FUTURE-AI: International Guidelines for Trustworthy & Deployable AI in Healthcare

Karim Lekadir
University of Barcelona
Artificial Intelligence in Medicine Lab (BCN-AIM)

Part 1-

Do We Need Guidelines for Trustworthy Medical AI?

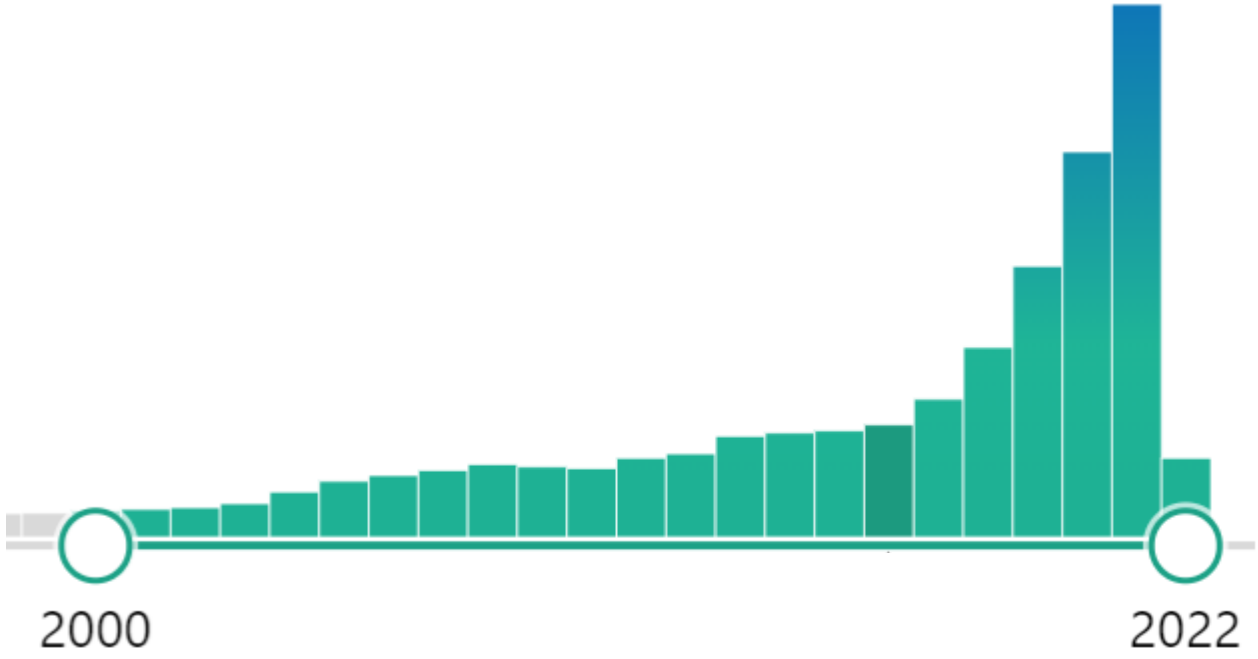
Medical AI in Research

Artificial intelligence papers in biomedical research

Source: PubMed (2000 – 2021)

2016: **8,551** papers

2021: **44,027** papers



Medical AI in Media

BBC | Sign in | Home | News | Sport | Reel | Worklife | Travel

NEWS

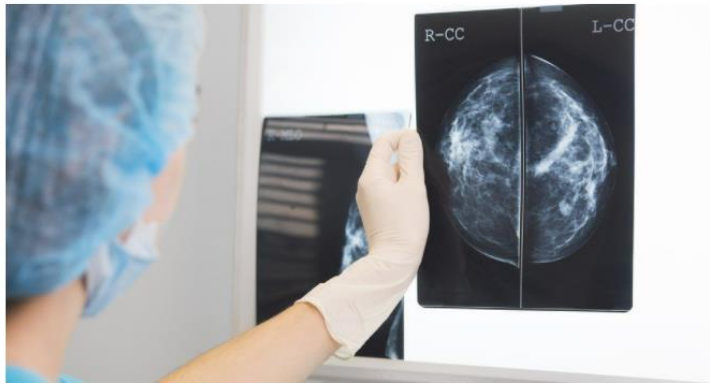
Home | Coronavirus | Climate | Video | World | UK | Business | Tech | Science | Stories | Entertainment & Arts

Health | Coronavirus

AI 'outperforms' doctors diagnosing breast cancer

Fergus Walsh
Medical correspondent
@BBCFergusWalsh

2 January 2020



The New York Times

Pandemic › | Covid-19 Updates | Coronavirus Map and Cases | Track State Mask Mandates | Analysis

A.I. Versus the Coronavirus

A new consortium of top scientists will be able to use some of the world's most advanced supercomputers to look for solutions.

