IHE EUROPE EXPERIENCE DAYS

VIENNA 2025

24-25 JUNE

Evaluation of AI algorithms in health

Sandie Cabon - Research Engineer, PhD

LTSI – UMR 1099 Inserm - Université de Rennes

sandie.cabon@univ-rennes.fr

Laboratoire Traitement du Signal et de l'Image LTSI

Université de Rennes

Inserm La science pour la santé From science to health

CHU Rennes

Testing and Experimentation Facility for Health AI and Robotics TEF-Health

IHE CATALYST

What is an AI system ?

What is an AI system ?

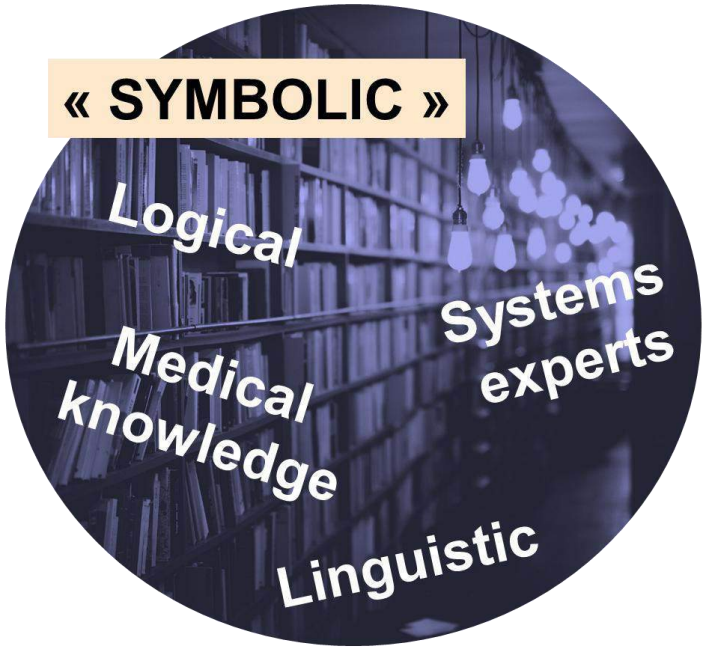
« a machine-based system designed to operate with varying levels of autonomy and, that may exhibit adaptiveness after deployment and that, for explicit or implicit objectives, **infers**, from the input it receives**, how to generate outputs** such as predictions, content, recommendations or decision that can influence physical or virtual environments »

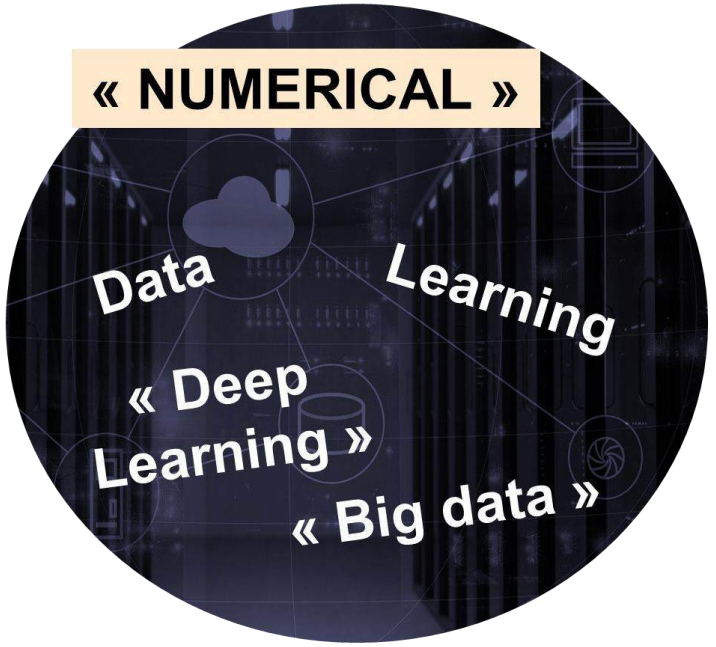
Source: N Alkhayat DG SANTE ‘AI Act / MDR webinar with TEF-Health’, 10. Jun 2024

What is an AI system ?

« a machine-based system designed to operate with varying levels of autonomy and, that may exhibit adaptiveness after deployment and that, for explicit or implicit objectives, **infers**, from the input it receives**, how to generate outputs** such as predictions, content, recommendations or decision that can influence physical or virtual environments »

Source: N Alkhayat DG SANTE 'AI Act / MDR webinar with TEF-Health', 10. Jun 2024





What use of AI in health ?

What use of AI in health ?

Precision Medicine
e.g., recommending treatment

Preventive
e.g., pharmacovigilance

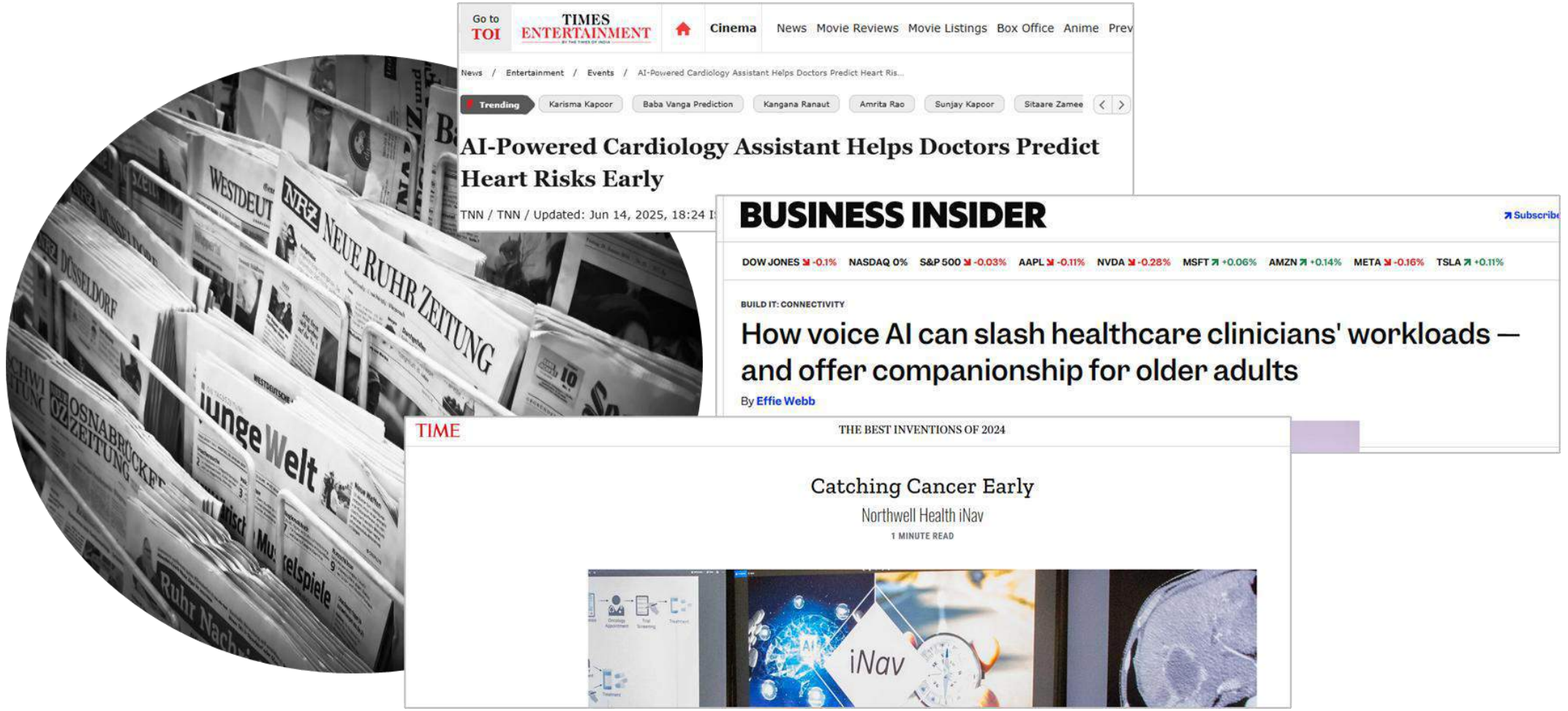
Robots /Assistant
e.g., guiding patients

