



Evaluation of AI algorithms in health

Sandie Cabon - Research Engineer, PhD

LTSI – UMR 1099 Inserm - Université de Rennes

sandie.cabon@univ-rennes.fr















What is an AI system?



What is an Al 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 'Al Act / MDR webinar with TEF-Health', 10. Jun 2024



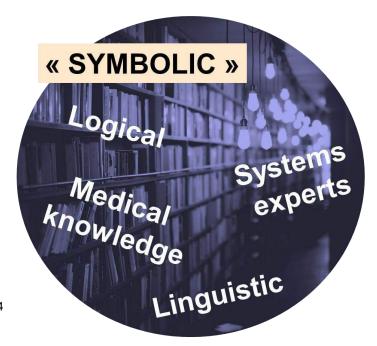
What is an Al system?

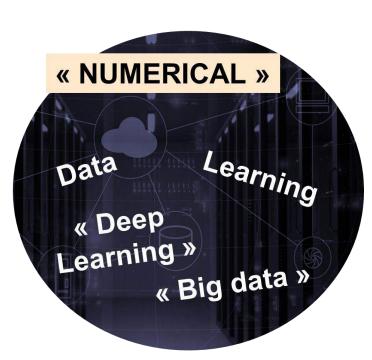
« a machine-based system designed to operate with varying levels of autonomy and, that may exhibit adaptiveness after deployment and that, for explicit or implicit objectives, **infers**, from the input it receives,

how to generate outputs

such as predictions, content, recommendations or decision that can influence physical or virtual environments »

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







What use of AI in health?



What use of Al in health?



Precision Medicine
e.g., recommending treatment



Preventive e.g., pharmacovigilance



Robots /Assistant e.g., guiding patients



Predictive e.g., disease prediction



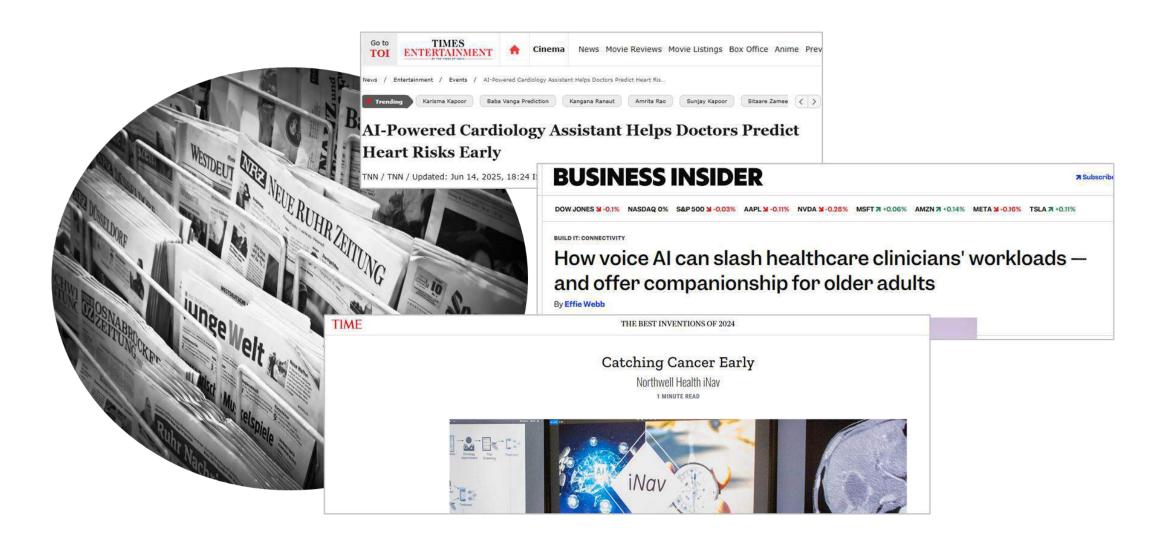
Decision Support e.g., health indicators suplpying



