Horizon Project IMPORTANT

IMPlementing geriatric assessment for dose Optimization of CDK 4/6-inhibitors in older bReasT cAncer patieNTs

Research and Innovation Action
HORIZON-MISS-2020-CANCER-01
GA 101104589

Duration: 60 months
Start date: 01/05/2023

Deliverable ID.:	D2.2	
Deliverable title:	IMPORTANT's technical requirements and specifications	
Planned delivery date:	30/10/2023	
Actual delivery date:	30/10/2023	
Deliverable leader:	SLC	
Contributing partners:	ORB, ORU, Care	
Dissemination Level:	X	PU = Public
		CO = Confidential
		CI = Classified

This project has received funding from the European Union's Horizon Europe research and innovation program under grant agreement No 101104589.

Project Partners

Num.	Partner Name	Short Name	Country
1 (C)	OREBRO LANS LANDSTING	ORB	SE
2	PANEPISTIMI PATRON	UPAT	GR
3	CARE ACROSS LTD	CARE	UK
4	UNIVERSITA DEGLI STUDI DI FIRENZE	UNIFI	IT
5	AZIENDAUNITA SANITARIA LOCALE TOSCANA CENTRO	LHUTC	IT
6	HELSINGIN JA UUDENMAAN SAIRAANHOITOPIIRIN	HUS	FI
7	INSTITUTE FOR MEDICAL TECHNOLOGY ASSESSMENT BV	IMTA	NL
8	SECURITY LABS CONSULTING LIMITED	SLC	IE
9	Circular Economy Foundation	CEF	BE
10	UNIVERSIDAD NACIONAL DE EDUCACION A DISTANCIA	UNED	ES
11	Elliniki Sinergazomeni Ogkologiki Omada	HECOG	GR
12	AKERSHUS UNIVERSITETSSYKEHUS HF	AHUS	NO
13	REGION UPPSALA	RUL	SE
14	FUNDACIO CLINIC PER A LA RECERCA BIOMEDICA	FCRB	ES
15	HOSPITAL CLINIC DE BARCELONA	HCB	ES
16	FACHHOCHSCHULE NORDWESTSCHWEIZ	FHNW	CH
17	Phaze Clinical Research & Pharma Consulting S.A.	PHAZE	GR
18	Bröstcancerförbundet	BCF	SE
19	EUNOMIA LIMITED	EUNL	IE
20	Örebro university	ORU	SE

Project Coordinator: Dr Antonis Valachis - Örebro Läns Landsting (ORB)

List of Authors

Name(s)	Partner
Thanos Karantjias	SLC
Theofanis Fotis	SLC
Antonis Valachis	ORB
Helena Hansson-Nylund	ORU
Mukhrizah Othman	ORB
Thanos Kosmidis	Care

Revision History

Date	Ver.	Author(s)	Summary of main changes
17/7/2023	0.1	Karantjias Thanos, Fotis Theofanis	Table of Contents
25/8/2023	0.2	Karantjias Thanos	Input in chapter 2
15/9/2023	0.3	Theofanis Fotis	Input in chapter 3
29/9/2023	0.4	Karantjias Thanos	Input in chapter 3 and 4
07/10/2023	0.5	Karantjias Thanos	Input in chapter 3 and 4
16/10/2023	0.6	Thanos Kosmidis, Aglaia Schiza, Helena Hansson-Nylund	Internal Review Process
23/10/2023	0.7	Karantjias Thanos, Fotis, Theofanis	Address Internal Review Comments
25/10/2023	0.8	Karantjias Thanos, Fotis Theofanis	Pre-final Version
29/10/2023	0.9	Thanos Kosmidis	Addition of DCT content to chapter 1
30/10/2023	1.0	Karantjias Thanos, Fotis Theofanis	Final Version

Table of Contents

List of Figures

List of Tables

Abbreviations

CRO	Contract Research Organization
DCT	Decentralized Clinical Trial
EC	European Commission
eCRF	Electronic Case Report Form
GCP	Good Clinical Practice
IEC	Independent Ethics Committees
IRB	Institutional Review Board
IoT	Internet of Things
MoSCoW	Must have, Should have, Could have, Won't have
RBAC	Role-Based Access Control
WP	Work-package

Executive Summary

This deliverable is the main outcome of T2.2 “IMPORTANT technical requirements for a decentralized clinical trial”, which focus on defining the functional and non-functional requirements for a DCT from a technical perspective. Specifically, it introduces the Decentralized Clinical Trials concept and it identifies IMPORTANT clinical sites technical specificities like systems, assets and procedures in place and addresses any limitation that may affect actions of WP3 and WP4.

To this end, D2.2 introduces the IMPORTANT services specifications, the main conceptual architectural layers, and phases, and derives the technical requirements for setting up a decentralized clinical trial for older people with breast cancer (e.g., age-appropriate user interfaces, interactions and data flows for patients, caregivers, clinicians, remote patient monitoring; telehealth platforms, web interfaces etc.).

In order to succeed, it takes, among other, the following steps:

- It identifies the main IMPORTANT stakeholders that will be invited to use the IMPORTANT solution.
- It adopts a proper methodology for describing and reporting the technical specifications (VOLERE), and the methodology to be followed to prioritize these specifications (MoSCoW), enabling authors to better describe the findings and outcomes of the designing phase in an organized and systematic manner.
- It identifies the main IMPORTANT phases (Data Capture Phase, Data Analysis & Cleansing Phase, Publishing Phase), structuring in parallel the core elements of the IMPORTANT architecture.
- For each IMPORTANT component, it introduces the concrete list of services, their interfaces required for implementation purposes, how these services are grouped and how these are mapped with the functionalities identified at the presentation of the IMPORTANT solution elements.
- Finally, it builds the mapping of identified services with the functional and non-functional technical requirements.

1 Introduction

1.1 Scope

The main scope of this deliverable is to define the functional and non-functional requirements for a Decentralized Clinical Trial (DCT) from a technical perspective. It reports on the finding of the task 2.2 “IMPORTANT technical requirements for decentralized clinical trial” which focuses on the identification of the IMPORTANT clinical sites’ technical specificities. These can be either systems, tools, assets, and/or procedures which are established.

The identification and reporting of technical specifications is based on a world-wide accepted and known methodology, the VOLERE methodology [1], which is introduced in section 2 of the document. The technical requirements and specifications are ranked and classified based on the initial stakeholders’ needs and the expertise of the partners based on the methodology MoSCoW [3], which is also introduced in section 2.

1.2 Contribution to other work packages

All identified requirements and technical specifications, among other, address any limitations that may affect actions of WP3 and WP4. These will be evaluated and further enhanced in WP4 and WP5.

1.3 Structure of the document

This document is structured in five main sections:

- **Section one** contains an introduction to the technical requirements and specifications and the deliverable structure.
- **Section two** introduces the Decentralized Clinical Trials concept.
- **Section three** introduces the main stakeholders of the IMPORTANT solution and defines the methodology used for analysing the specifications.
- **Section four** introduces the conceptual architecture of the IMPORTANT solution which consists of three different phases, the data capturing, the data analysis and cleansing, and the results publishing.
- **Section five** introduces the technical requirements and specifications identified for all IMPORTANT solution phases.
- **Section six** summarizes and draws conclusions.

2 Decentralized Clinical Trials (DCTs)

2.1 Context for IMPORTANT

The concept of Decentralized Clinical Trials (DCTs) has been evolving for several years and is one of the main innovations within IMPORTANT. This section aims to provide background information and context for DCTs, discuss their main characteristics, explore benefits and challenges, and outline key ethical considerations.

2.1.1 High-level description

The European Medicines Agency describes Decentralized Clinical Trials as follows: “Traditionally, clinical trials have been conducted at specific clinical trial sites, to which patients had to travel to. The aim of DCTs is to make it easier for patients to participate in clinical trials by reducing the need to travel to central trial sites. This approach has the potential to make clinical trials available to a wider demographic of participants and reduce drop-out rates” [4].

The FDA, in its draft guidance, describes a DCT as “a clinical trial where some or all of the trial-related activities occur at locations other than traditional clinical trial sites” [5].

2.1.2 History of DCTs

It is not easy to pinpoint exactly when a clinical trial was first conducted in a “Decentralized” manner. A feasibility study conducted in 2003 on a trial in patients with knee osteoarthritis [6]. A few small studies were conducted in the subsequent years, driven by academic and commercial interests. There was even an patent awarded to Boston University in 2007, which was originally filed in 2000, with the same key owner as the author of the subsequent, aforementioned, feasibility study [7].

It is generally considered that the first fully remote study was conducted in 2011 [8].

The evolution of DCTs continued, and a key inflection point was formed by the forces of COVID-19 in 2020 and later, which made hospital visits extremely complicated, and forced several traditional trials to shift towards the DCT model and include several decentralized elements.

2.1.3 Defining DCTs

The descriptions and definitions mentioned above are relatively vague, and this is not unintentional in a very diverse and broad ecosystem of clinical research. Below we attempt to capture the main elements of DCTs.

In a DCT, some parts of the trial are conducted outside of the strictly defined clinical physical settings (e.g., a hospital).

This often starts with the patient information and education about the trial; more importantly, it can encompass the entire recruitment and enrollment journey and experience, including the Informed Consent of the patient towards the trial’s protocol details and requirements.

Similarly, and beyond enrollment, a very common component of these trials' decentralized parts includes patients' travel: although it is not always possible to receive medical treatment (if this is part of the trial) away from a hospital, there are often other engagement paradigms which, in a DCT, are shifted elsewhere. For example, patient consent, data collection, monitoring of vital signs, and other activities, can be conducted from a distance, and often from the patient's home.

The data collection can take place in several ways, including:

- Traditional phone calls
- Smartphones and applications
- Wearable devices
- Internet-of-Things (IoT) devices.

Similarly, patient visits which do not require an administration of a medication or treatment (which would not be possible elsewhere) can often be conducted via telehealth.

At the same time, if a physical appointment is necessary, it can be done with a local healthcare professional via home visit or otherwise.

In the case of treatments that are normally administered in a clinical setting, they, too, are often redesigned to be administered elsewhere. For example, oral medications can be shipped in a controlled manner to a patient's home, or can even be distributed locally to the patient's home and, in collaboration with local healthcare professionals, administered outside of the main clinical site.

This is why Decentralized Clinical Trials (DCTs) are often called remote, siteless, virtual, or even digital – to capture the modality of interaction. In the majority of cases, though, they are "hybrid" in reality, because they encompass a combination of "traditional" and "decentralized" modalities.

2.1.4 Benefits of DCTs

The benefits of Decentralized Clinical Trials (DCTs) span all stakeholders; this section outlines the key ones.

