## Artificial Intelligence in Medicine: focus on Multi-Input Analysis



Alessandra Retico INFN Sezione di Pisa

Kick-off meeting, February 13, 2025

#### Outline

- Research context
- State of art
- The AIM\_MIA project
  - Objectives
  - Methods
  - Implementation
  - Connections
  - Expected impact
- People & Budget

A. Retico - Artificial Intelligence in Medicine: focus on Multi-Input Analysis

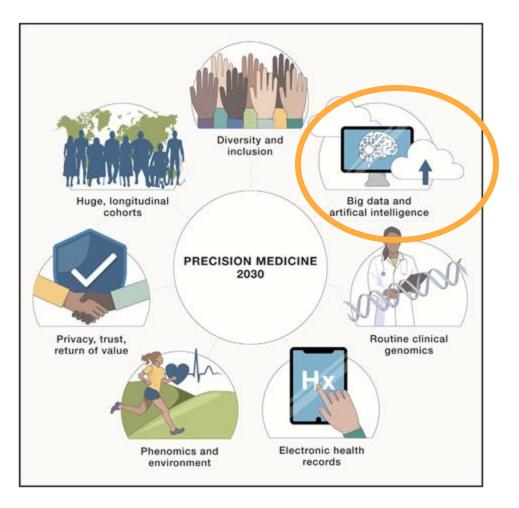


2/13/25 **2** 

# Artificial Intelligence in Medicine

- Artificial Intelligence (AI) is already *pervasive* in many domains of our daily life including the Healthcare sector.
- Al is expected to make an important contribution to achieving the goal of **precision medicine.**
- **Precision medicine** promises improved health by accounting for individual variability in genes, environment, and lifestyle.
- **Big data collections** and **advanced analytics approaches** (including Albased methods) are needed to fully exploit the potential of the large amount of digital information available today for each patient.

[Denny and Collins, Precision medicine in 2030—seven ways to transform healthcare. Cell 2021;184:1415–9. https://doi.org/10.1016/j.cell.2021.01.015.]



### Outline

- Research context
- State of art
- The AIM\_MIA project
  - Objectives
  - Methods
  - Implementation
  - Connections
  - Expected impact
- People & Budget

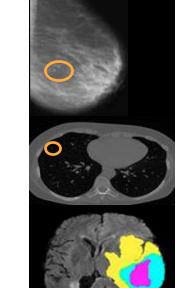
A. Retico - Artificial Intelligence in Medicine: focus on Multi-Input Analysis

2/13/25

3



## State of art, limitations and open issues



Since the 80s the scientific community has been developing AI-based systems to support clinicians (**Decision Support Systems, DSS**), especially in the field of diagnostic imaging, to facilitate the clinical workflows in radiological screening programs.

DSS are meant to support clinicians and not to replace them.

Several DSS are already available to clinicians as **CE-marked products** <u>https://radiology.healthairegister.com/</u>

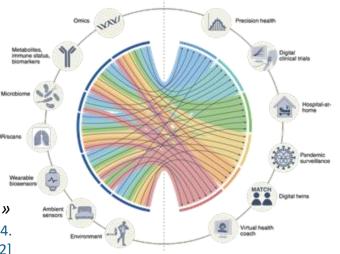
Nonetheless, most of them:

- lack in-depth validation on large and independent data samples
- need to move from narrow scopes (lesion identification/segmentation) to more clinically meaningful endpoints

**To make AI-tools predict relevant endpoints** (e.g. need for treatment, patient's survival, relapses after treatment) the **complementary information** encoded in omics data, electronic health records, imaging data should be exploited.

#### Al methods to mine multi-input data should be developed.

«... we are far better at collating and storing data, than we are at data analysis.» [Acosta et al (2022). Multimodal biomedical AI. *Nature Medicine*, 28(9), 1773–1784. https://doi.org/10.1038/s41591-022-01981-2]



#### Outline

- **Research context**
- State of art
- The AIM MIA project
  - **Objectives**
  - Methods •
  - Implementation
  - Connections
  - Expected impact
- People & Budget

A. Retico - Artificial Intelligence in Medicine: focus on Multi-Input Analysis

2/13/25

4



## State of art, limitations and open issues

Despite its relevance, how to effectively fuse the multi-modal medical data has received relatively little attention.

**Current literature** highlights the need for effective multi-input algorithms for medical data [Huang et al, 2020; Acosta et Al, 2022; Lipkova et al, 2022; Pei et al, 2023]

However, practical implementations are limited.

#### Data availability is crucial

- UK Biobank (since 2006), >500k participants, 30y of follow up:
- US "Million Veteran Program";
- US "All of Us Research Program" enrolled 400k diverse participants (target 1M participants)

#### Prediction of Pulmonary Embolism from CT and Electronic Medical Record (EMR)

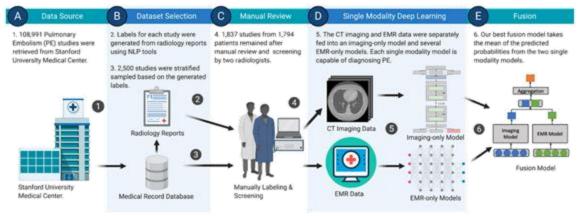


Figure 1. Overview of the workflow for this study. We extracted a total of 108,991 studies from Stanford Performance (AUC):

University Medical Center (A) and sampled a subset (B) for manual review (C). 1837 studies remained after screening by two radiologists and were used to train and evaluate our models. Single modality models were created (D) both as baselines for comparisons as well as components for the fusion models (E).

