Improving cancer diagnosis and prediction with AI and big data

**A Multimodal AI-based Toolbox and an Interoperable Health Imaging Repository for the Empowerment of Imaging Analysis related to the Diagnosis, Prediction and Follow-up of Cancer**

# Deliverable 6.7

# Best Practices for Wider Use and applicability

## WP 6 – System integration and pilot studies

29-03-2024

Revision 0.1.10

Status: Final

Grant Agreement n 952179

| DOCUMENT CONTROL | |
|---|---|
| **Project reference** | Grant Agreement number: 952179 |
| **Document name** | Best Practices for Wider Use and applicability |
| **Work Package** | WP 6 |
| **Work Package Title** | System integration and pilot studies |
| **Dissemination level** | PU |
| **Revision** | 0.1.10 |
| **Status** | Final |
| **Reviewers** | Magdalena Kogut, TLX; Gianna Tsakou, MAG |
| **Beneficiary(ies)** | UNS, KU, AUTH, MAG, ICCS, ED, BSC, TIS, FTSS, MED, CUT, CERTH |

*Dissemination level:*
*PU = Public, for wide dissemination (public deliverables shall be of a professional standard in a form suitable for print or electronic publication) or CO = Confidential, limited to project participants and European Commission.*

| AUTHORS | | |
|---|---|---|
| | **Name** | **Organisation** |
| **Document leader** | Tatjana Loncar-Turukalo | UNS |
| **Participants** | Niksa Jakovljevic | UNS |
| | Jasmina Boban | UNS |
| | Nebojsa Bozanic | UNS |
| | Ivan Lazic | UNS |
| | Mladen Kovacevic | UNS |
| | Elnabhani Shereen | KU |
| | Zaharias Lithin | KU |
| | Ghada Elkhoury | KU |
| | Sara Alabart Martínez | FTSS |
| | Ana Belen Callado Gil | FTSS |
| | Didier Dominguez | FTSS |
| | Alberto Gutiérrez Torre | BSC |
| | Andreas Charalambous | CUT |
| | Sotirios Chatzis | CUT |
| | Nikolina Dodlek | CUT |
| | Alexandra Kosvyra | AUTH |
| | Ioanna Chouvarda | AUTH |
| | Dimitris Filos | AUTH |
| | Olga Tsave | AUTH |

| | | |
|---|---|---|
| | Dimitris Fotopoulos | AUTH |
| | Caroline Barelle | ED |
| | Pablo Mezzon | ED |
| | Paris Laras | MAG |
| | Gianna Tsakou | MAG |
| | Andreas Charalambous | CUT |
| | Apostolos Dimoulakis | CERTH |
| | Dimitris Manolakis | CERTH |
| | Dimitris Trakos | CERTH |
| | Evangelos Politis | CERTH |
| | Giannis Aslanis | CERTH |
| | Georgios Foroglou | CERTH |
| | Maria Tsiobra | CERTH |
| | Michail Nikolaidis | CERTH |
| | Nikos Siopis | CERTH |
| | Paschalis Bizopoulos | CERTH |
| | Stelios Kokkas | CERTH |
| | Tasos Bachtses | CERTH |
| | Vasilis Charalabidis | CERTH |
| | Anastasios Tzelepakis | CERTH |
| | Zisis Sakellariou | CERTH |
| | Antonios Lalas | CERTH |
| | Konstantinos Votis | CERTH |
| | Pol Camps | WR |
| | Laia Juan | MED |
| | Magdalena Kogut | TLX |
| | Stavros Sykiotis | ICCS |
| | Ioannis Tzortzis | ICCS |
| | Ioannis Rallis | ICCS |
| | Aikaterini Angeli | ICCS |
| | Dimitrios Kalogeras | ICCS |
| | Anastasios Doulamis | ICCS |
| | Iason Katsamenis | ICCS |
| | Ioannis Vergardos | ICCS |
| | Matthaios Mpimpas | ICCS |
| | Athanasios Voulodimos | ICCS |
| | Georgios Kopsiaftis | ICCS |

| REVISION HISTORY | | | | |
|---|---|---|---|---|
| **Revision** | **Date** | **Author** | **Organisation** | **Description** |
| 0.1 | 1.02.2024 | Tatjana Loncar_Turukalo | UNS | Definition of table of contents |
| 0.1.1 | 15.02.2024. | Tatjana Loncar_Turukalo | UNS | Document revision and content added |
| 0.1.2 | 27.02.2024 | Gianna Tsakou | MAG | Comments and guidelines on the content |
| 0.1.3 | 05.03.2024 | Alberto Gutiérrez Torre | BSC | Input regarding the infrastructure |
| 0.1.4 | 10.03.2024. | Tatjana Loncar_Turukalo | UNS | Revision of content |
| 0.1.5 | 21.03.2024 | Caroline Barelle | ED | Input regarding the infrastructure |
| 0.1.6 | 22.03.2024. | Tatjana Loncar_Turukalo | UNS | Revision and restructuring the content |
| 0.1.7 | 26.03.2024. | Tatjana Loncar_Turukalo | UNS | Revision and restructuring of the content, discussion section |
| 0.1.8 | 27.03.2024. | Magdalena Kogut | TLX | Internal review |
| 0.1.9 | 29.03.2024. | Gianna Tsakou | MAG | Internal review |
| 0.1.10 | 29.03.2024 | Tatjana Loncar_Turukalo | UNS | Final editing based on the review feedback |

**Disclaimer and statement of originality**

*The content of this deliverable represents the views of the authors only and is their sole responsibility; it cannot be considered to reflect the views of the European Commission and/or the Consumers, Health, Agriculture and Food ExecutiveAgency or any other body of the European Union. The European Commission and the Agency do not accept any responsibility for use of its contents.*

*This deliverable contains original unpublished work except where clearly indicated otherwise. Acknowledgement of previously published material and of the work of others has been made through appropriate citation, quotation or both.*

# Table of Contents



# Table of figures

## Terms and Abbreviations

| Term | Description |
|---|---|
| AI Toolbox | AI solutions to improve cancer detection systems and enhance the clinical workflow |
| AI4HI | Group of 5 health imaging AI projects: INCISIVE, Primage, EuCanImage, ProCAncer-I, Chaimeleon. |
| Delphi | Delphi is a research survey technique used as a way of collecting data from respondents within their domain of expertise. It is a process used to arrive at a group opinion or decision by surveying a panel of experts. Experts respond to several rounds of questionnaires, and the responses are aggregated and shared with the group after each round. |
| Federated Data Sharing | The Data Providers share the Data by keeping them within their infrastructures (physical or virtual). After being pre-processed locally, they are made interoperable with other data existing in the INCISIVE Repository, through a data sharing mechanism. |
| Federated Learning | Means that AI model is trained in a distributed way using the required Federated nodes, i.e., the nodes that have the Data that matched with the user query. Each Federated or Central Node contains a particular set of Data that may be required for training, and it will not leave the node to ensure privacy. The model is trained in each Federated node, including Central Node, and then it is sent to the Central infrastructure to be merged to gather 'central knowledge'. This training-merging process can be repeated more than once for the same model as the more times this is done, the more robust is the solution. |
| Federated Node | Dedicated infrastructure (physical or virtual) in which the INCISIVE Data Provider stores the Data which they contribute to the INCISIVE Repository. |
| Federated Space | A virtual space formed by the composition of all Federated nodes, including the Central node, enabling the performance of centralized operations, such as the training of the AI models, access to Data utilized in the INCISIVE Platform, and visualizations of the results etc. |
| FUTURE-AI | An international, multi-stakeholder initiative for defining and maintaining concrete guidelines that will facilitate the design, |

| | |
|---|---|
| | development, validation and deployment of trustworthy AI solutions in medicine and healthcare based on six guiding principles: Fairness, Universality, Traceability, Usability, Robustness and Explainability. |
| Hybrid data sharing | Data sharing which takes place using both Federated data sharing and a Central node. |
| INCISIVE Data or Data | Anonymized medical data and images made available for re-use in the INCISIVE Repository, initially collected by INCISIVE Data Providers for the purposes of INCISIVE Project. |
| INCISIVE Platform or Platform | Platform (technical infrastructure integrating several components) provided by the Project which includes the Hybrid repository, AI development workspaces for AI training and the Inference services. |
| INCISIVE Repository | collection of INCISIVE Data and technical infrastructure (subset of the INCISIVE Platform) for the data's secure storage with the aim of making the data available for re-use by the Data Users under the conditions defined in this document |
| INCISIVE AI services | The process of using the AI models over new input data to obtain the targeted results, e.g., predictions or tumour segmentations. |
| MoSCoW | a prioritization technique used in management to reach a common understanding with stakeholders on the importance they place on the delivery of each requirement. The term MOSCOW itself is an acronym derived from the first letter of each of four prioritization categories: M - Must have, S - Should have, C - Could have, W – Will not have. |

| **Abbreviation** | **Description** |
|---|---|
| EC | European Commission |
| TCC | Technical and Clinical Committee |
| WP | Work Package |
| D | Deliverable |
| DIQCT | Data Integration Quality Check Tool |
| HCP | Health Care Provider |
| DP | Data Provider |
| UI | User Interface |
| UX | User Experience |
| TNM | Tumour Node Metastasis |
| HL7 | Health Level Seven International |
| FHIR | Fast Healthcare Interoperability Resources |

| | |
|---|---|
| PACS | Picture Archiving and Communication System |
| SNOMED | Systematized Nomenclature of Medicine |
| LOINC | Logical Observation Identifiers Names and Codes |
| API | Application Programming Interface |
| CDM | Common Data Model |
| ETL | Extract, Transform, Load |
| CI/CD | Continuous Integration/Continuous Delivery |
| LUPI | Learning using Privileged Information |
| XAI | Explainable AI |
| EHDS | European Health Data Space |

# 1 Introduction

## 1.1 Purpose and scope

This deliverable shares useful experiences and best practices that may serve wider use and have wider applicability for similar future research projects/initiatives. We layout the best practices related to INCISIVE major results and all technical, legal, ethical procedures and choices that have led to those. As other deliverables present full details on the final versions of INCISIVE infrastructure, Data Repository, AI toolbox, as well as multiple other results that cover all aspects of their development and sustainability, we focus here on challenges, lessons learned, as well as choices made and their justification. D6.7 summarizes all efforts within INCISIVE and their harmonization that facilitated delivery of the project results.

We extend this deliverable with shared experiences from collaboration with other AI4HI projects of similar scope: EuCanImage (No. 952103, the oncology imaging archive), ProCAncer-I (No. 952159, cancer: prostate), Chaimeleon (No 952172, cancers: prostate, lung, breast and colorectal), Primage (No. 826494, brain tumours in children: neuroblastoma and the diffuse Intrinsic pontine glioma). Acknowledging the value of collective wisdom and derivation of best practices from the wider AI4HI community as opposed to the narrower INCISIVE setting, in the past few months, INCISIVE initiated an effort to collect best practices and lessons learned from all five AI4HI projects. Several partners of these projects, including the Coordinators, joined forces under the lead of INCISIVE partner KU to share practices and approaches through seven focus groups, which we here briefly summarize.

## 1.2 Document structure

The Introduction section introduces the scope and purpose of the document, its structure and relation to other deliverables. In Section 2 the document summarizes the experiences and challenges, as well as best practices within INCISIVE project, and covers design of AI platform, AI services, standardization, data collection and image repository, and best practices from INCISIVE validation studies, legal and ethics management, dissemination and sustainability actions. Section 3 summarizes shared experiences across AI4HI projects and its focus groups. Section 4 provides overall discussion on the achievements, experiences and best practices. Section 5 concludes the document.

## 1.3 Relation with other deliverables

Deliverable D6.7 summarizes the best practices and results achieved during the INCISIVE projects, and their impact and sustainability after the project life-time. The highlighted experiences and challenges, choices made and their justification are thoroughly explained in other project deliverables. D6.7 is thus related to all major project deliverables, resulting from different WPs and all aspects from technical to legal, sustainability and dissemination actions: D2.2 INCISIVE User Requirements - Final Version, D2.3 INCISIVE Scenarios Definition, D2.5 INCISIVE System Design - Final Version, D2.6 INCISIVE Clinical Study Protocol; D3.1 INCISIVE Infrastructure and DevOps Environment Setup, D3.3. INCISIVE Infrastructure - Final Version, D3.4 Standardization Suggestions, D4.x INCISIVE AI Toolbox, Data Analytics and User Services, D5.3 INCISIVE Pan-European Repository of Health Images - Final Version, D6.3 INCISIVE integrated Prototypes - Final Version, D6.4 INCISIVE Training Material, D6.5 Evaluation of INCISIVE Blind studies,D7.1 Initial Data Management Plan, D7.2 Ethical Procedures and Ethics Letters, D8.3 Innovation Strategy - Second Version, D8.4 Preliminary Business Plan, D8.5 Preliminary Operational, Deployment and Sustainability Plan, D9.3 and D9.4 as Dissemination Activities Reports, D9.5 Clustering Events Proceedings and Raising Awareness Campaigns Results.

