Italian network of excellence for advanced diagnosis

INNOVA

Advanced Diagnostic Platform

WP2: advanced imaging generation

INNOVAre la diagnostica per immagini per il futuro che intravediamo: rendere disponibile il supporto dell'intelligenza artificiale nel rispetto della persona

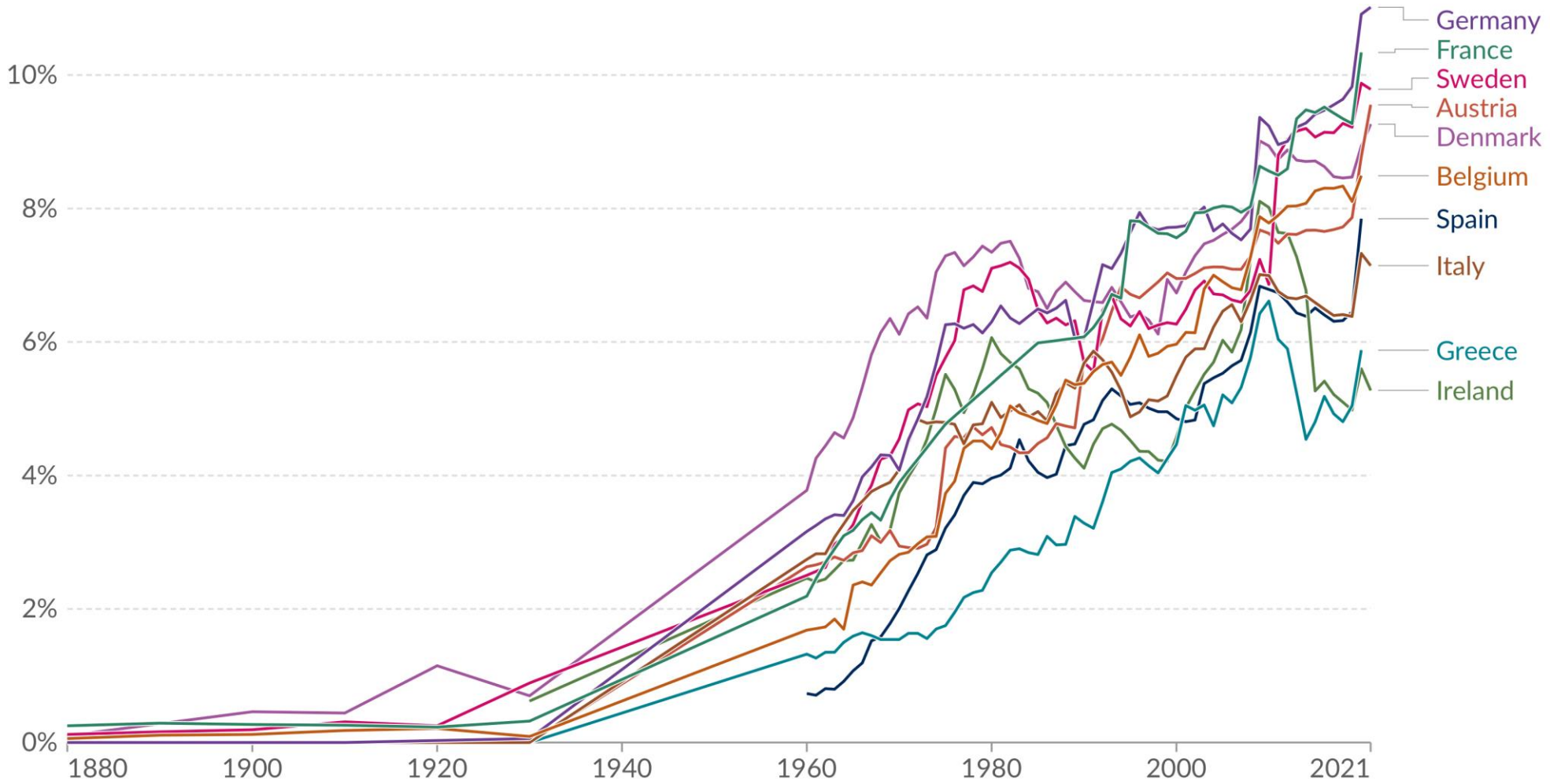
Gianmario Sambuceti e Sara Mora
IRCCS Ospedale Policlinico San Martino, Genova

Ministero della Salute

Funded by the European Union - Next Generation EU - and Ministero della Salute PNRR PNC-E3-2022-23683266 PNC-HLS-DA, INNOVA"
Finanziato da European Union - Next Generation EU - e Ministero della Salute PNRR PNC-E3-2022-23683266 PNC-HLS-DA, INNOVA"

Government health expenditure as a share of GDP, 1880 to 2021

This metric captures spending on government funded health care systems and social health insurance, as well as compulsory health insurance.



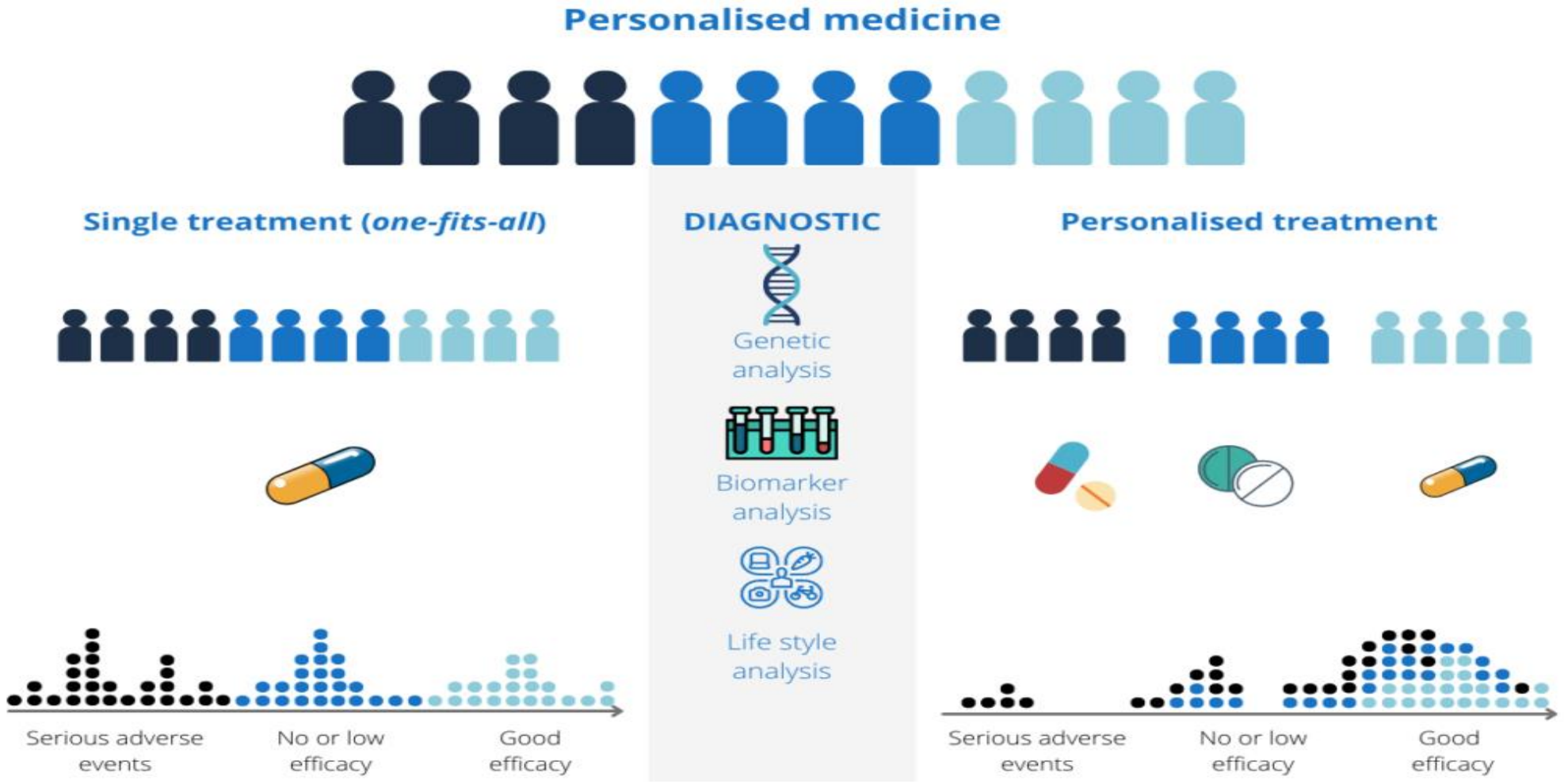
Data source: Our World In Data based on Lindert (1994), OECD (1993), OECD Stat OurWorldInData.org/financing-healthcare | CC BY
Note: Health spending includes final consumption of health care goods and services (i.e. current health expenditure). This excludes spending on capital investments.

Most research on healthcare sustainability focuses on the so-called “precision medicine”:

A series of activities capable of identifying the most appropriate therapeutic or preventive treatment for each individual patient.

The goals are:

To improve treatment effectiveness **To avoid useless spending for ineffective drugs**

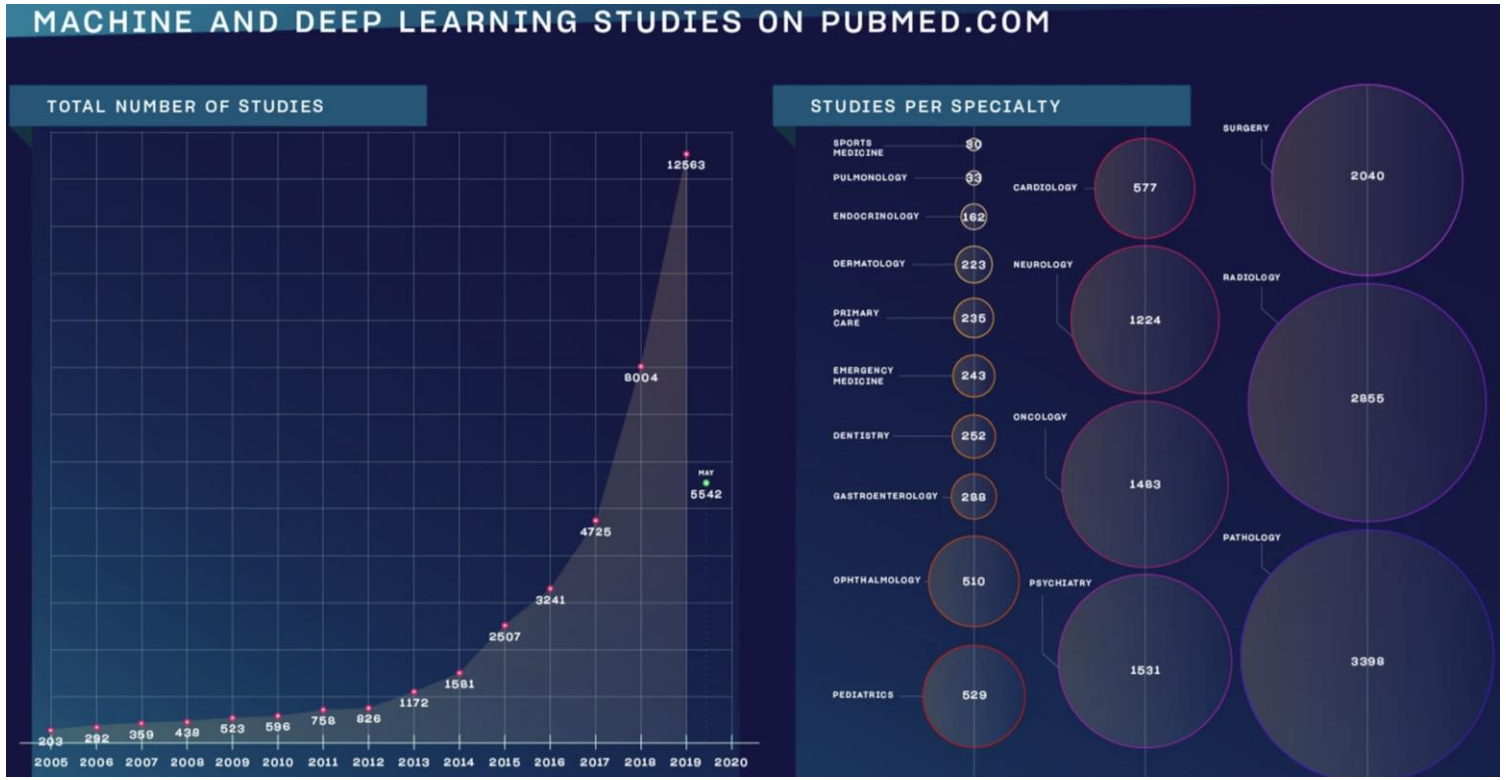


‘Artificial Intelligence’ was coined in 1956 at Dartmouth College, to indicate an emerging research field bringing together researchers on brain physiology

The ability of algorithms and approaches to carry out tasks by displaying intelligent, human-like behavior resembles the industrial revolution when machines provided ways to substitute, supplement, and amplify the manual work performed by humans

The high interest and activity developed by research in medicine and healthcare, is already exploitable in the clinical arena.

Makridakis S. Futures. 2017. https://doi.org/10.1016/j.futures.2017.03.006



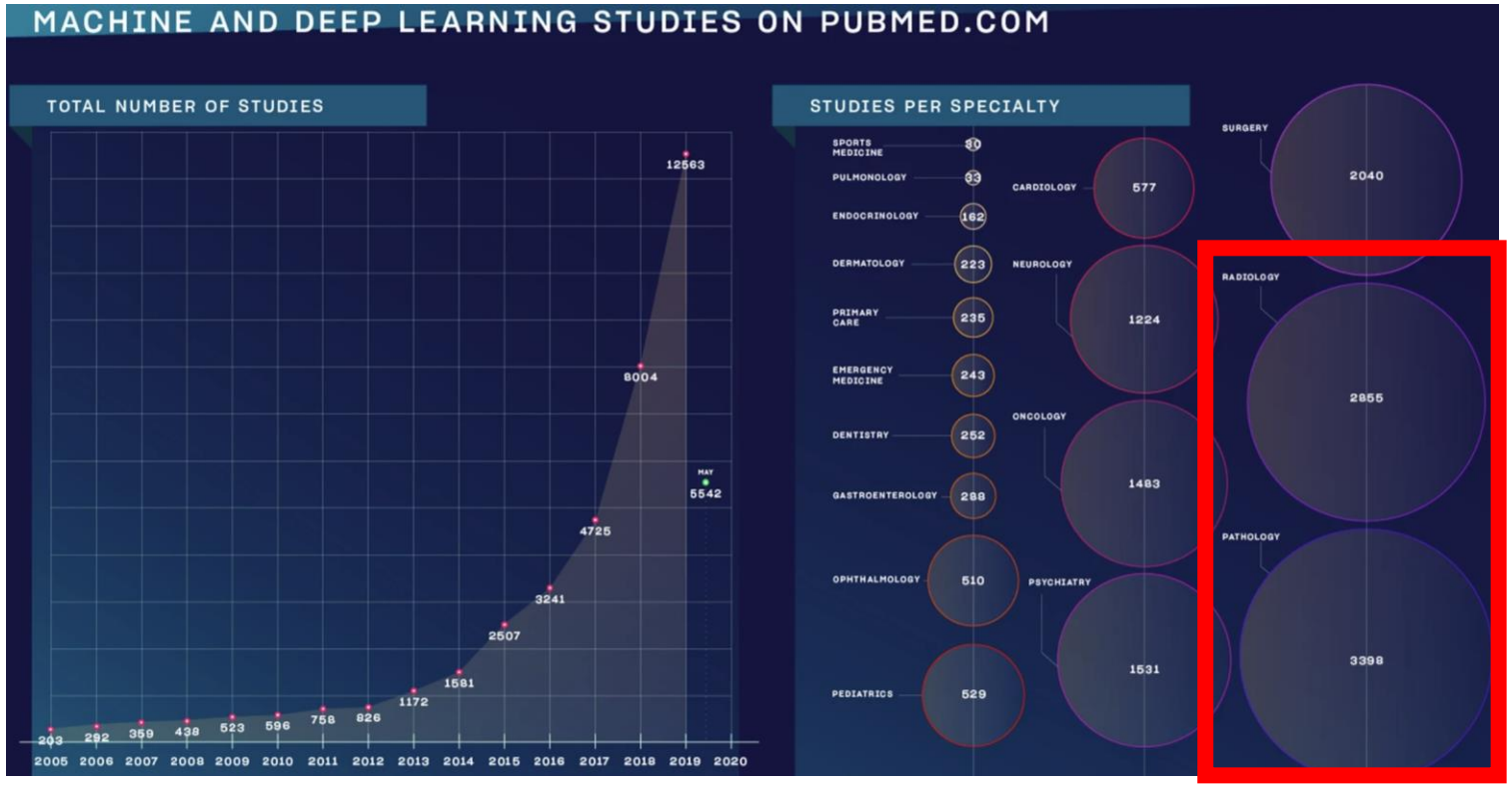
Meskó, B., Görög, M. A short guide for medical professionals in the era of artificial intelligence. npj Digit. Med. 3, 126 (2020). https://doi.org/10.1038/s41746-020-00333-