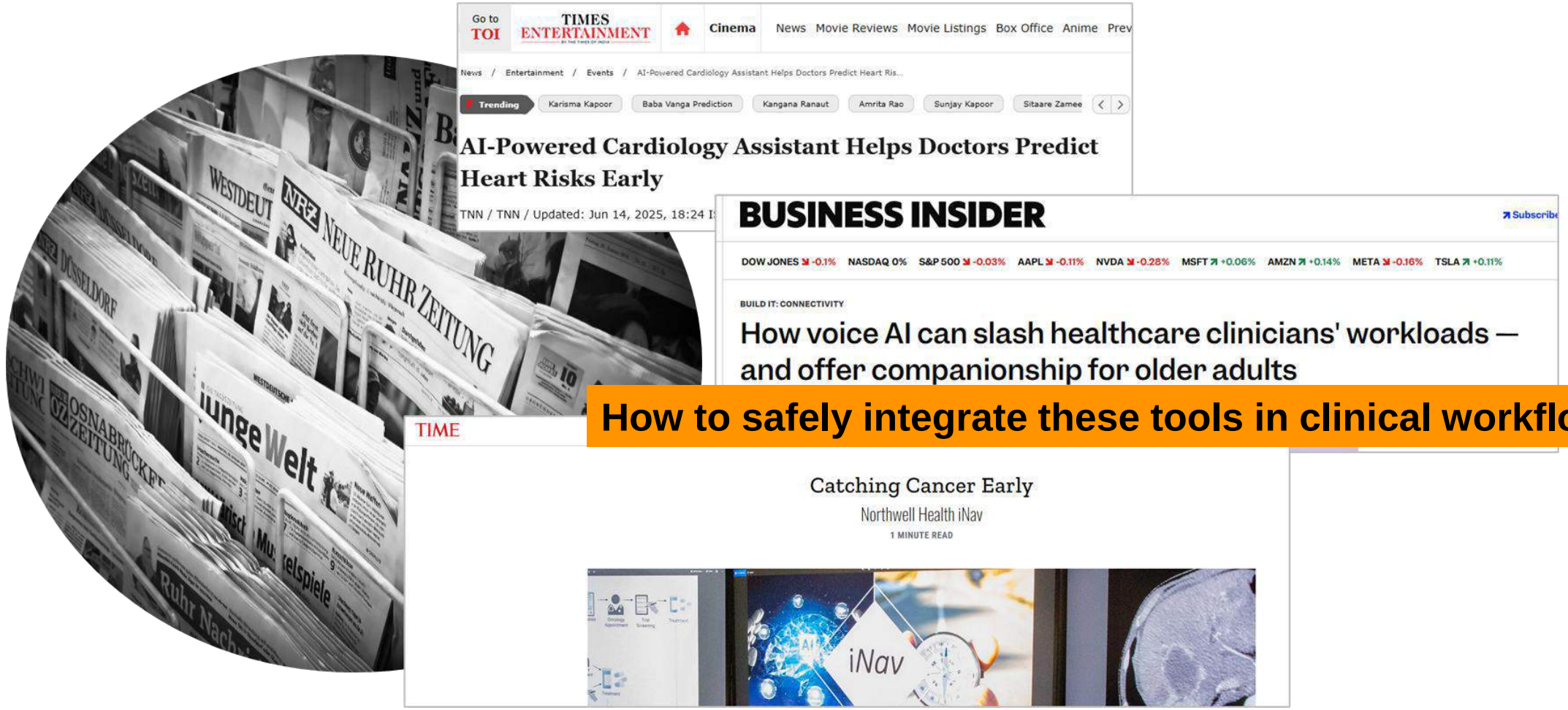
Predictive
e.g., disease prediction

Decision Support
e.g., health indicators suplpying

Assisted Surgery
e.g., specifying zone

Evaluation of AI algorithms in health





How to safely integrate these tools in clinical workflow?

IHE EUROPE EXPERIENCE DAYS

VIENNA 2025

24-25 JUNE

EU AI ACT – **The world's first regional regulation to enforce trustworthy AI**

EU Artificial Intelligence Act

IHE CATALYST

Conformity assessment

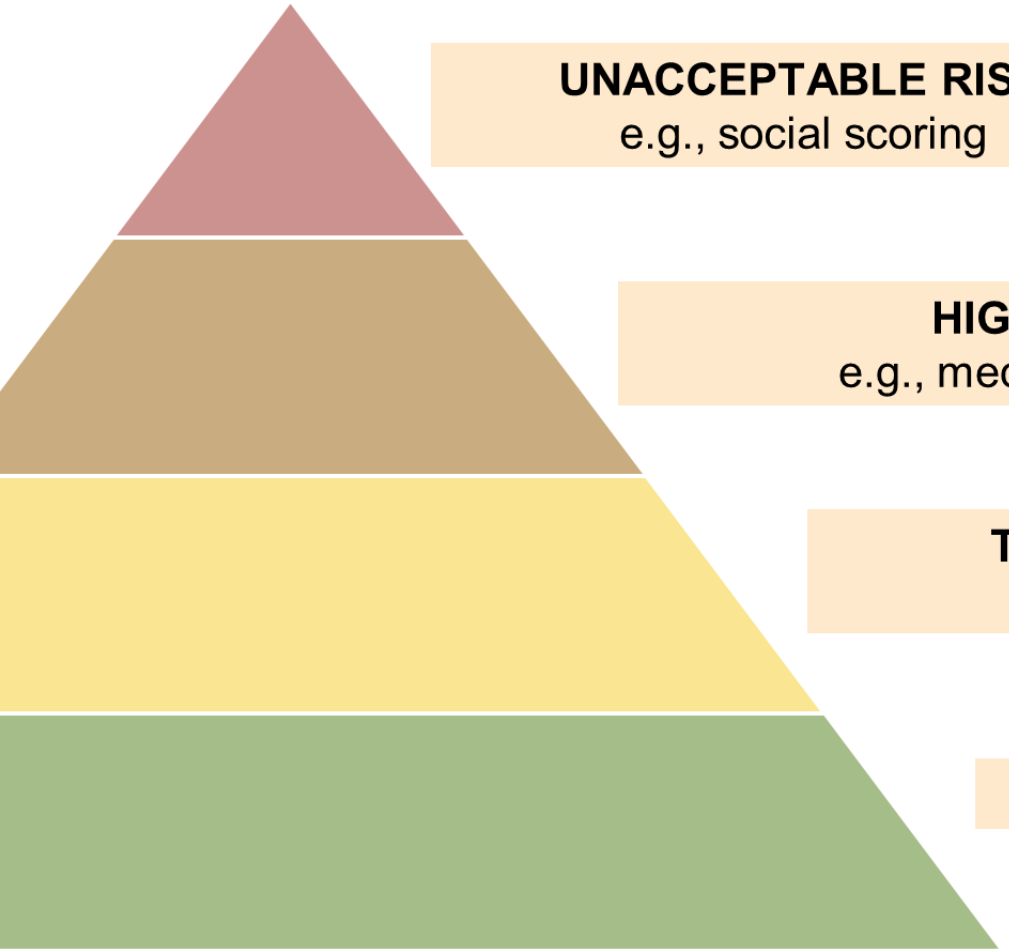
12 July 2024	1 Aug 2024	2 Feb 2025	2 Aug 2025	2 Aug 2026	2 Aug 2027	2 Aug 2030	31 Dec 2030
EU AI Act published in the Official Journal of the EU on **12 July 2024**	EU AI Act enters into force **1 Aug 2024** (Art.113)	Rules on subject matter, scope, definitions, AI literacy, and prohibitions come into effect on **2 Feb 2025** (Art.113(a))	Rules on notifications, GPAI models, certain enforcement issues, and penalties come into effect on **2 Aug 2025** (Art.113(b))	The EU AI Act applies, the general grace period for high-risk AI systems ends, and the bulk of the operative provisions come into effect, on **2 Aug 2026** (Arts.111(2) & 113)	Rules on high-risk AI systems under Art.6(1) come into effect on **2 Aug 2027** (Art.113(c))	The grace period for high-risk AI systems intended for use by public authorities ends on **2 Aug 2030** (Art.111(2))	

Grace period

EU Artificial Intelligence Act

Risk-based approach

Risk-based approach



UNACCEPTABLE RISK
e.g., social scoring

Prohibited

HIGH RISK
e.g., medical devices

Permitted but subject to compliance with AI requirements and ex-ante conformity assessment

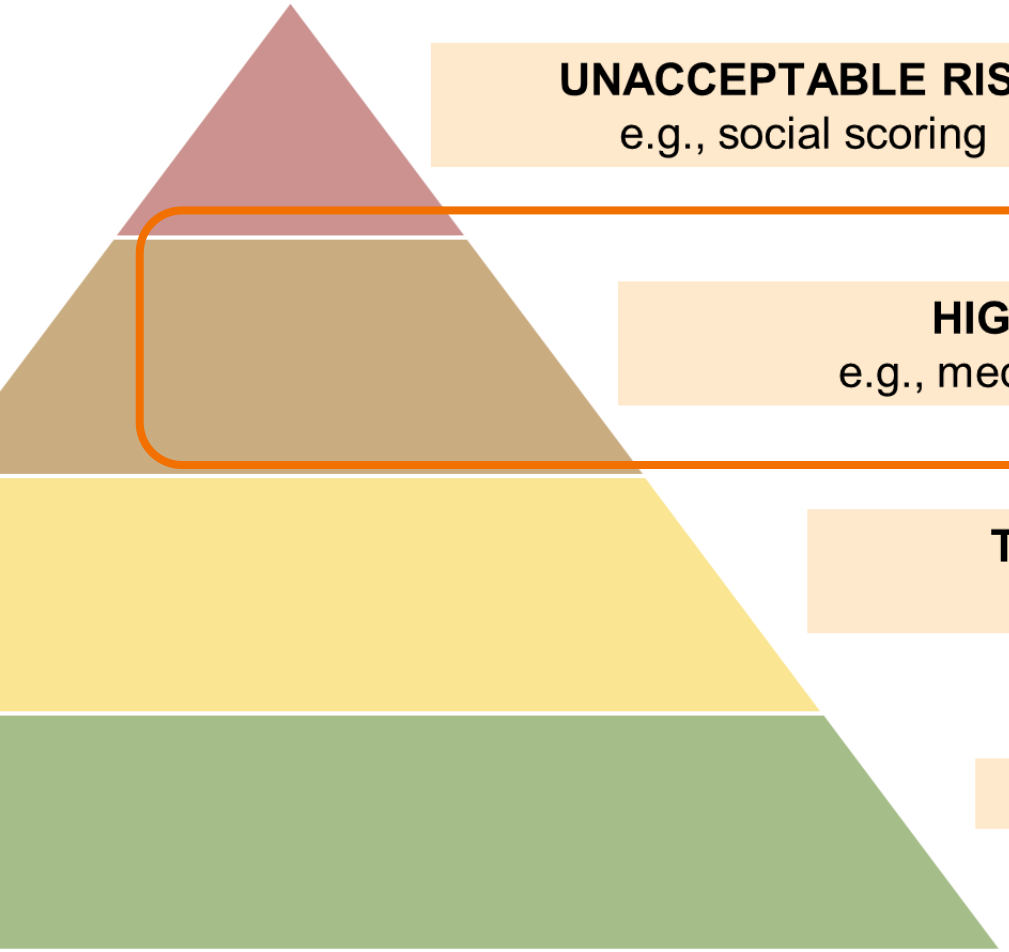
TRANSPARENCY RISK
e.g., deep fake

Permitted but subject to information/transparency obligations

MINIMAL OR NO RISK

Permitted with no restrictions

Risk-based approach



UNACCEPTABLE RISK
e.g., social scoring

Prohibited

HIGH RISK
e.g., medical devices

Permitted but subject to compliance with AI requirements and ex-ante conformity assessment