- CT alone: 79.1% EMR alone: 91.1%
- Late fusion: 94.7%

Huang, S.C. et al (2020). Multimodal fusion with deep neural networks for leveraging CT imaging and electronic health record: a case-study in pulmonary embolism detection. Scientific Reports. 10(1), 1–9. https://doi.org/10.1038/s41598-020-78888-w

#### Quantitative Imaging in Medicine and Surgery, Vol 14, No 1 January 2024

1115

#### Large Language Models (LLMs)

LLMs, like ChatGPT, which rely on Transformer architecture, can analyze images and are suitable to be fine-tuned to medical images and multiinput data analysis problems.

Tian, D. et al. (2024). The role of large language models in medical image processing: a narrative review. Quantitative Imaging in *Medicine and Surgery*, *14*(1), 1108–1121. https://doi.org/10.21037/qims-23-892

#### Table 5 Comparison of current functions of LLMs in the biomedical field

Current LLMs	ClinicalBERT	BioBERT	ChatGPT	BioMedLM	GeneGPT
Use of clinical notes	1	1	1	1	1
Biomedical text mining	N/A	1	1	1	1
Biomedical Q&A	N/A	N/A	1	1	1
Multimodal integration	×	×	1	N/A	1
Interactivity	×	×	1	1	1
Domain-specific	1	1	×	1	1

J, have the function; x, do not have the function. LLMs, large language models; GPT, Generative Pre-trained Transformer; BERT, Bidirectional Encoder Representations from Transformers; Q&A. Questions and Answers; N/A, not applicable



- Research context
- State of art
- The AIM\_MIA project
  - Objectives
  - Methods
  - Implementation
  - Connections
  - Expected impact
- People & Budget

A. Retico - Artificial Intelligence in Medicine: focus on Multi-Input Analysis

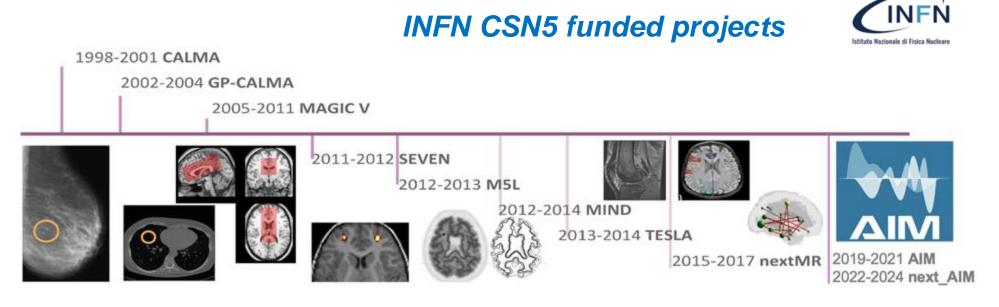


2/13/25

5

٠

## Previous projects @ INFN on this topic



## **INFN CNTT related activities**

- CNTT R4I → INFN Spin-off **DORIAN** (<u>https://www.dorian-tech.com/</u>), A. Chincarini (GE)
- CNTT R4I project **Deeplook** (A deep learning computed diagnosis support for digital breast tomosynthesis), G. Mettivier (NA)



- Research context
- State of art
- The AIM\_MIA project
  - Objectives
  - Methods
  - Implementation
  - Connections
  - Expected impact
- People & Budget

A. Retico - Artificial Intelligence in Medicine: focus on Multi-Input Analysis

2/13/25

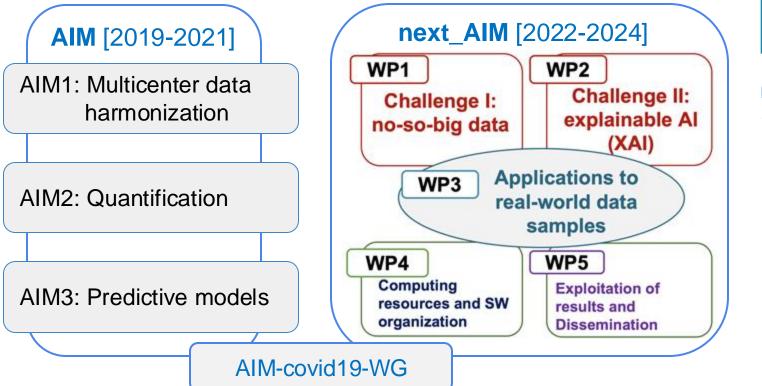
6



## Previous projects @ INFN on this topic

Artificial Intelligence to become the next revolution in medical diagnostics and therapy.

→ New image processing and data analysis strategies (including radiomics ML and DL), need to be developed and extensively validated.



Researchers from INFN divisions and University Departments collaborate closely with Clinicians and Medical Physicists of many Italian hospitals and IRCCS, and with international consortia sharing data



#### Resp. Naz.: A. Retico

13 Research Units: Bari (S. Tangaro) Bologna (D. Remondini) Cagliari (P. Oliva) Catania (M. Marrale) Ferrara (G. Paternò) Firenze (C. Talamonti) Genova (A. Chincarini) Lab. Naz. Sud (G. Russo) Milano (C. Lenardi) Napoli (G. Mettivier) Pavia (A. Lascialfari) Padova (A. Zucchetta) Pisa (M.E. Fantacci)



- Research context
- State of art
- The AIM\_MIA project
  - Objectives
  - Methods
  - Implementation
  - Connections
  - Expected impact
- People & Budget