2.1.4.1 Benefits for the Patient

The patient gets to stay at home longer and may require fewer interactions with the bureaucratic aspects of the healthcare system. Moreover, the framework of a DCT can allow more flexibility as it depends less on the physical colocation and synchronous exchange of information. This also means fewer costs for the patients, and less time spent commuting. For many patients, a DCT means more convenience and improved ability to continue living as normal despite having been diagnosed with a serious health condition.

2.1.4.2 Benefits for the Sponsor

Significant part of the complexity of the clinical trials is absorbed and streamlined in a DCT setting; from information, enrollment, engagement to overall communications, the sponsor can focus more on the actual study objectives. In some cases, a DCT can have a more favorable

cost structure as well. Moreover, the additional data points can be quite valuable when further analyses are required (always in alignment with the Protocol and Informed Consent).

2.1.4.3 Benefits for the Clinical Research Organization (CRO)

The CRO of a well-designed DCT can benefit from better structured data collection and more aligned data processes, given the flexibility of digitalized procedures and methodologies. Although this may come at a "cost" of reduced visibility into some of the patient-centric data input processes, the corresponding technologies and tools often provide several options as alternatives.

2.1.4.4 Benefits for the Clinical Trial Site

A clinic is a complicated operation. Many of its activities are beyond the medical or clinical scope, and include data capture and exchange, as well as procedural components that are not necessarily executed at the most appropriate time and place. A DCT can help the Sites improve their efficiencies and reduce the duplication of processes, while capturing more specific and structured data that can also help them deliver internal operational improvements.

2.1.5 DCTs: Challenges and barriers

Despite the several benefits, some of which have been outlined here, many challenges remain. The below bullet points identify a few generic ones, although additional ones can emerge depending on the specifics of each DCT.

- Stakeholders (including clinics, clinicians, CROs and patients) should be familiar with some technical tools and have access to sufficient and appropriate technologies and network resources in order to be able to participate in their own capacity.
- The design of a DCT requires significant focus on the technologies and tools, and experience with their integration in a clinical setting.
- In healthcare, clinical examination is the gold standard for many processes, and physical proximity is required.
- Clinical Trials are very lengthy and complex processes; this means that their evolution and maturity into DCTs is prolonged further, delaying their institutionalized adoption.
- In hybrid DCTs (which constitute the majority), the transition between physical and remote/virtual settings is continuous and often very complicated.
- The software and technological elements involved in DCTs may invertedly become influencers of its design, endpoints, and outcome.
- Managing more data from various sources can be a burden for all stakeholders, and the impact and importance of each data set or point is often under scrutiny (EU Recommendations on Decentralised Elements in Clinical Trials, PCWP/HCPWP joint meeting, 28 June 2023, EMA Amsterdam)

2.1.6 Ethical considerations for DCTs

In general, clinical trials overall implicate substantial ethical considerations. Their Decentralized versions have even further complications to consider. The growing maturity of DCTs has led some researchers to analyse towards initial analysis of considerations [9].

Starting with the recruitment and enrolment paths, is the information patients receive sufficient, if it is equivalent to that of a “traditional” clinical trial? Are they expected to be familiar with the technologies? Is the e-Consent variety of Informed Consent equivalent and always sufficient?

Furthermore, a fundamental ethical consideration is related to the access to, and familiarization with, these technologies: is there a bias introduced when any element of a trial is shifted to a Decentralized version? How does this affect the applicability of the trial, and the generalizability of its results?

A key set of consideration arises from the increasing breadth and variety of data collection modalities, especially when it is originated from the patient (for instance, directly via ePROMs, or indirectly via devices): what is the significance of these datapoints, and what impact can they have on the patient’s participation, health-related decisions, and trial outcomes? Also, how does the submission of ePROM datasets affect adverse event reporting – for instance, are these equivalent to clinician-reported for pharmacovigilance and other purposes?

Furthermore, what are the Data Protection and Privacy preservation requirements when it comes to DCTs? Since clinicians have unique access to the patient records and have a duty and obligation to include them in their decision making, how does the GDPR affect these processes, especially when it comes to non-clinical level data? What are the legal and security implications? How do the technical details of the implementation affect the trial design and its impact (e.g. around cloud computing, specific device types, data input requirements etc)?

Another set of concerns is related to the relationship between the patients and their healthcare professionals: does the natural distance between them, combined with the use of technology, alienate them, and thus deviate from the expected and necessary collaborative elements? Are asynchronous and virtual tools sufficient?

Similar to the above, what are the expectations of patients when they report anything directly or indirectly? Should clinicians be made aware of these reports, and if so, should their awareness be immediate and require prompt action? How can the proper prioritization and triaging take place?

Clearly, therefore, Decentralized Clinical Trials are not without their own, unique ethical complications. As more institutions become part of the conversation (including the European Commission and the European Medicines Association) there will be more guidelines and improved alignment for the benefit of patients and all other stakeholders.

3 Methodology

This section provides an overview of the overall process followed for the elicitation and analysis of the IMPORTANT technical requirements and specifications.

3.1 IMPORTANT Stakeholders

To successfully identify the IMPORTANT services and their characteristics, the project had at first to try to identify its different fictional user-types and characters. In order to succeed on this, we performed the following:

- First, we tried to understand their needs, experiences, behaviours, and goals, recognising that different people have different needs and expectations.
- Secondly, we adopted the role-based perspective, which is goal-directed, and it focuses on the user's behaviour. The different user types of the role-based perspectives are massively data-driven and incorporate data from both qualitative and quantitative sources.

Therefore, within IMPORTANT we identify four different main user-types:

- Research team, who represent the users in collecting and inputting patients' data into the eCRF (Clinical) which comprises of treating physician and study nurse at each clinical site. Research associate / research assistants who will perform an initial data cleaning of data captured at each site. Research associate and statistician who will perform data cleaning and analyses within the premises of the research repository at sponsor's site
- Patient, who represents the users in providing data for the project and engaging with the published results that are shared on social media platforms or from information provided by physician or investigator at each study site
- Sponsor, who represents the users in performing data cleansing and analyses within the premises of the research repository on-site. Sponsor and delegated personnel will have access to pseudonymized data kept securely at the sponsor's site in a research repository. Sponsor will also monitor the compliance of all sites with the protocol and GCP
- Contract Research Organization (CRO), who represents the users in reviewing and monitoring data collected for the project

3.2 Technical Requirements

Project requirements outline the minimum benchmark that the final output must satisfy to resolve the end user's needs and preferences. Such requirements could emerge from multiple sources (i.e., regulatory guidelines, ongoing domain specific trends, technical trends, etc.). Essentially, it contains the set features and functions necessary to meet the goal of the project.

It should be noted that this document focuses on the identification of technical requirements of any type and category. Specific requirements related to security and privacy are also reported and analysed in D2.4.

3.2.1 Requirements Process

The project adapts a Systems Engineering approach to assist in ensuring the IMPORTANT solution is suited to the practitioner-stakeholders for whom it is intended. The philosophy behind the approach and the role of the stakeholders' requirements is illustrated in the following figure.

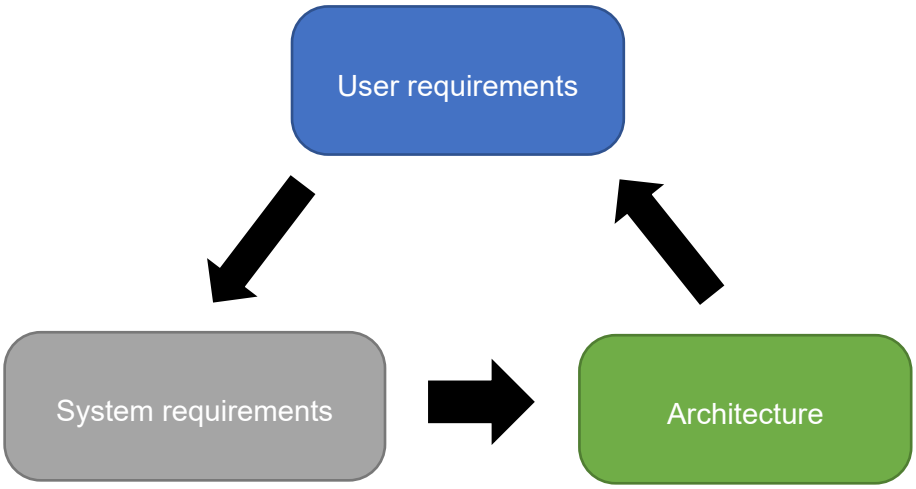

Figure 1. Systems Engineering development-cycle

IMPORTANT project, from its proposal writing phase, already acknowledges a high-level understanding of what counts as "user requirements". This allows the project to narrow down the scope and meaning of technical requirements, the identification process of which becomes easier. Technical requirements cover the technical issues that must be overcome for the successful development and operation of the final solution.

The process of collecting and documenting the technical requirements is quite straightforward and involves the following steps:

- **Gather inputs from all sources**: User (and/or business) requirements could arise from any touchpoint that interacts with the proposed solution. This list includes stakeholders, project managers, developers, and any type of end-user. Thus, the first course of action would be to collect feedback on what are the minimum expectations.
- **Perform Usage Analysis**: Upon identifying the basic technical requirements, it is time to profile the end-user and analyse their behaviour patterns in order to determine the level of performance the solution will have to deliver.
- **Create and validate a proper document**: Now that the foundational work is complete, it is all about compiling it into a detailed technical requirements document. Ideally and among others, it should contain:

- The end-user needs and expectations, along with the usability and applicability of the product in the practical world.
- Team structure detailing the skills and competencies required, along with contingency plans.
- Prototypes and the corresponding outcomes that users can anticipate upon the delivery of the solution.

Obviously, the proposed solution architecture will be based on the previous requirements defined. In the light of design experience, it is usual for the implementers to extract new features and functionalities from the user requirements and constrains.

However, implementers need also to balance the importance of the business needs with the real system requirements. This feedback loop may not be required or may be repeated more than once.

3.2.2 Prioritization of requirements

When designing and implementing the functionality of a system, it is important to prioritize the requirements focusing on the first development of the essential parts and remove the less significant ones if necessary due to the lack of time or resources.

In IMPORTANT, the requirements are ranked based on the initial stakeholders' needs and the expertise of the partners. It is necessary to prioritize what is essential for the operation of the product for the development. The prioritize methodology used as a reference to classify the requirements is MoSCoW [2][3].

MoSCoW was developed by Dai Clegg of Oracle UK in 1994 [10] and it gained popularity in the DSDM methodology (Dynamic Software Development Method). The MoSCoW method is a prioritization technique used in management, business analysis, project management, and software development to reach a common understanding with stakeholders on the importance they place on the delivery of each requirement - also known as MoSCoW prioritization or MoSCoW analysis, which is a fairly simple way to sort features into priority order – a way to help teams quickly understand from the customer's view what is essential for launching a product and what is not.

The term MoSCoW itself is an acronym derived from the first letter of each of four prioritization categories (**Must** have, **Should** have, **Could** have, and **Won't** have).