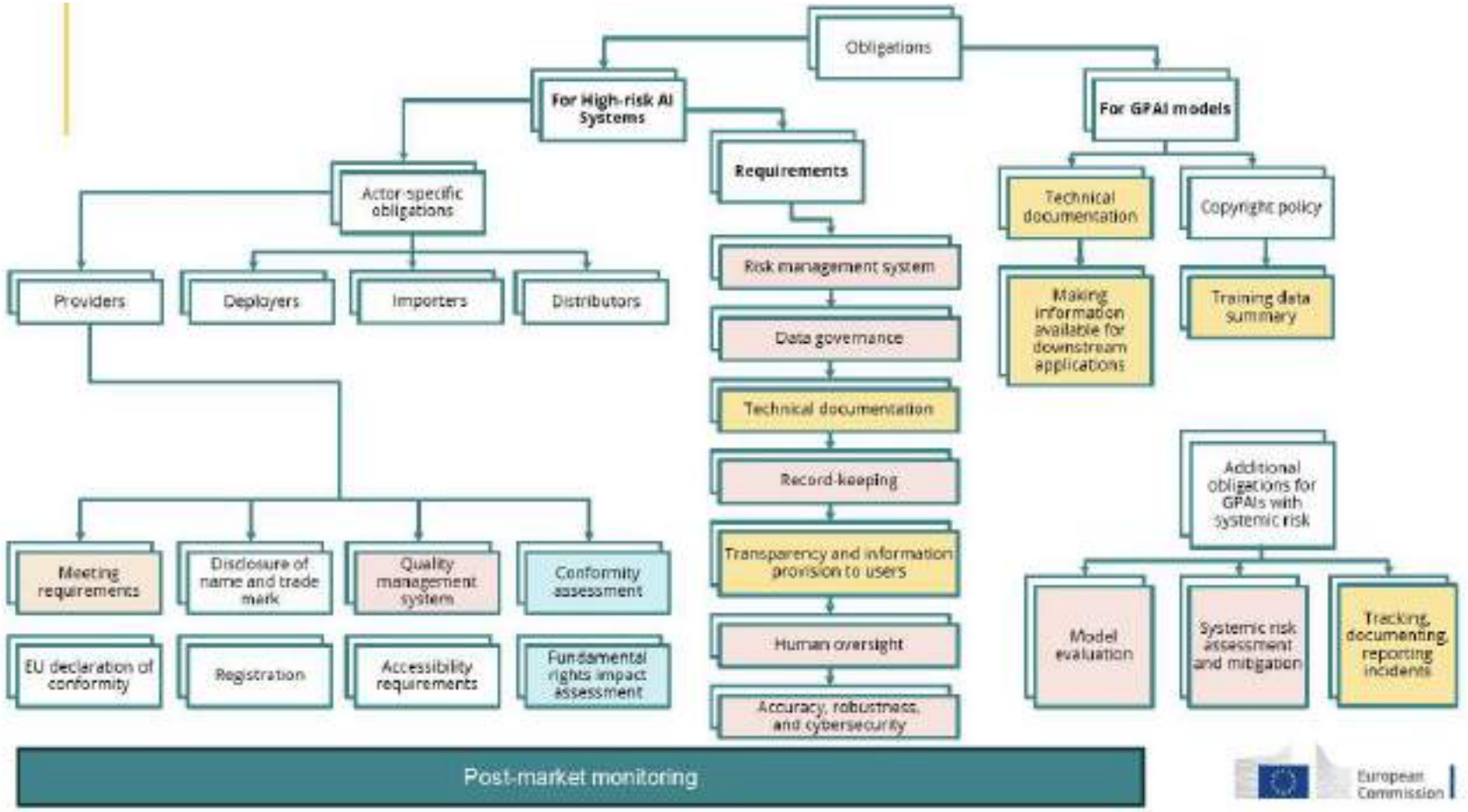
TRANSPARENCY RISK
e.g., deep fake

Permitted but subject to information/transparency obligations

MINIMAL OR NO RISK

Permitted with no restrictions

Requirements

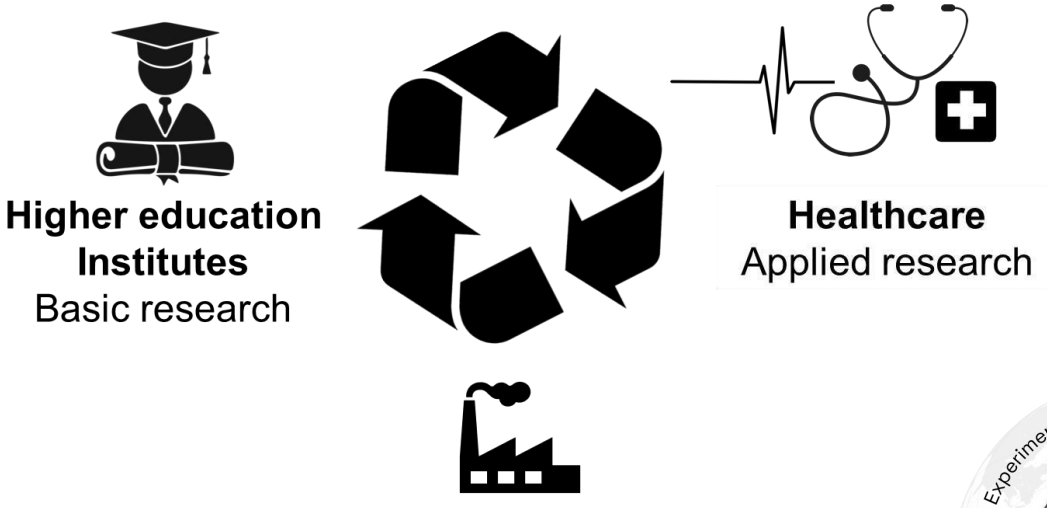


Testing and Experimentation Facilities (TEFs)

AI-Act: Defines the role of TEFs
As implementation instrument
TEFs provide ”technical and scientific support for providers and notified bodies“

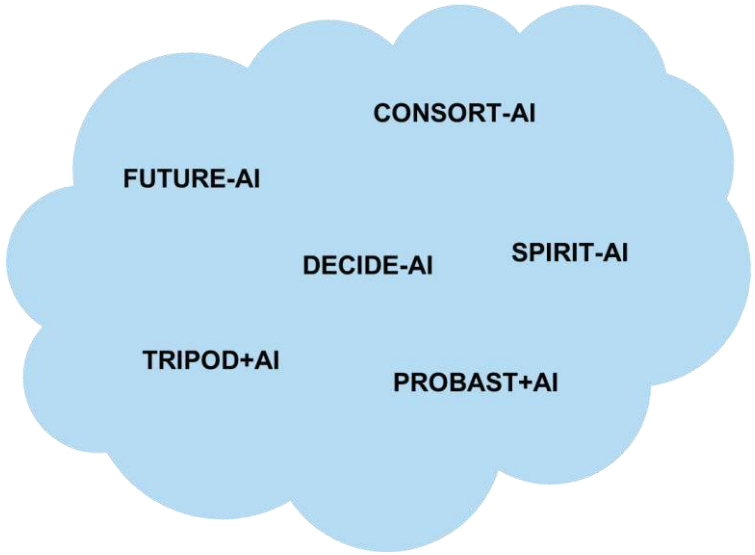
Four TEF projects
- Agri-Food: “agrifoodTEF”
- **Healthcare: “TEF-Health”**
- Manufacturing: “AI-MATTERS”
- Smart Cities & Communities: “Citcom.AI”

https://tefhealth.eu/home

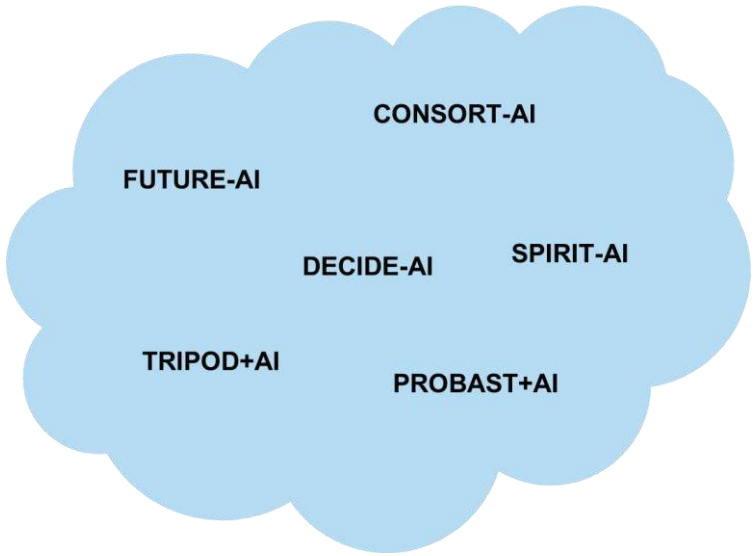


Health Specifics

Medical Device Requirments + Guidelines – **Health community initiatives**



Medical Device Requirments + Guidelines – **Health community initiatives**



RESEARCH METHODS AND REPORTING

OPEN ACCESS

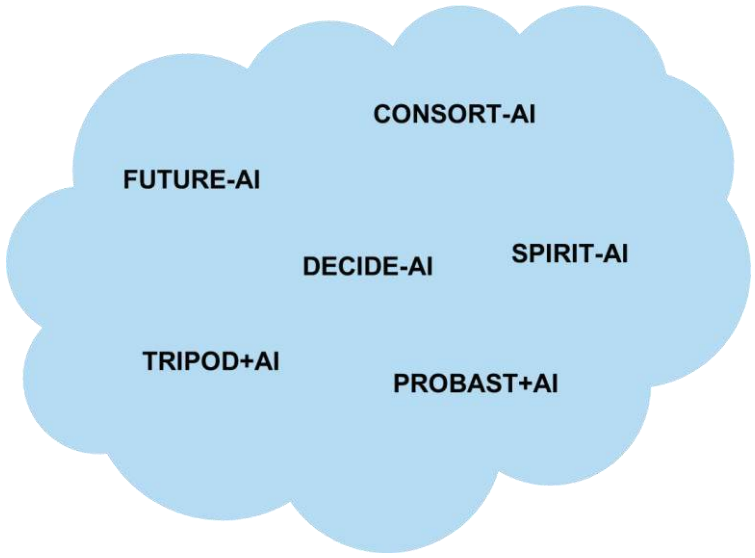
FUTURE-AI: international consensus guideline for trustworthy and deployable artificial intelligence in healthcare

Despite major advances in artificial intelligence (AI) research for healthcare, the deployment and adoption of AI technologies remain limited in clinical practice. This paper describes the FUTURE-AI framework, which provides guidance for the development and deployment of trustworthy AI tools in healthcare. The FUTURE-AI Consortium was founded in 2021 and comprises 117 interdisciplinary experts from 50 countries representing all continents, including AI scientists, clinical researchers, biomedical ethicists, and social scientists. Over a two-year period, the FUTURE-AI guideline was established through consensus based on six guiding principles—fairness, universality, traceability, usability, robustness, and explainability. To operationalise trustworthy AI in healthcare, a set of 30 best practices were defined, addressing technical, clinical, socio-technical, and legal dimensions. The recommendations cover the entire lifecycle of healthcare AI, from design, development, and validation to regulation, deployment, and monitoring.

