence domains: information, communication, content-creation and problem-solving, depending on the activities they have been able to do. The scores in each domain are basic, above basic, and below basic. There is one more gap with the ability to have software skills, meaning manipulate features such as word processing, advanced spreadsheet functions, create a presentation, or written code in a programming language. It can be shown that these software skills can be of importance on the professional side of the life.

In the last release, while 84 % of European people used the internet regularly in 2021, only 56 % possessed at least basic digital skills. 31 % with above basic digital skills and 58 % of individuals having at least basic software skills.

With no surprise, the survey shows that the skills indicators are strongly influenced by socio-demographic aspects.

For example, only 48 % of individuals living in rural areas possess at least basic digital skills, as opposite to the ones living in the cities (62 %).

It should be noted that youth is not a determinant of digital skills and growing up in a digital world does not automatically make one digitally competent. As demonstrated by the International Computer and Information Literacy Study, which assesses digital skills of students of 13-14 years old based on a competence test, rather than self-reporting, young people do not develop sophisticated digital skills just by growing up using digital devices.

The 2030 target of the Digital Compass is that at least 80 % of citizens have at least basic digital skills. But the current growth is only around 1 % per year.

The conclusion is that alongside the efforts to make to acculturate the population to digital, usability improvements can be a big help. That is globally the way consumer focused groups are exploring with the digital user centric approach, the smart identity, and the smart interface program of the user group.

One of the goal of the survey conducted, available in ETSI TR 103 438 [i.1] and ETSI EG 203 602 [i.2] was to try to understand if, regardless of their skills, users are properly informed of what they can do with their subscriptions and equipment, have enough knowledge about the services available, and if they know how to activate them, how to configure them, how to control them, and can easily do that.

Currently, setting up a smartphone is far from easy according to most respondents. More than half of them say that they do not know how to fully configure their smartphone, of which 15 % say they do not really know how to do it.

According to responders, some cellular parameters are difficult to set, from IEEE 802.11 [i.15] to voice over 4G (VoLTE), WIFI call (VoWIFI), etc.

34 % of people change the setting of their smartphone less than once a year.

-56 % of responders get online to the personal page of their fixe subscription never or rarely and it is quite the same for the mobile (53 %). Obviously, for these people, it is difficult to have a good level of knowledge and control.

A lot of people (42 %) do not really know the differences between the successive generations of GSM technologies, and 5G and is different level of implementation does not simplify the situation.

Developing the adoption and use of digital in business is a specific challenge. To succeed in its digital transformation, a company can ensure the large adoption of digital uses by its employees. However, they often find themselves confused by the gap between their digital experience in the private sphere and the one they know at work.

Therefore, it is interesting to note that an emerging market has recently appeared under the name of Digital Adoption Platform (DAP), which are solutions that guide the user step by step, over the application, a bit like a GPS.

The concept is not new. Almost twenty years ago, tooltips were offered to democratize integrated management software packages and office suites. However, by using Artificial Intelligence (AI), DAPs offer an approach that is no longer static but dynamic and contextualized. They come to the user's aid precisely when they need it.

Developing user maturity is one of the keys to success for digital transformation. It is one of the components of digital inclusion. Hence, for the purpose of wanting to see the maturity increase rapidly and prevent the widening of the digital divide, it is important to develop smart interface, based on the smart identity concept.

## 4.5 Usage Evolution and digital transformation of society

This clause provides some global data to draw the background.

World population stood at 7,85 billion at the start of April 2021.

At this date, there are 5,27 billion unique mobile phone users worldwide, which means that more than two-thirds of people on the planet now own a mobile phone, and the number of Internet users reach 4,72 billion, more than 60 % of the total world population (increasing by 7,6 % in one year).

More than 6 out of 10 people in the world are now online on internet, and the growth is of 7,6 % in 12 months.

Social media continues to be a key driver of internet adoption, and GWI's latest research indicates that nearly 99 % of internet users worldwide between the ages of 16 and 64 use every month a social network or messaging platform connected to the Internet. More than half a billion new users have joined social media platforms in the past 12 months, bringing the global total to 4,33 billion in April 2021.

3,8 million searches are performed on one of the most used search engines every minute.

Clearly, the digital technology is revolutionizing user's uses in terms of communication, learning and social relations. It is a whole ecosystem that is part of and reinvents daily lives, in private and professional lives.

If defining the digital transformation as a dematerialized economy based on the exchange of data, shows the rise in the past 10 years of the advent and quick development of:

- e-commerce;
- e-administration;
- e-payment;
- e-learning;
- music streaming and video on demand;
- telework;
- etc.

All this faculties of online interaction contributes to the inclusion of everyone in the social and economic life in modern society. That is why the EU included in Directive (EU) 2018/1972 [i.12], article 84 the obligation for Member States to ensure everywhere on their territory broadband access to the Internet at an affordable price and with sufficient speed to allow the use of the main online services described as:

- 1) e-mail
- 2) search engines enabling search and finding of all type of information
- 3) basic training and education online tools
- 4) online newspapers or news
- 5) buying or ordering goods or services online
- 6) job searching and job searching tools
- 7) professional networking
- 8) internet banking
- 9) eGovernment service use
- 10) social media and instant messaging
- 11) calls and video calls (standard quality)

It is mainly technological innovations that have enabled the development of these new services:

- the web of course, which is now considered to have entered its third generation (web3.0 or semantic web);
- the mobile that has benefited from the constantly increased performance of smartphones;
- GPS that allows people to move easily and whose precision is now reinforced by the WI-FI fingerprint;
- the Internet of things that allow people to monitor their house for example;
- the Next Generation of Network operating on the 4G+, 5G, and optical fibre; and so on.

These technologies allow the development of many new added value services, which services empowered people on their life.

If one considers the user point of view, the next step (already partially complete, is to be able to move from one service to another seamlessly, everywhere, with different devices, maintaining is liberty of choice.

All these new usages and digital environment introduce the question of the definition (content and scope) of digital identity(ies).

Civil identity, real identity, personal identity, biological identity, professional identity, sovereign identity. These multiple possibilities of identities, of fact or desired, show the importance of establishing precisely or surely depending on the context that a person is what they claim to be, see [i.13].

By determining that digital identity does not correspond to a transposition of civil identity into cyberspace, the exponential increase in the types of attributes allowed by dematerialization highlights the need to speak not of a digital identity but of digital identities.

Work on smart digital identity may need to integrate part of the personal identity but at the right level required to deliver the service based on the knowledge of the context of use and the wishes of the user.

That is called the User Profile, as described below.

## 4.6 User and Provider interaction for service delivery

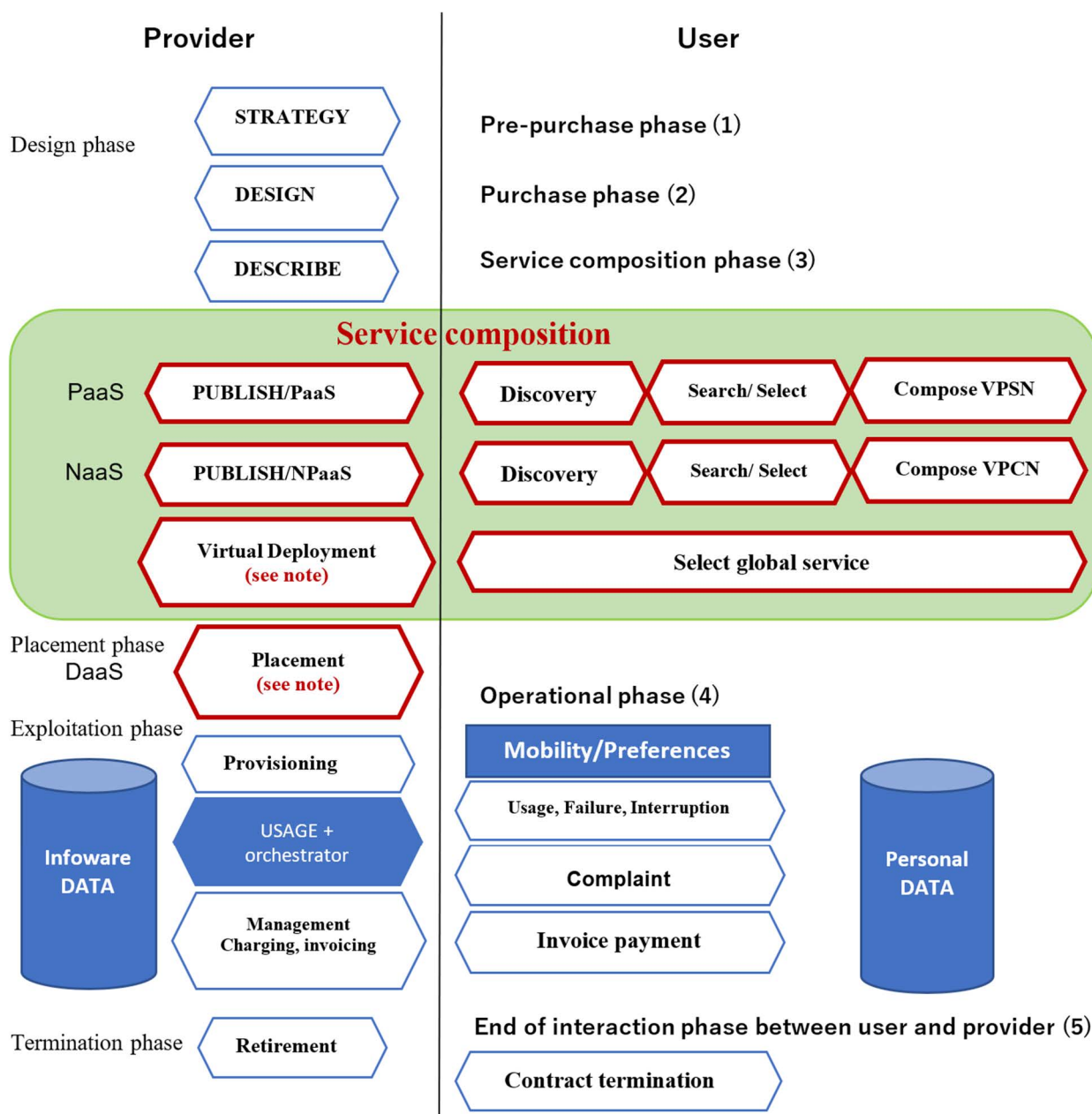
### 4.6.1 Overview

During the life cycle of a service, the provider/user interaction thinks differently. The user is at the centre of the architecture and wants personalized services. As a result, the provider offers different services allowing a composition in line with the desired preferences, location, and agenda of the user. This interaction is based on a service composition that is then a dynamic and flexible end-to-end session.

The "delivery" service supports all the service components, as for example in the "video streaming" service it has both the services of selection of a film according to these preferences, that of the method of payment, the mode of diffusion and the choice of the most appropriate access network.

The parallel of the supplier and user/customer view during the life cycle of a service will be presented and detailed in clause 4.6.2. Then, in order to take into consideration a "user centric" service, a composition of personalized services in the field of health is described, see clause 4.6.3.

## 4.6.2 Life cycle of digital services in provider side and user journey



NOTE: Deployment is separated from placement.

**Figure 1: The parallel of the provider and user/customer view during the service lifecycle**

Figure 1 illustrates the different phases on the provider side (A) and on the user side (B) during the life cycle of a digital service.

The specificity of the interaction between the A (provider) and the B (user) is supported by the services composition and data management (C).