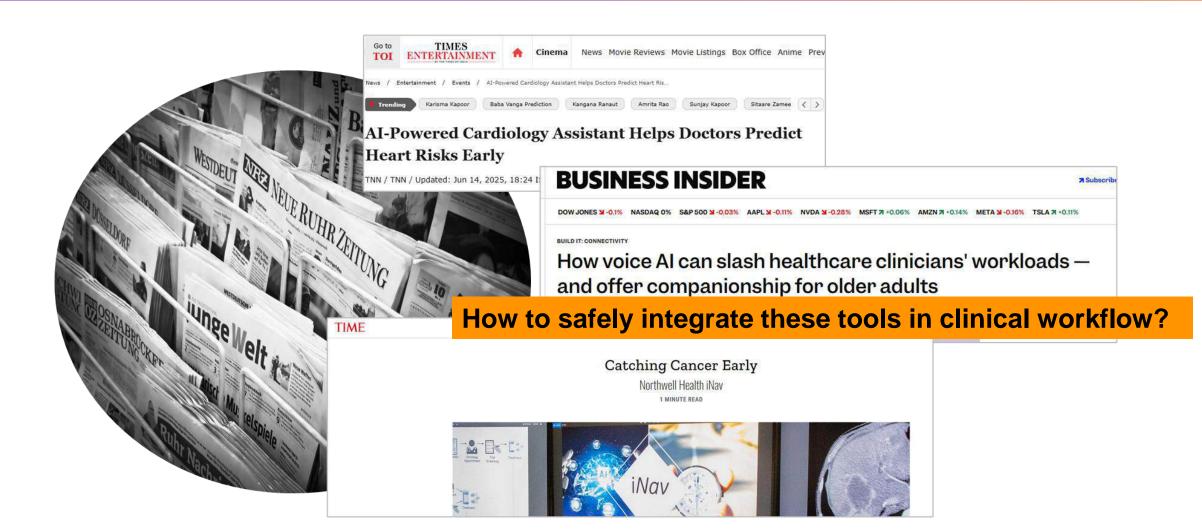
Assisted Surgery e.g., specifying zone



Evaluation of AI algorithms in health



Evaluation of AI algorithms in health







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











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



Grace period









UNACCEPTABLE RISK

e.g., social scoring

Prohibited

Risk-based approach

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



UNACCEPTABLE RISK

e.g., social scoring

Prohibited

Risk-based approach

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

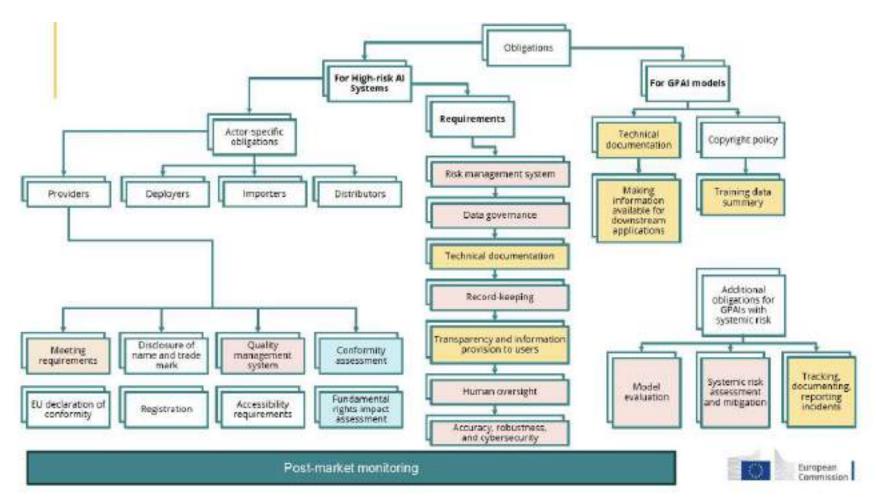
Permitted but subject to information/transparency obligations

MINIMAL OR NO RISK

Permitted with no restrictions



Requirements





Testing and Experimentation Facilities (TEFs)

Al-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.Al"

Higher education Institutes
Basic research

Healthcare Applied research

Industry
Small and medium-sized enterprises (SME:s)
Research and development
Marketing and sales

Mand Robotics

https://tefhealth.eu/home



Health Specifics



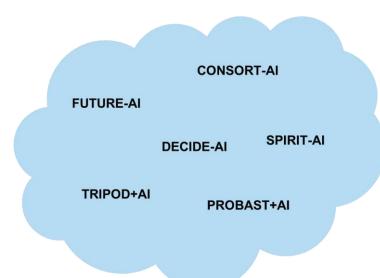


CONSORT-AI
FUTURE-AI

DECIDE-AI SPIRIT-AI

TRIPOD+AI
PROBAST+AI





RESEARCH METHODS AND EXPORTING

COMMON PROPERTY OF STREET AND EXPORTED AND EXPORTING

COMMON PROPERTY OF STREET AND EXP

Earth Lebate, "Abjund of Frange," Andonse & Portia, "Eart Glaciae," Colin Cestes,
Larin Prangiae, "Dura Wecker, "Estern W. Assaulte og, ""Prod. Price," Guy S. Callins, "
Indonesia, "General Patrice," Service Basen, "General Cestes, "Colon Cestes, "
Indonesia, "General Patrice," "Basen Schoole, "Basen Schoole, Graine,"
Indonesia, "Arrespong, "General Esternate," Less Moise Heles, "Lesson Cestes Millionico, II
Larin Blace From matt," M. Junge Landman, "Marcia Cestes Cestes," Robert Schoole, "
Marcia Cestes, "Marcia Basen, "Marcia Cestes Cestes," Notice Cassaulte,
Marcia General Legistra, "Marcia K. Tolouga, "Marcia Cestes Cestes, "Marcia Cestes Ind.
Marcia General Legistra, "Marcia Cestes, "Marcia Cestes Cestes, "Marcia Cestes Ind.
Marcia General Legistra, "Marcia Verena, "Martin De Bronn, "Valent G. Robert General,"
Welson Eccosed, "Mora on Cassa Procedy," "Bistone Cestes Cestes, "Street Legistra, "
Osol Parti. Richard Deadle, "Samp Majort," Sam Colon Cestes, "Sirvet Legistra, "Sam Procedy A Roger, "Zohad Sola Raddin," Marcia P. Nation P. Nationare."