# 2 INCISIVE Best practices

## 2.1 Shaping the INCISIVE platform

In the realm of healthcare technology development, creating solutions that truly address the needs of end-users is paramount. The journey of shaping the INCISIVE platform, a multimodal AI-based toolbox and health imaging repository for cancer analysis, provides valuable insights into the best practices and lessons learned in designing impactful healthcare technologies that are relevant and reflective of the clinical realities within the specific context.

To shape the INCISIVE platform effectively, several best practices were undertaken, in order establish a better interaction with stakeholders, offering a comprehensive view of the platform's development process, and the reflections on areas for improvement.

### 2.1.1 Engaging Stakeholders

At the core of the INCISIVE platform's design process lies the principle of stakeholder engagement. Recognizing the importance of understanding end-user needs, the project team initiated a series of activities to gather insights from healthcare professionals (HCPs) and other stakeholders. From UX design workshops to a Delphi study, each activity was meticulously planned to facilitate collaboration and co-creation.

### 2.1.2 The Choice of Activities

Firstly, we acknowledge the importance of early stakeholder engagement and plan to involve end-users from the outset of future projects to gather insights more comprehensively and contributing to a shared and sustained value creation. The decision to conduct UX design workshops stemmed from the recognition that user-centric design is critical for the success of healthcare technologies. It is also evident in the literature that such a stakeholder engagement creates the condition for the developed AI to be beneficial not strictly at the end-user's end but more broadly to benefit the society as a whole (Güngör, 2020).

By employing methodologies such as design thinking, the project team aimed to foster creativity and empathy, ensuring that the platform addressed real-world clinical challenges effectively. Additionally, the Delphi study was chosen as a method to gather expert opinions and prioritize features, guiding the platform's development based on consensus and evidence.

- **User-Centric Design Approach:** The platform's design was guided by a user-centric approach, involving extensive collaboration with stakeholders such HCPs and AI experts.

By organizing UI/UX workshops and a Delphi study, the project team gathered valuable insights into user needs, preferences, and challenges. This approach ensured that the platform's features and functionalities were aligned closely with real-world clinical requirements, enhancing usability and adoption.

- **Utilization of Design Thinking Methodologies:** Design thinking methodologies were employed during UI/UX workshops to facilitate problem-solving and innovation. By adopting a solution-based approach and encouraging brainstorming sessions, participants were able to generate ideas and prototype solutions collaboratively. This iterative process allowed the team to explore various possibilities and iterate on design concepts to create an intuitive and user-friendly platform.
- **Iterative Feedback Loops:** Continuous interaction and testing of the platform's user interface (UI) were conducted throughout the design process. This iterative feedback loop involved soliciting input from stakeholders, incorporating their feedback into design iterations, and testing prototypes to validate design decisions. By embracing an agile development methodology, the project team ensured that the platform evolved iteratively based on user feedback and emerging insights.
- **Identification and Prioritization of Features:** Through the Delphi study and MoSCoW prioritization technique, the project team identified and prioritized features essential for the platform's success. By engaging stakeholders in the process of feature prioritization, INCISIVE focused on implementing the most critical features while considering future scalability and sustainability. This approach helped streamline the development process and ensure alignment with project goals and user needs.
- **Addressing Implementation Barriers:** Potential implementation barriers, such as the need for explainability in AI-generated insights, were identified and addressed proactively. By acknowledging and mitigating these barriers early in the design process, the project team ensured smoother adoption and integration of the platform into clinical workflows. This proactive approach to barrier identification and mitigation enhanced the platform's overall effectiveness and usability.
- **Consideration of Long-Term Sustainability:** The design process extended beyond the project's lifespan, considering the sustainability and future-proofing of the platform. By prioritizing features based on their long-term impact and relevance, INCISIVE laid the groundwork for ongoing innovation and adaptation to future developments in AI and cancer care. This forward-thinking approach ensured that the platform remains relevant and valuable beyond the duration of the project.
- **Continuous Workshops:** The purpose of the UI workshops was to gather feedback on the full INCISIVE experience, including features that were currently integrated as well as those planned for future integration. These workshops were attended by consortium-internal individuals familiar with the project's scope and developments, including AI developers and HCP and Data Providers. The feedback collected during these workshops was

considered crucial for refining the platform's design and functionality. Building on the feedback received during the online workshops, in-person workshops were conducted during the first in-person INCISIVE Consortium meeting. These workshops provided an opportunity to present the feedback collected thus far, present new UI options based on suggestions received, and gather additional suggestions and information about the platform. While participants did not raise many concerns about the platform's ease of use, they did provide suggestions for missing functionalities or updates that could enhance the platform's user-friendliness. Based on the feedback received from both the online and in-person workshops, several updates and improvements to the INCISIVE UI were planned. These included the development of a Knowledge Base with glossary and examples of use, dedicated home pages for different user roles, the ability to download blockchain logs, visualization of inference results and explanations, and summary information on search results. Further interactions with end-users were planned to ensure that the platform continues to meet users' expectations and requirements.

### 2.1.3 Learnings from Stakeholder Interactions

Through interactions with HCPs and stakeholders, several key insights emerged, shaping the platform's design in profound ways. Firstly, understanding the intricacies of clinical workflows allowed the team to design features that seamlessly integrated into existing practices, minimizing disruption. Secondly, stakeholders articulated the importance of explainability in AI-generated insights, highlighting the need for transparency and trust in decision-making processes. Thirdly, definition of personas and related specific use case requirements provided valuable guidance, ensuring that the platform addressed the diverse needs of end-users effectively.

### 2.1.4 Study Protocol

These studies included a prospective observational study, a prospective feasibility study, and a data repository evaluation study. The methodological design for these studies has been informed by international good practices, the aims of the INCISIVE project, the expertise of the consortium and the context where the studies were going to be deployed. Key best practices and lessons learned from the study design include:

- **Comprehensive Methodological Design:** The D2.6 provides a detailed and thorough methodology for each evaluation study, including the prospective observational study, prospective feasibility study, and data repository evaluation study. This comprehensive approach ensures that all aspects of the evaluation process are carefully considered and addressed and appropriately integrated, from participant recruitment to data analysis.

- **Alignment with International Good Practices:** The methodology is informed by international good practices in clinical research and evaluation. By adhering to established standards and guidelines, such as those outlined by regulatory bodies and professional organizations, the Study Protocol Definitions ensure the validity, reliability, rigorous and ethical conduct of the evaluation studies.
- **Flexibility for Adaptation:** This flexibility enables the project to adapt to any changes that may arise during the development and deployment of the INCISIVE platform, ensuring that the evaluation studies remain relevant and effective.
- **Ethical Considerations:** The Study Protocol Definitions emphasize the importance of ethical considerations in conducting the evaluation studies. Providing a detailed description of the studies, including their objectives, methodologies, and expected outcomes, enabled recruitment sites to submit proposals to their Ethics Committees for approval. This ensures that the rights and well-being of participants were protected throughout the evaluation process and the obligations of the researchers were laid down in a transparent way.

### 2.1.5 Reflections on Areas for Improvement

While the development process of the INCISIVE platform was characterized by collaboration and innovation, there are areas where we recognize the potential for improvement:

- more frequent communication between technical and clinical partners would enhance collaboration and alignment of goals;
- managing stakeholder expectations and conducting thorough risk assessments are areas where more improvements could have been made, ensuring transparency and accountability throughout the project lifecycle.

In conclusion, the organization of UI workshops has been a critical component of the development process for the INCISIVE platform, providing valuable insights into user needs and preferences. By gathering feedback from both internal and external users, the project team has been able to refine the platform's design and functionality, ensuring that it meets the needs of its intended users effectively. Through ongoing collaboration and feedback, the INCISIVE platform is poised to make a significant impact in the field of cancer analysis and diagnosis.

The journey of shaping the INCISIVE platform offers valuable lessons for healthcare technology developers. By prioritizing stakeholder engagement, embracing user-centric design principles, and reflecting on areas for improvement, we can create solutions that truly make a difference in clinical practice. Moving forward, we remain committed to continuous learning and improvement, guided by the belief that impactful healthcare technologies are built through collaboration, empathy, and innovation.

## 2.2 Data Collection and Preparation

Best practices for data collection and preparation involve outlining guidelines and procedures to ensure that data is gathered, processed, and cleaned efficiently and accurately. The procedures that were followed in INCISIVE for the data collection and preparation, along with the challenges met and their mitigation actions are described in detail in the following sections.

### 2.2.1 Data Collection

An important process in the data collection is to **establish the data collection methods**. This involves the choice of appropriate techniques for collecting data (template + structure), as well as the definition of the types of data to be collected for the intended use cases. In INCISIVE, an iterative procedure was followed to identify the proper clinical and imaging data to be collected that includes the following steps: (a) identification, where an initial data structure and content was proposed (b) review, where this structure was reviewed and commented by the data collectors (c) merge, a consensus was extracted, (d) redefine, where workshops between technical and medical partners were conducted to resolve inhomogeneity issues, (e) standardize, where standardization of the content took place, and, (f) refine, the final structure with the requirements was extracted. This procedure is described in detail in the deliverable D1.3.

During this procedure, some challenges were encountered that needed to be addressed in order to overcome the problems and result the final schema. Referring to the data structure, some inconsistencies were identified in the way that data are collected and stored among the various sites participating in the project, but through the aforementioned procedure the consensus of variables to be included in the study was unified. Through the examination of data in the content standardization step, many inconsistencies were identified in terms of values provided. Each organization uses different protocols to store the data, resulting in many variations between the provided datasets. Some characteristic examples are: the units used for blood examination metrics, the reporting of cancer staging (e.g. TNM or overall staging) or the way that treatment is declared (e.g. drug commercial name against substance). Through this content standardization procedure, all Data Providers should follow the same protocols and provide data in a unified, consistent and homogenized way. Referring to the standardization of annotation labels, a common vocabulary was defined to be used through the segmentation and annotation procedure by all data collectors.

Disambiguation of the clinical terms and of the protocol, e.g. the concept of timepoints, as understood at different sites, or the modalities used in each cancer type and phase, was a necessary iterative procedure, and a shared experience.

### 2.2.2 Data Preprocessing

#### *2.2.2.1 Data transformation: de-identification*

Data de-identification consisted one of the main pillars of the INCISIVE project, since any medical data should first be de-identified, prior to its sharing to the INCISIVE Pan-European repository. Through this process and our efforts, we identified which of our actions helped us achieve a good result and could be considered as best practices for the future and they are described below:

- **Distinction between pseudonymization and anonymization of data and what they include:** One of the first initiatives was to identify what these terms mean and what each of these processes entail, according to existing legislations. It was a common misconception between the stakeholders of what each term meant, as well as what techniques would have to be applied at each case. Through early definition and clarification of these terms, what were their benefits and drawbacks, as well as documentation of the needs of the Data Providers (DPs) during the data sharing procedure, we managed to identify early on which processes, both legal and technical, should be followed and applied in the INCISIVE project. (See also Section 2.7).
- **De-identification protocol formulation:** It was really important to create the de-identification protocol early on in the project to speed up the data de-identification process. During this process, several protocols and standards were taken into account, heavily relying on the DICOM PS3.15 standard. The contribution of AI developers is considered critical in modifying the protocol and adjusting it in a way that data utility is still preserved, while guaranteeing data anonymity. The process of formulating the de-identification protocol should be an iterative process and involve all relative stakeholders, in order to reach to result that satisfies the specific needs of each project, while keeping in mind that the data must be usable while preserving the identity of the data subject. (See also section 2.7)
- **De-identification tools and ease of use:** Even though there is a plethora of available de-identification tools, the tool chosen to perform the de-identification of the medical images, was based on two pillars. The first one was a technical requirement, since the tool chosen should be able to support a pluggable de-identification protocol and should allow it to be modified if needed. The second one appealed to the healthcare professionals, as the selected tool should be one that is easy to install and use, or better, if it had already

been used by some medical professionals in the past. This could greatly reduce any time and effort the healthcare professionals would have to devote to learn any new tool. These factors were also the ones that drove the implementation of the INCISIVE's custom de-identification tool, at a later stage of the project. Hence, it is of great importance to present to the medical and technical teams all the available options and gather their feedback on which tool seems the easiest to use, which one better fits in their needs, and at the same time encapsulates the necessary functionalities, before coming to a decision.

- **Detailed de-identification guidelines:** The preparation of detailed guidelines on how to install and use the de-identification tool was proven to be a very effective solution. Even if the chosen tool had already been used by some medical professionals in INCISIVE, the preparation of a manual on how to use the specific tool in the context of the project was mandatory, as there were some steps that the healthcare professionals had not performed before. The de-identification of the data is part of a larger procedure, the data collection and preparation process, so the provided guidelines also contributed to maintaining coherence throughout the whole process. As a result, we consider the provision of guidelines, as detailed as possible, including visual representations of the different steps, as a necessary means to assist the medical professionals in their efforts and would recommend doing so in similar future endeavours.
- **Online trainings and continuous support to healthcare professionals:** Training of the healthcare professionals on how to use the tool to de-identify the data is highly recommended. Even with the preparation of extended and detailed guidelines, organizing workshops where live trainings on how to install and use the de-identification tool took place, proved to be really helpful for the medical professional. The live interaction between the technical and the medical teams, led to better understanding of the usage of the tool, helped to improve any of the guidelines, that were deemed a bit confusing or not well-described, as perceived from the professionals' side, while solving questions and issues that might come up. Support is expected even after the trainings and encouraging the professionals to reach out for any problem they encountered, created an open channel of communication between the different stakeholders, and dedicated solutions were provided. Continuous support and ongoing communication is of utmost importance in achieving a well-rounded result.

