

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 6.4

INCISIVE Training Material

WP6 – INCISIVE System Integration and Pilot Studies

20/12/2023

Revision 1.0

Status: < Final >

Grant Agreement n 952179



DOCUMENT CONTROL							
Project reference Grant Agreement number: 952179							
Document name INCISIVE Training Material							
Work Package	WP6						
Work Package Title	INCISIVE System Integration and Pilot Studies						
Dissemination level	СО						
Revision	0.1						
Status	< Final >						
Reviewers	DISBA and HCS						
Beneficiary(ies)							

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									
	Organisation								
Document le	ader		Shereen Nabhani	i -Gebara and Lithin 2	Zacharias	KU			
			Alberto Gutierrez	z Torre		BSC			
			Alexandra Kosvyr	a		AUTH			
			Caroline Barelle			ED			
			Chrysostomos Sy	mvoulidis		TIS			
			Gianna Tsakou			MAG			
			Paris Laras			MAG			
			Sofia Theodorido	u		MAG			
Dorticiponto			Stefanie Charalar	MAG					
Participants			Tatjana Loncar Tu	UNS					
			Paschalis Bizopou	CERTH					
			Zisis Sakellariou	CERTH					
			Giovanni Mazzeo	CeRICT					
			Olga Tsave	AUTH					
			Chara Stefanou	ADAPTIT					
			Olalla Aramburu			MDT			
			Ana Belen	FTSS					
			REVISION H	HISTORY					
Revision	Date		Author	Organisation		Description			
0.1	02/06/2023	Lit	nin Zacharias	KU	ТоС				
0.1.1	07/06/2023	Lithin Zacharias		KU	Version for internal review by Shereen Nabhani-Gebara				

0.1.2	16/06/2023	Lithin Zacharias	KU	Draft for peer review			
0.1.3	20/07/2023	Lithin Zacharias	KU	Draft incorporating input from peer reviewers			
0.1.4	15/08/2023	Lithin Zacharias	KU	Second draft for peer review			
0.1.5	09/09/2023	Lithin Zacharias	KU	Draft incorporating input from peer reviewers			
0.1.6	25/10/2023	Lithin Zacharias	KU	Third draft for peer review			
0.1.7	15/11/2023	Lithin Zacharias	KU	Final draft for review			
0.1.8	28/11/2023	Lithin Zacharias	KU	Final version ready and submitted for internal peer review			
0.1.9	19/12/2023	Lithin Zacharias	KU	Incorporated inputs from internal peer reviewers			
1.0	20/12/2023	Lithin Zacharias	KU	Final version ready for external peer review			

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 Executive Agency 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

1	Exec	utive summary	5
2	Intro	duction	6
-	21	Purpose and scope	0
	2.2	Document structure	7
	23	Relation with other deliverables	7
	2.0		
3	Usag	e Guidelines	8
4	Data	Providers	9
	4.1	Data Sharing and Preparation with INCISIVE	10
	4.2	Data Preparation	11
	4.3	Installation Guidelines for the Data Integration Quality Check Tool	23
	4.4	Utilizing the Quality Tool: Instructions and Video Tutorial	37
	4.5	The INCISIVE Storage Functionality	48
5	Data	Users: AI Developers/Researchers	. 60
-	5.1	The INCISIVE Search Engine	61
	5.2	Model Training	68
	5.3	The INCISIVE workspace and AI engine	69
6	Med	ical Personnel	. 73
	6.1	AI Service Selection	74
	6.2	The INCISIVE AI Services	75
7	Cond	lusion	. 79
~			00
ð		ANNEYUDE 1. Data da idantification and Americation Table	. 80
	8.1	ANNEXURE 1 – Data de-identification and Annotation Tools	80
	8.1.1 Do id	EXLETTION TOURS	00
	De-la	enujication (CTP Diculvi Anonymizer)	80
	A11110	LULIUTI (TT STVAR)	91
	0.1.2 Do id	IIVUSIVE TOUS	55
	De-lu Somi	cutomatic appotation tool	93
	Senti		90

1 Executive summary

The ongoing development of the INCISIVE project (<u>https://INCISIVE-project.eu/</u>) has resulted in significant progress towards its goals. A clear understanding of how to effectively use the INCISIVE platform is described in this manual. This manual serves as a complete guide to using the INCISIVE platform, catering to three key user groups: Data Providers, AI Developers/Researchers, and Medical Personnel.

Within the manual, tailored guidance is given for each user group, highlighting key functionalities relevant to their roles. Data Providers receive instructions on securely contributing and sharing health data, emphasizing data privacy and security protocols. AI Developers/Researchers are equipped with insights into utilizing the platform's tools for AI solution development, including data pre-processing and model training. Medical Personnel gain a comprehensive guide for utilizing the platform in their research or clinical workflows, improving their decision-making and patient care plans.

By delivering this manual, the aim is to empower Data Providers, AI Developers/Researchers, and Medical Personnel with the knowledge and tools to effectively leverage the INCISIVE platform.

2 Introduction

This manual offers detailed, step-by-step instructions on how to effectively use the INCISIVE platform, catering to three distinct user groups: Data Providers, AI Developers/Researchers, and Medical Personnel. By following this information, users will be trained with the necessary knowledge and skills to leverage the full potential of the INCISIVE platform in their respective domains.

For Data Providers, this manual presents clear instructions on how to contribute and share health data securely and efficiently on the INCISIVE platform. It details the data submission process, emphasizing the significance of adhering to data privacy and security protocols to ensure the confidentiality and integrity of the shared data.

AI Developers/Researchers will discover a detailed walkthrough of how to access and utilize the various tools and resources presented by the INCISIVE platform to develop cutting-edge AI solutions for cancer-related tasks. The manual provides insights into the platform's AI development capabilities, such as data pre-processing, model training, and evaluation, allowing developers to harness the platform's potential for building accurate and robust AI algorithms.

In the INCISIVE platform, Medical Personnel and healthcare professionals have a significant role in applying AI services for cancer diagnosis, prediction, and follow-up. They use AI services developed to gain insights and predictions related to specific types of cancer. By incorporating these AI findings into their research or clinical workflows, they enhance decision-making and patient care plans.

2.1 Purpose and scope

The purpose of this document is to provide a comprehensive manual that guides users on effectively utilizing the INCISIVE platform. It caters to three distinct user groups: Data Providers, AI Developers/Researchers, and Medical Personnel. The manual aims to empower users with the necessary knowledge and skills to leverage the platform's capabilities in addressing the challenge of health data availability and developing trustworthy AI tools for cancer diagnosis, prediction, and follow-up.

2.2 Document structure

This document is divided into 8 main sections. The first section (Section 1) provides an executive summary of the work done. The current section (Section 2) sets the introduction, document's purpose, and scope, document structure and relation with other deliverables. Section 3 furnishes details regarding the instructions to be followed prior to utilizing the manual. Sections 4, 5, and 6 provide instructions tailored to various service users. Section 7 presents the conclusion, and finally, Section 8 enumerates the annexures.

2.3 Relation with other deliverables

This deliverable provides an elaborate description of the training materials within INCISIVE, making it pertinent to all the work packages.

3 Usage Guidelines

By adhering to these guidelines, users can effectively utilize the manual and access the assistance tailored to their specific user group on the INCISIVE platform, serving as a roadmap for efficient navigation and utilization of platform functionalities.



4 Data Providers



4.1 Data Sharing and Preparation with INCISIVE

Follow the below steps to prepare and share your data:

Data Preparation: After acquiring health data, follow these steps to prepare it for sharing:



De-identification/Anonymization

Use the INCISIVE data preparation toolbox to remove or anonymize any identifying information from the data.

Ensure the privacy of individuals by adhering to data protection regulations.

Data Quality Check Integration

Utilize the INCISIVE data preparation toolbox to integrate data quality checks.

Verify the consistency, accuracy, and completeness of the data.





Annotation

Employ the INCISIVE data preparation toolbox to annotate the data.

Enhance the value of the data by labelling relevant features and ensuring proper organization.

Data Upload and Storage: Once the data is prepared, follow these steps to upload and store it within INCISIVE:



4.2 Data Preparation

To prepare the data, follow the steps provided below:

Data Structure

Step 1: Create folders as per cancer type



Step 2: Make sure the imaging modality is ready as per the cancer.



Step 3: Set up the patient folder



- Open the parent folder (Image 1)
- Create another folder with patient ID name (Image 2 & 3)
- Copy or drag the imaging modality in the folder (Image 4)

Templates

Step 1: Copying of the Excel template



- Copy the excel template from <u>templatesProspective_Pilots</u>
- Insert one row per patient or many rows for the same patient.
- Please pay attention to the value ranges proposed in the second raw of each template.
- Fill in the values of each field based on its range. Examples:

- Field "Ki67", value range: "%", value inserted: "61".
- Field "Max tumor diameter", value range: "in cm", value inserted: "9".
- Field "Distant metastases", value range: "0: no 1: yes", value inserted: "0".
- Moreover, decimal numbers should be inserted with dot, not comma (e.g., '5.2')
- For the fields that you do not have information about, leave them blank

Step 2: Saving the Excel file

· → • ↑	KU73861 (\\kufs.kingston.ac.uk\hc	me\Staff\1) (H:) > Desktop > Breast C	ancer 🗸	õ	 Search Brea 	ast Cancer	
rganize 💌	New folder					800 -	3
😭 Data	* Name	Date modified	Туре	Size			
Download:	*	03/11/2022 10:52	File folder				
Pictures	*						
Desktop							
Final D7.2							
INCISIVE_L	лінті						
WORK PLA	N DO						
Microsoft Ex	cel						
	11-						
OneDrive - K	ingst						
This PC							
	•						
File nam	e	_					
Save as typ	e: Excel Workbook (*.xlsx)	1. To					
Autho	rs: Excel Workbook (* xlsx)	dd a tag	Title: Add a tit	tle			
	Excel Binary Workbook (* visb)	(sism)					
	Excel 97-2003 Workbook (*.als)						
Hide Folders	XML Data (*.xml)	-8	1	Tools 👻 🔽	Save	Cancel	
inde i ordero	Single File Web Page (*.mht, *.mh	tml)		L			
	Web Page (*.htm, *.html)						
	Excel Macro-Enabled Template (*.)	(tm)					
	Excel 97-2003 Template (*.xit)						
	Text (Tab delimited) (*.txt)						
	Unicode Text (*.txt)						
	XML Spreadsheet 2003 (*.xml)						
	CSV /Comma delimited) (* ccv)	(-305)					
	Formatted Text (Space delimited) (.prn)					
	Text (Macintosh) (*.txt)						
	Text (MS-DOS) (*.txt)						
	CSV (Macintosh) (*.csv)						
	CSV (MS-DOS) (*.CSV)						

- The templates should be stored in '.xls' format.
- Please try NOT to modify the template structure (rearrange, add or delete tabs or columns).

Timepoints

A	autoSave 💽 off	田 ら、 C、 =		Template	• ~
Fi	le Home	Insert Page Lay	out Formulas	Data Review View	Help
Pa	Cut Copy ~ Copy ~ Copy ~ Copy ~ Clipboard	Calibri B I U ∼	- 11 - A^ / · ⊞ - <u> </u>		eb Wrap Text E Merge & Center → nent 5 Pro
N	4 -	× ~ fx			
	А	В	с	D	E
1	Patient ID	Diagnosis	After 1 st treatmen	1st Follow Up	2nd Follow Up
2	NHS_10001				
3	NHS_10002				
4					
5					
6					
7					
8					
9					
11					
12					
13					

The timepoints defined are the following:

- 1. **Diagnosis (Baseline):** will contain all initial exams and information collected during the diagnosis of the pathology. (Month: 0)
- 2. After 1st treatment (Timepoint1/TP1): will contain exams and information collected during the assessment of the first cycle of treatment. The first cycle of treatment can be considered a surgery or a therapy/ combination of therapies that the clinician decided is the best approach for the patient. In case there are multiple examinations, you will allow for multiple entries under the tag 1st treatment but characterized and distinguished by an additional field. (Month: 1-4 approximately, depending on the applied intervention)
- 3. 1st Follow Up (Timepoint 2/TP2): will contain exams and information collected in two cases: (a) the patient is in remission, and this is a routine follow up, or (b) the patient is in relapse; the therapeutic procedure continues and this timepoint corresponds to a second cycle of treatment. (Month:5-8 approximately, depending on the applied intervention or the cancer/country specific pathway, guidelines)

2nd Follow Up (Timepoint 3/TP3): will contain exams and information collected in two cases: (a) the patient is in remission, and this is a typical follow up or (b) the patient is in relapse; the therapeutic procedure continues and this timepoint correspond to a third cycle of treatment. (Month: 9-12 approximately, depending on the applied intervention or the cancer/country specific pathway)

Non-Dicom Images



In cases that images are provided in another format (e.g., jpg, png), then they should be provided in a directory named as: Patient Id/Modality/Timepoint.

Where:

Patient Id	as defined above
	MRI=Magnetic resonance imaging
	PETCT=Positron emission tomography-computed tomography
	CT=Computed tomography
Modality	MMG=Mammography
	US = Ultrasound
	HI = Histopathological Image
	Xray= Radiography

	BL = Baseline/Diagnosis
Timepoint	TP1= After 1 st Treatment
	TP2 = FollowUp1
	TP3 = FollowUp2

Data Preparation Flow

Step 1: Collect the data



Start collecting the imaging data of patients from each cancer type separately and organize them as mentioned above.

Step 2: De-identify the imaging data

Perform the de-identification and export the pseudonymization mapping table following the instructions provided below.



• Fill in the template

By using the pseudonymization mapping table, insert the clinical data to the template using the patient id provided by Clinical Trial Processor (CTP) Anonymizer Tool for each patient and following the instructions for the ingestion of data in the templates. Transfer the template to the folder where the de-identified data is, so the dataset follows the predefined structure.

• Run the Quality Check Tool

Run the quality check tool following the instructions provided in the next module. After running the quality check tool, correct any issues identified from the tool and run the tool again. Repeat until no error shows up.

• Annotate the images

Annotate the images by following the instructions provided below and using the annotation type and labelling as defined in the tables below per cancer type and modality. You should have 1 radiologist/medical doctor for image annotation and another one to validate the results. There are no restrictions on the annotators' experience. However, each data provider should state the names of the annotators and provide a robust rationale behind their suitability to perform the annotations.

Lung:

X-Rays			СТ	PETCT		
Annotation Procedure	Labels	Annotation Procedure	Labels	Annotation Procedure	Labels	
Bounding Box	1: Suspicious, 2: Problematic	Contouring	1: Malignant, 2: Benign	Contouring	1: Malignant, 2: Benign	

Additional annotation guidelines:

• PET/CT annotation - separated CT and PET images from one PETCT examination are fused in the annotation tool (to accurately perform the annotation on the fused images) and only ONE NIFTY file is a result of the annotation process. This resulting annotation file should be placed in both CT and PET series folder, not only in one of them, to avoid confusion. Please be aware of those problems arise in the anonymization procedure if you attempt to upload fused images (which are considered to be secondary captures).

• For CT scans, if several image series with different windows have been stored and examined (for example lung and soft tissue window), only one series should be uploaded, the one on which the annotation has been done.

- Please annotate only primary malignant lesions.
- Please do not annotate:
 - secondary lesions,
 - metastatic lymph nodes,

- bone lesions,
- pleural effusion,
- pleural thickening,
- radiation-induced pneumonitis,
- radiation-induced pulmonary fibrosis,
- atelectasis or any other identifiable abnormalities.
- If osteosclerotic lesion is part of the primary lesion, it should be annotated.

• A benign lesion should be contoured in green, but when in contact with the primary malignant lesion, should be contoured in red.

• Ipsilateral radiation-induced scar should be annotated as benign lesion, if feasible.

• Special attention should be paid when annotating cavitating lung lesions (i.e., malignant mass lesions with an air-filled cavity inside them) and cystic airspaces in order not to "annotate" the cavity or the airspace.

• Post-surgery scars should not be annotated.

• Do not annotate any suspicious regions on imaging modalities if the lesion cannot be seen on them (for example only based on some background information).

Breast:

Ultrasound		MRI		СТ		Mammography		РЕТСТ	
Annota	Labels	Annota	Labels	Annotat	Labels	Annota	Labels*	Annota	Labels
tion		tion		ion		tion		tion	
Proced		Proced		Procedu		Proced		Proced	
ure		ure		re		ure		ure	

Boundi	1:	Contou	1:	Boundin	1:	Contou	1:	Contou	1:
ng Box	Benign,	ring	Benign,	g Box	Benign,	ring	Benign,	ring	Benign,
	2:		2:		2:		2:		2:
	Suspiciou		Suspiciou		Suspicio		Suspiciou		Suspiciou
	s or		s or		us or		s or		s or
	Indeterm		Indeterm		Indeterm		Indeterm		Indeterm
	inate,		inate,		inate,		inate,		inate,
	3:		3:		3:		3:		3:
	Malignan		Malignan		Malignan		Malignan		Malignan
	t		t		t, 4:		t, 4:		t
					Macrocal		Calcificati		
					cification		on, 5:		
					S		Surgical		
							clip,		
							6:		
							Axillary		
							lymph		
							node		

* Labels meaning:

- Benign label
 - o refers only to benign LESIONS, not to benign (micro)calcifications
- Suspicious or indeterminate lesion label:
 - do not contour the whole volume of the corona radiata (spicules), only contour the core of the lesion
- Malignant lesion label:
 - do not contour the whole volume of the corona radiata (spicules), only contour the core of the lesion
- Calcification label:
 - o do not annotate benign calcifications, only suspicious calcifications indicating malignancy,
 - o do not annotate calcifications along the blood vessels

- Axillary lymph node:
 - should be labeled only if clearly suspicious or malignant based on their appearance in image.
- Additional findings to be annotated as malignant lesions:
 - skin infiltration and mammilla infiltration

Another relevant information:

- Surgical scars are not annotated, unless some malignancy is visible in the area (in that case label is malignant/suspicious region)
- Skin infiltration and mammilla retraction should be labeled as malignant lesion
- Lymph edema in the skin after irradiation should not be annotated.

Colorectal:

	MRI	СТ		РЕТСТ		
Annotation Procedure	Labels Annotation Procedure		Labels	Annotation Procedure	Labels	
Contouring	1: Benign 2: Malignant, 3: Lymph Node (only for DWI images)	Contouring	1: Benign 2: Malignant,	Contouring	1: Benign 2: Malignant,	

Additional annotation guidelines:

- All malignant and benign lesions should be annotated
- Fibrotic tissue should be labelled as a benign lesion
- T2WI is the main type of image to be annotated.
- Annotate T2WI only in axial view and use semi-automatic annotation tool option while annotating in ITK Snap. This will improve the annotation quality of other views which will be done automatically.
- Use Full field of view T2WI

- PET/CT annotation should be done on CT and saved in both series folder (CT series and PET series) suggestion by UNITOV to follow the same principle like in lung cancer.
- Metastasis and lymph nodes should not be annotated
- Extramural vascular invasion and tumour deposits for rectal cancer should not be annotated
- there is no need to indicate if the axial view for rectal cancer is not always indeed axial, as it depends on the cancer.