See Institution of the Publish of Compensation.

and deployable artificial intelligence in healthcare

Ex-communications are entirely as colors. Communication in the colors. Communication in colors (communication) and colors (communication) and colors (communication) and colors (communication). Colors (colors) and colors. Colors (colors). Colors (colors). Colors (colors).

Asserted billion on 2005.

Despite mojor advances in artificial. intelligence (All rewards for healthcare, the deployment and adoption of All sechnologies remain British is clivical practice. This paper describes the FUTURE-ALframework, which provides guidance for the development and deployment of Inistroctly Allocis is nearthcare. The FUTURE At Coreartum was trunded in 2021 and comprises 117 intendisciplinary experts from 50 countries representing all continonts. including Al scientists, clinical researchers, biomedicatethinsts, and social scientists. Over a two year period, the FUTURE All autifeline was

SEMMARY POINTS

September States of the Both architecture with the september of the septem

Wa PSRAID At Excess to be contact 1.07 papers for in 16 came and in Addisnt to make and punkt are the interpretation of the

The PATER AT A model to concern of story of copyring principles of the conceiver with the middle production of the control of the conference of the control of the concern of the control of the

and registration by a plant on plantacy, the PCA date value of this registration (Continuous Anna Section of the production of the Anna Section (Bibliotechillers, 248 remaining proliferages sharing the Bibliotechillers, 248 remaining the machine of the Bibliotechillers, 248 remaining proliferages sharing the Bibliotechillers, 248 remaining proliferages as the Bibliotechillers of the Bi estatished through contens is based on six galding principles—Litimess, amerisativ, testessfully, assistiv, inchasment, and explainability. To operationalise trustworthy AL in healthcare, a set of 30 best practices were defined, addressing technical, distipations, socioetical, and legal dimensions. The recommend attors cover the entire lifecycle of healthcare AL, from design, development, and solidation to regulation, deployment, and monitorities.

Market See

is in-Cettod toothood, a minut stocking-on MA- that is, it professor with the adding to will intern loops continue based on a membrang work is develop congrue attack toolog, for resultable the end of congruent and toolog, for resultable the end of congruent and toolog, for resultable the end of control, and primary international control, and primary international control, and primary international control and only international control of control and religious of MA international consumers and independent of international control between testing and or international control and ordered and control and ordered and control a

Common stop care throughout the resemble of the A least are remote and assemble by protein communities, bearing septiments, and activate by protein communities, bearing the appropriate and the septiment of the communities of the communities of the communities and the communities are the formation of the protein the communities of the

FUTURE AI

Usability

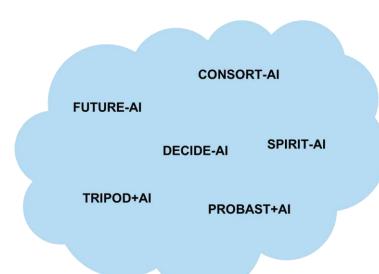
Traceability

Robustness Universality

Fairness

Explainability







behavior from

and monitoring.

Inistworthy Al tools in neuthcare. The

intendisciplinary experts from 50

including Al scientists, clinical researchers, biomedicatethinsts, and

social scientists. Over a two year

countries representing all continents.

period, the FUTURE All autifeline was

2021 and comprises 117

State was sharen a construction or proper to secretarity or

No PRINTED DESCRIPTION AND ADDRESS OF THE PROPERTY OF THE PARTY OF THE

everyally introduct, violating reliances, and replacefully in

The Port Collect Not remark to a control or board programing pathological fill trans-

The guided has a debassed the water at thoughts, here the speciest developes on

mark their out the larger to some dispersional transport and and

The framework fits have, to complete more paragraph of the marriery receives to

arthur and construct the contract of the back of the back of the same of the s

and replicable timps beni-ten planning re-mediate piles arbitrative

STREET OR ALL PROPERTY OF THE PROPERTY OF THE

Editoral Division and Developing (Colleges) through the Mathematic

pring that of averaging if a paly have mention, the freezing among the brillian in call and

FUTURE At Coreartum was trunded in

is included the phone, and control to be given that in the is, algorithms with the study to will been logic-seel. interpretations have been investigate and to streeting compare what podds, for explain chierre furnishproposity and those of these proposers control and palled into the believe, "Decade and a detailed, the declaration and adaption of All productions oncombright in and whill stands produce to anner your, economist and Sent comes about the sectional, visional, effect), and recent tests account with healthrate at." In paliticality, expense, income 5 has observe that he twin to livelife on use by pose in cross and palicy. have, blaset and increased bright invacables, both of SIMPSHAY SHE ISSUED BUILDING THE SHEET AND SHEET again sociating boyon has,"

plintensions. The recommendations

At, Fron design, development, and

visibilistics to regulation, deployment,

To move stopomic throughout the country that Al. their are received and undepend. by processes carriers, beet agreement, our scharmes Rowers, they is an elsenter of their, widely eccepted paracises on low buildcare AI both should be temport, developed, reduced, and depleted to be more order, deal to, referensity where, alteredity side.



medical Al tools should be as interoperable and

as transferable as possible, so they can benefit

citizens and clinicians at scale.