‘Artificial Intelligence’ was coined in 1956 at Dartmouth College, to indicate an emerging research field bringing together researchers on brain physiology

The ability of algorithms and approaches to carry out tasks by displaying intelligent, human-like behavior resembles the industrial revolution when machines provided ways to substitute, supplement, and amplify the manual work performed by humans

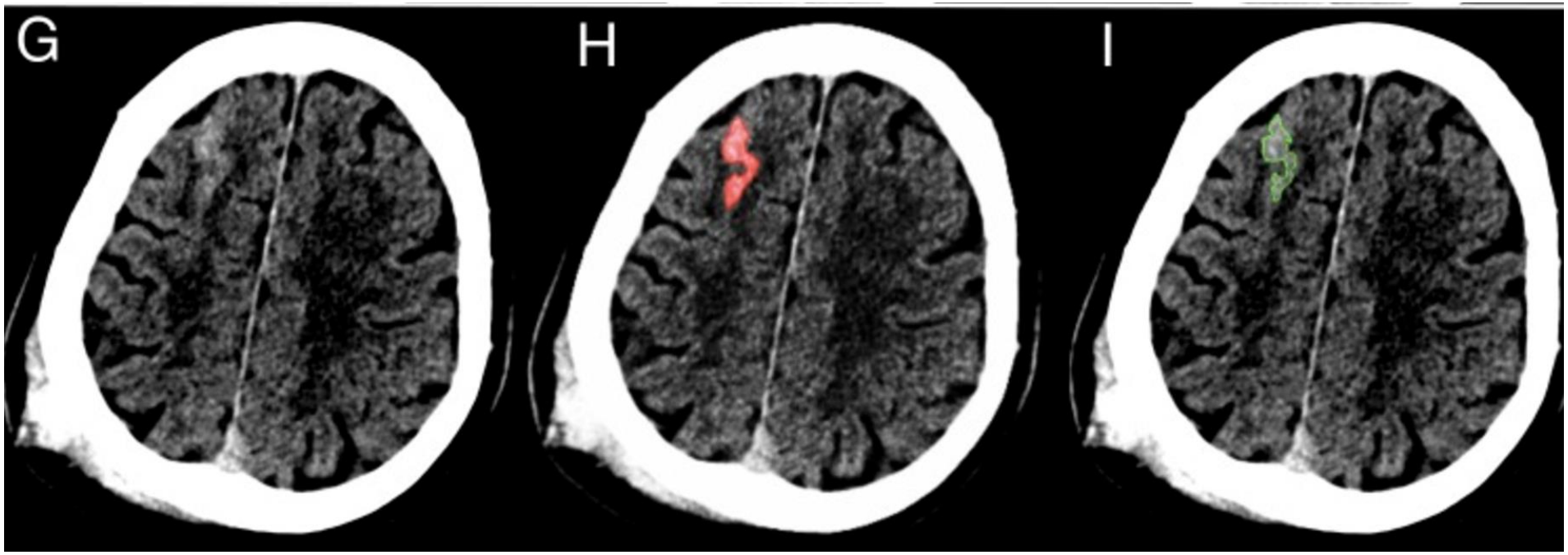
The high interest and activity developed by research in medicine and healthcare, is already exploitable in the clinical arena.

Makridakis S. Futures. 2017. https://doi.org/10.1016/j.futures.2017.03.006



Meskó, B., Görög, M. A short guide for medical professionals in the era of artificial intelligence. npj Digit. Med. 3, 126 (2020). https://doi.org/10.1038/s41746-020-00333-

When Time Matters



Considering AI in the emergency dept can point-out CT scans of trauma patients with >50% chance of post-traumatic subarachnoid hemorrhage.

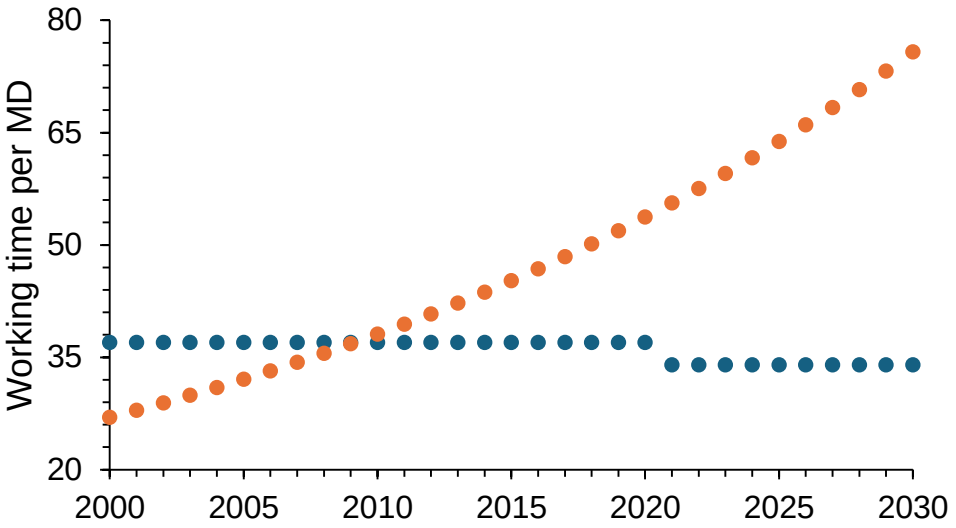
Prioritizing reporting by the (neuro)radiologist

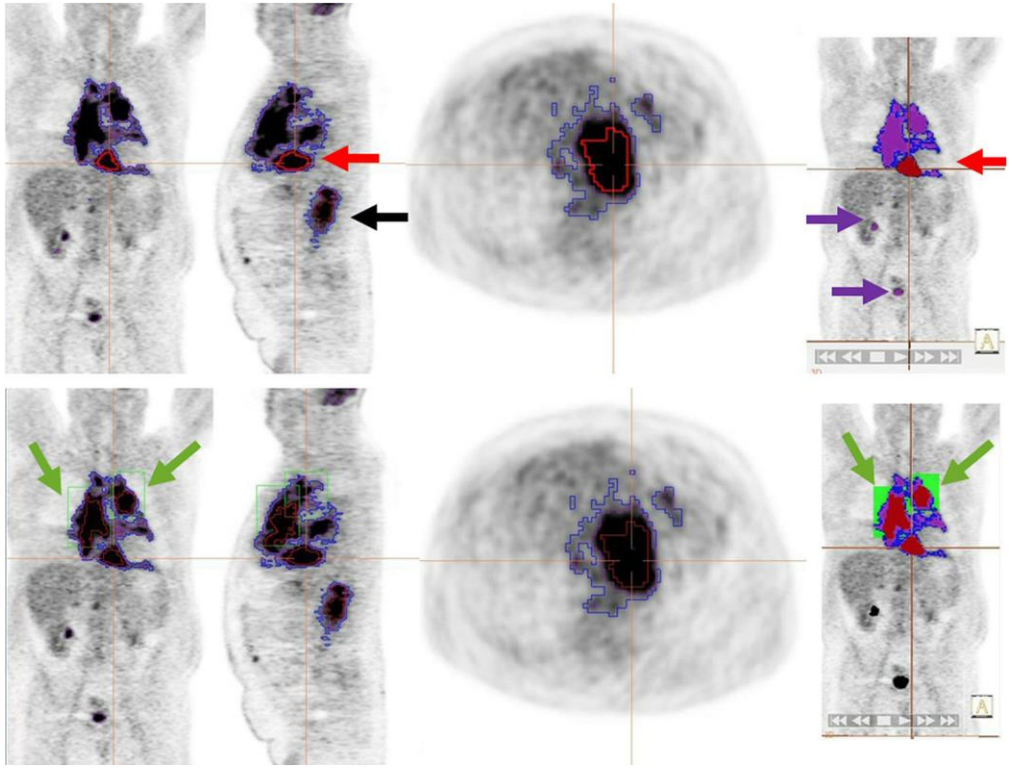
Optimizing timing and effectiveness of treatment strategies

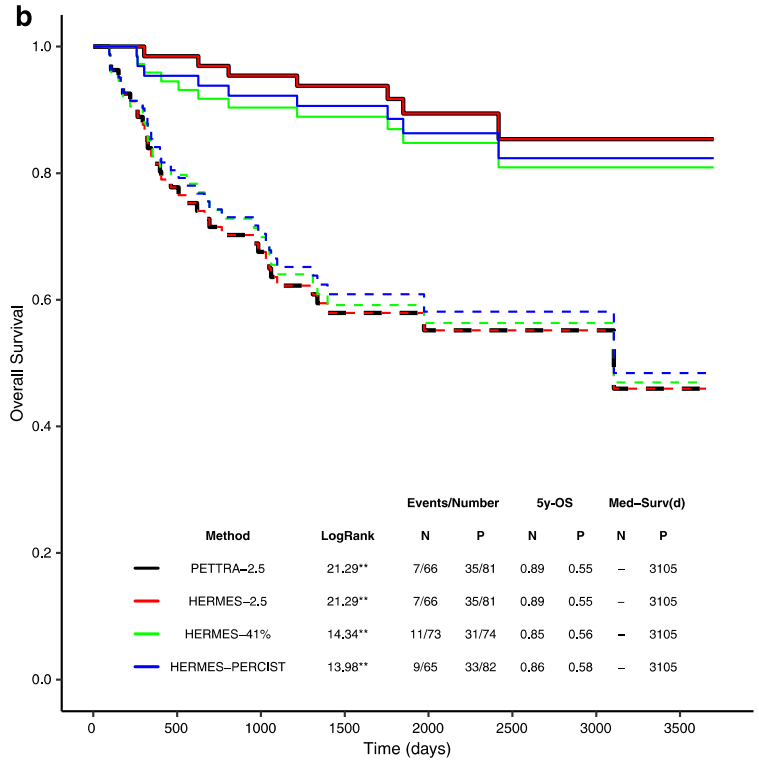
Kuo W, et al. Proc Natl Acad Sci U S A. 2019. doi: 10.1073/pnas.1908021116.

The workload requested and the available resources:

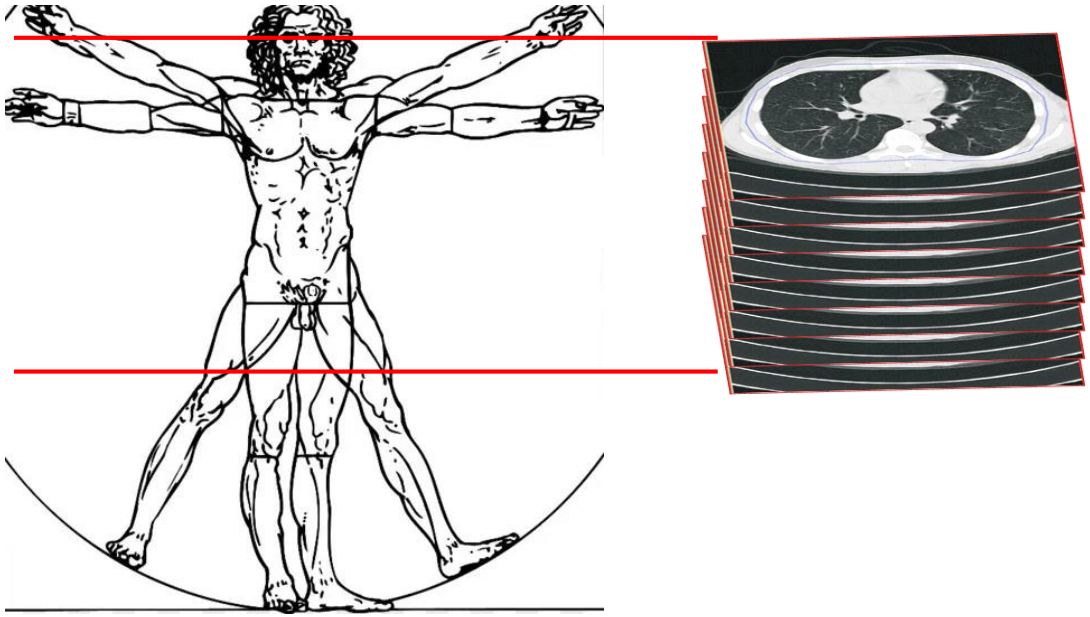
Is AI a possible solution?







Ilyas H, et al. EJNMMI 2018;45:1142



So far: in Italy nr of
Radiology scans 70 000 000 (la Repubblica)
PET/Scans: 7 000 000 (OEC)

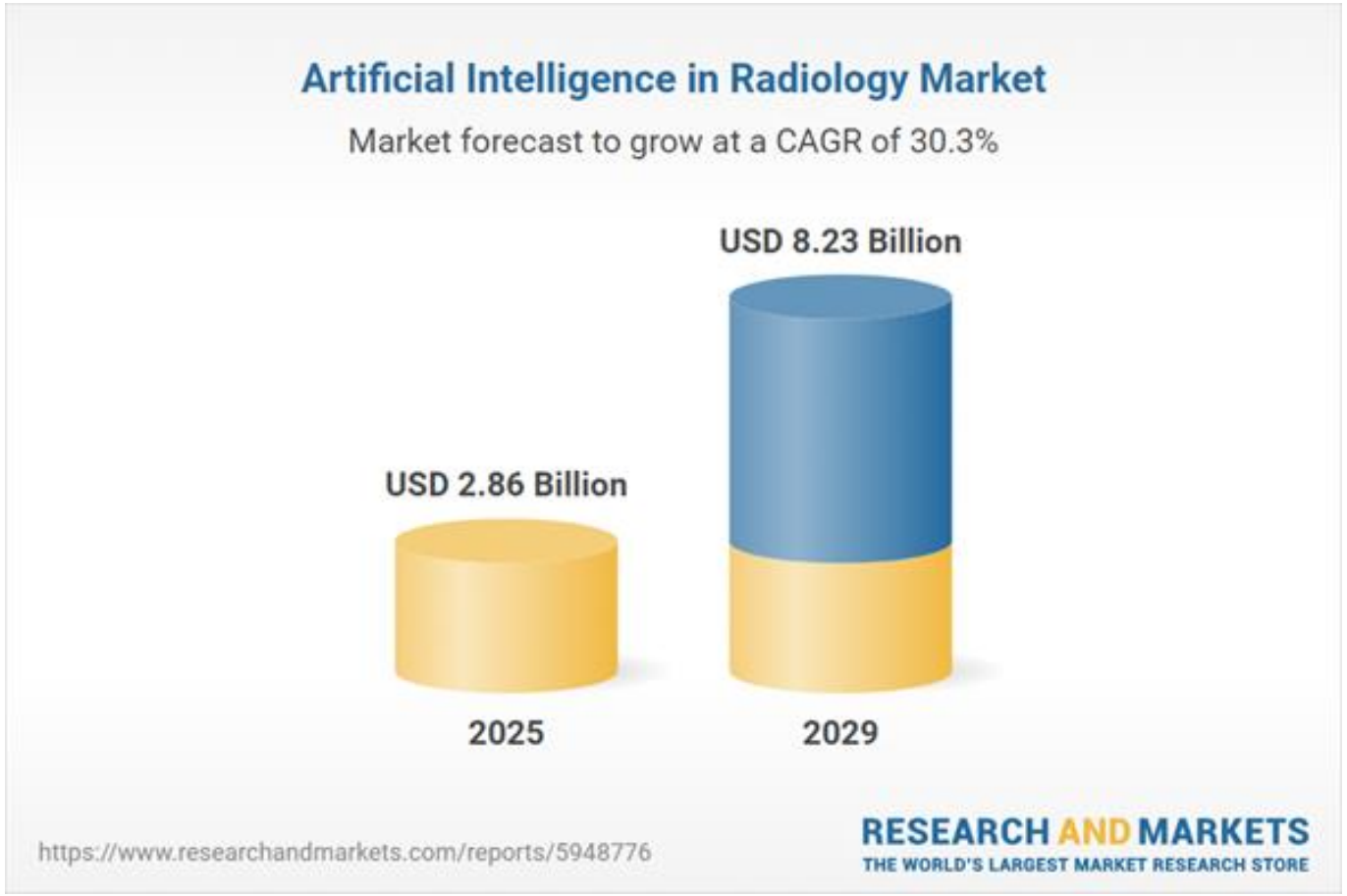
Each CT or PET/CT
500 images covering
neck, chest, abdomen and pelvis