Prostate:

	MRI
Annotation Procedure	Labels
Contouring	1: Benign 2: Malignant

Additional annotation guidelines:

- The annotation sequence is the axial, high resolution (i.e., with a small field of view) T2W sequence
- Do not annotate DWI sequences
- Annotate by a contour all visible lesions (malignant and benign) in the images, both large and small ones.
- For the largest malignant lesion introduce the information needed in the Excel template
- Benign lesions include identifiable PIRADS 1 & 2 lesions, as well as other benign lesions (PBH nodules). The rest, non-annotated gland is considered to be healthy tissue.

4.3 Installation Guidelines for the Data Integration Quality Check Tool

Windows

Step 1: Install Docker Engine

- Install windows subsystem for Linux
 - a) Open PowerShell as administrator

All Apps Documents Web More	• •	<i>ह</i> र
Best match		
Command Prompt App		
Search the web		Command Prompt
	>	Арр
℅ command prompt	>	
$\mathcal P$ command line	>	□ ⁻ Open
𝒫 command prompt admin	>	Run as administrator
	>	Den file location
Settings		-🏳 Pin to Start
Replace Command Prompt with Windows PowerShell in the Win + X	>	-🏳 Pin to taskbar
☆ Turn high contrast on or off	>	
Manage app execution aliases	>	
다고 Change Narrator's keyboard layout	>	
Use Sticky Keys to press one key at a time for keyboard shortcuts	>	
		o 🛱 🧰 🧿 🖬 🚱 🖬

- 4	

b) Copy and paste the following command to Enable the Windows Subsystem for Linux :

"dism.exe /online /enable-feature /featurename:Microsoft-Windows-Subsystem-Linux /all /norestart"





c) Copy and paste the following command to Enable Virtual Machine feature:

"dism.exe /online /enable-feature /featurename:VirtualMachinePlatform /all /norestart"

Administrator: Command Prompt -		<
terminate, -t <distro> Terminates the specified distribution.</distro>		^
unregister <distro> Unregisters the distribution and deletes the root filesystem.</distro>		
C:\WINDOWS\system32>dism.exe /online /enable-feature /featurename:Microsoft-Windows-Subsystem-Linux /all /nores	tart	
Deployment Image Servicing and Management tool Version: 10.0.19041.844		
Image Version: 10.0.19042.1165		
Enabling feature(s) [=======] The operation completed successfully.		
C:\WINDOWS\system32>dism.exe /online /enable-feature /featurename:VirtualMachinePlatform /all /norestart		
Deployment Image Servicing and Management tool Version: 10.0.19041.844		
Image Version: 10.0.19042.1165		
Enabling feature(s) [======] The operation completed successfully.		
C:\WINDOWS\system32>		v

d) Restart your computer



e) Download the Linux kernel update package from here





g) Copy and paste the following command to Set WSL 2 as your default version

"wsl --set-default-version 2"



h) Install your Linux distribution of choice from here

i)Click on 'Get'. Then, the Microsoft store opens.

j) Click on 'Get' again. The ubuntu 20.04 distribution is downloading. Wait until this procedure is over

and then click on 'Launch'

G install docker windows 10 - Avo: X 😸 Install Docker Decktop on Windo X 🚦	hestal WSL on Windows 10 (Min 🗴 🚦 Anoministic co Ulumbu 20.04 UTS 🗴 🕂	• - 5 ×
← → C	kaclivetab=pivotoven/ewitab	x O * O E
 Τρησιμοποιούμε cookie για να βιλητώσουμε την εμπαρία σα Μοτοsoft αρχική σελίδα κοι 	lawari haw - C X	Autocuv cookie
ubuntu ^a Ubuntu ^a Ubuntu ^a Ubuntu Ubuntu Ubuntu Ubuntu Ubuntu Ubuntu Ubuntu Ubuntu Ubuntu Ubuntu Ubuntu Ubuntu	Ubunct 20.04 LfS rec Ubunct 20.04 LfS Cr Ubunct 20.04 LfS The second sec	
Διαθέσιμο σε	Available on	
Περιγραφή Unumu 2004 11% or Windows allows you to use Ub inducing basis, kit, all gat and many more. Please note that Windows 10 5 does not support run To launch, use "ubunhu2004" on the command line is Suit Menu.	Description Ubuntu 200 LIIS on Windows allows you to use Liburitus Terminal and run Ubuntu command line utilities Including baits via you of and many more. Prese more that Windows 10.5 does not support running this app. To launch, use "ubuntu2004" on the command line prompt (omdexe), or click on the Ubuntu tile in the Sart Meru.	
🕼 wsLupdate_x64msi 🗠	To use this feature, one first needs to use "Tum Windows features on or off" and select "Windows Subsystem for Linux", click OK, reboot, and use this app.	Show all X

/licrosoft Store	-		×
- Home Gaming Entertainment Productivity Deals	, ○ Search	8	
This product is installed.			
Ubuntu 20.04 LTS Cannical Group Limited ・ Developer tools > Utilities ピ Share Ubuntu 20.04 LTS on Windows allows you to use Ubuntu Terminal and run Ubuntu command Line utilities including bash, ssh, git, apt and many more. More PEG 3			
Overview System Requirements Reviews Related			
PC			
Description			
Ubuntu 20.04 LTS on Windows allows you to use Ubuntu Terminal and run Ubuntu command line utilities including bash, ssh, git, apt and many more.			
Please note that Windows 10 S does not support running this app.			
To launch, use "ubuntu2004" on the command-line prompt (cmd.exe), or click on the Ubuntu tile in the			

K) A terminal open and you will then need to <u>create a user account and password for your new Linux</u> <u>distribution</u>. Type a username and click enter. Type a password and click enter. Retype your password and hit enter.



Step 2: Install Docker Desktop for Windows:

a. Download docker desktop installer from here

Install Docker Desktop on Windows



Welcome to Docker Desktop for Windows. This page contains information about Docker Desktop for Windows system requirements, download URL, instructions to install and update Docker Desktop for Windows.

Oownload Docker Desktop for Windows

Docker Desktop for Windows

System requirements

Your Windows machine must meet the following requirements to successfully install Docker Desktop.

WSL 2 backend Hyper-V backend and Windows containers

WSL 2 backend

- Windows 10 64-bit: Home or Pro 2004 (build 19041) or higher, or Enterprise or Education 1909 (build 18363) or higher.
- Enable the WSL 2 feature on Windows. For detailed instructions, refer to the Microsoft documentation.
 - b. Double-click Docker Desktop Installer.exe to run the installer.

c. When prompted, ensure the Enable Hyper-V Windows Features or the Install required Windows components for WSL 2 option is selected on the Configuration page



d. Follow the instructions on the installation wizard to authorize the installer and proceed with the install.

e. When the installation is successful, click Close to complete the installation process.

f. If your admin account is different to your user account, you must add the user to the **docker-users** group. Run **Computer Management** as an administrator and navigate to **Local Users and Groups** > **Groups** > **docker-users**. Right-click to add the user to the group. Log out and log back in for the changes to take effect.

Step 3: Start Docker Desktop and Run the Quality tool



a) Double click on the whale icon on Desktop and accept the terms and conditions



C:\Windows\system32>docker pull -a akosvyra/incisive
firsttry: Pulling from akosvyra/incisive
c549ccf8d472: Pull complete
d2b31eaeac06: Pull complete
f0e96c9b73ad: Pull complete
1aa80c77a06c: Pull complete
ec37d50e9db8: Pull complete
db5efacbb2e7: Pull complete
3f2c49ad7f7b: Pull complete
e07b06bf45c6: Pull complete
8c119a585b0e: Pull complete
89c3b4037654: Pull complete
a5816cc3c570: Pull complete
29f03281e438: Pull complete
78e8adc852af: Pull complete
Digest: sha256:e47c659e16cde69af6c4bf02451d0211cc0a9f220a6d3c6fa2c9e6b26dd747e4
latest: Pulling from akosvyra/incisive
c549ccf8d472: Already exists
d2b31eaeac06: Already exists
f0e96c9b73ad: Already exists
1aa80c77a06c: Already exists
ec37d50e9db8: Already exists
db5efacbb2e7: Already exists
3f2c49ad7f7b: Already exists
e07b06bf45c6: Already exists
8c119a585b0e: Already exists
89c3b4037654: Already exists
1f6077792c27: Pull complete
29f03281e438: Pull complete
78e8adc852af: Pull complete
Digest: sha256:311ee117dbf68128f55d88354e6034d793a91be337319c1de45cdda58e35d509
Status: Downloaded newer image for akosvyra/incisive
docker.io/akosvyra/incisive
C:\Windows\system32>

f. Type: sudo docker run -d --rm -p 3838:3838 -v {path-todata}:/home akosvyra/diqct:latest

Example: sudo docker run -d --rm -p 3838:3838 -v E:\INCISIVEdata:/home akosvyra/diqct:latest

Mac

Step 1: Install docker engine

In this link you can find the instructions on how to install docker:

https://docs.docker.com/desktop/mac/install/

Step 2: Download image and Run the Quality Tool

In a terminal type: docker pull akosvyra/diqct:latest

Wait until the download is complete

```
::\Windows\system32>docker pull -a akosvyra/incisive
firsttry: Pulling from akosvyra/incisive
c549ccf8d472: Pull complete
d2b31eaeac06: Pull complete
f0e96c9b73ad: Pull complete
1aa80c77a06c: Pull complete
ec37d50e9db8: Pull complete
db5efacbb2e7: Pull complete
3f2c49ad7f7b: Pull complete
e07b06bf45c6: Pull complete
8c119a585b0e: Pull complete
89c3b4037654: Pull complete
a5816cc3c570: Pull complete
29f03281e438: Pull complete
78e8adc852af: Pull complete
Digest: sha256:e47c659e16cde69af6c4bf02451d0211cc0a9f220a6d3c6fa2c9e6b26dd747e4
latest: Pulling from akosvyra/incisive
c549ccf8d472: Already exists
d2b31eaeac06: Already exists
f0e96c9b73ad: Already exists
1aa80c77a06c: Already exists
ec37d50e9db8: Already exists
db5efacbb2e7: Already exists
3f2c49ad7f7b: Already exists
e07b06bf45c6: Already exists
8c119a585b0e: Already exists
89c3b4037654: Already exists
1f6077792c27: Pull complete
29f03281e438: Pull complete
78e8adc852af: Pull complete
Digest: sha256:311ee117dbf68128f55d88354e6034d793a91be337319c1de45cdda58e35d509
Status: Downloaded newer image for akosvyra/incisive
docker.io/akosvyra/incisive
C:\Windows\system32>
```

Ubuntu

Step 1: Set up the repository

In a terminal type:

```
1. sudo apt-get update
2. sudo apt-get install \
    ca-certificates \
    curl \
    gnupg \
    lsb-release
3. sudo mkdir -p /etc/apt/keyrings
4. curl -fsSL https://download.docker.com/linux/ubuntu/gpg |
    sudo gpg --dearmor -o /etc/apt/keyrings/docker.gpg
```

Step 2: Install Docker Engine

In a terminal type:

```
    sudo apt-get update
    sudo apt-get install docker-ce docker-ce-cli containerd.io
    docker-compose-plugin
```

Step 3: Download image and Run the Quality Tool

In a terminal type: sudo docker pull akosvyra/diqct:latest Wait until the download is complete

C:\Windows\system32>docker pull -a akosvyra/incisive
firsttry: Pulling from akosvyra/incisive
c549ccf8d472: Pull complete
d2b31eaeac06: Pull complete
f0e96c9b73ad: Pull complete
1aa80c77a06c: Pull complete
ec37d50e9db8: Pull complete
db5efacbb2e7: Pull complete
3f2c49ad7f7b: Pull complete
e07b06bf45c6: Pull complete
8c119a585b0e: Pull complete
89c3b4037654: Pull complete
a5816cc3c570: Pull complete
29f03281e438: Pull complete
78e8adc852af: Pull complete
Digest: sha256:e47c659e16cde69af6c4bf02451d0211cc0a9f220a6d3c6fa2c9e6b26dd747e4
latest: Pulling from akosvyra/incisive
c549ccf8d472: Already exists
d2b31eaeac06: Already exists
f0e96c9b73ad: Already exists
1aa80c77a06c: Already exists
ec37d50e9db8: Already exists
db5efacbb2e7: Already exists
3f2c49ad7f7b: Already exists
e07b06bf45c6: Already exists
8c119a585b0e: Already exists
89c3b4037654: Already exists
1f6077792c27: Pull complete
29f03281e438: Pull complete
78e8adc852af: Pull complete
Digest: sha256:311ee117dbf68128f55d88354e6034d793a91be337319c1de45cdda58e35d509
Status: Downloaded newer image for akosvyra/incisive
docker.io/akosvyra/incisive
C:\Windows\system32>

In a terminal type: sudo docker run -d --rm -p 3838:3838 -v {path-todata}:/home akosvyra/diqct:latest
4.4 Utilizing the Quality Tool: Instructions and Video Tutorial

To utilize the quality tool effectively, kindly follow the instructional video (<u>here</u>) that lasts approximately 8 minutes or follow the step-by-step instructions provided below.

1. The quality tool is a web application and thus the user has to open a web browser (e.g., google chrome or Mozilla Firefox) and type: 127.0.0.1:3838 in the URL address. A web page appears with a login window. The user needs to insert the username and password that have already been sent and click on the 'ok' button:



2. To start the quality check of the data, the user must click on the button 'Select a folder' button located on the upper left side of the page, in order to select the data needed to be analyzed.

INCISIVE - quality check			
Select a folder	Structure and Codification Content Validity		
Clase Completeness	In this tab, errors related to the template structure an not and, (ii) if the patient ids have the proper encoding	nd Platent codification are reported. Specifically, this component reports: (i) whether the initial structure of the template in tem ong format and if there are duplicate patient ids in the template. The user has to correct the reported errors and continue.	rms of tabs and columns position in the template is kept or
Template-Image Consistency DICOM De-identification Protocol	Show 10 v entries		Search:
	X1	÷ X2	
		No data available in table	
DICOM Overall Patient Evaluation	Showing 0 to 0 of 0 entries		Previous Nex
	Designed & Implemented by Laboratory of Computing, Medica	al Informatics and Biomedical - Imaging Technologies, School of Medicine, Aristotle University of Thessalonik	i, Greece

3. The directory tree of the computer appears, and the user needs to select the parent folder where the data lies, the images and the excel file (note: the excel file does not appear in this window) and click on the button 'Select'.

<u>File E</u> dit <u>V</u> iew Hi <u>s</u> tory <u>B</u> ookmarks	Tools Help		×
单 附 Inbox - alexandra.kosvyra ×	User_guide_QualityTool_× 0.0.0.0:3838/ ×		
$\epsilon \rightarrow c$ 0 t	≥ ↔ 0.0.0.0:3838	☆ 🗢 또 원	
M 🖂 🦸 🚾 🗣 🤀			
INCISIVE - quality check	=		
	Please select a folder	×	
Select a loider	Structure and Codification		
Clinical Metadata Integrity	Create new folder	of the template in terms of tabs, and columns position in the template	
Case Completeness	Directories	Content ported errors and continue.	
	▼		
I Template-Image Consistency	Salouta		1
	Show 10 v entries	lea foi-poi277 Search:	
COM De-identification Protocol	Colorantel Propertie	Pe 001-000228	
IE DICOUValidation	data	₩ 001-000229	
	Documents	<u>₩ 001-000230</u>	
I= DICOM Analysis Requirements	🕨 🗁 expJan	🗠 001-000231	
	Showing 0 to 0 of 0 entrice	📴 001-000232 Previous Net	
DICOM Overall Patient Evaluation	🕨 🗁 january	1000233 E E E E E E E E E E E E E E E E E E	
	🕨 🗁 nov	Rep 001-000234	
	▶ 😂 novResults	<u>≥</u> 001-000235	
	🕨 🗁 phd		
	PhDFinalResults		
	RshinyDockerProv08		
	🕨 🗁 Zoom		
	Downloads		
	JupEnv		
	Designed & Implemen	sity of Thessaloniki, Greece	
		Cancel Select	

4. After that, the first component of the quality tool named *Clinical Metadata Integrity* starts to run. When it is finished the results appear in two tabs named *Structure and Codification* & *Content Validity*.

In the first tab, errors related to the template structure and Patient codification are reported. Specifically, this component reports: (i) whether the initial structure of the template in terms of tabs and columns position in the template is kept or not and, (ii) if the patient ids have the proper encoding format and if there are duplicate patient ids in the template.

INCISIVE - quality check											
Select a folder	Structure an	Structure and Codification Content Validity									
Clinical Metadata Integrity											
Case Completeness	↓	This day, a first readed on the impact structure and readed contraction are reported, depending mit or other impacts to use of the tempate in terms of lates and counting position in the tempate is lead of the set of the tempate in terms of lates and counting position in the tempate is lead of the set of the tempate is lead of the set of the tempate is lead of the set of the tempate is lead of the tempate is lead of the set of the tempate is lead of the set of the tempate is lead of tempate is lead of the tempate is lead of the tempate is lead of tempate i									
I Template-Image Consistency											
L DICOM De-identification Protocol	Show All	· entries	Search:								
I≞ DICOM Validation		Aspect	Result								
LE_DICOM Analysis Requirements		All	All								
DICOM Overall Patient Evaluation	1	Tabs Integrity	Tabs follow the proper sequence								
	2	GeneralInfoTab Columns	All columns are in place								
	3	BaselineTab Columns	Columns: Unnamed: 1, needs to be removed or renamed								
	4	TimepointsTab Columns	All columns are in place								
	5	TreatmentTab Columns	All columns are in place								
	6	HistologyTab Columns	All columns are in place								
	7	LabResultsTab Columns	All columns are in place								
	8	GeneralInfoTab PatientIdEncoding	01-000226								
	9	BaselineTab PatientIdEncoding	All patients ids are properly encoded								
	10	TimepointsTab PatientIdEncoding	All patients ids are properly encoded								
	11	TreatmentTab PatientIdEncoding	All patients ids are properly encoded								
	12	HistologyTab PatientIdEncoding	All patients ids are properly encoded								
	13	LabResultsTab PatientIdEncoding	All patients ids are properly encoded								
	14	GeneralInfoTab DuplicatePatientIds	No duplicates found								
	15	BaselineTab DuplicatePatientIds	No duplicates found								
	16	TimepointsTab DuplicatePatientIds	001-000224								
	17	TreatmentTab DuplicatePatientIds	No duplicates found								
	18	HistologyTab DuplicatePatientIds	No duplicates found								
	19	LabResultsTab DuplicatePatientIds	No duplicates found								
	Showing 1 to	10 of 10 entries	Denvire an 4 Mové								

Some common errors in this step might be:

i.Line 3: A new column was inserted and has to be deleted

ii.Line 8: Wrong patient codification. The patient id should be 001-000026 instead of 01-000026.

iii.Line 16: A duplicate patient id was identified in Timepoints tab. This may also result from inserting the same label two times for a patient. The only label that can be inserted multiple times is label that represents 'other' option.

The user must open the template and correct the above errors.

In the second tab the user can see errors related to the template content are reported. Specifically, this component checks if the standards and terminologies proposed for the allowable values of all fields of the templates are followed. It reports, for each patient separately in a different row, the fields of each tab that do not comply with the proposed value range.