#### *2.2.2.2 Data annotation tool*

This section outlines the lessons learned and recommendations from our experience in data annotation tools and the development of a semi-automatic annotation tool for cancer imaging.

We highlight what was done well, offer recommendations for similar projects, discuss challenges faced and solutions implemented.

- **Thorough exploration and utilization of established tools:** We invested significant time and effort in exploring a wide range of available annotation tools, thoroughly examining their functionalities and suitability for our project requirements. Leveraging ITK Snap for both manual and semi-automatic annotation was a strategic decision, as it provided a familiar interface for Data Providers and offered comprehensive functionalities tailored to cancer imaging annotation tasks. As with the de-identification of data, its annotation was important to begin early in the project. For future endeavours, we strongly recommend leveraging established tools with proven functionalities and user familiarity to streamline adoption and minimize training needs, accompanied by comprehensive support materials and training resources. Future projects should also allocate sufficient time and resources for exploring available annotation tools and collaborating closely with stakeholders to understand and prioritize requirements effectively.
- **Workshops for defining the annotation labels:** Dedicated workshops were held in WP4 for each cancer type, involving both the medical experts and the technical teams, aiming to define the most appropriate imaging modalities, type of annotations and labels that would be assigned. Medical experts' help was valuable in specifying what would be the techniques that would be used for each cancer type and what would be the correct label for each category of lesions. Best practices for annotating data were shared in these workshops between the medical professionals, helping each other and providing insights. The outcomes of these workshops were set of annotation guidelines valuable for framing the work that needed to be carried out, and the organization of such workshops should be considered by relative stakeholders in future attempts.
- **Provision of annotation tool guidelines:** For the annotation procedure, as part of the data preparation pipeline, detailed guidelines were given, while including official documentation and tutorial videos, that accompanied the tool, to facilitate seamless adoption and usage by users with varying levels of technical expertise. Where applicable, it is strongly recommended to provide the official documentation of the existing tools, since they have been prepared by the team that developed the tool and constitute a complete guide for the tool. However, the additional provision of guidelines can act complementary to the existing documentation and can assist medical experts in learning the basics of the tool sooner.
- **Effective communication, training and support:** Workshops, demonstrations, and one-on-one sessions played a crucial role in familiarizing users with the annotation tools and providing ongoing support. Comprehensive support measures, including documentation,

tutorial videos, and online trainings, were instrumental in addressing user queries and ensuring smooth tool adoption and usage. The organization of the trainings were specifically tailored to the needs of medical experts, focusing on practical aspects and real-world applications of annotation tools. Such methods are highly recommended, since they create an open channel of communication between the medical and the technical teams that contributes to faster and better issues resolutions.

- **Close collaboration for implementing a new annotation tool:** One of the main outcomes of data annotation task, was the development of a new data annotation tool. To achieve this goal, collaboration with data providers and medical experts was pivotal in gathering exhaustive requirements, ensuring that the developed tool would address both technical and usability needs effectively. Through this process, and with the acquired knowledge that there are resource constraints when implementing such software tools, the INCISIVE has succeeded in developing a browser-based annotation tool ensuring accessibility and minimizing the need for additional computational resources. Such solution should be considered for similar future endeavours, aiming to assist the medical experts in their everyday life, minimizing the need for having cutting edge resources.

### 2.2.3 Data Quality Assurance

#### *2.2.3.1 Ensure data quality*

Data quality refers to the degree to which data meets the requirements set for its intended use. Its main purpose that data are error-free and relevant to the use cases that they need to serve. Achieving data quality is very important as it affects the reliability of a repository and the effectiveness of AI applications. In INCISIVE, to achieve high data quality, a procedure was followed that includes the following steps:

- **Set the quality metrics that need to be checked.** Data quality metrics are quantitative measures used to assess the reliability of data and they can vary between different domains and projects. Based on the needs and requirements of INCISIVE, the following metrics were selected to be assessed: Consistency, Accuracy, Validity, Completeness, Uniqueness, Integrity.
- **Set the requirements that data should follow to achieve high quality metrics.** This step refers to the definition of the requirements that data should comply with in order to be considered as a high-quality dataset, and other metrics for a basically usable dataset. For each one of the metrics mentioned above, a respective set of rules was defined prior to the data collection, so the data collection can conform to these rules. These rules were

derived from the iterative procedure for identification of relevant clinical and imaging data, as described in section 2.2.1.

- **Perform a quality control upon the creation of the dataset at local level.** This step checks if the data conforms to the predefined rules. In INCISIVE, a Data Integration Quality Check Tool (DIQCT) was developed from scratch. This tool is used by the user prior to the data upload and performs a comprehensive check on the data that includes checks on the structure, content, integrity and consistency among entries, checks for missing mandatory information, checks for duplicate and valid entries, annotation consistency and quality, as well as the de-identification of images both in the DICOM header level and in the image level (burnt-in information). The tool provides reports, through a user interface (UI), with the identified errors and proposes corrective actions that the user should follow in order to provide a high-quality dataset.

This procedure proved to be a successful way to properly define the quality standards and identify quality-related errors in the datasets, as well as support the data provider in correcting the identified errors. However, during the implementation of this procedure many issues and challenges were encountered:

- **Need for a revision**: the requirements set in the beginning of the project in some cases were not precise enough resulting in inconsistencies in the data collection and had to be revised after the first round of data collection. This implies that a more thorough investigation should have taken place from the beginning to avoid such inconsistencies between datasets coming from different sources.
- **Conformity of data providers to the requirements**: many errors identified through the quality control process seem to originate from Data Providers non abiding to the set of data collection requirements. As a derivative of this issue, the compliance of the users with the request for manual correction after the quality control was not always as high as expected, perhaps finding the procedure time-consuming and tedious. One way to address this issue was to make the DIQCT more user-friendly by implementing a user interface so the user can navigate through the various errors that are presented in a graphic and easy-to-understand way. Another mitigation plan that should have been implemented is that the data correction, in cases that this is applicable, should be done automatically by a data curation tool. Such a tool could perform minor corrections, such as converting strings to ASCII code, converting strings to numbers or converting inserted values into the value range based on a similarity check. This way the user would proceed to the major corrections only in a more dedicated way.

#### *2.2.3.2 Data Cleaning and Curation*

Data cleaning and curation are essential steps in ensuring the quality and usability of datasets for research project. Adhering to best practices in these areas helps maintain data integrity and facilitates downstream analyses. Here, we outline the approach taken and highlight key practices employed in the context of our research project. In this section, we reflect on the data cleaning and curation process, highlighting what was executed effectively, recommendations for similar projects, challenges encountered, and alternative approaches considered.

**Best practises that were followed:**

- **Proactive Identification of Issues:** We demonstrated a proactive approach in identifying potential data quality issues by conducting a thorough exploration of the database.
- **Automated Data Validation:** The implementation of a meticulously designed script enabled automated detection and handling of common data anomalies, streamlining the data cleaning process.
- **Iterative Testing and Refinement:** Prior to full-scale implementation, we conducted iterative testing on a smaller subset of the dataset, ensuring the effectiveness and robustness of our data cleaning procedures.
- **Effective Communication and Collaboration:** Seamless communication channels were established with key stakeholders, facilitated by a central coordinating entity, T6.3 and T6.4 task leader UNS, which acted both as data providers and technical personnel, expediting issue resolution and enhancing collaboration.
- **Complexity of Annotation Quality:** Complex issues related to the quality of annotation files required manual intervention by domain experts. To address this challenge, reports detailing such issues were generated and communicated to data providers for resolution.
- **Quarantine Mechanism for Non-Compliant Data:** Designating a temporary repository for non-compliant data allowed for segregation and further validation, ensuring that only clean, validated data were integrated into the main repository.

**Additional Recommendations:**

- **Establish Clear Data Guidelines:** Clearly define data uploading requirements, formatting standards, and folder structures at the outset of the project to facilitate adherence by data providers.
- **Invest in Automation:** Develop robust scripts or tools for automated data validation and cleaning to expedite the process and minimize manual intervention.
- **Prioritize Communication:** Establish effective communication channels and designate a central coordinating entity to facilitate seamless communication between project stakeholders.
- **Implement Quality Assurance Protocols:** Incorporate quality assurance protocols to ensure the accuracy and reliability of annotation files and other critical data components.

- **Allocate Resources for Manual Intervention:** Recognize the potential need for manual intervention by domain experts for complex data issues and allocate resources accordingly.
- **Continuous Quality Monitoring:** Implement continuous quality monitoring mechanisms to detect and address data anomalies in real-time, minimizing the need for post-upload corrections.

#### *2.2.3.3 Annotation Quality Assurance*

Image annotation instructions for all cancer type and each image modality have been set by INCISIVE Data Providers radiologists within WP4 workshops (described in D4.1, and in 2.2.2.2) during the first project year. These workshops have defined:

- what imaging modalities are to be collected for each cancer type
- type of morphological changes to be annotated for each image modality/cancer type
- type of annotation (contour, bounding box) and labels to be used

During the following AI development and validation steps, we have identified some good practices in the image annotation procedures:

- **Medical experts and AI developers are both to be included in the organization of the annotation definition workshops**. This will ensure that most of the different physiological conditions that can be encountered in clinical practise are analysed and their annotation resolved.
- **The scope of the annotation work should be scaled realistically, analogous to the time and efforts available**. Setting the goals too ambitious might result in incomplete annotations, and more errors in the process. Partial and incomplete annotations within a single image are confusing in the AI model learning step.
- **Annotation instructions should be clear and detailed, leaving no ambiguities to the users**. Make sure that certain relevant phenomena cannot be treated under more different labels. Attention should be paid to some accompanying tissue changes with specific instructions about if and how these are to be annotated.
- **In the contour type of annotation, where applicable, define what should be treated as a border** (e.g. How loose the lesion border should be?)
- **Advanced annotation tool options should be used when annotating 3D modalities to facilitate faster annotation process.**
- **In MRI images clear definition of the sequence type and view to be annotated are relevant for consistent annotation process.**

In INCISIVE some annotation inconsistencies were noticed only in the training stage of the AI models. They were only partly due to mastering the technical aspects and annotation options in ITK Snap, and partly due to ambiguities on how to annotate some morphological representations

not specifically mentioned or discussed by radiologists in the initial WP4 workshops. **The lessons learned within this process are:**

- Clear understanding is needed from medical professionals on the relevance of consistent annotations for successful development of an AI model.
- The annotation process for each cancer type should have been continuously monitored, including a living document with encountered problem and ambiguities. This would ensure timely actions and less efforts invested in the annotation corrections.
- Involving more Data Providers and more annotators introduces more variability in the annotation process – assessment and analysis of these effects can be resolved with more annotators per image, if possible.
- More inconsistencies were noticed in the cancer types where radiologists initially defined more labels, making the annotation process more complex and error prone.

More details on different annotation errors and their impact have been described in D4.2. The refined annotation guidelines have been reported in D6.3, as part of the training material to serve as a future reference for Data Providers that will opt to contribute to the INCISIVE repository.

### 2.2.4 Documentation Supporting Data Collection

In order to ensure smooth and efficient data collection process the guidelines have been prepared for each of the data pre-processing steps. The importance of the thorough guidelines has already been highlighted in previous sections, here we summarize the best practices from INCISIVE project related to thorough documentation of the data collection process that include:

- The **step-by-step guidelines for data providers** should be made available as soon as possible, so that they know exactly what data types are to be collected and how to achieve data homogeneity.
- **Clear definition of patient timeline and relative time of imaging and clinical examination** with respect to diagnoses/relapse is mandatory, to ensure the unique time reference.
- **Definition of data structure**: folder structure, file and folder naming conventions and mandatory data types and fields.
- **Providing for image harmonization**: some minimum requirements for DICOM imaging data should be defined (for different imaging modalities and each cancer type) and a separate guide for non-DICOM images is needed (e.g. histopathology images)
- **Clear and precise definition of image annotation procedures and labels** used
- **Step by step install and usage instructions**: guidelines should be provided on how to install and use tools for image de-identification, data quality checking, and annotation.

The lesson we have learned in INCISIVE is **that not all the errors and all possible patient pathways could be efficiently envisioned**. For these reasons during the project, the guidelines had to be

fine-tuned and revised in each data collection cycle encompassing some unforeseen cases and experiences accumulated in large data volumes. In the final project stage, we consider that we have managed to include all relevant parameters, achieving that both guidelines and developed tools could accommodate most of the real life situations.