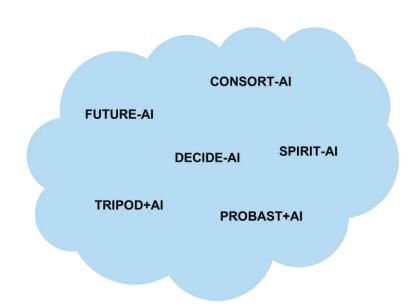


as the objected in the empirisate for solid strategic and a freedom as a few fronts.

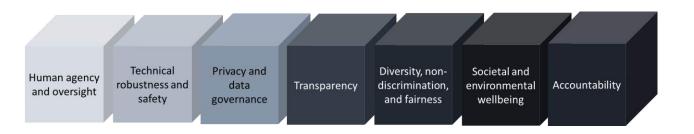
at the wall beday purplet ment that it is become by treatment of the

SEMANNY POINTS





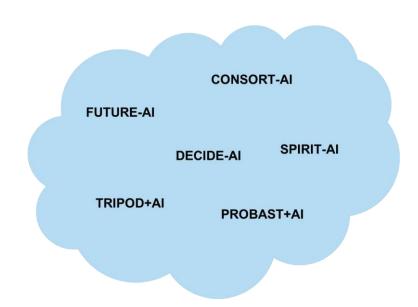
Specific to AI for health evaluation



Many criteria for Trustworthy AI in health

→ concrete definitions and methods/metrics to
evaluate each are still in development





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

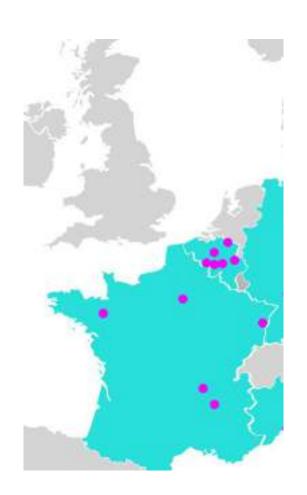


€ 60 million (2023-2027)

