



# Intelligenza artificiale e sistemi di supporto decisionale



**Prof. Ing. Federico Cabitza, PhD**  
Dipartimento di Informatica, Sistemistica e Comunicazione  
Università degli Studi di Milano-Bicocca

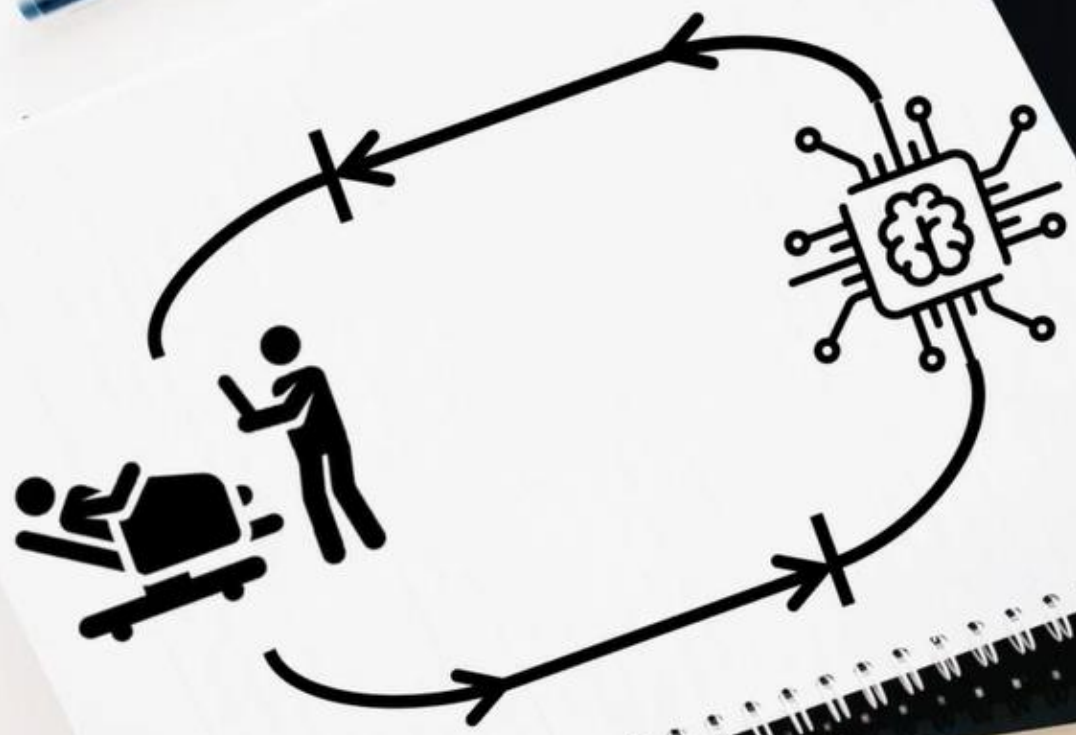




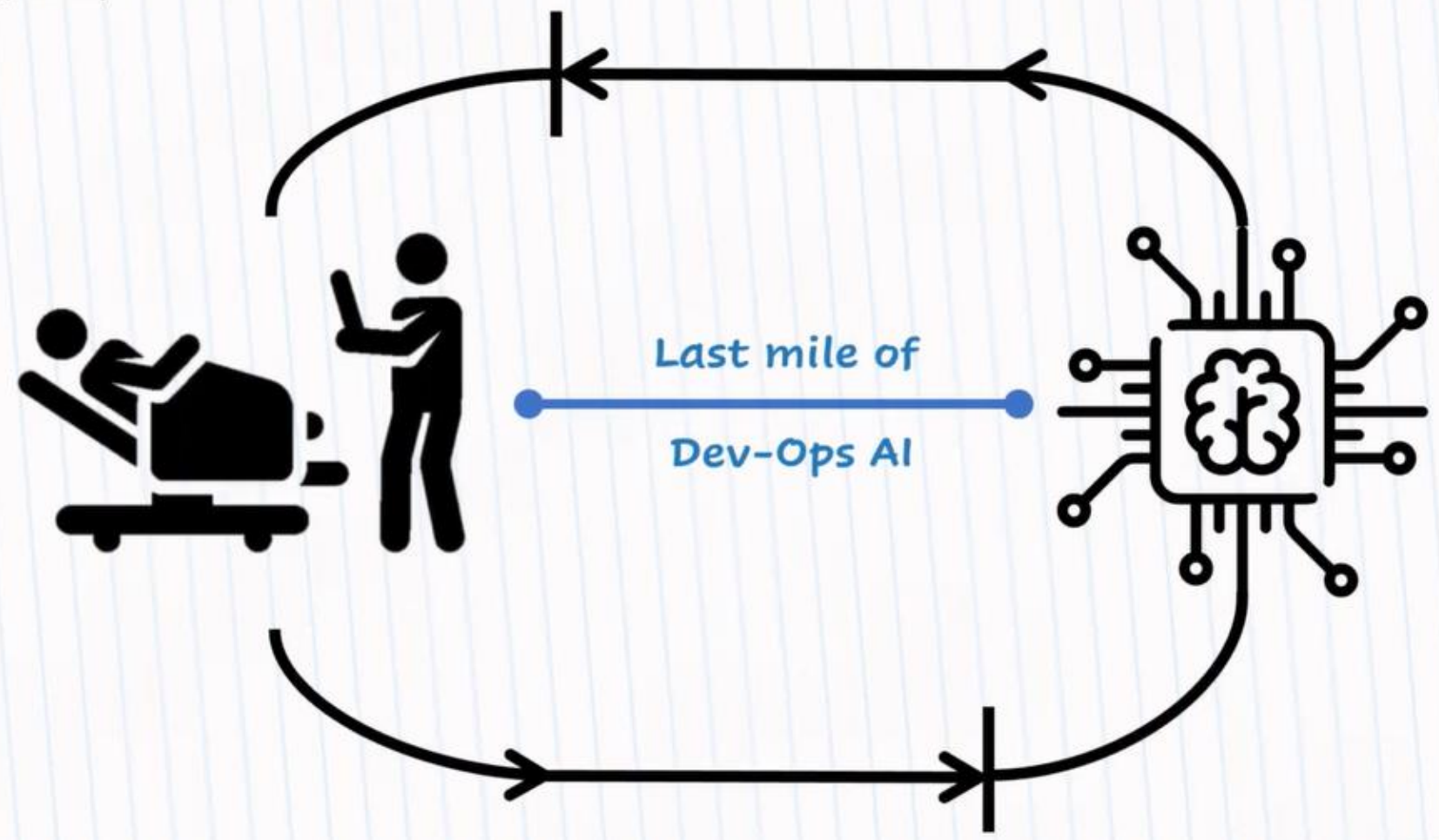
A futuristic robot with a glowing eye and a hand pointing at a grid of brain scan images. The robot is white and grey, with a large, circular, glowing eye on its head. Its right arm is extended, with its hand pointing towards a specific brain scan image in a grid. The background is a dark blue grid of many brain scan images, each showing a cross-section of a brain with white matter tracts. The text "Intelligenza artificiale e sistemi di supporto decisionale" is overlaid on the right side of the image in a bold, red font.