**A. On the provider side** according to its strategy will have:

- *Design phase:* the phase that includes the following steps:
  - "Design": represents service design.
  - "Describe": describes the offered services.



- "Publish/ PaaS" which publishes application services on a PaaS platform and in a catalogue/repository.
- "Publish/ NPaaS" which publishes network services on an NPaaS platform and in a catalogue/repository.
- "Virtual deployment" which allows to design VPSN and VPCN according to QoS and independently of the hardware.
- *Placement phase* with the "placement" step which consists of placing the components in virtual machines according to the QoS requested by the service and the QoS offered by the equipment (memory and CPU).
- *Exploitation phase* that includes the following steps:
  - Provisioning step during which the control plan reserves the resources.
  - Usage represents the service operation in the data plan.
  - Management step during which the management plan reacts to changes.
- *Termination phase* with the finalization step which consists of withdrawing the service when the service is no longer operated.

**B. On the user side** will have:

- *Pre-purchase phase* which is the knowledge phase of supplier offers.
- *Purchase phase* that takes place after the user's choice.
- *Service composition phase* which depending on the case will have:
  - "Discovery" to discover the services offered by user's environment depending on the user location, profiles, etc. (train station, town hall, airport).
  - "Search/Select" to choose the application services offered by the provider's platform (PaaS).
  - "Compose VPSN" which allows to have the composition of services personalized for the desired application. The service composition should be obtained by a service combination from basic or composed services. This composition offers the possibility to create new services answering to the personalized needs of users with a service logic.
  - "Discovery" to discover the network services offered by its environment (WIFI, Ethernet, 4G, 5G).
  - "Search/Select" to choose the network services offered by the provider's platform (NPaaS).
  - "Compose VPCN" which allows to have the composition of network services for the connection "networks".
  - "Select global service" when the service composition is proposed by the provider.
- *Operational phase* with the following steps:
  - "Usage" in a context of continuity of service (mobility and preferences), allowing to automatically compensate for any interruptions and interruptions.
  - "Complaint" when QoS criteria (availability, reliability, time, capacity) is not respected.
  - "Invoice payment".
- *End of the interaction phase* between user and provider which is signified by a letter of termination.

### C. About the interaction between the user and the provider.

In the digital world focused on the personalization of services, the difference lies in the phase of service composition. Indeed, *the* interactions between suppliers and users were during pre-purchase phases (1), purchase (2) and the operational phase (4). Whereas *today* with the user centric approach, the user can also be an actor in the service requested through selectivity, discovery in the phase of composition of services (3), shown in green in figure 1.

**In figure 1, the green pad highlights the composition of services.**

A service can be an elementary service or to represent a composition of elements (micro-services) that meet the same objective(s). In a "user-centric" approach, Service-Oriented Architecture (SOA) makes it possible to organize a set of isolated micro-services into a set of interconnected services, accessible by API (Application Programming Interface).

The specificity of the digital world and of the "user centric" approach is centred on the personalization of services that induces the interaction and collaboration between the user and the provider to have an end-to-end service "anytime, anywhere through any means". This end-to-end service is concretized by a service composition "application service, network service, equipment service" representing the "global service" called "delivery service".

**In figure 1 the elements in blue** highlight the specificities of context (digital and user centric) namely:

- the actors at the data level:
  - "Infoware data" contains data managed by providers.
  - "Personal data" are the data that are under the responsibility of the user. They are usually personal in nature, but some may be visible from the outside.
- the dynamicity at the "usage" step on the provider side, which means that the control plane, the data plan, and the management plan collaborate through an orchestrator.
- the "mobility and preferences" properties at the grade of the "usage" step on the user side, which means that the continuity of services is ensured regardless of the user's movements and preferences.

#### 4.6.3 User-centric service (Health use case)

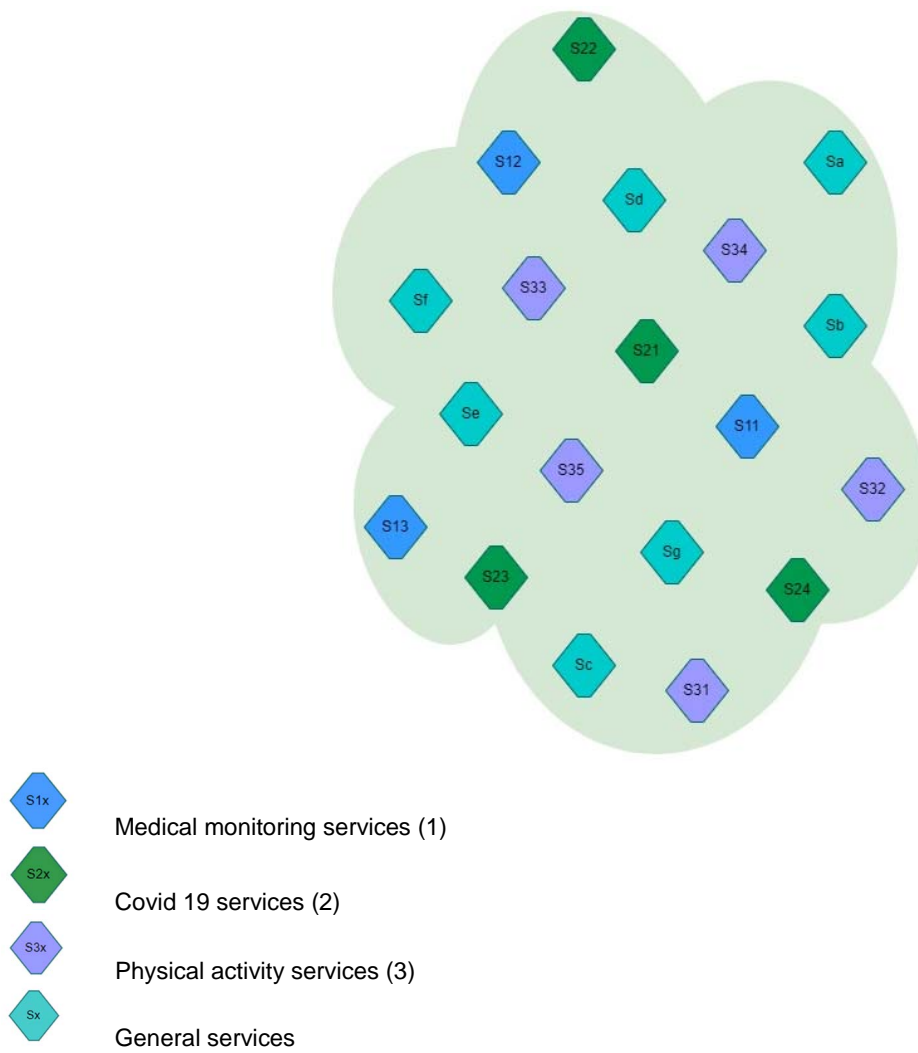
To implement an example of a user-centric service, the choice is on a typical day from the period just gone through, which was heavily impacted by the COVID-19 epidemic. Because of the many phases of lockdown that was experienced, users got into the habit of teleworking, thus reducing the frequency of their movements. This sedentary lifestyle coupled with a lack of physical exercise can be the cause of many diseases.

In this chosen "user centric" service, in the present document are described the services related to:

- the medical follow-up of the patient (1);
- prevention and monitoring of epidemics (2);
- the user's well-being (3).

Therefore, a health profile of the user in the role of patient is described, with sensitive information (vital parameters of the user such as temperature, heart rate, blood pressure, oxygen saturation) captured from the patient's internal resources and disseminated information (vaccination) stored in the user's knowledge base.

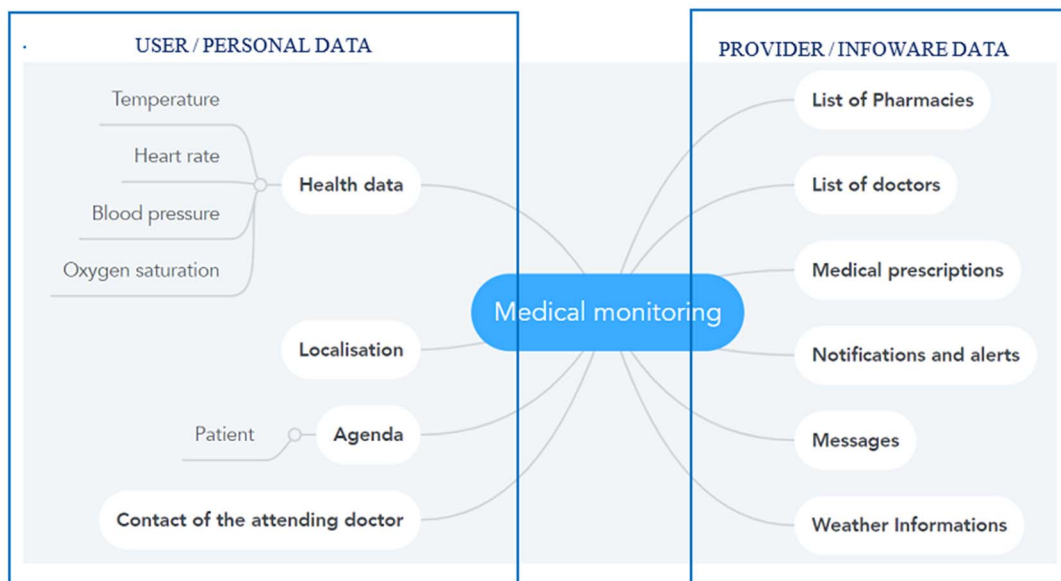
Providers could offer services (1), (2) and (3) through a platform (figure 2). The information (resulting data) is grouped to the right of the diagrams (figures 3, 4 and 5).



**Figure 2: PaaS platform for the user centric health service**

The services associated with the patient's medical follow-up could be the following:

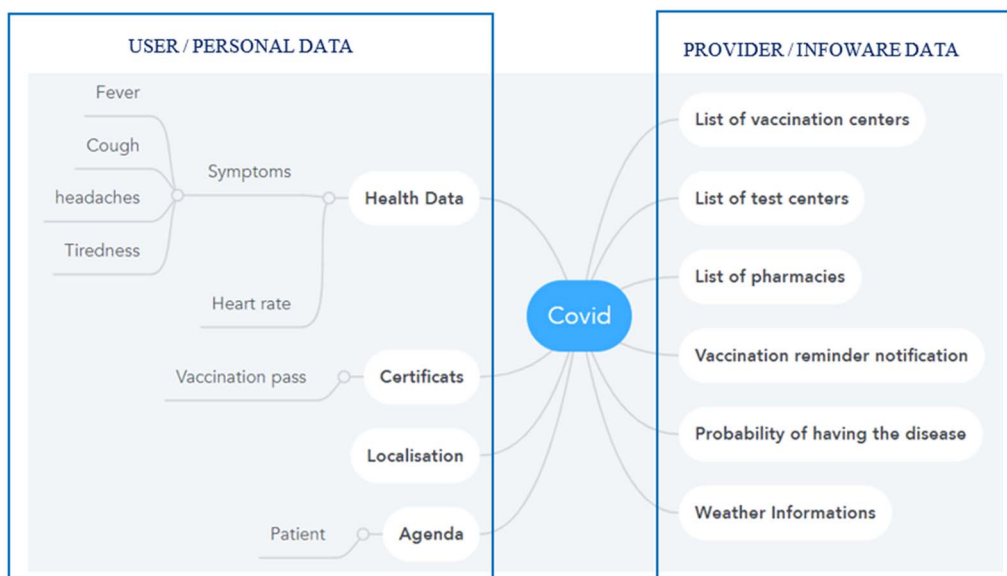
- S11: Secure messaging
- S12: Notification and alert system
- S13: Medical prescriptions
- Sa: List of geolocated pharmacies
- Sb: List of geolocated doctors and urgency services
- Sc: Appointment booking platform
- Sd: Teleconsultation
- Se: Geolocation
- Sf: Health advice
- Sg: Weather