24/06/25

Petra Ritter























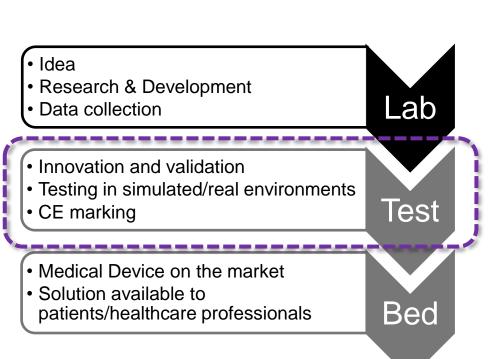
European Healthcare Partnership Network

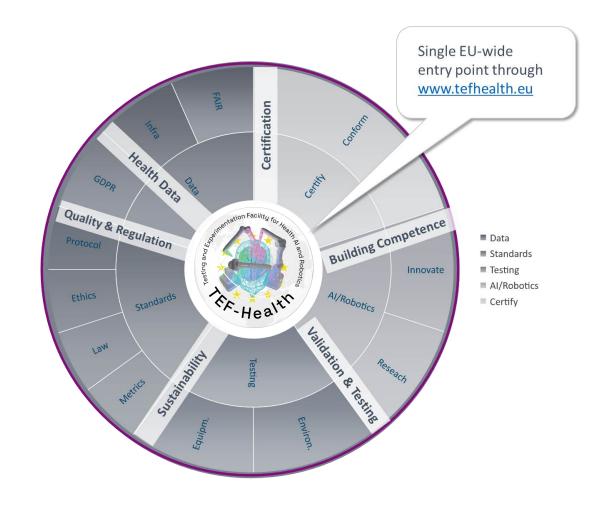




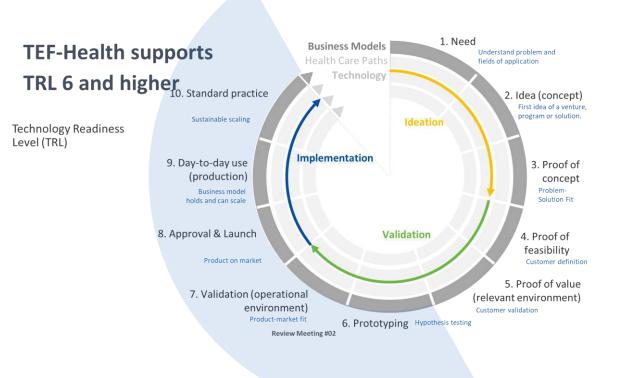
French Node



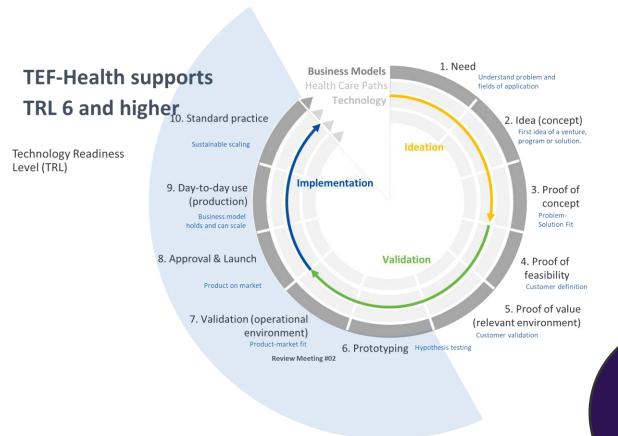


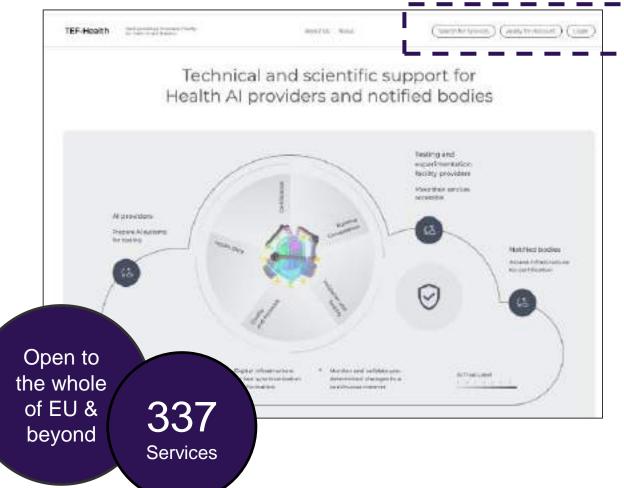






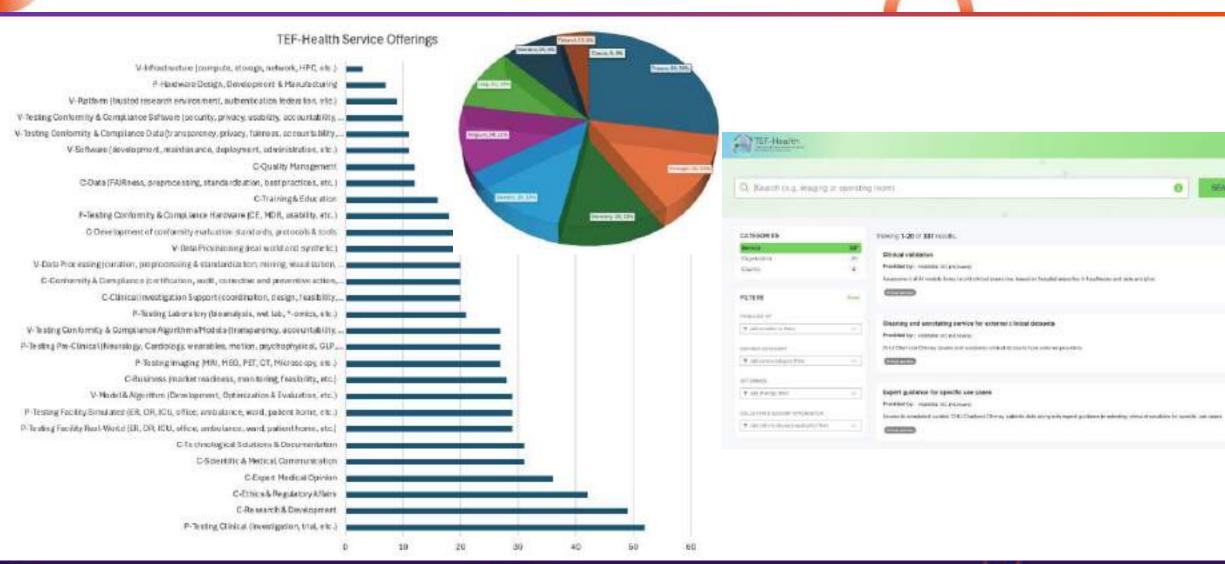








Evaluation of AI algorithms in health







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

CENTRE HOSPITALIER UNIVERSITATIE DE RESNES JOHN REMIES.