News | Opinion | Sport | Culture | Lifestyle

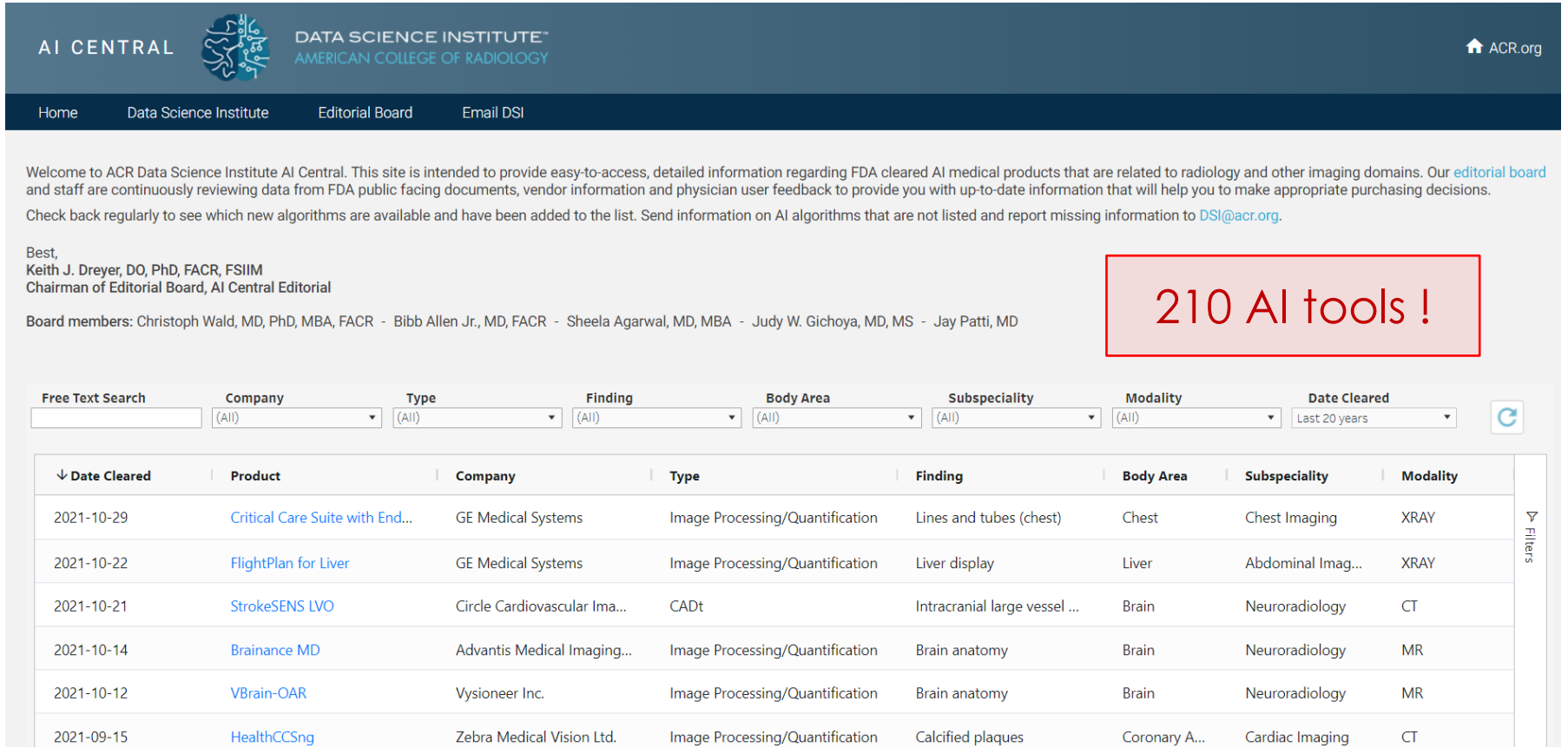
Coronavirus World UK Environment Science Global development Football **Tech** Business Obituaries

Artificial intelligence

AI system as good as experts at recognising skin cancers, say researchers

Deep learning-based system could be further developed for smartphones, increasing access to screening and aiding early detection of cancers