A. Retico - Artificial Intelligence in Medicine: focus on Multi-Input Analysis

2/13/25

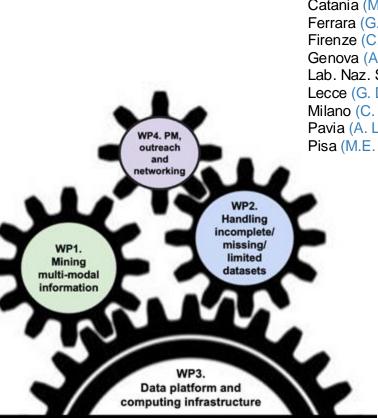


## Artificial Intelligence in Medicine: focus on Multi-Input Analysis (AIM\_MIA)

General goal: to take a step forward in the development and validation of AI-based tools for medical data analysis

## **Objectives**

- Mining multi-modal information
- 2. Handling incomplete/missing/limited datasets
- 3. Development of a dedicated data and computing platform





Resp. Naz.: A. Retico

12 Research Units: Bari (S. Tangaro) Bologna (D. Remondini) Cagliari (P. Oliva) Catania (M. Marrale) Ferrara (G. Di Domenico) Firenze (C. Talamonti) Genova (A. Chincarini) Lab. Naz. Sud (G. Russo) Lecce (G. De Nunzio) Milano (C. Lenardi) Pavia (A. Lascialfari) Pisa (M.E. Fantacci)



- Research context
- State of art
- The AIM\_MIA project
  - Objectives
  - Methods
  - Implementation
  - Connections
  - Expected impact
- People & Budget

A. Retico - Artificial Intelligence in Medicine: focus on Multi-Input Analysis

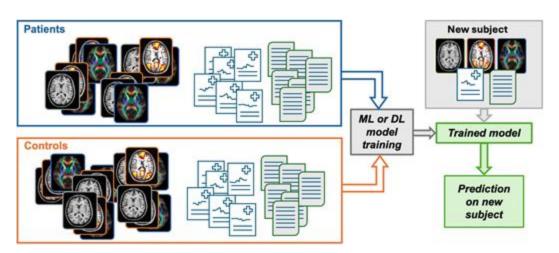
2/13/25

8



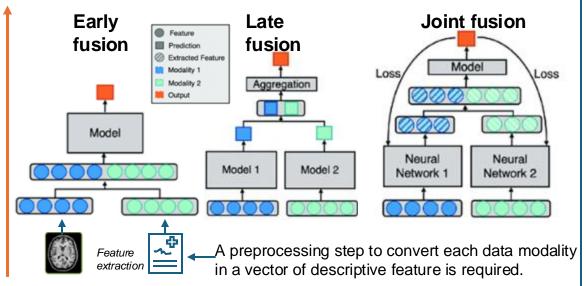
## Methods: multi-input analysis (WP1)

- Data from different sources (image modalities, text and clinical scales, omics data) should be combined
  - → Multimodal Fusion



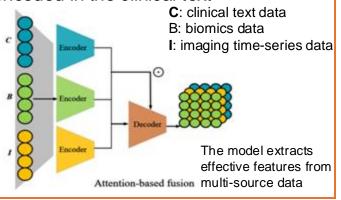
#### Feature-based approach

Model architectures for different fusion strategies:



#### **Advanced DL methods**

Based on the attention mechanism of the Transformers, the attention-based fusion approach enable to weight the image features according to the information encoded in the clinical text





- Research context
- State of art
- The AIM\_MIA project
  - Objectives
  - Methods
  - Implementation
  - Connections
  - Expected impact
- People & Budget

A. Retico - Artificial Intelligence in Medicine: focus on Multi-Input Analysis



## Methods: multi-input analysis (WP1)

🔆 covidcxr - Hackathon

Clinical Features:

PaO2, SaO2, pH,

Sex. Cough.

Age, Temp\_C, WBC,

CRP, LDH, D dimer,

DifficultyInBreathing.

Input 13x1

Dense (8)

LeakyReLU

Dropout

Concatenate

**Dense** (16)

LeakyReLU

**Dense** (16)

LeakyReLU Dense (8)

LeakyReLU

Dense (8)

LeakyReLU

Dense (1)

RespiratoryFailure

Previous results of the collaboration on this topic

#### RX images + clinical data

Input 1024x1024

Conv2D

**Residual Block** 

**Residual Block** 

**Residual Block** 

Residual Block

**Residual Block** 

Residual Block

GlobalAveragePooling

The AIM-WG team achieved the 4<sup>th</sup> place with

a 74% accuracy in predicting patient

prognosis on the hackathon test set

Conv2D

LeakyReLU

BatchNorm

Conv2D

LeakyReLU

BatchNorm

Add

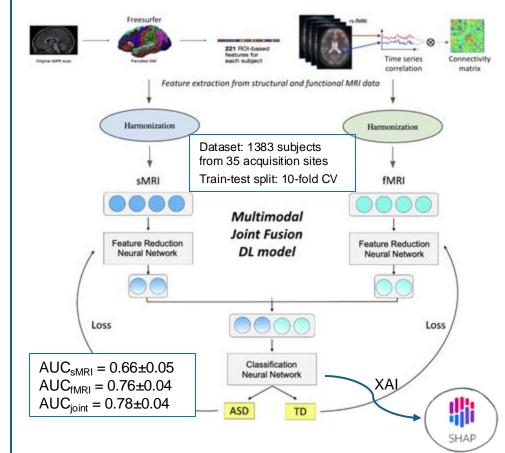
Dataset:

axPooling

~ 1100 CXR for training

~ 500 CXR for test

#### Structural MRI + functional MRI data



Lizzi F, Brero F, Fantacci ME, Lascialfari A, Paternò G, Postuma I, Oliva P, Scapicchio C, Retico A (2024). A multi-input deep learning model to classify COVID-19 pneumonia severity from imaging and clinical data. IWBBIO, 1–12.