Figure 2. MoSCoW priority

The categories are typically understood as:

- **Must have**

  Requirements labelled as MUST have to be included in the current delivery time-box in order for it to be a success. If even one MUST requirement is not included, the project delivery should be considered a failure. It is good to have clarity on this before a project begins, as this is the minimum scope for the product to be useful. MUST can also be considered an acronym for the Minimum Usable SubseT.

- **Should have**

  SHOULD have requirements are also critical to the success of the project, but are not necessary for delivery in the current delivery time-box. SHOULD requirements are as important as MUST, although SHOULD requirements are often not as time-critical or there may be another way to satisfy the requirement so that it can be held back until a future phase. Therefore, it could be considered SHOULD are features that are not critical to launch, but are considered to be important and of a high value to the user.

- **Could have**

  Requirements labelled as COULD are desirable but not necessary, which could improve user experience or customer satisfaction for little development cost. These will typically be included if time and resources permit.

- **Won't have**

  Requirements labelled as WON'T have been agreed by stakeholders as the least-critical, lowest-payback items, or not appropriate at that time. As a result, WON'T requirements are not planned into the development schedule for the delivery time-box. WON'T requirements are either dropped or reconsidered for inclusion in later phases or projects. This, however, doesn't make them any less important. Alternately described as "Would like to have" in the future.

All requirements are important, but they are prioritized to deliver the greatest and most immediate business benefits early. Developers will initially try to deliver all the Must have,

Should have and Could have requirements but the Should and Could requirements will be the first to be removed if the delivery timescale looks threatened.

Thus, this ranking helps everyone (stakeholders, project manager, designer, developers) understand the most important requirements, in what order to develop them, and what not to deliver if there is pressure on resources.

3.3 Technical Specifications

IMPORTANT specifications follow the VOLERE requirements process [1], adapted to meet the project research activities and allow agile refinement. The VOLERE requirements specification process ("volere" is the Italian word for "to wish" or "to want") was developed to answer the need for a common language for discovering requirements and connecting them to solutions. The language needs to be understandable by business-people, customers, business analysts, engineers, designers, suppliers, testers or anyone else whose input is needed.

All these people have different skills and, not surprisingly, different views on what is important. A language intended for all these people must recognise the differences in peoples' viewpoints and yet have a consistent way of communicating and tracing the relevant knowledge.

Since the introduction of the first version in 1995, the VOLERE requirements techniques have been used on projects in a wide variety of domains such as banking, air traffic control, retail, aviation, government, real-time control, business analysis, manufacturing, just to name a few. The seemingly contradictory characteristics of rigour and flexibility have made the techniques popular as an aid to discovering, understanding, writing, and communicating requirements.

There are also management advantages like consistent input to estimating, risk management, monitoring, benefit analysis and the basis for reuse.

VOLERE makes use of established principles, models and practices and builds traceable connections between them. These connections provide a common thread between business or domain requirements, systems analysis models/deliverables, design models / components / deliverables, code, and testing. The resulting framework is one that can be used regardless of modelling notation, methodology, degree of agility, development lifecycle or tool usage.

There are three groups of interconnected components:

- The ***Requirements Knowledge Structure*** is concerned with how different items of requirements knowledge relate to each other and how to trace requirements from one level to another.
- The ***Requirements Process*** is concerned with procedures and activities for how to discover, populate and disseminate the requirements for knowledge.
- The ***Requirements Stakeholders*** are the input for determining how much of the knowledge structure needs to be populated for each project. The project team uses knowledge of the stakeholders to determine the order in which the work needs to be done and the appropriate level of detail.

The thinking requirements practitioner works to achieve the appropriate balance between requirements knowledge, process and stakeholders. The purpose of VOLERE is to provide structure and guidance in achieving the appropriate balance for each project. The detailed contents of each one of these component groups are as follows:

- Requirements Knowledge Structure
    - a requirements knowledge model that acts as a filing system for requirements knowledge
    - a requirements template
    - atomic requirements structure
    - requirements traceability
    - levels of requirements
- Requirements Process
    - a generic requirements process for discovering requirements knowledge
    - how to get started depending on your project characteristics
    - requirements trawling/discovery techniques description and guidance
    - quality assessment techniques
    - goal analysis techniques
    - requirements estimation
    - requirements specification audit techniques
    - prioritisation techniques
- Requirements Stakeholders
    - stakeholder analysis techniques
    - role, knowledge, person analysis
    - roles, responsibilities and commitment
    - stakeholder feedback techniques
    - conflict resolution techniques
    - requirements viewpoints

One key-methodology technique is to follow the VOLERE Requirements Specification Template, which is intended for use as a basis for requirements specifications. It provides sections for each of the requirement types appropriate to today's software systems.

The VOLERE template is a compartmentalized container for requirements and it categorizes the requirements into types that prove useful for the purpose of recognition and elicitation. These types are:

- Functional requirements: describe what the product has to do or which processing actions it must take
- Non-functional requirements: properties that the functions must have, such as performance and usability
- Project constraints: restrictions on the product due to the budget or the time available to build the product
- Design constraints: restrictions on how the product must be designed
- Project drivers: business-related forces
- Project issues: conditions under which the project will be carried out

Whereas the template is a guide to what to write about, the requirements shell or snow card, is a guide to how to write it. Individual requirements have a structure – a set of attributes, where each attribute contributes something to the understanding of the requirement, and to the precision of the requirement, and thereby to the accuracy of the product's development.

Based on this adapted methodology, each IMPORTANT requirement shell specification is identified as follows:

Table 1: IMPORTANT Service Specification Template

Service Contract Name	**<Name>**
Specification ID	Unique identifier with the following taxonomy: <COMPONENT_ID>-<TYPE>-<NUMBER> For example: FoU-F-010 refers to the FoU component, is a functional requirement (F) and has a ten based number for initial value. Once attributed an identifier cannot be changed or reused to represent a different concept.
Specification Type	Functional (F) Security (S) Privacy (P) Look and Feel (LF) Usability (U) Performance (PF) Operational and Environmental (O) Maintainability and Support (M)

Category / Grouping	Main category of the requirement
Dependencies	Dependencies that the requirement have from other ones
Stakeholders	Based on the identified Stakeholders (in paragraph 2.1)
Priority	(MoSCoW values) considering the corresponding the classification of the specification
Description and Rationale	Brief Description of the module / service / component
Tool(s) / Solution(s)	Project tool / solution that shall / can be used

Towards this, for each IMPORTANT service, its interface operation specification is identified as follows:

Table 2: Service Interface Operation Specification

Operation Name	*Name*
Operation ID	<Specification ID> . <Operation ID>……
Description	(Provide a description for the operation – not more than one small paragraph) i.e., this operation will support the… For all these, it will be mandatory to provide…
Input	(Describe if any)
Output	(Describe if any)

4 Conceptual Architecture & Usage Flow

The conceptual architecture of the IMPORTANT solution depicts the three different main phases and process that the project recognizes, as depicted in the following figure. These phases have as follows:

- Data capture phase
- Data analysis and cleansing phase
- Publishing phase

In the data capture phase, the gathering process is executed for the research team, the sponsors and the CROs to be able to prepare and maintain adequate and accurate source documents and trial records of the patients.

These records are properly and securely exported to feed the data analysis and cleansing phase, in which an initial data cleansing is performed. Moreover, in this phase data is carefully aggregated and further analysis is performed from experts. Obviously, cleansing and analysis may run multiple times in a secure and protected environment / lab, till all participating stakeholders reach and produce the “expected” outcomes and results.

Upon analysis and cleansing completion, the produced results are ready for publishing. Therefore, the results publishing phase runs getting feedback from the data analysis and cleansing one. In this phase provision and dissemination of the study results, which consist of aggregated results after suitable statistical analysis, to many different sources and channels takes place.

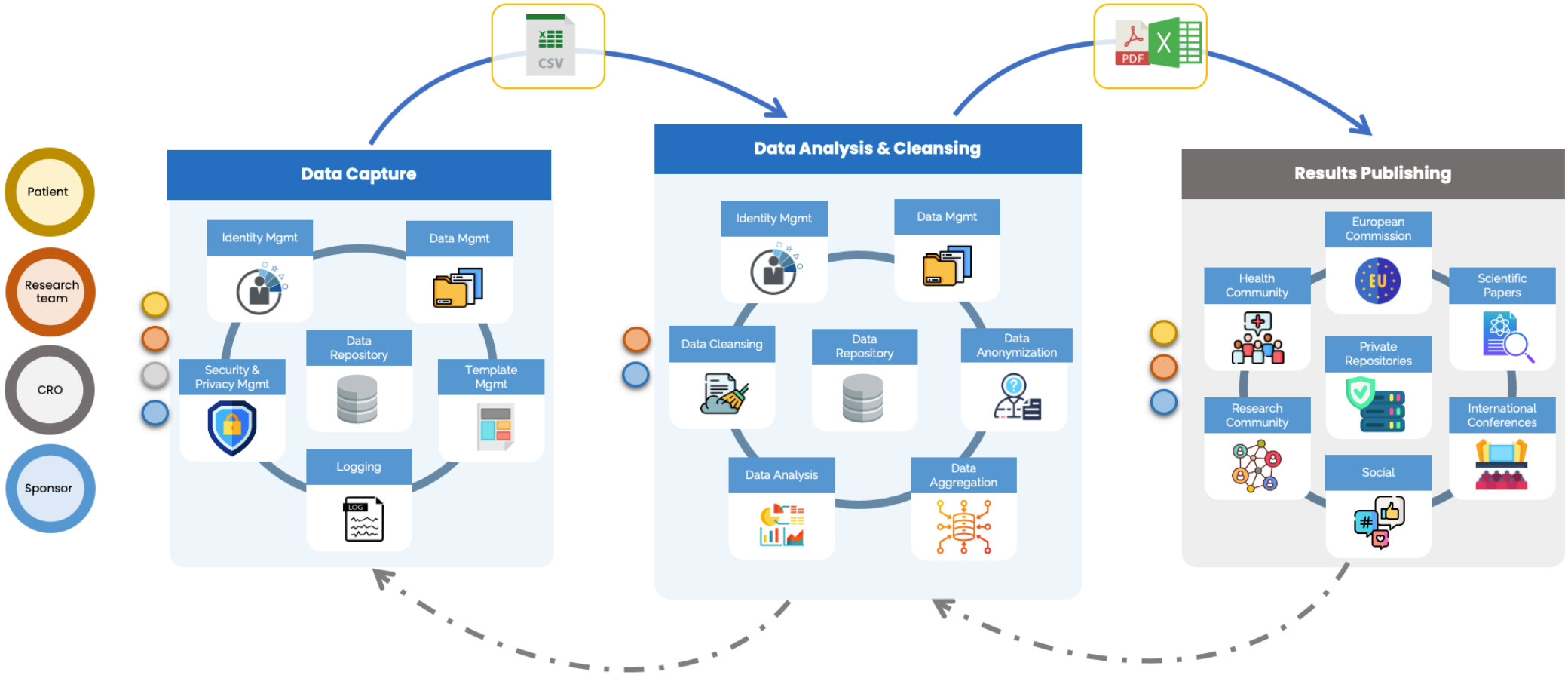

Figure 3: IMPORTANT main phases and processes

The following paragraphs analyse in more detail these three phases.