FDA-Cleared AI



Cleared but not approved

https://aicentral.acrdsi.org

Low-Risk AI

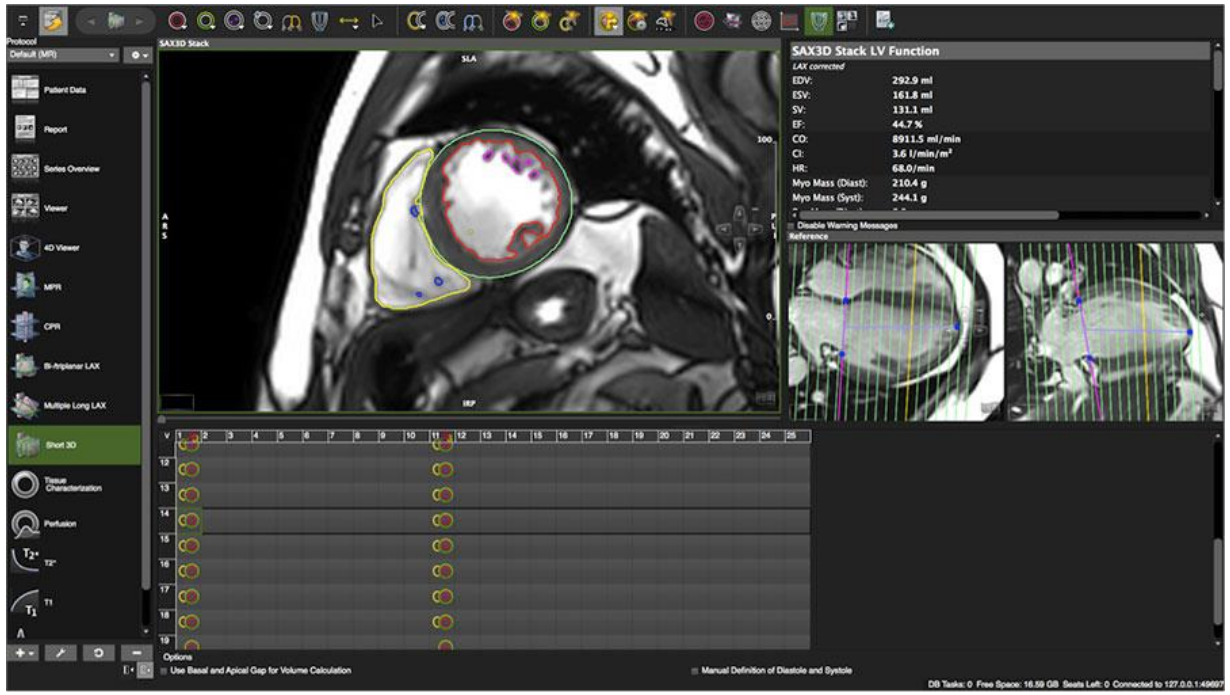
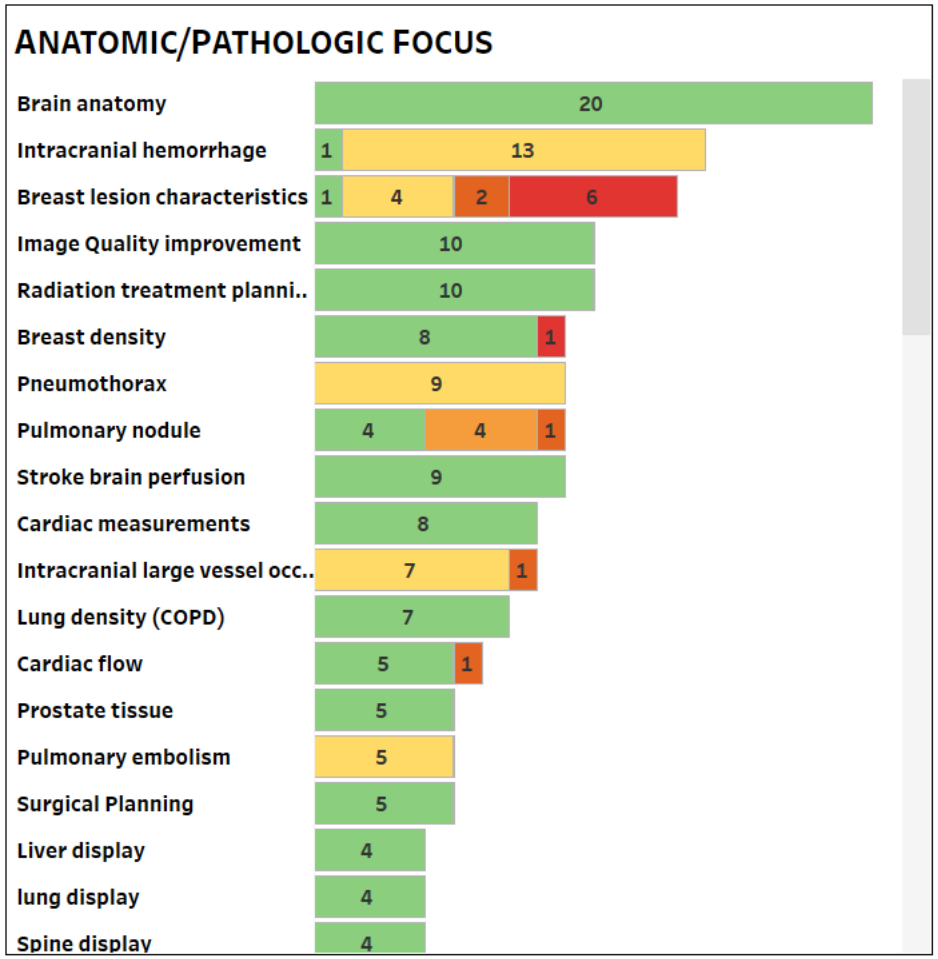
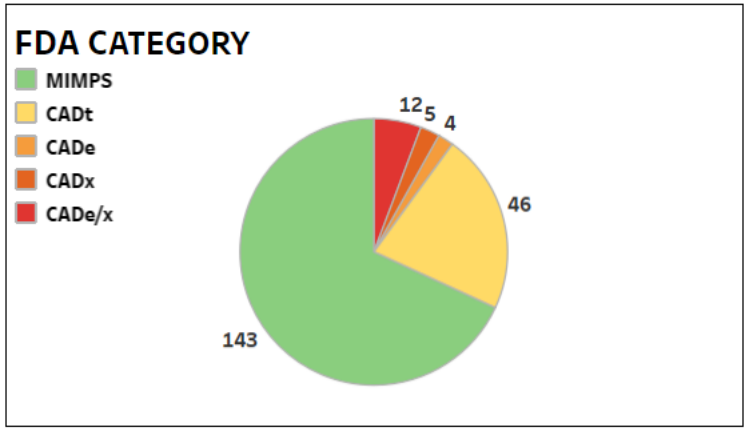