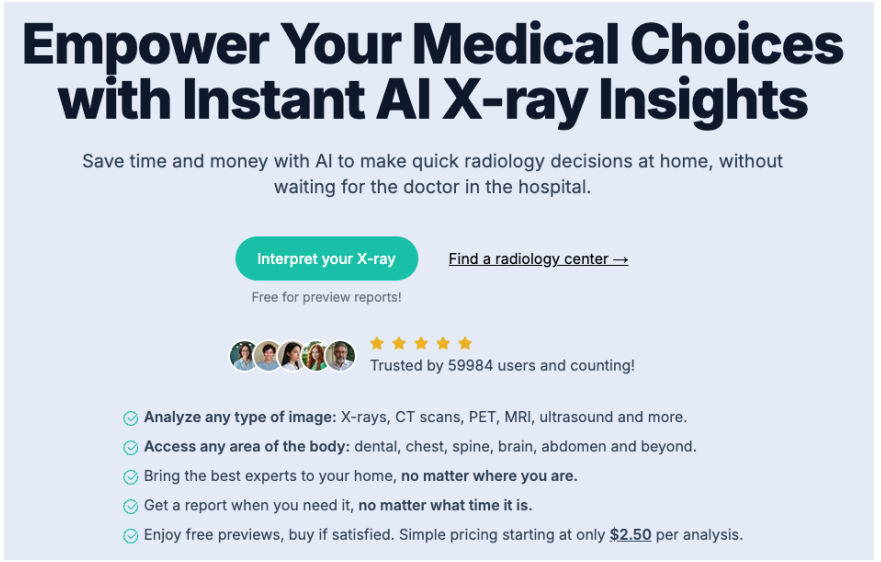
Time allowed for interpretation and reporting
45-60 minutes
(***Karantanis et al, Nuklearmedizin. 2014 doi: 10.3413/Nukmed-0613-13-07)***

Nr of reports in charge of each physician: 12
(**6000 images**)

Nr of paintings present at the Louvre: **7500**

Pervasiveness of Artificial Intelligence in Diagnostic Imaging





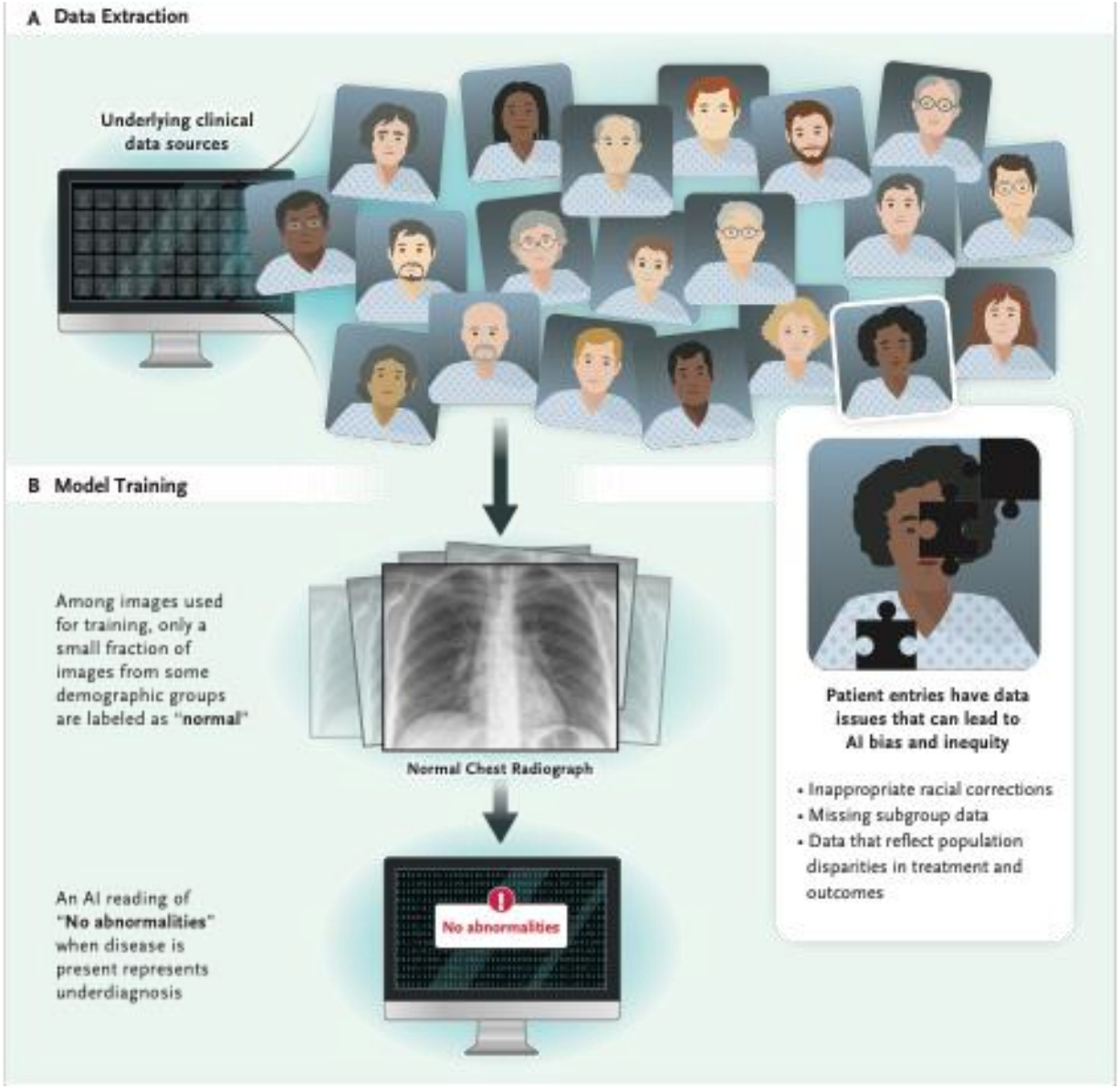
Are there ethical issues?

Every «skill» is dependent upon education and previous experience.

For AI, this experience not only reflects the amount of data used for learning and training.
It also asks for a full representation of the complex and heterogeneous reality the analysis is developed for.

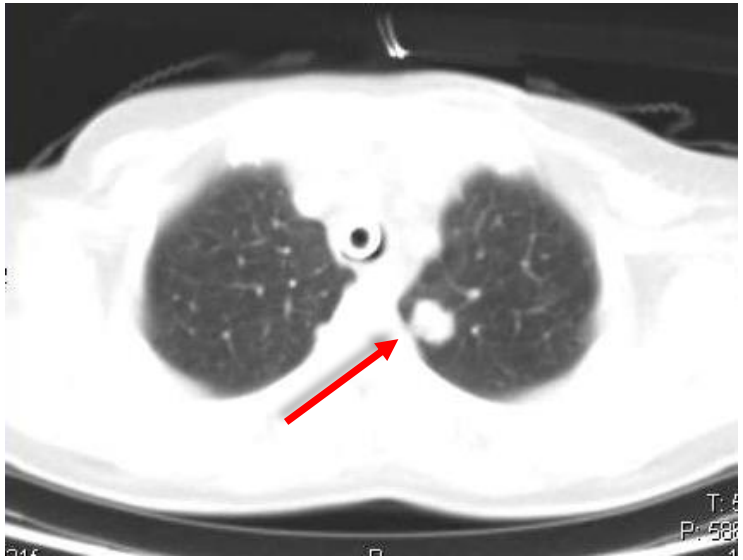
Io Galileo … sono stato dichiarato dal Sant'Uffizio come veementemente sospettato di eresia, cioè d'aver tenuto e creduto che il Sole sia centro del mondo e imobile e che la Terra non sia centro e che si muova; […] con cuor sincero e fede non finta abiuro, maledico e detesto li sudetti errori e eresie, e generalmente ogni e qualunque altro errore , eresia e setta contraria alla Santa Chiesa

When statistics matters: the impact of Bayes theorem on AI performance (1)

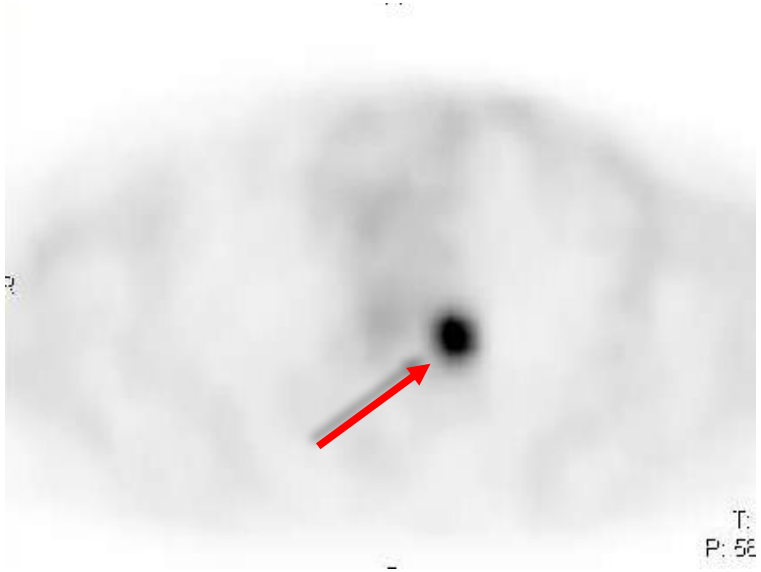


When statistics matters: the impact of Bayes theorem on AI performance (2)

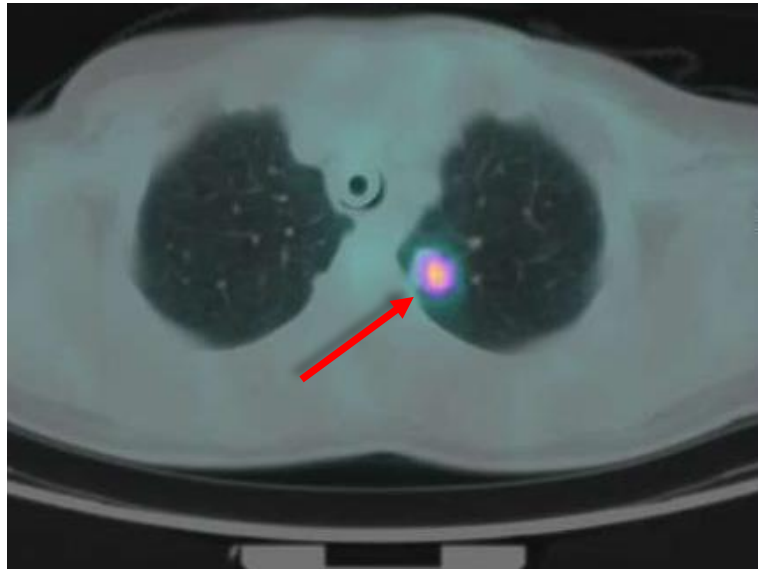
Diagnosis of an unexpected pulmonary nodule documented by a CT performed after trauma



X ray CT

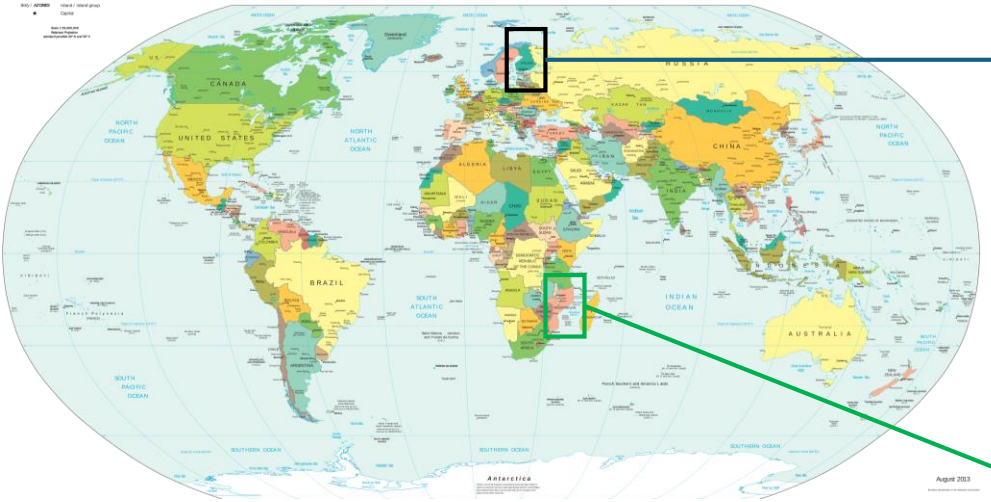


FDG PET

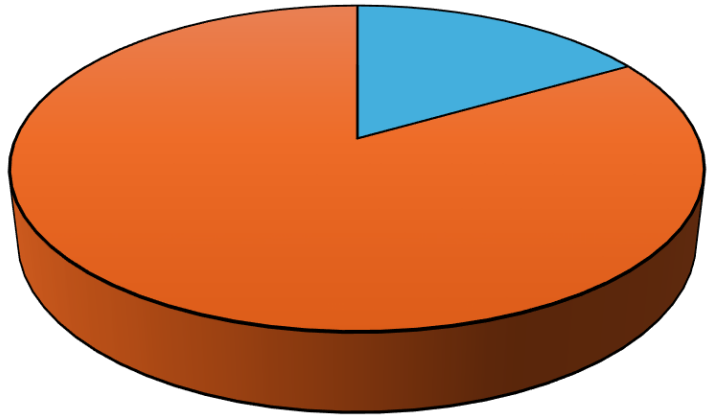


FDG PET/CT

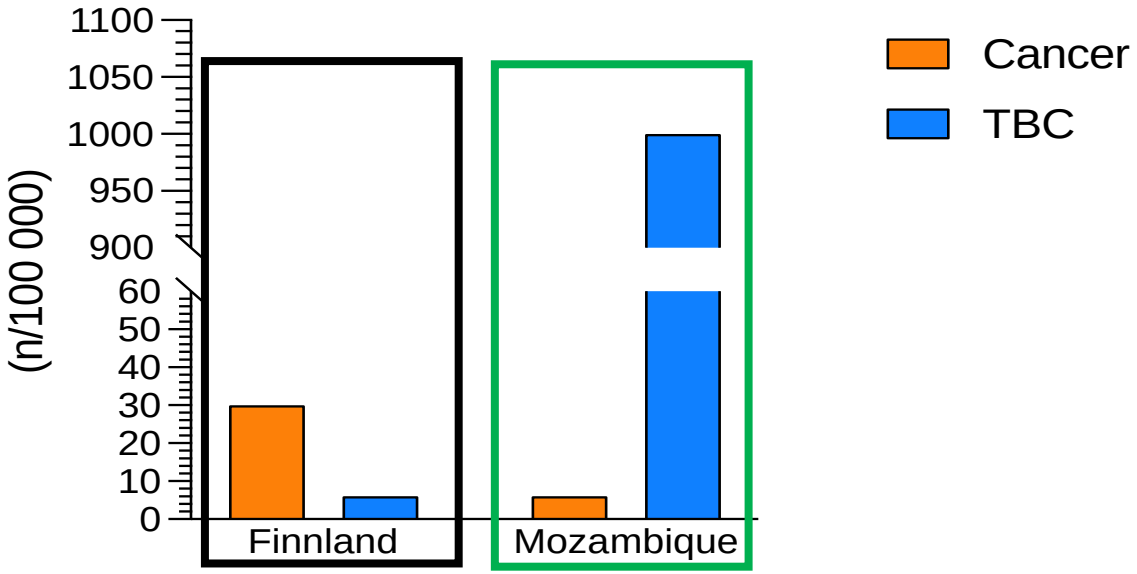
When statistics matters: the impact of Bayes theorem on AI performance (3)

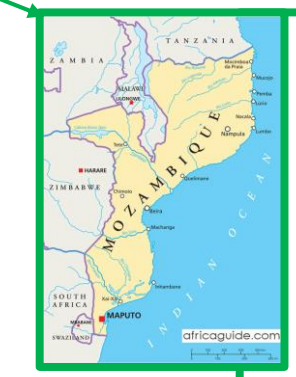


Training in Finnland

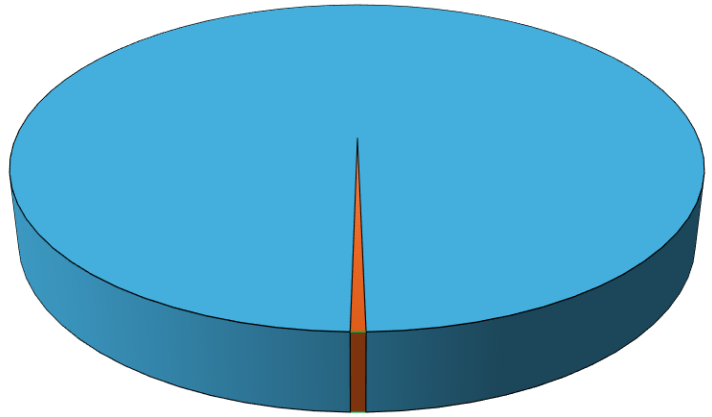


Differential prevalence of cancer and TBC





Training in Mozambique



(Lung Cancer): CA Cancer J Clin Oncol. 2024 doi: 10.3322/caac.21834

(TBC): Lancet. 2019. doi: 10.1016/S0140-6736(19)30308-3.