INCISIVE - quality check												
Select a folder	Structure and Codification Content											
Clinical Metadata Integrity	In this tab, errors related to the tem	plate content are reported. Specif	ically this component checks if the standards and terminologies pro	nosed for the allowable va	lues of all fields of the templa	ites are followed. It reports, for each patient						
Case Completeness	separately in a different row, the fields of each tab that do not comply with the proposed value range. The user has to correct the reported errors and continue.											
I Template-Image Consistency												
PICOM Anonymization Protocol	Show All v entries					Search:						
토 DICOM Validation	Patient ID GeneralInfo	Baseline	Timepoints	Treatment	Histopathology	LabResults						
E DICOM Analysis Requirements	All	All	All	All	All	All						
DICOM Overall Patient Evaluation	1 001-000225	Max Tumor Diameter, Distant metastasis location, Ti, Ni	Lymph nodes max short axis diameter in TP:2.0,Distant metastasis location in TP:2.0,Ni in TP:1.0 2.0,Mi in TP:2.0	Type of CT in TP:0.0		insulin in TP:0.0 2.0,creatinine in TP:0.0 2.0,potassium in TP:0.0 2.0,sodium in TP:2.0,HDL in TP:0.0						
	2 01-000226											
	3 001-000227	Max Tumor Diameter	Max Tumor Diameter in TP:1.0 2.0 3.0,Distant metastasis location in TP:2.0 3.0	Type of CRT in TP:0.0		HCT in TP:0.0,insulin in TP:2.0,creatinine in TP:2.0,potassium in TP:0.0 2.0,sodium in TP:2.0						
	4 001-000228	Max Tumor Diameter	Distant metastasis location in TP:1.0			HCT in TP:0.0 1.0,insulin in TP:0.0,potassium in TP:0.0 1.0						
	5 001-000229	Ti				HCT in TP:0.0 1.0,potassium in TP:0.0 1.0						
	6 001-000230	Distant metastasis location,Ni	Lymph node stations in TP:1.0 2.0,Lymph nodes max short axis diameter in TP:2.0,Distant metastasis location in TP:1.0 2.0 3.0 4.0,Ni in TP:1.0 2.0,Mi in TP:1.0 2.0 3.0 4.0	Type of CT in TP:0.0 1.0 2.0 3.0		HCT in TP:0.0 1.0,insulin in TP:0.0,creatinine in TP:1.0,potassium in TP:0.0 1.0,sodium in TP:1.0						
	7 001-000231		Distant metastasis location in TP:1.0 2.0			HCT in TP:0.0 1.0,creatinine in TP:1.0,potassium in TP:0.0 1.0,sodium in TP:0.0 1.0,SGPT in TP:0.0						
	8 001-000232	Distant metastasis location, Ti	Distant metastasis location in TP:1.0 2.0 3.0, Ti in TP:1.0 2.0 3.0	Type of CT in TP:1.0 2.0		HCT in TP:0.0 1.0,insulin in TP:0.0,creatinine in TP:0.0,potassium in TP:0.0 1.0,sodium in TP:1.0						
	9 001-000233	Ni	Lymph nodes max short axis diameter in TP:1.0 2.0,Distant metastasis location in TP:1.0 2.0,Ti in TP:1.0 2.0,Ni in TP:1.0 2.0	Type of CRT in TP:0.0 1.0		HCT in TP:0.0 2.0,creatinine in TP:0.0,potassium in TP:0.0 2.0,sodium in TP:0.0 2.0,calcium in TP:0.0						
	10 001-000234	Max Tumor Diameter				HCT in TP:0.0,potassium in TP:0.0,sodium in TP:0.0						
	11 001-000235					HCT in TP:0.0,insulin in TP:0.0						
	Showing 1 to 11 of 11 entries					Previous 1 Next						

The user needs to go back to the template and correct the errors reported by the tool. These errors mean that the inserted values do not comply with the proposed value ranges and guidelines. Examples given in the image below.



5. Then the user must click on the second component named Case Completeness. This component presents an overview of the data provided, what modalities in each timepoint, as well as the percentage of mandatory fields that are present for each patient. No action is required by the user in this stage, but

this step needs to be executed because it is a dependency for the next step. Please be aware that for patients with coding errors or identified duplicates, there are no results available at this stage.



6. The third component named *"Template-Image Consistency"* checks if the file structure of the images is compliant to the template and performs a proper renaming of the studies' folders so they can be stored in a unified and commonly understandable way, which is a prerequisite in INCISIVE. In case that there is an inconsistency between the template and the folders a message appears for the specific patient. If the folders were renamed properly, no message appears.

INCISIVE - quality check	≡		
Select a folder Clinical Metadata Integrity Case Completeness	This betw	component checks if the file structure of the images is compliant to the template and performs each the template and the folders a message appears for the specific patient. If the toders were	a proper remaining of the studies' folders so they can be stored in a unified and predefined way. In case that there is an inconsistency remained property, no message appears. The user has to correct the reported errors and continue.
Template-Image Consistency	Results of	the Template-Image Consistency and I	Renaming check
LICOM De-identification Protocol	Show All ~ entries		Search:
E DICOM Validation	patientID	≑ path	÷ error message
E DICOM Analysis Requirements	All	All	All
DICOM Overall Patient Evaluation	1 001-000224	N/A	No valid timepoints index in csv template (or patient empty entry)
	2 001-000225	/home/saloufa/Documents/ColorectalProspectiv	ve/001-000225 # of studies in template do not match # of folders in dir
	3 01-000226	N∕A	No folder with that patient ID
	4 001-000227	/home/saloufa/Documents/ColorectalProspectiv	ve/001-000227 Folder already renamed 001-000227_CT_TP3

41

The errors that might appear are as in the image above:

- i.No valid timepoints index in csv template (or patient empty entry): which means that the previous step did not produce a line for this patient, either because the patient id is not correct, or nothing was inserted in the template.
- ii.#of studies in template do not match # of folders in dir: which means that the insertions for the modalities in the template do not agree with the folders inside the patient directory. So, for the patient 001-000025, from the previous step the user can see that in baseline a CT and an MRI is inserted, however the folder contains only CT. The template must be corrected accordingly.
- iii.No folder with that patient ID: which means that for this patient there is no available folder with images, either because the patient id is wrong, or the folder is missing.

iv.Folder already renamed 001-000227_CT_TP3: Means that the renaming was done properly.

The user needs to go back to the data and correct the errors reported by the tool.

7. The user then must click on the fourth component named "*DICOM de-identification Protocol*". This component checks whether the de-identification protocol, already defined in INCISIVE, is properly applied in the metadata included on the imaging files. The user is given the ability to select the de-identification protocol that that was applied and is related to the phase of the data collection (retrospective or prospective)



The user can see in a bar chart the 10 most common issues that are related to the de-identification of the images. In the Y axis the name of the Dicom metadata field is shown whereas the number of images which present the respective issue appears in the X axis. More details on the output of the de-identification check process are provided in a tabular format below. In that table, the exact metadata field in terms of group, element, and name, where a problem was identified is mentioned. The current value of each field is mentioned in the column named "value" whereas in the column named "Problem" a sort of description of the issue is reported. Finally, in the last column, named "filename" the path of the image is mentioned ("filename").

INCISIVE - quality check	=													
Select a folder Clinical Metadata Integrity Cose Commissionees	This component choose whether the do-deembcation protocol, annually defined in IHCIDIVE, is properly applied in the metadata included on the imaging thes													
Template-Image Consistency	Most frequent e	rrors												
E DICOM De Identification Protocol														
Select de-identification Protocol Retrospective	StritsDesorption 12 13 14 14 14 14 14 14 14 14 14 14 14 14 14													
	DICOM													
	ProtocolName													
		0 5 10 15	20 25 30 35 4	80 45 50 55	60 65 70 75 80 number of images	85 90 95 100 11	05 110 115 120 125 130							
	Results of Show All • entrie	the De-identifica	ation Protocol				Search:							
	group	element ((name	• value	Problem	filename	+							
	All	All	All	All	All	All								
	81 0018	1030	ProtocolName	IDOSE THORAX ABDOMEN BODY	This field should have been removed	E/INCISIVE/qualityToolCheck/Colo 000224_CT_TP1/Senes-1/Image-1	rectalProspective/001-000224/001- .dom							
	111 0018	1030	ProtocolName	IDOSE THORAX ABDOMEN BODY	This field should have been removed	E./INCISIVE/qualityToolCheck/Colo 000224_CT_TP1/Series-1/Image-1	rectalProspective/001-000224/001- .dom							

The users are also able to apply filters on this table, by typing the respective information in the first row of the table.

Res	Results of the De-identification Protocol											
	group	element	name	🔷 value	Problem	filename 🔶						
	All	All	Access 🛞	All	All	All						
20	0008	0050	AccessionNumber	1611293332707533	This field should have been removed	E:/INCISIVE/qualityToolCheck/ColorectalProspective/001-000224/001- 000224_CT_BL/Series-1/image-1.dcm						
201	0008	0050	AccessionNumber	1397548021897190	This field should have been removed	E:/INCISIVE/qualityToolCheck/ColorectalProspective/001-000224/001- 000224_CT_TP1/Series-1/Image-1.dcm						
202	0008	0050	AccessionNumber	2026243997257304	This field should have been removed	E:/INCISIVE/qualityToolCheck/ColorectalProspective/001-000224/001- 000224 CT_TP2/Series-1/Image-1.dcm						

If errors appear in this step, the user must communicate with the respective partner and resolve the issue. For example, a user by mistake adds a new dicom file in a folder without having previously applying any de-identifaction process and the quality tool reports the issue and prompts the user to correct it.

Res	Results of the De-identification Protocol										
	group	element +	🗧 name	\$ value	Problem 4	filename					
	All	All	All	All	All	μı					
20	0008	0050	AccessionNumber	3059906632743265	This field should have been removed	E:/INCISIVE/qualityTooICheck/ColorectalProspective/001- 000224/001-000224_CT_BL/Series-1/1-001.dcm					
41	0010	1010	PatientsAge	034Y	This field should have been removed	E:/INCISIVE/qualityTooICheck/ColorectalProspective/001- 000224/001-000224_CT_BL/Series-1/1-001.dcm					
40	0010	0040	PatientsSex	М	This field should have been removed	E:/INCISIVE/qualityTooICheck/ColorectalProspective/001- 000224/001-000224_CT_BL/Series-1/1-001.dcm					

8. The user then must click on the fifth component named DICOM Validation. This component checks whether the dicom files are valid, in terms of missing any attributes or have any unexpected ones. The component is based on https://dclunie.com/dicom3tools/dciodvfy.html. The check is done by comparing the contents of the DICOM file against the modules and attributes required by the DICOM standard for the SOP class of the given dataset. This component does not check for consistency between multiple files (e.g., consistent values of Study or Series level attributes. The most common validation problems are depicted in a bar-chart (see figure below-upper) whereas in the bottom part of the figure below, the information of the validation issues is depicted for each patient, where most of the issues observed, for each timepoint. The user can interact with the chart and filter the visualization.



Finally, details are presented in a tabular format, and the user can filter the data based on the value of each column.

Details of the Validation Analysis

Show	All • entries				Search:
	patientID	timepoint	series	🔷 filename 🤇	Problem
	All	All	All	All	All
1	001-000224	001-000224_CT_BL	Series-1	Image-1.dcm	Illegal root for UID - 001.258986172133644653782258737943497575221 in (0x0002
2	001-000224	001-000224_CT_BL	Series-1	Image-1.dcm	Illegal root for UID - 001.258986172133644653782258737943497575221 in (0x0008
3	001-000224	001-000224_CT_BL	Series-1	Image-1.dcm	Illegal root for UID - 001.205294404651239147689287871998109593221 in (0x0020
4	001-000224	001-000224_CT_BL	Series-1	Image-1.dcm	Illegal root for UID - 001.217662178487597544487989934299246462086 in (0x0020
5	001-000224	001-000224_CT_BL	Series-1	Image-1.dcm	Illegal root for UID - 001.44771394957271111611120522683925807177 in (0x0020
6	001-000224	001-000224_CT_BL	Series-1	Image-1.dcm	Value invalid for this VR - (0x0002
7	001-000224	001-000224_CT_BL	Series-1	Image-1.dcm	Value invalid for this VR - (0x0008
8	001-000224	001-000224_CT_BL	Series-1	Image-1.dcm	Value invalid for this VR - (0x0020
9	001-000224	001-000224_CT_BL	Series-1	Image-1.dcm	Value invalid for this VR - (0x0020
10	001-000224	001-000224_CT_BL	Series-1	Image-1.dcm	Value invalid for this VR - (0x0020
11	001-000224	001-000224_CT_BL	Series-1	Image-1.dcm	Dicom dataset contains invalid data values for Value Representations
12	001-000224	001-000224_CT_BL	Series-1	Image-1.dcm	Missing attribute Type 2 Required Element= <patientbirthdate> Module=<patient></patient></patientbirthdate>
13	001-000224	001-000224_CT_BL	Series-1	Image-1.dcm	Missing attribute Type 2 Required Element= <patientsex> Module=<patient></patient></patientsex>
14	001-000224	001-000224_CT_TP1	Series-1	Image-1.dcm	Illegal root for UID - 001.224747158546386685720357197760935290233 in (0x0002
15	001-000224	001-000224_CT_TP1	Series-1	Image-1.dcm	Illegal root for UID - 001.224747158546386685720357197760935290233 in (0x0008
16	001-000224	001-000224 CT TP1	Series-1	Image-1.dcm	Illegal root for UID - 001.43545773974238219514176551007868722964 in (0x0020

The validation of the dicom files may be crucial for some applications, which do not operate correctly in case an invalid dicom file is provided. However, it is worthy to mention here, that the user is not able to perform any corrective actions in this step, since the editing of the dicom tags is performed by the deidentification tool. This functionality was included to provide the user with a validation report in those case where some applications do not provide the expected results because of the provision of invalid dicom files.

9. The user must click on the sixth component named *DICOM Analysis Requirements*. This component is divided into two different requirement checks. This first one checks whether some analysis requirements are met. In INCISIVE such requirements were defined for the thickness of the slices, or the availability of the information related to the pulse sequence use. The availability of this information allows the technical developers to use data of expected quality for analysis and the training of the algorithms.

INCISIVE - quality check	=											
Select a folder	imag	e requirements Ann	otations									
Clinical Metadata Integrity												
Case Completeness		This component checks whether some requirements are met with the gost of allowing the technical developers to use data of expected quality for analysis and the training of the algorithms										
Template-Image Consistency												
E DICOM De Identification Protocol	Re	sults of the	DIC	OM Analysis	Re	equirements						
Lat. DICOM Validation	Show	All • entries										Search:
E DICOM Analysis Requirements		group	¢	element	¢	name	¢	value	¢	Problem	¢	filename 0
DICOM Overall Patient Evaluation		All		All		All		All		All		All
	1	0008		103E	5	SeriesDescription				Not able to check, field is missing from file		E:INCISIVE/qualityToolCheck/ColorectalProspective/001-000224/001- 000224_MR_TP2/Series-101/image-1.dom
	2	0008		103E	ę	SeriesDescription				Not able to check, field is missing from file		E3INCISIVE/gualityToolCheck/ColorectalProspective/001-000224/001- 000224_MR_TP2/Series-101/image-10.dcm
	з	0008		103E	5	SeriesDescription				Not able to check, field is missing from file		E3INCISIVE/qualityToolChock/ColorectalProspective/001-000224/001- 000224_MR_TP2/Series-101/image-11.dcm
	4	0008		103E	5	SeriesDescription				Not able to check, field is missing from file		E:/INCISIVE/qualityTooICheck/ColorectalProspective/001-000224/001- 000224 MR_TP2/Series-101/image-2.dom

The second dimension of the requirements check is related to the availability of annotation files. The user can see in a tabular format in which folders any annotation files exist. The annotation file must be inside the corresponding 'Series' folder and only one annotation file must exist in that folder. In case an annotation file is found outside the series folder a red sign appears while in the case of multiple annotation file in the same folder the respective rows are painted in orange.

INCISIVE - quality check				
Select a folder	image requirements Annotations			
Clinical Metadata Integrity Case Completeness Template-Image Consistency	This component checks if the	re are annotations in images		
LE DICOM De-identification Protocol	Show All • entries			Search:
LE DICOM Validation	patientID	timepoint	\$ series	♦ filename ♦
LE DICOM Analysis Requirements	All	All	All	Ali
DICOM Overall Patient Evaluation	1 001-000224	001-000224_CT_BL	Series-1	001-000001-BL_PETCT.nil.gz
	2 001-000224	001-000224_CT_BL	Series-1	Untitled nil gz
	3 001-000224	001-000224_CT_TP1	Series-1	Untitled.nii.gz
	4 001-000224	001-000224_CT_TP2		Untitled nii gz
	5 001-000224	001-000224_CT_TP3	Series-1	Untitled.nii.gz
	Showing 1 to 5 of 5 entries			Previous 1 Next
	Designed & Implemented by Laboratory of Com	nputing, Medical Informatics and Biomedical - Imaging Technologies	, School of Medicine, Aristotle University of Thessalon	iki, Greece

The user needs to go back to the data and place the respective file into the correct folders.

10. Finally, the user must click on the fifth component named *DICOM Overall Patient evaluation*. This component checks if there are duplicate images in the dataset and if some prerequisites for the analysis are present. In this respect, the tool checks whether the expected imaging modalities, the slice thickness and in the case of MR images the pulse sequence are provided for each cancer type. The green and the red color reflects that the requirement is met or not, respectively. Finally, in case a requirement is partially met, the respective cell is painted in orange.

Select a folder					
Clinical Metadata Integrity		there are duplicate images in the dataset and it some p	prerequisites for the analysis are present,	modalities in each cancer type, slice thicknes	s and pulse sequence
Case Completeness					
Template-Image Consistency	Results of the DICO	A overall patient evaluatio	n		
DICOM De-identification Protocol	Show All • entries				Search:
LE DICOM Validation	patientID	modalities	duplicates	sliceThickness	Pulse Sequence
LE DICOM Analysis Requirements	All	All	All	AI	All
DICOM Overall Patient Evaluation	1 001-000224	all required imaging modalities are provided	No duplicates found	Slice Thickness requirement is covered	WR data are provided and required pulse sequences are not found
	2 001-000225	at least one required imaging modalities are provided	No duplicates found	Slice Thickness requirement is covered	No MR data are provided
	3 001-000226	all required imaging modalities are provided	No duplicates found	Slice Thickness requirement is covered	Mix data are provided and required pulse sequences are not found
	4 001-000227	at least one required imaging modalities are provided	No duplicates found	Slice Thickness requirement is covered	No MR data are provided
	5 001-000228	at least one required imaging modalities are provided	No duplicates found	No requirement for the provided modalities	WR data are provided and required pulse sequences are not found
	6 001-000229	all required imaging modalities are provided	No duplicates found	Slice Thickness requirement is covered	MR data are provided and required pulse sequences are not

The user needs to go back to the data and correct the errors reported by the tool.

Then the user needs to run the pipeline again to ensure that all errors are corrected.