**Figure 3: Medical monitoring services related data**

The COVID-19 epidemic, also known as Coronavirus, has been affecting the world since 2020. To facilitate the daily life of people with or without COVID-19, services optimizing the accessibility of tests, vaccines, medical care, can be offered. The PaaS platform could offer the following services:

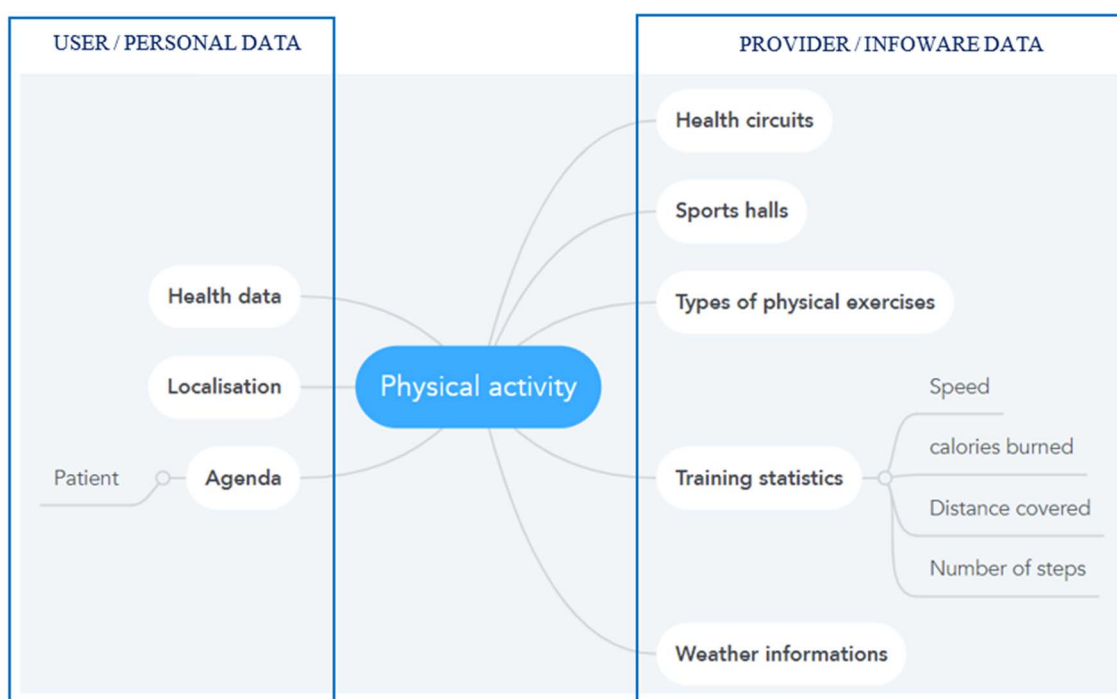
- S21: List of geolocated vaccination centre
- S22: List of geolocated testing centre
- S23: Certificates (vaccination)
- S25: Vaccination reminder notifications
- Sa: List of geolocated pharmacies
- Sc: Appointment booking platform
- Se: Geolocation
- Sg: Weather



**Figure 4: Covid 19 Services (2) related data**

To ensure the well-being of users through physical activity, the PaaS platform could offer the following services:

- S31: Geolocated health circuits
- S32: Geolocated gyms
- S33: Personalized physical exercises
- S34: Training statistics
- S35: Online shop
- Se: Geolocation
- Sf: Health advice
- Sg: Weather (and air pollution)



**Figure 5: Physical activity Services (3) related data**

This user-centric service in the field of health shows that the user will oversee entering and managing all the data relating to their health (left side of the diagrams, see figures 3, 4 and 5). It is up to them to decide which data can be private or disseminated.

---

## 5 From User Digital Clone to Smart Identity

### 5.1 The Modelling

Modelling the digital ecosystem leads to abstracting from the real world. This abstraction is based on four mental operations:

- 1) Simplification, by retaining the essential resources of the digital ecosystem, namely "equipment, networks, services".
- 2) Generalization, retaining common properties related to data structure.

- 3) Selection, by isolating the specific feature of the digital world, which is the N-to-N connection, the linking of the different roles of the user with the different appropriate services.
- 4) Schematization, analysing the objectives and positioning of the user (at the centre of the architecture). The purpose is to render the whole as a data system.

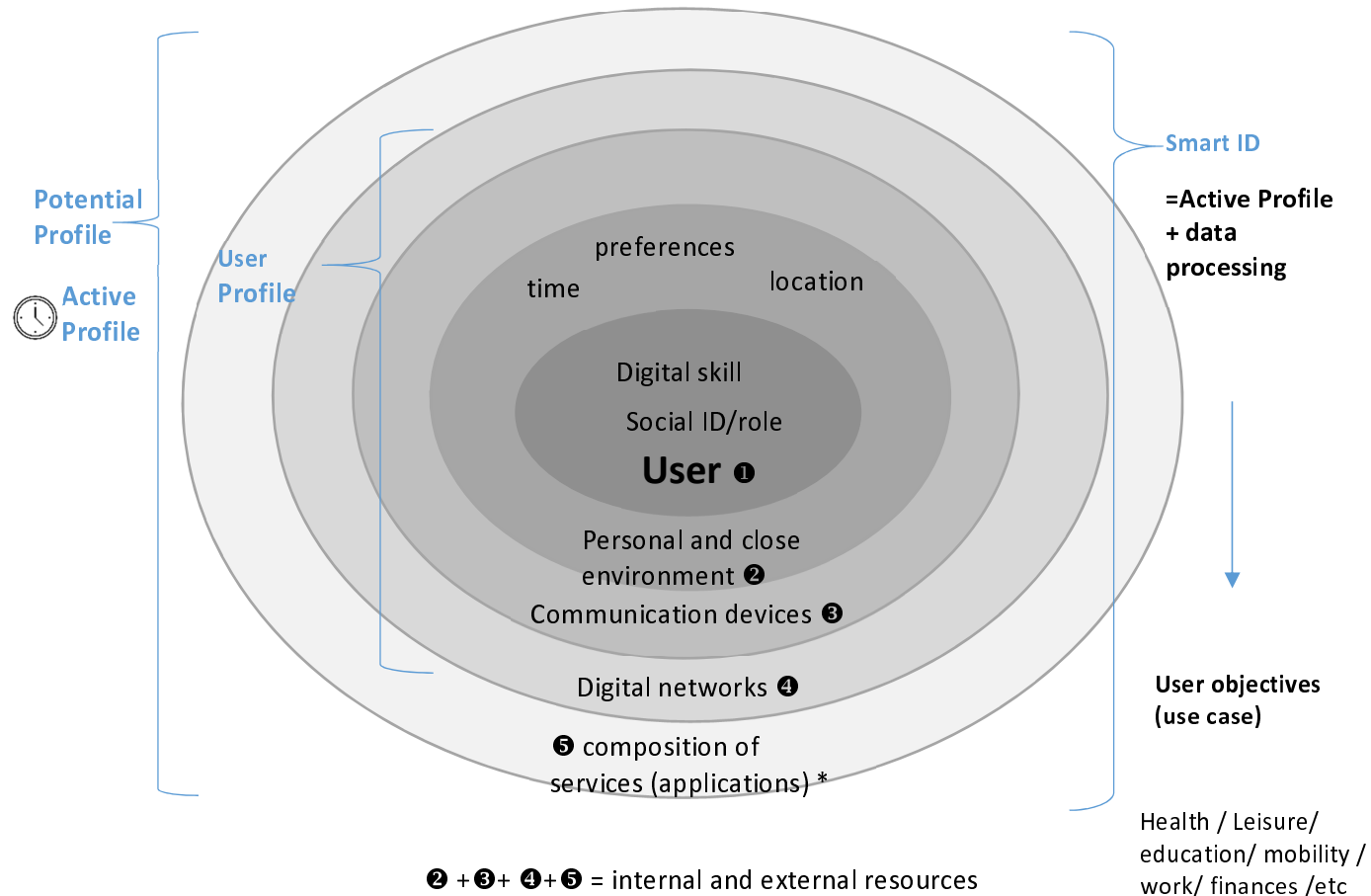
The result of these four operations makes it possible to:

- 1) Generate a representation of the user's digital clone, i.e. their image in the ecosystem (clause 5.2).
- 2) Propose the different profiles constituting the knowledge base (clause 5.3).
- 3) To describe the associated models (clause 5.4) the User Profile model and the smart ID model which is the aim of this study. This identity is intended to be intelligent because it addresses an identity in full consciousness.

The following clauses will return to data processing (clause 6) and the new technologies used for smart identity (clause 7).

## 5.2 User Digital Clone representation

Digital clone = User knowledge base (all profiles) + data processing



NOTE: \* digital services examples: weather, mobility offers, payment, agenda, domotic, etc.

Figure 6: User Digital Clone

Figure 6 shows:

- 1) In accordance with the user centric approach, the user is at the centre of the architecture and therefore at the centre of the representation of the digital world recorded in:
 

Ellipse sector ❶, the user first has:

  - Social ID: civil identity i.e. name, age, sex, job, etc.
  - Personal data (physiological data, bank data, school data, etc.).
  - digital skills: level of expertise in the digital scope.
- 2) The characteristics of the user centre to know in priority their agenda, location, preferences represented in:
 

Ellipse sector ❷:

  - Location and time.
  - Preferences and special needs.
  - Personal and close environment in the real world (access to internal resources).
- 3) The result of the modelling makes it possible to represent the resources by concentric ellipses translating levels of visibility for the connection of the different components (equipment, networks, services):
 

Ellipse sector ❸, communicating device:

  - Personal IoT (connected watch, tensiometer, connected scale, etc.).
  - Communicating systems: smartphone, box, hot spot WIFI, etc.
  - External IoT (rain gauge, pollution sensors, etc.).
- 4) All external resources (edge, switch, server) of the connections:
 

Ellipse sector ❹, digital communication networks (PAN, LAN, WAN): 4G/5G, public WIFI, Wired network.

Ellipse sector ❺, applications necessary to meet the user' needs. An application is described as services composition (weather, agenda, mobility offer, payment, etc.).
- 5) On the left of the figure are positioned:
  - The Potential Profile grouping all the ellipses from the informational point of view.
  - The Active Profile (see definition in clause 5.3.4.1).
- 6) On the right of the figure are positioned:
  - Smart ID is Active Profile and data processing (clause 5.3.4).
  - User Objectives. Finality of delivered service(s).

In conclusion, the digital clone includes all the data mentioned in figure 6 (user knowledge base) plus the data resulting from data processing (clause 6).

## 5.3 Profiles: user knowledge base

### 5.3.1 Overview

All the elements of the user digital clone will be described through profiles that will feed the user knowledge base. In this clause are described the main profiles allowing the construction of the smart identity, namely User Profile which is at the centre of the identity (clause 5.3.2), Potential Profile which is the clone (clause 5.3.3) and Active Profile which is the identity at a specific moment in the life of the user (clause 5.3.4).



The present document addresses the Information Model.