When statistics matters: the impact of Bayes theorem on AI performance (4)

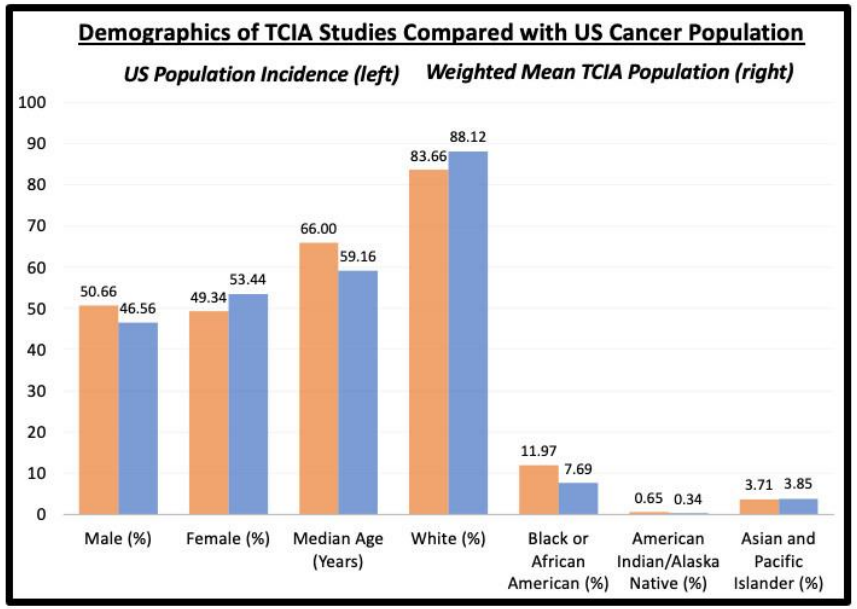
RSNA

Disparities in the Demographic Composition of The Cancer Imaging Archive

Radiologic images from TCIA are increasingly being used to train artificial intelligence-based medical tools, but its demographic composition is unknown.

Our investigation revealed an overrepresentation of younger, White, and female patients within the database compared with the broader U.S. cancer population.

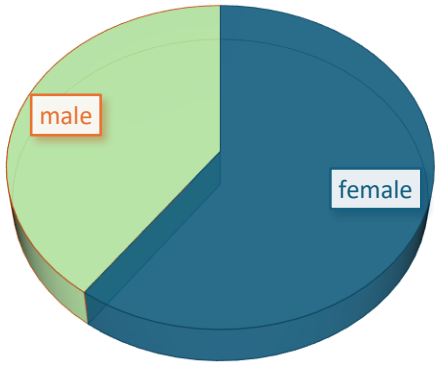
This raises concerns about potential biases in artificial intelligence models trained with TCIA data, which could lead to flawed generalizations and misinformed clinical decision-making, ultimately perpetuating cancer health disparities.



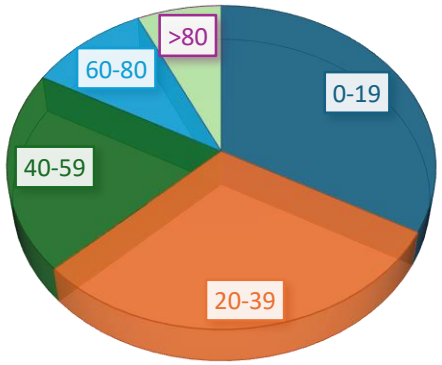
Dulaney A and Virostko J. Published Online: January 19, 2024
https://doi.org/10.1148/rycan.230100

Radiology: Imaging Cancer

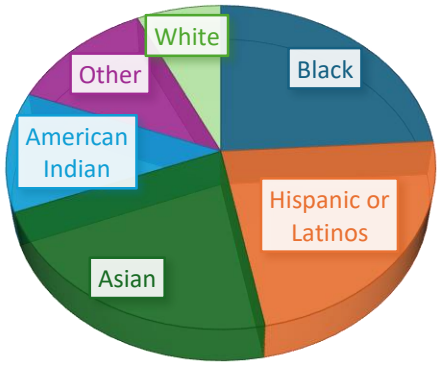
GENDER IDENTITY



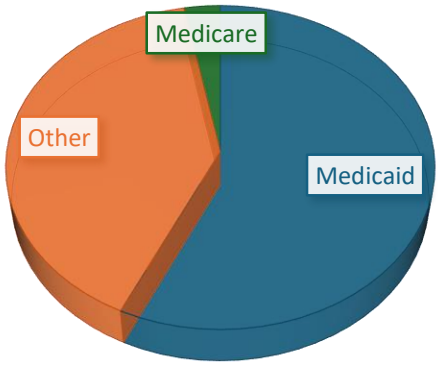
AGE (YEARS)



RACE AND ETHNIC GROUP



INSURANCE TYPE



EUROPEAN
COMMISSION

Brussels, 21.4.2021
COM(2021) 206 final

2021/0106 (COD)

Proposal for a

REGULATION OF THE EUROPEAN PARLIAMENT AND OF THE COUNCIL

LAYING DOWN HARMONISED RULES ON ARTIFICIAL INTELLIGENCE (ARTIFICIAL INTELLIGENCE ACT) AND AMENDING CERTAIN UNION LEGISLATIVE ACTS

{SEC(2021) 167 final} - {SWD(2021) 84 final} - {SWD(2021) 85 final}

TITLE III
HIGH-RISK AI SYSTEMS
CHAPTER 1
CLASSIFICATION OF AI SYSTEMS AS HIGH-RISK

Article 7

(b) Are defined High Risk the AI systems that pose a risk of harm to the health and safety, or a risk of adverse impact on fundamental rights.

Article 13

Transparency and provision of information to users

1. High-risk AI systems shall be sufficiently transparent to enable users to interpret the system's output and use

2. High-risk AI systems shall be accompanied by clear information relevant, accessible and comprehensible to users, specifying:

For these systems, it should be specified: …

performance as regards the persons or groups of persons on which the system is intended to be used;

any information in terms of training, validation and testing data sets, taking into account the intended purpose

the human oversight measures, including the technical measures able to facilitate the interpretation of the outputs of AI systems;

EUROPEAN
COMMISSION

Preamble
Comma 45

Brussels, 21.4.2021
COM(2021) 206 final

2021/0106 (COD)

Proposal for a

REGULATION OF THE EUROPEAN PARLIAMENT AND OF THE COUNCIL

LAYING DOWN HARMONISED RULES ON ARTIFICIAL INTELLIGENCE (ARTIFICIAL INTELLIGENCE ACT) AND AMENDING CERTAIN UNION LEGISLATIVE ACTS

{SEC(2021) 167 final} - {SWD(2021) 84 final} - {SWD(2021) 85 final}

For the development of high-risk AI systems, certain actors, such as providers, notified bodies and … and researchers, should be able to access and use high quality datasets within their respective fields of activities ….

…

For example, **in health, the European health data space will facilitate non-discriminatory access to health data and the training of artificial intelligence algorithms on those datasets, in a privacy-preserving, secure, timely, transparent and trustworthy manner, and with an appropriate institutional governance**.

Relevant competent authorities, including sectoral ones, providing or supporting the access to data may also support the provision of high-quality data for the training, validation and testing of AI systems.

AI is a planetary item

Governments should support collective development of international rules for the governance of AI.

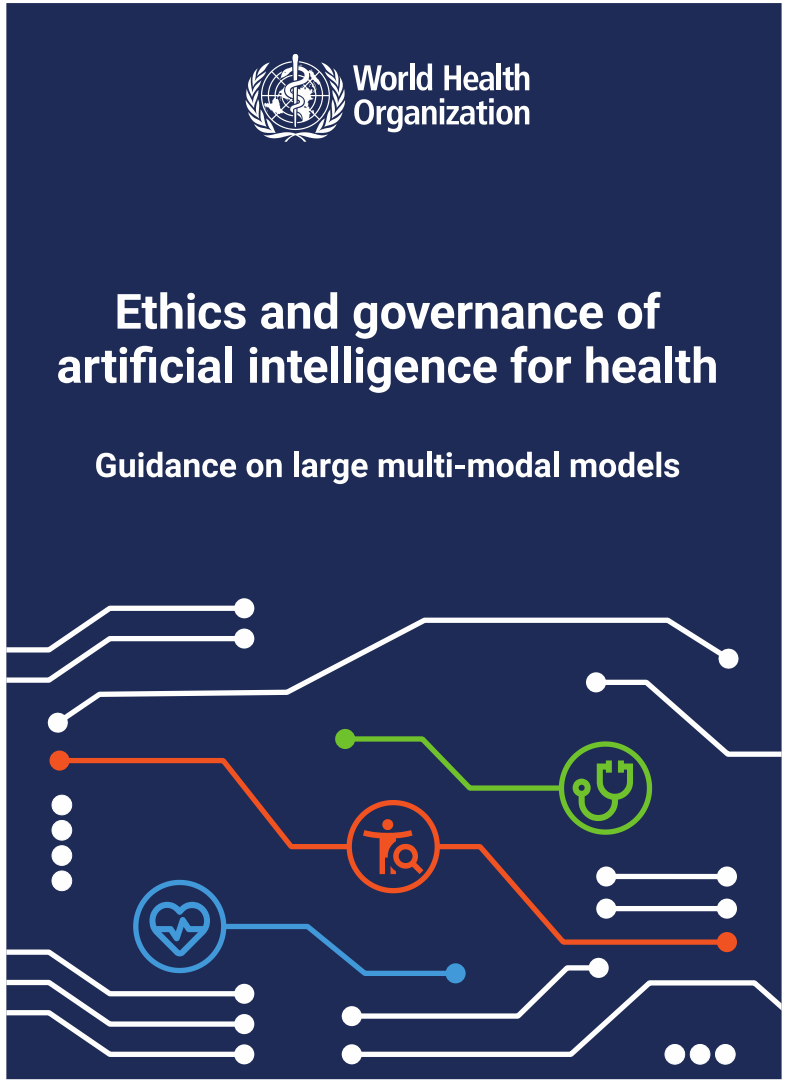
Whatever the form of governance, it must not be shaped solely by high-income countries or by high-income countries …, as that approach would leave most of humanity, in low- and middle-income countries, without a role or voice in shaping international governance of AI.

AI and advancement in medical science advancements

LMMs do not generate knowledge and have no moral or contextual reasoning … thus, it can produce useful information, yet it is not a substitute for knowledge production by humans.

AI should not de-humanize clinical care

Should AI technology reduce the contact between a provider and a patient, it could reduce the opportunities for clinicians to promote health and could undermine general supportive care, such as human–human interactions when people are often most vulnerable (1). Generally, there is concern that clinical care could be “de-humanized” by AI.



WHO report rev. 2024

Protect autonomy:
Humans should remain in control of health-care systems and medical decisions. …
People understand the role that AI systems play in their care.
Data privacy and confidentiality are protected by appropriate legal frameworks for data protection.

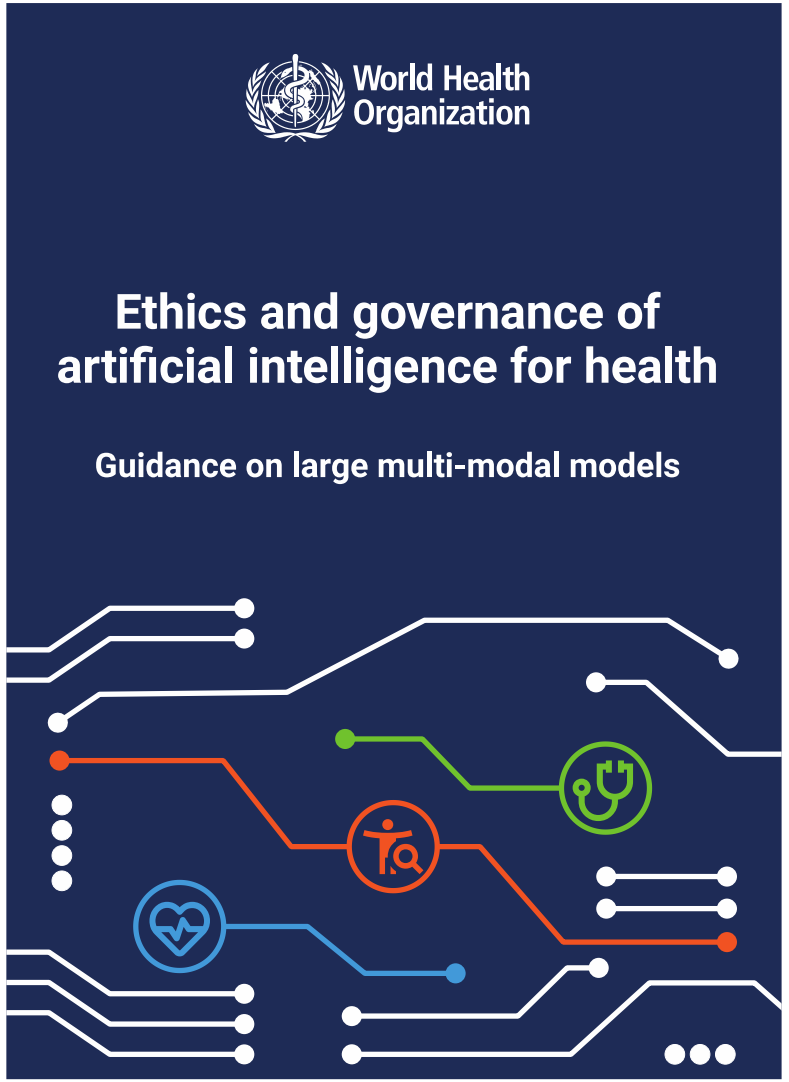
Promote human well-being, human safety and the public interest:
Designers of AI satisfy regulatory requirements for safety, accuracy and efficacy for well-defined uses or indications.

Ensure transparency, "explainability" and intelligibility:
AI technologies should be intelligible or understandable to medical professions, patients, users and regulators.