Saponaro S, Lizzi F, Serra G, Mainas F, Oliva P, Giuliano A, Calderoni S, Retico A. (2024). Deep learning based joint fusion approach to exploit anatomical and functional brain information in autism spectrum disorders. Brain Informatics, 11(1), 2. https://doi.org/10.1186/s40708-023-00217-4

2/13/25

9

#### Outline

- **Research context**
- State of art
- The AIM MIA project
  - **Objectives**
  - **Methods**
  - Implementation
  - Connections
  - Expected impact
- People & Budget

A. Retico - Artificial Intelligence in Medicine: focus on Multi-Input Analysis

2/13/25

10



## Methods: handling incomplete/unbalanced/ limited datasets (WP2)

#### incomplete

A high proportion of missing data may affect multimodal data collections.

Simply excluding patients with missing data:

- reduces the dataset dimensionality
- may lead to a selection bias

 $\rightarrow$  Data imputation techniques infer missing values from the existing part of the data.

#### unbalanced

It happens very often that one class of subjects is under-represented with respect to the other one(s)

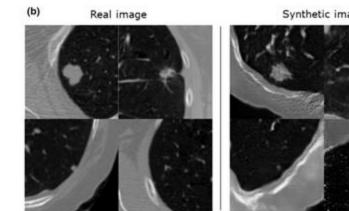
→ Data resampling techniques (over-sampling and under-sampling); **balancing data** in training: stratified kfold CV, adjusting the cost function (e.g. setting different weights for different classes); reformulate the problem as an anomaly detection one

#### limited

Collecting large annotated datasets still represents a major challenge in medical data analysis

 $\rightarrow$  Traditional data augmentation techniques and synthetic data generation

FILE, ID	AGE	na	DX GROUP	L Mean Thickness	T Mean Thickness	L Cortes Vol	R Cortex Vul		FLE.ID	AGE	no	DK GROUP	L Mean Thickness	T Mean Dickness	L Cortes Vel	R Cortes Val
Cellach_0051457	22.0	107	1	2.54	2.67	321132	317005		Callech. 0051467	22.0	107		2.56	2.67	321132	317005
Caltech, 0051458	39.2	83	1	2.45	. Net	206419	205456		Callech,0051458	28.2	193	1	2.65	2.43	296419	256456
Callech_0051458	22.8	106		2.71	2.73	307157	310540		Callech,0051458	22.8	106	1	2.71	2.73	307167	310540
Callach, 0051461	\$7,7	- 99		2.58	2,61	263933	209417		Caltech_0051481	37.7			2.59	2.61	263933	209417
Caltech, 9051464	20.9	101		2.76	2.78	361572	394085		Caltech, 0051464	20.8	101	1	2.76	2.78	381572	394085
Callach, 0051472	17.8	125	.1	Net	2.77	345452	345834		Callech, 0051472	17.5	125		2.83	2.77	345432	345834
Callerth, 0051474	20.0	100	1	Net	2.63	294059	309524		Callech, 0051474	20.9	100	1	2.42	2.65	298058	309524
CMU_A_0050854	24	85		2.68	2,21	287015	287815		CMU, a. 0050654	. 24	- 95		2.68	2.71	287010	287815
CMU, a. 0050659	87	109		2.72	2.72	330374	330325		CMU_x.0050609	27	108		2.72	2.72	990376	990335
CMU, a. 0050660	25	Nati	-1	2.74	2.74	279281	284870	data	CMU_A_0050660	25	115	-1	2.74	2.74	279281	254870
CMU + 0050663	21	101	1	2.63	2.67	292910	296389	uutu	CMU + 0050643	21	101	-1	2.63	2.47	292910	296389
CMU, #, 0050664	21	100	1	2.54	2.54	242763	261800	imputation	CMU_A_0050664	21	109	- 1	2.58	2.58	262753	261800
CMU, a. 0050665	33	108		2.55	2.67	237432	237280		CMU, a. 0050665	30	109	- 1	2.55	2.57	237432	237280
CMIJ_A_0050686	31	107	-1	2.19	2.56	315076	311584		CMU_a_0050666	31	107	-1	2.59	2.54	515076	311384
CMU, a. 0050668	25	Net	-1	2.65	2.65	Net	250974		CMU_A_0050668	25	110	-4	2.65	2.65	250916	250974
CMU, b. 0000043	21	123	1	2.45	2,64	257398	Nels		CMU & 0050643	21	123	1	2.65	2.64	257398	256875
CMU & 0050645	20	124	. 1	2.58	2.59	264307	260633		CMU & 0050645	20	124	1	2.54	2.59	264307	260833
CMU, b. 0050651	- 24	196	+	2.47	2.45	306868	306113		CMU, h. 0050651	39	116	1	2.47	2.45	306868	306173
KK0_0050014	8.48	108	-1	2.79	2.81	351158	355034		KK0_0050014	8.46	108	-1	2.79	2.81	351108	355034
KHQ.0050815	10.8	105		2.46	2.62	Net	219226		KKI_0050815	10.62	105		2.46	2.62	207923	219226
KKQ_0050818	9.73	119	15	2.61	2.69	271211	275917		KKI_0050818	0.73	118		2.41	2.69	271211	275917
800,0050817	9.97	Nati		2.61	2,60	299531	299597		KKI_0050817	9.97	110		2.81	2.60	298531	299592
800,0050818	11.8	- 98	-1	2.67	2.69	277604	205714		KKI_0050818	11.79	. 98	-1	2.67	2.69	277604	206714
KKG_0050815	9.71	101	-4	2.61	2.62	296599	295811		KKI_0050015	9.71	101	-4	2.61	2.62	296599	295511
KKQ_0050821	11.2	114	-1	2.75	2.75	288554	288993		KN0,0050821	11,17	114	-1	2.75	2.75	288554	288993
800,0050822	12.4		- 1	2.74	2.72	287400	289102		KX0_0050822	12.43		-1	2.74	2.72	287403	289102
XXX 0050823	11.4	120		2.86	2.65	314365	310731		KX0.0050823	11.37	120	1	2.86	2.83	314385	310731