4.5 The INCISIVE Storage Functionality

1. Usage

To use the Federated Storage functionality, data can be imported into user's organization's Federated Node or the INCISIVE Central Node. Here are the steps for each:

2.2.1 Federated Node Storage

2.2.1.1 Transfer Data to Federated Node

- Connect to the Federated Node either directly, by logging in through the operating system's graphical environment, or remotely through a network and software configuration that has been made by your organization, depending on the Node's configuration.
- Connect the storage device that will be used to copy data from, (e.g., external, or internal hard drive, USB sticks or network drives)
- Copy data from the storage device to the Federated Node's "/data" directory. This directory can be found at the root directory of the Federated Node's file system.

2.2.2. Central Node Storage

2.2.2.1 Request Access to the Central Node



Once connected to the Central Node, you can transfer data to it using the following steps:

1. For Linux:

- Ensure that your local computer has a directory named "data" containing the data you want to transfer.
- Navigate to the parent directory of the "data" folder.
- Connect to the SFTP server.
- Navigate to the directory that contains the "data" directory on the SFTP server.
- Run the command "put -r data/" in the SFTP interface to transfer the directory and its contents to the SFTP server.
- Wait for the transfer to complete without interrupting the process.

2. For Windows:

• Once connected to the SFTP server, navigate to the directory where you want to upload the files.

- Drag and drop the local files onto the remote file panel in the SFTP client.
- The files will be uploaded to the SFTP server.
- The uploaded data can now be used by the INCISIVE platform and the ETL (Extract, Transform, Load) procedures. The ETL process ensures data harmonization and the inclusion of FAIR (Findable, Accessible, Interoperable, Reusable) principles. The data is processed locally without leaving your premises and is indexed in the INCISIVE federated data storage repository.
- The INCISIVE platform also includes a search engine that allows retrospective and prospective datasets stored in the federated repository to be searched using distributed queries. The search engine provides an API for federated AI model development and training. To access the search engine, you can log in through the provided login page using your username (email) and password.

Notably, each organization will have its own separate storage volume assigned to the SFTP, so individuals making use of the Central Node will not have to worry about disrupting another organization's data. The subsequent instructions outline the necessary steps to utilize the Central Node.

2.2.2.2 Install VPN Client on your computer

1. For Linux

a. Open a new Terminal session.

You can open a new session by pressing 'Ctrl + Alt + T' at the same time.

b. Enter the following commands to install the required components:

sudo apt-get install openvpn
sudo apt-get -y install network-manager-openvpn
sudo service network-manager restart

2. For Windows

A VPN client is pre-installed in Windows 10 & 11 platforms. For older versions of Windows (7,8) download the OpenVPN client from here. (https://openvpn.net/community-downloads/ <u>Community Downloads - Open Source VPN | OpenVPN</u>)

2.2.2.3. Install SFTP Client on your computer

1. For Linux

- For All popular distributions of Linux, OpenSSH client is already installed
- Otherwise, execute the following command: sudo apt-get install opensshclient

2. For Windows

For all versions of Windows, download the WinSCP client from here.
 <u>WinSCP :: Official Site :: Download</u>

2.2.2. 4. Connect to Central Node VPN

1. For Linux

Extract the private key and the base64 thumbprint from the .pfx file that was received in
 Step 2.2.1. There are multiple ways to do this, the command below will use OpenSSL to achieve this:

openssl pkcs12 -in "filename.pfx" -nodes -out "exported_key.txt"

The exported_key.txt file will contain the private key and the thumbprint for the Certificate Authority, and the Client certificate. Be sure to use the thumbprint of the client certificate.

Open profileinfo.txt in a text editor to get the thumbprint of the client (child) certificate.

Select the text including and between "-----BEGIN CERTIFICATE-----" and "-----END CERTIFICATE-----" for the child certificate and copy it. You can identify the child certificate by looking at the subject=/ line.

b. Open the vpnconfig.ovpn file found in /usr/bin/openvpn/config and find the section shown below. Replace everything between "cert" and "/cert" with the copied thumbprint from the above step.

P2S client certificate

Please fill this field with a PEM formatted cert

<cert>

\$CLIENTCERTIFICATE

</cert>

- c. Open the profileinfo.txt in a text editor. To get the private key, select the text including and between "-----BEGIN PRIVATE KEY-----" and "-----END PRIVATE KEY-----" and copy it.
- d. Open the vpnconfig.ovpn file found in /usr/bin/openvpn/config and find the section shown below. the private key replacing everything between "key" and "/key".

P2S client root certificate private key

Please fill this field with a PEM formatted key

<key>

\$PRIVATEKEY

</key>

- e. Don't change any other fields. Use the filled in configuration in client input to connect to the VPN.
- f. To connect using the command line, type the following command:

sudo openvpn --config <name and path of your VPN profile file>&

2. For Windows

- a. Locate and double click the .pfx file that was received in Step 2.2.2.1
- b. On the File to import page, don't make any changes. Select Next.

- c. On the Private key protection page, input the password for the certificate, or verify that the security principal is correct, then select Next.
- d. On the Certificate Store page, leave the default location, and then select Next.
- e. Select Finish. On the Security Warning for the certificate installation, select Yes. You can comfortably select 'Yes' for this security warning because you received the certificate through a secure channel.
- f. The certificate is now successfully imported.
- g. Navigate to your Window's VPN Client and open it
- h. On the configuration dialogue, locate and select the vpnconfig.xml file that was downloaded in Step 2.2.2.1
- i. Select Connect

2.2.2.5. Connect to SFTP Server

Once your computer has been connected to the VPN of the Central Node, it is now able to connect to the SFTP Server, below follow the steps to connect:

1. For Linux

a. Open a new Terminal session.

You can open a new session by pressing 'Ctrl + Alt + T' at the same time.

b. Run the following command to connect to the SFTP Server:

sftp -P <port> foo@<host-ip>

Where <port> is the SFTP port that was received in Step 2.2.2.1

foo is the SFTP username that was received in Step 2.2.2.1

and <host-ip> is the SFTP IP address that was received in Step 2.2.2.1

- c. After executing the command, the SFTP interface will request a password. Insert the password that was received in Step 2.2.2.1
- If successfully connected to the SFTP, your terminal command line will change to "sftp>" from here on commands are given to the SFTP interface, which not an exact match to 1:1 Linux bash commands. From here you can execute all needed commands to transfer data to the SFTP.
- e. To exit the SFTP interface and terminate the connection, type the exit command

2. For Windows

- a. Launch WinSCP.exe, which was installed in Step 2.2.2.3
- b. Select "New Site" from the left column
- c. In the Right Column, select the following:

File Protocol: "SFTP"

Host Name: <host-ip>

Username: <username>

Password: <password>

d. Press the "Login" button on the bottom of the dialogue screen.

All the above credentials must have been already received in Step 2.2.2.2.

Completing the above steps will connect your WinSCP client to the SFTP server.

From there, you can choose to interact with the SFTP interface through the WinSCP interface.

2.2.2.6. Transfer Data to Central Node

• Once your local computer has connected to the SFTP interface, you are ready to begin the data transfer process. Follow the below steps to upload data to the Central node:

1. For Linux

- a. The local computer should a named "data" which contains the data that you intend to transfer. Please validate that the files you have prepared for upload are structured as described in section 2.3.
- b. Navigate to the parent directory of the "data" folder
- c. Connect to the sftp
- d. Navigate to the directory that contains the "data" directory.
- e. Validate that you are connected to the SFTP interface, by locating the "sftp>" prefix in your command line
- f. Run the following command in the sftp interface to transfer a directory and its contents to the SFTP Server:

put -r data/

This command assumes two pre-conditions:

- I. The Central Node's SFTP to which you have connected also has an existing directory named "data" in the working directory in which you are running the command.
- g. Executing the above command will take some time to transfer the files, please wait for it to complete and do not interrupt the process.
- h. In the case that the file-transfer process gets interrupted, you can attempt to transfer the files in chunks.
- i. Congratulations! Your organization's data can now be used by the INCISIVE platform.

2. For Windows

- a. Once you are connected to the SFTP server, you will see content of default remote directory (typically a home directory of your account) on remote file panel.
- Please validate that the files you have prepared for upload are structured as described in section 2.3

- c. If you want to upload the files to different directory, navigate there first.
- d. Now you can simply drag the local files and drop them on remote file panel. By default, transfer settings dialog will appear. Typically, you do not want to change any settings, so just press Copy.
- e. Uploading will proceed. Once it finishes, you will see uploaded file in remote file panel.
- f. Congratulations! Your organization's data can now be used by the INCISIVE platform and the ETL procedures.

An ETL procedure in your premises ensure the harmonization of the data in a Common Data Model to then enable their sharing in an interoperable manner and to allow federated AI. The data is processed locally and never has to leave your premises, neither to be processed nor to be ingested into your local storage.

This step is the basis of the inclusion of the FAIR principles in the INCISIVE platform. The INCISIVE Common Data Model builds on open standards based on OMOP and SNOMED CT for the data semantic. FHIR and PACS are used to define the data structure. It ensures data findability, accessibility, interoperability and reusability.

When the ETL procedure is finalised, the data are indexed in the INCISIVE federated data storage repository.

2. 3. File Structure

The Federated Nodes of each Data Provider and the storage volumes used in the Central Node adopt a similar directory structure. This structure allows efficient data ingestion by our Federated Storage mechanisms.

2. 3.1 Directory structure for the Node

Each node employs a directory structure that follows this pattern: /data/INCISIVE-dp-{ORG}, where {ORG} represents your Data Partner (such as auth, goc, uns, etc.). For instance, for AUTH, the directory structure looks like this:



fhir/add: This is the designated folder for placing Excel files that you want to ingest into the FHIR server. You simply move your Excel files into this folder, and the ingestion process will commence.

fhir/processed: After the FHIR server ingestion is complete, the ETL Tool will sort the results into two categories, 'success' and 'error', represented by two respective folders. Files related to successful ingestion will be in the 'success' folder, while any issues or errors that occurred during the ingestion process will be reflected in the 'error' folder.

pacs/add: This is where you need to add patient folders containing DICOM and NIFTI images for ingestion into the PACS server. Just like with the FHIR server, you move the desired folders into this directory to start the ingestion process.

pacs/processed: Similar to the 'fhir/processed' folder, the 'pacs/processed' folder contains two sub-folders, 'success' and 'error'. The ETL Tool will categorize the results of the PACS server ingestion into these folders. Successful ingestions will be represented in the 'success' folder, and any errors that occurred during the process will be in the 'error' folder.

Remember to always check the 'processed' folders after each ingestion process to ensure that the data was successfully ingested and to address any issues as early as possible.

2. 3.2 Folder Standardization for Ingestion

Before you can ingest data into the pacs server, it must first be standardized by the data-quality check tool. This tool generates output in a specific structure:

/{patient}/{patient-study-timeline}/Series-{number}/Untitled.nii.gz

/{patient}/{patient-study-timeline}/Series-{number}/image-1.dcm

For instance, a valid subject structure to be processed by the ETL Pacs application would look like:



2. 3.3 Ingesting Data

To ingest data into the PACS directory, you need to move the data from the patient directories into the pacs/add subdirectory. Given the size and quantity of these files and folders, it's recommended to use the mv command on Linux for this task. Note that depending on your permissions, you might need to use sudo with this command. For instance, to move all patient folders starting with 001-0000 to the ingestion directory, you could execute the following command:

sudo mv /data/new_data/001-0000* /data/incisive-dp-uns/pacs/add/

After this, the ETL Tool will automatically start processing and ingesting the data. Please be aware that this process can take some time to complete, depending on the amount of data being ingested. Ingesting data into the FHIR server follows a similar process to the PACS directory, with a slight difference. Instead of moving patient directories, you will be moving Excel files. To do this, you place your Excel files into the fhir/add subdirectory. So, if you have an Excel file in /data/new_data, you could use the following command to move it for ingestion: sudo mv /data/new_data/Breast_Cancer.xls /data/incisive-dp-uns/fhir/add/

Same as occurs during the pacs data ingestion, the ETL Tool will start the processing and data ingestion automatically.

2. 4. Data update

Updating data at the patient level is straightforward. Simply add the patient's folder to be ingested again. The ETL Tool will then verify if the patient already exists in the FHIR and/or PACS systems, and if so, it will replace all existing data with the new information.

5 Data Users: AI Developers/Researchers



5.1 The INCISIVE Search Engine

The registration procedures and usage of the search engine is detailed below.

Step 1: Registration

Click on the link: <u>http://portal.INCISIVE.magrnd.com/aiservices</u>



Figure 1. Login page of the search engine

Step 2: Sign up/Log in

As a first-time user, please sign up. If you are not, kindly log in using your email ID and password. This page is developed to make sure that only authorized people can use the search engine, and it adheres to important privacy rules called General Data Protection Regulation (GDPR) to keep your data safe.

Enter your first name *	Enter your email *	\sim
Enter your last name *	Enter your password *	0
Enter your organization *	Repeat your password *	Ø
Role *		-
□ I have read and accept the Terms and Conditions		
I have read and accept the Privacy Policy		

Figure 2. Sign up page

Step 3: Home page

When you log in, you will be directed to the home page of the search engine. You will be able to access the below information there:

- List of data providers who provide data for the platform.
- Number of data providers already registered in the INCISIVE platform
- The number of datasets already available in the repository
- Statistical information

MCISIVE											FAQ admin@incisive-project.er
Home 💿	11										
Data Providers	Home Welcome to INCISIVE. You can use the link on the	e et NGUYC. You can use the link on the lint to navigate to the different pages.									
Al Services											
Search	Data providers	Searches		Workspaces		Users		Services		Executed servi	ces
Al Infrastructure											
Administration											
Logs	2	66		17		2		11		66	
	3	55		17		2				00	
	Patients matched			AI Engines				Models			
	7231			8				8			
	Last Used AI Services							All Services	Al Engine		
	Name	Status	Workspac	ce Inference Filename	Created On		Finished On				
	[BREAST] Pipeline	succeeded	breast_ca	incer_pipeline.zip	2023-07-19 10:	51:37	2023-07-19 10:57:18				
	[BREAST] Birads Classification with XAI	succeeded	external_e	data_mg.zip	2023-07-18 07:	16:50	2023-07-18 07:21:37		You can quickly add an r from here. Click the butt	new Al Engine on to start.	
	(BREAST) Birads Classification with XAI	succeeded	external_t	data_mg.zip	2023-07-14 12	49:06	2023-07-14 12:56:24				
	[BREAST] Pipeline	succeeded	breast_ca	incer_pipeline.zip	2023-07-13 09:	02:02	2023-07-13 09:06:13				
	[LUNG] CT Scan Staging	failed	external_e	data_lung_staging.zip	2023-07-12 14:	25:03	2023-07-12 15:28:02		Add New		
	My Search Results							All Search Results	Al Model		
	Query	Status	Results o	ollected	Started On		Finished On				
	My Al Models							All Al Models	You can quickly add an i	new Al Model	
	Name	Merge Type	Ai Engine		Created At		Updated At		from here. Click the butt	on to start.	

Figure 3. Home page of the search engine

You can use the link on the left to navigate to the different pages.

Step 4: Data providers

It describes about various Data Providers registered in the INCISIVE platform. It serves as a summary of various sources of cancer-related information available on the platform, making it convenient for users to find the specific data they want. This page contains valuable information on:

- Data Providers' locations
- The types of cancer they cover
- The number of cases and datasets they share
- Other relevant information

WICISIVE								
Home)	Dete servidere						
Data Providers		ta providers effect the sources from which data becomes available in INCISIVE. The list below presents all Data Providers currently registered in INCISIVE together with details about each one.						
Al Services								
Search		auth						
Al Infrastructure		umon Notario No						
Administration		Auth Comrt						
Logs		A drain visat						
		Integrippential incluing magnets.com/ scenarti 70 a scenarti magnets.com/ scolorectei, lung						
		UG3 Grace Britowartow UGA contar Admin waart http://jortal.idoibie.magind.com/						
		Lus sola s						

Figure 4. Data provider page of the search engine

Step 5: Search page of the search engine

This helps in finding answers to the question by enabling specific information from data base. One can use specific filters to find cases that match one's selected criteria. There are three types of main filters:

- Patient information
- Cancer information
- Data information

MCISIVE					FAQ adminglincisive projectieu (
Home 💿	Search				
Data Providers	The search functionality allows you to query the data available in INCISIVE. Each search being p	erformed is distri	ibuted among data providers and the results are aggregated by INCISIVE.		
Al Services					
Search					C Search
Q Search	Patient Information				
Search results	Gender		Age From	Age To	
Al Infrastructure	Choose *				
Administration					
Logs	Cancer Information				
	Cancer Type		Cancer Stage	Months of Observation	
	Choose -		Choose		
	image Modality			Treatment/Therapy	
	Choose -			Choose	
	Data Information				
	Data Provider / Research group		Dataset country of origin	Datasets will full cases	Genomic data available
					Q Search



Once users have selected the filters that best describe their query, they can run the search by clicking on the search button. Additional parameters that can help further specify the query can be considered upon request, but only if they are related to the search.

Step 6: Search result

The search results page displays a list of processed searches performed by users. On this page, users can view the status of their search, see the results that have been found, and even review the specific query they made with the selected parameters from the search screen.

WICISIVE				FAQ admin@incisive.project.eu 🖪			
Home							
Data Providers	Search results The search results are the anomated results from INCISIVE data providers for w	ch results					
Al Services	,						
Search	Details						
Q Search							
Search results	Query ::	Status :	Date started +	Date finished :			
Al Infrastructure	Proved annous Managementals, income modelles in data annuiden une	THICKED.	2022 07 14 16/12/47	2022 07 14 16 14 22			
Administration	breast cancer, manningraphy image mouting in data providers dita	r morne.	2023/07/14 10:13/0	2020071910.1932			
Logs							
	Breast cancer, Mammography image modality in data providers uns	FINSHED	2023-07-14 16:12:05	2023-07-14 16:12:54			
	Breast cancer in data providers uns	FINISHED	2023-07-10 10.03:16	2023-07-10 10:03:47			
	Lung cancer, with cancer stage IVB in data providers auth, uoa	FINISHED	2023-07-03 19:41:40	2023-07-03 19:41:47			

Figure 6. Search results page

Clicking on a specific search from the list enables to examine detailed information about the patients that were found in that search. This feature permits to closely examine the individual cases that matched the search criteria.

RISIVE							FAQ admin@incisive.project.eu
Home 📀 Data Providers	Search results	results					
Al Services	Overview						
Search Q. Search	Query	Status	Results Collected		Date Started	Date Finished	
Al Infrastructure Administration Logs	Breast cancer, Mammography image modality in data pr a unit	ovider FINISHED	727		2023-07-14 16:13:47	2023-07-14 16:14:32	
	Details						
	Id	Gender			Institution		0
	005-000006	female			uns		0
	005-000008	female			uns		0
	005-000010	female			uns		0
	005-000020	female			una		8

On this page, users can easily examine the number of patients returned by their search and view basic information about each patient, such as their ID, gender, and the data provider source.

For more detailed information about an individual patient, users can click on their record. This action will prompt a new request to retrieve additional clinical data, such as TNM (TNM Classification of Malignant Tumours) staging, and data from their studies, like DICOM (Digital Imaging and Communications in Medicine) images.