4.1 Data Capture Phase

In data capture phase the following stakeholders participate:

- The research team at each clinical site
- The patients,
- The sponsors, and
- The CROs

In this phase the research team will undertake to capture data from patient in the electronic Case Report Form (eCRF). This is achieved with the creation and publication of specific questionnaires in which the patients will be invited to complete. Clinical enforces strict security mechanisms for allowing access to their systems. Specifically, two-factor authentication is utilized for patients to gain access to the questionnaires.

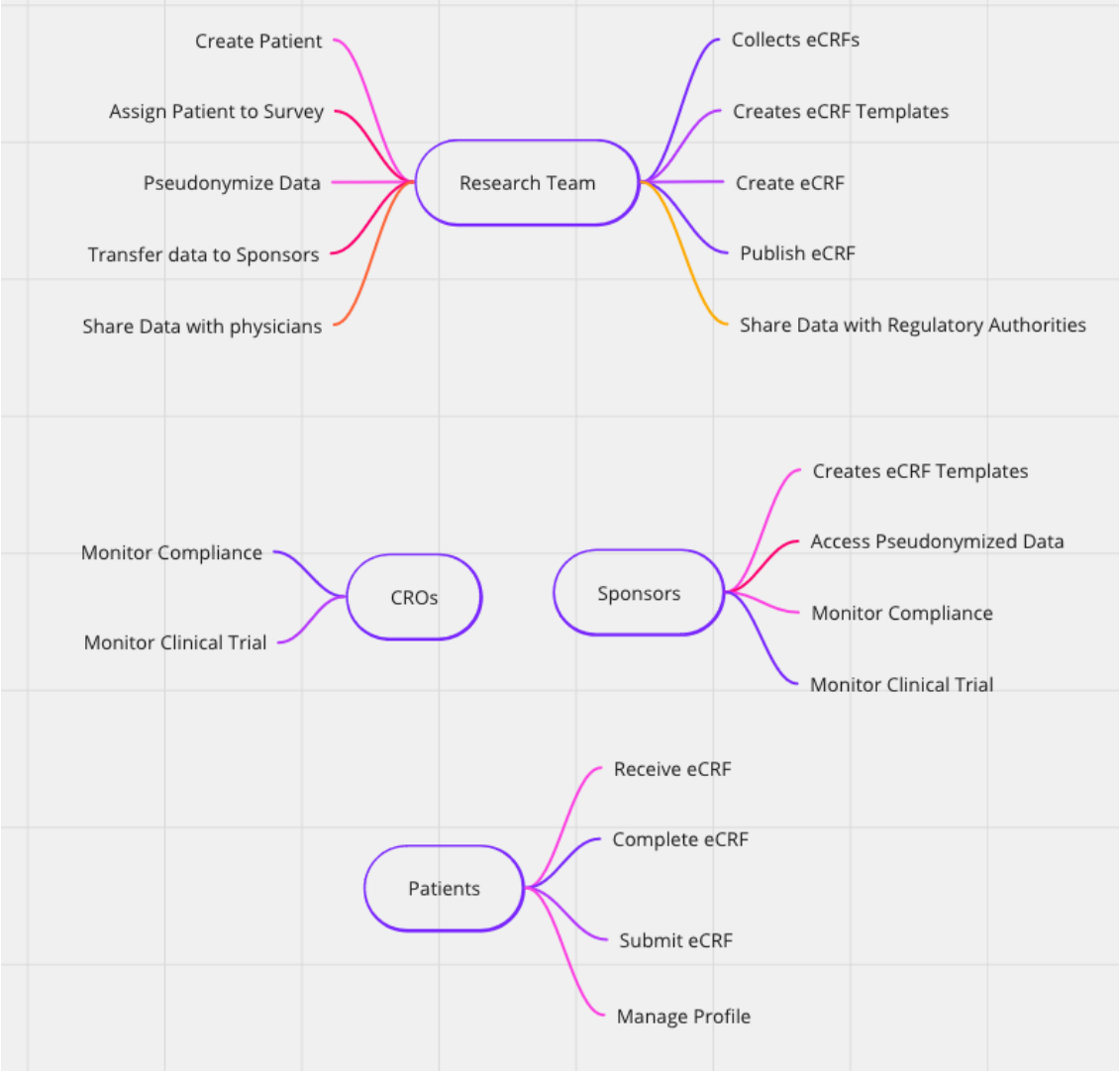

Figure 4: Actions and Responsibilities of Stakeholders in Data Capture phase

All activities / events are properly and in real time logged in 'Audit Logs'. Each enrolled patient who participates in the process is assigned with a unique pseudonymised ID. To this end, patient's identification is only visible to each site's eCRF.

The next step is for the research team to prepare and maintain adequate and accurate source documents and trial records, which include all observations and other data pertinent to the investigation on each trial participant. At this step the data captured will be managed and kept securely at each site. Considering this, the research team will allow site trial-related monitoring, audits, IRB/IEC review and regulatory inspections. Direct access for review will be provided to monitor, auditor and regulatory inspector (e.g., EMA, FDA) for the CRF and all source documents / data. These include progress notes, copies of laboratory and medical test results.

To ensure confidentiality of records and personal data, only pseudonymized data will be transferred to the sponsor. This will be realized with the proper use of the participant's identification number instead of the participant's name. The code is only available at the site

and must not be forwarded to the sponsor/CRO. In case that participants' records will be forwarded (e.g., for SAE processing), personal data that can identify the participant will be redacted by the site prior to forwarding.

Access to the participant files and clinical data will be again strictly limited; personalized treatment data may be given to the participant's personal physician or to other appropriate medical personnel responsible for the participant's welfare. Data generated at the site because of the trial need to be available for inspection on request by the participating physicians, the sponsor's representatives, by the IRB/IEC, and the regulatory authorities. Research team needs to make sure that no records will be transferred to another location or party without written notification to the sponsor/CRO.

In this data capture phase, sponsors are responsible to provide template management for the eCRF in Clinical. Specifically:

- Each sponsor will gain access to pseudonymised data at all eCRF sites.
- They will undertake the responsibility to monitor compliance of all sites with the protocol and GCP.
- In collaboration with the CRO they will develop a systematic, prioritized, risk-based approach to monitoring of this clinical trial.

Finally, CROs will develop a systematic, prioritized, risk-based approach to monitoring of this clinical trial in collaboration with the sponsor. CRO will also monitor compliance with the protocol and GCP.

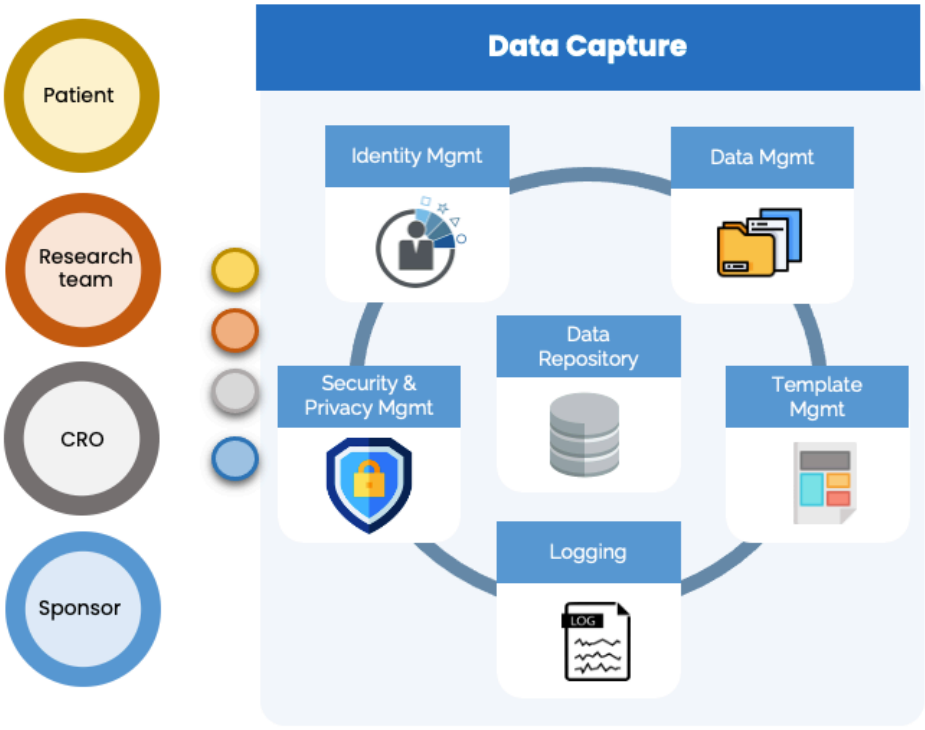

Figure 5: Data Capture phase

From the technical point of view the data capture phase could be realized as a generic module of the IMPORTANT solution, which consists of the following main entities:

- The Identity Management component, which undertakes the responsibility to collect and organize the data, implementing methods for gathering and acquiring data from different sources and in parallel structuring and categorizing data in a meaningful way for easy retrieval, analysis, and utilization. It is the component that undertakes the responsibility to enforce strong authentication and authorization mechanisms such as two (2) factor authentication and role-based privileges across the components of this module.
- The Data Management component, which undertakes the responsibility to collect and organize the data, implementing methods for gathering and acquiring data from different sources. This component allows the building of comprehensive and reliable datasets, enabling stakeholders to derive valuable insights, and support effective data analysis and interpretation.
- The Template Management component, which allows the creation and proper management of survey templates to be sent to patients and start capturing data in an organized, systematic, and seamless manner.
- The Security & Privacy Management component, which on one hand implements security mechanisms at all architectural layers in order to prevent among others SQL injection and cross site scripting attacks, while on the other hand it implements and enforces the various policies based on rules, terms and conditions.
- The Audit & Logging component, which records the occurrence of all events, the time at which they occurred, the responsible user or service, and the impacted entity.
- The Data Repository component, which is a vital aspect of the data capturing module, focusing on the physical / virtual storage of data in a secure and accessible manner.

4.2 Data Analysis & Cleansing Phase

In this phase the following stakeholders participate:

- The research team at each clinical site, and
- The sponsors

Specifically, the research team will be responsible for data management at each site, performing an initial data cleansing of data captured at each site. This is required in order to ensure that data are pseudonymized. As already mentioned in the data capture phase, the research team needs to make sure that exported data are pseudonymized before forwarding it to the sponsor. Even at this phase strict security mechanisms for allowing access to data are properly enforced. Specifically, two-factor authentication is utilized in order for any system user to gain access to captured and exported data.