### 2.2.5 User Training and Process Monitoring

Medical experts from data provider sites have been involved in the data collection setup: the development and evaluation of the INCISIVE clinical data templates and selection of imaging data types relevant for diagnoses and follow up for each cancer type. Despite their everyday practice, collecting the data in a structured and homogenized way is a difficult task. The possibilities of accidental errors are increased when dealing with high data volumes, due to intensive work load on data collection with multiple clinical data entries and new tool for image annotation. Based on INCISIVE experiences some good practices towards more consistent and less error prone data collection are:

- **User trainings should ensure data quality level, privacy preservation and required data structure.** The trainings of medical experts should facilitate the knowledge transfer and improve their understanding of data quality parameters and data structuring conventions, and relevance of these technical details to achieve truly reliable, searchable and reusable repository.
- **Trainings should be focused on the guidelines and developed tools assisting the data collection**. The clinicians should understand how these tools monitor adherence to the predefined data structure and predefined quality parameters. Understanding of these tools and their function will facilitate faster error correction and removal of inconsistences in data structure.
- **For imaging modalities, the training process should include live demonstrations of the advanced tools options** supporting faster annotation of 3D modalities (e.g. MRI or CT scans which contain multiple 2D images)
- The **monitoring of data collection process has to be continuous**, to ensure smooth progression. Keeping tabulated and updated information on data collection progress provides valuable feedback for multiple dependent processes.
- **Data uploading to the central repository should follow a staged approach**. Data is uploaded to a temporary “quarantine” folder, where it is first validated and curated and then transferred, by the responsible partner, to the central data repository (Figure 1).
- **The continuous and targeted e-mail communication with Data Providers** proved as an efficient tool to resolve all issues in error reports, provide continuous support, and more efficient upload process.

The data collection process and INCISIVE pilot studies, both needed continuous support and supervision. **There are several lessons we have learned during these studies:**

- Clear understanding between medical professionals and data and system engineers is crucial in order to design the truly beneficial AI toolbox. In INCISIVE we had a few quite delayed interactions from some clinicians related to service design and UI interfaces, which took much more efforts to be resolved at the later project stages.
- Live meeting and live training sessions are much more efficient, as more interactions and hand-on experience facilitate faster adoption of AI tools, and more useful feedback to system and service designers.
- Collecting clinical data from hospital information systems, where these are usually stored within reports or otherwise incoherently, has been very time-consuming and demanding, since the reporting standards may differ from those used within Data Providers' hospitals. It is questionable if less clinical data volume demanded would result in more clinical data collected.

Training sessions have as well been organized for external parties, to present the process of data collection to potential data donors. Refined and final version of data collection and image annotation guidelines have been published in D 6.4 *Incisive Training Material,* together with the user manual for INCISIVE platform usage.

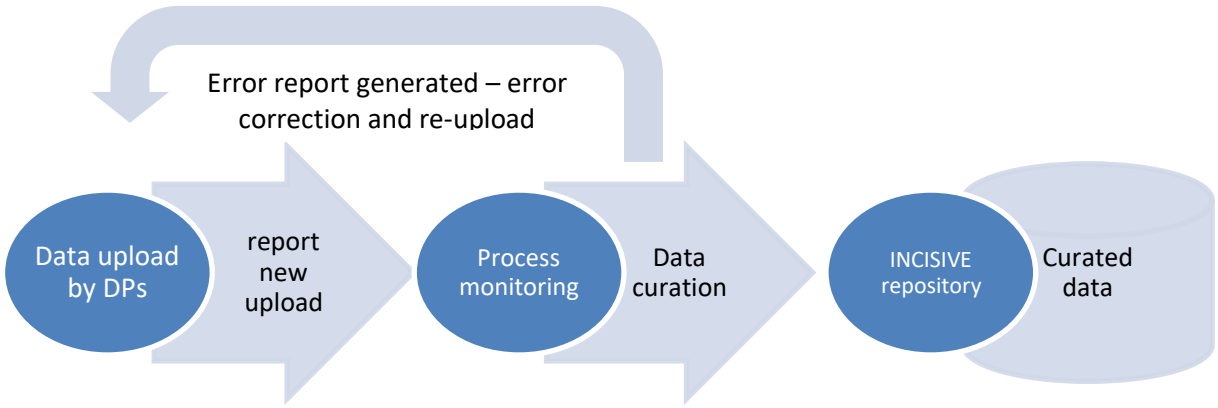


**Figure 1: Data collection monitoring process**

## 2.3 Alignment with Corresponding Standards-Standardization Activities

The clinical data in INCISIVE have been collected within an Excel file that comprises all relevant clinical parameters and terms to be reported. Each DP uploaded these Excel sheets with accompanying images to their Federated Node or in the Central Node in a de-identified way using a patient ID. This mechanism makes it possible to store and send data about multiple patients from the same DP within the same file. Four different Excel file templates have been generated, one for each cancer type with different clinical data considered relevant to describe the diagnoses, characterisation and follow-up of disease. For this reason, it has been decided to create four different HL7 FHIR messages, one for each cancer type.

- IHE standard has been taken into account to define the INCISIVE use cases.
- Medical Images were standardized with DICOM standard and are stored in PACS.
- Clinical data was semantically encoded in a way that gets everyone within a system to speak the same language and understand the meaning of the data. Templates contain input fields about clinical elements and laboratory elements, it was distinguished between these two types of data and used SNOMED CT or LOINC according to the medical concepts. A total of 486 terms using SNOMED CT and 156 terms using LOINC have been encoded. Clinical data was syntactically structured with HL7 FHIR messages for each cancer type. A total of 67 CodeSystem, 13 ValueSets, and 11 StructureDefinitions have been created.
- The need for a clinical report output for each cancer type when running AI services for model inference was standardized with the HL7 CDA standard.
- To store all HL7 FHIR messages, an HL7 FHIR server has been enabled and the necessary queries have been defined to search for these clinical data using the FHIR API and for the medical images the necessary queries have been defined to search for the images via the PACS RESTfull API.

With DICOM and HL7 FHIR, INCISIVE achieves a **Common Data Model** (CDM) that allows receiving, storing, and processing information from multiple sources, multiple Data Providers, as a one standard way.

To share our experiences and best practices related to standardisation we have elaborated a HL7 FHIR INCISIVE Implementation Guide and the Logical Model for each cancer type, see https://fhir.incisive-project.eu/index.html. In order to be aligned with other European Cancer projects, HL7 FHIR is elaborating a European Cancer Implementation Guide where INCISIVE have been participating in order to be aligned as much as possible. In summary, our recommendation is that when HL7 FHIR publishes this European Cancer Implementation Guide, all the European cancer projects should try to review all their HL7 FHIR implementations and adapt it to this common guideline.

For more detail, see INCISIVE_D3.4_StandardizationSuggestions_v1.0_FinalVersion.pdf.

## 2.4 INCISIVE Hybrid Data Repository

Building the INCISIVE hybrid repository of health data involved aggregating and integrating data and resources from multiple sources while ensuring data quality, reliability, and security. It included adopting a strategy to ensure quality assurance during its development but also having in mind its maintenance. To this end, ensuring the quality of INCISIVE repository consisted in:

- taking particular care to data uploading, storage and sharing constraints/requirements, technicalities and processes,
- soliciting feedback from users and stakeholders, conduct regular workshops, meetings and incorporate lessons learned from past experiences to drive ongoing optimization efforts
- ensuring continuous improvement and iterative refinement over time based on the user's feedback.

The strategies/best practices that have been followed in terms of data uploading, storage and sharing as well as with respect to the integration of the different components of the infrastructure are highlighted in the following sections. In some cases, recommendations based on lesson learned are provided.

### 2.4.1 Best Practices for Data Uploading, Storage and Sharing

- **ETL process for different data types** - our strategy involves using two distinct ETL tools tailored for specific data formats. The first tool processes and ingests clinical data from Excel files into the FHIR Server, handling each patient's data carefully through the FHIR server's REST API. The second tool manages the ingestion of imaging data, such as DICOM and NIFTI files, into the PACS server. This dual-tool approach ensures comprehensive coverage of clinical and imaging data, enabling a complete view of patient information on the consortium's platform.
- **Resilience in data ingestion -** a key learning was how to manage unexpected data variations without halting the entire ingestion process. We developed robust error-handling mechanisms to ensure minor issues in clinical data or imaging files don't disrupt the overall system, maintaining data integrity and accuracy.
- **Advanced logging for issue tracking** - realizing the importance of identifying and resolving ingestion issues quickly, we implemented a detailed logging system. This not only helps in swiftly addressing problems but also aids in refining the ingestion process over time.
- **Data Compression and Transfer Efficiency -** our experience highlighted the importance of efficient data management, especially for large-scale data transfers. Employing data compression techniques, particularly for image previews, significantly reduces data size, enhancing system performance and user experience during patient data retrieval.

- **Simplifying FHIR Data Retrieval** - The complexity of the FHIR standard necessitated a solution to ease the retrieval of linked resources. By enriching the ingested data with metadata, we've made it easier to search and retrieve comprehensive patient information, improving usability for all search results
- **Enhancing Federated Search with the Esthesis Component -** A pivotal advancement in our integration efforts is the development of the Esthesis component, a sophisticated tool designed to enhance federated search across the platform. With an instance of its device agent deployed on each node, Esthesis enables the simultaneous retrieval of data from these nodes, playing a crucial role in monitoring and facilitating asynchronous communication with the central platform. This innovation not only streamlines the process of accessing distributed data but also significantly boosts the system's efficiency and responsiveness. By bridging the gap between disparate nodes, Esthesis empowers users with real-time insights and control over their data, marking a key milestone in our journey toward creating a cohesive, interoperable ecosystem.

It is worth noting that the care has been taken about **data security and compliance** in line with strict privacy regulations. We have ensured that all data is handled in a compliant manner. This includes promptly removing unauthorized or temporary data from our database, thus protecting patient confidentiality and meeting legal standards.

### 2.4.2 Integration of AI Engines

- **Containerization -** containerization of the applications is a must for any kind of big system. It is a standard de facto practice currently and it is encouraged to maintain services independent from the infrastructure. This is really important when speaking about AI as the packages keep on evolving and versioning of some packages should be maintained to ensure model compatibility.
- **Create development guides and good practices -** to avoid many issues in integration, we recommend having a clear guideline that states how the software to be integrated should be developed. The guide should cover coding standards, recommended packages, how to use a central code repository (if available and needed) and expected input and output formats. Specially the former requires a big amount of communication, however, once the guide is clear many integration issues are avoided. This was especially important for the integration of the AI Engines as all the input/output needed to be standard to make the services interoperable.
- **Include Continuous Integration/Continuous Delivery (CI/CD ) automation while integrating -** even though it requires more effort than manual checking an integration at the beginning, CI/CD tools enable later faster integration and checking. In order to avoid manually checking every AI Engine to integrate, later in the project, a tool that checked input and outputs was created. With this tool we could check the AI Engines faster and

provide feedback to the AI Developers to correct their code. Ideally these kind of tools should be triggered when a new version of the AI Engine is drafted, this way the repository would tell the developer that they need to correct their code without burdening the integrator in the process.

- **Standardize component interactions through APIs -** as software evolves and developer tools change, it is mandatory to define, for each type of component to be integrated, a clear and standard API. In the case of the AI Engines, we defined an API for the execution of the different use cases and let the developers adapt to it. This way, the developer internally can use whatever tool they are prefer while keeping the compatibility with the infrastructure.
- **Model and code pipelining -** there is the possibility of two different AI Engines to require the same preprocess over the data, therefore it is useful to isolate these processes so that the output of it can be used by multiple applications. On the other hand, there are models that could depend on the results from others, as it is our case. We recommend to use an execution framework such as ArgoWorkflows that enables pipelines.
- **Data Staging Infrastructure -** the process of building a model is not monolithic and it requires of experimentation. Similarly, projects like INCISIVE which take on the challenge of harmonizing very different data sources, require a lot of testing to make sure the process is solid before carrying it out. Because of this fact, it is useful for AI Developers and Data Scientists to have a separate, data storage, processing & staging system where intermediate results such as processed, anonymized or harmonized versions of data can be stored. It's also crucial to separate this data infrastructure from any "live" infrastructure for data federation and learning, as introducing variations in structure, formatting or naming to these systems can cause issues, while at the same time, performing these actions are an inherent part of experimentation with data.
- **Size your resources before starting to work in a federated repository -** even though it is hard, try to set up a minimum requirement list for each federated node. Contrary to cloud environments in which you can scale if you pay more, this kind of system may include federated nodes inside of a hospital premises, with their rules and their space to hosts computers. It must be clear that AI applications require of resources (specially GPUs) to train and perform inference.
- **Federated Learning and Data visualization -** even though Federated Learning is about securing the data by not showing it to the developers, in the end developers need to have a glimpse on the data to understand how to work with it and to debug. This could be done with aggregations that fulfil the privacy requirements and also with the collaboration of the Data Providers if they have a Data Manager that can check particular pieces of information.
- **Federated Learning and Privacy -** federated learning was born from the necessity of training models without sharing the data. This is a perfect fit for healthcare applications

as patient data is very sensitive. However, take into account that to set such system you need an expert on the topic as there are many aspects to cover. The naïve federated learning we followed (McMahan, 2017) was good enough for the project purposes and the "safe" project environment created through the various data sharing agreements that all involved parties signed and committed to. However, it is not considered secure in some scenarios as you need to assume that there are no attackers on the client side nor in the server side. If it is not possible to assume this, security measures need to be put in place, such as secure aggregation and attack detection. Sadly, these kind of measures work on detriment of the model accuracy as usually the measures imply adding noise to the original data to avoid leakages. Finally, even though the process is secured, the final model information leakage should be tested.