The ACIFO method turns out to be recursive and will be applied to the informational dimension. Thus, it will have data collection, internal and intra-data communications through the APIs of use, control and management. An Informational production will gradually enrich the clone, thanks to the Processing Features (AI, automation, inference, etc.). And organizational choices will be made based on security and storage optimization (cloud, fog, dew).

In fact, with the introduction of Artificial Intelligence (AI), the information model is enriched with additional data, collected every day and which will refine the user's knowledge (for example data from sensors, connected objects in their environment, etc.). It is this learning that will provide a better understanding of the needs and a contextualization of the compound services.

This is why the data will be described in the User Profile model (clause 5.3.2.2) and the data processing (which enriches this data) will be detailed in the Potential Profile model (clause 5.3.3.2).

## 5.3.2 User Profile

### 5.3.2.1 User Profile definition

The User Profile is the informational representation of the complete image of the user in the digital ecosystem including the description of personal information (social identities -roles-, digital skills), the description of user centric characteristics (preferences, location, agenda) as well as the description of internal and external resources (equipment, networks and services).

### 5.3.2.2 User Profile Model

The User Profile model is an information component with an Architecture (figure 7A), Communications (APIs), data (I), Functionality (services (figure 7B)) and Organization (cf. ACIFO).

In accordance with the representation of the user digital clone (clause 5.2), the User Profile architecture is described by the following six components (figure 7A):

- Personal information, used to identify the user (name, age, gender, etc.).
- Data related to the user's field of activity (profession, parent, finance, leisure, health, etc.).
- Internal and external resources used, including equipment, networks and services required to meet the user's needs.
- User agenda, which records their activities according to time (timing).
- Location, as there is mobile.
- Personal preferences by location and time slots (agenda), but also those improving accessibility to different services (e.g. language, interface properties), considered as general preferences.

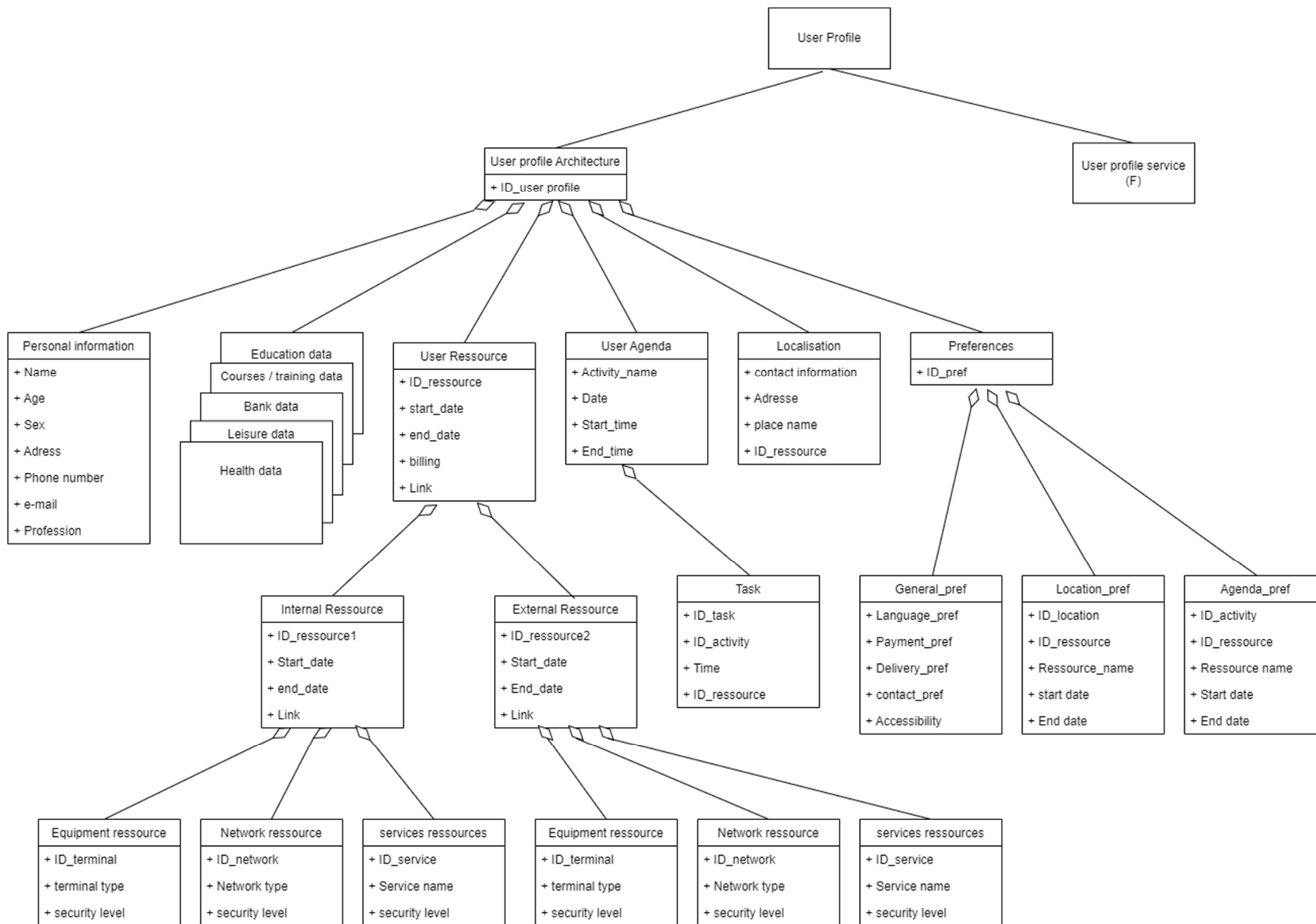
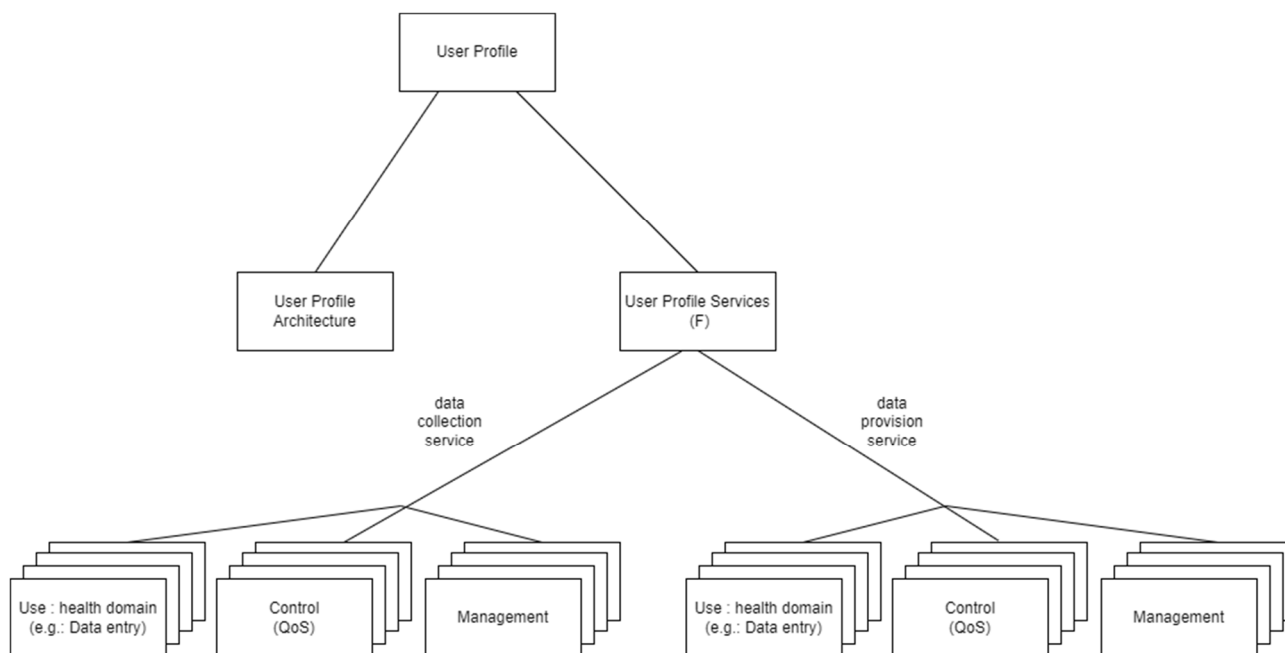


Figure 7A: User Profile Architecture

The User Profile service corresponds to the processing of the user's data either to collect them or to provide them during the various solicitations. Each of these services has three (3) APIs, namely:

- The use, whose basic function is the entry of data relating to a field of activity of the user.
- Management, which manages the usage service.
- Control, which refers to QoS.

There are several uses (one per activity) and each has its own control and management part.



**Figure 7B: User Profile Services**

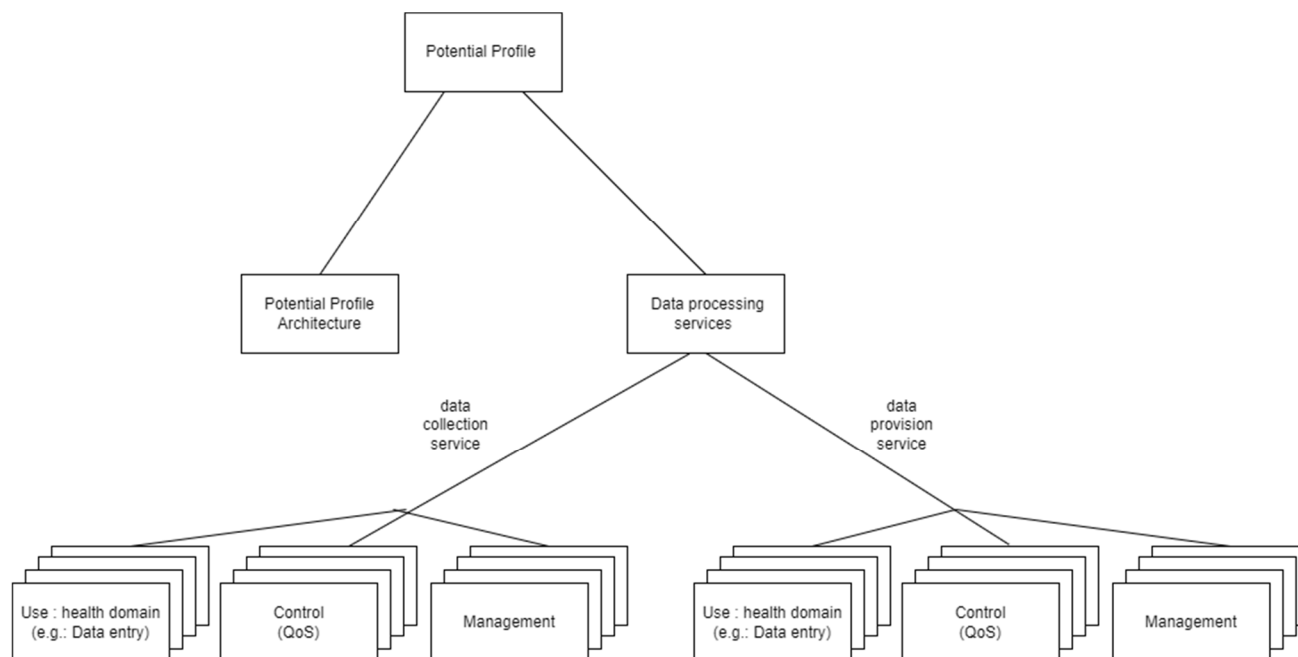
### 5.3.3 Potential Profile

#### 5.3.3.1 Potential Profile definition

Dynamic profile which, based on the User Profile, describes, by anticipating needs, all possible sessions (connections between the user and the requested service) according to location, agenda and activity, enriched by data processing (AI, machine learning). In other words, it represents the clone.