Furthermore, users have the choice to modify the search results. For example, they can delete a single patient by clicking the trash button next to their information. Also, there is a button labelled "FILTER RESULTS" that allows users to apply specific filters. When this button is clicked, a popup will appear, providing options to filter the results based on a percentage (randomly removing patients) or by using a previous workspace as a filter. This way, users can refine and customize their search results to match their specific needs.

005-000020	temaie	uns	8
005-000028	female	un	8
005-000031	Filter results by remaile	uni	0
005-000037	e previoat volta female FLTER	uns	8
005-000851	female	uni	0
005-000852	female	uni	0
005-000857	female	uns	8
		Items per page: 10 -	1 - 10 of 727 < >
		GO TO WORKSPACE	FILTER RESULTS DELETE

5.2 Model Training

In this section the steps to train AI models are explained.

Steps to train AI models

Once the prepared data is uploaded and stored, the INCISIVE platform enables to train AI models.

Follow these steps to train models with the available data:



Data Query and Selection

Utilize the search functionality within INCISIVE to find relevant datasets for your AI model development.

Choose datasets specific to the cancer type you are focusing on.



Model Training

Train AI models using the selected datasets.

Follow the guidelines and instructions provided by INCISIVE to optimize and refine your AI models.

5.3 The INCISIVE workspace and AI engine

In this section the concept of workspace, how it is used, the creation of the AI Engines and the training procedure are detailed.

AI Engines

The AI Engines must be registered to be used in the system. Now the docker container that has inside all the code implementing the different use cases must be provided to the administrator, so it is uploaded and made available on the system. Figure 7 shows the screen that is presented when you are creating an AI Engine. This screen requests the name of the AI Engine, the name of the docker container to be used, its version, the type of data query to be used, the description of the AI Engine and the available use cases in the AI Engine. Notice that data query is not used now. When choosing from the provided use cases, you encounter what is shown in Figure 8. For every use case, you can provide a default JSON format, which needs to be uploaded to the platform as showed in the illustration. The AI Developers/Researchers, especially for each AI Engine, define this configuration. Ultimately, upon ensuring all is accurate, upon clicking "Save," you enable the AI Engine. Subsequently, it becomes accessible for utilization in the designated use cases.

MCISIVE			FAQ admin@incisive-project.eu
Home 📀			
Data Providers	Al Engines		
AI Services	Engine Details		
Search	Name*	Container Name*	Owner*
Al Infrastructu <mark>r</mark> e			
B Workspaces			
Al Engines	Data type*	Role type*	Description*
🖏 Al Models	Choose *	Choose -	
Administration	Data considerations	Tri	Ethics
Logs			
	Caveats	Metrics	License
	Ai Engine Version		
	Container Version*	Description*	Functionalities*
			Choose -
	Explains*		

Figure 7. AI Engine specification

Ai Engine Version		
Container Version*	Description*	Functionalities* Choose
Explains* Choose		
Train an Al Model	亡 TRAINING FROM SCRATCH CONFIGURATION FILE	
E Fine tune an Al Model	亡」 TRAINING FROM PRETRAINED MODEL CONFIGURATION FILE	E
Evaluate an Al Model		
Merge Al Models		
Inference		

Figure 8. AI Engine use case default configuration provision

Workspaces

In the "Search Results" section of the INCISIVE search engine, after performing a search and reviewing the displayed dataset, you will notice an option to "Create Workspace" within the accessible screen. Now you can establish a connection between a specific search result and a designated AI Engine.

Figure 9 illustrates the screen that appears upon selecting "Create workspace." In this screen, users can specify the workspace's name, description, and desired type. As demonstrated in Figure 10, the option to create three distinct workspaces is available.

- Training from scratch: the AI Engine will do is to train a completely new model (from the initialization process that the AI Engine contains) with the data selected.
- Training from a pretrained model: the AI Engine will get the specified model in the UI and train using the model as a starting point (usually called model fine tuning).

Evaluating from a pretrained model: the AI Engine will get the specified model in the UI and use it to do inference over the provided data to obtain a score of how good the model for that particular set of data.

Create workspace		
Data to use		
Search	Total Results	Data Providers
Breast cancer in data providers uns	727	1
Workspace Details		
Name*	Description*	Type* Choose -
		CREATE WORKSPACE CANCEL

Figure 9. Workspace configuration basics

Workspace Details		
Name*	Description*	Type* Train an Al Model
Workspace Al Engine & models		
Ai Engine* Lung CT Scan Staging * The Al Engine to sused in your workspace	Ai Engine Version* 2.0.0 - The AI Engine Version to be used in your workspace	Output Al Model Name* Lung2Model The name of the ai model to be created
AI Engine Configuration		
Configuration* { "param1": 1, "param2": 2 } The Al Engine to be used in your workspace		

Figure 10. Workspace configuration basics: selecting the type of workspace

After selecting the type of workspace, you are offered with extra information depending on the type. Figure 11 shows the particular case of doing a "Training from scratch". In this case you can select which AI Engine you are using, the name of the model you want to create (Model1 in this example) and the type of model. Notice that the AI Engines may be able to create more than one type of model, e.g., more of one type of neural network architecture. For those cases, the model

type field is available. In this case, where the MedT model type is utilized, you would specify it accordingly.

Next to that is the AI Engine configuration, which is a JSON text file that defines the parameters to be sent. These parameters are dependent on the AI Engine and defined by the AI Developers/Researchers.

Once the workspace configuration is complete, the scheduling of AI Engine training takes place, and you are presented with a screen similar to the one depicted in Figure 11.



Test Test

Workspace details	
Name	Test
Description	Test
Created on	2023-07-14 16:42:31
Started on	
Finished on	
Data query	Breast cancer in data providers uns
Last message on	2023-07-14 16:42:36

Figure 11. Workspace detail screen
6 Medical Personnel



6.1 AI Service Selection

This section outlines steps to access specialized services for Medical Personnel and healthcare professionals, allowing them to gain insights and predictions about specific types of cancer on the INCISIVE platform.

Steps to utilize AI services

If you are a medical researcher aiming to utilize AI services, follow these steps:



6.2 The INCISIVE AI Services

The INCISIVE AI services covers the services that are offered in the platform.

Usage

The services are intended to be run as easy as possible so that they are convenient to use. First, go to the Services tab as shown in Figure 12. In this screen you will be able to see all the available services in the platform and to start the process by clicking on Run.



Figure 12. Service list screen

After selecting a service, you will be presented with the screen shown in Figure 13, where you will be able to upload the data to analyse by clicking on Click to select data file and then to run the service by clicking on Run service. Currently, the input should be provided in a zip file containing the data to analyse.

Service [BREAST] Birads Classification
Expected Output:
Provided by ICCS
L UPLOAD DATA
A The AI tools available on the platform are research results and have not been the subject matter of clinical evaluation or conformity assessment within the meaning of the Medical Device Regulation 2017/745. The tools can be used exclusively for research. Any predictions or annotations generated by the tools must not be used for diagnostic or therapeutic decisions.
RUN SERVICE CANCE

Figure 13. Service data upload screen

The service is now scheduled to run, as shown in Figure 14. Notice that the first stage of the service is Scheduled. This means that your request is now on a queue waiting to be processed. When it is executing, the status will change to Pending while the system is executing the internals and then to Running when the query is being run. Notice that it is not mandatory to wait on this screen. The home screen will notify you when your query is done as seen in Figure 15.

Service [BREAST] Bi You can monitor the progress You can leave this page and ge Expected Input: Expected Output:	rads Classific	ation is screen. Once your æ results from the H	results are ready, you ome page, by selectir	can download then g the service under	n by clicking on the "D "My Services".	ownload results' bi	utton below.		
			Progress					Results	
Ē				\bigcirc		\bigcirc	<u></u>		\odot
SCHEDULED		PENDING		RUNNING		Done	DOWNLOAD		Display
									Delete Service

Figure 14. Service status screen

Last Used AI Services								
Name	Status	Workspace Inference Filename	Created On	Finished On				
[BREAST] Birads Classification with XAI	succeeded	external_data_mg.zip	2023-07-14 13:49:06	2023-07-14 13:56:24				
[BREAST] Pipeline	succeeded	breast_cancer_pipeline.zip	2023-07-13 10:02:02	2023-07-13 10:06:13				
[LUNG] CT Scan Staging	failed	external_data_lung_staging.zip	2023-07-12 15:25:03	2023-07-12 16:28:02				
[BREAST] Birads Classification	succeeded	breast_cancer_pipeline.zip	2023-07-12 13:13:38	2023-07-12 13:14:06				
[BREAST] Pipeline	succeeded	breast_cancer_pipeline.zip	2023-07-12 11:35:34	2023-07-12 11:40:27				

Figure 15. Service notification in the home

Finally, the results will be downloadable by clicking in Download results as seen in Figure 16. This will provide the results in a zip file. Remember that this is not the final UI and that it might change in the future. In the case of this service, one sample result is the one that can be seen in Figure 17.

Service [BREAST] Birads You can monitor the progress of your You can leave this page and get back Expected Input: Expected Output:	Classification service on this screen. Once your to your service results from the H	esults are ready, you can download ther ime page, by selecting the service under	n by clicking on the "Download resu "My Services".	its" button below.		
		Progress			Results	
	→ ∑	\rightarrow	\rightarrow	<u></u>	\odot	
SCHEDULED	PENDING	RUNNING	Done	DOWNLOAD	Display	
					Delete Se	ervice

Figure 16. Service results screen



Figure 17. Resulting image of this service

7 Conclusion

This manual has been designed as a comprehensive guide tailored for Data Providers, AI Developers/Researchers, and Medical Personnel to efficiently utilize the powerful capabilities of the INCISIVE platform. By tackling the critical challenge of health data availability and emphasizing the development of reliable AI tools for cancer diagnosis, prediction, and follow-up procedures, this manual equips users with detailed instructions, insights, and tools specific to their roles. Through secure data sharing, AI solution development, and medical research guidance, the INCISIVE platform attempts to revolutionize cancer-related healthcare. Empowering users with the expertise and resources outlined in this manual nurture a collaborative environment aimed at advancing accurate diagnoses, timely predictions, and improved follow-up procedures in the realm of cancer care.

8 ANNEXURES

8.1 ANNEXURE 1 – Data de-identification and Annotation Tools

8.1.1 External Tools

De-identification (CTP DICOM Anonymizer)

1.1 Installation of CTP DICOM Anonymizer

1. Download the CTP Dicom Editor installer: <u>http://mirc.rsna.org/download/DicomEditor-</u> <u>installer.jar</u>

 \times

2. Double click on the installer. A popup window, like in the following image, will appear:

DicomEditor is a stand-alone tool for This program allows you to upgrade t SWinREAgent Recovery Anonymizer System.sav hp Windows Doeb In: Windows Doeb In: Windows Doeb In: Windows Doeb In: Windows Doeb In: Windows Doeb In: BB B- SWinREAgent Recovery Anonymizer System.sav hp Windows Doeb In: BB B- Stem.sav Doeb In: BB B- Doeb In: BB B- Stem.sav Doeb In: BB B- Stem.sav Doeb In: BB B- Stem.sav Doeb In: BB B- Stem.sav Doeb In: BB B- Doeb In: BB B- Stem.sav Doeb In: BB B- Doeb In: BB B-			DicomEditor Installe	r	
This program allows you to upgrade t Look In: Windows (C:) SWinREAgent Recovery Anonymizer System.sav hp Windows hpswsetup Program Files OneDriveTemp Program Files Program Data Folder name:	DicomEditor is a stand-alone tool for	🛃 Select a directo	ry in which to install the progran	n ×	
SWinREAgent Recovery Anonymizer System.sav hp Windows hpswsetup Program Files OneDriveTemp Program Files (x86) PerfLogs Users ProgramData Folder name:	This program allows you to upgrade th	Look In: 🖲 Window	s (C:) 💌		
np Windows hpswsetup Program Files OneDriveTemp Program Files (x86) PerfLogs Users ProgramData Programe:		SWinREAgent Anonymizer	Recovery		
Folder name:		hp hpswsetup	Windows Program Files Program Files		
Folder name:		PerfLogs	Users		
Folder name:		ProgramData			
		Folder <u>n</u> ame: C:			
				Open Cancel	

3. Make sure the default 'Folder name' is 'C:\'. Click Open. If the installation was complete, you must see something like this:

DicomEditor Installer

– 🗆 🗙

DicomEditor Installer

DicomEditor is a stand-alone tool for viewing, editing, and anonymizing DICOM objects. This program allows you to upgrade the **DicomEditor** program or install a new one.



Click OK.

1.2 Installation CTP Dicom Editor

You will need to install this tool only if you are a Data Provider whose one or more patients have multiple different patient IDs

For the installation of the CTP Dicom Editor, please refer to 1.1

1.3 De-identifying a patient

Naming conventions of the de-identified files

Patient Ids:

- Patients should be given an id with the following structure: XXXYYYYYY where (i) XXX is a 3-digit code referring to the pilot site and (ii)YYYYYY is a 6-digit serial number (e.g., 000001, 000002).
- The 3-digit numbers for each pilot site are: AUTH: 001, UOA: 002, HCS: 003, GOC: 004, UNS: 005, UNITOV: 006, DISBA: 007, IDIBAPS: 008, VISARIS: 009
- These ids are automatically generated by the CTP tool
- CTP tool exports a log file (excel file) with the mapping of the real patients with the new ids. Each pilot site must keep this file locally and secured. This file will never be asked by any other partner of the project.
- Each time you prepare a batch of data check for duplicate patients to avoid uploading data of the same patient twice.

IMPORTANT NOTE: Some Data Providers have stated that some of their patients have exams from different hospitals, thus it is possible for a patient to have two (or more) different patient IDs. In this case, those Data **Providers need to install another tool name DICOM Editor**, courtesy of the CTP. This tool is necessary to modify the different patient IDs, so that each patient has only a unique patient ID.

If you are such a Data Provider, please follow the next steps (using the Dicom Editor), otherwise ignore them.

The following procedure assumes you have organized the raw patient files as follows: 'C:\...**PatientID**\Study\Series\Image' and all patients under the same root folder.

- 1. Identify the patient(s) with multiple IDs.
- 2. Place all Studies of this patient(s) under the same folder, so that each patient appears only once under the root folder.
- 3. Navigate to 'C:\DicomEditor'
- 4. Double-click DicomEditor
- 5. In the 'Directory' tab, on the left pane navigate to the root folder.
- 6. Double click the root folder to expand it
- 7. Click the root folder
- 8. Make sure the 'Include subdirectories' checkbox is checked (left pane bottom left button) and the file type is '.dcm' (left pane bottom right button).
- In case your files don't have the extension .dcm, click on the file type button, delete what is written, type '*' and then click OK

Directory Viewer Editor Anonymizer Help			
Directory	Refresh	C5. 💌	Results
		^	CONT-87318/6.000000-LIVER-KIDNEYTFL2DAXIAL-3003811-05.dcm
45			UK 1680. Modifying: C.%. IPseudo-PH-DICOM-Data/9894340094/04-16-1997-MRI ABDOMEN PELVIS CONT-87318/6.000000-LIVER-KIDNEYTIFL2DAXIAL-3003811-06.dcm
Directory tab			1661. Modifying: C.\ IPseudo-PHEDICOM-Data/9894340094/04-16-1997-MRI ABDOMEN PELVIS CONT-87318/6.000000-LIVER-RODNEYTFL2DAXIAL-30038/1-07.dcm
			1652. Moding: C: IPseudo-PHI-DICOM-Data/9894340694/04-16-1997-MRI ABDOMEN PELVIS CONT-87318/6.000000-LIVER-ROINEYTIFL2DAXIAL-30038/1-08.dcm
Pasuco PHI DICOM Data	Root folder		1653. Modifying: C:\ IPseudo-PHLDICOM-Data/9894340694/04-16-1997-MRI ABDOMEN PELVIS CONT-87318/s 000000-LIVER-R0LINEY III-L2LIAXVAL-30038(1-09.dcm OK
 ► 292821900 ► 3209648408 ► 330803062 			1694. Modifying: C.\ Pseudo-PH-DICOM-Data/9894340694/04-16-1997-MRI ABD ONEN PELVIS CONT-87318/6.000000-LIVER-RDNEYTIFL2DAXIAL-30038/1-10.dcm
 □ 3642991863 □ 4025360156 □ 571403387 			1685. Modifying: C: P9seudo-PHI-DiCOM-Data/9894340894104-18-1997-MRI ABDOMEN PELVIS CONT-87318/6.000000-LIVER-HIDNEYTIFL2DAXIAL-30038/11-11.dcm OK
 ► 6412074217 ► 6451050561 ► 6614203005 	Right pane after 🛛 🗖		1686. Modifying: C.\ Pseudo-PH-DICOM-Data/9894340694/04-16-1997-MRI ABDOMEN PELVIS CONT-87318/6.000000-LIVER-KIDNEYTIFL2DAXIAL-30038/1-12.dcm OK
 □ 60742/471 □ 60742/473 □ 725997752 □ 735947728 	the procedure has		1667. Modifying: C1 Pseudo-PHI-DICOM-Data/9894340694/04-16-1997-MRI ABDOMEN PELVIS CONT.87318/6.000000-LIVER-RDNEYTIFL2DAXIAL.30038(1-13.dcm OK
 ≥ 8155012288 ≥ 8198244859 ≥ 8548158246 	been completed		1688. Moditying: C1. Pseudo-PH-DICOM-Data/9894340894/04-16-1997-MRI ABDOMEN PELVIS CONT-87318/6.000000-LIVER-KIDNEYTIFL2DAXIAL-30038/1-14.dcm OK
 ■ 8732322741 ■ 8834647407 ■ 8090190720 			1689. Modifying: C1 IPseudo-PHEDICOM-Data/9894/340694/04-16-1997-MRI ABDOMEN PELVIS CONT.87318/6.000000-LIVER-KIDNEYTIFL2DAXIAL-30038/1-15.dcm OK
► □ 9186822008 ► □ 9894340594			1680. Modifying: C1 IPseudo-PH-DICOM-Data/9894340894/04-16-1997-MRI ABDOMEN PELVIS CONT-87318/6.000000-LIVER-KIDNEYTIFL2DAXIAL-3003811-16.dcm OK
			1691. Modifying: C1 Pseudo-PH-DiCOM-Data/9894340694/04-18-1997-MRI ABDOMEN PELVIS CONT-87318/8.000000-LIVER-RODNEYTIFL2DAXAL-3003811-17.dcm OK
	File type	button	1692. Modifying: C11 IPseudo-PHI-DICOM-Data/9894340694/04-16-1997-MRI ABDOMEN PELVIS CONT-87318/6.000000-LIVER-RODNEYTIFL2DAXAL-3003811-18.dcm OK
		Ţ	Test. Modifying: C11 Preudo PHEDICOM. Data/8694340094/04-16-1997-MRI ABDOMEN PELVIS CONT-873166.000000.LIVER-RDNEYTFL2DAXAL.3003811-19.dcm OK OK
		V	
1			Change names of modified files SetPatientIDs button
Cinclude subdirectories		.dcm	Use SOPIUD for name Clear preamble Anonymize

- 10. Click on the 'SetPatientIDs' button on the bottom right
- 11. Wait until the results appear on the right pane.