- **Federated Learning and non identically distributed data (non-ID) -** given the nature of data, data from different providers can be different. This might not affect a human radiologist, but it might have a big impact on the accuracy of the models produced. This fact is accentuated in federated learning, as we are building models independently in each DP, we might find a situation in which the models do not converge or that are biased towards some concrete population. In order to address these kinds of problems, there are works such as FedMA (Wang, 2020) merging function (https://github.com/IBM/FedMA), even though it is still an open research topic.

## 2.5 AI Toolbox

The design and development of the INCISIVE AI toolbox was a challenging procedure. Multiple actors with diverse were involved in every step of the process, from the data providers to the technical teams and finally to the end-user, HCP. It could be argued that the implementation of such a toolbox was an interdisciplinary, collaborative effort, which highlighted both good practices, as well as challenges during the process. In this section, we will briefly touch upon both aspects, to serve as a guideline for future similar initiatives.

- **Meetings with health care providers (HCP) early in the project were both beneficial and tricky.** A unique characteristic of the INCISIVE AI toolbox is that the Data Providers are also the end-users. INCISIVE organized cancer-type specific workshops with HCPs/Data Providers early in the project's lifecycle to a) understand the needs in the clinical workflow and b) identify the available data. The outcome of the workshops was a list of clinical challenges and respective data and annotation needs. These workshops were an opportunity for medical professionals and technical partners to create a common understanding for the interconnection between data provision and AI service development. To that end, the workshops were highly beneficial. At the same time, however, some of the desired clinical challenges were either hard to solve, or entailed

significant data collection complexity, which put a burden on the AI developers. As a future guideline, organizing such workshops early in the project should be considered a good practice, as long as the respective data types are clearly defined and there is a possibility to collect this data within a given time framework.

- **Data harmonization is the most important part of AI medical service development**. One of the main challenges that the AI developers faced during model development was the high degree of harmonization required between different Data Providers. One cannot expect that Data Providers from different institutions will use the same equipment for data generation. Therefore, harmonization of medical images between Data Providers is one of the most important steps in the development pipeline, and can "make or break" the effectiveness and generalization capabilities of a model.
- **Developing a good model is not enough to create a good AI service**. One of the most important lessons learnt from the INCISIVE AI toolbox was during the integration phase. During cross-consortium meetings during the integration period, a need for additional features to the AI services arose. These features did not require any changes to the models, but mainly "reshaping" information to be better presented to the medical professionals. Through these meetings, we learned how important user experience and information presentation are, especially when the background of the developing team and the end-user differ (engineering - healthcare).

### 2.5.1 Clinical and Imaging Data Analysis

Some selected, more detail observations and lessons learned related to both imaging and clinical data analysis and model development include:

- The **identification of labelled datasets** was one of the main challenges of the model development phase. In particular, models that require annotated images together with clinical data during specific periods limit the availability of an adequate number of data. While INCISIVE constitutes a great initiative for the creation of big multi-site repositories for cancer research, the data collection process was slow, and thus the development of specific models was based on existing open datasets while retraining and testing was performed when an adequate number of data from the INCISIVE repository were available. As a result, for the development of AI models it is recommended to start the development early by identifying existing datasets which cover the requirements for each model, since the data collection process even if it is well designed, might lead to unexpected delays affecting the development process.
- However, even in the case where open datasets were searched, it was identified that there is a **lack of large repositories** that can be used to develop predictive models which address specific clinical challenges. And thus, before the definition of the clinical challenges and

the goal of the predictive models, the AI developers must check the availability of open data or possibility for collection of the required data within the project.

- The **generalizability of the AI models** is a great challenge for all developers. It is a common approach to start the development of a model using a well-defined dataset and as the data collection evolves, the additional data can be used for the retraining of the model or for testing purposes. As found during the development phase, models presented good performance in the training data while when tested on external datasets, the performance was dramatically reduced. This problem originated from the fact that data originated from different clinical sites. Multi-site imaging data are inherently inhomogeneous since they may have been collected using different vendor machines, acquisition protocols, and resolutions, or the region of interest (ROI) might differ. Thus, the harmonization of such data is of paramount significance. Actually, this harmonization pre-processing step was one of the major challenges during the implementation phase. The identification of such parameters as well as confounding factors was also challenging. Apart from image harmonization techniques, such as Hounsfield Units conversion, histogram equalization, or the use of GANs, the radiomics were found to differ significantly. In this respect, feature harmonization approaches were adopted, such as feature scaling and ComBat. Overcoming those harmonization issues has the potential to develop generalizable models and thus maximize the potential for higher acceptance by clinicians.
- Furthermore, the fact that training data might include additional information not always present in INCISIVE data led AI developers to **identify approaches for privileged learning** to use as much information as possible that might be useful and improve prediction performance. Thus, Learning using Privileged Information (LUPI) was adopted, demonstrating its value on model prediction. In addition, some models, such as the lung metastasis prediction model, require a follow-up period which is not always present, even in the training dataset. In this respect, in order to increase the number of data that can be used for training, semi-supervised techniques were used. Their integration in the model development process was found to increase the specificity of the model. Addressing the harmonization issues described above confirmed the generalizability of the developed models.
- One additional challenge found during the testing phase was the **clinical validation of the models**. While the validation of models that address segmentation tasks might be straightforward, the validation of models, such as the prediction of metastasis, is challenging and time-consuming. In particular, for the metastasis prediction model, doctors were asked to evaluate the baseline patient data, including imaging or clinical ones, and to predict whether the cancer is metastatic. The clinical validation results demonstrated that the progression of the tumor in a 2-year period is not always obvious, highlighting the importance of the use of predictive models to accomplish this task. On the other hand, the fact that ML models appear as black boxes for clinicians reduces their

clinical interpretation, while the use of current explainability methods are not able to overcome this limitation. As such, new XAI methods need to be investigated.

- In addition, the **integration of AI tools**, which require not only imaging data but also segmentation files or clinical data, is challenging. This issue arises from the fact that segmentation files might include characteristics not identified previously, such as mask misalignment. In this respect, additional pre-processing steps were required to correct the segmentation masks to the target imaging modality. This process is not possible in INCISIVE data due to the fact that the segmentation files were in NIfTI format while the images were in DICOM. As a result, the alignment between file types, is of paramount significance in order to facilitate correction between the imaging and the annotation files.

Finally, the use of de-identified data is a prerequisite for AI developers, and thus, this de-identification step always initiates the analysis pipeline. Actually, AI developers do not have access to personal data at all. However, the **de-identification process must be designed carefully** by applying the DICOM standard to ensure that no information loss relevant for AI development will occur. This aspect becomes more apparent when information is added as free text in specific DICOM tags. For example, in MRI examinations, there is no universal nomenclature to describe a pulse sequence, such as T1 or T2W weighted sequences. In INCISIVE, this information was deleted during retrospective data collection, and thus, AI developers were unable to process MRI data to develop AI models. Especially in the case of prostate cancer where features from T2 and DWI sequences were fused, it became challenging to retrieve and combine these two series for all cases. While there are tags that could be used to infer the type of sequence, the association is not always correct. Therefore, it is highly recommended that a nomenclature of terms be created to describe different pulse sequences, and this information must be preserved during the de-identification process.

### 2.5.2 Experiences and Lessons Learned Related to AI Explainability Algorithms

Throughout the development of these explainability methods, a myriad of insights has been gleaned, encompassing learned lessons, identified advantages, faced limitations, and recognized best practices. These elements collectively contribute to enhancing the efficiency of the explainability process within AI applications in healthcare. By reflecting on these experiences, the aim is to distil valuable guidance that can streamline the development of transparent, understandable, and trustworthy AI services across different cancer diagnostic tasks. This distilled knowledge not only sheds light on the intricacies of implementing explainability in AI but also underscores the critical importance of a thoughtful, strategic approach to the development and application of AI services in healthcare.

Best practices for algorithm selection and relevant limitations:

- Lung Cancer

- Chest X-rays: The LIME method was selected for chest X-ray explainability due to the complexity of custom-made TensorFlow models, featuring a multi-branch structure. This complexity necessitated a model-agnostic approach, with LIME emerging as the most promising option after extensive research. Although the Layer-Wise Relevance Propagation (LRP) algorithm was considered, its implementation was deemed too complex due to a lack of sufficient information in the current literature. Additionally, while Grad-CAM theoretically could apply to models with a multi-branch structure, in practice, the use of standard libraries does not feasibly allow access and analysis of certain intermediate layers when the model architecture is non-sequential.
- Cancer Staging: For the task of cancer staging, both LIME and Anchors were employed. These methods agree on identifying significant features for predicting cancer staging. LIME assigns numerical weights to each feature, enhancing clarity on their contribution to the model's outcome. Conversely, Anchors facilitate interpretable rules, simplifying outcomes for non-technical individuals, such as health experts. Anchors shine with its use of interpretable rules and reinforcement learning techniques, proving effective for complex, nonlinear model predictions thanks to its model-agnostic nature and parallelization capability. However, Anchors' effective application demands careful hyperparameter tuning and meticulous perturbation function design, highlighting the need for a deep data understanding and the potential necessity for data discretization.
- Metastasis Risk: The flexibility of LIME and Anchors in providing insights across various predictive tasks is noteworthy, effectively used in assessing metastasis risk to demonstrate their ability to enhance model transparency and communicate results efficiently. The combination of LIME and Anchors exemplifies the benefits of using both to address different explainability aspects, offering a nuanced view of the model's decision-making process.

- Breast Cancer
  - BIRADS Classification and Density Classification: Grad-CAM is lauded for its ability to provide global explanations, aiding in understanding how deep learning models reach predictions by evaluating the contribution of each image region. This global explanatory capability is invaluable for gaining an overview of the model's decision-making process, particularly useful in breast cancer diagnostic applications like BIRADS and density classifications. However, Grad-CAM's effectiveness can be influenced by the selection of the last convolutional layer within the model's architecture. While it offers valuable visualizations, they may not provide a fully comprehensive or precise explanation of the model's decision-making process. Strategic planning of explainability methods and modular programming allowed for the efficient use of the same explainability algorithm, Grad-CAM, in both diagnostic tasks, showcasing the

algorithm's versatility and the value of foresight in the selection and implementation of explainability methods.

- **XAI requirements questionnaire**: A critical step towards enhancing the effectiveness and adoption of AI in healthcare has been the development of a targeted questionnaire aimed at understanding the specific requirements of healthcare personnel regarding the explainability of AI models. This approach has proven instrumental in bridging the gap between the technical aspects of AI services and the practical, day-to-day needs of medical professionals. By gathering direct insights into what healthcare providers value and need in terms of AI explainability, it has become possible to tailor the implementation of explainability methods more closely to their expectations and workflow requirements. This focused feedback mechanism has facilitated the selection and refinement of explainability approaches that are not only technically sound but also intuitively understandable and practically useful for medical staff. Consequently, this alignment of AI services with the actual needs of healthcare practitioners ensures that the deployment of these technologies leads to more transparent, comprehensible, and effective diagnostic processes, ultimately contributing to better patient care and outcomes.

In the endeavour to integrate explainability methods into AI services, several overarching lessons and challenges have been identified. The task underscores **the importance of sometimes selecting model-agnostic approaches** due to the intricate nature of certain AI models, which feature complex and non-sequential architectures. This complexity necessitates explainability solutions that can accommodate a wide array of model structures without requiring extensive customization. Moreover, **achieving a balance between technical precision and the accessibility of explanations** has emerged as a pivotal consideration, ensuring that insights derived from AI are comprehensible and actionable for all stakeholders, including those without technical backgrounds. The process also highlighted the need **for planning explainability methods and modular programming in enhancing the development efficiency of explainability mechanisms**. These strategies facilitate the adaptable application of explainability across different diagnostic tasks, emphasizing the need for careful consideration of each method's benefits and limitations. Furthermore, creating a questionnaire to understand the healthcare personnel's requirements regarding the explainability of AI models facilitated the implementation of methods that adapted to their needs. Collectively, these insights forge a pathway toward more transparent, understandable, and effective AI diagnostic in healthcare, advocating for a thoughtful and strategic approach to the development and implementation of explainability methods.

## 2.6 INCISIVE Validation Studies

### 2.6.1 Pre-validation Studies

During pre-validation studies, activities took place to evaluate INCISIVE platform robustness and data quality validation prior to the pilot phase (feasibility studies). These subtasks were positioned as a bridge between development and pilot phases. The overarching goal was to incorporate feedback from this phase into the INCISIVE prototypes that are then used in pilot studies. The overall aim was to develop and establish a checklist or roadmap ensuring system reliability and feasibility across the data repository, platform, and AI toolbox.