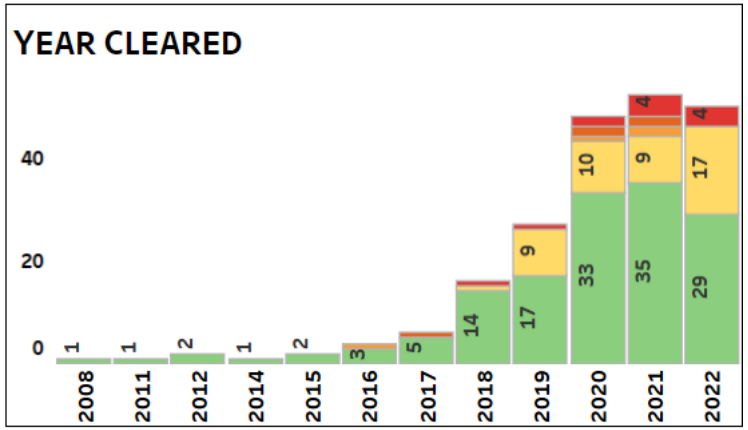
Image segmentation:

- ✓ Clinicians can verify and correction the segmentations
- ✓ Limited risk for patient harm
- ✓ Enormous gain in productivity

FDA-Cleared AI







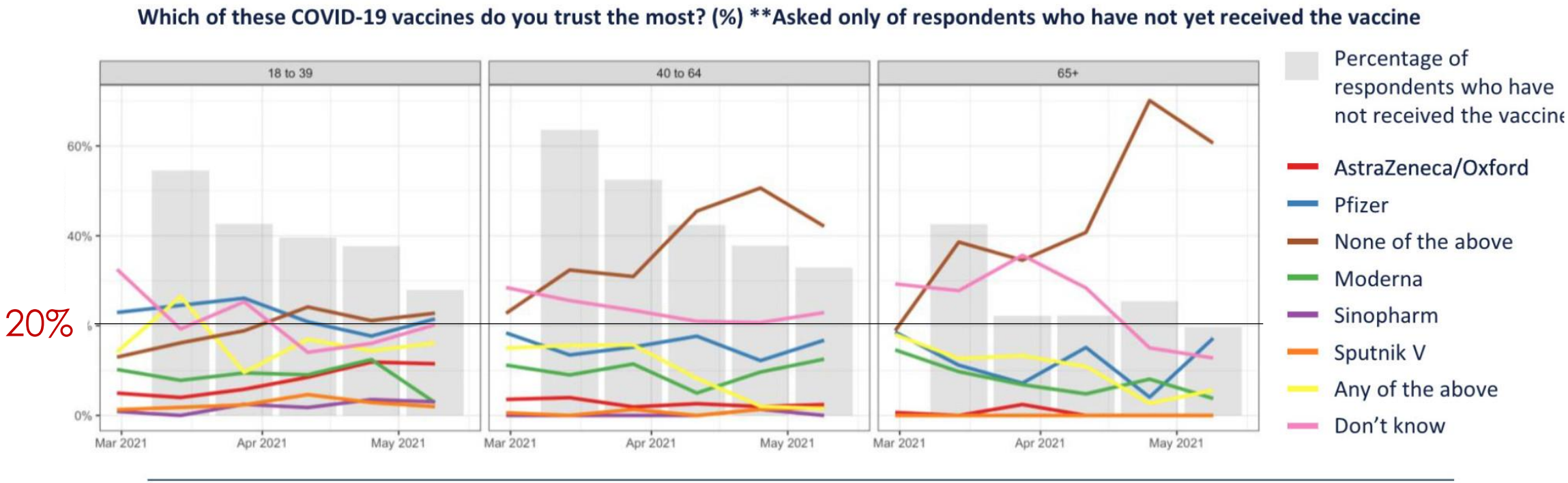
https://aicentral.acrdsi.org

Would You Trust an **AI Tool** that Decides on Your Diagnosis and/or Treatment?

Would You Trust a **New Vaccine** Developed Against a New Infectious Disease?

Covid-19 Vaccines

Trust about 6 months after vaccination started



Global attitudes towards a COVID-19 vaccine. Report Imperial College London, May 2021

Would **You** Trust an AI Tool that Decides on Your Diagnosis and/or Treatment?

You

The physician

The social carer

The health centre

The data protection officer

The health authorities

Your fellow citizens

Your insurance

Part 2-

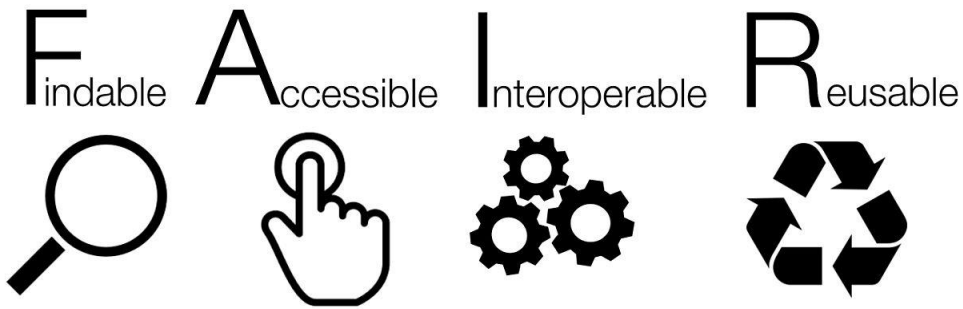
Success Story:

FAIR Guidelines for Data Management

Guidelines

- What is a guideline (Oxford Languages):
  - A general rule
  - Principle
  - Piece of advice.
- Other synonyms:
  - Best practice
  - Code of practice.