[Chlap P et al, A review of medical image data augmentation techniques for deep learning applications. Journal of Medical Imaging and Radiation Oncology, 65(5), 545–563 (2021)]

Synthetic image

#### Outline

- **Research context**
- State of art
- The AIM MIA project
  - **Objectives**
  - **Methods** •
  - Implementation
  - Connections
  - Expected impact
- People & Budget

A. Retico - Artificial Intelligence in Medicine: focus on Multi-Input Analysis



2/13/25

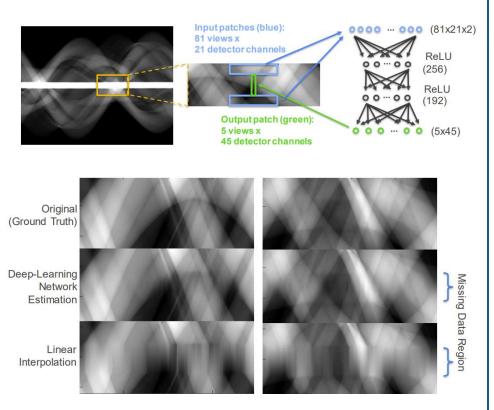
## 11

## Methods: handling incomplete/unbalanced/ limited datasets (WP2)

A selection of the proposed activities...

#### Sinogram inpainting

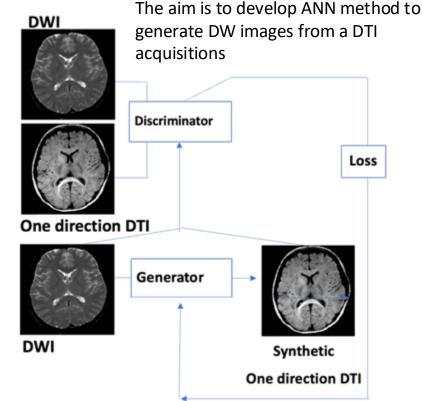
The aim is to develop deep learning methods for inpainting sinograms affected by metallic implant problems.



Metal-Artifact Reduction Using Deep-Learning Based Sinogram Completion: Initial Results, Bernhard E. H. Claus, Yannan Jin, Lars A. Gjesteby, Ge Wang, Bruno De Man, Fully3D-2017

#### Generation of highly-resolved diffusion MR images (DWI) with ANN

DWI with multiple b-values and high number of gradient directions allow a unique insight into human brain microstructure. However, extremely long acquisition time are required.





- Research context
- State of art
- The AIM\_MIA project
  - Objectives
  - Methods
  - Implementation
  - Connections
  - Expected impact
- People & Budget

A. Retico - Artificial Intelligence in Medicine: focus on Multi-Input Analysis

2/13/25

12



## Methods: data platform (WP3)

## The FAIR Guiding Principles for scientific data management

Scientific data should be: Findable, Accessible, Interoperable, and Reusable

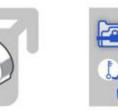
[Wilkinson, M., Dumontier, M., Aalbersberg, I. *et al.* The FAIR Guiding Principles for scientific data management and stewardship. *Sci Data* **3**, 160018 (2016). https://doi.org/10.1038/sdata.2016.18]

 $\rightarrow$  Development of a dedicated data platform for medical data management

It will be based on an open-source imaging IT platform developed by Washington University for neuroimaging data analysis: XNAT Thanks to its **extensibility**, XNAT can be used to support a wide range of imaging-based projects.



Extensible Neuroimaging Archive Toolkit https://www.xnat.org/



Full DICOM Integration and Anonymization: Get image data in, and keep PHI out.

▲ ど

Secure Access & Permission Control: You decide who does what with your data.



Integrated Search & Reporting: Report on your image and clinical data together.



Pipeline Processing: Use the power of highperformance computing on your data.

Modular Extensibility:

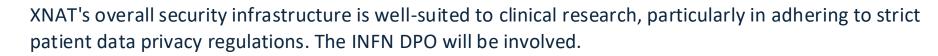
Expand the capabilities

of your XNAT to meet

your needs.



Developer Community: Benefit from an active and engaged set of XNAT power users.





- **Research context**
- State of art
- The AIM MIA project
  - **Objectives** •
  - **Methods** •
  - Implementation
  - Connections 0
  - Expected impact  $\bullet$
- People & Budget

A. Retico - Artificial Intelligence in Medicine: focus on Multi-Input Analysis

2/13/25

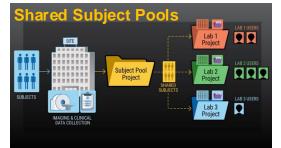
13



## Methods: data platform (WP3) **Preliminary work @ INFN-PI**

First implementation of a platform prototype to collect, store and analyze CT images and data of phantoms

Planned collaboration antm with AIFM



		$\sim$							
Subject	Phantom_name	Journoor	CTDIvol	ReconstructionAlgorithm	Iterativ	vePercentage	PixelSpacing	SliceThickness	ConvolutionKernel
CAT_fc18_2_aidrSTD	Catphan	TOSHIBA	2	AIDR3D	40		0.428	1	FC18
CAT_lung_2_asir10	Catphan	GE	2	ASIR	1		0.406	1.25	LUNG
RAD_lung_4_FBP	Radiomik	TOSHIBA	4	FBP	0		0.781	1	LUNG
RAD_lung_7_asir10	Radiomik	GE	7	ASIR	10		0.703	1.25	LUNG
						<u> </u>			