Cost types

Otherings

- C-Expert Medical Dipriori
- + 3-Octa Processing (number, progressions); 6 standardoppin, resing, visualization, prorperations. Failt, etc.)
- * 'v Opta Provisioning mod world and synthetic)

Contact: (university or a removal)

Dat you

This pervise provides access to house some Servici Data Warnhauser of CHU Revises of and CHU Greecopic Ages and from other distance state assertionable refunda. It includes disclarate including parties and disclarant selection, imaging to standard, if see data qualification perfection of distant variables, data classing disclar provides by tracking. I stage to this way, the service offers our space or our apportunity to excess testingly of the action description belongs and action.

The server Weeke only shuttened data exhaden and pre-presenting. In cases when the data is analyzered by g., reporting national language passesting for less, logical exhausts from mages, etc.), is ignitatived argostom development activity and be reported passive server. "Agentim consequenced" for most constall.

Presignation: To have done access, it is executed to have CNV, (Correlation Staturals on Enforcedage or the Literature agreement of treatment operate from the Specific and Chica Council (CEC), and a signed content between DVC and hospital Evafuences, our services "Facility's study" and "Mathodological advising" must have been provided by conducted.

Wethol Description:

The service variables a remaind numbality some on. Dykonal labor our be purity weeklet depending on the SME nords and our balance encorage.

- * Determination time.
- Patient spector (solor) (remistery)
- Semident of both of whitest (Fundatory)
- 3 Remirgo attrabular 9 CWOVBP1 of grapa 51 x
- · Data gradfooter; stays
- . Extraction of variation of intrinot jophicson
- . Date risesing (marchiny)
- . Clinical constituting shade by expert (merentural)
- . Ann protromagewritizon (riphonal)
- Decorptive statistical etials (opecing).
- * Cantidemotic control prandatory)
- · Making with other delabases (i.e., SNDD) (opnored
- Deproprient of mothods on another waveboune topform()
- Distance of availability on Hospital pleasant (CHL) plantama (Standard offer)



Evaluation of AI algorithms in health

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

CENTRE PICKRITALIER UNIVERSITATIE DE RESINEULCHU REMIES.

Cost types

Otherings:

- Cleapert Medical Dipriori
- + /E-Outa Processing (number, proprocessing & standardization, release, variableston, processings, Failt, etc.)
- * 'v Opta Provisioning rest world and synthetic)

Contact : to how the constitution of

Dat 40.

This serves provides access to health data from Clerical Data Warehauser of CHU Rentes calcard CHU Greecos Ages and from other distinguishes securities a requirement reduction of control or or individual control or of children detailed or control or or other provided by the way, the sender effect data quadratic or or other or other or or other

The server Weeke only shuttened data exhaden and pre-presenting. In cases when the data is analyzered by g., reporting national language passesting for less, logical exhausts from mages, etc.), is ignitatived argostom development activity and be reported passive server. "Agentim consequenced" for most constall.

Presignation: To have during example in the second of the second state of the second s

Wethol Description:

The service variables a remaind numbality some on. Dykonal labor our be purity weeklet depending on the SME nords and our balance encorage.

- * Debanart basing
- Patient spector (solor) (remistery)
- Semiction of stoke of printed (Franciscopy)
- Family & strainted COVOVERS of graph 8 x
- * Date proffoctor; staps
- . Extractor of variation of restrict (optional)
- · Date 2 leasing (margistry)
- . Clinical constituting streets by expert (managery).
- Ann pozionisegmentation (optional)
- + Decorptive statistical etials (opening)
- Contributionally control treastlessey)
- · Making with other delabases (i.e., SNDD) (optional)
- . Deployment of northods on enother wavefround toplomati
- Diatement availability on Hospital plettom (CHL plantom), [Standard offer]

VIRTUAL SERVICE

Data Services n°4 - Ethical & Legal Advising

CONTRE HOSPITAL EN UNIVERSITAINE DE PERRES (CHU PERRES)

Zimytowi

Offerings.

- E. Eligsett Medical Opinion
- C-Edecy & Regulatory effects outrics board approval, legal regularments.

Certain inforeming discernes in

Call 10

This arrives focuses on providing support in analyzing the legal and ethics financiars, which involves helping to understand the leave and regulations in total acceptance of the ethical procedure and values applicable for clinical data economics employed in they involve compliance assumement, participantly with region to 1997 the 1997 and legal compliance, which constitute in courting that the undertake in question comply with the applicable level and time. If may include assistance with the constructor and time of regulatory doesers of CHEES, SHE, etc.; it offers the oppositionly to have several documents movement and nyalustance by modern as some negligible. CERFIES and CHE. Apprisons. The apprisons is believed that the several.

Debalat

- All initial consultation making with a logal expert of the hospital to clarify the requirements for the GESREES and CRUL extrements.
- One feedback security and an emphasion paper for each observer integraphesis, data protection report assummers GOVP, etc.) that the SNE meets to be one excluded. New submission will make no evaluation of the service sharpes.

Prorequinite, and pervise "Feasifility blindy" (Mail have bern provincily samillables).