FAIR Guidelines



>8,000 citations

SCIENTIFIC DATA

Amended: Addendum

OPEN

SUBJECT CATEGORIES

» Research data

» Publication characteristics

Comment: The FAIR Guiding Principles for scientific data management and stewardship

Mark D. Wilkinson *et al.*#

Received: 10 December 2015

Accepted: 12 February 2016

Published: 15 March 2016

There is an urgent need to improve the infrastructure supporting the reuse of scholarly data. A diverse set of stakeholders—representing academia, industry, funding agencies, and scholarly publishers—have come together to design and jointly endorse a concise and measureable set of principles that we refer to as the FAIR Data Principles. The intent is that these may act as a guideline for those wishing to enhance the reusability of their data holdings. Distinct from peer initiatives that focus on the human scholar, the FAIR Principles put specific emphasis on enhancing the ability of machines to automatically find and use the data, in addition to supporting its reuse by individuals. This Comment is the first formal publication of the FAIR Principles, and includes the rationale behind them, and some exemplar implementations in the community.

FAIR Guidelines

Box 2 | The FAIR Guiding Principles

To be Findable:
F1. (meta)data are assigned a globally unique and persistent identifier
F2. data are described with rich metadata (defined by R1 below)
F3. metadata clearly and explicitly include the identifier of the data it describes
F4. (meta)data are registered or indexed in a searchable resource

To be Accessible:
A1. (meta)data are retrievable by their identifier using a standardized communications protocol
A1.1 the protocol is open, free, and universally implementable
A1.2 the protocol allows for an authentication and authorization procedure, where necessary
A2. metadata are accessible, even when the data are no longer available

To be Interoperable:
I1. (meta)data use a formal, accessible, shared, and broadly applicable language for knowledge representation.
I2. (meta)data use vocabularies that follow FAIR principles
I3. (meta)data include qualified references to other (meta)data

To be Reusable:
R1. meta(data) are richly described with a plurality of accurate and relevant attributes
R1.1. (meta)data are released with a clear and accessible data usage license
R1.2. (meta)data are associated with detailed provenance
R1.3. (meta)data meet domain-relevant community standards

FAIR Guidelines

www.go-fair.org

Example services that supply globally unique and persistent identifiers

- Identifiers.org provides resolvable identifiers in the form of URIs and CURIEs: **http://identifiers.org**
- Universally unique identifier: **https://en.wikipedia.org/wiki/Universally_unique_identifier**
- Persistent URLs: **http://www.purlz.org**
- Digital Object Identifier: **http://www.doi.org**
- Archival Resource Key: **https://escholarship.org/uc/item/9p9863nc**
- Research Resource Identifiers: **https://scicrunch.org/resources**

FAIR Impact

Indicators

1. Time spent
2. Cost of storage
3. Licence costs

4. Research retraction
5. Double funding
6. Cross-fertilization

7. Potential economic growth (as % of GDP)

FAIR Impact

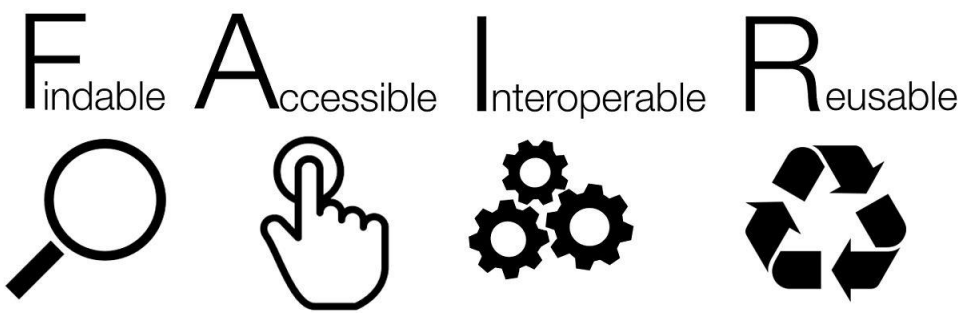
EU CALL: Clinical validation of AI solutions for treatment and care

TOPIC ID: HLTH-2021-DISEASE-04-04

"Proposals have to ensure that resulting data comply with the FAIR[2] principles."

[2] FAIR data are data, which meet principles of findability, accessibility, interoperability, and reusability.

FAIR Strengths



- ✓ Compact (4 principles, 15 items)
- ✓ Agreed upon internationally (sort of)
- ✓ Complements data regulations on privacy and security
- ✓ Backed by authorities