Foster responsibility and accountability:
Regulatory principles are applied upstream and downstream of the algorithm by establishing points of human supervision.

Ensure inclusiveness and equity:
AI is designed and shared to encourage the widest possible, appropriate, equitable use and access, irrespective of age, sex, gender identity, income, race, ethnicity, sexual orientation, ability or other characteristics.
AI is available for use not only in high-income settings but also in low- and middle-income countries.

Promote AI that is responsive and sustainable:
AI is consistent with the wider promotion of the sustainability of health systems, the environment and workplaces.



WHO report rev. 2024

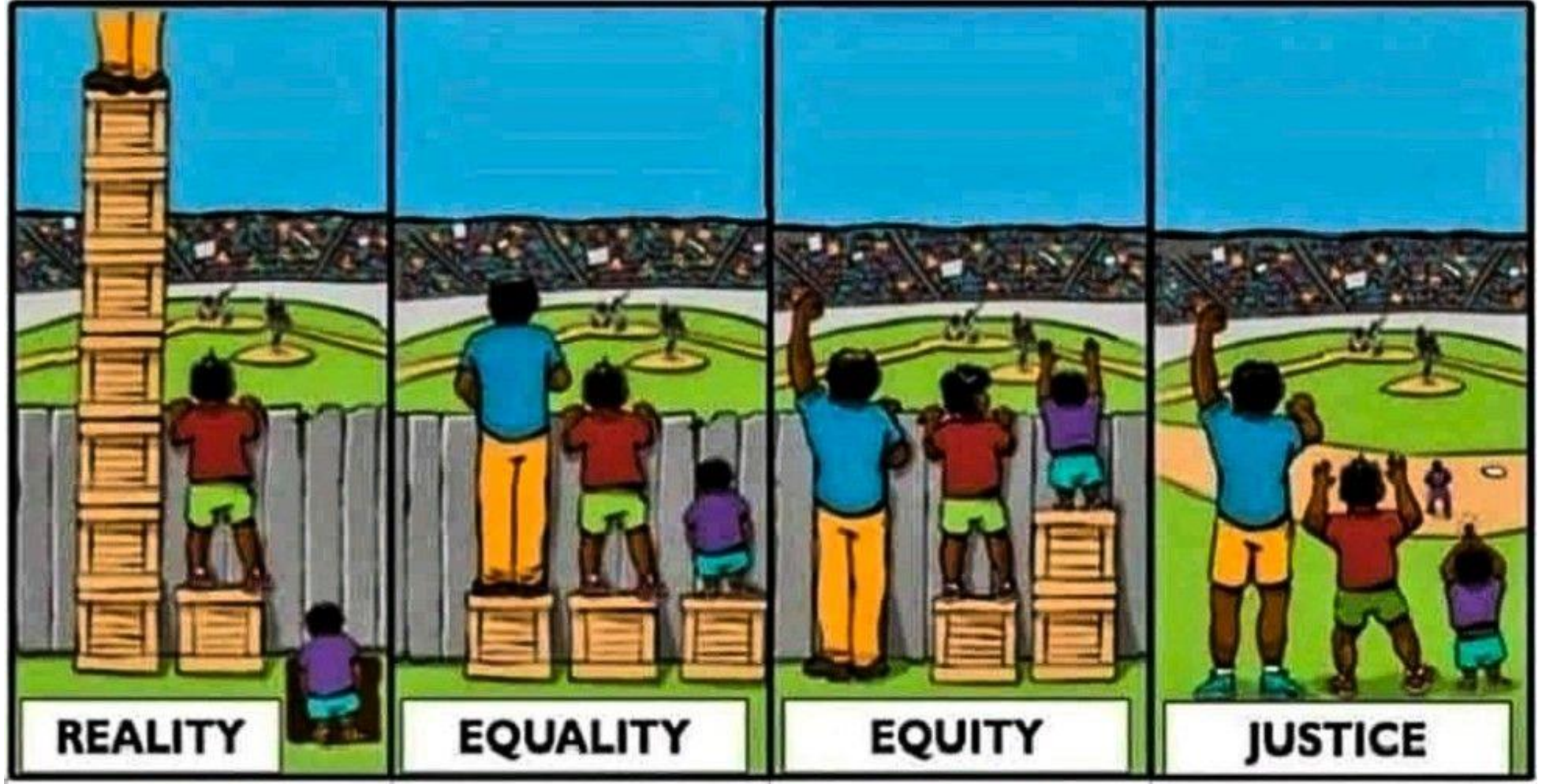
E, l’articolo 32 della ns Costituzione recita:

La Repubblica tutela la salute come fondamentale diritto dell'individuo e interesse della collettività, e garantisce cure gratuite agli indigenti.

Nessuno può essere obbligato a un determinato trattamento sanitario se non per disposizione di legge.

La legge non può in nessun caso violare i limiti imposti dal rispetto della persona umana.

La prevenzione sanitaria e l’assistenza
sanitaria devono quindi essere ispirate a:
Universalismo,
Equità,
Sostenibilità.



One gets more than is needed, while the other gets less than is needed. Thus, a huge disparity is created.	The assumption is that everyone benefits from the same supports. This is considered to be equal treatment.	Everyone gets the support they need, which produces equity.	All 3 can see the game without supports or accommodations because the cause(s) of the inequity was addressed. The systemic barrier has been removed.

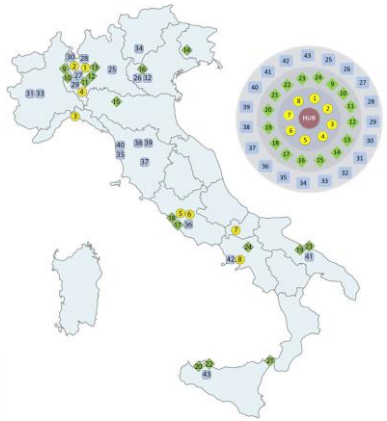
La performance diagnostica di analisi AI sulle immagini (e su ogni altro tipo di dato) è fortemente influenzata dall'esperienza dell'algoritmo.

L'esperienza dell'algoritmo dipende fortemente dalla sorgente delle basi dati a lui accessibili.

Le basi dati attualmente utilizzate sono
- Spesso (se non sempre) proprietarie del fornitore dell'algoritmo
- Caratterizzate da una aderenza indefinibile con la realtà etnica, sociale e geografica affrontata nei nostri Ospedali.

È necessario sviluppare un data base che il più possibile aderente alla realtà clinica vissuta da ogni operatore

Italian network of excellence for advanced diagnosis (**INNOVA**)



Ministero della Salute
DIREZIONE GENERALE DELLA RICERCA E DELL'INNOVAZIONE IN SANITA'

PIANO NAZIONALE PER GLI INVESTIMENTI COMPLEMENTARI AL PIANO NAZIONALE DI RIPRESA E RESILIENZA

PROGRAMMA «ECOSISTEMA INNOVATIVO DELLA SAUTE»

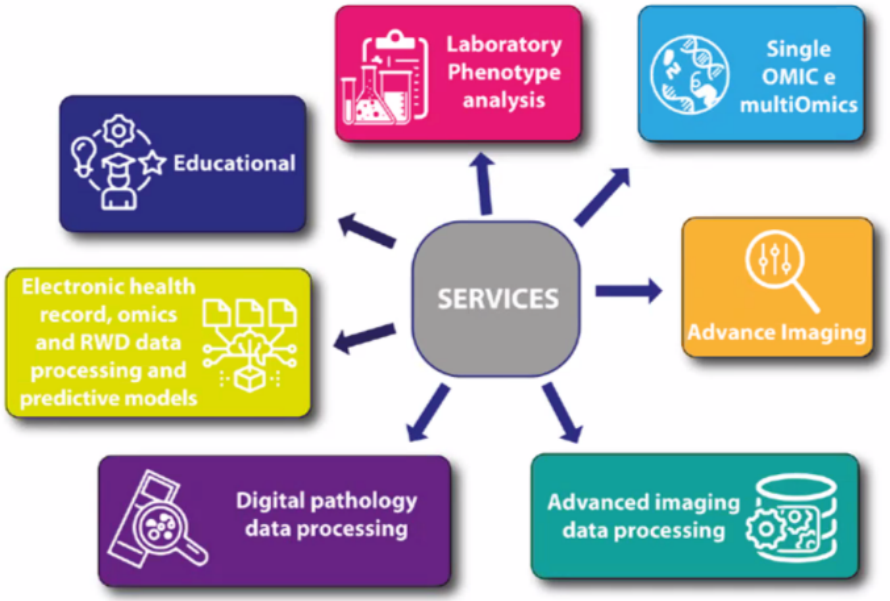
GENERAL OBJECTIVE

Creating a **diffused diagnostic platform providing access to cutting-edge technologies and expertise** available to the Italian life science R&D community.

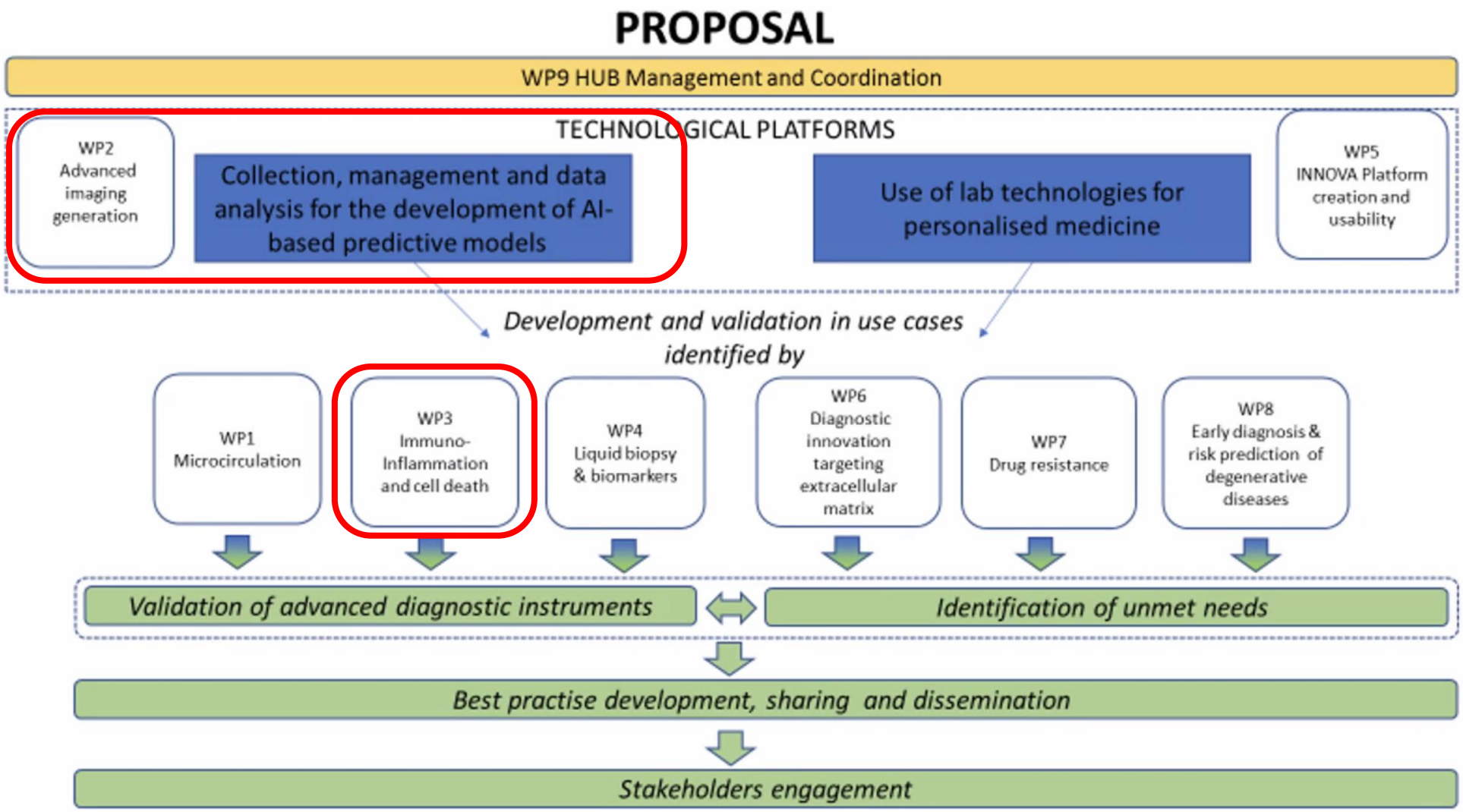
SPECIFIC OBJECTIVES

1. generating a "catalogue" of advanced diagnostic tools available to academia, national health system and industry;
2. establishing offer and access sites across the INNOVA members;
3. assisting and guiding users in the choice of the right approach to obtain reliable, relevant, and reproducible results;
4. establishing collaborations/partnership with industry players.

The catalogue will progressively absorb the **new advanced diagnostic approaches** generated by INNOVA research activities.



Project duration: first 4 years financed by MoH + 4 years



Italian network of excellence for advanced diagnosis (**INNOVA**)

Ministero della Salute
DIREZIONE GENERALE DELLA RICERCA E DELL'INNOVAZIONE IN SANITA'

PIANO NAZIONALE PER GLI INVESTIMENTI COMPLEMENTARI AL PIANO NAZIONALE DI RIPRESA E RESILIENZA

PROGRAMMA «ECOSISTEMA INNOVATIVO DELLA SAUTE»

WP2: advanced imaging generation

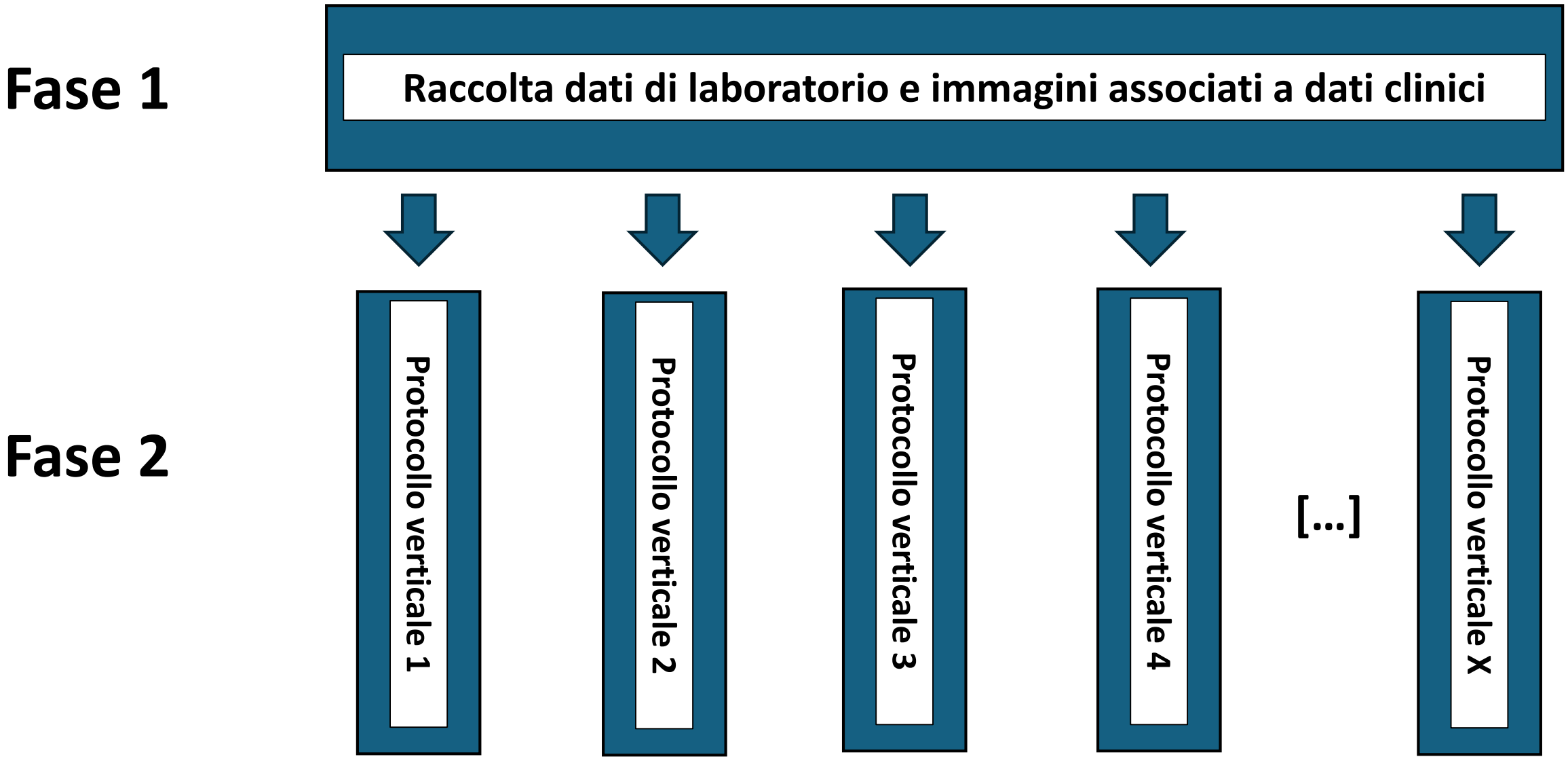
What *(Aims)*

WP2 is dedicated to the development of a Service distributed throughout the National Healthcare System to enhance the informative content of diagnostic imaging.