For this, a methodological approach was designed to serve the above scope emphasizing on data quality and harmonization and user's feedback, model evaluation and robustness, and system robustness.

The quality of the collected clinical data was extensively examined with respect to completeness, validity, integrity, fairness and consistency of the data in the corresponding Excel templates. In clinical data erroneous insertions during data collection arose as the most significant source of error, mainly due to human factors.

Image validation was based on certain criteria such as:

- adherence to the predefined hierarchical folder structure
- variability of DICOM attributes - DICOM tags were analyzed for potential sources of error and uniformity of the datasets whereas the selection of the attributes was based on literature sources and clinical guidance by experts in the field (HCPs, radiologists). Radiomics-based analysis was considered as the association between the radiomics features and DICOM attributes studied with statistical analysis. Diversity in attributes may result inhomogeneity of the imaging data that in turn may bias the development of the ML models affecting their accuracy and generalizability.
- annotations - the identification of unspecified labels previously undefined during workshops and HCPs guidance dedicated to each cancer type
- anonymization issues
- presence of image duplicates - duplication removal was based both on image content and DICOM tags
- identification of privacy related issues – privacy preservation was tested to investigate and identify presence of images that constitute DICOM reports.

In general, the data were evaluated in 4 different stages, corresponding to the 4 distinct data collection studies: retrospective, training, observational and feasibility studies. The evaluation was performed with Data Integration Quality Check Tool in-house developed tools for overall quality assessment in the central infrastructure.

The experience gained from the pre-validation process can be further exploited when evaluating data quality and harmonization, especially when data is aimed to be used for model development, evaluation and fine-tuning. The key points include the following:

- Validation studies gave rise to the observation that significant challenges exist regarding data trustworthiness when it comes to quality and reliability in cancer repository as in the case of other repositories dealing with medical/imaging data as well.
- The greatest challenge was dealing with multiple Data Providers and cancer types, multiple image modalities and timepoints as well as multisource data and data types.
- During the design phase, it is important to promote the use of standards and avoid over-complicated data schemas.
- Well-defined guidelines during data collection are of great importance.
- The use of quality checks and automated curation helps in most of the problems that arose.
- In the repository, it is important to facilitate an iterative procedure of data quality refinements with automation and human oversight.
- Human factors play a significant role in data collection process (e.g. erroneous insertions during data collection)
- Familiarization of health care professionals with the data collection rules can significantly improve data uniformity and quality.
- Some errors may be avoided from the outset by user training and understanding. In INCISIVE, annotation workshops helped in identifying inconsistencies and disambiguating annotation issues between radiologists and AI developers.
- Feedback and interaction between HCPs and AI developers should take place (in INCISIVE we have implemented error monitoring and reporting activities).
- Finally, despite data quality and reliability efforts, errors, diversity of attributes and sources of bias in real life data are facts which trustworthy AI systems need to recognize and account for.

### 2.6.2 Evaluation of the INCISIVE AI Toolbox – Observational Study

INCISIVE observational pilot study – was envisioned as a prospective data collection aiming at quantitative evaluation of the final INCISIVE AI Toolbox (described in D4.3). In the period of M25 to M35 the Data Providers had an extensive task to recruit the patients, collect and pre-process the data related to the baseline, diagnostic, time point, and follow these patients in time in order to collect as much as possible additional information to facilitate evaluation. Fair performance evaluation requires new data, unseen by the models during the training and model selection process. Some good practices in INCISIVE experience during this study are:

- **Data collection process should start as early as possible**, as the amount of available data is largely determined by the pace of new cases for each cancer type in collaborating hospitals.

- **Additional time has to be allocated for informing and consenting patients.** Despite the study is not interventional, patients are stressed and more vulnerable, as being faced with severe diagnoses.
- **As data collection takes place in real settings, accumulated delays are common.** Good mitigation strategy used in INCISIVE was to use all delayed data from the previous collection stages for model evaluation, not used during the model training and selection. Thus, at each project stage we took advantage of all eligible INCISIVE data resources (data not previously used), including open datasets.
- **Both development and evaluation of the models** in the setting with multiple data providers, enriched with extensive open data sets, **stressed the importance of data quality at the input and relevance of heterogeneous data volumes**. Experience gained during the observational study facilitated further improvements in all developed tools and routines related to quality assurance:
  - Data Integration Quality Check Tool (DIQCT) final improvements to track more rigorously adherence to data collection guidelines to ensure structured and standardizes way of clinical data upload.
  - development of the curation scripts, the automated tools to control the folder structure, correct certain folder structure issues and generate error report as a feedback to data provider with instructions on error correction.

In INCISIVE we have been dealing with four cancer types, disease characterization using different imaging modalities and different morphological presentations. These ambitious goals required collection of multitude of different types of data from multiple sources, which was not always perfectly aligned in time with the AI model development milestones. **What we have learned from the flow of the observational study is:**

- data collection process has to be conceived and initiated as early as possible in the project lifetime, **with ensured technical support related to data manipulation in partners acting as Data Providers,** as these skills facilitate smooth data collection progress
- **data harmonization step is crucial for the development of AI models with increased generalization capabilities** and attention should be placed to the timely harmonization of the imaging modalities.
- **Data curation tools and data and image quality assurance processes are crucial in the model development, and have significantly improved overall quality of the INCISIVE repository**. While these tools should be developed preferably before data collection onset, it is very difficult to predict all error sources in realistic settings, thus fine tuning these tools during data collection is usually inevitable.
- having in mind duration of both the project and the observational study, **focus on one or two cancer types** might have resulted with more aligned data types and more mature models.

### 2.6.3 Feasibility Study

This section focuses on the best practices employed during the feasibility study to ensure a streamlined process. By documenting and sharing these experiences, we aim to provide valuable insights that can inform and enhance the success of similar endeavors in the future.

- **Conducting biweekly meeting.** One of our best practices involved conducting bi-weekly meetings with all partners involved to review progress on tasks. This approach ensured that everyone was on the same page, fostered a deeper understanding of progress, and provided crucial support to those who required assistance. These regular meetings served as a platform for collaboration, alignment, and mutual support, ultimately enhancing our collective efficiency and effectiveness.
- **Obtaining ethical clearance in the early stage**. All partners were proactive in obtaining ethical clearance during the early stage of the INCISIVE project, highlighting their commitment to ethical research practices. This collaborative effort ensured that approval was secured for all studies conducted, meeting our target efficiently. Hence, it was easy to establish a foundation of trust and integrity, enhancing the credibility and validity of all research studies.
- **Ethical oversight and collaboration with data providers.** As part of the collaboration with Data Providers, tasks include clarifying certain details, such as terminology or specific requirements. Additionally, facilitating the translation of approvals or relevant documentation into English, when necessary, ensures clear communication and understanding. Moreover, assistance was provided in the process of resubmitting ethics amendments by coordinating with relevant partners and following up on procedures to ensure compliance with ethical standards. Close communication and collaboration with Data Providers are integral to addressing any ethical concerns promptly and effectively.
- **Brochure for accelerating patient recruitment.** The team devised a brochure aimed at expediting the patient recruitment process. Feedback from Data Providers revealed that patients visiting clinics or hospitals often lacked the time to read lengthy participant information sheets thoroughly. To address this, a concise brochure was prepared, summarizing the study's key concepts. Placing the brochures at the entrance of the door where patients first visit made it convenient for them to access the information. The catchy titles of the brochures helped grab their interest and encouraged them to learn more about the study and give consent. This approach proved effective in capturing patient interest and conveying essential study information efficiently.
- **Conducting training for healthcare professionals.** Through collaboration with UNS, hands-on training sessions were conducted in advance to familiarize HCPs with the INCISIVE platform. This proactive approach equipped them with the necessary skills and knowledge to effectively navigate the platform during the actual feasibility study. Planning such programs well in advance ensures that HCPs are adequately prepared, enabling them to

perform their tasks more proficiently. By providing this sort of comprehensive training, we enhance the capability of HCPs to contribute effectively to the study's success.

- **Utilizing Microsoft outlook forms.** To streamline the consent process and expedite document handling, digital signatures were integrated into Microsoft Outlook forms. This approach eliminated the need for manual signing, scanning, and emailing of consent forms to the research associate at KU. Furthermore, the evaluation tools for the feasibility study were incorporated into these forms. By leveraging this method, the workflow was simplified, and outputs were conveniently generated in Excel format, facilitating easy analysis and time-saving benefits. Additionally, the monitoring of the Incisive user experience was enhanced by creating another form. This form allowed HCPs to provide updates on the progress of INCISIVE testing. These advancements highlight a best practice, demonstrating a time-saving solution that could prove beneficial for future projects.
- **Providing training handbooks.** The preparation of handbooks for both Healthcare Professionals and Data Providers significantly contributed to the smooth execution of the feasibility study. Offering comprehensive training handbooks to HCPs offers numerous advantages. Firstly, it serves as a valuable refresher, ensuring that HCPs retain essential information covered during their training sessions. Additionally, the handbook provides a systematic, step-by-step guide for conducting testing procedures. By detailing each task, it enables HCPs to navigate the testing process smoothly, minimizing errors and uncertainties. Moreover, the handbook serves as a reliable reference tool during testing sessions, offering guidance and clarification to HCPs encountering challenges.
  Similarly, for Data Providers, the handbook provided instructions on pre-study preparations and outlined their role in supporting HCPs during the study. This comprehensive approach facilitated clear communication and boosted confidence among both HCPs and Data Providers, thus playing a crucial role in ensuring the success and efficiency of the testing process.
- **Development of monitoring tables.** Creation of monitoring tables facilitated the monitoring and tracking of patient recruitment and feasibility study progress. By providing everyone with access, the table ensured transparency and allowed partners to easily visualize the study's advancement.
- **Enhancing study insights with qualitative interviews alongside quantitative measures.** Incorporating face-to-face interviews and qualitative components into the feasibility study was essential. While quantitative tools provided valuable data, they alone could not fully capture the full experiences of participants. To delve deeper and obtain qualitative insights, individual qualitative interviews were organized. These interviews offered a more profound understanding of the real experiences and perspectives of the HCPs, enriching the study with valuable qualitative data. This hybrid approach, combining quantitative with qualitative interviews, ensured a comprehensive and holistic analysis, capturing both numerical data and human experience.

- **Presenting results at conferences and awareness events.** Presenting the results at conferences and awareness events was another one of our best practices. This approach helped in obtaining suggestions and identifying areas for improvement. It provided a platform for sharing insights, receiving feedback, and fostering collaboration with stakeholders. Overall, presenting findings in awareness events facilitated continuous learning and refinement of our strategies.

## 2.7 Legal and Ethics Management and Data Sharing Practices

The best practices faced by INCISIVE in the work on the definition of the data sharing legal framework, as well as lessons learned from this process are outlined in detail in D7.3. They include:

- **Need for thorough discussion about manner and implications of selected approach to de-identification of health personal data is paramount. Moreover, caution in sharing even anonymous data is advised, as advancement of technology may lead to new means of re-identification of previously anonymized information.** The terms "anonymous information" and "pseudonymized data" are clearly distinguished by the applicable law (General Data Protection Regulation). However, in practice the understanding of the terms 'anonymized' and 'pseudonymized' both among privacy practitioners and in the medical community is not uniform. Moreover, the translation of the legal requirements for data anonymization into concrete technical standards is not an easy task. At the same time, implications of the assessment of a data set as 'personal data' (or non-personal data) under GDPR framework are crucial, as they determine the applicability of GDPR or the lack of it. Still, the paradox is that if the data is considered anonymized during its initial submission and it is determined that GDPR does not apply to it, the data theoretically can be shared with a wider public. If this sharing is done without any controls and restrictions, this may lead to the risk of re-identification. With this in mind, INCISIVE legal framework was designed to protect all the contributed data (even if considered as 'anonymous' during the data submission), taking into account the sensitivity of the shared medical information and potential of linking the data to the patient (for example, with use of future technologies). In the experience of INCISIVE, the 'Registered Access Plus' model was considered as more appropriate than a totally 'open' one to ensure that the Data is used for ethically and legally permissible purposes only. This solution was accepted by the INCISIVE Data Providers and Data Users.
- **The perception of GDPR roles (controller, processor or joint controller) of the partners and participants in data sharing, also in the context of research projects, is not straight forward.** The project also faced discussions regarding GDPR status of the individual stakeholders in the data sharing. In particular, for the sharing of Data between a defined group of beneficiaries (i.e. Data Providers and Data Users) bound by a clear purpose of

implementation of the DoA and using defined means of processing, the joint controllership arrangement seemed most appropriate. In the experience of the project, this agreement model proved effective to provide legal safeguards and yet flexible enough to allow necessary sharing of Data between the partners. However, as the pool of the Data Providers is planned to expand to Data Providers from external organizations, the roles need to be changed, as described in D7.3.