Part 3-

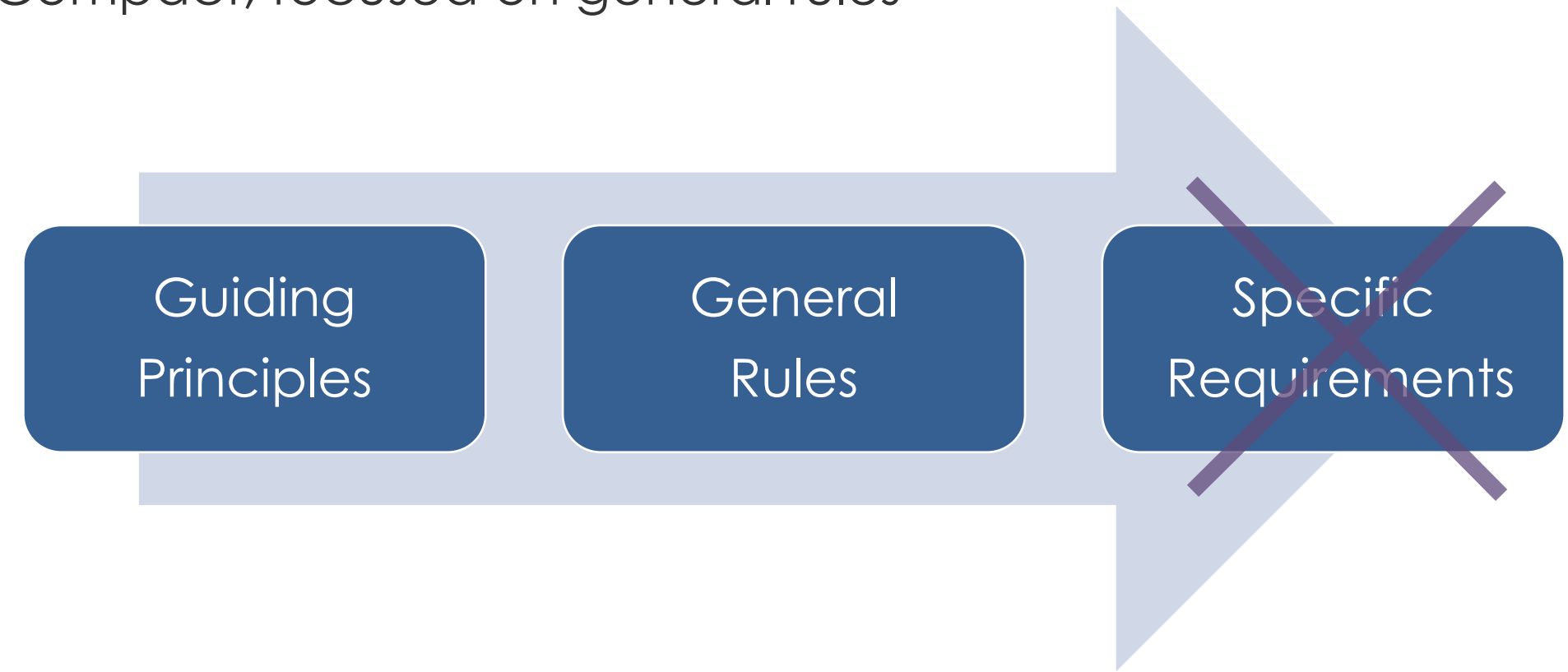
FUTURE-AI Guidelines

FUTURE-AI Promoters

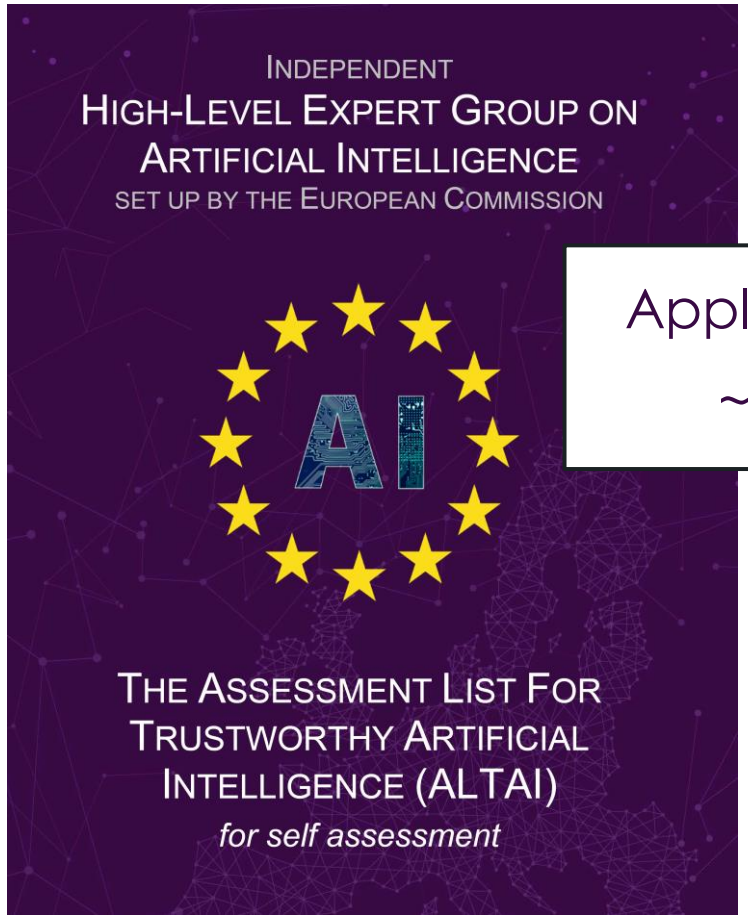
FUTURE-AI Approach

Similar to FAIR:

Compact, focused on general rules



ALTAI Checklist



Apply to AI in general

~140 questions

FG-AI4H Good Practices

INTERNATIONAL TELECOMMUNICATION UNION

TELECOMMUNICATION STANDARDIZATION SECTOR

STUDY PERIOD 2017-2020

FG-AI4H-N-031

ITU-T Focus Group on AI for Health

Original: English

WG(s):	Plenary	E-meeting, 15-17 February 2022

DOCUMENT

Source:	Editors DEL2.2
Title:	DEL2.2 Update: Good practices for health applications of machine learning: Considerations for manufacturers and regulators

Reporting Guidelines

BMJ Open

Protocol for development of a reporting guideline (TRIPOD-AI) and risk of bias tool (PROBAST-AI) for diagnostic and prognostic prediction model studies based on artificial intelligence