#### 5.3.3.2 Potential Profile Model

The Potential Profile Model (figure 8) information component presents an architecture associated with services (see ACIFO). These services will enrich the knowledge the Potential Profile has of the user by different data processing operations that the Potential Profile will process later.



**Figure 8: Potential Profile Model**

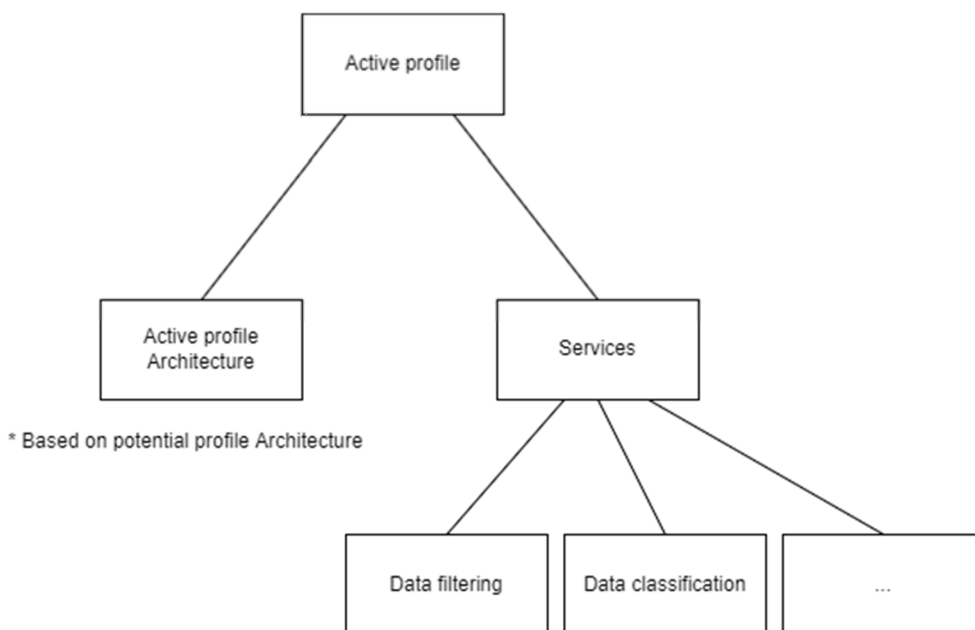
## 5.3.4 Active Profile

### 5.3.4.1 Active Profile definition

Is the User Profile at a precise moment, in a precise place according to the course of the activity. It is therefore the description of all the components of the relationship between the user and the requested service at a given time for a given activity.

### 5.3.4.2 Active Profile Model

Active Profile Model (figure 9) is based on the informational architecture of the Potential Profile with as its main services the filtering and classification of Potential Profile data for a given time slot and a given activity.

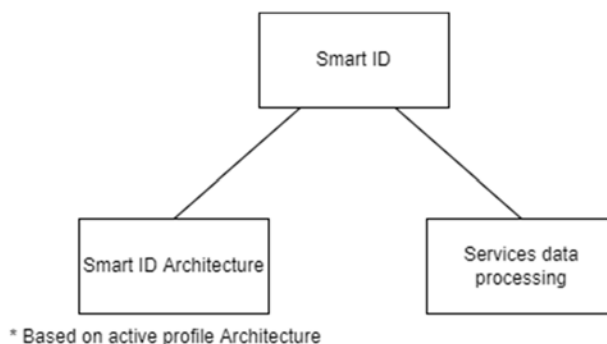


**Figure 9: Active Profile Model**

## 5.4 Smart ID

### 5.4.1 Smart ID Model

Smart ID Model (figure 10) is based on the informational architecture of the "Active Profile". It is an informational component with its own services. Depending on the time slot and the activity chosen (in Active Profile), it will enrich the data (by data processing) in order to have decision-making data and recommendations useful to the user.



**Figure 10: Smart ID Model**

### 5.4.2 Smart ID: Informational Use Cases

#### 5.4.2.1 Overview

The use case personal of Bérénice, a 25-year-old girl currently on an internship as a Research Engineer at IMT. Her daily life is very punctuated by studies, work, shopping, outings with colleagues/friends, not to mention sports sessions to keep fit whenever the weather allows.

Its User Profile architecture (clause 5.4.2.2) is powered by the information collection service. Then, the Potential Profile (clause 5.4.2.3) will enrich the health context.

The Active Profile will be filled in according to two time slots chosen from Bérénice's agenda. The first (clause 5.4.2.4) corresponds to a remote medical examination and the second (clause 5.3.2.5) to his fitness session.

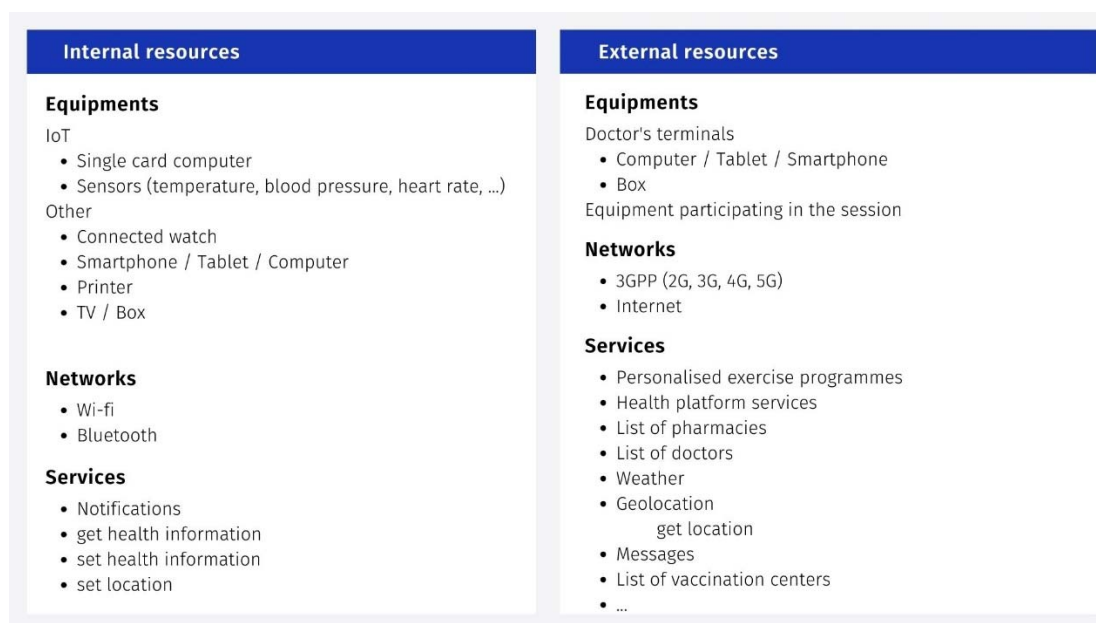
The Smart ID of the first time slot (clause 5.4.2.6) and smart ID of the second time slot (clause 5.4.2.7) finalize these informational use templates.

### 5.4.2.2 User profile architecture



**Figure 11: Personal Information**

Depending on the scenario are chosen the templates of the social ID, the main activity (trainee) and the health section (figure 11) are filled in.



**Figure 12: Resources**

The two internal and external resource templates (figure 12) list the equipment, networks and services that will participate in the session opened by Bérénice. Each resource will be filled in by the "resource profile" template that will be described during the PoC.



**Figure 13: Template Localization**

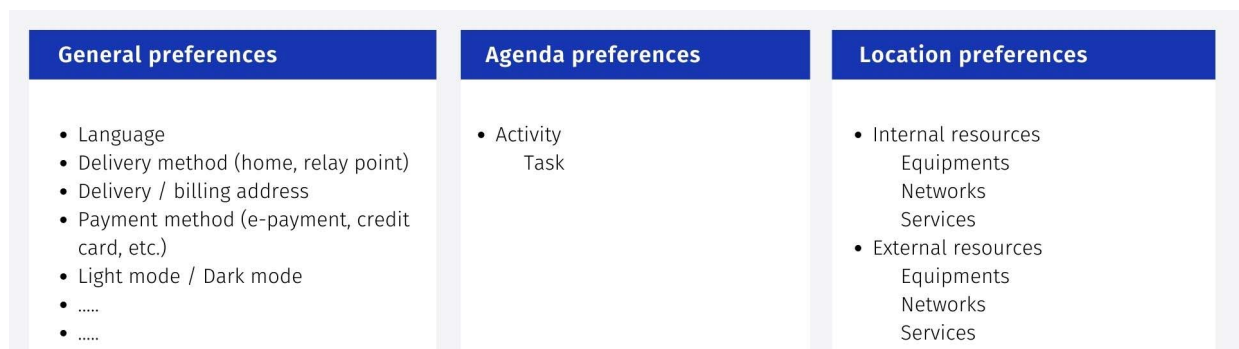
The user's location will be updated through the geolocation service (figure 13). The request will be made by a get location and updated by a set location. The periodicity of the update will be described during the PoC.



**Figure 14: Template Agenda**

The calendar template (figure 14) will be filled in with the precision desired by the user. Associated with this calendar, the "task" templates will contain the information describing each task.

It should be noted that the services of the Potential Profile will be able to complete the agenda following learning techniques (AI). Assistance in carrying out the task and decision-making aids may also be provided by appropriate services (figure 7A).



**Figure 15: Preference Templates**

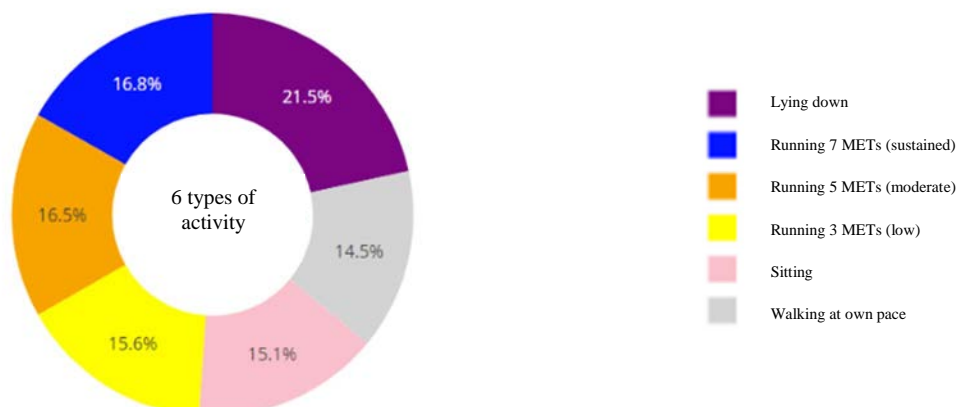
The preferences templates (figure 15) allow to finalize the customization and to select the resources preferred by the user according to the place and the activity.

#### 5.4.2.3 Enrichment of the User Profile by the Potential Profile

As mentioned above, the services of the Potential Profile will enrich the user's knowledge. In this scenario, the internal health department of Bérénice's connected watch records her physical activity throughout the day (inactivity time, number of steps, etc.) to suggest sports sessions when she has spent the majority of her day working (physical inactivity time).