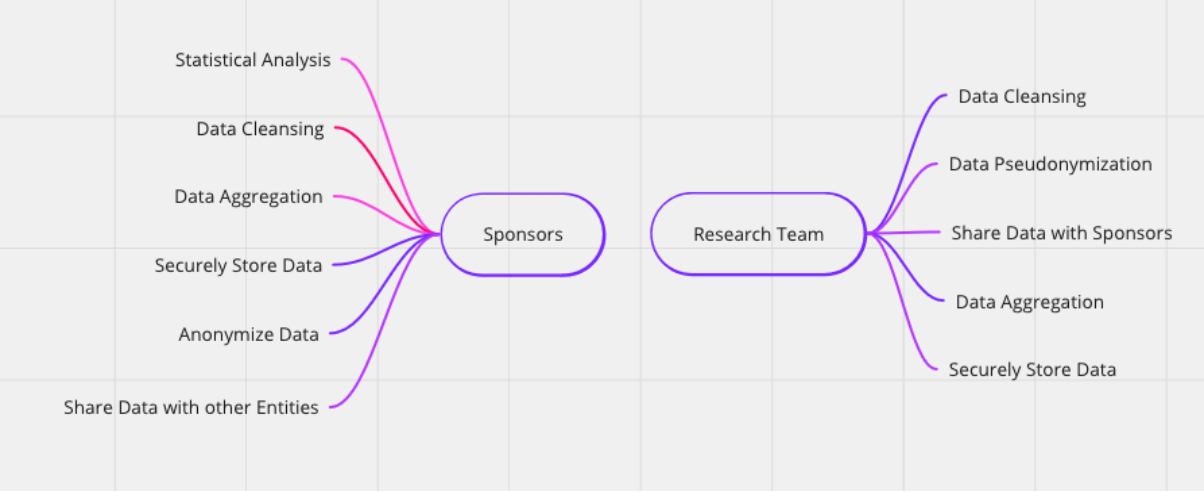

Figure 6: Actions & Responsibilities of Stakeholders in Data Analysis and Cleansing phase

On the other hand, sponsors will be responsible for data management of aggregated pseudonymized data received from study sites. To this end, sponsors undertake the performance of the statistical analysis of the IMPORTANT trial.

Data cleansing will be performed by sponsor and the appointed research team. In this process, pseudonymized data will be kept securely at the sponsor's site in a research repository where only sponsor and delegated personnel will be able to gain access. Data cleansing and analysis will be specifically performed within the premises of the research repository.

For analysis outside the premises of the research repository, data will be required to be anonymized. Again, sponsor will undertake the responsibility to anonymize the specific data. For this data, access will be ensured through the research repository for consortium partners that will perform additional and further analysis.

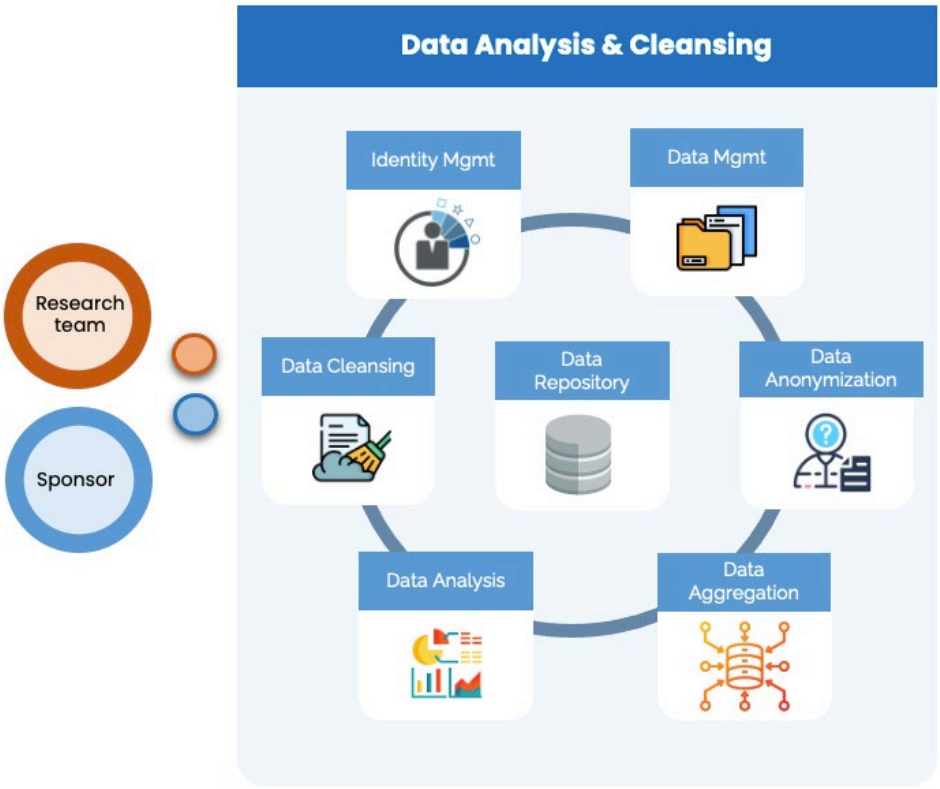

Figure 7: Data Analysis & Cleansing phase

As previously, the data analysis and cleansing phase in the IMPORTANT conceptual architecture is considered as a main module which consists of the following components:

- The Identity Management component, which enables the members of the research team and sponsors to be able to gain access to the data analysis and cleansing services based on their role and privileges.
- The Data Management component, which undertakes the responsibility to organize the data, implementing methods for structuring and categorizing data in a meaningful way for easy retrieval, analysis, and utilization.
- The Data Analysis component, which standardizes the range of continuous initial variables, computes the covariance matrix to identify correlations, create feature vectors to decide which principal components to keep and recast the data along the principal components' axes. For data analysis external tools and solutions may be utilized as well such as excel application, SPSS (Statistical Package for the Social Sciences) tools, etc.
- The Data Anonymization component, which enforces the data transformation and masking, implementing various anonymization types, and replacing the original values with randomly generated data based on the selected anonymization type.
- The Data Aggregation component, which implements processes for grouping data based on specific attributes (i.e., age, location, gender, etc.) dealing also with the quality of the data being aggregated.

- The Data Cleansing component, which undertakes the responsibility to execute processes for logical and system wised cleansing, such as business errors, empty values, removal of white spaces, duplicate records, parsing errors, etc.
- The Data Repository component, as previously, focuses on the physical / virtual storage of data in a secure and accessible manner.

4.3 Results Publishing

In this last phase the following stakeholders participate:

- The research team at each clinical site
- The sponsors, and
- The patients

Specifically, the research team and sponsors undertake the provision of the study results, which consist of aggregated results after suitable statistical analysis, to the European Commission. The research team at each site will make sure that data, concerning their participants who are directly relevant to the questions posed in the trial, is not published until the first publication of the analysis of the primary endpoint. Together with the sponsors will engage with health and research community, following results publication.

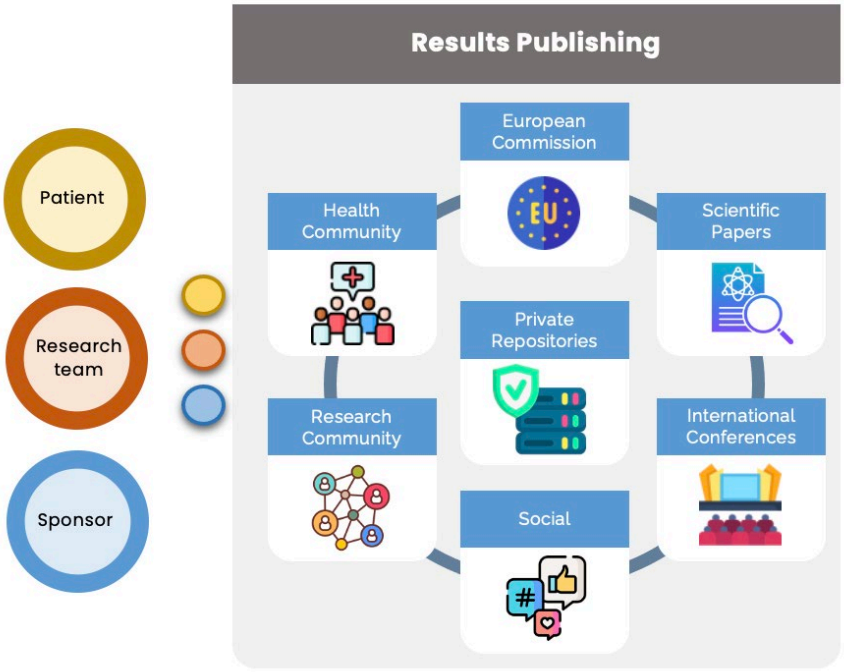

Figure 8: Results Publishing phase

The research team will also engage the patients in the trial through sharing of published results on targeted social media platforms. Finally, the research team and the sponsors will participate and disseminate the results in international conferences after the first publication of the analysis of the primary endpoint.

In this phase sponsor will be responsible for scientific publication of analysed data from the study, making sure that authorship and contributor-ship is discussed and agreed in the IMPORTANT Trial Steering Committee. Sponsors will also ensure that data is not released prior to the first publication of the analysis of the first endpoint, either for trial publication and/or oral presentation purposes, without the permission of the Trial Steering Committee. Published results will be kept at private site of the publisher and results presented at conferences will be kept at the organisers' repository.

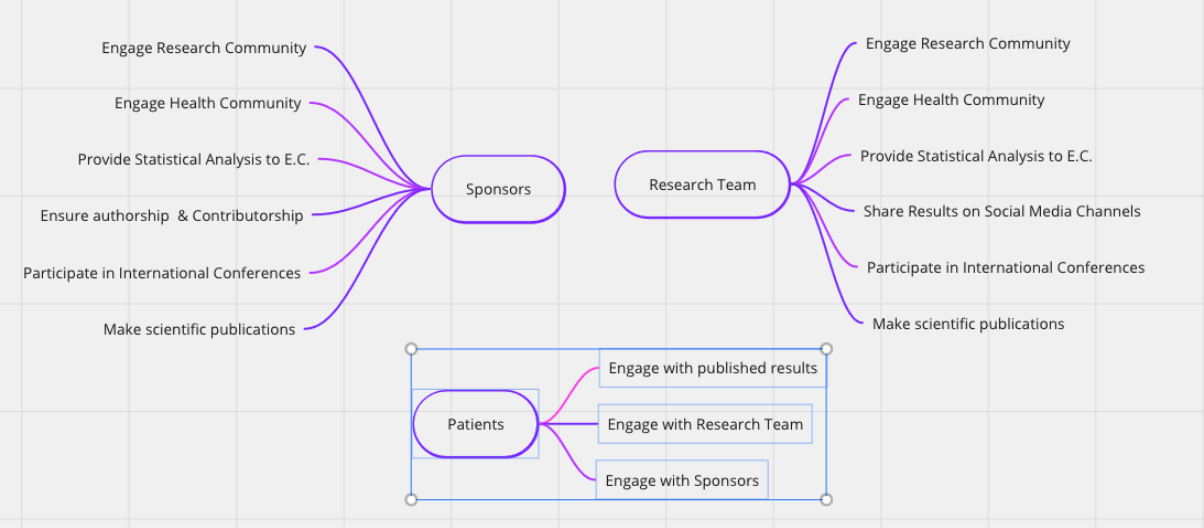

Figure 9: Actions and Responsibilities of Stakeholders in Publishing Results phase

Patient will be kept informed of all this process and will be able to engage with the published results that are shared on social media platforms or from information provided by physician or investigator at each study site.