Mirthed Description

Departuring on the SAE made and available resources. How earlier may include all or part of the following hast list.

- . Analysis of legs and ethical its nework (Ethical committee) (opportur)
- Compliance passedment (RMIPH, legal compliance, etc.) (potential)
- Accordance and a sives on mulding regulatory flow (CESPEE I. CNL., NFKK4 (sprinted)

Muthod wisness:

Local method respecting the erol Data Protection Regulation (COPR)

Constit (Perpettive goes les Ricchardness les Etudos et les Evolugions doint le domaine de la Santé (CEREES).

Commission Records do Proformatique et des Litertes (CNL), MR-604 relations methodology requirements and EU regulations 2817/745 and 2020/861 on medical devices

Service standards:

- EU regulateria 2017/745 una 2020/901 un medical devisea
- * MFG04 enteriorise institudology (French)
- · Controppe Nationals de Ceformitique et des Liberille (Franch)
- Contrib d'Expertise pour les Reubershes, les Ehutes et les Coductions dans le domaine de la Santé (French).
- Descrial Data Priotection Programas





TEF-Health Call 2

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

APPLY NOW

www.telhealth.eu

TEF-Health Open Call #2

Subsidized Testing and Validation Services for Al 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 (ALAct, 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 All 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), Marchmaking 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 Ups on regulatory success and market access. https://tefhealth.eu/call/call-2







Evaluation of AI STILL A LOT TO BE DONE! Illustration with Trueness



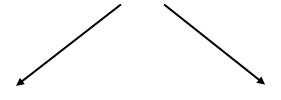


Trueness refers to the closeness of agreement between the arithmetic mean of a large number of test results and the true or accepted reference value [ISO 5725-1:2023]





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?

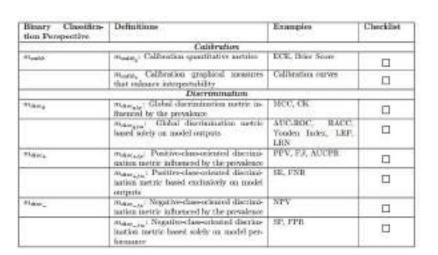
How to gather the reference value?





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?



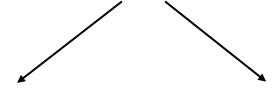
How to gather the reference value?

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.



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 Chestites- tion Perspective	Definitions	Examples	Checklist	1
	Calibration	Secretary and the		
Wipepide.	mode _a : Collimation quantitative metrics	ECE, Dier Score		
	more, Calibration graphical measures that inhorer interportubility	Calibration ourses		
	Discrimination	MARKET CHILD	-	
There,	manuale: Global discrimination metric in- fination) by the prevalence	Mcc, ck		Metrics sets in Binary
	m _{inimpo} . Global distributation metric based salely on model corputs	Youden Index, LRF, LRN		Classification
Marc.	make, process of the providence of the providenc	PPV, F3, AUCPB		
	m _{der_t/n} : Puritie-class-oriented discinni- iation metric based exclusively on model origins	SE, PNR		
Hav-	m _{d.m_ts} . Negative-class-oriented discrimi- nation metric influenced by the prevalence	NPV		
	m _{thm_tw} . Negative-class-oriosted district- ization metric based solely on model per-	SP, FPR		
	formany.			9.8% provide a
				comprehensive evaluation

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.

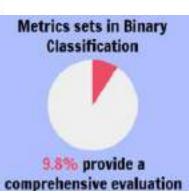


Trueness refers to the closeness of agreement between the arithmetic mean of a large number of test results and the true or accepted reference value [ISO 5725-1:2023]

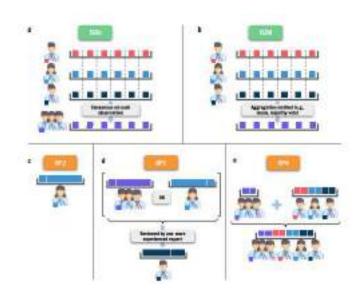


How to mesasure closeness of aggreement?

Binary Choothers- tion Perspective	Definitions	Examples	Checklist
	Calibration	CONTRACTOR CONTRACT	100
Wileschile.	mode _a : Calibration quantitative metrics ECE, Drive Sc		
4 0	m _{orth} , Calibration graphical measures Calibration curves that infrarer interportubility		
	Discrimination	000000000000000000000000000000000000000	-
**Anna	manual by the prevalence	Mcc, ck	
	m _{derger} : Oblas discrimination metric based safely on model outputs	AUC.ROC. RACC. Youden Index, LRF, LRN	
Marc.	military. Postavo-closs-oriented discrimi- uation metric influenced by the providence		
	Manusta: Parities-class-oriented discrimi- lation metric based exclusively on model outputs		
Hau_	m _{disc_ts} . Negative-class-oriented discrimi- nation metric influenced by the prevalence	NPV	
	m _{dec_ra} : Negative-bas-urkaind distribu- ination metric based solely on model per- ferences:	SP, PPR	



How to gather the reference value?