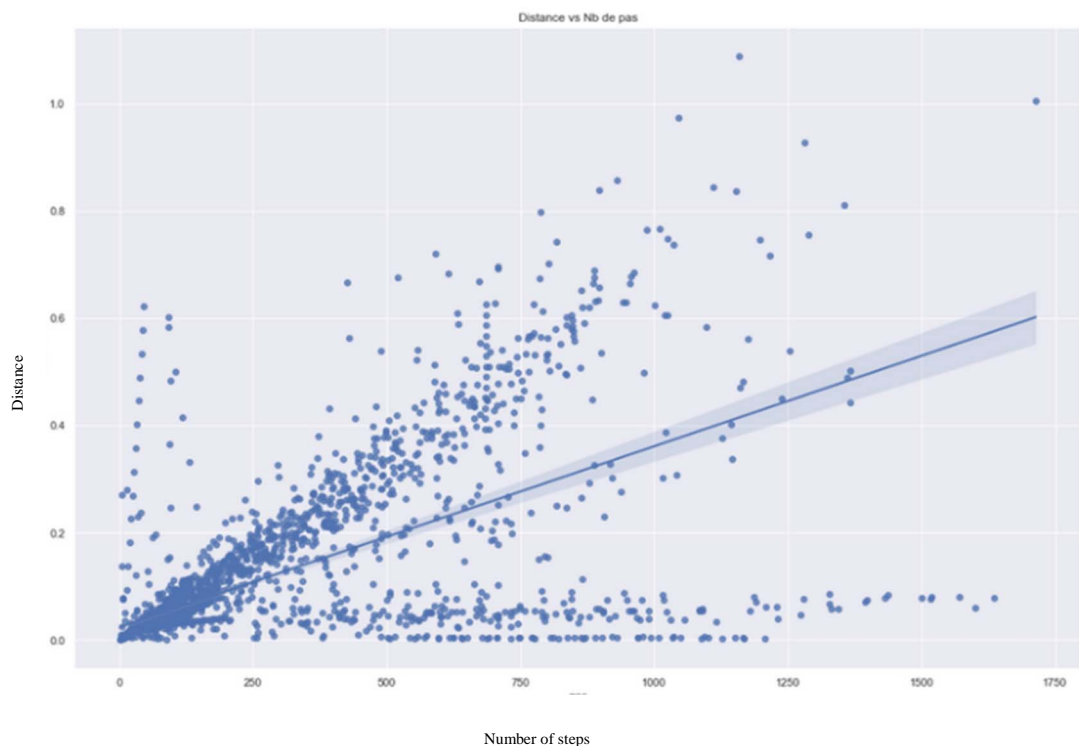
After each sports session of Bérénice, information about her constant health and training statistics is collected. These are processed by artificial intelligence techniques (which will be specified during the PoC) to feed the personalized training programs to which she is subscribed.

For example, using the data retrieved from the connected watch, a classification of the user's activities can be made : sleeping, running (low, moderate, sustained pace), sitting, walking (figure 16).



**Figure 16: Distribution by Activity**

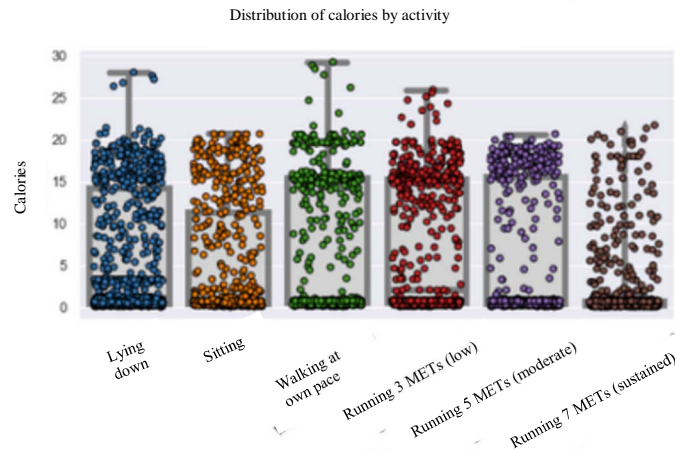
An evaluation of the distance travelled can be made based on the number of steps taken (figure 17).



**Figure 17: Distance travelled according to the number of steps taken**

One can also obtain the distribution of calories by activity. See figure 18.

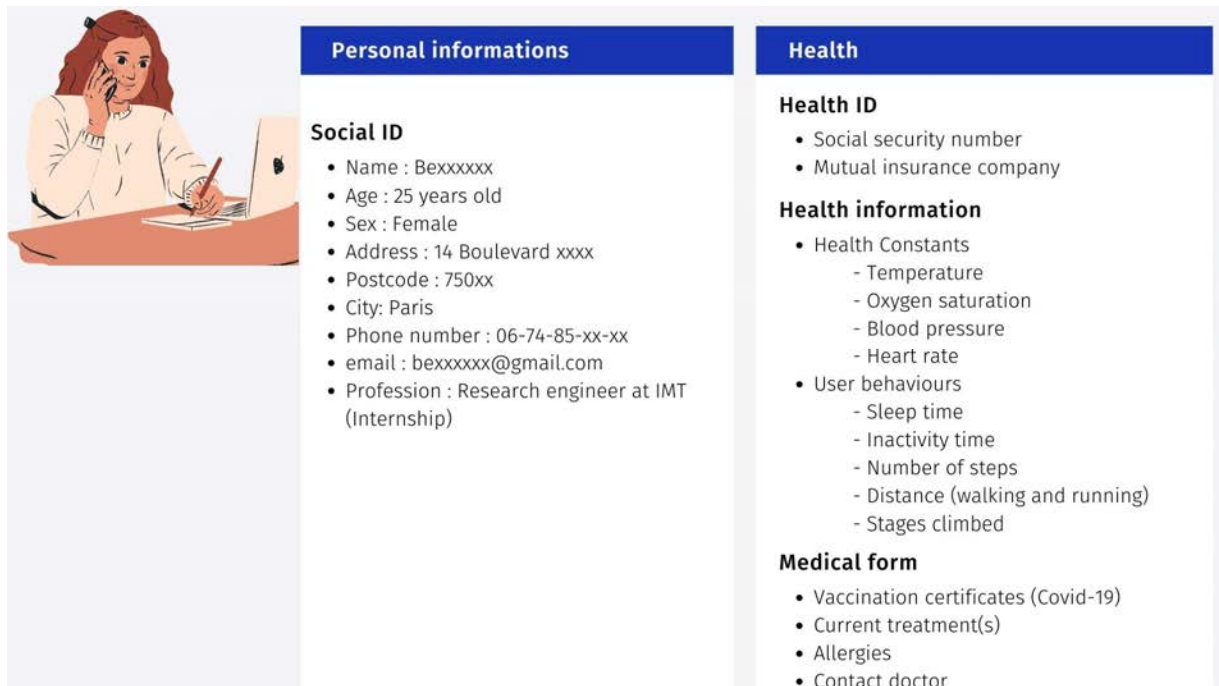




**Figure 18: Calorie Distribution by Activity**

The digital secretary will be able to use these results to suggest to the user to do a physical activity based on calories.

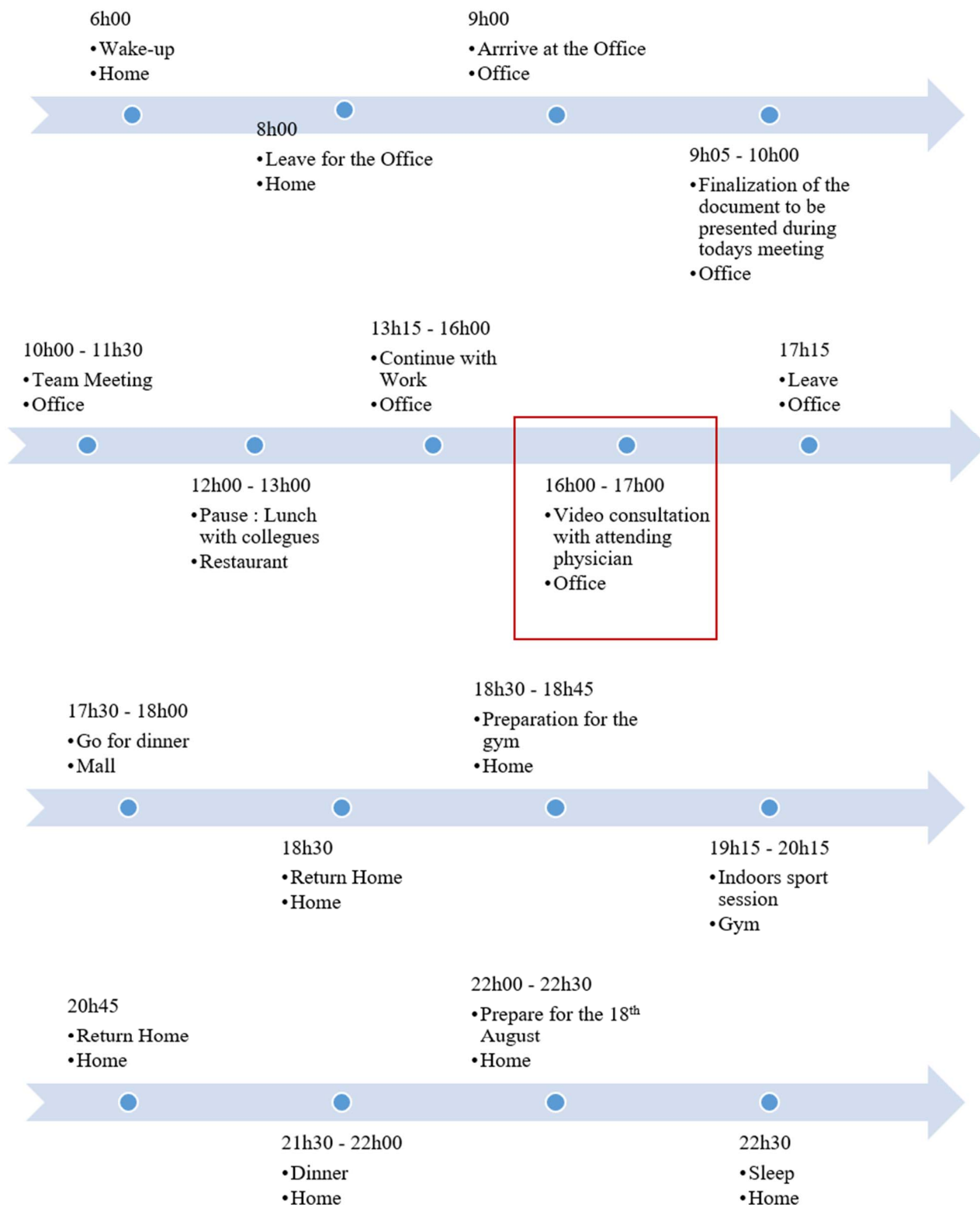
This gives the following potential architecture profile for the information in the health section of figure 19.



**Figure 19: Potential Profile - health category**

#### 5.4.2.4 Active profile for the first time slot

Bérénice received a notification of exceeding the threshold of her health parameters and, worried, she decided to make an appointment with her attending physician for August 17, 2022. Here is what her agenda indicates for this day.



**Figure 20**

From the User Profile, the Active Profile service from the "calendar template" (for the first time slot) will select the "health template" with useful information for its remote medical visit (figure 21). As for the "resource template", it will select all the active resources of the session of this medical visit.

The different "resource templates" will be filled in during the PoC.

The "location template" clearly indicates the fact that the visit will be from Bérénice's workplace.