Gary S Collins [1,2] Paula Dhiman [1,2] Constanza L Andaur Navarro [3] Jie Ma [1] Lotty Hooft,[3,4] Johannes B Reitsma,[3] Patricia Logullo [1,2] Andrew L Beam [5,6] Lily Peng,[7] Ben Van Calster [8,9,10] Maarten van Smeden [3] Richard D Riley [11] Karel GM Moons[3,4]

nature medicine

OPEN

Guidelines for clinical trial protocols for interventions involving artificial intelligence: the SPIRIT-AI extension

Samantha Cruz Rivera[1,2,3], Xiaoxuan Liu[3,4,5,6,7], An-Wen Chan[8], Alastair K. Denniston[1,3,4,5,6,9] ✉, Melanie J. Calvert[1,2,3,6,10,11,12], The SPIRIT-AI and CONSORT-AI Working Group*, SPIRIT-AI and CONSORT-AI Steering Group and SPIRIT-AI and CONSORT-AI Consensus Group

Radiology: Artificial Intelligence

Checklist for Artificial Intelligence in Medical Imaging (CLAIM): A Guide for Authors and Reviewers

John Mongan, MD, PhD • Linda Moy, MD • Charles E. Kahn, Jr, MD, MS

From the Department of Radiology and Biomedical Imaging, University of California–San Francisco, San Francisco, Calif (J.M.); Department of Radiology and Center for Advanced Imaging Innovation and Research, New York University School of Medicine, New York, NY (L.M.); and Department of Radiology, University of Pennsylvania, 3400 Spruce St, 1 Silverstein, Philadelphia, PA 19104 (C.E.K.). Received March 4, 2020; revision requested March 5; accepted March 5. **Address correspondence to** C.E.K. (e-mail: *ckahn@rsna.org*).

Evaluation Guidelines

JAMIA Open, 3(3), 2020, 326–331
doi: 10.1093/jamiaopen/ooaa033
Advance Access Publication Date: 8 September 2020
Perspective

Perspective

Evaluating artificial intelligence in medicine: phases of clinical research

Yoonyoung Park[1], Gretchen Purcell Jackson[2,3], Morgan A. Foreman[1], Daniel Gruen[1], Jianying Hu[4] and Amar K. Das[1]

Open access Review

BMJ Health & Care Informatics

Evaluation framework to guide implementation of AI systems into healthcare settings

Sandeep Reddy ,[1] Wendy Rogers,[2] Ville-Petteri Makinen,[3] Enrico Coiera ,[4] Pieta Brown,[5] Markus Wenzel,[6] Eva Weicken,[6] Saba Ansari,[7] Piyush Mathur ,[8] Aaron Casey,[3] Blair Kelly[7]

JACR Journal of the American College of Radiology

Regulatory Frameworks for Development and Evaluation of Artificial Intelligence–Based Diagnostic Imaging Algorithms: Summary and Recommendations

David B. Larson, MD, MBA[a], Hugh Harvey, MBBS[b], Daniel L. Rubin, MD, MS[c], Neville Irani, MD[d], Justin R. Tse, MD[e], Curtis P. Langlotz, MD, PhD[f]

FUTURE-AI Guidelines

Summarise current knowledge into FUTURE-AI guidelines

- Compact as FAIR → 6 principles
- Focus on general rules → 30 rules
- Through consensus → 95+ international experts
- Inter-disciplinary → AI scientists, clinicians, ethics & legal experts, social scientists, regulators, etc

Literature Review

Robustness, safety, security, Fairness transparency, Traceability, accountability generalisability, Explainability, Usability responsible AI

Literature Review

Robustness, safety, security, **F**airness transparency, Traceability, accountability generalisability, **E**xplainability, **U**sability responsible AI

FEU

Literature Review

Robustness, safety, security, **F**airness

transparency, **T**raceability, accountability

generalisability, **E**xplainability, **U**sability

responsible AI

FEURT

Literature Review

Robustness, safety, security, **F**airness

transparency, **T**raceability, accountability

Universality, **E**xplainability, **U**sability

responsible AI

FEURTU

Final Acronym

Robustness, safety, security, **F**airness

transparency, **T**raceability, accountability

Universality, **E**xplainability, **U**sability

responsible AI

FEURTU → FUTURE

Medical AI Shall Be:

FAIR	UNIVERSAL	TRACEABLE	USABLE	ROBUST	EXPLAINABLE
					

Methodology

(1) Create working groups for each guiding principle

(2) Review literature then propose initial guidelines

(3) Organise a survey with 80+ international experts

(4) Gather feedback on guidelines & disagreements

(5) Obtain a final list of 30 consensus guidelines

Consortium

www.future-ai.eu

karim.lekadir@ub.edu

Fairness

1. Define sources of bias (at the design phase)
2. Collect data according to sources of bias (when possible)
3. Evaluate bias (when data permits it)
4. Mitigate bias (if bias is identified)

Universality

1. Define clinical settings (and variations across settings)
2. Use proven standards (clinical, technical, etc)
3. Perform external validation (with external data, sites, evaluators)
4. Apply generalisation measures (if needed)

Traceability

1. Define requirements for traceability (what to trace, when, how, by whom)
2. Provide documentation (general, technical, scientific)
3. Implement tools for monitoring (quality control, periodic evaluations, model updates, logging)
4. Implement human oversight and governance (quality control, decision over-rule, tool management)

Usability

1. Define intended use and user requirements
2. Provide training (materials and activities)
3. Evaluate usability (with diverse end-users)
4. Evaluate utility & safety (in real-world practice)