FUTURE AI

Fairness	**Usability**	**Traceability**
Robustness	**Universality**	**Explainability**

Medical Device Requirments + Guidelines – **Health community initiatives**



RESEARCH METHODS AND REPORTING

OPEN ACCESS

FUTURE-AI: international consensus guideline for trustworthy and deployable artificial intelligence in healthcare

Despite major advances in artificial intelligence (AI) research for healthcare, the deployment and adoption of AI technologies remain limited in clinical practice. This paper describes the FUTURE-AI framework, which provides guidance for the development and deployment of trustworthy AI tools in healthcare. The FUTURE-AI Consortium was founded in 2021 and comprises 117 interdisciplinary experts from 50 countries representing all continents, including AI scientists, clinical researchers, biomedical ethicists, and social scientists. Over a two-year period, the FUTURE-AI guideline was established through consensus based on six guiding principles—fairness, universality, traceability, usability, robustness, and explainability. To operationalise trustworthy AI in healthcare, a set of 30 best practices were defined, addressing technical, clinical, socio-ethical, and legal dimensions. The recommendations cover the entire lifecycle of healthcare AI, from design, development, and validation to regulation, deployment, and monitoring.

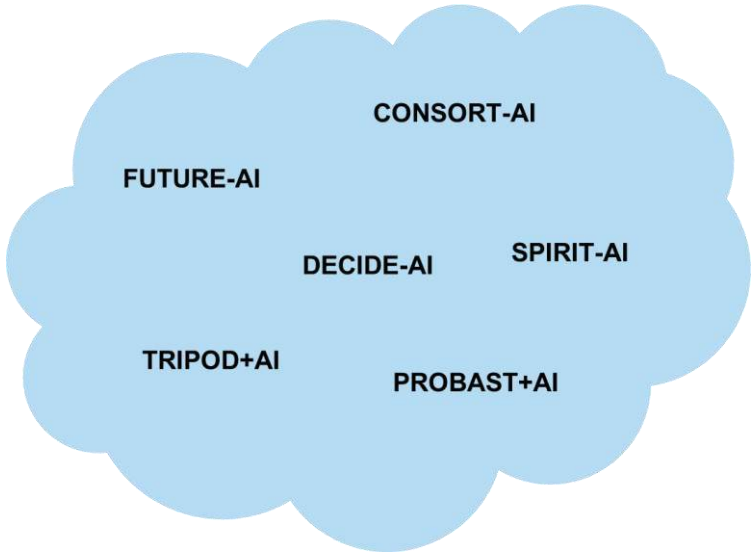
FUTURE AI

Fairness | **Usability** | **Traceability**

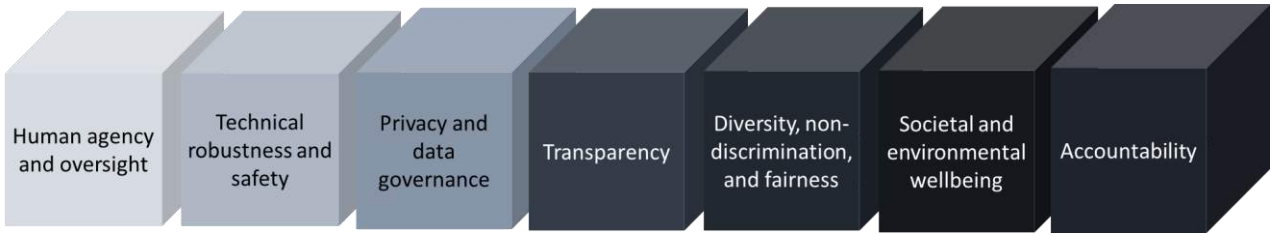
Robustness | **Universality** | **Explainability**

Example :
Depending on the intended radius of application, **medical AI tools should be as interoperable and as transferable as possible**, so they can benefit citizens and clinicians at scale.

Medical Device Requirments + Guidelines – **Health community initiatives**

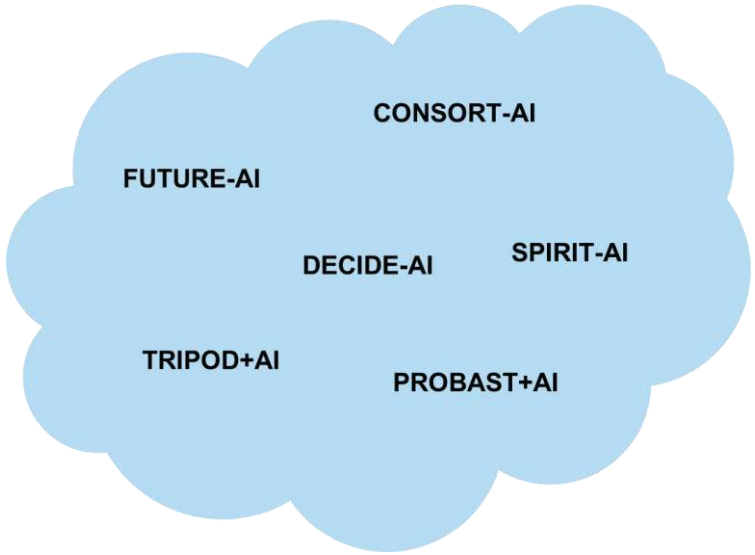


Specific to AI for health evaluation

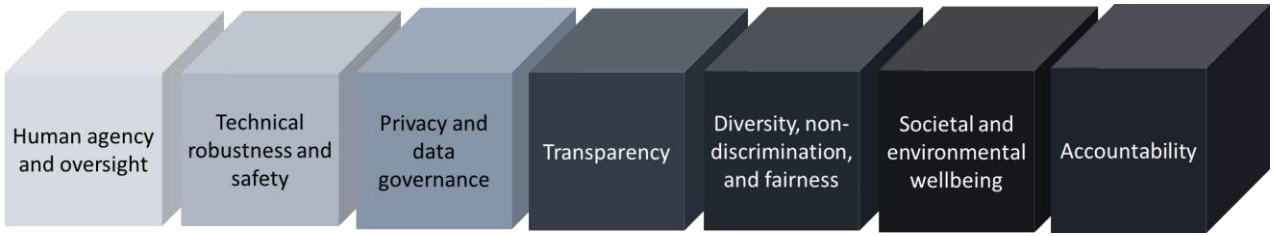


Many criteria for Trustworthy AI in health
→ concrete definitions and methods/metrics to evaluate each are still in development

Medical Device Requirments + Guidelines – **Health community initiatives**



Specific to AI for health evaluation



Many criteria for Trustworthy AI in health
→ concrete definitions and methods/metrics to evaluate each are still in development

TEF-Health Project

Objectives

- Creating a European **interdisciplinary** network to gather **complementary expertise**
- Implement **services to facilitate market access** for these technologies, taking into account regulatory requirements (certification, standardization, code of conduct, etc.)
- Testing and evaluation of AI solutions (software and hardware) in **real or realistic, large-scale environments**.
- **Support regulatory policies** for these solutions by developping concrete methodologies (certification, standards, etc.)

Consortium

GERMANY

Berlin	CHARITÉ, BPWT, VdTÜV, KI Park
Braunschweig	PTB
Erlangen	FAU
Heidelberg	EITH SI GmbH
München	TUM, FHG, EIT HEALTH EV

BELGIUM

Charleroi	ISPPC , CETIC, BIOWIN
Liège	WSL
Mol	VITO
Mons	MULTITEL
Namur	Uliamur, POLEM

FRANCE

Grenoble	CHUGA, UGA
Lyon	HCL
Paris	LNE, CEA, EIT Health Fr, HDH
Rennes	CHU RENNES
Strasboug	IHUS

ITALY

Genova	IGG, IIT
Lecce	INNOVAAL
Milano	POLIMI
Pavia	UNIPV
Povo	FBK
Roma	ISS

PORTUGAL

Coimbra	IPN, CHUC EPE
Lisboa	SPMS, InnoStars
Porto	CHSJ, HCP

CZECHIA

BRNO	MU

FINLAND

Helsinki	METROPOLIA, HELSINKI, HUS

SLOVAKIA

Bratislava	UK BA, STUBA
Martin	UNIM
Zilina	UNIZA

SWEDEN

Boras	RISE
Stockholm	KI

THE NETHERLANDS

Rotterdam	EIT Health BeNe

PAN-EU

Brussels	EBRAINS

€ 60 million (2023-2027)

Petra Ritter

Director of the Brain Simulation Section at Charité University Medicine Berlin,
Director International Affairs at Charité University Medicine Berlin
TEF-Health Lead and Coordinator