Accession # XNAT\_S00003 Date Added 2024-03-25 16:30:49 (camilla

Subject Details: RAD\_lung\_4\_FBP

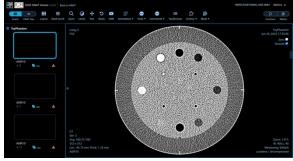
Projects

Parameters used for the acquisition

Exper

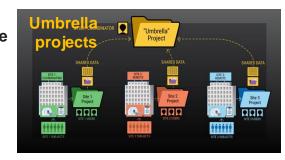
TOSHIBA

Searchable fields to make specific queries and retrieve subsets of data

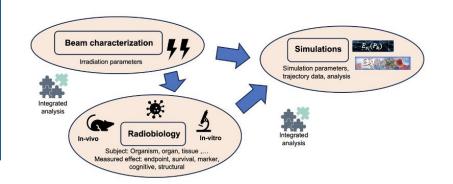


CT images of phantoms in the DICOM standard format

Design of a platform prototype to collect, store and analyze the heterogeneous data acquired in the PNRR ECS THE project







### Outline

- Research context
- State of art
- The AIM\_MIA project
  - Objectives
  - Methods
  - Implementation
  - Connections
  - Expected impact
- People & Budget

A. Retico - Artificial Intelligence in Medicine: focus on Multi-Input Analysis

2/13/25

14



## Methods: datasets to be shared

Only public data will be stored in the XNAT platform

Dataset ID	Target pathology/ condition/task	Approximate sample size	Data types	link
ABIDE	Autism Spectrum Disorders (ASD)	> 2000 subjects including subjects with ASD and controls	demographic, clinical, sMRI, fMRI	https://fcon_1000.projects.nitrc.o rg/indi/abide/
ADNI	Alzheimers' disease	>3000 participants	clinical, biochemical.sMRI, fMRI, DWI, PET	https://adni.loni.usc.edu/
AIBL	Alzheimers' disease	>3000 participant	clinical, biochemical, sMRI, fMRI, DWI, PET	https://aibl.org.au/
GBM	De novo glioblastoma	630 subjects with many missing data	demographic, clinical, genetic, sMRI, DWI	https://www.cancerimagingarchiv e.net/collection/upenn-gbm/
Lung-PET- CT-Dx	Lung cancer	355 participants, 436 studies, 1295 series	РЕТ СТ	https://www.cancerimagingarchiv e.net/collection/lung-pet-ct-dx/
MMIST ccRCC	Kidney cancer	618 Patients	clinical CT, MRI, WSI, Genomics	https://multi-modal-ist.github.io/ datasets/ccRCC/
CPTAC-PDA	Pancreatic Ductal Adenocarcinoma	168 patients	US CT, MR, PET, histopathological images, clinical info	https://www.cancerimagingarchiv e.net/collection/cptac-pda/
CT-MAR	Metal artifact reduction in CT reconstruction	14000 cases (1773 head + 12227 body) for training 1000 cases for testing	set of CT image pairs and sinogram pairs generated from NIH DeepLesion dataset	https://www.aapm.org/GrandCha llenge/CT-MAR/
BLUES	Covid-19 pneumonia: COVID Bluepoint Lung Ultrasound	63 patients (33 COVID-positive and 30 COVID-negative) 362 videos corresponding to 31,746 frames	patient characteristics, lung ultrasound (US) videos, symptoms, comorbidities, blood test data, vital parameters, and the PCR test result testing for COVID-19	<u>https://github.com/NinaWie/COVI</u> <u>D-BLUES</u>
SARCOMA	Soft-tissue sarcoma	51 patients	RTSTRUCT, CT, MR (T1-weighted, T2-weighted with fat-suppression) , PT	https://www.cancerimagingarchiv e.net/collection/soft-tissue-sarco ma/



Gantt

#### Outline

- Research context
- State of art
- The AIM\_MIA project
  - Objectives
  - Methods
  - Implementation
  - Connections
  - Expected impact
- People & Budget

A. Retico - Artificial Intelligence in Medicine: focus on Multi-Input Analysis



#### 2/13/25 15

## Implementation

Mining multi-modal information WP1 T1.1 Feature-based approach to multi-input analysis T1.2 Integration of multi-parametric and multi-modal imaging data T1.3 AI solutions for heterogeneous data analysis Handling incomplete/missing/limited datasets WP2 T2.1 Traditional approaches for data curation and augmentation T2.2 Medical Image Data Generation T2.3 Data inpainting with CNN Data platform and computing infrastructure WP3 T3.1 Definition of requirements and user roles T3.2 Realization and maintenance of the data platform prototype T3.3 Integration of data processing pipelines and output storing T.3.4 SW organization and repository T3.5 Data collection Project management, outreach and collaborations with external parties WP4 T4.1 Project management and networking T4.2 Collaboration with AIFM and medical associations T4.3 Outreach

				Sec	di co	oinv	olte	•	timeline								
BA	во	CA	ст	FE	FI	GE	LE	LNS	мі	PI	PV	20	25	20	26	20	27
						ο							M1.1		M1.2		M1.3
x	x	ο			x	x	x	x			x						
	x				o	x	x	x			x						
			0		x	x	x	x	x	x	x						
		ο											M2.1		M2.2		M2.3
		x					ο		x								
	ο			x	x						x						
		x	x	ο													
										ο			M3.1		M3.2		M3.3
x										ο	x						
x	x	x				x				o	x						
		х				ο				x							
	x					x				x	ο						
	0	х								x							
										0		M4.1		M4.2			M4.3
x	x	x	x	x	x	x	x	x	x	0	x						
х	x	х	х	х	x	х	х	x	х	0	x						
x x Leg	x x end: WP	x x	x x	x x	o x	o x	x x	x x	x x	x o	x x						