Obviously, this phase cannot be considered as technological module in the IMPORTANT solution but only a phase in which several dissemination processes can run in parallel to one or more of the following sources:

- European Commission
- Health Community
- Research Community
- Journal Publications
- International and/or National Conferences and Workshops
- Social Media Channels

5 Technical Requirements and Specifications

The following paragraphs analyse and provide in detail the technical requirements and specifications of the IMPORTANT solution based on the introduction of components in Section 3. The exact requirements and specifications of all IMPORTANT modules are introduced in the following tables, based on the VOLERE methodology.

5.1 CRF Template Management

5.1.1 Specification Overview

The following table introduces the CRF template management specification.

Table 3: CRF Template Management Specification

Service Contract Name	**CRF Template Management**
Specification ID	IM-F-010
Specification Type	Functional (F)
Category / Grouping	Data Capture - Template Management - Data Repository - Data Management - Identity Management
Dependencies	None
Stakeholders	Research Team, CROs
Priority	Should Have
Description and Rationale	This specification refers to the design sub-phase of the data capture main phase. In this, Case Report Forms (CRFs), which are specialized documents used for gathering data in clinical research, need to be produced in terms of templates. This means that the end-user with the proper privileges and rights is allowed to manage all the templates to be used and published to patients.
Tool(s) / Solution(s)	SmartTrial

5.1.2 Interface Specifications

The following tables introduces the interfaces of the CRF template management specification.

Table 4: createTemplate Interface

Operation Name	*createTemplate*
Operation ID	IM-F-010.1
Description	This interface implements the creation process of a new template
Input	List of Question Types, Inserted Values
Output	New Template

Table 5: editTemplate Interface

Operation Name	*editTemplate*
Operation ID	IM-F-010.2
Description	This interface implements the edit process of an existing template
Input	List of existing Templates
Output	New version of Template

Table 6: deleteTemplate Interface

Operation Name	*deleteTemplate*
Operation ID	IM-F-010.3
Description	This interface implements the delete process of an existing template
Input	List of existing Templates
Output	None

Table 7: assignTemplate Interface

Operation Name	*assignTemplate*
Operation ID	IM-F-010.4
Description	This interface implements the assignment process of an existing template
Input	List of existing Templates, List of Patients' Identifiers
Output	Reference of template with List of Patients

Table 8: deAssignTemplate Interface

Operation Name	*deAssignTemplate*
Operation ID	IM-F-010.5
Description	This interface implements the de-assignment process of an already assigned survey
Input	Assigned Template, List of Patients' Identifiers
Output	Removal of one or more patients from the assigned CRF template

5.2 CRF Management

5.2.1 Specification Overview

The following table describes the CRF management specification.

Table 9: CRF Management Specification

Service Contract Name	**CRF Management**
Specification ID	IM-F-020
Specification Type	Functional (F)
Category / Grouping	Data Capture - Template Management - Data Repository

	- Data Management - Identity Management
Dependencies	CRF Template
Stakeholders	Patients, Research Team, CROs, Sponsors
Priority	Must Have
Description and Rationale	This specification refers to the management of one or more Case Report Forms (CRFs), which gather the results of a clinical research. This means that the end-user with the proper privileges and rights is allowed to complete the survey, and review the answers, while the system automatically stores each completed survey in the data repository of the data capture module.
Tool(s) / Solution(s)	SmartTrial

5.2.2 Interface Specifications

The following tables introduce the interfaces for the CRF management specification.

Table 10: publishSurvey Interface

Operation Name	*publishSurvey*
Operation ID	IM-F-020.1
Description	This interface implements the publication process of an already assigned template to a specific group / list of patients
Input	Assigned Template, List of Patients' Identifiers
Output	None

Table 11: submitCompletedSurvey Interface

Operation Name	*submitCompletedSurvey*
Operation ID	IM-F-020.2
Description	This interface implements all the required mechanisms that allow patients to complete and successfully submit the assigned survey
Input	Assigned Template, Inserted Values / Answers

Output	Complete

Table 12: reviewCompletedSurvey Interface

Operation Name	*reviewCompletedSurvey*
Operation ID	IM-F-020.3
Description	This interface implements all the required mechanisms that allow the research team, sponsors and/or CROs to review the values inserted from each patient in regard to a specific assigned to them survey
Input	Assigned Template, Patient's identifier
Output	Completed Survey

5.3 Provisioning

5.3.1 Specification Overview

The following table describes the provisioning specification.

Table 13: Provisioning Specification

Service Contract Name	**Provisioning**
Specification ID	IM-F-030
Specification Type	Security (S), Functional (F)
Category / Grouping	Data Capture - Identity Management
Dependencies	None
Stakeholders	Patients, Research Team, CROs, Sponsors
Priority	Must Have
Description and Rationale	Provisioning will undertake the management and the maintenance of the IMPORTANT users to access the resources available in the data capture phase. Provisioning in this phase will include among others creation, update, and deletion of the user accounts.

Tool(s) / Solution(s)	SmartTrial

5.3.2 Interface Specifications

The following tables introduce the interfaces for the provisioning specification.

Table 14: createUser Interface

Operation Name	*createUser*
Operation ID	IM-F-030.1
Description	This operation will be responsible for creating a new IMPORTANT user account
Input	User Account Entity's Attributes
Output	New User Account

Table 15: updateUser Interface

Operation Name	*updateUser*
Operation ID	IM-F-030.2
Description	This operation will be responsible for updating specific details of an existing IMPORTANT user account.
Input	User Account identifier, User Account Entity's Attributes
Output	User Account

Table 16: deleteUser Interface

Operation Name	*deleteUser*
Operation ID	IM-F-030.3
Description	The deleteUser operation will be responsible for deleting an existing IMPORTANT user account, respecting the policies of the platform, and deleting all the data captured for this user
Input	User Account identifier, Data Gathered for this User Account

Output	None

Table 17: associateWithOrganization Interface

Operation Name	*associateWithOrganization*
Operation ID	IM-F-030.4
Description	The deleteUser operation will be responsible for deleting an existing IMPORTANT user account, respecting the policies of the platform, and deleting all the data captured for this user
Input	User Account identifier, Data Gathered for this User Account
Output	None

Table 18: removeAssociation Interface

Operation Name	*removeAssociation*
Operation ID	IM-F-030.5
Description	The removeAssociation operation will be responsible of removing the specific association type between a user and an organization account
Input	User Account identifier, Data Gathered for this User Account
Output	None

5.4 Authentication

5.4.1 Specification Overview

The following table introduces the authentication specification.

Table 19: Authentication Specification

Service Contract Name	**Authentication**
Specification ID	IM-S-040
Specification Type	Security (S)

Category / Grouping	Data Capture Module - Identity Management component Data Analysis & Cleansing Module - Identity Management component
Dependencies	None
Stakeholders	Patients, Research Team, CROs, Sponsors
Priority	Must Have
Description and Rationale	Authentication in the context of an IMPORTANT user accessing either the IMPORTANT modules and/or any other application included in the corresponding phases will undertake to answer on the question “who the current user is”. A full authentication protocol will probably be required in the project’s case that will allow to additionally include a number of attributes about this user (i.e., email address). This is required across the IMPORTANT solution. Specifically, for the data capture, and the data analysis and cleansing phases in which the security and privacy criticality is very high, two (2) factor authentication will be enforced in order for a user to properly gain access to the services of the aforementioned modules.
Tool(s) / Solution(s)	SmartTrial

5.4.2 Interface Specifications

The following tables introduce the interfaces required for the realization of the authentication specification.

Table 20: userLogin Interface

Operation Name	*userLogin*
Operation ID	IM-S-040.1
Description	This operation will allow all end-users, regardless their type, to be able to gain access on the data capture module and its services
Input	User Account Credentials
Output	Valid Session

Table 21: userLogout Interface

Operation Name	*userLogout*
Operation ID	DC-F-040.2
Description	This operation will allow all end-users, to sign-out the data capture services and functions
Input	User Identifier, List of Active Sessions
Output	None

Table 22: twoFactorAuthentication Interface

Operation Name	*twoFactorAuthentication*
Operation ID	IM-S-040.3
Description	This operation will allow all end-users, to sign-in utilizing two different channels and sources of truth
Input	User Identifier, User Email Address
Output	Validation Token

5.5 Authorization

5.5.1 Specification Overview

The following table describe the authorization specification.

Table 23: Authorization Specification

Service Contract Name	**Authorization**
Specification ID	IM-S-050
Specification Type	Security (S)
Category / Grouping	Data Capture Module - Identity Management component Data Analysis & Cleansing Module - Identity Management component

Dependencies	Authentication
Stakeholders	Patients, Research Team, CROs, Sponsors
Priority	Must Have
Description and Rationale	Authorization refers to the process of verifying what a user has access to. In this data capture phase, authorization will occur after identity is successfully validated through the authentication process. Authorization will be determined through the use of policies and rules, which can be used with Role-Based Access Control (RBAC)
Tool(s) / Solution(s)	SmartTrial

5.5.2 Interface Specifications

The following tables describe the interfaces required for the authorization specification.

Table 24: createRole Interface

Operation Name	*createRole*
Operation ID	IM-S-050.1
Description	The role will be a collection of permissions, applied to users. Using roles will make it easier for the data capture module to add, remove, and adjust permissions than assigning permissions to users individually. Therefore, this operation will allow the creation of a new Role in which permissions can be properly assigned.
Input	Role Entity's Attributes
Output	Role Entity

Table 25: updateRole Interface

Operation Name	*updateRole*
Operation ID	IM-S-050.2
Description	This operation will allow the update of an existing Role in which permissions can be properly modified and configured.