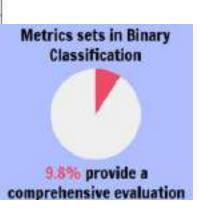
Poiron, A., Cabon, S., & Cuggia, M. (2024). How Trueness of Clinical Decision Support Systems Based on Machine Learning Is Assessed?. *Digital Health and Informatics Innovations for Sustainable Health Care Systems*, 813-817.



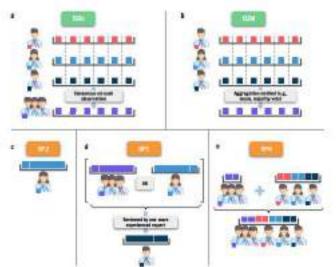
Trueness refers to the closeness of agreement between the arithmetic mean of a large number of test results and the true or accepted reference value [ISO 5725-1:2023]

How to mesasure closeness of aggreement?

Binary Choothers- tion Perspective	Definitions	Examples	Checklist
	Calibration	CONTRACTOR CONTRACT	10.
Winglib.	mode _a : Calibration quantitative metrics	ECE, Bine Score	
4 0	more, Calibration graphical measures that enhance interpretability	Calibration curves	
	Discrimination	100 CE 0 CUIT	-
**Anna	manual by the prevalence	Mcc, ck	
	$m_{\rm cheque}$. Obliad discrimination metric based safely on model outputs	AUC.ROC. RACC. Youden Index, LRF, LRN	
Marca.	man, Postive-class-microtel discrimi- nation metric influenced by the providence	PPV, P3, ACCEB	
	m _{der_t/m} : Purities-class-oriented discrimi- lation metric based exclusively on model outputs	SE, PNR	
Hau_	man_in. Negative-class-oriented discrimi- nation metric influenced by the prevalence	NPV	
	man_p Negative-bas-urlanted distribu- ization metric based solely on model per- formance	SP, PPR	



How to gather the reference value?



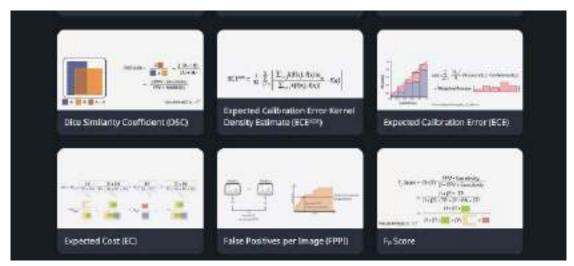


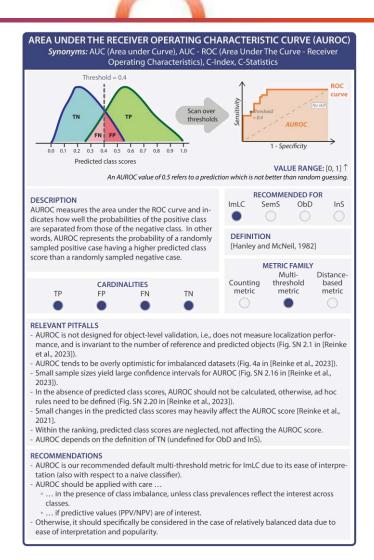
Poiron, A., Cabon, S., & Cuggia, M. (2024). How Trueness of Clinical Decision Support Systems Based on Machine Learning Is Assessed?. *Digital Health and Informatics Innovations for Sustainable Health Care Systems*, 813-817.



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.



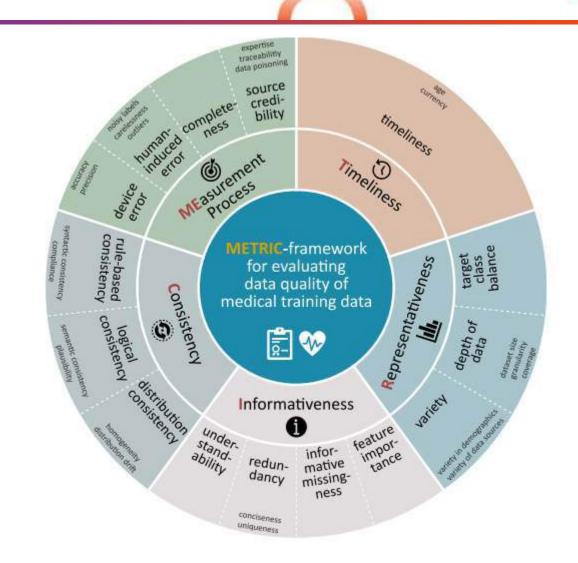




Data quality – The METRIC framework

- 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

u Signal et de l'Image

Sandie Cabon - Research Engineer, PhD

LTSI – UMR 1099 Inserm - Université de Rennes

sandie.cabon@univ-rennes.fr

VIENNA 24-25 THANK YOU!













Evaluation of AI algorithms in health