- o Task coordination
- x Task participation

#### Outline

- Research context
- State of art
- The AIM\_MIA project
  - Objectives
  - Methods
  - Implementation
  - Connections
  - Expected impact
- People & Budget

A. Retico - Artificial Intelligence in Medicine: focus on Multi-Input Analysis

2/13/25

16



## Implementation: Milestones

	Dea	adline	Milestone	Description											
31/12/	/2025	M1.1	Developm	ment of a DL-based analysis pipeline for feature extracted from multiple sources and release to the SW rep											
31/12/	/2025	M2.1	Implement	tation and release to the SW repository of traditional data augmentation and data imputation techniques.											
31/03/	/2025	M3.1a	Initializatic	on of the shared SW repository and release of guidelines for contributors											
31/12/	/2025	M3.1b	Release of allowed da	an internal note reporting the guidelines to set up the data platform (definition of the requirem ata types)	ients, user roles,										
31/03/	/2025	M4.1	Kick-off me	eeting organization and release of a plan of collaboration with external parties											
ſ	31/	/12/26	M1.2	Definition of a generalizable AI-based pipeline that integrated multiparametric or multimodal images											
	31/12/26 M2.2			Inpainting of metal artifact corrupted sinograms with NN											
	31/	/12/26	M3.2	Instantiation of a XNAT platform prototype and integration of data sets											
r	30/	/06/26	M4.2	Identification of joint research or dissemination activities to promote the networking among researchers from different Institutions and associations											
ſ	31/12/27 M1.3		M1.3	Definition of an AI-based pipeline that integrates images and features extracted from heterogeneous data											
	31/	/12/27	M2.3	GAN generation of medical images											
	31/	/12/27	M3.3	Integration of data processing pipelines and output storing											
	31/12/27 M4.3			Organization of a workshop to discuss the results of the project with stakeholders											

## Outline

- Research context
- State of art
- The AIM\_MIA project
  - Objectives
  - Methods
  - Implementation
  - Connections
  - Expected impact
- People & Budget

A. Retico - Artificial Intelligence in Medicine: focus on Multi-Input Analysis

2/13/25

17



## Connections with external entities and other projects

- Collaboration with Italian Association of Medical Physicists (AIFM)
  - A **5-year research framework agreement** has been renewed in January 2024 between INFN and the Italian Association of Medical Physicists (AIFM). This agreement foresees the collaboration between the parties for the synergistic realization of common research objectives in the healthcare field.
  - Roles of INFN and associated researchers in AIFM boards and initiatives:
    - A. Retico, M.E. Fantacci, M. Marrale and C. Lenardi are members of the Research Committee of AIFM
    - C. Talamonti coordinates the AI working group of AIFM
    - A. Lascialfari is the chair of the Research Committee of AIFM
    - C. Lenardi is in the Executive Board of AIFM
  - Example of joint activities
    - Organization of joint webinars INFN-AIFM on research topics related to the use of AI in the medical field:
      - e.g. in 2024 <u>https://fisicamedica.it/formazione/agenda-eventi-formativi/come-affrontare-insieme-le-sfide-dellia/</u>
    - Participation of AIFM medical Physicists to INFN projects
    - Joint organization of workshop, training opportunities and dissemination events (e.g. Bright-Nights)
      - <u>https://fisicamedica.it/formazione/agenda-eventi-formativi/workshop-aifm-infn-oltre-la-diagnosi-e-la-terapia-la-ricerca-che-rivoluziona-la-clinica/</u>

### Collaboration with other associations

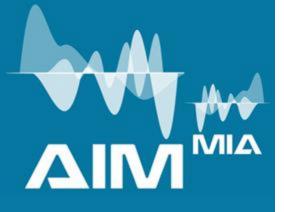
- A. Chincarini is a member of the
  - Executive Board of European Alzheimer's Disease Consortium (EADC, <u>https://eadc.online/</u>)
  - Neurological Study Committee of the Italian Association of Nuclear Medicine (AIMN, <a href="https://aimn.it/">https://aimn.it/</a>)
- C. Testa is a member of the Executive Board of the Italian Chapter of the International Society of Magnetic Resonance in Medicine (AIRMM, <u>https://www.ismrm.it/it/airmm/</u>)

• ..





#### WORKSHOP AIFM-INFN Oltre la Diagnosi e la Terapia: La Ricerca che Rivoluziona la Clinica



- Research context
- State of art
- The AIM\_MIA project
  - Objectives
  - Methods
  - Implementation
  - Connections
  - Expected impact
- People & Budget

A. Retico - Artificial Intelligence in Medicine: focus on Multi-Input Analysis



2/13/25 18

# Expected impact

- The AIM\_MIA project is expected to deliver:
  - robust and effective analysis pipelines to make predictions about the health status of individuals, by extracting and combining via multi-input AI-based tools the complementary and heterogeneous information provided by different data sources (e.g. images, diagnostic tests, phenotypic and genetic data);
  - many scientific publications on peer-reviewed journals on this topic;
  - a **data platform prototype** able to collect and share samples according to FAIR principles and GDPR compliant.
- The acquired knowledge in the field will be shared within the scientific community, after evaluating its potential for transferability and exploitability together with INFN CNTT
- The **data platform prototype** will serve as a fundamental reusable/repurposing resource for several additional research studies in the field
- The network of INFN researchers, technologists and associated personnel interested in AI developments for medical applications will be consolidated and further expanded