Input	Role Entity's Identifier & Attributes
Output	Role Entity

Table 26: deleteRole Interface

Operation Name	*deleteRole*
Operation ID	IM-S-050.3
Description	This operation will allow the deletion of an existing Role.
Input	Role Entity's Identifier
Output	None

Table 27: assignPermissions Interface

Operation Name	*assignPermissions*
Operation ID	IM-S-050.4
Description	Granting of roles to a data capture requires the assignment of permissions to the last mentioned of one or more give system roles.
Input	List of Actions, List of Services, List of Roles
Output	None

Table 28: updatePermissions Interface

Operation Name	*updatePermissions*
Operation ID	IM-S-050.5
Description	This operation will allow to update existing permissions to services account for one or more give system roles.
Input	List of Actions, List of Services, List of Roles
Output	None

Table 29: grantRole Interface

Operation Name	*grantRole*
Operation ID	IM-S-050.6
Description	This operation will allow to grant a specific role to a specific user account.
Input	Role Entity Identifier, User Account Identifier
Output	None

Table 30: removeRole Interface

Operation Name	*removeRole*
Operation ID	IM-S-050.7
Description	This operation will allow to remove an already assigned role from a specific user account.
Input	Role Entity Identifier, User Account Identifier
Output	None

5.6 Profile Management

5.6.1 Specification Overview

The following table introduces the profile management specification.

Table 31: Profile Management Specification

Service Contract Name	**Profile Management**
Specification ID	IM-FP-060
Specification Type	Functional (F), Privacy (P)
Category / Grouping	Data Capture Module - Identity Management component - Data Management component - Data Repository component

	Data Analysis and Cleansing Module - Identity Management component - Data Management component - Data Anonymization component - Data Repository component - Data Analysis component - Data Repository component
Dependencies	None
Stakeholders	Patients, Research Team, CROs, Sponsors
Priority	Should Have
Description and Rationale	Profile management will ensure that the user's personal settings are applied to the IMPORTANT modules and applications, regardless of the location and end point device. Profile management will be enabled through a profile optimization service that will provide an easy, reliable way for managing these settings, providing a consistent experience through the maintenance of a single profile that follows the user. This will also partly be applied to the Organization Entity, which is needed also to maintain a specific (business-oriented) profile within the IMPORTANT solution.
Tool(s) / Solution(s)	SmartTrial

5.6.2 Interface Specifications

The following tables describe the interfaces required for the realization of the profile management specification.

Table 32: createUserProfile Interface

Operation Name	*createUserProfile*
Operation ID	IM-FP-060.1
Description	The operation will allow the creation of a user profile in the IMPORTANT solution. By default, this profile will be created simultaneously with the creation of the corresponding user account.
Input	User Profile Entity's Attributes
Output	User Profile Entity

Table 33: updateUserProfile Interface

Operation Name	*updateUserProfile*
Operation ID	IM-FP-060.2
Description	The operation will allow the update of an existing user profile in the IMPORTANT solution.
Input	User Profile Entity's Attributes
Output	User Profile Entity

Table 34: deleteUserProfile Interface

Operation Name	*deleteUserProfile*
Operation ID	IM-FP-060.3
Description	The operation will allow the deletion of an existing user profile in the IMPORTANT solution, respecting the internal policies and dependencies.
Input	User Profile Entity's Attributes
Output	User Profile Entity

Table 35: userProfileAnonymization Interface

Operation Name	*userProfileAnonymization*
Operation ID	IM-FP-060.4
Description	The operation will allow the anonymization of specific profile attributes.
Input	User Profile Entity's Attributes
Output	Random Data Generated

Table 36: userProfileCleansing Interface

Operation Name	*userProfileCleansing*
Operation ID	IM-FP-060.5
Description	The operation will allow the data cleansing of one or more profiles.
Input	User Profile Entity's Attributes
Output	Updated User Profile

5.7 Data Analytics

5.7.1 Specification Overview

The following table introduces the data analytics specification.

Table 37: Data Analytics Specification

Service Contract Name	**Data Analytics**
Specification ID	IM-F-070
Specification Type	Functional (F)
Category / Grouping	Data Capture Module - Data Management component - Data Repository component Data Analysis and Cleansing Module - Data Management component - Data Repository component - Data Analysis component
Dependencies	Data Anonymization, Data Aggregation, Data Cleansing
Stakeholders	Research Team, CROs, Sponsors
Priority	Must Have
Description and Rationale	Data analytics implements the process of inspecting, transforming, and modelling the IMPORTANT data with the goal of discovering useful information, informing conclusions, and supporting decision-making. Specifically, statistical analysis and clinical study report are performed giving answers to trials objectives/endpoints based on

	the data captured in the first step of the IMPORTANT trial scenario. This includes among others the identification of patterns, the detection of anomalies, the enforcement of predictive models, etc.
Tool(s) / Solution(s)	SPSS, Stata, R

5.7.2 Interface Specifications

The following tables describe the interfaces required for the realization of the data analytics specification. It should be noted that these interfaces refer only to the proper retrieval of data, since the actual data analysis is performed by the research team enforcing custom algorithms and analysis patterns.

Table 38: importDataset Interface

Operation Name	*importDataset*
Operation ID	IM-F-070.1
Description	The operation will implement functions which will allow the end users to import preferred datasets utilizing the data models already implemented.
Input	Datasets
Output	(None)

Table 39: exportDataset Interface

Operation Name	*exportDataset*
Operation ID	DC-F-070.2
Description	The operation will implement functions which will allow the end users to export one or more selected datasets
Input	Dataset identifier
Output	Datasets in various format types and files

Table 40: deleteDataset Interface

Operation Name	*deleteDataset*
Operation ID	IM-F-070.3
Description	The operation will allow the deletion of one or more selected datasets
Input	Dataset Identifier
Output	none

Table 41: findByAttribute Interface

Operation Name	*findByAttribute*
Operation ID	IM-F-070.4
Description	The operation will allow the end user to search in a preferred dataset with one or more specific attribute values
Input	Dataset Identifier, Preferred Attribute, Attribute Value
Output	Filtered Datasets

5.8 Reporting

5.8.1 Specification Overview

The following table introduces the reporting specification.

Table 42: Reporting Specification

Service Contract Name	**Reporting**
Specification ID	IM-F-080
Specification Type	Functional (F)
Category / Grouping	Data Capture Module - Data Management component - Data Repository component Data Analysis and Cleansing Module - Data Management component - Data Repository component

	- Data Analysis component
Dependencies	Data Anonymization, Data Aggregation, Data Cleansing
Stakeholders	Research Team, CROs, Sponsors
Priority	Must Have
Description and Rationale	Reporting services will enable IMPORTANT solution to generate and provide reports and their definitions. This specification will include the definition of the required IMPORTANT elements that will concretely define the data source connections, queries to retrieve data, expressions, parameters, images, text boxes, tables, and other. These reports and among other may be one of the following types: - Aggregation reports - Query reports - Drill-down reports - Combo tree reports
Tool(s) / Solution(s)	SmartTrial, FoU Drive

5.8.2 Interface Specifications

The following tables describe the interfaces required for the realization of the reporting specification.

Table 43: generateReport Interface

Operation Name	*generateReport*
Operation ID	IM-F-080.1
Description	This operation will enable the IMPORTANT solution or an individual user to generate a specific report
Input	Datasets, Report Template
Output	Report

Table 44: exportReport Interface

Operation Name	*exportReport*
Operation ID	IM-F-080.2
Description	This operation will be responsible of exporting an existing report from the IMPORTANT solution to a specific file format (i.e. csv, file)
Input	Report Entity Identifier
Output	Specific File Type

Table 45: shareReport Interface

Operation Name	*shareReport*
Operation ID	IM-F-080.3
Description	This operation will enable the user to forward the report to other recipients
Input	Report Entity Identifier, User Account Identifier
Output	None

5.9 Audit & Logging Management

5.9.1 Specification Overview

The following table introduces the audit and logging management specification.

Table 46: Audit and Logging Management Specification

Service Contract Name	**Audit and Logging Management**
Specification ID	IM-SP-090
Specification Type	Security (S), Privacy (P)
Category / Grouping	Data Capture Module - Identity Management component - Data Management component - Data Repository component

	- Security & Privacy Management component - Template Management Data Analysis and Cleansing Module - Identity Management component - Data Management component - Data Cleansing component - Data Anonymization component - Data Repository component - Data Aggregation component - Data Analysis component
Dependencies	Identity Management component
Stakeholders	Patients, Research Team, CROs, Sponsors
Priority	Must Have
Description and Rationale	Audit and Logging management service will undertake to describe all the activities and processes used to generate, collect, centralize, parse, store, archive, and dispose of system-generated log data. Therefore, this service will be used to handle all the logs generated by the IMPORTANT modules, components, and its users.
Tool(s) / Solution(s)	SmartTrial, FoU Drive

5.9.2 Interface Specifications

The following tables describe the interfaces required for the realization of the audit and logging management specification.

Table 47: collectLogInterface

Operation Name	*collectLog*
Operation ID	IM-SP-090.1
Description	This operation will allow the system to collect log data and store it. These logs will be aggregate data from different parts, phases, modules, and components of the IMPORTANT environment
Input	Log Data, System/Component Identifier, User Account Identifier

Output	None

Table 48: filterLog Interface

Operation Name	*filterLog*
Operation ID	IM-SP-090.2
Description	This service will allow the proper filtering of logs. This, among other, includes the following: - Search-By-Text - Search-By-Date - Search-By-Number - Advance-Search
Input	Filter Options
Output	Log Data

Table 49: anonymizeLog Interface

Operation Name	*anonymizeLog*
Operation ID	IM-SP-090.3
Description	This service will allow the proper anonymization of logged data enforcing the privacy policy of each case
Input	Log Data, System/Component Identifier, User Account Identifier
Output	Updated Log Data

5.10 Visualization

5.10.1 Specification Overview

The following table introduces the visualization specification.

Table 50: Visualization Specification

Service Contract Name	**Visualization**
Specification ID	IM-UP-100
Specification Type	Usability (U), Privacy (P)
Category / Grouping	Data Capture Module - Identity Management component - Data Management component - Data Repository component - Security & Privacy Management component - Template Management Data Analysis and Cleansing Module - Identity Management component - Data Management component - Data Cleansing component - Data Anonymization component - Data Repository component - Data Aggregation component - Data Analysis component
Dependencies	None
Stakeholders	Patients, Research Team, CROs, Sponsors
Priority	Must Have
Description and Rationale	Dashboards are used to track KPI metrics that are important to user's they are gaining access to. In the IMPORTANT solution case it's important to quickly obtain visibility into changes regarding the data received from the various components, alerts and notifications, etc. Therefore, advanced visualization services will allow the users to build a powerful way to report and keep track of their preferred data.
Tool(s) / Solution(s)	Smarttrial, SPSS, Stata, R

5.10.2 Interface Specifications

The following tables describe the interfaces required for the realization of the visualization specification.