9 countries - 52 partners

CHU et IHU

Research Centre

Metrology Lab

European Healthcare Partnership Network

Univsersity

French Node

- Idea
- Research & Development
- Data collection

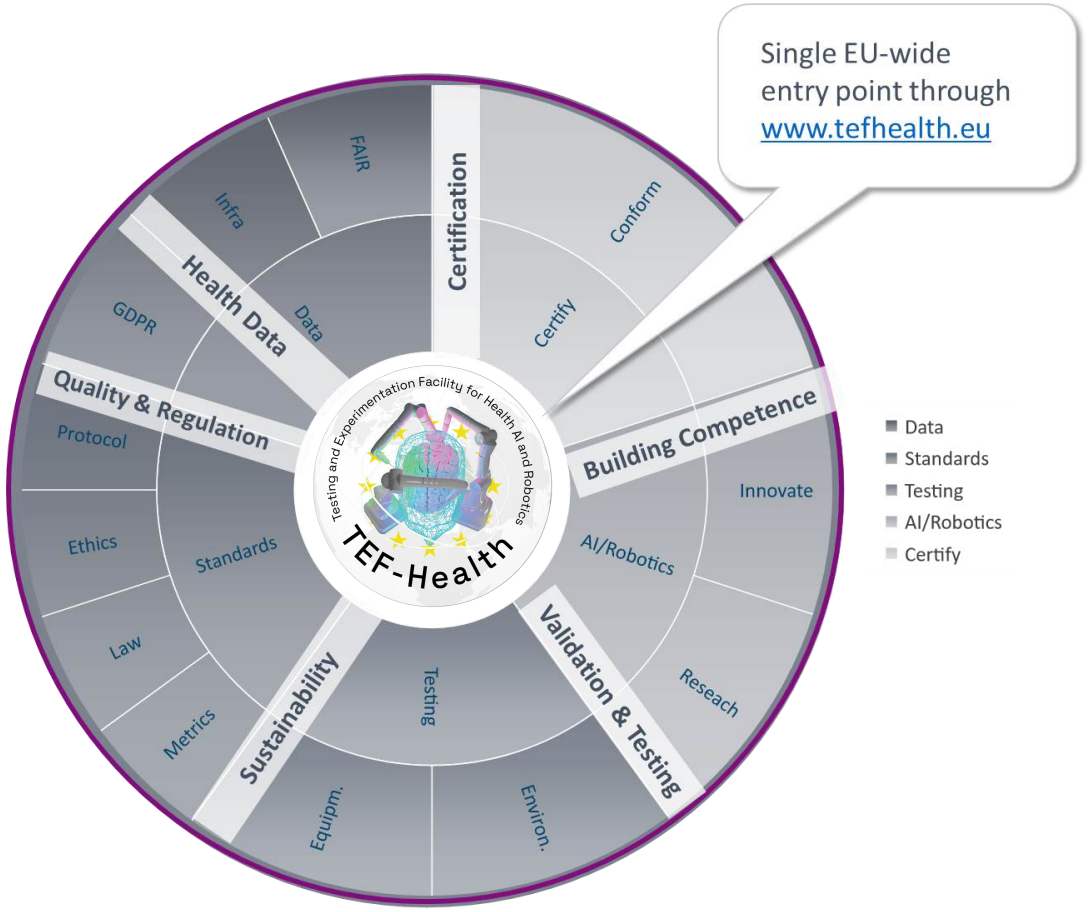
Lab

- Innovation and validation
- Testing in simulated/real environments
- CE marking

Test

- Medical Device on the market
- Solution available to patients/healthcare professionals

Bed



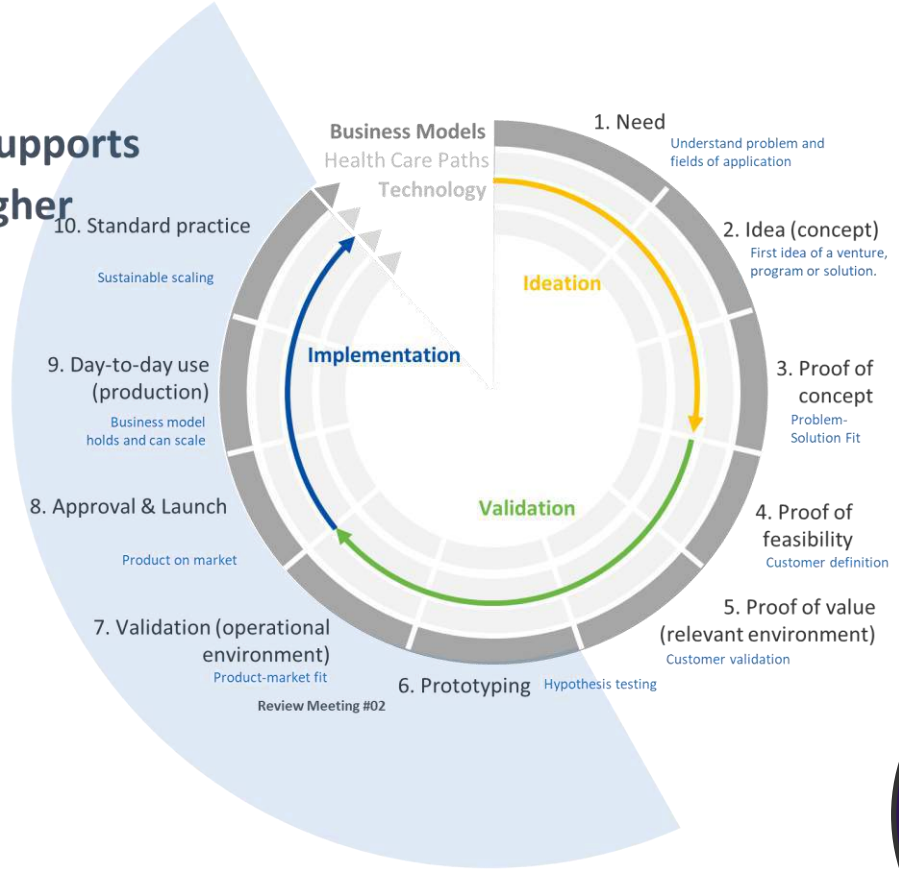
TEF-Health supports TRL 6 and higher

Technology Readiness Level (TRL)

Evaluation of AI algorithms in health

TEF-Health supports TRL 6 and higher

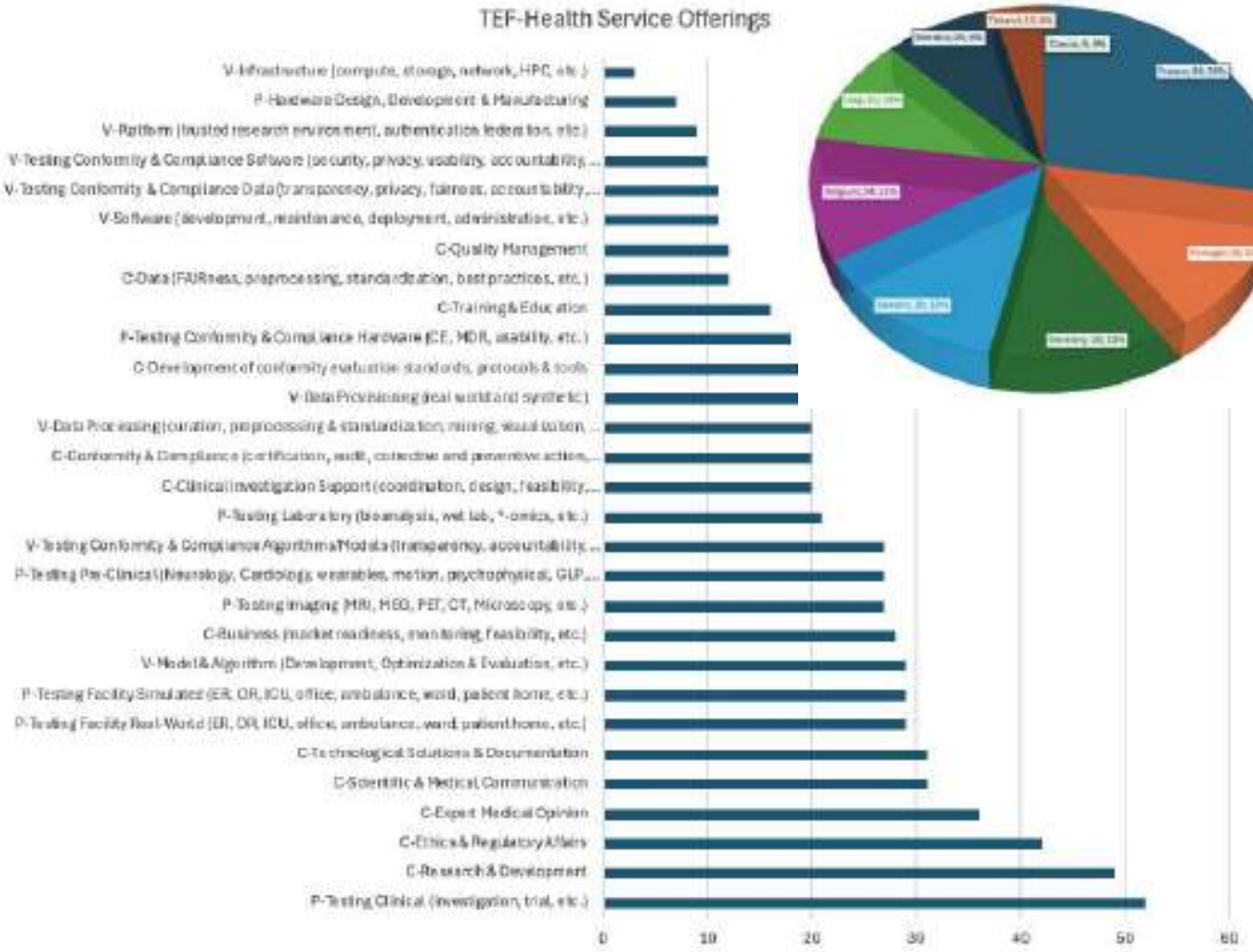
Technology Readiness Level (TRL)



Open to the whole of EU & beyond

337 Services

Evaluation of AI algorithms in health



Data Services n°5 - Provision of real-world data from clinical data warehouse and analyses

CENTRE HOSPITALIER UNIVERSITAIRE DE RENNES (CHU RENNES)

Overview

Offerings :

- CI-Expert Medical Opinion
- CI-Data Processing (curation, preprocessing & standardization, mining, visualization, anonymization, FAIR, etc.)
- CI-Data Provisioning (real-world and synthetic)

Contact : sph-health@chu-rennes.fr

Call #0

This service provides access to health data from Clinical Data Warehouses of CHU Rennes (eHOP) and CHU Grenoble-Alpes (and from other clinical data warehouse networks. It includes cohort building (patient and document selection, mapping to standard…) and data qualification (extraction of clinical variables, data cleaning, clinical consistency checking…) steps. In this way, the service offers researchers an opportunity to access healthcare data and/or descriptive statistical analyses.