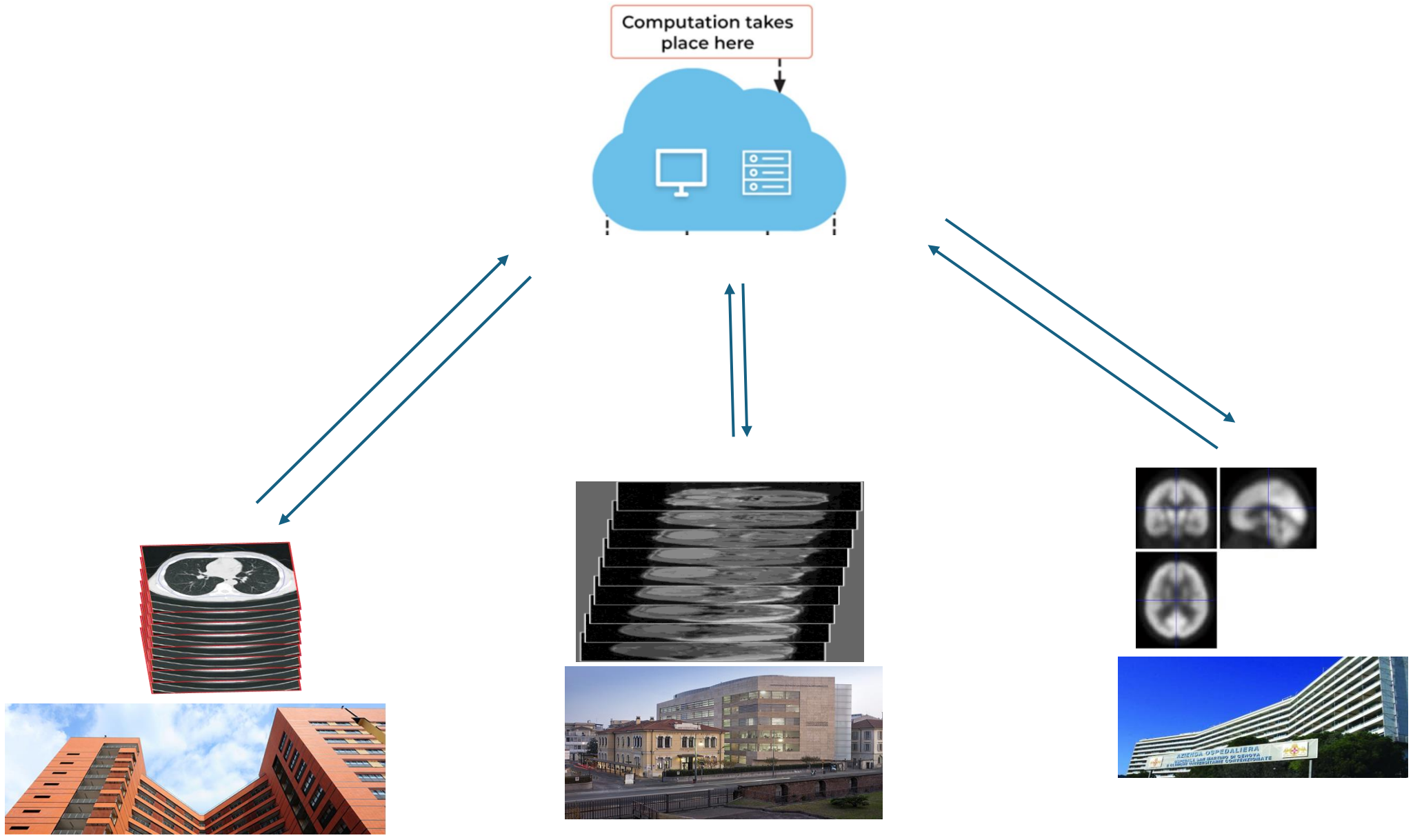
This function will be pursued approaching the following activities:

- developing new metrics and features,
- facilitating data interoperability
- improving the analytics tools needed to hybridize imaging and lab medicine
- disseminating info to imaging providers and consumers

Struttura di INNOVA: WP2 e WP3



Cloud Computing



AI, rispetto della privacy e aderenza al GDPR del settembre 2018 (1)

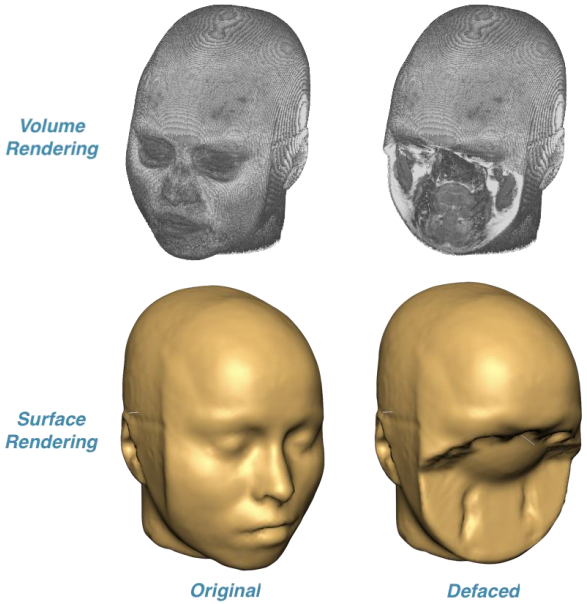
I dati:

1. sono raccolti per uno *scopo esplicito, determinato, legittimo*. I dati personali sono un'estensione del diritto umano fondamentale definito privacy
2. devono essere trattati in modo lecito, equo e trasparente.
3. *devono essere pertinenti, adeguati e limitati a quanto sufficiente per raggiungere lo scopo per il quale i dati sono trattati*.
4. …
5. …
6. *saranno trattati in modo da garantire la sicurezza contro trattamenti non autorizzati e perdite o danni accidentali.*

AI, rispetto della privacy e aderenza al GDPR del settembre 2018 (2)

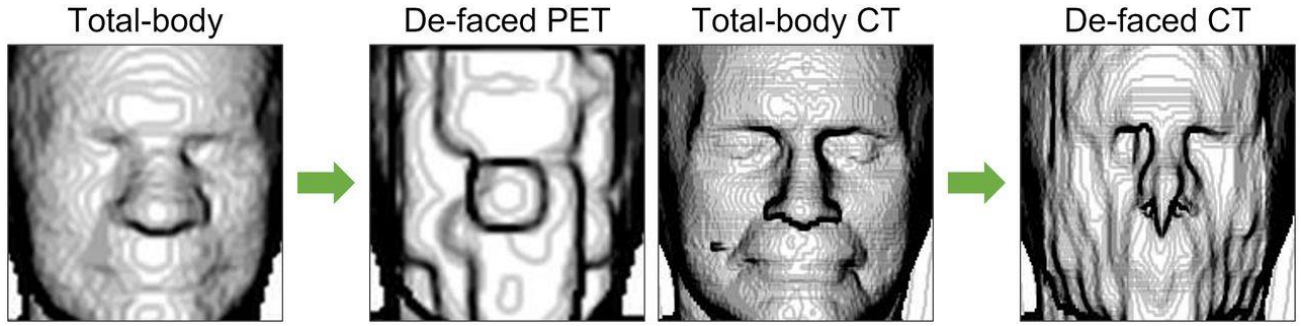
(l'anonimizzazione non è sufficiente)

Per studi focalizzati sul cervello



https://brainvoyager.com/

Per studi estesi a «testa/collo»

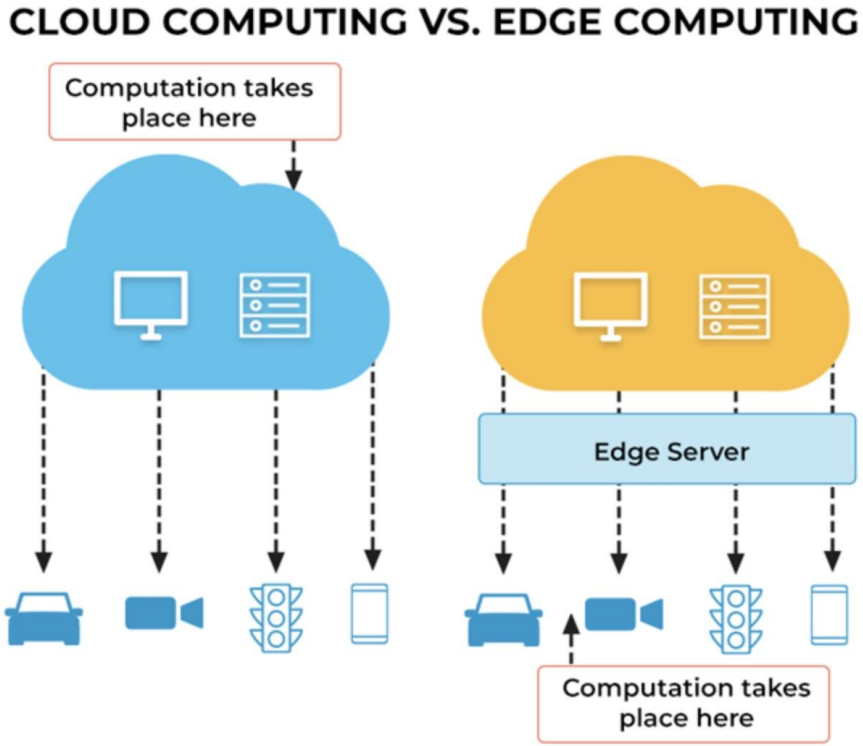


Selfridge AR et al Journal of Nuclear Medicine 2023, 64: 1304;

Ma … questo non è sufficiente: Ogni DPO consultato ha confermato che sito di provenienza, sesso, età e data dell'esame possono identificare un paziente, soprattutto in caso di patologie rare.

edge computing vs cloud computing

- l’edge computing avvicina l'elaborazione dei dati al luogo in cui vengono generati e archiviati
- il cloud computing si affida a una infrastruttura computazionale centrale



esempio: una telecamera di sicurezza ha un computer edge che consente di elaborare i dati direttamente sul dispositivo invece di inviare le riprese a un server cloud

Edge computing: vantaggi e svantaggi

vantaggi:

- **Minore latenza**:tempi di risposta più rapidi
- **Riduzione del traffico dat**i: meno dati inviati al cloud, con conseguente risparmio di banda
- **Rispetto del GDPR**: i dati sensibili possono essere trattati localmente
- **Sicurezza informatica** (per costruzione)
- **Maggiore affidabilità:** continua a funzionare anche in caso di connessione intermittente.
- Miglioramento delle prestazioni per le applicazioni in tempo reale

svantaggi :

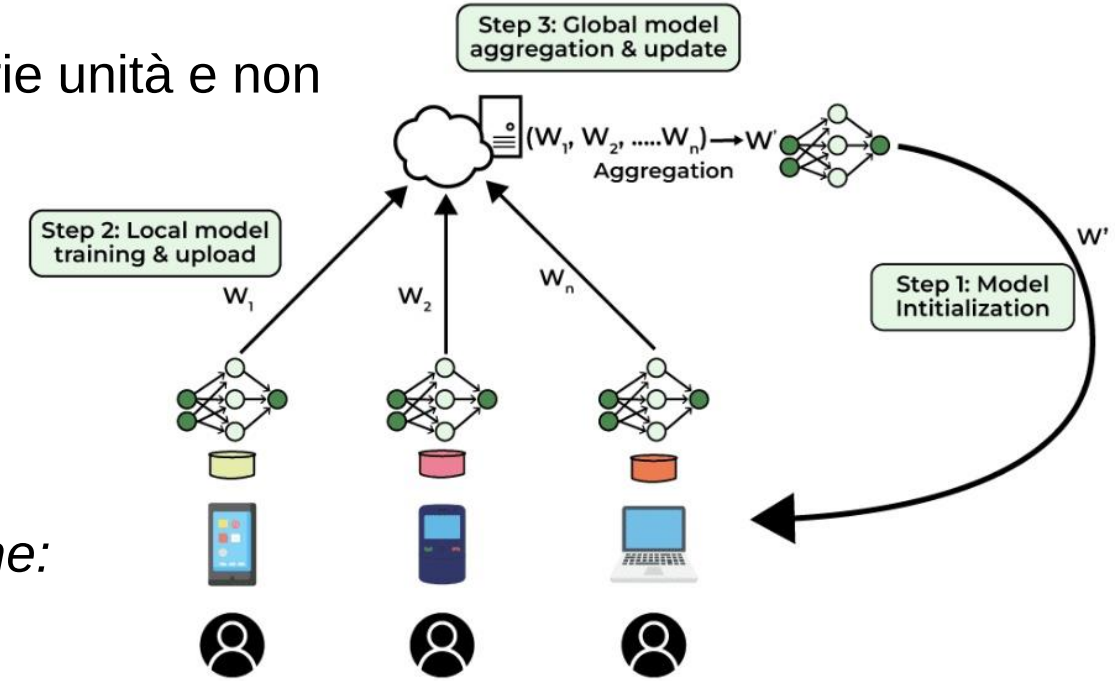
- non è possible usare tutti i dati prodotti in centri di cura lontani (come in un datacenter centralizzato) per programmare l'intelligenza artificiale

Federated learning

è la soluzione tecnologica che permette di usare i dati senza condividerli

come funziona:

- una rete neurale è una funzione con molti parametri detti pesi
- allenare una rete neurale consiste nello stimare i pesi utilizzando un archivio di dati annotati
- nel learning federato gli archivi dati rimangono nelle varie unità e non vengono condivisi
- a essere condivisi sono i pesi



in pratica il federated learning è un insieme di software che:

✓ lavorano separatamente sui dati in ogni centro

✓collaborano in modo criptato

✓aggiornano i risultati condividendo i parametri di ottimizzazione

L’**edge computing** è un modello di elaborazione dei dati in cui il calcolo e l’elaborazione avvengono **vicino alla fonte dei dati**, ossia **"al bordo" (edge)** della rete, piuttosto che in un data center centrale o nel cloud.