This procedure is followed by all Data Providers.

The following procedure assumes you have organized the raw patient files as follows: 'C:\...\PatientID\Study\Series\Image' and all patients under the same root folder.

- 1. Navigate to 'C:\Anonymizer'
- 2. Open the program.properties file with a text editor
- 3. On the line: 'SITEID=XXXXX', replace the XXXX with the value corresponding to your data provider's name (e.g. If your data provider's name is AUTH and the corresponding value is 001, then SITEID=001)
- 4. Save and close the file
- 5. Double-click Anonymizer
- 6. Click 'Admin' (tab on top-left)
- 7. Click 'Filter' (tab on top left)
- 8. Uncheck the 'Remove Secondary Capture Files' (checkbox on the bottom-left)
- 9. Click 'Save' (button the bottom right corner)



- 10. Click 'Directory' (tab on top)
- 11. Make sure the 'Include subdirectories' checkbox is checked (left pane bottom left button) and the file type is '.dcm' (left pane bottom right button).
- 12. In case your files don't have the extension .dcm, click on the file type button, delete what is written, type '*' and then click OK

ctory	Refresh C: Results C:Manymizer/Storag
P → Pause P4 COCH D2a → → → D504480 → → → D504480 → → → D504480 → → → D504480 → → → D504800 → → → D5048000 → → → D50480000 → → → D504800000 → → → D5048000000 → → → → D5048000000 → → → → D5048000000 → → → → → → → → → → → → → → →	<pre> the second se</pre>

- 13. Navigate to the root folder containing all the patients from the left panel
- 14. Double click the root folder to expand it
- 15. Click 'Import' (button bottom-right)



- 16. The de-identified files are located in the 'C:\Anonymizer\Storage'
- Navigate to 'C:\Anonymizer\Storage'. The structure of the de-identified folder(s) will look like this:
 '00100001/Study-CT-19971005-292/Series-3'

1.4 Exporting the pseudonymization mapping table

CTP takes care of mapping each patient to a pseudo id and provide that information in a pseudonymization table, in the form of an Excel sheet, which can be saved on the Data Provider's computer.

- 1. If you have closed the Anonymizer program, open it once again.
- 2. Click the 'Admin' (tab on top-left)
- 3. Click 'Index' (tab on top left)
- 4. Click the 'List' button (button on the middle bottom)



5. Click the 'Save' button. The list shown in the following image depicts the mapping between each patient and their given pseudo id (generated by CTP). More specifically the two columns in the left show the de-identified patient name and id respectively, while the next two columns show the real patient's name and id respectively.

				Patient Index List
ON-PatientName	ANON-PatientID	PHI-PatientName	PHI-PatientID	
00001	001030002	DAVIDSCB*9000CLAS	1001.000000-37718	
000001	001030003	DAVIDSON-DODOLAS	1002.000000-53238	
000014	001030005	SEXTOP-CARNER	3203648408	
000006	001030007	MARTIN~CEAD	339633062	
000038	001030009	NALLACE*SEELLY	3642991863	
0100010	001030011	POOLE-OLIVIA	4025360156	
000012	001030013	IEROV-BICOLE	571403367	
000014	001030015	ERASLEY*ANDREW	6415974217	
00016	001030017	MIERS MARK	6451050561	
00018	001030019	YOUNO-ELISADETE	6616238035	
00020	001030021	RIVERA-MICHAEL	6670427471	
00022	001030023	KADIMAN" SCOTT	6774625273	
00024	001030025	REDGERS-SORCEL	7255997752	
00025	001030027	HEROTES-ALITTYR	7361647728	
00028	001030029	HOWER-BOOMET	8155012268	
00030	001030031	SABCHEZ*GRACE	#189244869	
00032	001030033	HEYER*STEPSAKIE	8548156246	
000.34	001000035	UNITS TRAVES	0/32222/41	
00035	001030031	THE REPORT	0034547407	
Nov 25	001000033	COMES-RECEARC		
00042	001030041	10.0175.0 1040.05	3103022330	
00012	001090043	HILLIGH-DEVENU	20112021	
				Save button
				Ĥ

6. Click 'Save' on the pop-up window

ORD-Patienthame AM Million	NCN-PatientID 100003 100005 100005 100005 100005 100005 100001 10000 10000 10000 10000 10000 10000 10000 1000 1000 1000 1000 1000 1000 1000 1000 1000 1000 1000 10	PRT-PatientName Dorbservoxula Dorbservoxula Sister-voxula Sister-voxula Sister-voxula Sister-voxula Sister-voxula Sister-voxula Sister-voxula Sister-voxula Sister-voxula Sister-voxula Sister-voxula Sister-voxula Sister-voxula Sister-voxula Sister-voxula Sister-voxula Sister-voxula Sister-voxula Sister-voxula	PHI-PatientID Idd:.0000-30118 Idd:.0000-3018 Idd:.0000-3018 Idd:.0000-3018 Idd:.0000 I	ave E Anonymizer ata ta [for [for [for [for [for [for [for [for []]]) Nersocijet) nejskral jegi jeroperties jeroperangeoperties	× atomic a
000011 0011 000014 0011 000014 0011 000014 0011 000014 0011 000014 0011 000014 0011 000014 0011 000014 0011 000015 0011 000016 0011 000017 0011 000018 0011 000019 0011 000019 0011 000019 0011 000019 0011 000014 0011 000015 0011 000016 0011 000017 0011 000018 0011 000019 0011 000014 0011 000015 0011	100000 1000000 10000	DAVIDERS ¹¹ -000LAS BIRTIN ¹² -000LAS BIRTIN ¹² -000B BIRTIN ¹² -000B WILLOC ¹² -000D WILLOC ¹² -000D MILLOC ¹	1003.00000-07138 1000.00000-03239 009444449 209420344 209420344 4005100344 4005100344 4005100344 4005100344 4005100344 4005100344 4005100344 4005100344 4005100344 4005100344 4005100344 4005100344 400510034 400510004 400510000000000000000000000000000000000	ave Is: Anonymizer ala [conces] conces] conces] conces]) filler.script) solp.hmi) egit.properties) eropran.properties	
00001.00001 0011 0000014 0011 0000015 0011 000015 0011 000015 0011 000015 0011 000015 0011 000015 0011 000016 0011 000017 0011 000018 0011 000019 0011 000019 0011 000019 0011 000019 0011 000019 0011 000019 0011 000019 0011 000019 0011 000019 0011 000019 0011	100005 100005 100005 100001 100001 100001 100003 100005 100005 100005 100005 100005 100005 100005 100005 100005 100005 100005 100005	DAVESSE">CONLAS SERTIS-CANES SERTIS-CANES SERTIS-CANES SERTIS-CENT SERTIS-CENT SERVICE	1003.0000-5333 2009-64444 2009-64444 344279348-3 442279425-4 002530255 9714-8384 4645397427 4645197427 4645197427 4645197427 4645197428 4645197428 464519745555555555555555555555555555555555	ane In Anonymiter ata [ga [cP] crage) Siter script) solg.timi) poly.progettes) progettes	×
0000-6 0011 0000-5 0012 0000-5 0012 0000-5 0012 0000-5 0012 0000-5 0012 0000-5 0012 0000-5 0012 0000-2 0012 0000-2 0012 0000-2 0012 0000-3 0012 0000-3 0012 0000-3 0012 0000-3 0012 0000-3 0012 0000-3 0012 0000-3 0012 0000-3 0012 0000-3 0012 0000-3 0012 0000-3 0012 0000-3 0012 0000-3 0012 0000-3 0012 0000-3 0012 0000-3 0012	1,00005 1,00007 1,00007 1,00001 1,00001 1,00001 1,00001 1,00001 1,00002 1,00002 1,00002 1,00002 1,00002 1,00002 1,00005 1,0005 1,	SEXTO-CAMERIE SALACE-SEXTAC SALACE-SEXTAC 2004/COLT/DA INSIGN-VERCOLE IREASEV-ASSIGN IREASEV-ASSIGN INSIGN-SEXTAR SALACESES SA	1005444404 334933064 4603340254 9403340254 9403340254 94033500254 9403350045 9403350045 9403350045 9403350045 94034728 935042849 935042849 935042849 935042849 935042849	ave In: Anonymizer als [gs] CP] Crage) Herscript) solp.thmi) solp.thmi) solp.thmi) soptanopertes	× Seter
000005 000005 000005 00000 000005 00000 000005 00000 000005 00000 000005 00000 000005 00000 000005 00000 000005 00000 000005 00000 000005 00000 000005 00000 000005 00000 000005 00000 000005 00000 000005 00000 000005 00000 000005 00000 000005 00000 000005 00000	1 090007 1 1090019 1 090019 1 090019 1 090019 1 090019 1 090029 1 090029 1 090029 1 090029 1 090029 1 090029 1 090029	NUTTO-CEND SALLACE" SERVICE POLL-CLIVEA IBBN: VECCULE BESSLE'-RESOLUT RESSLE'-RESOLUT RESSLE'S AND RESSLE'S AND RESSLE AND RESSLE RES	339530662 3622336026 373423861 645357487 645357487 645357487 6475427473 6476427473 64744728 64744728 64744728 615501228 615501228 615501228 61550124 6155012000000000000000000000000000	ave In: Anonymizer ata [parios] CP] Crage) filer.script) help.time) polp.troperties) program.properties	×
00003 0013 00013 0013 00013 0013 00014 0013 00015 0013 00016 0013 00017 0013 00018 0013 00019 0013 00019 0013 00019 0013 00019 0013 00019 0013 00019 0013 00019 0013	1 890009 100011 100011 100013 100013 100013 100013 100023 100029 100029 100029 100029 100029	NALLOCATORILLY DOUGHOUTSY HIBBLY RECOLL HEASTLY: ANDERN MERSLY MARK MERSLY MARK MERSLY MARK MERSLY MARK MERSLY MERSL MERSLY MERSL MERSLY MERSL MERSLY MERSL MERSLY MERSL	34627934623 4025346234 64155774237 64535774237 6453259005 64704272713 0774455273 1285947755 1285947755 1285947755 128594266 119359424 119325646 119723529743	ave In: Anonymizer ata [cranos] CP] Crage) litter.script) bolp.thmi) popt.properties) program.properties	
100 0000000 00000000 0000000000 000000000000000000000000000000000000	1990011 1990013 1990013 1990017 1990019 1990021 1990029 1990029 1990029 1990023 1990023	POOL "CATADA INDEL "RECOLL RESULT- ADDREM RESULT- ADDREM RESULT- ADDREM RESULT- ADDREM RESULT- ADDREM RESULT- ADDREM RESULT- STREAM RESULT- STREAM RESULT- RESULT- STREAM S	4003500256 972403847 645357407 645350061 6470420075 6470420075 726464728 6470427473 7265467728 015001228 015001228 015001228 015001228 015001228	ave In: Anonymizer ata [dramos] GP] CP]) filter.script) help.timi) koj4j.properties) program.proporties	×
00013 0013 00014 001 00015 001 00016 001 00017 001 000018 001 000019 001 000019 001 000019 001 000019 001 000019 001 000019 001 000019 001 000019 001 000019 001 000019 001 00019 001	1 00013 1030015 1030017 1030017 1030021 1030021 1030022 1030025 1030025 1030025 1030023	INNER VERCULE BEASLEY (* ANDERN KENERSYNNE LODOR ELLIDARTYS KENERSYNNE KENERSYNE KENERSYNE KENERSYNE KENERSYNE KENERSYNE KENERSYNE KENERSYNE KENERSYNE	173453841 641597417 6453350041 66143530041 6614389035 6674453273 7361447728 615301228 6154135644 6154135644 6154135644	ave In: Anonymizer ata [ga [CP] [torage) filter.script) telp.timi) telp.timi) telp4j.properties) program.proporties	× • • • • • • •
000014 0011 000016 0011 000018 0011 000019 0011 000019 0011 000029 0011 000028 0011 000029 0011 000029 0011 000028 0011 000029 0011 000029 0011 000029 0011 000029 0011 000029 011 000029 011 000029 011 000029 011 000029 011 000029 011 000029 011 000029 011 000029 011 000029 011 000029 011 000029 011 000029 011 000029 011 000029 011 000029 011	1 000015 1 000015 1 000019 1 000021 1 000029 1 000029 1 000029 1 000029 1 000029 1 000023 1 000033	HRANEY-ANDRY MILLIN-THUR KUNDA-BLIDARTH KUNDA-HLIDARTH KUNDAR-KUNDAR KUNDAR-KUNDAR HRANER-KUNDAR HRANER-KUNDAR HRANER-KUNDAR HRITZ-STREAM	6415974217 6452500642 6616428003 6676427803 6776425273 9877730 7355997750 7366487728 6155012288 6155012288 6155012288 6155012288 6155012288 6155012288 6164155646 617503222745	ave In: Anonymizer ata [branes] GP [torage) filter.script) selp.thmi) selp.thmi) selp.thmi) program.proporties	×
000016 0013 000018 0013 000020 0013 000024 0013 000026 0013 0000276 0013 000018 0013 000019 0013 000019 0013 000019 0013 000019 0013 000019 0013 000019 0013 000019 0013 000019 0013 000019 0013	1090017 1090023 1090023 1090025 1090025 1090025 1090025 1090025 1090023 1090023	NEERS HOUR LOOK-PEILARETH RATURA HEIGENER BUTHON-SCOTT BUTHON-SCOTT BUTHON-SCOTT BUTHON-SCOTT BUTHON-SCOTT BUTHON-SCOTT BUTHON-SCOTT BUTHON-SCOTT BUTHON-SCOTT BUTHON-SCOTT BUTHON-SCOTT	6451050061 6664398035 6670427471 7255997755 7361647728 8155012288 81350248865 8135248865 8135248865 8735125229	avé In: Anonymizer ata [branos [go [CP [torage) filter.script) help.thml) top4.properties) program.proporties	× 38118 •
000018 0013 000030 0013 000023 0013 000024 0013 000028 0013 000038 0013 000039 0013 000030 0013 000034 0013 000035 0033 000038 0013 000038 0013	1090019 1090023 1090023 1090025 1090027 1090027 1090029 1090023 1090023 1090025	VORG-BLILDERTE KYVER-BLILDERTE BUTHON-SCOTT KURCHES-SCORE KURCHES-GUNE HURCHES-GUNE MUTH-STREMATE JAMES-TECHNS		ave Jac Anonymizer ata [branes [gs [CP [torage) filter.script) holp.html) kop4j.properties) propram.proporties	×
000020 0011 000723 0011 000724 0011 000724 0011 000728 0011 000732 0011 000732 0011 000734 0011 000734 0011	1090021 1090023 1090027 1090027 1090029 1090029 1090031 1090033 1090033	REVERA-MECEDARL RADINAR-SCOTT RECORDS-SCOTT RECORDS-VALUEAR HERALES-VELLEAR HERALES-VELLEAR META-STRENATE JANES-TRENATE JANES-TRENAT	erro427411 er74825273 Sev 735599732 2 7361447728 2 815502288 2 815502488 2 81550246 2 8752322745 22	In: Anonymizer ata [branes [gs [CP [torage) litter.script) help.thmi) kog4j.properties) program.properties	• 6 3 1 382
000022 0011 00074 0011 000026 0011 000028 0011 000030 0012 000030 0011 000030 0011 000036 0011 000036 0011	1030023 1030025 1030027 1030027 1030031 1030033 1030035	RAUTHON-SCOTT REDURSS-GAMERI HERALISS-WILLIAM HOMER-SEONEL SANTHER-GAME MELER-STEPSANTE JANES-TRENAS	6774625273 Sev 7255597732 20 736164772# 20 815501228# 20 8159218840 20 81548156246 20 8732322741 04	III: Anonymizer ala [branes [cp [torage) filter.script) help.thmi) help.thmi) help.thmi) program.properties	
000024 0011 000026 0011 000028 0011 000030 0011 000030 0011 000036 0011 000036 0011	1030025 1030027 1030029 1030031 1030035 1030035	REDGERS*SAMEL HERALES*VELLAN HOMER*REDREY HERES*GEREALE HERES*GEREALE AMELS*GEREALE	1255997752 7361647728 8155012288 8189244869 8548156246 8732322741	ata [branes [lgs [CP [torage) filter.script) holp.timi) log4j.properties) program.proporties	
000026 0011 000028 0011 000038 0011 000032 0011 000036 0011 000036 0011	1030027 1030029 1030031 1030033 1030033	HEBALES-VELLEAN HEOVER-SCONEL EXECUEZ-GRACE HEVER-STEPENNE LAHES-THEBAS	734144772# 015501220# 0199244869 0548156246 0732322741	branes [ligs [CP [torage) help.timi) kog4j.properties) program.properties	
000028 0011 000030 0011 000032 0011 000034 0011 000036 0011	1000029 1000031 1000033 1000033 1000033	HOOVER-ROOMEL BANCHEZ-GRACE MELER-STEPISACIE JAMES-THEMAS	8155012288 8189344849 8548156246 8732322741	ops [] CP [] torage	1004j.properties program.proporties	
000030 0011 000032 0011 000034 0011 000036 0011	1090031 1090033 1090035	EJACHEZ-CEJACE NEVER-STEPEAKIE JANES-TRENAS	8189344869 8548156246 8732322741	CP [program.properties	
000032 0011 000034 0011 000036 0011 000038 0011	1090033 1090035	MEXER-STEPHANIE JAMES-THEMAS	8548156246 8732322745	torage	1 broth mich ober and	
000034 0011 000035 0011 000038 0011	1030035	JANES-TECHAS	8752322745	too age		
000036 0031	700001			normanitar lar		
00038 0011		FULLER-ALICIA	8824547497	com anonyminer script		
	1000039	JANES WICHAEL	6969193730	com-accoputaci, acripi		
000040 0011	1000041	TECHTISCH-MURCON	9189822998 File 1	ame: Indecidar		
000042 0011	1000045	MILLER-BEVERLY	9894340694 Eden	of Type: All Files		
						Tan I canad
						Sava Carda
						☆
						and the second
						save button