Table 51: getContent Interface

Operation Name	*getContent*
Operation ID	IM-UP-100.1
Description	This operation will allow the system to collect data from all modules and based on the role and privileges of the user who is successfully logged in to visualize the proper content
Input	Data Query
Output	Data Model

5.11 GDPR Compliance

5.11.1 Specification Overview

The following table introduces the GDPR compliance specification.

Table 52: GDPR compliance Specification

Service Contract Name	**GDPR Compliance**
Specification ID	IM-PS-110
Specification Type	Privacy (P), Security (S)
Category / Grouping	Data Capture Module - Identity Management component - Data Management component - Data Repository component - Security & Privacy Management component - Template Management Data Analysis and Cleansing Module - Identity Management component - Data Management component - Data Cleansing component

	- Data Anonymization component - Data Repository component - Data Aggregation component - Data Analysis component
Dependencies	Identity Management component
Stakeholders	Patients, Research Team, CROs, Sponsors
Priority	Must Have
Description and Rationale	This specification deals with the mechanisms that need to be implemented in order to consider and apply the GDPR framework and the specific IMPORTANT privacy needs. This specification relates to almost all components of the IMPORTANT modules, which process sensitive data and mostly with the data anonymization, management, and repository
Tool(s) / Solution(s)	Smarttrial, FoU Drive

5.11.2 Interface Specifications

The following tables describe the interfaces required for the realization of the GDPR compliance specification.

Table 53: encryptData Interface

Operation Name	*encryptData*
Operation ID	IM-PS-110.1
Description	This operation will allow to encrypt specific attributes within a given data set (Attribute Based Encryption – ABE). Most of the times these data include among other: - Names - Emails - Addresses - Specific health data
Input	List of Attributes, Encryption algorithm, Dataset
Output	Encrypted Dataset

Table 54: manageUserConsent Interface

Operation Name	*manageUserConsent*
Operation ID	IM-PS-110.2
Description	This operation will allow the system to properly manage (get, apply, securely store, force update) the consent of all end-users, regardless their type and role within the IMPORTANT solution
Input	User Account Identifier, Consent Types, Terms and Conditions, Privacy Policy, Cookies Policy
Output	None

Table 55: managePrivacyPolicy Interface

Operation Name	*managePrivacyPolicy*
Operation ID	DC-PS-110.3
Description	This operation will allow the proper management (create, update, publish, remove) of the Privacy policy
Input	User Account Identifier, Privacy Policy
Output	None

Table 56: manageCookiesPolicy

Operation Name	*manageCookiesPolicy*
Operation ID	IM-PS-110.4
Description	This operation will allow the proper management (create, update, publish, remove) of the Cookies policy
Input	User Account Identifier, Cookies Policy
Output	None

Table 57: rightToBeForgotten Interface

Operation Name	*rightToBeForgotten*
Operation ID	IM-PS-110.5

Description	This is the operation that will allow the user to request to erase his/her personal data and delete his/her account.
Input	User Account Identifier, User Consent Policy
Output	None

Table 58: anonymizeData Interface

Operation Name	*anonymizeData*
Operation ID	IM-PS-110.6
Description	This operation will allow the allow the anonymization of personal data within a given data set. Anonymization in IMPORTANT solution may be realized either by encrypting or removing personally identifiable information from data sets so that the users whom the data describes remain permanently anonymous.
Input	User Account Identifier, Dataset, Sensitive Attributes
Output	Anonymized Dataset

5.12 Policy Management

5.12.1 Specification Overview

The following table introduces the policy management specification.

Table 59: Policy Management Specification

Service Contract Name	**Policy Management**
Specification ID	IM-F-120
Specification Type	Functional (F)
Category / Grouping	Data Capture Module - Identity Management component - Data Management component - Data Repository component - Security & Privacy Management component - Template Management

	Data Analysis and Cleansing Module - Identity Management component - Data Management component - Data Cleansing component - Data Anonymization component - Data Repository component - Data Aggregation component - Data Analysis component
Dependencies	None
Stakeholders	Patients, Research Team, CROs, Sponsors
Priority	Could Have
Description and Rationale	The Policy Management service will be used to capture decision logic as a business rule, which is then automated across the IMPORTANT solution's applications and modules. Instead of embedding rules as code within the various internal modules, with Policy Management Service, the rules will be externalized and managed away from the application code. This will enable the logic to be leveraged by multiple modules and changed independently from the governing one.
Tool(s) / Solution(s)	Smarttrial, FoU Drive

5.12.2 Interface Specifications

The following tables describe the interfaces required for the realization of the policy management specification.

Table 60: enforcePolicy Interface

Operation Name	*enforcePolicy*
Operation ID	IM-F-120.1
Description	This operation will allow the enforcement of one or more specific policies within the various business and IMPORTANT solution services
Input	List of Rules, List of arguments, Service identifier, User Account Identifier
Output	None

5.13 Data Management

5.13.1 Specification Overview

The following table introduces the data management specification.

Table 61: Data Management Specification

Service Contract Name	**Data Management**
Specification ID	IM-FSP-130
Specification Type	Functional (F), Security (S), Privacy (P)
Category / Grouping	Data Capture Module - Data Repository component - Data Management component Data Analysis and Cleansing Module - Data Repository component - Data Management component
Dependencies	Data Repository, Data Cleansing, Data Aggregation, Data Anonymization
Stakeholders	Research Team, CROs, Sponsors
Priority	Must Have
Description and Rationale	The data management service will undertake the responsibility to implement all related services in order to allow Research Team members, CROs and/or Sponsors to share captured data with other entities such physicians, regulatory authorities, etc. It should be noted that the Research team needs to make sure that no records will be transferred to another location or party without written notification to the sponsor/CRO.
Tool(s) / Solution(s)	SmartTrial, FoU Drive

5.13.2 Interface Specifications

The following tables describe the interfaces required for the realization of the data management specification.

Table 62: anonymizeData Interface

Operation Name	*anonymizeData*
Operation ID	IM-FSP -130.1
Description	This operation will allow the anonymization of data collected, enforcing the privacy policy of sharing with other entities. Each entity may be allowed to view different parts of the dataset
Input	Dataset, Data model, Anonymization needs (columns)
Output	Anonymized Dataset

Table 63: performDataCleansing Interface

Operation Name	*performDataCleansing*
Operation ID	IM-FSP -130.2
Description	This operation will allow the data cleansing of the dataset captured
Input	Cleansing rules, Dataset, Data model
Output	Dataset

Table 64: aggregateData Interface

Operation Name	*aggregateData*
Operation ID	IM-FSP -130.3
Description	This operation will allow the aggregation of data based on rules (i.e., User Account Identifier)
Input	List of Aggregation Rules, List of arguments, User Account Identifier, Datasets
Output	Updated Datasets

Table 65: pseudonymizeData Interface

Operation Name	*pseudonymizeData*
Operation ID	IM-FSP -130.4
Description	This operation will allow the pseudonymization of Patient Identity when sharing or further analysis needs to be performed
Input	User Account Identifier, Datasets
Output	Updated Datasets

Table 66: shareData Interface

Operation Name	*shareData*
Operation ID	IM-FSP -130.5
Description	This operation will allow the sharing of data with other entities
Input	Datasets, User Account Identifier
Output	Datasets

6 Summary and Conclusions

This deliverable is the main outcome of T2.2 “IMPORTANT technical requirements for decentralized clinical trial”, which specifies the components comprising the IMPORTANT trail case, along with the structuring principles of the solution.

Its main scope is to provide the technical requirements and specification of services, systems, tools, assets, and procedures which are already established, drawing the main conceptual architectural, its main components and all the entities that are included in each one of them.

Based on this report, the IMPORTANT trial case recognizes three different phases:

- Data capture phase, in which data capturing processes from patients run. Moreover, in this phase pseudonymization of data occur before this is exported and shared with other entities.
- Data analysis and cleansing phase, in which specific data management services are executed at each site. This means that pseudonymization, anonymization, data cleansing and aggregation is performed in order for this data to be properly analysed and generate results.
- Publishing phase, in which all involved stakeholders undertake the responsibility to publish and disseminate the results in various entities, and channels such as the European Commission, the health and research communities, international conferences etc.

This document identifies 13 major specifications and a big number of relative interfaces to address functional, non-functional, security, privacy, usability, performance and infrastructure requirements. All identified requirements and technical specifications seriously affect the actions that need to be properly taken in WP3 and WP4 tasks.

In the future, the same approach will be applied in any additional but parallel and relevant to the trial projects.

References

[1] S. Robertson and J. Robertson, "*Mastering the Requirements Process*", Addison-Wesley, 1999

[2] Bittner, Kurt; Spence, Ian, "*Use Case Modeling*", Addison-Wesley Professional, ISBN 978-0-201-70913-1

[3] "MoSCoW Analysis (6.1.5.2)", A Guide to the Business Analysis Body of Knowledge (2nd ed.), International Institute of Business Analysis, ISBN 978-0-9811292-1-1, 2009

[4] European Medicines Agency. (2022, December 19). Facilitating decentralised clinical trials in the EU. https://www.ema.europa.eu/en/news/facilitating-decentralised-clinical-trials-eu

[5] U.S. Food and Drug Administration. Decentralized Clinical Trials for Drugs, Biological Products, and Devices Draft Guidance Webinar June 20, 2023

[6] McAlindon T, Formica M, Kabbara K, LaValley M, Lehmer M. Conducting clinical trials over the internet: feasibility study. BMJ. 2003 Aug 30;327(7413):484-7. doi: 10.1136/bmj.327.7413.484

[7] McAlindon, Timothy E., and Karim AN Kabbara. "Method for conducting clinical trials over the internet." U.S. Patent No. 7,251,609. 31 Jul. 2007

[8] Orri M, Lipset CH, Jacobs BP, Costello AJ, Cummings SR. Web-based trial to evaluate the efficacy and safety of tolterodine ER 4 mg in participants with overactive bladder: REMOTE trial. Contemp Clin Trials. 2014 Jul;38(2):190-7. doi: 10.1016/j.cct.2014.04.009

[9] PETRINI, Carlo, et al. Decentralized clinical trials (DCTs): A few ethical considerations. Frontiers in Public Health, 2022, 10: 1081150

[10] Clegg, Dai; Barker, Richard, "*Case Method Fast-Track: A RAD Approach*", Addison-Wesley. ISBN 978-0-201-62432-8, 1994

[11] Beynon-Davies, Williams, "*The diffusion of information systems development methods*", Journal of Strategic Information Systems 12 p. 29-46, 2003

[12] Sjaak Brinkkemper, Saeki and Harmsen, "*Assembly Techniques for Method Engineering*", Advanced Information Systems Engineering, Proceedings of CaiSE'98, Springer Verlag, 1998

[13] Abrahamsson, Salo, Ronkainen, Warsta, "*Agile Software Development Methods: Review and Analysis*", VTT Publications 478, p. 61-68, 2002