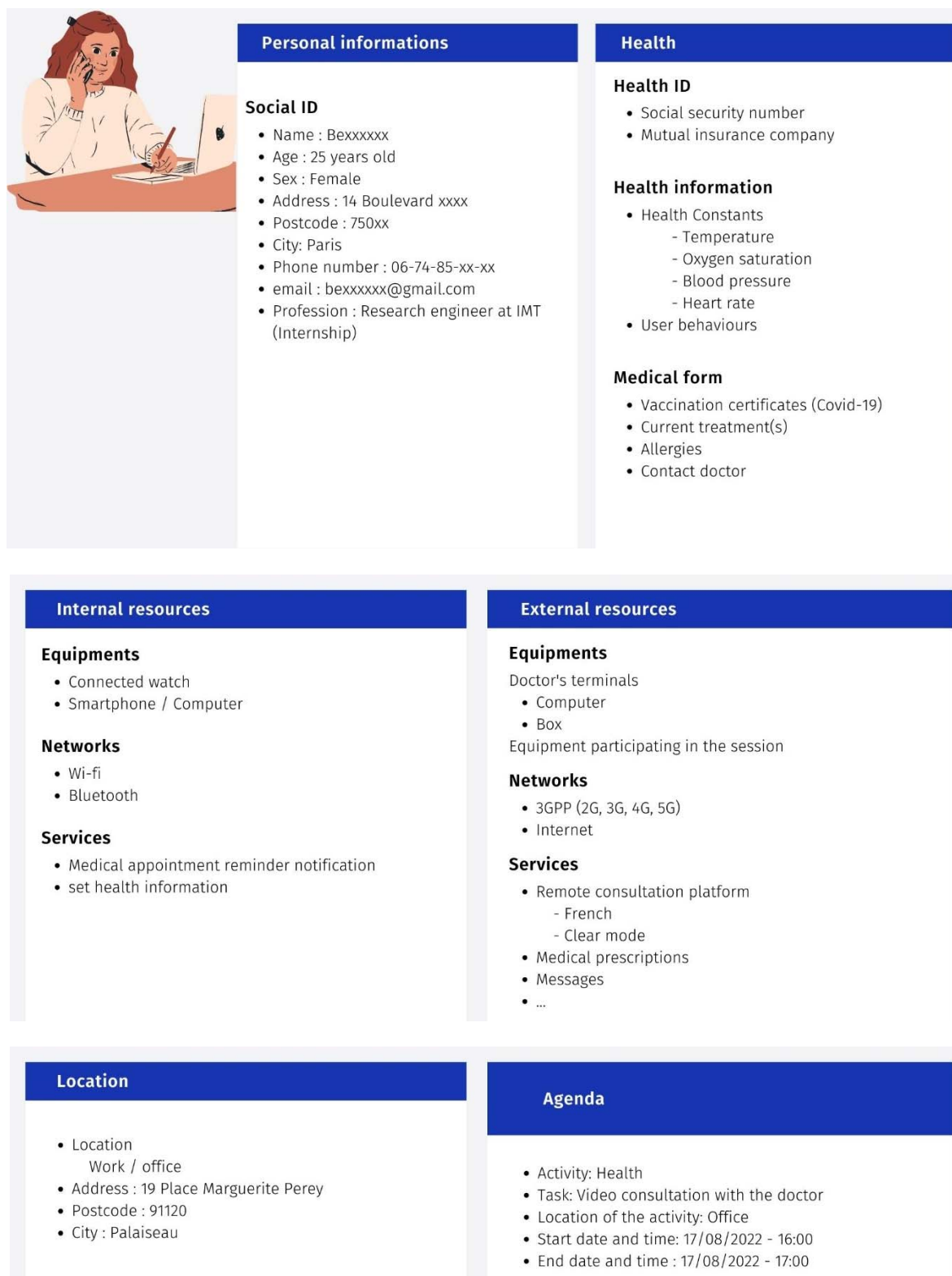


Figure 21: Active Profile at 4:00 p.m. on August 17<sup>th</sup>, 2022

5.4.2.5 Active Profile for the second time slot

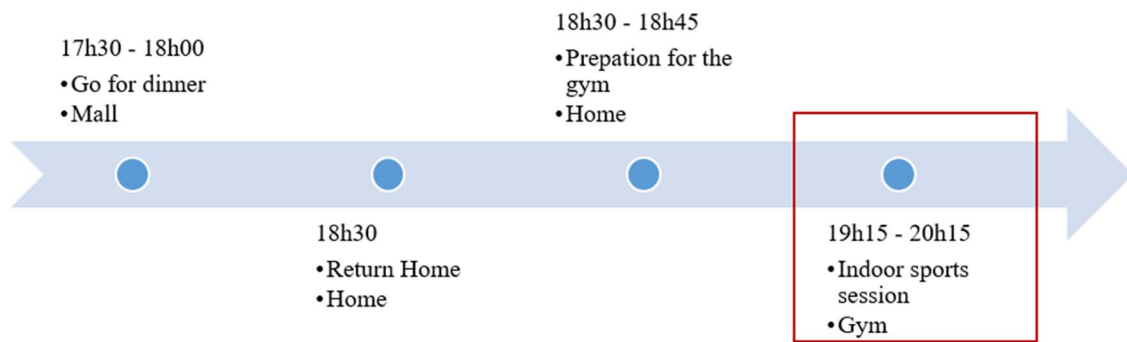


Figure 22

For the second time slot, the "Active Profile template" selects information about physical activity at the "health template" (figure 23). At the level of resources, those concerning the linking of resources used for this activity.


	<p><b>Personal informations</b></p> <p><b>Social ID</b></p> <ul style="list-style-type: none"> <li>• Name : Bexxxxxx</li> <li>• Age : 25 years old</li> <li>• Sex : Female</li> <li>• Address : 14 Boulevard xxxx</li> <li>• Postcode : 750xx</li> <li>• City: Paris</li> <li>• Phone number : 06-74-85-xx-xx</li> <li>• email : bexxxxxx@gmail.com</li> <li>• Profession : Research engineer at IMT (Internship)</li> </ul>	<p><b>Health</b></p> <p><b>Health ID</b></p> <ul style="list-style-type: none"> <li>• Membership card (gym pass)</li> </ul> <p><b>Health information</b></p> <ul style="list-style-type: none"> <li>• Health Constants</li> <li>• User behaviours                             <ul style="list-style-type: none"> <li>- Sleep time</li> <li>- Inactivity time</li> <li>- Number of steps</li> <li>- Distance (walking and running)</li> <li>- Stages climbed</li> </ul> </li> </ul>
	<p><b>Internal resources</b></p> <p><b>Equipments</b></p> <ul style="list-style-type: none"> <li>• Connected watch</li> <li>• Smartphone</li> </ul> <p><b>Networks</b></p> <ul style="list-style-type: none"> <li>• Bluetooth</li> </ul> <p><b>Services</b></p> <ul style="list-style-type: none"> <li>• Physical activity reminder notification</li> <li>• set location</li> <li>• get health information</li> <li>• set health information</li> </ul>	<p><b>External resources</b></p> <p><b>Equipments</b></p> <p>Equipment participating in the session</p> <p><b>Networks</b></p> <ul style="list-style-type: none"> <li>• 3GPP (2G, 3G, 4G, 5G)</li> <li>• Internet</li> </ul> <p><b>Services</b></p> <ul style="list-style-type: none"> <li>• Weather</li> <li>• Geolocation                             <ul style="list-style-type: none"> <li>- get location</li> </ul> </li> <li>• Music</li> <li>• Personalized exercise programs                             <ul style="list-style-type: none"> <li>- Clear mode</li> <li>- French</li> </ul> </li> <li>• ...</li> </ul>
<p><b>Location</b></p> <ul style="list-style-type: none"> <li>• Location                             <ul style="list-style-type: none"> <li>Sports hall</li> </ul> </li> <li>• Address : 30 bd Vaugirard</li> <li>• Postcode : 75015</li> <li>• City : Paris</li> </ul>	<p><b>Agenda</b></p> <ul style="list-style-type: none"> <li>• Activity: Health</li> <li>• Task: Sports session</li> <li>• Location of activity: Sports hall</li> <li>• Start date and time: 17/08/2022 - 19:15</li> <li>• End date and time: 17/08/2022 - 20:15</li> </ul>	

Figure 23: Active profile at 7:15 p.m. on August 17<sup>th</sup>, 2022

#### 5.4.2.6 Smart ID for the first time slot

In accordance with the Smart ID Model, the information to be considered is that of the Active Profile (figure 21) with the enrichment of the Potential Profile (figure 19) compared to the health constants that will be consulted during the remote medical visit.

It is on this information that the data processing services of the Smart ID will operate.

The following set of services (not exhaustive) will be applied to the Active Profile of the first time slot (figure 21):

- 1) Verification of the instructions for the opening of the session concerning the remote medical examination
- 2) Opening of the session
- 3) Depending on the doctor's request, information on constants will be provided with possibly statistics on previous values of these constants
- 4) In the section of analyses, classifications and filters can be carried out at the request of the doctor
- 5) Etc.

The algorithms and datasets used will be detailed during the PoC.

#### 5.4.2.7 Smart ID for the second time slot

In accordance with the smart ID model, the information to be considered is that of the active profile (figure 23) with the enrichment of the potential profile (figure 19) compared to the monitoring of physical activities during sports sessions.

It is on this information that the data processing services of the smart ID will operate.

All of the following services (non-exhaustive) will be applied to the active profile of the second time slot (figure 21):

- 1) A request to the weather service can be made to choose the location of the physical activity (outside or in the gym)
- 2) Parameters reflecting physical behaviour are saved
- 3) Statistics on these same parameters are provided to recommend a personalized exercise schedule
- 4) Etc.

The algorithms and datasets used will be detailed during the PoC.

---

## 6 Data processing

### 6.1 Data categorization

Building a digital clone requires collecting and processing a large amount of heterogeneous data.

These data can be classified according to their main characteristics, which can be grouped according to following 4 dimensions.

- Temporality.

The data to be exploited have a long lifespan. Some may be volatile, others more stable. When they are changeable or scalable, they can be predictable.

- Source.

The origin of the data impacts the way in which it can be collected and processed. Some data is easy to obtain (open data, API) others need to be acquired (and may induce acquisition or usage costs) Still others can be inferred from other data already available. For many of them, the prior agreement of the user for their use is necessary, and the methods of their conservation fall within the framework of the regulations on the protection of data and freedoms.

- Quality.

Quality in the broad sense covers the subsidiary notions of reliability, accuracy, and sensitivity.

- Relevance.

The relevance depends on the field considered and the service to be done. The regulations govern the use of data and take into account the relevance of the collection and processing of data according to the service to be provided.

Some data is transverse and necessary to satisfy most use cases, others specific to the user objectives.

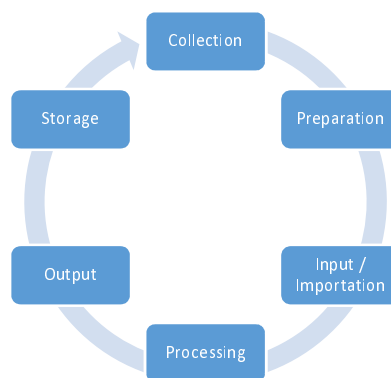
## 6.2 Data processing steps

With the emergence of new technologies and the proliferation of connected objects, the world around users has become increasingly connected, resulting in the production of large amounts of data. This data, whose volume is constantly increasing, is now everywhere and reflects human activities and behaviour.

It is therefore necessary to process it to extract information because, when grouped, compared and scrutinized, it can reveal previously unsuspected trends.