This service involves only structured data extraction and pre-processing. In cases where the data is unstructured (e.g., requiring natural language processing for text, feature extraction from images, etc.), a specialized algorithm development service will be required (see our service "Algorithm development" for more details).

Prerequisite : To have data access, it is essential to have CNIL (Commission Nationale de l'Informatique et des Libertés) approval, a favorable opinion from the Scientific and Ethics Council (CSE), and a signed contract between DVE and hospital. Furthermore, our services "Feasibility study" and "Methodological advising" must have been previously conducted.

Method Description:
The service includes a standard mandatory base set. Optional tasks can be jointly selected depending on the RWE needs and available resources:

- Dataset building
- Patient selection (cohort) [mandatory]
- Selection of data of interest [mandatory]
- Mapping to PHRS/OMOP standards [optional]
- Data quantification steps
- Extraction of variables of interest [optional]
- Data cleaning [mandatory]
- Clinical consistency check by expert [mandatory]
- Anonymization/pseudonymization [optional]
- Descriptive statistical study [optional]
- Confidentiality control [mandatory]
- Matching with other databases (i.e., SNDS) [optional]
- Deployment of methods on another warehouse [optional]
- Dataset availability on Hospital platform (CI-Kit platform) [standard offer]

Data Services n°5 - Provision of real-world data from clinical data warehouse and analyses

CENTRE HOSPITALIER UNIVERSITAIRE DE RENNES (CHU RENNES)

Offerings :

- C-Expert Medical Opinion
- C-Data Processing (curation, preprocessing & standardization, mining, visualization, anonymization, FAIR, etc.)
- V-Data Provisioning (real world and synthetic)

Contact : lst-health@chu-rennes.fr

Call #5

This service provides access to health data from Clinical Data Warehouses of CHU Rennes (eHOP) and CHU Grenoble Alpes (and from other clinical data warehouse networks). It includes cohort(s) building, patient and document selection, mapping to standard (...) and data quantitative extraction of clinical variables, data cleaning, clinical consistency checking (...) steps. In this way, the service offers researchers an opportunity to access healthcare data and/or descriptive statistical analyses.

This service involves only structured data extraction and pre-processing. In cases where the data is unstructured (e.g., requiring natural language processing for text, feature extraction from images, etc.), a specialized algorithm development service will be required (see our service "Algorithm development" for more details).

Prerequisite : To have data access, it is essential to have CNIL (Commission Nationale de l'Informatique et des Libertés) approval, a favorable opinion from the Scientific and Ethics Council (CSE), and a signed contract between SME and hospital. Furthermore, our services "Feasibility study" and "Methodological advising" must have been previously conducted.

Method Description :
The service includes a standard mandatory base set. Optional tasks can be partly selected depending on the SME needs and available resources.

- Dataset building
- Patient selection (cohort) [mandatory]
- Selection of data of interest (mandatory)
- Mapping to PHRS/OMOP standards [optional]
- Data quantification steps
- Extraction of variables of interest [optional]
- Data cleaning [mandatory]
- Clinical consistency check by expert [mandatory]
- Arm (optional)segmentation (optional)
- Descriptive statistical study (optional)
- Confidentiality control (mandatory)
- Matching with other databases (i.e., SNDS) (optional)
- Deployment of methods on another warehouse (optional)
- Dataset availability on Hospital platform (CHU platform) [standard offer]

VIRTUAL SERVICE

Data Services n°4 - Ethical & Legal Advising

CENTRE HOSPITALIER UNIVERSITAIRE DE RENNES (CHU RENNES)

Offerings :

- C-Expert Medical Opinion
- C-Ethics & Regulatory Affairs (ethics board approval, legal requirements, ...)

Contact : lst-health@chu-rennes.fr

Call #2

This service focuses on providing support in analyzing the legal and ethical frameworks, which involves helping to understand the laws and regulations in force, as well as the ethical principles and values applicable for clinical data warehouse exploitation. It may involve compliance assessment, particularly with regard to RNIPH and legal compliance, which consists in ensuring that the activities in question comply with the applicable laws and rules. It may include assistance with the construction and filing of regulatory dossiers (CESREES, CNIL, etc.). It offers the opportunity to have several documents reviewed and evaluated by experts as some help for CESREES and CNIL submissions. This service is tailored by the SME needs.

Details :

- An initial consultation meeting with a legal expert of the hospital to clarify the requirements for the CESREES and CNIL submissions.
- One feedback session and an annotated report for each document (study protocol, data protection impact assessment, GDPR, etc.) that the SME needs to revise and resubmit. New submission will result in a re-evaluation of the service charges.

Prerequisite : our service "Feasibility study" must have been previously conducted.

Method Description :
Depending on the SME needs and available resources, this service may include all or part of the following task list:

- Analysis of legal and ethical frameworks (Ethical committee) [optional]
- Compliance assessment (RNIPH, legal compliance, etc.) [optional]
- Assistance and advice on building regulatory files (CESREES, CNIL, MR004) [optional]

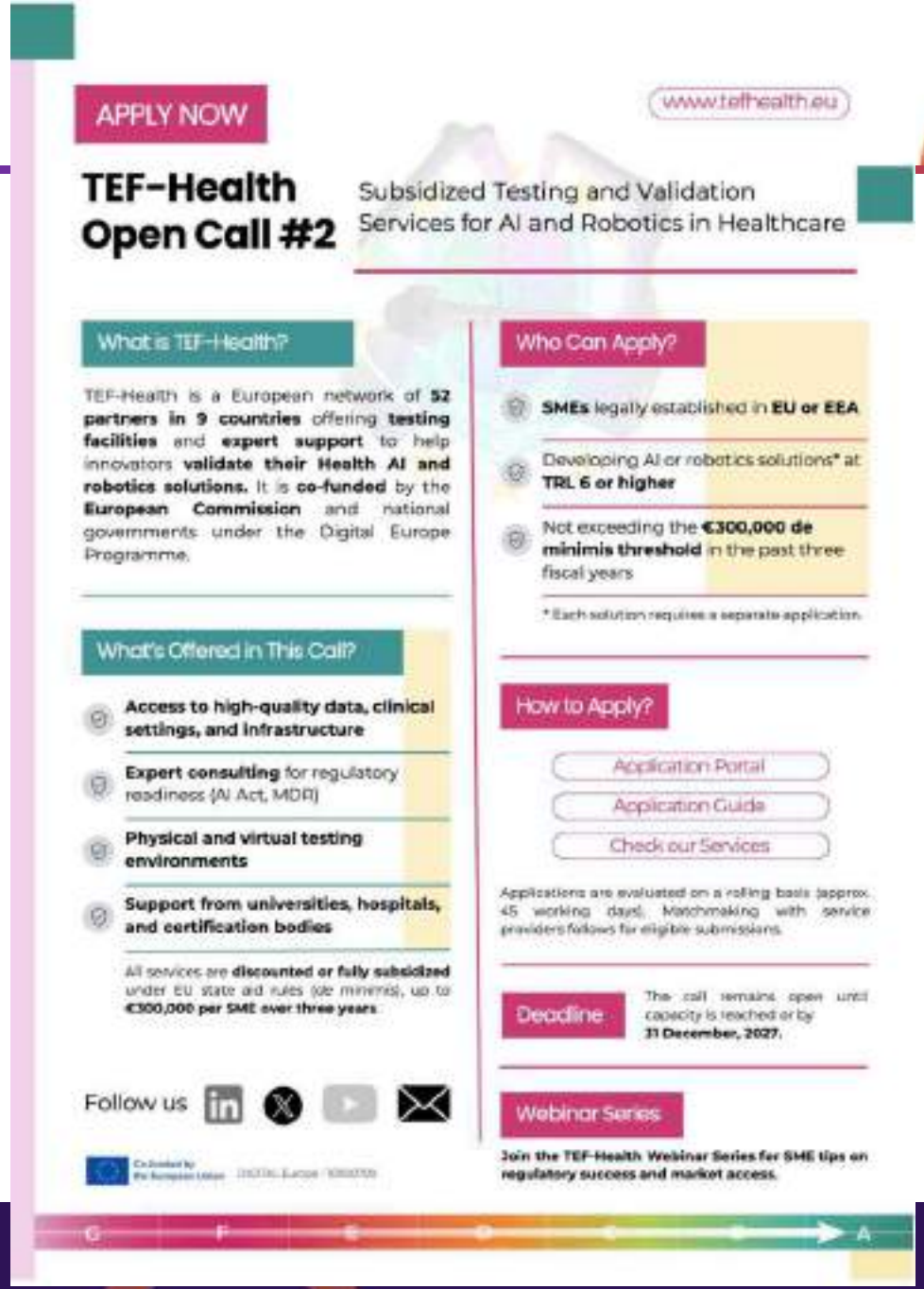
Method reference :
Local method regarding General Data Protection Regulation (GDPR).
Comité d'Expertise pour les Recherches, les Etudes et les Evaluations dans le domaine de la Santé (CEREES).
Commission Nationale de l'Informatique et des Libertés (CNIL). MR-004 reference methodology requirements and EU regulations 2017/745 and 2020/561 on medical devices.