- **Striking the right balance between the privacy of the Data and its usability is a multifaceted task**. It is a known dilemma that too stringent security measures likely lead to data non-use. Yet, from the GDPR compliance perspective, privacy-by-design principle requires to minimize the amount of data processed. To tackle these conflicting positions, INCISIVE conducted multiple consultations and discussions between the beneficiaries regarding scope of the data required for AI Models training and the manner of privacy compliant data sharing. One of the most heavily discussed aspects was the requirement for the data/images to be visible (viewable) to the AI developers. In INCISIVE, the interviewed Data Providers were not in general concerned about visibility of the de-identified images and Data to the approved Users, assuming they use the Data for legitimate purposes within the boundaries of the Platform. The solution which was selected was a restricted sandbox hosted in each Federated Node, through an installation of an optional viewer component. Additional safeguards may be added in the future, if required. A lesson which may be learned from INCISIVE experience is to start these discussions early in the project and to dive deep into specific technical requirements of the users during the product design studies.
- **There are limitations stemming from the ethics approvals and consents of the patients.** One of the most important requirements for the Data Providers to be able to participate in the data sharing is ensuring that the Data can be contributed in compliance with statutory requirements. This typically entails securing the approval of the hospital's bioethics committee and consents from the patients (unless other GDPR legal basis applies). In INCISIVE, the Data Providers obtained ethics approvals, which allowed them to submit Data to the Project for the implementation of Action (see Deliverable D7.2). We can conclude that it was helpful for the Data Providers to rely on coordinated efforts for drafting ethics applications and patient consent forms based on a common template.
- **The evolving regulatory landscape may have significant implications for research projects and decisions on sustainability, in particular in relation to data sharing platforms and repositories.** During the course of INCISIVE project, the legal landscape has significantly changed, due to adoption of Data Governance Act, Data Act and other EU acts. Further changes are expected when European Health Data Space regulation is passed. Those requirements had to be carefully considered during project's sustainability planning. As a result of these discussions and practical challenges in establishing a

sustainable stand-alone data repository, the project has closely looked at synergies with other existing projects and is working on collaborated with EUCAIM to ensure compliant and sustainable sharing of the project data.

**Other lessons learned which may be helpful for data providers, data users and other partners involved in data sharing repositories include:**

- Design and development of the data sharing framework is a gradual and tiered process, which involves not only analysis of the legal framework, but also a thorough understanding of the purpose and functioning of the data repository, as well as expectations and limitations of data providers (including patient representatives) and data users.
- Legal work, data collection and AI training in the project cannot be considered in silos. Frequent discussions, in person meetings and workshops, as well as questionnaires collecting views and needs of the involved stakeholders have been proven to bring the discussion forward and help with the decision making process.
- Including legal team in various discussions about the technical design of the data sharing solutions and seeking legal input is a requirement for implementation of the privacy-by-design principle.
- The builders of the repository and also data providers need to understand and – if needed- openly discuss the possible limitations and consequences of selected method of de-identification of the data (pseudonymization vs anonymization). The use of the term 'anonymized data' if done incorrectly can have important legal and data usability implications.
- Data providers need to have a good understanding of the whole 'life cycle' of the collected medical data to be able to appropriately prepare consents and ethics applications. This means that planned use of the data within the project and beyond it, especially potential re-use of the data for other purposes, should be discussed early in the project and should be described in the platform documents.
- To design and test the repository intended for storage and use of personal data, the beneficiaries need to consider putting in place agreement(s) for the transfer and use of the data. Those agreements need to be tailored to the specific circumstances of the project. Also, as the circumstances (for example, place of storage of data) may evolve, thus the project needs to anticipate the need to update the agreements, if needed.

## 2.8 INCISIVE Dissemination and Sustainability

### 2.8.1 Sustainability and Exploitation of INCISIVE's Main Results

To guarantee its financial self-sustainability, the INCISIVE AI tools can be commercialised in the future, after the required further development. According to the project's preliminary business plan, it will be primarily sold as a cloud-based solution through a recurrent subscription, targeting hospitals in the large EU countries where the INCISIVE consortium has a presence (e.g., Spain, Italy, Germany) as the initial customer groups. The AI tools developed during the project will require further development after the end of the grant in order to be services that are ready to be commercialised. Therefore, project partners are working to devise an exploitation strategy that will consist of several phases: initially, after the required further development is done, the INCISIVE AI toolbox should undergo conformity assessment as a medical device and afterwards can be offered to radiologists, who will be the main users of our solution and may act as brand ambassadors to help introduce INCISIVE in the market. Accordingly, we will discuss how the value proposition of the AI services needs to be adapted to serve the needs of these stakeholder types. Further phases of the exploitation plan will include the sale of INCISIVE's AI services within the AI for medical health.

Concerning the Data Repository sustainability, the main lessons learned, and best practices are:

- Ensure project initiatives align closely with established guidelines and core objectives to maximize impact and sustainability.
- Maintain a clear focus on overarching project goals and objectives throughout transitions and adaptations.
- Recognize the need for transitioning existing projects into sustainable entities, such as EUCAIM, to ensure continued impact and relevance beyond initial funding periods.
- Leverage transitions as opportunities to refine and improve project structures and objectives in alignment with evolving needs and standards.
- Prioritize adherence to ethical, legal, and privacy standards to ensure data integrity, security, and user trust.
- Emphasize the importance of enhancing data accessibility to facilitate research, innovation, and collaboration within relevant fields.
- Implement mechanisms to streamline data sharing while maintaining compliance with ethical and legal requirements.
- Recognize the significance of interoperability in fostering collaboration and innovation across multiple projects and initiatives.
- Harmonize project components and standards to ensure seamless interoperability and accessibility.

- Secure funding mechanisms to support the long-term sustainability of projects and initiatives beyond initial implementation phases.
- Maintain technical sustainability by preserving original project architectures and components post-project completion.
- Consider releasing components under open-source licenses to promote transparency, accessibility, and community engagement.
- Prioritize secure data storage practices to safeguard sensitive information and promote trust among users and stakeholders.
- Encourage the reuse of data through transparent and accessible mechanisms to maximize its value for research and innovation purposes.

### 2.8.2 Dissemination of INCISIVE's Main Results

INCISIVE has developed effective communication tools about the project and conducted a wide range of Communication and Dissemination (C&D) activities aimed at fostering general awareness and disseminating the project's outcomes. The involvement of INCISIVE partners in C&D activities has been excellent. They have made significant contributions to WP9 by publishing scientific articles in peer-review journals and conferences, actively participating in external events, sharing project-related content through various channels including websites and social media, and utilizing project materials properly. Simultaneously, the project has consistently nourished its website and associated portals with high-quality content, and established a dynamic network of professionals through strategic use of social media, mailing lists, events, and meetings.

#### *2.8.2.1 Dissemination Best Practices – Events*

INCISIVE partners engaged in various events aimed at enhancing the project's visibility, drawing stakeholders to the consortium, and disseminating the project's results. Additionally, they co-organized clustering activities with the Artificial Intelligence for Health Imaging (AI4HI) cluster.

The dissemination activities were addressed to the main target stakeholders and audiences defined at the project's beginning and described in the D.9.2. Communication & Dissemination Plan. These are healthcare professionals, medical imaging data providers, the industry, the academic and research community, and European health policymakers and national agencies. Additionally, the project also addressed messages to the general public, patient associations, NGOs supporting cancer patients and civil society, relevant European projects and major European and international associations and institutions of radiology.

In total, INCISIVE partners have participated and make contributions in more than 65 events and more than 10 presentations in industrial events. They have also co-organized more than 15 joint

actions with the AI4HI projects and more than 10 clustering events. Among all these actions, we would like to highlight five events in which INCISIVE results were presented to multiple audiences:

- **"AI in Cancer: Unleashing Opportunities, Overcoming Challenges" workshop event**: Madrid, 7th of November, 2023. 240 professionals from more than 20 countries registered for this workshop event coorganized by INCISIVE, EUCAIM, and the AI4HI cluster. The workshop joined clinicians, researchers, AI developers, lawyers, innovation managers, and other professionals interested in oncology and AI working in hospitals, universities, research institutes, research infrastructures, and private companies. During the event, speakers and attendants jointly explored how AI powered by health data sharing can transform cancer challenges into remarkable opportunities.
- **'Health data sharing and AI in cancer imaging – empowering AI-driven solutions for cancer diagnosis, treatment and follow-up' event**: Belgrade, 17th and 18th of November 2022. More than 60 healthcare practitioners and researchers from the field of oncology and AI joined this INCISIVE clustering event. The participants had the chance to discuss about the need of AI for cancer diagnosis, treatment, and follow-up and to deepen their knowledge about different aspects that were being considered within the INCISIVE project, such as the process of data sharing and its benefits, the legal framework and challenges involved, medical image (dicom) de-identification tools, and image annotation tools and guidelines, among others.
  The meeting was a great opportunity to raise awareness in the region about the solutions of the project and to receive feedback from professionals at the different INCISIVE workshops. The event also included sessions about other European-related initiatives clustered under the umbrella of the AI4HI cluster.
- '**Research to Reality – Digital Solution for European Challenges' event:** Brussels, 5$^{th}$ February, 2023. INCISIVE's coordinator, Gianna Tsakou, was one of the invited panelists by the European Commission for the session 'AI serving Healthcare', focused on how advanced digital tools, in particular data, AI, visualisation and new hardware, are benefiting patients, hospitals, intermediaries and governments. The session was organized in the context of "Research to Reality – Digital Solution for European Challenges", a promising event under the Belgian Presidency in cooperation with the European Commission. The event was attended by many policy-makers, as well as researchers, industry and many more stakeholders. During the panel meeting, the INCISIVE coordinator Gianna Tsakou shared experiences gained during INCISIVE's lifetime on ethical health data sharing and AI development. She also discussed how collaboration with the EUCAIM project is contributing to moving the project results from Research to Reality and full deployment.
- **Mini-symposium at the Annual International Conference of the IEEE Engineering in Medicine and Biology Society:** Sydney, 24$^{th}$- 27$^{th}$ July, 2023. Representatives of the AI4HI

cluster and EUCAIM organized a mini-symposium entitled 'The European Cancer Imaging Initiative – Status, Challenges and Opportunities'. The event provided details on the achievements of the cluster in creating large interoperable cancer imaging repositories, and their experiences in developing novel AI-based models. It also offered hands-on experience on the use of the FUTURE-AI guidelines and reported on the challenges addressed by the EUCAIM project. It was an excellent opportunity to disseminate the cluster's joint efforts and foster networking with Asian and Australian audiences.

- **Awareness event on INCISIVE and TRANSITION EU funded projects: Cyprus, 27th February, 2024.** The event aimed to foster dialogue and understanding around these innovative initiatives, driving forward progress and collaboration in our community.

#### *2.8.2.2 Dissemination Best Practices – LOI*

Best practices concerning the letter of intent initiative to promote and engage data sharing from potential data providers are:

- Leverage events as opportunities to connect with potential collaborators and stakeholders who have shown interest in the project or relevant activities.
- Capitalize on the momentum generated by events to initiate discussions and solicit participation through mechanisms like Letter of intent (LOIs).
- Use reminders and additional communication to reinforce the importance of collaboration and the benefits of participation.
- Recognize the value of collaborative networks within larger project clusters or consortia for expanding reach and fostering partnerships.
- Share relevant documents, such as LOIs, with other projects or initiatives within the same cluster to capitalize on shared objectives and interests.
- Utilize project websites and online platforms to disseminate important documents and information, such as LOIs, to a wider audience.
- Distribute content through various channels, including partner networks, social media, and relevant stakeholders, to maximize visibility and engagement.

# 3 AI4HI – Shared Experiences and Practices

During the project duration INCISIVE had an extensive collaboration with other AI4HI projects of similar scope: EuCanImage (No. 952103, development of an oncology imaging archive), ProCAncer-I (No. 952159, cancer: prostate), Chaimeleon (No 952172, cancers: prostate, lung, breast and colorectal), Primage (No. 826494, brain tumours in children: neuroblastoma and the diffuse Intrinsic pontine glioma). This collaboration was well defined and initiated from the very beginning to facilitate sharing experiences on the most critical common topics.

## 3.1 AI4HI Focus Groups

The AI4HI continuous collaboration was structured into eight focus groups that met monthly or bimonthly to cover major common concerns related to:

- ELSI: Ethical, Legal and Social Issues
- Metadata Models (Imaging/Non-Imaging Data**)**
- Data Storage/Curation/Management
- AI Development
- AI Validation
- Clinical group
- Annotation group
- Dissemination and Exploitation group

**Participation in these focus groups brought many valuable shared experiences and exchange of opinions and good practices.** Discussions on approach to the **Ethical, Legal and Social Issues** have paved the solid way facilitating all projects to move forward with data collection in a privacy preserving way, respecting all legal and ethical considerations. The best practice from all the projects, including INCISIVE, is inclusion of legal advisor as a consortium partner, as the challenges met needed to be thoroughly analysed and a tailored solution had to be proposed having in mind all legal and ethical constraints. The discussions in the ELSI group led to involvement of the legal and ethical partners is various joint publications and seminars, allowing them to share best practices between the projects.