Artificial intelligence and machine learning (the study of computer algorithms that automatically improve with experience) will be used to extract meaning, determine better outcomes, and enable faster decision making from a variety of data sources.

Before the actual processing of the data, the raw data needs to first be collected and translated into usable information. The raw data is collected, prepared (filtered, sorted), processed, analysed, stored, and presented in a readable format.



**Figure 24: Data processing steps**

Figure 24 describes the following steps.

- Collection.

Data comes from all available sources (social networks, web, connected objects, sensors, etc.), including online data warehouses. It is important that the available data sources are reliable and structured so that the data imported (and subsequently used as information or knowledge) is of the highest possible quality.

- Preparation.

Data preparation, also known as 'pre-processing' or data cleaning, is a process of sorting and filtering raw data to remove unnecessary and inaccurate data. This is to ensure that only the highest quality data is fed into the processing unit.

The purpose of this step is to eliminate "bad" data (redundant, incomplete, or incorrect data) in order to start assembling high quality information so that it can be used in the best possible way to reach the purposes.

- Input or import.

Data is converted into a machine-readable format and fed into the processing unit. This may be data input from a keyboard, scanner, or other input source.

- Processing.

In this main process step, the data is subjected to various methods using Machine Learning and Artificial Intelligence algorithms to generate a desirable output. This step may vary slightly from one process to another depending on the source of the processed data (data lakes, online databases, connected objects, etc.) and the intended use of the output.

- Output

The data is finally transmitted and displayed to the user in a readable form (graphs, tables, vector files, audio, video, documents, etc.). This output can be stored and further processed in the next data processing cycle.

- Storage

The final step in the data processing cycle is storage, where data and metadata are kept for later use. This allows information to be quickly accessed and retrieved when needed, but also to be used directly in the next data processing cycle because in Smart Identity, data processing is done continuously with user-generated data.

## 7 New technologies for smart identity

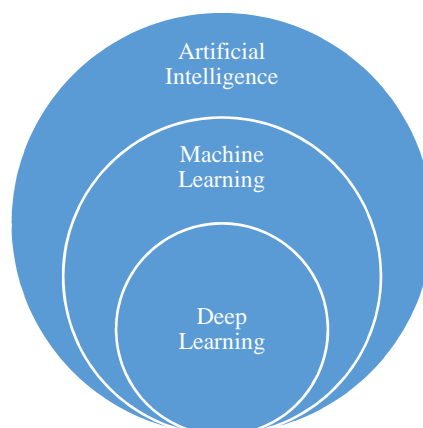
### 7.1 Overview

Several technologies are used to process Smart Identity. This clause focuses on the contribution of AI (clause 7.2), data security and data storage on the user side (clause 7.3) and the issue of data localization (clause 7.4).

### 7.2 Artificial Intelligence (AI)

Artificial intelligence is a computer science discipline closely linked to other sciences: mathematics, logic and statistics, which serve as its theoretical basis; human sciences (cognitive sciences, psychology, philosophy, linguistics, etc.) and neurosciences, which help to reproduce components of human intelligence by biomimicry; and finally, hardware technologies, which serve as the physical support for the execution of AI software.

It encompasses machine learning and deep learning.



**Figure 25: Artificial Intelligence**

Artificial Intelligence (figure 25) systems are based on data, which is the fundamental element that allows Machine Learning and Deep Learning algorithms to learn and improve. It is therefore at the heart of AI.

And one needs AI because it is far more efficient in terms of speed and accuracy than traditional data processing methods.

Given the huge amount of data produced by users daily, artificial intelligence and its techniques will be used to extract meaning, determine better outcomes, and enable faster decision making from various data sources.

## A. Machine Learning

Machine Learning is therefore the branch of artificial intelligence and computer science that focuses on using data and algorithms to enable the machine to mimic the way humans learn, gradually improving its accuracy.

For Smart Identity, the main Artificial Intelligence functions necessary for data processing are classification, regression, clustering, and predictive analysis. There are many Machine Learning algorithms that can be used to implement these functions, including:

a) KNN: K-Nearest Neighbour algorithm

Also known as the "k-nearest neighbour algorithm", KNN is a supervised machine learning algorithm generally used for classification and regression problems. Its principle is to calculate the distance between the point to be classified and the other points, and to take the k nearest neighbours according to this distance. The point will belong to the majority class.

(+) Easy to implement.

(-) The algorithm can become much slower as the number of observations and independent variables increases.

b) K-Means

Also known as the "k-means algorithm", K-Means is an unsupervised machine learning algorithm used for classification problems. It is used to divide N elements of the dataset into homogeneous groups/clusters.

c) Random Forest

Algorithm used for classification and regression problems. It allows to build decision trees on different subsets to form a forest of decision trees. The method generally used is bagging, which consists of training each decision tree formed with a part of the data and a part of the variables.

d) Naïve Bayes

Algorithm based on Bayes' theorem. The predictors are assumed to be independent ( $P(c|x) = P(x|c) P(c) / P(x)$ ). This is an efficient algorithm for learning problems on large datasets.

e) Principal Component Analysis (PCA)

Principal Component Analysis is an unsupervised machine learning algorithm that summarizes the information contained in a large dataset into a number of synthetic variables called "principal components".

The idea is then to be able to project these principal components onto the nearest hyperplane to have a simple representation of the data.

f) Logistic Regression

A statistical model often used for classification and predictive analysis. It allows the study of the relationships between a set of qualitative variables  $X_i$  and a qualitative variable  $Y$  to estimate the probability of an event occurring or not.

g) Gradient Boosting

A learning method for turning weak learners into strong learners. It is generally used for its accuracy and speed in handling complex and large data. The principle is as follows:

- Formation of decision trees
- Assigning an equal weight to each observation/predictor
- Increase the weight of each observation that is complicated to classify and decrease the weight of those for which classification is not a problem
- Idea: Make improvements to the predictions of the first tree. The second tree will grow on the weighted data.

The final model will be the combination of the trained trees.



## B. Deep Learning

To the extent that more specific processing is required in the PoC, one will also use Deep Learning, which is a subset of machine learning that does not require structured data and is based on layers of neural networks. These neural networks attempt to simulate the behaviour of the human brain, even though they are far from capable of doing so, allowing them to 'learn' from large amounts of data.

This approach is particularly suitable for complex tasks, where not all aspects of the objects to be processed can be categorized in advance. With deep learning, the system itself identifies the discriminating features of the data, without necessarily needing to be categorized in advance. The system does not need to be trained by a person. It assesses the need to modify the classification or create new categories based on the new data.

## 7.3 Data storage and security

Data storage security involves protecting storage resources and the data stored on them - both on-premises and in external data centres and the cloud - from accidental or deliberate damage or destruction and from unauthorized users and uses. It is an area that is of critical importance to enterprises because many data breaches are ultimately caused by a failure in data storage security.

Well-designed data storage security is also mandated by various compliance regulations such as PCI-DSS and the EU's General Data Protection Regulation (GDPR), thus adding legal weight to storage security demands. In general, good data storage security minimizes the risk of an organization suffering data theft, unauthorized disclosure of data, data tampering, accidental corruption, or destruction, and seeks to ensure accountability and authenticity of data as well as regulatory and legal compliance.

The relevant international standard for storage security is ISO/IEC 27040 [i.7], which calls for the application of physical, technical, and administrative controls to protect storage systems and infrastructure as well as the data stored within them. It notes that these controls may be preventive; detective; corrective; deterrent; recovery; or compensatory in nature. The bottom line is that ISO/IEC 27040 [i.7] describes best practices that ultimately set the minimum expectations for storage security.

At the highest level, data storage security standards, guidelines and regulation seek to ensure "CIA" - confidentiality, integrity, and availability:

- **Confidentiality:** Keeping data confidential by ensuring that it cannot be accessed either over a network or locally by unauthorized people is a key storage security principle for preventing data breaches.
- **Integrity:** Data integrity in the context of data storage security means ensuring that the data cannot be tampered with or changed.
- **Availability:** In the context of data storage security, availability means minimizing the risk that storage resources are destroyed or made inaccessible either deliberately - say during a DDoS attack - or accidentally, due to a natural disaster, power failure, or mechanical breakdown.

Identity and Access Management (IAM) refers to the collection of policies, processes and systems which support binding an individual (or in some cases a system) to a set of permissions within an organizations or businesses system. IAM is a part of overall process of ensuring data storage security.

These permissions may allow the individual to:

- Perform functions (such as altering and industrial control process).
- Access data (such as staff records).
- Administer the system.

An access management system is comprised of several technical components, include directory services, authentication components and the parts of the system that consume authentication and authorization information.

Identity and access management can be broken down into the following broad areas:

- **Policy** - the strategy governing who is authorized to access systems, data, or functionality, how they can request access, when their access should be revoked and whether any particular operations should require multiple users to collaborate.

- Identity management - how establish the identity of a person, both at point of first contact and subsequent interactions with systems or processes.
- Privileged user management - the additional processes and controls to put in place to safeguard the most sensitive operations in the system.
- Architectural design - secure design of the computer systems that support the above areas.
- Operations and monitoring - the supporting processes and technology to identify and enable investigation of breaches of policy or controls.

If identity and access management procedures and controls are badly designed or implemented, they can give attackers an easy way to gain access to an organizations system which could appear legitimate. The security of all the aspects of identity include:

- Ensuring that a new user is who they say they are, and the level of trust and access given them is commensurate with their personal and professional background.
- Binding an identified user to an identity within the system with an appropriate method of authentication.
- Ensuring that the authentication method give confidence that when an identity is used, it is being used by the member of staff whose identity that has been previously validated.
- Applying the principle of least privilege to limit the access or functionality that different users have.

The key thing to remember is that standards, guidelines, and regulations are designed to help ensure that security is effective. Attaining compliance does not mean that an organization is completely secure, but it is very rare that measures taken to ensure compliance would make an organization less secure than they otherwise would be.

## 7.4 Data localization (Cloud, Edge, Dew)

As seen previously, to make a digital clone it is necessary to have a knowledge base representative of the potential profile of the user. This raises the question of the location of the data for the constitution of this knowledge base.

The heterogeneous nature of these data (see clause 6.2) leads to the identification of data sources, their location, and their nature.

Driven by big data projects, the development of the IoT and new NGN type network architectures, especially mobile 5G, the cloud computing concept, that's to say the on-demand availability of computer system resources, data storage and computing power, now come in Edge and Dew computing. depending on the nature of the data to be collected and processed, some computing actions can be done locally, sometimes without connection, or remotely.

A multi-cloud approach is now possible and is necessary to make the digital clone in a fog of heterogeneous data. It is thus possible to improve cloud infrastructure capabilities, costs, security, resilience, energy efficiency, and continuity of service.

---

## 8 Conclusion and next steps

The present document describes the digital clone with all the necessary elements for modellization and implementation of the Smart ID. The intelligence of this Smart ID is achieved by AI new technologies.

These elements are described in a theoretical framework needed to implement a Proof of Concept (PoC) to apply on a specific use case, as described in clauses 4.6.3 and 5.3.2 (health management).

The next step is to provide, for the chosen use case, the Smart ID. The presentation of this Smart ID is expected to be provided in ETSI TR 103 875-2 [i.14], the Proof of Concept of the Smart ID, not published yet at the time of publication of the present document.

The final step is to describe and implement the Smart Interface based on Smart ID for a smart interaction between the user and the digital ecosystem.

---

## History

<b>Document history</b>		
V1.1.1	November 2022	Publication