#### Outline

- **Research context**
- State of art
- The AIM\_MIA project
  - Objectives  $\bullet$
  - Methods •
  - Implementation 0
  - Connections
  - Expected impact
- **People & Budget**

A. Retico - Artificial Intelligence in Medicine: focus on Multi-Input Analysis



# People & Budget

#### 2025 > Globale > Gruppo V > Esperimento AIM\_MIA > Riassuntivo assegnazi

**Budget 2025** 

#### **People 2025**

2/13/25

19

85 people, 27.45 FTE (~0.3 FTE per person)

SEZIONE	DICED	CATORI	TECN	
SEZIONE				
BA (10 PERSONE - 2.75 FTE)	2.4 fte	9 pers.	0.35 fte	1 pers.
	2.75	i fte / 10 p	oers. (me	dia 0.28)
	5.8 fte	10 pers.	0.15 fte	2 pers.
BO (12 PERSONE - 5.95 FTE)	5.95	i fte / 12 p	bers. (me	dia 0.50)
	1.6 fte	3 pers.	0 fte	pers.
CA (3 PERSONE - 1.6 FTE)	1.6	60 fte / 3 p	bers. (me	dia 0.53)
	2.15 fte	4 pers.	0 fte	pers.
CT (4 PERSONE - 2.15 FTE)	2.1	5 fte / 4 p	bers. (me	dia 0.54)
	1.9 fte	7 pers.	0.2 fte	1 pers.
FE (8 PERSONE - 2.1 FTE)	2.1	0 fte / 8 p	bers. (me	dia 0.26)
	1.8 fte	5 pers.	0 fte	pers.
FI (5 PERSONE - 1.8 FTE)	1.8	0 fte / 5 p	pers. (me	dia 0.36)
	3.1 fte	4 pers.	0 fte	pers.
GE (4 PERSONE - 3.1 FTE)	3.1	0 fte / 4 p	pers. (me	dia 0.78)
	1 fte	2 pers.	0 fte	pers.
LE (2 PERSONE - 1 FTE)	1.0	0 fte / 2 p	bers. (me	dia 0.50)
	0.8 fte	2 pers.	0.2 fte	1 pers.
LNS (3 PERSONE - 1 FTE)	1.0	0 fte / 3 p	oers. (me	dia 0.33)
	1.4 fte	11 pers.	0 fte	pers.
MI (11 PERSONE - 1.4 FTE)	1.40	) fte / 11 p	oers. (me	dia 0.13)
	2.2 fte	7 pers.	0.95 fte	7 pers.
		N. 144.	are (mo	dia 0 22)
PI (14 PERSONE - 3.15 FTE)	3.15	5 fte / 14 p	Jers. (me	ula 0.23)
. , ,	3.15 0.95 fte	6 pers.		
PI (14 PERSONE - 3.15 FTE) PV (9 PERSONE - 1.45 FTE)	0.95 fte		0.5 fte	3 pers.

Sez. & Suf.		MISS				со					CON			SEM			TF				PL				MAN			INV			:-sw		AP			ERVIZ		тот		
		SjD	ot.	Ant.			Dot.	Ant.	5	Sj D	oot.	Ant.	5	j Dot	. Ant		SjD	ot. /	Ant.		Sj	Dot.	Ant.	2	Sj Dot.	Ant.	S	j Dot. /	Ant.	Sj	Dot.	Ant.	Sj Dot	t. Ant.	Sj	Dot. Ani	1		Dot	Ant
BA	3.0				1.0	1																															4			
	3.0				1.0					_	_		_	+	+	+	_	_						_				+ +					_	-			4.0	-	-	<u> </u>
во	3.0				1.0	1																															4	1		
	3.0				1.0					_	_		_	_	_	$\downarrow$		_																		_	4.0	-	_	<u> </u>
CA	5.0				2.0	1																								0.5							7.5	1		
	5.0				2.0					_	_		_	_	_															0.5							7.5	-	_	_
ст	3.0				2.0	1																															5	1		
	3.0				2.0					_			_	_	_				_						_												5.0	-	_	<u> </u>
FE	2.0				2.0																																4			
	2.0				1.0																																3.0	-		
FI	5.0				2.0																																7			
	3.5				1.0																																4.5	-		
GE	2.0	2			1.5	1																		2.0						0.5							6		1	
	2.0	2.0			1.5																			2.0						0.5							_	2.0	)	
LE	2.0				0.5																																2.5			
	2.0				0.5																																2.5			
LNS	2.0				1.0																																3			
	2.0				0.0	1.0																															2.0	1.0	)	
мі	2.5				2.5																																5			
	2.5				1.0																																3.5			
PI	7.0				2.0																10									0.5							9.5		1	
	6.0				2.0																7.0									0.5							8.5	7.0	)	
PV	5.0				2.0																						2.0										9			
	3.0				1.0																						2.0										6.0			
	41.5	2			19.5															0	10			2			2			1.5							66.5	12	2	
			4	3.5				19.5				0			(				0				10			2			2			1.5		0		(	)			78.5
TOTALE	37	2	0	0	14	1	0	0												0	7	0	0	2	0	0	2	0	0	1.5	0	0					56.5	10.0	0.0	0.0
			3	9.0				15.0				0.0			0.0				0.0				7.0			2.0		:	2.0			1.5		0.0		0.0	)			66.5

Computing resources on the national INFN infrastructure ReCaS

AIM_MIA - GPU: 9000 GPU-hours suddivise su 2 x GPU A100 (>	
40GB di VRAM) per training modelli Al	

6.50

Collaboration with the AI\_INFN project, which is developing easy-to-access GPU resources for DL applications