Participating and actively contributing to the **AI4HI data storage/curation/management** has enabled INCISIVE benchmarking with other EU health data repositories, sharing best practices and adopting a common roadmap in the perspective of EHDS. In particular, it has allowed:

- collaboratively addressing challenges, driving innovation, and achieving shared goal of enhancing healthcare through data-driven insights,

- highlight the good of different approaches for data storage: central versus federated versus hybrid,
- Ensure right choices with respect to structural and semantic standards for INCISIVE common data model in the perspective of repository and data interoperability,
- Share best practices and common understanding with respect to legal compliance in the frame of GDPR anonymization/de-identification/pseudo-anonymization,
- Adopt a common roadmap in the perspective of the EHDS,
- Increase our dissemination activities and visibility. One of joint publications: "Data infrastructures for AI in medical imaging: a report on the experiences of five EU projects", DOI: 10.1186/S41747-023-00336-X received a certificate for the third highest number of article downloads in European Radiology Experimental in 2023 that reflects the high interest of the scientific and health communities with regards to our common activities.

Another solid example is the **AI4HI technical validation working group** that discussed the differences and similarities between technical and clinical perspectives in the technical validation of AI. A questionnaire was collaboratively designed and used to collect feedback on a number of AI validation topics by both clinical and technical consortium members. This work resulted in a report and a paper (not yet published) reporting a number of observations regarding for example inhomogeneities in the way biases are examined, or in the way training/testing datasets are handled for technical validation. It also highlighted the need for common AI understanding, better AI training and team multidisciplinarity in AI development and validation. The axes of discussion were revolving also around trustworthy AI, having in mind the FUTURE-AI initiative, that proposed guidelines for trustworthy AI based on six guiding principles: Fairness, Universality, Traceability, Usability, Robustness and Explainability (https://future-ai.eu/). At a later stage, the group was merged with the Clinical Validation group, to benefit more from multidisciplinary participation, and discussions continued around the ways XAI and fairness issues are conceptualised and implemented in the projects, and around potential publications tackling the challenges of AI validation.

During the first 2 years of the projects life time **Annotation AI4HI group** was mainly active, when in each project relevant decisions had to be made regarding the protocols for data acquisition, annotation and pre-processing. Through the regular meetings the group has summarized approaches and experiences used by all projects into a document with the final aim to propose ontologies, standards, tools and recommendations to help the medical imaging community to perform automatic high-quality annotations in medical images. The initial document summarizes different types of data annotation, issues and challenges in the medical imaging domain: (1) heterogeneity in data formatting, (2) meta data quality and harmonization, (3) image quality and harmonization, (4) human-based annotation, (5) AI based annotation, (6) metric to evaluate quality of annotations, (7) metrics to evaluate inter-reader variability, (8) most commonly used

annotation tools in medical image community. This document as well shares the practices applied in each of the five projects.

In the second project duration half, very important joint discussions within **AI development group** enabled sharing good practices related to each step during the AI model development from data pre-processing to model evaluation. These steps and suggested approaches to be used are summarized in a questionnaire that was circulated gathering different aspects and opinion from AI developers in contributing AI4HI projects.

## 3.2 AI4HI Overall Experience, Challenges and Lessons Learned

***An additional effort, beyond the work envisioned by the INCISIVE project*, has been the initiative of Kingston University, INCISIVE partner, to design, conduct and summarize the results of the study that would convey the experiences from members of the special AI4HI groups.**

Experiences and lessons learnt were collected via interviews with a total of 24 stakeholders actively engaged within the AI4HI cluster (data scientists, clinicians, technical coordinators and legal partners). The interviews tackled the overall experience of participants with their respective projects as well as more specific questions on the processes and logistics of the work with a focus on identifying challenges and lessons learnt. Thematic analysis followed the interviews and consisted of identifying themes that emerged from the interview itself through a process of open coding and categorization. The most prominent themes were as follow:

- **Theme 1: Speaking the same language**
  - Understanding each others was the most challenging part of the project according to all participants and meanings of different principles and terminologies varied significantly across disciplines.
- **Theme 2: Alignment of needs across disciplines**
  - Bridging the gap among all participants from different backgrounds and the proper alignment of needs and expectations in the early phases of the project was considered crucial for the delivery of an effective and useful AI model and is application.
- **Theme 3: Best practices and lessons learnt -** Important recommendations were shared around good practices and lessons learnt as described below:
  - Presence of legal partners: involving legal partners from the early stages of the research proposal and treating them as researchers on the project and not only consultants. The importance of teaming up with legal experts who will help shaping the research questions and strategies was highly emphasized by all participants.
  - Exchanging knowledge: constant communication to share knowledge and skills was highly praised. Adopting a meeting strategy that fosters involvement of participants

for less time and keep stakeholders well engaged via in-person or more focused meetings when possible was recommended.

- Adopting proactive approaches and choosing solutions (standardized when possible) in addition to putting in place dynamic contingency plans were measures used to facilitate the workflow.
- Acknowledging and accounting for local and regional study-sites differences. Opting for harmonized solutions with no contradictions instead of uniform ones was proposed to particularly overcome differences in laws and regulations across the study sites.
- Inter-professional collaboration and team work was described as successful when members are committed, research-oriented and appreciate the role of technology in advancing clinical practice.
- Anticipating the need for new specialties that can bridge the gap between technical and clinical experts and getting them employed meaningfully in new roles to help advance the digital space in clinical departments.
- The final recommendation entails the importance of reusing and recycling the work done to ensure continuity of the project and avoid replication of efforts or suboptimal use of resources.

# 4 Discussions

INCISIVE project delivers three main outcomes: 1) INCISIVE clinical and imaging data, 2) INCISIVE AI toolbox, both facilitated through 3) INCISIVE hybrid data repository. To pave the way to these resources there were multiple pathways thoroughly evaluated and examined under difficult constraints of dealing with very sensitive data, requiring the continuous commitment to preservation of privacy, legal and ethical concerns. We have staged our progress, developing more and more improved prototype gathering all experiences collected from multiple data collection sites and AI model development process. In order to share our experience and lessons learned, we have delivered the material with thorough explanations of all components comprising INCISIVE hybrid repository: data preparation and curation tools, data storage and search functionalities, the data sharing mechanism, and the federated space facilitating integration and deployment of AI engines.

In this deliverable we share lessons learned and some best practices behind INCISIVE results related to all comprising segments, covering it from multiple perspectives: system and data engineering, clinical, legal and business. Additional attention is placed on best practices in dissemination of INCISIVE results on multiple levels: from scientific communities to wider audience.

Best practices on the INCISIVE system design relying on prioritizing stakeholder engagement and embracing user-centric design principles indicate a need for a clearer communication channels between technical and clinical partners to align their expectations and alignment of goals. In order to ensure a satisfying result, the stakeholders must be continuously engaged in the other stages of system development to express not only their needs, buts as well to offer their expertise ensuring these expectations are met in the final prototype.

INCISIVE data make central contribution to the scientific community given the ongoing work toward the European Health Data Space (EHDS). With 9 Data Providers from 4 countries we have faced heterogeneous data streams, that needed clear data collection methodology, well defined data structure and definition of the relevant data types. Through refinement, redefinition and continuous improvements, aligning with relevant medical standards, INCISIVE teams have managed to achieve data harmonization and deliver clinical data templates paired with accompanying imaging data with detailed information on patient journey, from diagnoses, treatment to follow up. In this process extensive work on de-identification, annotation tools, definition of data quality metrics, data quality monitoring, data and annotation curation have been done. All of these aspects engineered by the technical teams, for the use of medical personnel share some common messages as best practices: 1) establishing clear guidelines; 2) ensure continuous communication with Data Providers and data status monitoring; 3) organize (live) workshops to ensure Data Providers know how to use all tools and understand all steps in

data collection process and data quality metrics ; 4) invest in automation by developing robust and user-friendly tools to enforce adherence to data structure and data quality; 5) apply staged data upload and quarantine mechanism for non-compliant data.

For developed clinical and imaging data templates, INCISIVE offers the tailored solutions such as Data Integration Quality Check Tool (DIQCT) and web based medical image annotation tool, creating a solid example of comprehensive medical data collection tools with quality assurance.

However, it is worth noting that the critical link in data collection process is adherence of Data Providers to data collection guidelines, their careful error examination and correction. In INCISIVE experience, Data Providers where technical and medical personnel were teamed, provided for smooth and reliable data collection.

Adherence to standardization with DICOM and HL7 FHIR, INCISIVE achieves a Common Data Model (CDM) that allows receiving, storing, and processing information from multiple sources, multiple data providers in a standard way (experiences shared in the HL7 FHIR INCISIVE Implementation Guide and the Logical Model for each cancer type, see https://fhir.incisive-project.eu/index.html). We have invested our effort to support development of the HL7 FHIR European Cancer Implementation Guide which is our recommendation for all future similar efforts in cancer data collection.

INCISIVE hybrid data repository as an infrastructure facilitating hybrid data storage, data sharing, data search engine and federated space for integration and development of AI models. This deliverable offers some best practices to ensure journey toward creating a cohesive, interoperable ecosystem. INCISIVE technical teams have summarized successful practices in implementation of data ingestion and federated search for medical data of different types with all constraints/requirements. The BSC team provided a through set of best practices on integration of the developed INCISIVE AI models, and the challenges of development of new models in the federated space.

INCISIVE experiences in development of AI Toolbox highlights importance of data quality and data harmonization for development of solid AI models. There is a clear need to enrich the communication between users (healthcare professionals) and data engineers, that would bridge the gap between the expectations and needs in healthcare and the corresponding model development. It is very important to ensure that data providers understand the need to invest their efforts to facilitate good quality data with integrated expert knowledge in order to support service development. In the final step, between a good AI model and good AI service, there is a user interface that shapes the information presentation, being critical to user comprehension of the quality of the services offered.

INCISIVE validation studies have confirmed importance of early stage ethical clearance, timely patient recruitment and realistic estimate of numbers of patients that can be recruited in certain

time framework. The successful study conduction has been ensured through various supporting tools: training handbooks and live training sessions, Microsoft Outlook forms with integrated digital signatures to streamline the consent process and expedite document handling. These forms allowed healthcare professionals to provide updates on the progress of INCISIVE prototype testing. These advancements highlight a best practice, demonstrating a time-saving solution that could prove beneficial for future projects.

The continuous support offered by legal partner TLX has been significant catalysist for the data collection and data analysis process, leading the way ensuring ethical and legal constraints are met. There was a clear need to understand all privacy related terms and processes, GDPR roles of different partners, providing the optimal data usability under privacy constraints and different ethical clearance requests. In the ever reshaping legal space, with new acts and new regulations, there is a need for a solid basis that would maximize potential for sustainability of the INCISIVE repository and data sharing platform.

Regarding the INCISIVE repository sustainability, we have managed to transition INCISIVE into sustainable entities, EUCAIM, to ensure continued impact and relevance beyond initial funding periods. We have managed to leverage transition as opportunity to refine and improve in alignment with evolving needs and standards.

INCISIVE boosts effective communication tools aimed at fostering general awareness and disseminating the project's outcomes. The D&C team has innovatively and consistently nourished its website and associated portals with high-quality content, and established a dynamic network of professionals through strategic use of social media, mailing lists, events, and meetings. At the same time INCISIVE partners made significant contributions to dissemination by publishing scientific articles in peer-review journals and conferences, participating in external events, sharing project-related content through websites and social media.

The wider scope of AI4HI projects have confirmed that challenges met on the INCISIVE pathway were common. Teaming the forces and discussing different approaches through focus groups has been not only beneficial for each of the projects, but as well has connected wider community to summarize good practices and support future initiatives. There are some common challenges identified in all projects, such as communication across different disciplines, and alignment of their needs and expectations. On the other hand, some good practices were highlighted by all AI4HI projects such as investing efforts in data harmonization, producing solutions aligned with standards, inter-professional team work, and engaging the legal advisors as partners were among the most relevant.

# 5 Conclusions

This deliverable summarizes INCISIVE best practices, experiences and lessons to be shared with scientific community to ensure smooth continuity, knowledge transfer and further advancing of the INCISIVE results and avoid replication of similar efforts or suboptimal use of resources. INCISIVE has evolved towards completing its ambitious goals, providing for its major outputs, but as well resulting in lots of accumulated experience, summarized in this deliverable, that we consider valuable for future similar initiatives.

INCISIVE experience, aligned with the wider AI4HI perspective, shares the common views on the need of more cohesive inter-disciplinary collaboration, towards bridging the gap between technical and clinical experts to facilitate development of trustworthy AI solution in healthcare and enable these advances enter the digital space in clinical departments.

# 6 References

Güngör, H. (2020). Creating Value with Artificial Intelligence: A Multi-stakeholder Perspective. *Journal of Creating Value 6(1)*, 72-85.

McMahan, B. M. (2017). Communication-efficient learning of deep networks from decentralized data. *Artificial intelligence and statistics, PMLR*, pp. 1273-1282.

Wang, H. Y. (2020). Federated learning with matched averaging. *arXiv preprint* , arXiv:2002.06440.