In pratica:

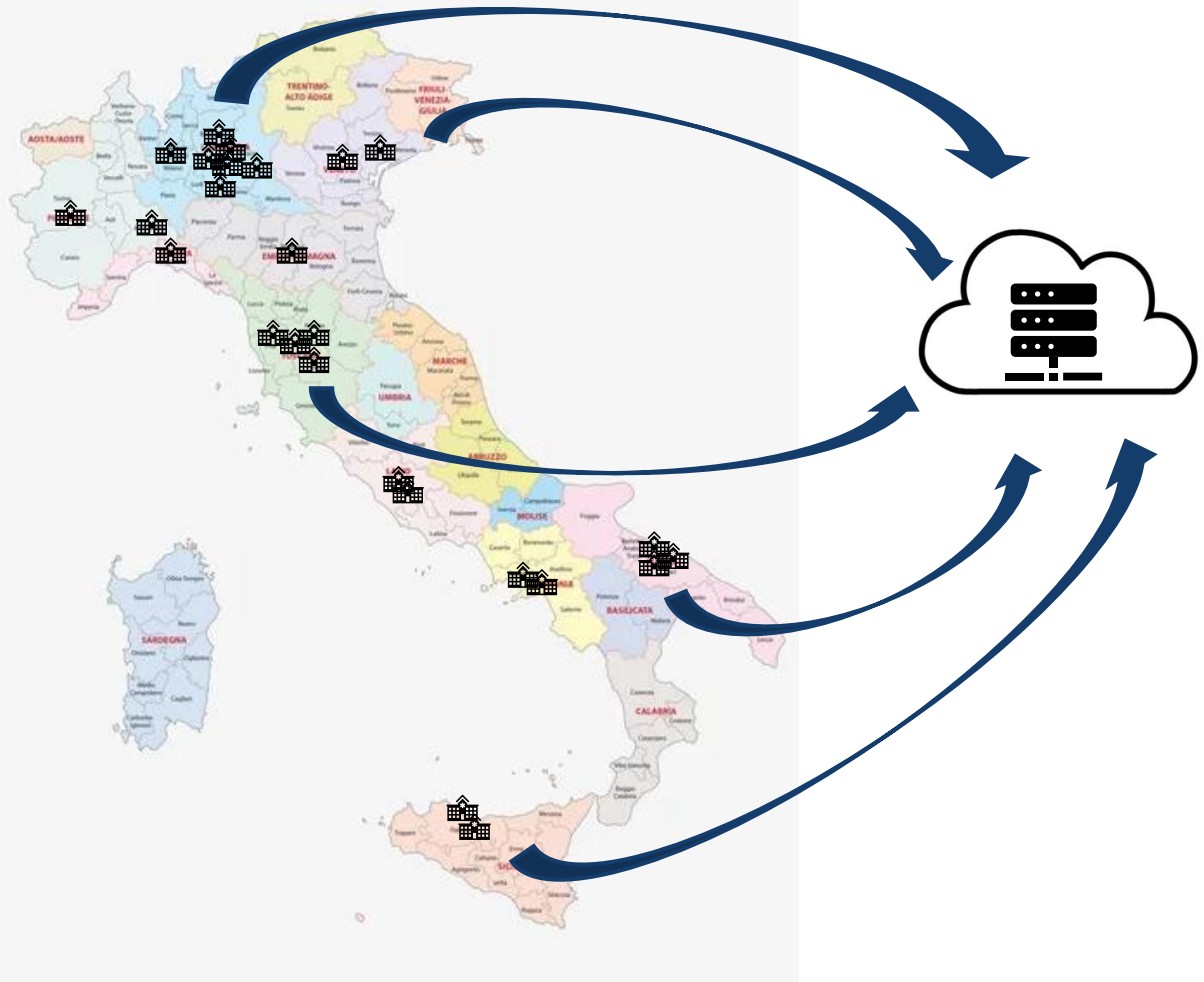
Invece di inviare tutti i dati a un server remoto per essere elaborati, l’edge computing permette di processare le informazioni **localmente**, ad esempio su dispositivi come:

- sensori IoT,
- router intelligenti,
- gateway,
- piccoli server locali.

Vantaggi principali:

- **Minore latenza**: risposta più veloce, utile in applicazioni critiche come veicoli autonomi o medicina.
- **Riduzione del traffico dati**: meno dati inviati al cloud, con conseguente risparmio di banda.
- **Maggiore privacy**: i dati sensibili possono essere trattati localmente.
- **Maggiore affidabilità**: continua a funzionare anche in caso di connessione intermittente.

ARCHITETTURA



ARCHITETTURA FEDERATA

+

EDGE COMPUTING

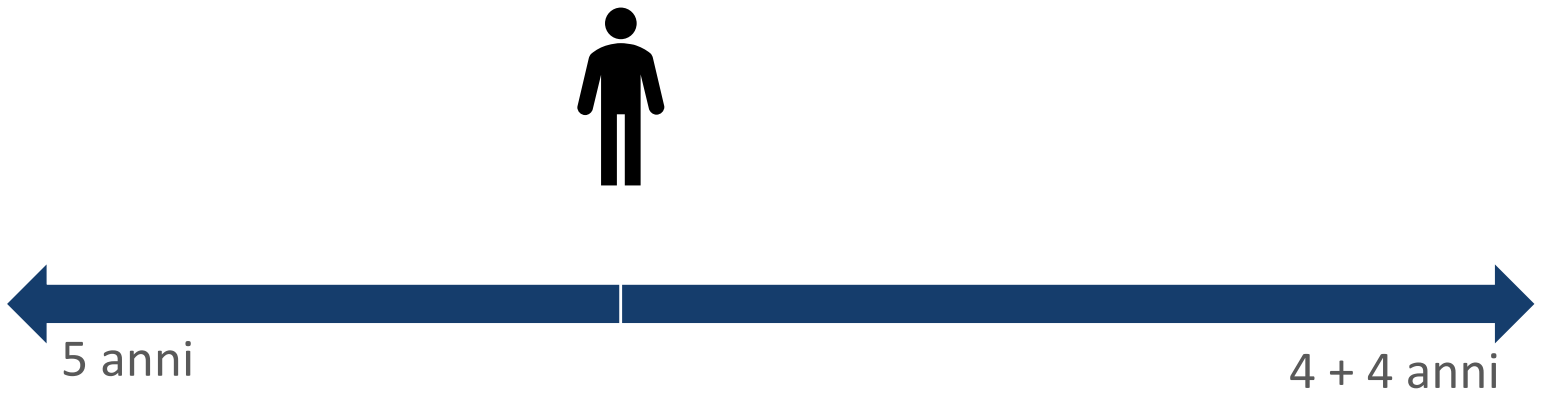
Xia, Qi, et al. "A survey of federated learning for edge computing: Research problems and solutions." *High-Confidence Computing* 1.1 (2021): 100008.

CRITERI DI INCLUSIONE

Centri coinvolti nel INNOVA-WP2 (Fase 1)

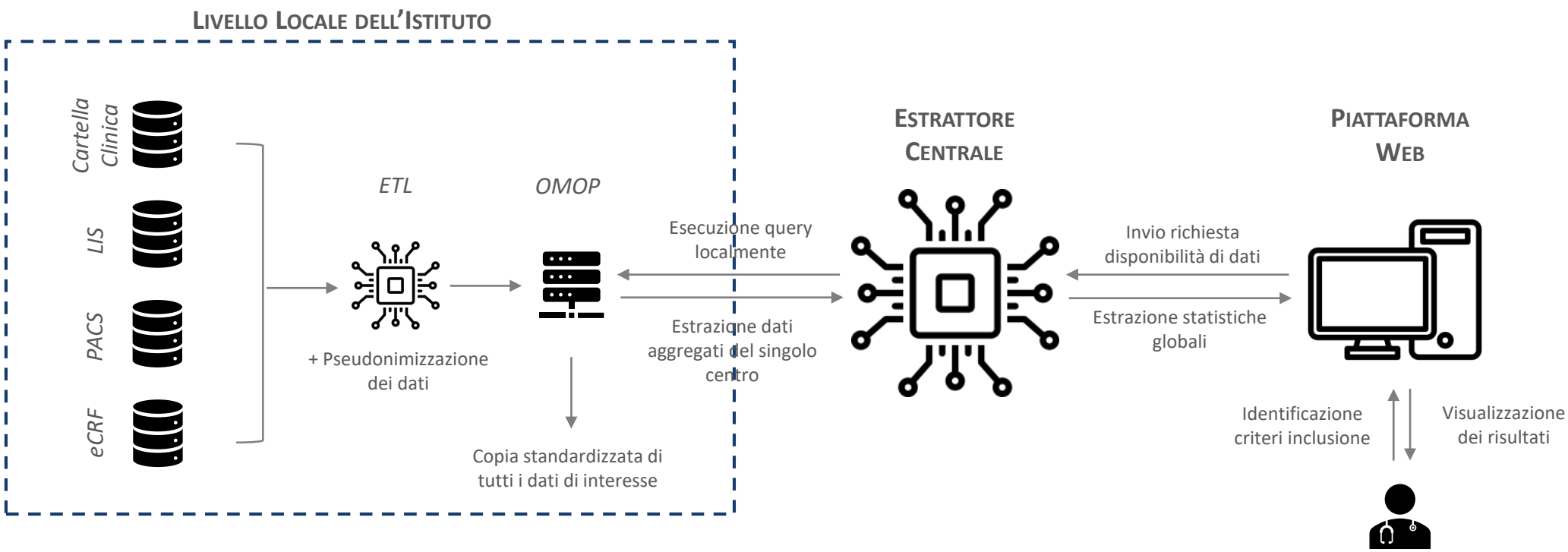
- Radiografia (RX)
- Tomografia computerizzata (CT)
- Ecografia (ECO)
- Risonanza magnetica (MRI)
- Scintigrafia o Tomografia ad emissione di fotone singolo con o senza CT (scintigrafia, SPECT o SPECT/CT)
- Tomografia ad Emissione di Positroni con CT o MRI (PET/CT o PET/RM)
- Optical coherence tomography (OCT)

VARIABILI DATASET CLINICO



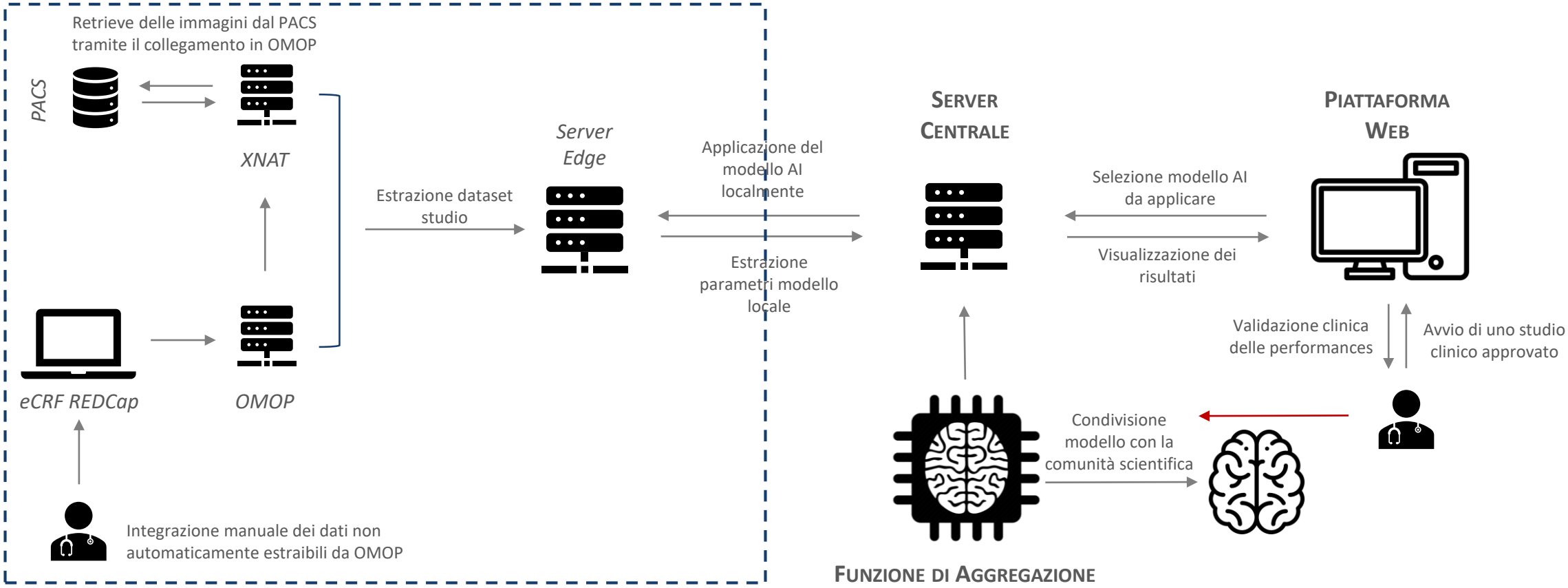
- Età
- Sesso
- Codice delle patologie riportate sulla scheda di dimissione ospedaliera (SDO) dei ricoveri
- Codice delle eventuali esenzioni dal ticket sanitario
- Risultati di tutte le indagini di imaging
- Risultati esami di laboratorio

FASE 1: Interoperabilità Tecnica



Kristoffersen, Espen Saxhaug, et al. "The Norwegian PraksisNett: a nationwide practice-based research network with a novel IT infrastructure." *Scandinavian Journal of Primary Health Care* 40.2 (2022): 217-226.

FASE 2: Interoperabilità Tecnica



Cremonesi, Francesco, et al. "The need for multimodal health data modeling: A practical approach for a federated-learning healthcare platform." Journal of Biomedical Informatics 141 (2023): 104338.

FASE 1: Interoperabilità Semantica

Dati Strutturati

Dati Non Strutturati

Possibile applicazione di tecniche di NLP per estrazione dati di interesse

Dati Imaging

In OMOP solo puntatori alle immagini nel PACS (non il dato stesso)

Pseudonimizzazione

- De-identification
- XNat

Mapping Vocabolari Standard

https://athena.ohdsi.org/search-terms/start
https://github.com/OHDSI/Vocabulary-v5.0

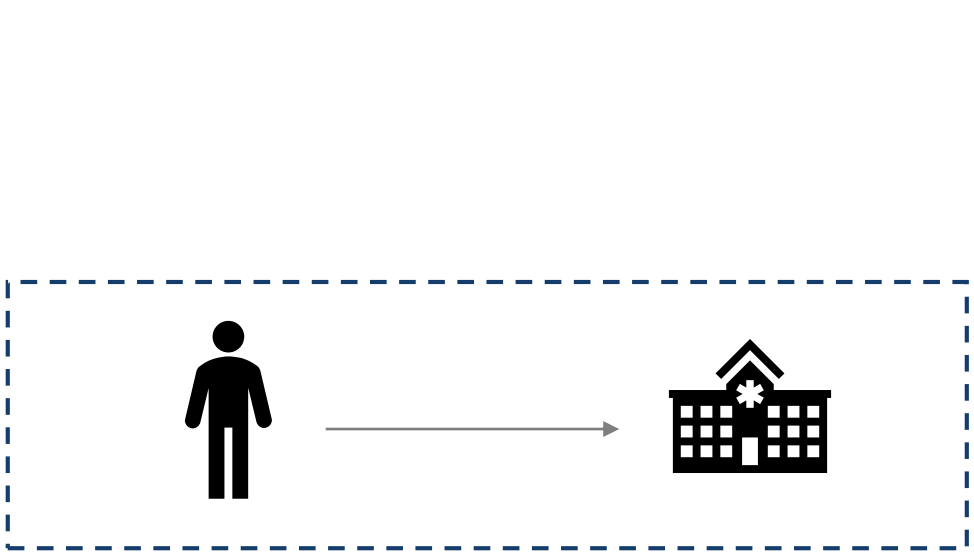
Domain	For Standard Concepts
Condition	SNOMED, ICDO3
Procedure	SNOMED, CPT4, HCPCS, ICD10PCS, **ICD9Proc**, OPCS4
Measurement	SNOMED, **LOINC**
Drug	RxNorm, RxNorm Extension, CVX
Device	SNOMED
Observation	**SNOMED**
Visit	CMS Place of Service, ABMT, NUCC