Service standards :

- EU regulations 2017/745 and 2020/561 on medical devices
- MR-004 reference methodology (French)
- Commission Nationale de l'Informatique et des Libertés (French)
- Comité d'Expertise pour les Recherches, les Etudes et les Evaluations dans le domaine de la Santé (French)
- General Data Protection Regulation

TEF-Health Call 2

EU/EEA startups and **SME**s can apply for **price reductions on Services** offered by TEF-Health



APPLY NOW

www.tefhealth.eu

TEF-Health Open Call #2

Subsidized Testing and Validation Services for AI and Robotics in Healthcare

What is TEF-Health?

TEF-Health is a European network of **52 partners in 9 countries** offering **testing facilities** and **expert support** to help innovators **validate their Health AI and robotics solutions.** It is **co-funded** by the **European Commission** and national governments under the Digital Europe Programme.

What's Offered in This Call?

- **Access to high-quality data, clinical settings, and infrastructure**
- **Expert consulting** for regulatory readiness (AI Act, MDR)
- **Physical and virtual testing environments**
- **Support from universities, hospitals, and certification bodies**

All services are **discounted or fully subsidized** under EU state aid rules (de minimis), up to **€300,000 per SME over three years**.

Follow us

Who Can Apply?

- **SMEs** legally established in **EU or EEA**
- Developing AI or robotics solutions* at **TRL 6 or higher**
- Not exceeding the **€300,000 de minimis threshold** in the past three fiscal years

* Each solution requires a separate application.

How to Apply?

Application Portal
Application Guide
Check our Services

Applications are evaluated on a rolling basis (approx. 45 working days). Matchmaking with service providers follows for eligible submissions.

Deadline

The call remains open until capacity is reached or by **31 December, 2027.**

Webinar Series

Join the TEF-Health Webinar Series for SME tips on regulatory success and market access.

G F E D C B A

https://tefhealth.eu/call/call-2

Evaluation of AI

STILL A LOT TO BE DONE !

Illustration with Trueness

Work with Alex Poiron, PhD student - Coming paper

Trueness refers to the **closeness of agreement** between the arithmetic mean of a large number of test results and the **true or accepted reference value** [ISO 5725-1:2023]

Work with Alex Poiron, PhD student - Coming paper

Trueness refers to the **closeness of agreement** between the arithmetic mean of a large number of test results and the **true or accepted reference value** [ISO 5725-1:2023]

How to mesasure closeness of aggreement ?

How to gather the reference value ?

Work with Alex Poiron, PhD student - Coming paper

Scoping Review

Trueness refers to the **closeness of agreement** between the arithmetic mean of a large number of test results and the **true or accepted reference value** [ISO 5725-1:2023]

How to mesasure closeness of aggreement ?

How to gather the reference value ?

Binary Classification Perspective	Definitions	Examples	Checklist
	Calibration		
	a_{calib_q}: Calibration quantitative metrics	ECE, Brier Score	☐
	a_{calib_g}: Calibration graphical measures that enhance interpretability	Calibration curves	☐
	Discrimination		
a_{discg}	$a_{discg_{a,b}}$: Global discrimination metric influenced by the prevalence	MCC, CK	☐
	$a_{discg_{a,ib}}$: Global discrimination metric based solely on model outputs	AUC-ROC, BACC, Youden Index, LRP, LRN	☐
a_{disc_+}	$a_{disc_{+,b}}$: Positive-class-oriented discrimination metric influenced by the prevalence	PPV, F1, AUCPR	☐
	$a_{disc_{+,ib}}$: Positive-class-oriented discrimination metric based exclusively on model outputs	SE, TNR	☐
a_{disc_-}	$a_{disc_{-,b}}$: Negative-class-oriented discrimination metric influenced by the prevalence	NPV	☐
	$a_{disc_{-,ib}}$: Negative-class-oriented discrimination metric based solely on model performance	SP, PPR	☐

Poiron, A., Cabon, S., & Cuggia, M. (2024). How Trueness of Clinical Decision Support Systems Based on Machine Learning Is Assessed?. *Digital Health and Informatics Innovations for Sustainable Health Care Systems*, 813-817.

Work with Alex Poiron, PhD student - Coming paper

Scoping Review

Trueness refers to the **closeness of agreement** between the arithmetic mean of a large number of test results and the **true or accepted reference value** [ISO 5725-1:2023]

How to mesasure closeness of aggreement ?

How to gather the reference value ?

Binary Classification Perspective	Definitions	Examples	Checklist
	Calibration		
	m_{calib_q}: Calibration quantitative metrics	ECE, Brier Score	☐
	m_{calib_g}: Calibration graphical measures that enhance interpretability	Calibration curves	☐
	Discrimination		
m_{disc_g}	$m_{disc_{g,b}}$: Global discrimination metric influenced by the prevalence	MCC, CK	☐
	$m_{disc_{g,u}}$: Global discrimination metric based solely on model outputs	AUC-ROC, BACC, Youden Index, LRP, LRN	☐
m_{disc_+}	$m_{disc_{+,b}}$: Positive-class-oriented discrimination metric influenced by the prevalence	PPV, F1, AUCPR	☐
	$m_{disc_{+,u}}$: Positive-class-oriented discrimination metric based exclusively on model outputs	SE, TNR	☐
m_{disc_-}	$m_{disc_{-,b}}$: Negative-class-oriented discrimination metric influenced by the prevalence	NPV	☐
	$m_{disc_{-,u}}$: Negative-class-oriented discrimination metric based solely on model performance	SP, PPR	☐

Poiron, A., Cabon, S., & Cuggia, M. (2024). How Trueness of Clinical Decision Support Systems Based on Machine Learning Is Assessed?. *Digital Health and Informatics Innovations for Sustainable Health Care Systems*, 813-817.

Work with Alex Poiron, PhD student - Coming paper

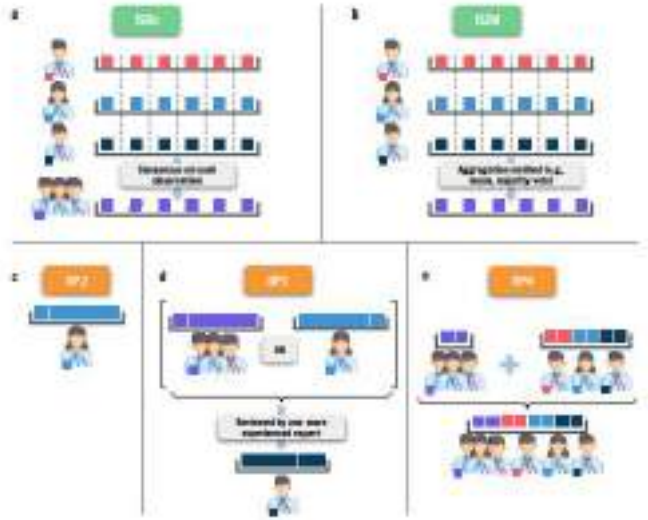
Scoping Review

Trueness refers to the **closeness of agreement** between the arithmetic mean of a large number of test results and the **true or accepted reference value** [ISO 5725-1:2023]

How to mesasure closeness of aggreement ?

Binary Classification Perspective	Definitions	Examples	Checklist
	Calibration		
	m_{calib_q}: Calibration quantitative metrics	ECE, Brier Score	☐
	m_{calib_g}: Calibration graphical measures that enhance interpretability	Calibration curves	☐
	Discrimination		
m_{disc_g}	$m_{disc_{g,b}}$: Global discrimination metric influenced by the prevalence	MCC, CK	☐
	$m_{disc_{g,u}}$: Global discrimination metric based solely on model outputs	AUC-ROC, BACC, Youden Index, LRP, LRN	☐
m_{disc_+}	$m_{disc_{+,b}}$: Positive-class-oriented discrimination metric influenced by the prevalence	PPV, F1, AUCPR	☐
	$m_{disc_{+,u}}$: Positive-class-oriented discrimination metric based exclusively on model outputs	SE, TNR	☐
m_{disc_-}	$m_{disc_{-,b}}$: Negative-class-oriented discrimination metric influenced by the prevalence	NPV	☐
	$m_{disc_{-,u}}$: Negative-class-oriented discrimination metric based solely on model performance	SP, PPR	☐

How to gather the reference value ?



Poiron, A., Cabon, S., & Cuggia, M. (2024). How Trueness of Clinical Decision Support Systems Based on Machine Learning Is Assessed?. *Digital Health and Informatics Innovations for Sustainable Health Care Systems*, 813-817.

Work with Alex Poiron, PhD student - Coming paper

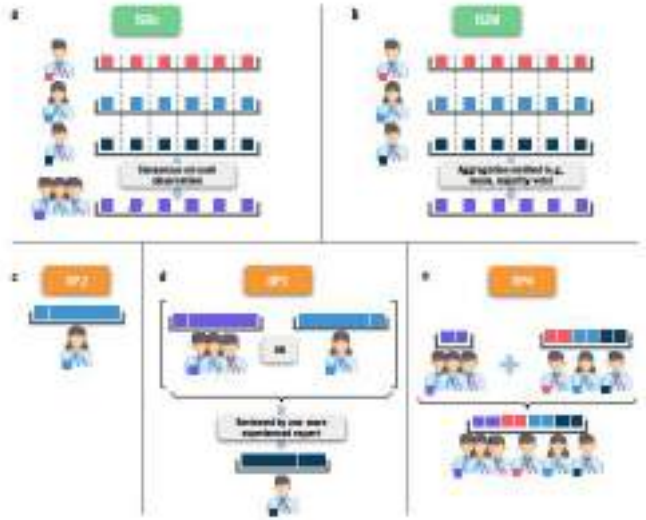
Scoping Review

Trueness refers to the **closeness of agreement** between the arithmetic mean of a large number of test results and the **true or accepted reference value** [ISO 5725-1:2023]

How to mesasure closeness of aggreement ?