7. The result is an Excel sheet named "index.xls" and is located in 'C:\Anonymizer'

I data 2/7/2021 2:00 μμ Φάκελος αρχείων I libraries 1/7/2021 11:47 πμ Φάκελος αρχείων I logs 2/7/2021 159 μμ Φάκελος αρχείων SCP 1/7/2021 11:47 πμ Φάκελος αρχείων Storage 2/7/2021 2:02 μμ Φάκελος αρχείων Anonymizer 1/7/2021 11:47 πμ Φάκελος αρχείων Anonymizer 1/7/2021 11:47 πμ Αρχείο SCRIPT 285 KB i dicom-anonymizer 2/7/2021 1:59 μμ Αρχείο SCRIPT 285 KB i filter 1/7/2021 11:47 πμ Brave HTML Docu 13 KB I Index 2/7/2021 2:24 μμ Φύλλο εργασίας τ 6 KB I log4j 1/7/2021 11:47 πμ Aρχείο PROPERTIES 2 KB Index 2/7/2021 2:24 μμ Φύλλο εργασίας τ 6 KB I program 2/7/2021 2:01 μμ Αρχείο PROPERTIES 1 KB I solotex 2/7/2021 2:01 μμ Φύλλο εργασίας τ 1 KB	Όνομα	x	Ημερομηνία τροποποί	Τύπος	Μέγεθος
Ibbraries 1/7/2021 11:47 πμ Φάκελος αρχείων Iogs 2/7/2021 1:59 μμ Φάκελος αρχείων SCP 1/7/2021 11:47 πμ Φάκελος αρχείων Storage 2/7/2021 2:02 μμ Φάκελος αρχείων Anonymizer 1/7/2021 11:47 πμ Αρχείο SCRIPT 285 KB dicom-anonymizer 2/7/2021 1:59 μμ Αρχείο SCRIPT 285 KB filter 1/7/2021 11:47 πμ Apχείο SCRIPT 0 KB Maine 1/7/2021 11:47 πμ Brave HTML Docu 13 KB Index 2/7/2021 2:24 μμ Φύλλο εργασίας τ 6 KB Iog4j 1/7/2021 11:47 πμ Apχείο PROPERTIES 2 KB Index 2/7/2021 2:24 μμ Φύλλο εργασίας τ 6 KB orgaram 2/7/2021 2:01 μμ Apχείο PROPERTIES 1 KB rogaram 2/7/2021 2:01 μμ Φύλλο εργασίας τ 1 KB Index 2/7/2021 2:01 μμ Φύλλο εργασίας τ 1 KB	📜 da	ta	2/7/2021 2:00 µµ	Φάκελος αρχείων	
I logs 2/7/2021 1:59 μμ Φάκελος αρχείων SCP 1/7/2021 11:47 πμ Φάκελος αρχείων Storage 2/7/2021 2:02 μμ Φάκελος αρχείων Storage 2/7/2021 2:02 μμ Φάκελος αρχείων Anonymizer 1/7/2021 11:47 πμ Αρχείο JAR 164 KB Gicom-anonymizer 2/7/2021 1:59 μμ Αρχείο SCRIPT 265 KB filter 1/7/2021 1:296 μμ Αρχείο SCRIPT 0 KB V help 1/7/2021 1:47 πμ Brave HTML Docu 13 KB D index 2/7/2021 2:24 μμ Φύλλο εργασίας τ 6 KB 4 1094j 1/7/2021 11:47 πμ Αρχείο PROPERTIES 1 KB Dindex 2/7/2021 2:24 μμ Φύλλο εργασίας τ 6 KB 4 1094j 1/7/2021 11:47 πμ Αρχείο PROPERTIES 1 KB Dindex 2/7/2021 2:24 μμ Φύλλο εργασίας τ 1 KB	📕 lib	raries	1/7/2021 11:47 πμ	Φάκελος αρχείων	
SCP 1/7/2021 11:47 πμ Φάκελος αρχείων Storage 2/7/2021 2:02 μμ Φάκελος αρχείων Anonymizer 1/7/2021 11:47 πμ Αρχείο SCRIPT 285 K8 dicom-anonymizer 2/7/2021 1:39 μμ Αρχείο SCRIPT 0 K8 filter 1/7/2021 1:36 μμ Αρχείο SCRIPT 0 K8 plintex 2/7/2021 2:36 μμ Φώλο εργασίας τ 6 K8 glog4j 1/7/2021 1:47 πμ Aρχείο PROPERTIES 2 K8 glog4j 1/7/2021 1:47 πμ Αρχείο PROPERTIES 2 K8 glog4j 1/7/2021 1:47 πμ Αρχείο PROPERTIES 2 K8 glog4j 1/7/2021 1:24 μμ Φύλο εργασίας τ 6 K8 glog4j 2/7/2021 2:24 μμ Φύλο εργασίας τ 1 K8 glog4s 2/7/2021 2:24 μμ Φύλο εργασίας τ 1 K8	📙 log	gs	2/7/2021 1:59 µµ	Φάκελος αρχείων	
Storage 2/7/2021 2:02 μμ Φάκελος αρχείων 22 Anonymizer 1/7/2021 11:47 πμ Aρχείο JAR 164 KB a dicom-anonymizer 2/7/2021 1529 μμ Aρχείο SCRIPT 285 KB a dicom-anonymizer 2/7/2021 12:36 μμ Aρχείο SCRIPT 0 KB a filter 1/7/2021 11:47 πμ Brave HTML Docu 13 KB b Index 2/7/2021 2:24 μμ Φύλλο εργασίας τ 6 KB 4 log4j 1/7/2021 11:47 πμ Aρχείο PROPERTIES 2 KB 4 program 2/7/2021 2:24 μμ Φύλλο εργασίας τ 6 KB 2 program 2/7/2021 2:24 μμ Φύλλο εργασίας τ 1 KB	📜 sc		1/7/2021 11:47 πμ	Φάκελος αρχείων	
Ω Anonymizer 1/7/2021 11:47 πμ Αρχείο JAR 164 KB □ dicom-anonymizer 2/7/2021 1:59 μμ Αρχείο SCRIPT 285 KB □ filter 1/7/2021 1:236 μμ Αρχείο SCRIPT 0 KB ♥ help 1/7/2021 11:47 πμ Brave HTML Docu 13 KB □ Index 2/7/2021 2:24 μμ Φώλο εργασίας τ 6 KB ∅ log4j 1/7/2021 1:47 πμ Αρχείο PROPERTIES 2 K8 ☑ rogram 2/7/2021 2:24 μμ Φώλο εργασίας τ 1 KB □ sindex 2/7/2021 2:24 μμ Φύλο εργασίας τ 1 KB	📜 Ste	orage	2/7/2021 2:02 µµ	Φάκελος αρχείων	
□ dicom-anonymizer 2/7/2021 1:59 μμ Αρχείο SCRIPT 285 KB □ filter 1/7/2021 1:236 μμ Αρχείο SCRIPT 0 KB ♥ help 1/7/2021 11:47 πμ Brave HTML Docu 13 KB ■ Index 2/7/2021 2:24 μμ Φύλλο εργασίας τ 6 KB 10g4j 1/7/2021 1:47 πμ Aρχείο PROPERTIES 2 KB program 2/7/2021 2:21 μμ Αρχείο PROPERTIES 1 KB -Sindex 2/7/2021 2:01 μμ Αρχείο PROPERTIES 1 KB	🚉 An	onymizer	1/7/2021 11:47 πμ	Αρχείο JAR	164 KB
Initiation 1/7/2021 12:36 μμ Αρχείο SCRIPT 0 KB Phelp 1/7/2021 11:47 πμ Brave HTML Docu 13 KB Index 2/7/2021 2:24 μμ Φύλλο εργασίας τ 6 KB Index 2/7/2021 2:21 μμ Αρχείο PROPERTIES 2 KB Index 2/7/2021 2:01 μμ Αρχείο PROPERTIES 1 KB Index 2/7/2021 2:01 μμ Αρχείο PROPERTIES 1 KB Index 2/7/2021 2:01 μμ Φύλλο εργασίας τ 1 KB	📄 dia	com-anonymizer	2/7/2021 1:59 µµ	Αρχείο SCRIPT	285 KB
Υ help 1/7/2021 11:47 πμ Brave HTML Docu 13 KB P index 2/7/2021 2:24 μμ Φύλλο εργασίας τ 6 K8 Iggin 10/7/2021 11:47 πμ Apχeio PROPERTIES 2 K8 Iggin program 2/7/2021 2:24 μμ Apχeio PROPERTIES 1 K8 Index 2/7/2021 2:24 μμ Φύλλο εργασίας τ 1 K8	📄 filt	ter	1/7/2021 12:36 µµ	Αρχείο SCRIPT	0 KB
Index 2/7/2021 2:24 μμ Φύλλο εργασίας τ 6 KB Iog4j 1/7/2021 11:47 πμ Αρχείο PROPERTIES 2 KB program 2/7/2021 2:01 μμ Αρχείο PROPERTIES 1 KB - \$Index 2/7/2021 2:01 μμ Αρχείο PROPERTIES 1 KB	🦁 he	lp	1/7/2021 11:47 πμ	Brave HTML Docu	13 KB
Ιοg4j 1/7/2021 11:47 πμ Αρχείο PROPERTIES 2 KB program 2/7/2021 2:01 μμ Αρχείο PROPERTIES 1 KB - \$Index 2/7/2021 2:24 μμ Φύλλο εργασίας τ 1 KB	斗 Inc	dex	2/7/2021 2:24 µµ	Φύλλο εργασίας τ	6 KB
2/7/2021 2:01 μμ Aρχείο PROPERTIES 1 KB 3μ ~\$Index 2/7/2021 2:24 μμ Φύλλο εργασίας τ 1 KB	🥥 log	g4j	1/7/2021 11:47 πμ	Αρχείο PROPERTIES	2 KB
🚈 ~\$Index 2/7/2021 2:24 μμ Φύλλο εργασίας τ 1 KB	🥥 pro	ogram	2/7/2021 2:01 µµ	Αρχείο PROPERTIES	1 KB
	- 🖬 ~\$	index	2/7/2021 2:24 µµ	Φύλλο εργασίας τ	1 KB

X Carbo Poliziono B J J J T Trougerstee A A B J J T Trougerstee A A B J J A B A B J A B A B A B A B A B A B A B A B A B A B A B A B A B A B A B A B A B A B A A B A B A A B A A B A A B A	11 - År Ar E E 2 Δ × Δ - 5 E E 01 PHI-PatientName E E E 04/05/00+00/06L85 DAVID5/00+00/06L85 DAVID5/00+00/06L85 E 04/05/00+00/06L85 SEXTOR-CARMEN MARTINCIAD MARTINCIAD 04/05/00+00/06L85 SEXTOR-CARMEN MARTINCIAD DOI: 10+04	Prin Petention 1002 1002 1001 1000 1001 1000 1002 1000 1002 1002 10	ύλωση κειμένου νειοη και στοίχιση στι η ΑΝΟΝ-StudyDate 20040716 20100219 19971005 19910815 199540411	ка́лтро - Б АНОН-Ассеззіон 000001 000001 000002 000003 000004	9 сто 120 - 201 Аргіцьіс 20130713 20130713 20130713 20130713 20130713 20130713 20130713 20130713 20130713	Moseowingen Mogen Moseowingen <th>μοποίηση tiverκα ~ κ λ</th> <th>Στυλ εελιών ~</th> <th>Εισαγωγ Διογραφ Μορφοπ Κολε</th> <th>ή ~ ή ~ ιοίηση ~ λ</th> <th>∑ ~ Ta</th> <th>Αντροποία το το</th> <th>О Ейреал & етиλоуп ~</th> <th>P</th>	μοποίηση tiverκα ~ κ λ	Στυλ εελιών ~	Εισαγωγ Διογραφ Μορφοπ Κολε	ή ~ ή ~ ιοίηση ~ λ	∑ ~ Ta	Αντροποία το	О Ейреал & етиλоуп ~	P
Construction C	C ID PHI-PatientName DAVIDSON*DOUGLAS DAVIDSON*DOUGLAS SEXTON*CARMEN MARTIN*CHAD WALACE*SHELLY DOCIEGOLIBIA	E 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2	Participanti and a statution of the state of	F ANON-Accession 000001 000002 000002 000003 000004	G PHI-StudyDete 20130713 20130713 20130713 20010705 19610705	итб броис ч ше т 5 Ут 19 Ин-Ассеззіол 20130714E864535 20130714E864535 18667928 20010706E403961 1965070705513320	Ι	J	K Mopeon Kohd	i i L	M	ολτράρεσμα ~ Επεξεργοσία Ν	σπιλογή ~ Ο	P
Telegon 6 Tegentities A X F A X F A A A A A <	PRI FSI C DPHI-PatientName DAVIDSON*DOUGLAS DAVIDSON*DOUGLAS SEXTON*CARMEN MARTIN*CHAD WALLACE*SHELLY DOCIECOLINIA	D PHI-PatientiD 1001.00000-37718 1002.00000-53238 3209648408 339833062 3642991663 642991663	m E ANON-StudyDate 20040716 20100219 19971005 19910815 19540411	5 ANON-Accession 000001 000002 000002 000003 000004	G PHI-StudyDate 20130713 20130713 19990923 20010705 19610705	H PHI-Accession 20130714E864535 20130714E864535 18687928 20010706E403961 18607025513320	Ĩ)	K	L	M	ExcEspycoia N	0	P
A B A ANOVA Failed A B ANOVA Failed 022000051 0022000001 002000001 0022000001 002000001 0022000001 002000000 0022000001 002000000 0022000001 002000000 0022000005 002000000 002200005 002000005 002000005 002000005 002000005 002000005 002000005 002000005 0020005 002005 002005 0005 0005 0005 000	C DAVIDSON*DOUGLAS DAVIDSON*DOUGLAS DAVIDSON*DOUGLAS SEXTON*CARMEN MARTIN*CHAD WALLACE*SHELLY DOCIEGOLINA	D PHI-PatientID 1001.000000-37718 1002.00000-53238 3209648408 339833062 3642991663 4005360156	E ANON-StudyDate 20040716 20100219 19971005 19910815 19540411	F ANON-Accession 000001 000002 000002 000003 000004	G PHI-StudyDate 20130713 20130713 19990923 20010705 19610705	H PHI-Accession 20130714E864535 20130714E864535 18687928 20010706E403961 19650702E513230	I	J	K	L	м	N	0	P
A B 02200001 02200001 02200001 02200001 02200001 02200001 02000002 02200001 02000003 02200001 02000003 02200001 02000003 02200001 02000004 02200001 02000006 02200001 02000006 02200007 02000008 02200007 02000008 02200007 02000008 02200007 02000008 02200007 02000009 02200007 02000008 02200007 02000009 02200007	C ID PHI-PatientName DAVIDSON*DOUGLAS DAVIDSON*DOUGLAS SEXTON*CARMEN MARTIN*CHAD WALLACE*SHELLY DOCIECOLINIA	D PHI-PatientID 1001.000000-37718 1002.000000-53238 3209648408 339833062 3642991663 4005360156	E ANON-StudyDate 20040716 20100219 19971005 19910815 19540411	F ANON-Accession 000001 000001 000002 000002 000003 000004	G PHI-StudyDate 20130713 20130713 19990923 20010705 19610705	H PHI-Accession 20130714E864535 20130714E864535 18687928 20010706E403961 196507076E13320	1	J	К	L	М	N	0	P
Another 022000003 022000003 022000003 022000003 022000003 022000003 022000003 022000003 022000003 022000004 022000005 022000005 022000005 022000005 022000006 022000005 022000006 022000007 022000006 022000007 022000007 022000007 022000008 022000007 022000009 022000009 022000009 022000009	ID PHI-PatientName DAVIDSON^DOUGLAS DAVIDSON^DOUGLAS SEXTON^CARMEN MARTIN^CHAD WALLACE^SHELLY BOOLEADLINIA	PHI-PatientID 1001.000000-37718 1002.00000-53238 3209648408 339833062 3642991663 4035360156	ANON-StudyDate 20040716 20100219 19971005 19910815 19540411	ANON-Accession 000001 000001 000002 0000003 0000003	PHI-StudyDate 20130713 20130713 19990923 20010705 19610705	PHI-Accession 20130714E864535 20130714E864535 18687928 20010706E403961 19650702E513330								
02200001 02200001 02200001 02200001 02200002 02200003 0200002 02200003 02000003 02200004 02000004 02200005 02000006 02200005 02000006 02200005 02000006 02200007 02000008 02200007 02000008 02200007 02000008 02200007 02000008 02200007 02000008 02200007 02000008 02200007 0200009 02200007	DAVIDSON^DOUGLAS DAVIDSON^DOUGLAS SEXTON^CARMEN MARTIN^CHAD WALLACE^SHELLY BOOLEADLINA	1001.000000-37718 1002.000000-53238 3209648408 339833062 3642991663	20040716 20100219 19971005 19910815 19540411	000001 000001 000002 000003 000003	20130713 20130713 19990923 20010705 19610705	20130714E864535 20130714E864535 18687928 20010706E403961 19610707E512320								
0x2000001 0x2000002 0x2000002 0x2000002 0x2000003 0x2000004 0x2000004 0x2000005 0x2000005 0x2000005 0x2000006 0x2000005 0x2000006 0x2000005 0x2000006 0x2000005 0x2000006 0x2000007 0x2000007 0x2000006 0x2000008 0x2000007 0x2000009 0x2000009 0x20000090 0x2000009	DAVIDSON^DOUGLAS SEXTON^CARMEN MARTIN^CHAD WALLACE^SHELLY BOOLEAOLBRA	1002.00000-53238 3209648408 339833062 3642991663 4005360156	20040716 20100219 19971005 19910815 19540411	000001 000001 000002 000003 000004	20130713 20130713 19990923 20010705 19610705	20130714E864535 20130714E864535 18687928 20010706E403961 19610707E512320								
02200003 02200003 02200003 02200003 0200003 02200003 0200003 02200004 02000004 02200005 02000005 02200005 02000006 02200005 02000006 02200007 02000006 02200007 02000006 02200007 02000006 02200007 02000007 02200007 02000008 02200007 02000009 02200007	DAVIDSON^DOUGLAS SEXTON^CARMEN MARTIN^CHAD WALLACE^SHELLY	1002.000000-53238 3209648408 339833062 3642991663 4005360156	20100219 19971005 19910815 19540411	000001 000002 000003 000004	20130713 19990923 20010705 19610705	20130714E864535 18687928 20010706E403961								
022000002 022000003 7022000023 022200005 702200003 022200005 702200005 02200005 702200005 02200005 702200005 02200005 702200005 02200005 702200005 02200005 702200005 02200005 702200005 022200007 702200005 022200007 702200005 022200007 702200005 022200007 702200005 022200007 702200005 022200007 702200005 022200007	SEXTONACARMEN MARTINACHAD WALLACEASHELLY	3209648408 339833062 3642991663	20100219 19971005 19910815 19540411	000001 000002 000003 000004	20130713 19990923 20010705 19610705	20130714E864535 18687928 20010706E403961 19610707E512320								
002000002 002000001 002000013 002000001 002000014 0022000001 002000004 0022000005 002000004 0022000005 002000000 0022000007 002000000 0022000007 002000000 0022000007 002000000 0022000007 0020000000 0022000009 0020000000 00220000010 0120000000 00220000010	SEXTONACARMEN MARTINACHAD WALLACEASHELLY	3209648408 339833062 3642991663	19971005 19910815 19540411	000002 000003 000004	19990923 20010705 19610705	18687928 20010706E403961 19610707E512320								
0 0x2200003 0x2200004 0 0x2200005 0x2200005 1 0x2000005 0x2200005 0 0x2000005 0x2200005 0 0x2000007 0x2200007 0 0x20000007 0x2200007 0 0x2000000 0x2200007 0 0x2000009 0x2200009 0 0x2000009 0x22000019	MARTINACHAD WALLACEASHELLY	339833062 3642991663	19971005 19910815 19540411	000002	19990923 20010705 19610705	18687928 20010706E403961								
0220000033 02200004 02200004 022000051 20200004 02200005 40200006 02200005 60200006 02200007 60200006 02200007 60200006 02200007 7 02200006 9 02200007 9 02200007 9 02200007 9 02200007 9 02200007 9 02200007 9 02200007 9 02200007 9 02200007 9 02200007	WARTINACHAD WALLACEASHELLY	339833062 3642991663	19910815 19540411	000003	20010705	20010706E403961								
0 02200004 02200005 1 02200006 02200005 3 02200006 02200007 5 00200000 02200007 8 02200000 02200007 9 02200000 02200009 0 02200009 02200009	WALLACE^SHELLY	3642991663	19910815	000003	20010705	20010706E403961								
0 1020200004 002200005 2 1020200006 002200007 4 1020200006 002200007 6 102000008 002200007 8 102000008 002200009 0 102000009 002200009 0 102000009 002000010	WALLACE*SHELLY	3642991663	19540411	000004	19610705	195107075512320								
1 2 000000005 002000005 002000007 5 002000007 002000008 002000008 7 002000008 002000009 9 0 000000009 9 0 000000009 0 000000	ROOLEAOLIN/A	4025260156	19540411	000004	19610705	196107075512220								
22 002000005 00200005 33 44 002000006 002200007 55 16 002000007 002200007 17 18 002000008 002200009 19 002000008 002200009 10 002000009 002200009 10 002000001 012 11	DOOLEAOLIN/IA	4025260156				19010/070312330								
3 002000006 002000007 15 002000007 002000008 16 002000008 002000008 17 002000009 002000009 19 002000009 0020000010 10 0022000010 10	POOLE-OLIVIA	4023300230												
4 002000006 002200007 5 002000007 00200008 7 8 002000008 002000009 9 00 002000009 002000009 1 002000009			19900626	000005	19931025	978B4101								
5 6 002000007 002000008 7 8 002000008 00200009 9 9 002000009 00200009 10 002000009 002000010	HENRY^NICOLE	571403367												
7 7 8 002000008 002000009 9 0 002000009 002000010 11	05101511100511	Concerner 2	20140120	000006	20190711	20190712E924533								
8 002000008 002000009 9 002000009 002000009	BEASLEY^ANDREW	6415974217				(()))								
9 0 002000009 002000010	AN/ DCAMADY	CATAOTOLEA	19781222	000007	19880609	65385522								
0 002000009 002000010	MITERSTWARK	0451050501	TOFOOT 12	000000	00000000	105107205522422								
1	VOLINGAELIZABETH	6614339035	19580512	000008	19010728	19010729E532433								
	TOUNGPELIZABETH	0014238035	10060937	000000	20001017	93483700								
22 002000010 002000011	DIVERAAMICHAEL	6670437471	19900027	000005	20001017	03403733								
3	RIVERACINICHAEL	00/042/4/1	19990426	000010	20000526	20083529								
4 002000011 002000012	KALIEMANASCOTT	6774825273	19990420	000010	20000320	20003329								
5	NOTION SCOTT	UT HOLDETS	20091216	000011	20130911	201309125801911								
002000012 002000013	RODGERSASAMUEL	7255997752												
7	Sector Specific L		19610525	000012	19620109	19620110E237653								
002000013 002000014	MORALES*WILLIAM	7361647728												
9			20100723	000013	20130721	20130722E607303								
index (+)						1 4						-	_	