Robustness

1. Define sources of data variations
2. Train with representative data
3. Test robustness (against real-world variations)
4. Apply robustness measures (if needed)

Explainability

1. Define need, goals and approaches for explainability
2. Test explainability (with end-users)
3. Estimate uncertainty (calibrated)

General

1. Engage inter-disciplinary stakeholders
2. Implement measures for data protection
3. Select appropriate evaluation metrics
4. Comply with medical AI regulations
5. Implement security measures against attacks
6. Investigate and address ethical issues
7. Investigate and address social issues

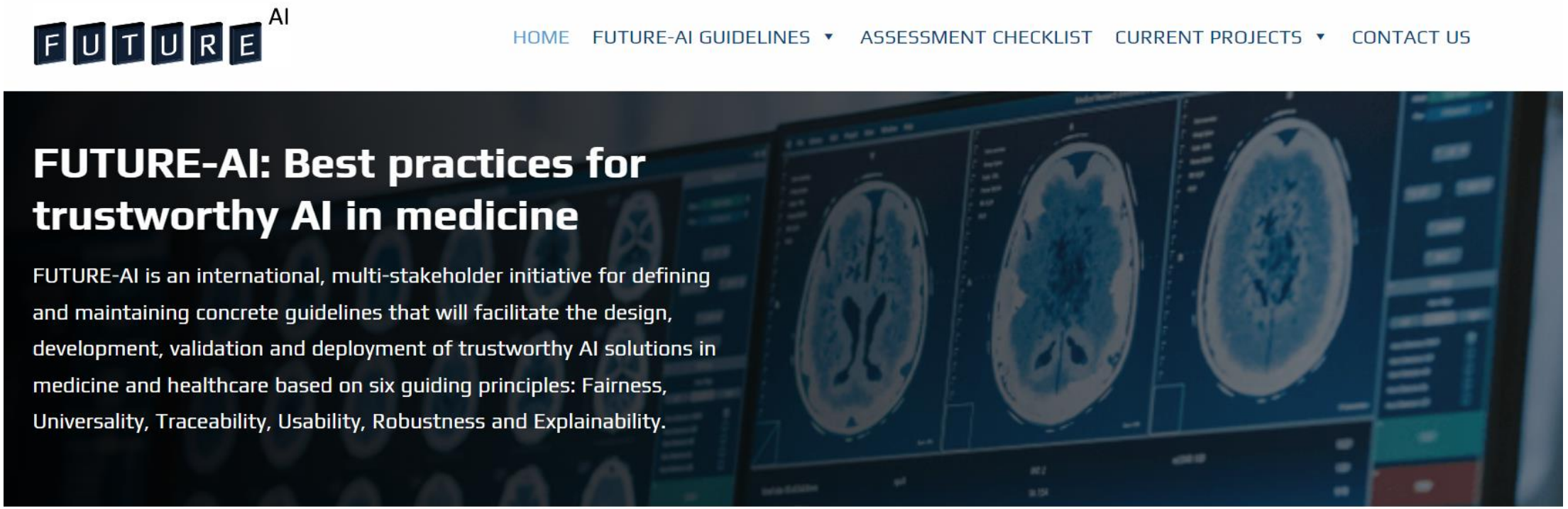
FUTURE-AI Paper

FUTURE-AI: International consensus guidelines for trustworthy and deployable artificial intelligence in healthcare

Abstract:

Despite major advances in the field of medical artificial intelligence (AI) over the years, the deployment and adoption of AI technologies remain limited in real-world clinical practice. In recent years, concerns have been raised on the technical, clinical, ethical and legal risks associated with medical AI. To increase adoption in the real world, it is essential that medical AI tools are trusted and accepted by patients, clinicians, health organisations and authorities. This paper addresses an important gap in the field of medical AI, by delivering the FUTURE-AI framework as the very first international, consensus guidelines for trustworthy and deployable AI in healthcare. We established a concise, yet actionable list of 30 guidelines organised according to six guiding principles of trustworthy AI, *i.e.* Fairness, Universality, Traceability, Usability, Robustness and Explainability. We employed an iterative process based on a modified Delphi approach, with a large-scale multi-stakeholder, multi-disciplinary, multi-society and multi-national consultation that involved a consortium of over 90 inter-disciplinary experts from all continents. The derived guidelines are holistic and cover technical, clinical, legal and socio-ethical dimensions of trustworthy AI, as well as the entire lifecycle of medical AI, from design, development and validation to regulation, deployment, and monitoring. The FUTURE-AI framework is intended to guide AI developers in the process of building medical AI tools that will be deployed, adopted and used in real-world practice. Furthermore, researchers are encouraged to take the guidelines into account in proof-of-concept stages to facilitate future translation towards clinical practice of medical AI.

FUTURE-AI



www.future-ai.eu

karim.lekadir@ub.edu

Next Steps

- Integrate feedback from FG-AI4H members
- Submit the paper as a contribution to the FG-AI4H
- Compare the FUTURE-AI and FG-AI4H deliverables, identify commonalities, mismatches and complementarities
- Discuss further (e.g. March meeting)

Funding

EuCanImage (2020-2024)
Grant # 952103

RadioVal (2022-2026)
Grant # 101057699

www.future-ai.eu

karim.lekadir@ub.edu

Acknowledgment

www.future-ai.eu

karim.lekadir@ub.edu