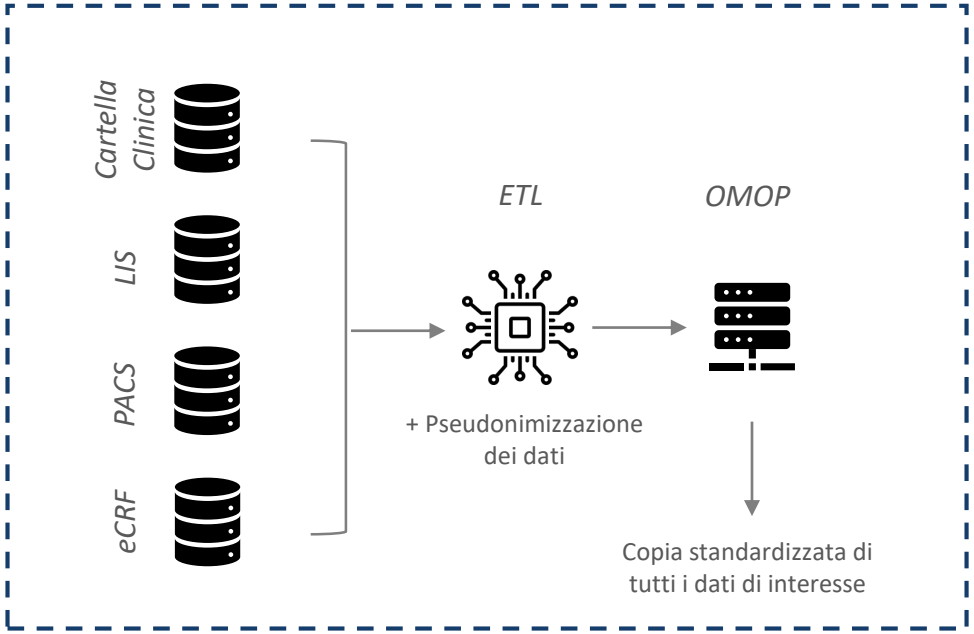
https://ohdsi.github.io/TheBookOfOhdsi/StandardizedVocabularies.html

MAPPING VOCABOLARI STANDARD

Variable Name	Variable Description	Variable Category	Variable type	Domain	Field	ID	Name	Class	Domain	Vocabulary
Sesso	Sesso	Anagrafica	M/F/O	PERSON	Gender_concept_id	*8507*	*MALE*	Gender	Gender	Gender
Codice ICD9 SDO	Codice delle patologie riportate sulla SDO	Anamnestica	VARCHAR	CONDITION_OCCURRENCE	Condition_concept_id	*44825230*	*Plasma cell leukemia, in remission*	Condition	Condition	ICD9CM
Codice Esenzione	Codice delle esenzioni dal ticket sanitario	Anamnestica	VARCHAR	OBSERVATION	Observation_concept_id					
Modality	Tipo di Imaging DICOM Tag: (0008,0060) e (0008,0061)	Diagnostica per immagini	MR/CR/CT/US/NM/…	IMAGING_OCCURRENCE	modality_concept_id	*4234381*	*Magnetic resonance imaging unit*	Physical Object	DEVICE	SNOMED
PLT	Piastrine	Esami di Laboratorio	NUMBER	MEASUREMENT	measurement_concept_id	4267147	Platelet count	Procedure	Measurement	SNOMED
					unit_concept_id	9444	billion per liter	Unit	Unit	UCUM

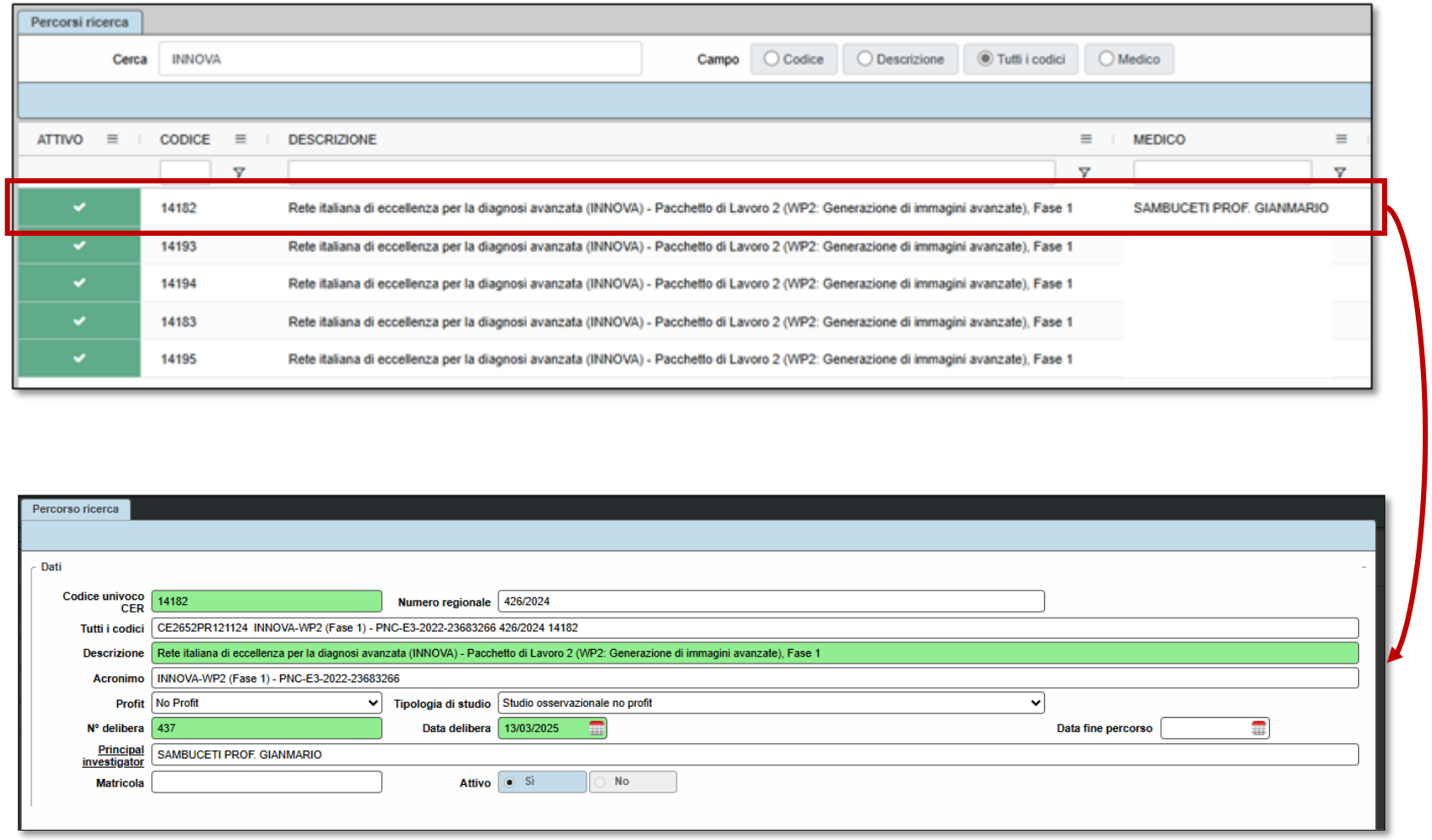
TRA EVENTO E FLUSSO INFORMATIZZATO



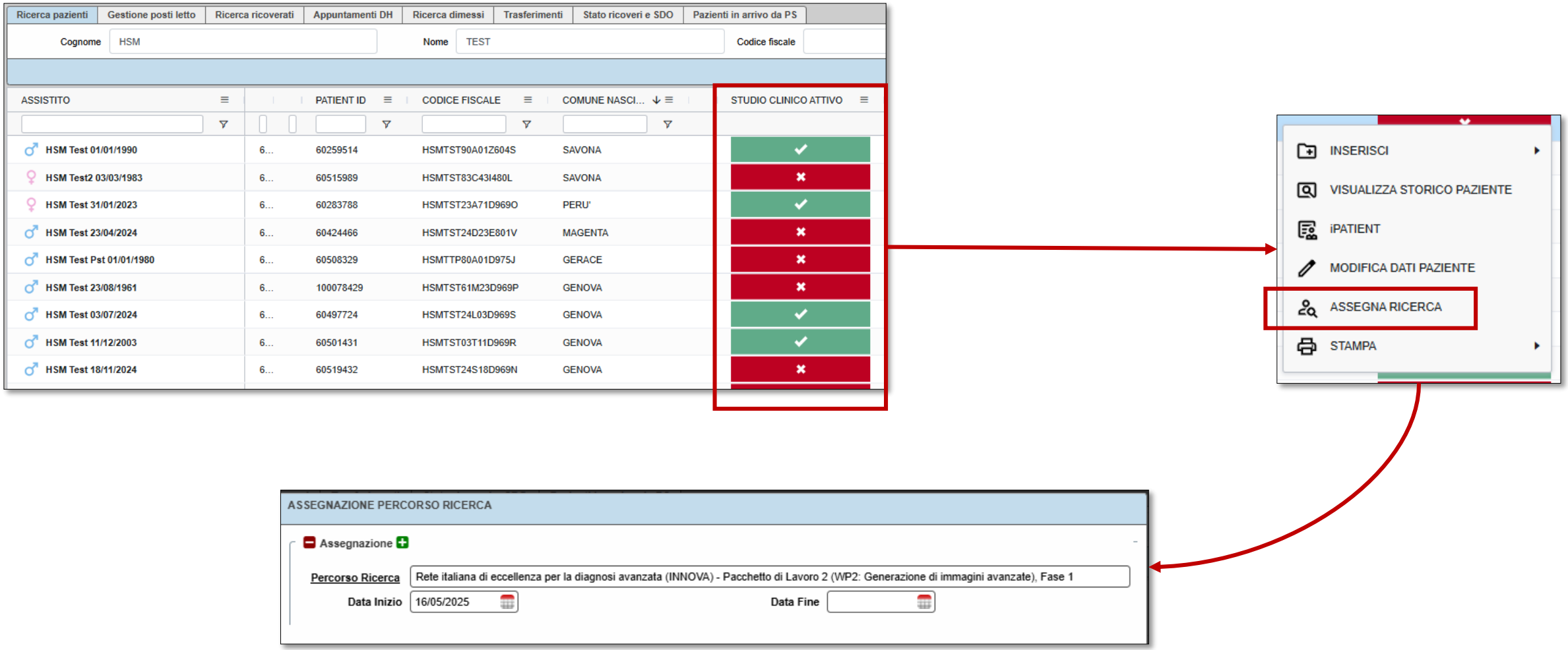


SERVE UN TRIGGER!

CENSIMENTO STUDI di RICERCA su CCE

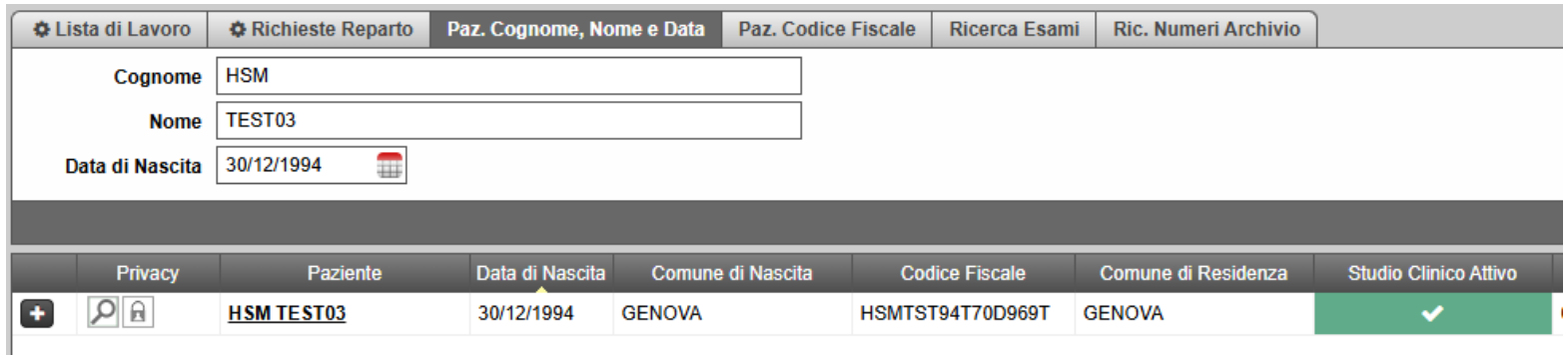


ARRUOLAMENTO PAZIENTE

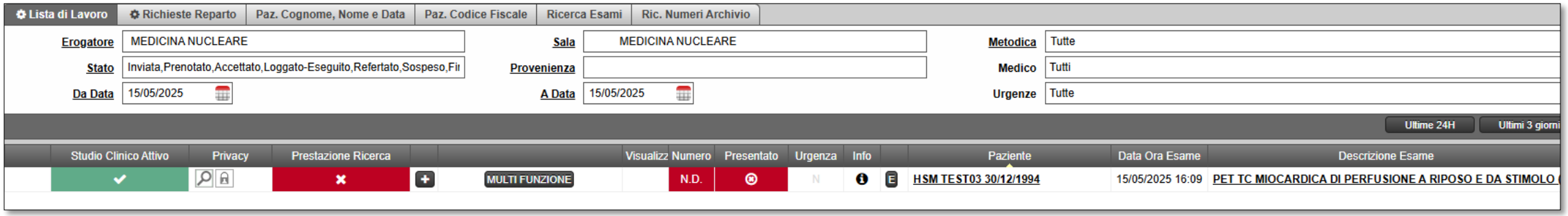


ACCETTAZIONE PAZIENTE

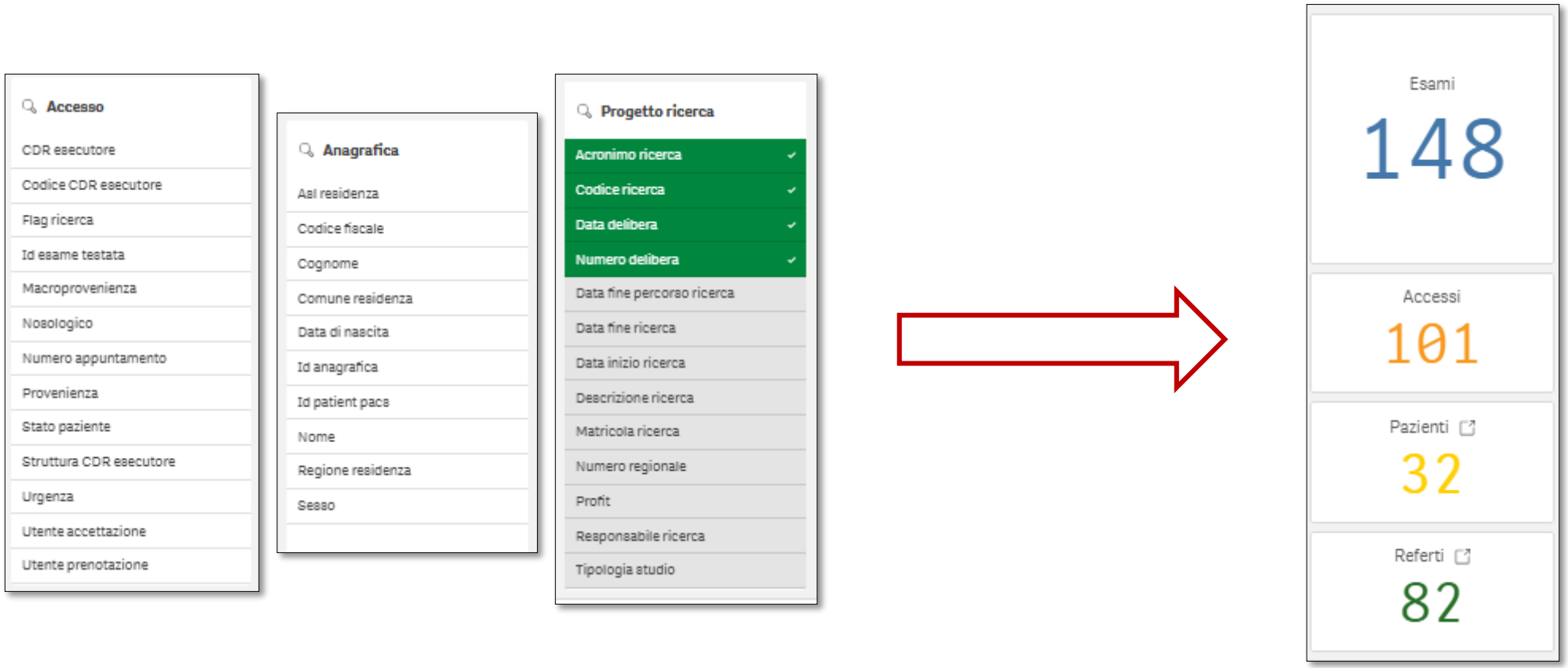
Maschera di Ricerca Paziente con i Dati Anagrafici:



Worklist della Giornata:



MONITORAGGIO CON BI



GRAZIE PER L’ATTENZIONE!