1.5 Adding new patients with unique patient id

In case you want to add new patients to de-identify follow the steps below:

- 1. Add the patient folder(s) to the root folder containing all patients.
- 2. Navigate to 'C:\Anonymizer'
- 3. Double-click Anonymizer
- 4. Click 'Directory' (tab on top)
- 5. Navigate to the root folder containing all the patients from the left panel
- 6. Double click the root folder to expand it
- 7. Navigate to the new patient folder
- 8. Select it and click the 'Import' button
- 9. Repeat steps 7 and 8 for all new patient folders
- 10. Click 'Admin' (tab on top left)
- 11. Click 'Index' (tab on top left)
- 12. Click the 'List' button (button on the middle bottom)
- 13. If you don't see the newly added patient(s) on the list, please click the 'List' button again.
- 14. In case you have the .xls file containing the pseudonymization table open, close it.
- 15. Click the 'Save' button
- 16. Click 'Save' on the pop-up window
- 17. The result is an Excel sheet named "index.xls" and is located in 'C:\Anonymizer'

<u>1.6 Adding new patients where some patients have different patient IDs</u>

In case you want to add new patients and some of them have different ids among their studies, follow the next steps:

- 1. Create a new folder (e.g., 'New patients') and place all the patient folders inside that folder. All patient folders should be named after the patients' id. The patient folder for patients with multiple id's should be named after one of the different patient ids.
- 2. Navigate to 'C:\DicomEditor'
- 3. Double-click DicomEditor
- 4. In the 'Directory' tab, on the left pane navigate to the new folder you created at the beginning of this section (e.g., 'New patients')
- 5. Double click this folder to expand it
- 6. Click the new (in this example the 'New Patients') folder
- 7. Click on the 'SetPatientIDs' button on the bottom right
- 8. Wait until the results appear on the right pane.
- 9. Follow the steps in section 1.5 Adding new patients with unique patient id

1.7 Adding new Studies for an old patient with unique ID

Assuming a patient with a unique ID has performed new exams

and they must be added to the temporary infrastructure, follow the next steps:

- 1. Add the Study(ies) folder inside the corresponding patient folder
- 2. Navigate to 'C:\Anonymizer'
- 3. Double-click Anonymizer
- 4. Click 'Directory' (tab on top)
- 5. Navigate to the root folder containing all the patients from the left panel
- 6. Double click the root folder to expand it
- 7. Navigate to the patient folder containing the new Study(ies)
- 8. Select it and click the 'Import' button

1.8 Adding new Studies for an old patient and the Studies contain a new different ID for the patient

Assuming new Studies come up for a patient and those studies have been performed outside of the Data Provider's premises, thus they contain a different id for the patient, follow the next steps to de-identify the new Studies.

- 1. Create a new folder and name it, e.g., 'New studies for patient David Douglas'. Inside that folder create another folder. The name of the inner folder must be the patient id as seen on the pseudonymization table for that patient (excel file that was exported during the first de-identification)
- 2. Inside the folder named after the patient id, place the new studies
- 3. Open DicomEditor
- 4. Navigate to the outer folder (e.g., 'New studies for patient David Douglas ')
- 5. Click that folder
- 6. Click the 'SetPatientIDs' button
- 7. Copy the folder containing the studies and place it under the root folder containing all the patients, inside the correct patient folder
- 8. Open Dicom Anonymizer by double-clicking Anonymizer (on the 'C:\Anonymizer' directory)
- 9. Navigate to the root folder (double click to expand it) and then to the patient folder containing the new studies
- 10. Double-click the patient folder to expand it
- 11. Click the new studies folder
- 12. Click the 'Import' button to de-identify the new studies
- 13. The de-identified new studies are located in the 'C:\Anonymizer\Storage' inside the correct patient folder

Annotation (ITK-SNAP)

This document shows how to install and use the ITK-SNAP. Currently supported platform for this guide is 64-bit Windows 10. Other platforms such as Linux and MacOS are supported with the only difference the procedure of downloading and installing ITK-SNAP (you can find more information in the following <u>link</u>.

2.1 Installation

- 1. Download <u>https://sourceforge.net/projects/itk-snap/files/itk-snap/3.8.0/itksnap-3.8.0-</u> 20190612-win64.exe/download
- 2. Double-click itksnap-3.8.0-20190612-win64.exe
- 3. Follow the installation instructions on the ITK-SNAP Setup Wizard.
- 4. Download label-descriptions.zip from Microsoft teams WP5/T5.2/label-descriptions.zip and unzip.

2.2 Use

A helpful video regarding this procedure is the following: <u>https://www.youtube.com/watch?v=ZE6634gpbto</u>. Moreover, there is a video library for ITK-SNAP manual and semi-automatic training here: <u>http://www.itksnap.org/pmwiki/pmwiki.php?n=Videos.SNAP3</u>.

- 1. Run ITK-SNAP from the Windows Start Menu.
- 2. Click 'File' (from menu top-left)
- In 'Image Filename' navigate to 'C:\Anonymizer\Storage\SiteID-PatientID-Modality-Timepoint' (e.g., 001-00002-CT-TP1)
- 4. Choose the corresponding DICOMDIR, or any DICOM file from the series.
- 5. Click 'Next' and 'Next'.
 - a. If this steps outputs an error, then click 'Back' and choose 'Generic ITK Image' in 'File Format'.
 - b. If the imaging modality is 2D (e.g., x-rays, mammography) then choose 'DICOM Single Image` in 'File Format'.
- 6. Keep note of:
 - a. SiteID is completed from each data provider (IDIBPAS, GOC, DISBA, UNITOV, UNS, UOA, HCS, Visaris, AUTH that corresponds to a SiteID, e.g., 001)
 - b. PatientID is created by CTP (e.g., 000001, 000002)
 - c. Modality is completed from each data provider (e.g., MRI, PETCT, CT, MMG, US, HP)

- d. Timepoint (BL, TP1, TP2, TP3)
- 7. From 'Menu', 'Segmentation' click 'Import Label Descriptions...' and choose file depending on the disease and imaging modality (e.g., lungcancer-ctscan.txt).
- 8. Follow either the 'Bounding box procedure/manual contouring' or the 'Segmentation' procedure.
- 9. Bounding box/manual contouring procedure
 - a. Select the polygon tool (button on top-left)
 - b. Adjust the box around the Region of Interest using the mouse (clicks)
- 10. Segmentation procedure
 - a. Select the segmentation tool (button with snake icon on top-left)
 - b. Adjust the box around the Region of Interest using the mouse (click and drag)
 - c. Click 'Segment 3D' (button on middle-left)
 - d. Adjust the 'Lower Threshold' and 'Upper Threshold' (sliders on middle-right) in such a way that you will see the surroundings are little as possible and your ROI the most
 - e. Click 'Next' (button on bottom-right)
 - f. Adjust 'Bubble radius' (slider on middle-right)
 - g. Click on ROI in one of the views in the middle
 - h. Click 'Add Bubble at Cursor' (button on top-right)
 - i. Repeat steps 12 and 13 as many times as possible (depending on the size of the ROI)
 - j. In case of a mistake in previous steps click the 'Delete Active Bubble' (button on middleright)
 - k. Click 'Next' (button on middle-right)
 - I. Click the 'Play' (button with play icon on middle-right)
 - m. Click 'Finish' when the results from the previous are satisfactory
- 11. Click 'File' (menu from top-left)
- 12. Click 'Save Segmentation Image' from the dropdown menu
- 13. Save as 'C:\Anonymizer\Storage\SiteID-PatientID-Modality-Timepoint\Untitled.nii.gz'

8.1.2 INCISIVE Tools

De-identification Tool

3.1 Installation and Use

Regarding the use of the tool, this is straightforward:

- Visit <u>https://dicom-de-identification-and-curation-tool.INCISIVE.iti.gr/.</u>
- The following screen appears:



The following visual indications and options are available and explained:

- **Input:** The following input elements allow the user to select the files/directories of the DICOM files that need to be de-identified/curated. The total file size of the loaded data is also shown.
 - **Load directory with DICOM files:** This input directory element allows the user to load his DICOM files when they are stored in nested directories.
 - Load session file: This input file element allows the user to load his session.json file that is useful in cases where de-identification sessions occur in different timepoints and some state needs to be preserved through time. It contains an array with object with the following fields:
 - patientPseudold: this helps in keeping the correspondence between the original and the de-identified patient IDs.

- secondsOffset: this contains a random number in the range of a day in seconds, that is added to the time-date related DICOM tags when the 'offset' option is chosen for 'date processing'.
- daysOffset: this contains a random number in the range of a decade in days, that is added to the date related DICOM tags when the 'offset' option is chosen for 'date processing'.
- **De-identification options:** The following checkbox elements allow the user to apply deidentification using a modification of the NEMA recommendations⁷.
 - **Retain safe private:** This checkbox element allows the user to keep the UIDs as mentioned in the 'Rtn Save Priv. Opt.' column of the NEMA recommendations.
 - **Retain UIDs:** This checkbox element allows the user to keep the UIDs as mentioned in the 'Rtn UIDs Opt.' column of the NEMA recommendations.
 - **Retain device identity:** This checkbox element allows the user to keep the UIDs as mentioned in the 'Rtn Dev. Id. Opt.' column of the NEMA recommendations.
 - Retain patient characteristics: This checkbox element allows the user to keep the UIDs as mentioned in the 'Rtn Pat. Chars. Opt.' column of the NEMA recommendations.
 - Date processing: This dropdown element allows the user to choose the date processing:
 - offset the dates as mentioned in the in the 'Rtn. Long Modified Dates Opt.' column of the NEMA recommendations.
 - keep the dates as mentioned in the in the 'Rtn. Long Full Dates Opt.' column of the NEMA recommendations.
 - remove the dates as mentioned in the in the 'Rtn. Long Full Dates Opt.' column of the NEMA recommendations by replacing 'K' with 'X'.
 - Retain Descriptors: This checkbox element allows the user to keep the description DICOM tags as mentioned in the 'Clean Desc. Opt.' column of the NEMA recommendations.
- Curation options: The following checkbox elements allow the user to apply curation.

- Patient ID postfix: This text field allows the user to input a postfix to the Patient ID DICOM tag for differentiation purposes e.g., when the files are being used in a multi-institutional study.
- Save path: (Currently disabled) This text field element allows the user to customize the export directory structure by using the DICOM tags (without the comma) separated by the forward slash '/' (this field is required to end with '/'). For example, the shown screenshot (as also loaded by default) applies the INCISIVE protocol directory which consists of:
 - 00100010: Patient Name (e.g., Jane Doe)
 - 00080060: Modality (e.g., MR)
 - 00200011: Series Number (e.g., 401)

• Output:

- **Save de-identified files:** This button element allows the user to download the processed DICOM files as a zip file into his local disk.
- Files processed: This element shows the number of files that have currently been processed after pressing the 'Save de-identified files' with respect to the amount of the total files loaded.
- DICOM tag values removed: This element shows the number of DICOM tag values that were removed after processing.
- **DICOM tag values replaced:** This element shows the number of DICOM tag values that were replaced after processing.

3.2 Video Tutorial

You can find a video tutorial for the INCISIVE de-identification tool, here: <u>dicom-de-identification-</u> <u>and-curation-tool-screencast.mp4</u>

Semi-automatic annotation tool

The semi-automatic annotation tool is a web-based client-side application being implemented within T5.2 of the INCISIVE project. It has the following features:

- Privacy oriented (since no data leave from the user's side)
- Easy to use (build with familiar web user interface)
- Only a modern browser is needed (there is no installation procedure since the user only needs to visit a link)

The last feature is one of the major differences of the tool compared with the currently used tool for annotation within INCISIVE: ITK-Snap. This feature should be considered along with the fact that there are usually constrained policies in environments such as hospitals and clinics disallowing installation of 3rd party tools.

4.1 Usage

Regarding the use of the tool, this is straightforward:

- Visit <u>https://semi-automatic-annotation-tool.INCISIVE.iti.gr/</u> with your browser.
- You will see the following screen:



The following visual indications and options are available and explained:

- Canvas on the right: This is where the loaded images will be shown.
- Model options:
 - Load model: This dropdown element allows the user to select the pretrained model to be used for the semi-automated annotation procedure depending on the modality and the disease type.
 - Model loaded: This text element serves as a visual indication of the model downloading progress.
 - Input shape: This text element shows the input shape that the loaded model accepts.
 - **Prediction shape**: This text element shows the prediction shape that the loaded model outputs (when empty then it is inferred it is the same as the input).
- Input options:
 - **Input element top:** This input element allows the user to load a set of DICOM or nifti files that need to be annotated.

- **Input element bottom**: This input element allows the user to load a set of nifti files that consists of the masks (e.g., created during a previous annotation session).
- Index: This range element allows the user to navigate between the loaded images and shows the total number of loaded images.
- HxW: This text element shows the height and width of the loaded images. If the loaded images have different shape than the one required by the pretrained model then they are automatically resized.
- Reset image value: This button element allows the user to 'normalize' the pixel values for better visualization. It does not change the original pixel values of the image.
- Inference options:
 - Predict current image: This button element allows the user to create the prediction from the loaded pretrained model for the currently viewed image and visualize the output to the canvas.
 - Predict all images: This button element allows the user to create predictions from the loaded pretrained model for all the loaded images and visualize the output to the canvas.
 - List of annotations: This list (currently having None and Lesion) consists of the type of annotations that the user may choose from. The colored rectangles on the left are clickable and change the corresponding annotation type.
 - Brush size: This button element allows the user to change the radius of the annotation brush for manually annotating in a more fine-grained manner or larger structures.
 - Save predictions to disk: This button element allows the user to save the annotated predictions to his local device in order to be used in a subsequent annotation session or when the annotation session is finished.

• Training:

• **Trainable (not currently used):** This text element shows whether the pretrained model is available for training on the browser. This is an experimental feature that is not currently being used and can be ignored.

Indicative workflow for end-user

- 1. Visit <u>https://semi-automatic-annotation-tool.INCISIVE.iti.gr/</u>.
- 2. Select the model depending on the task.
- 3. Load the file(s).
- 4. Predict.
- 5. Correct the results if needed.
- 6. Train only if correction took place.
- 7. Save the predictions.

4.2 Protocol for AI developers

Al developers who want to make available their pretrained model through the semi-automatic annotation tool should follow these instructions:

- 1. Develop and train a model in TensorFlowJS, Tensorflow or PyTorch having the following considerations in mind:
 - 1.1. fast inference (<1 second/data sample is recommended) and
 - 1.2. small size (100MB max for a saved model is recommended).
- 2. Convert the model to TensorFlowJS.
 - 2.1. For TensorFlow

https://www.tensorflow.org/js/tutorials/conversion/import_saved_model

- 2.2. For PyTorch <u>https://github.com/pbizopoulos/comprehensive-comparison-of-deep-learning-models-for-lung-and-covid-19-lesion-segmentation-in-</u> ct/blob/becf52d7cdb74eaad994fda2bba503460857e386/main.py#L582
- 3. Test the converted model by applying inference on a data sample.

- 4. Upload the model (model.json and shard files) and the example data on a server (e.g. GitHub, AWS etc.).
- Create a pull request on https://github.com/pbizopoulos/semi-automatic-annotation-tool tool that adds the model configuration with the following fields:
 - classNames: ["Array that", "contains", "the names of the classes"],
 - exampleDataUrl: "Optional: URL or local file of an example data of this model",
 - loss: "Optional: loss used for training (can be any of <u>https://js.tensorflow.org/api/latest/#Training-Losses</u>)",
 - machineLearningType: "type of machine learning algorithm for the model (can be any of 'image classification', 'image segmentation')",
 - metrics: ["Optional: can be any of https://js.tensorflow.org/api/latest/#Metrics"],
 - modelDownloadUrl: "URL or local filepath of model.json",
 - modelUploadUrl: "Optional: URL of remote server or local directory",
 - name: "name of the configuration (e.g. indicate imaging modality and disease)",
 - optimizer: "Optional: optimizer used for training (can be any of https://js.tensorflow.org/api/latest/#Training-Optimizers)",
 - projectUrl: "Optional: URL of the project that has created the model"

4.3 Video tutorial

You can find a video tutorial here: <u>semi-automatic-annotation-tool-screencast.mp4</u>

END