Binary Classification Perspective	Definitions	Examples	Checklist
	Calibration		
	m_{calib_g}: Calibration quantitative metrics	ECE, Brier Score	☐
	m_{calib_v}: Calibration graphical measures that enhance interpretability	Calibration curves	☐
	Discrimination		
m_{discr_g}	$m_{discr_{g,i}}$: Global discrimination metric influenced by the prevalence	MCC, CK	☐
	$m_{discr_{g,u}}$: Global discrimination metric based solely on model outputs	AUC-ROC, BACC, Youden Index, LRP, LRN	☐
m_{discr_+}	$m_{discr_{+,i}}$: Positive-class-oriented discrimination metric influenced by the prevalence	PPV, F1, AUCPR	☐
	$m_{discr_{+,u}}$: Positive-class-oriented discrimination metric based exclusively on model outputs	SE, TNR	☐
m_{discr_-}	$m_{discr_{-,i}}$: Negative-class-oriented discrimination metric influenced by the prevalence	NPV	☐
	$m_{discr_{-,u}}$: Negative-class-oriented discrimination metric based solely on model performance	SP, PPR	☐

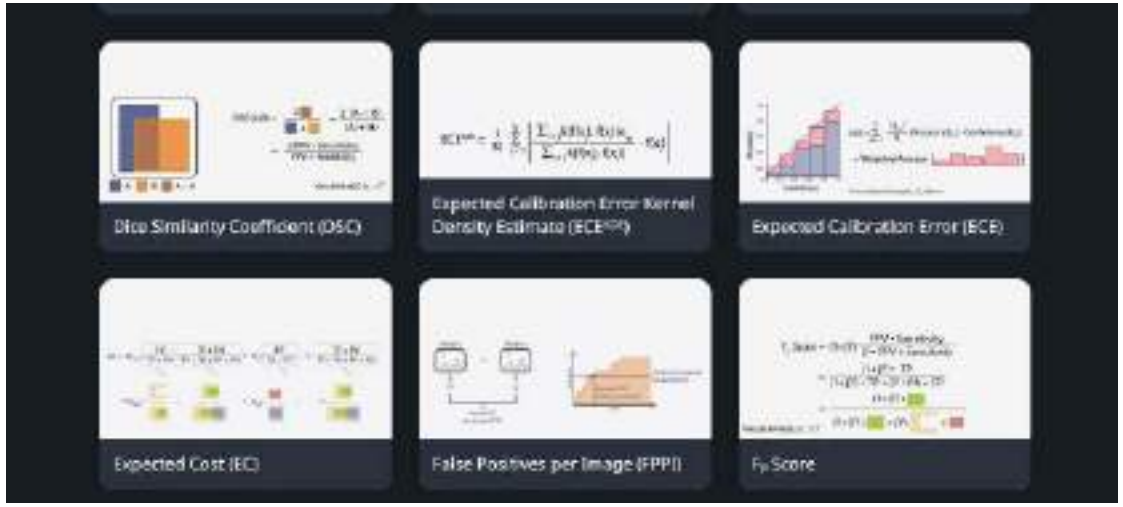
How to gather the reference value ?

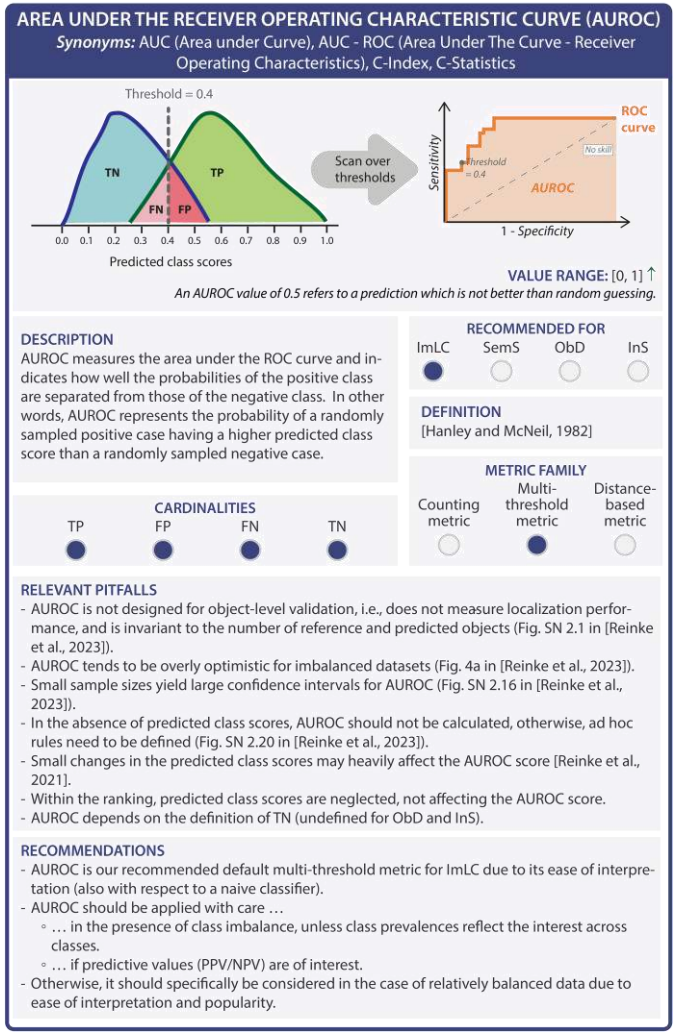


Poiron, A., Cabon, S., & Cuggia, M. (2024). How Trueness of Clinical Decision Support Systems Based on Machine Learning Is Assessed?. *Digital Health and Informatics Innovations for Sustainable Health Care Systems*, 813-817.

Performance of AI for Medical Image analysis METRICS Reloaded

Maier-Hein, L., Reinke, A., Godau, P., Tizabi, M. D., Buettner, F., Christodoulou, E., ... & Jäger, P. F. (2024). Metrics reloaded: recommendations for image analysis validation. *Nature methods*, *21*(2), 195-212.

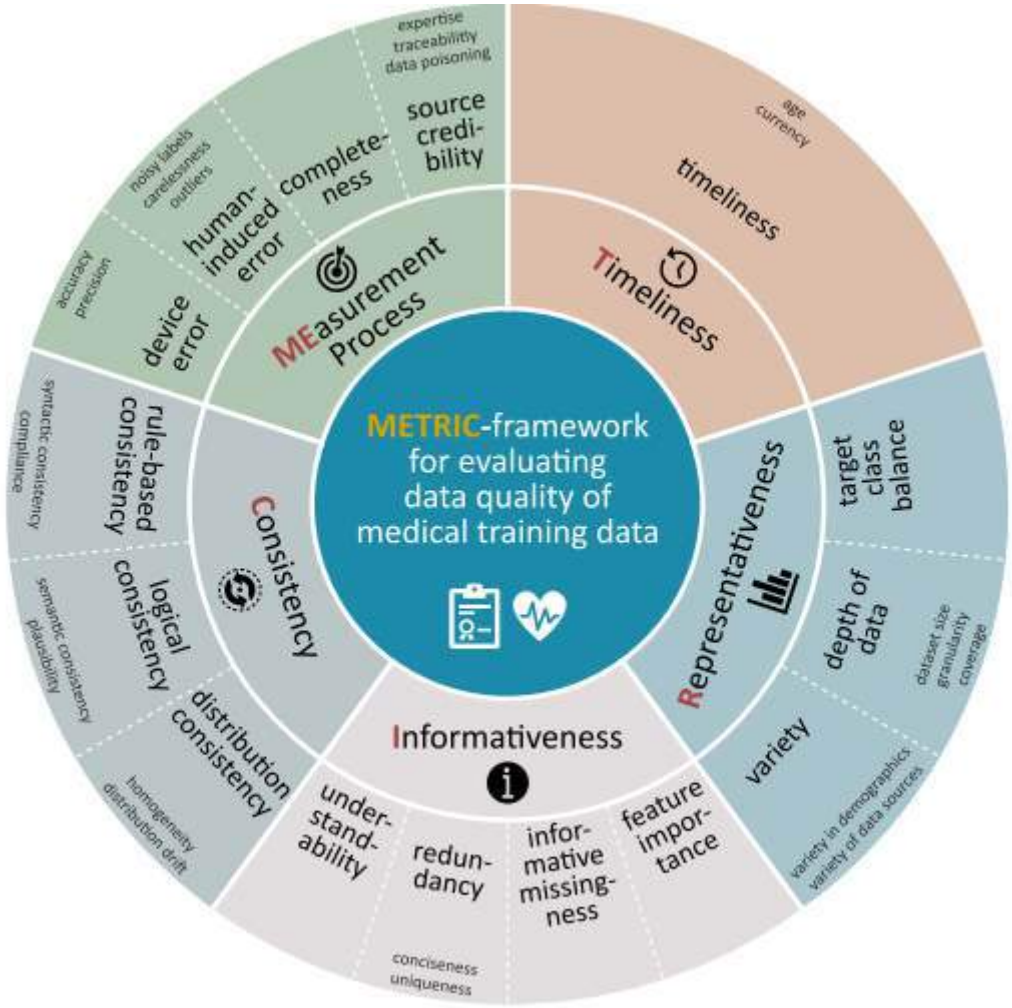




Data quality – The METRIC framework

- Quantitative
- Qualitative

Schwabe, D., Becker, K., Seyferth, M., Klaß, A., & Schaeffter, T. (2024). The METRIC-framework for assessing data quality for trustworthy AI in medicine: a systematic review. *NPJ Digital Medicine*, *7*(1), 203.



Evaluation of AI algorithms in health

THANK YOU !

Sandie Cabon - Research Engineer, PhD

LTSI – UMR 1099 Inserm - Université de Rennes

sandie.cabon@univ-rennes.fr

u Signal et de l'Image

Université de Rennes

Inserm

La science pour la santé
From science to health

CHU Rennes

Testing and Experimentation Facility for Health AI and Robotics

TEF-Health

IHE CATALYST

Evaluation of AI algorithms in health