# Intelligenza artificiale e sistemi di supporto decisionale

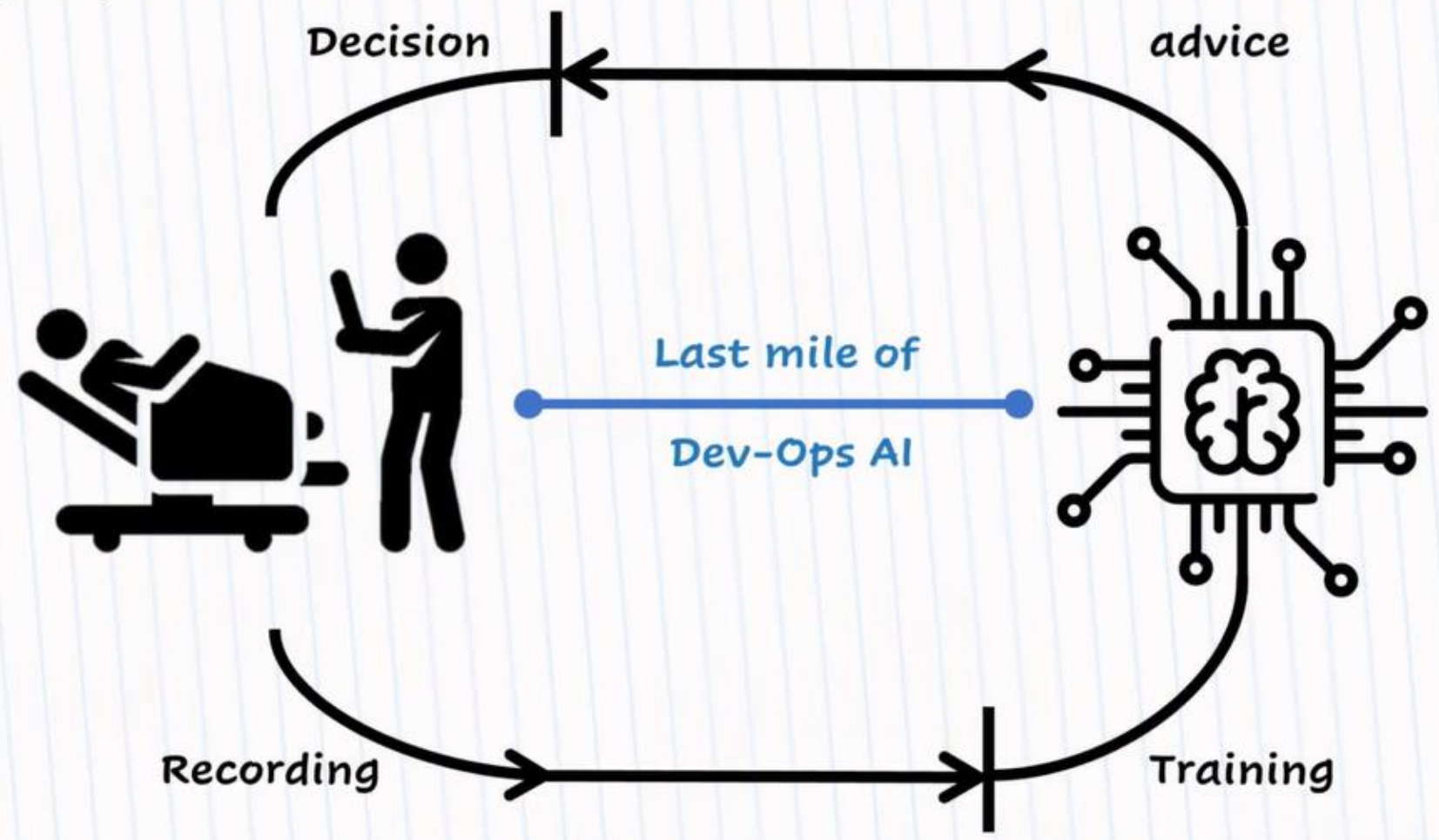




closed loop of human-machine mutual influence



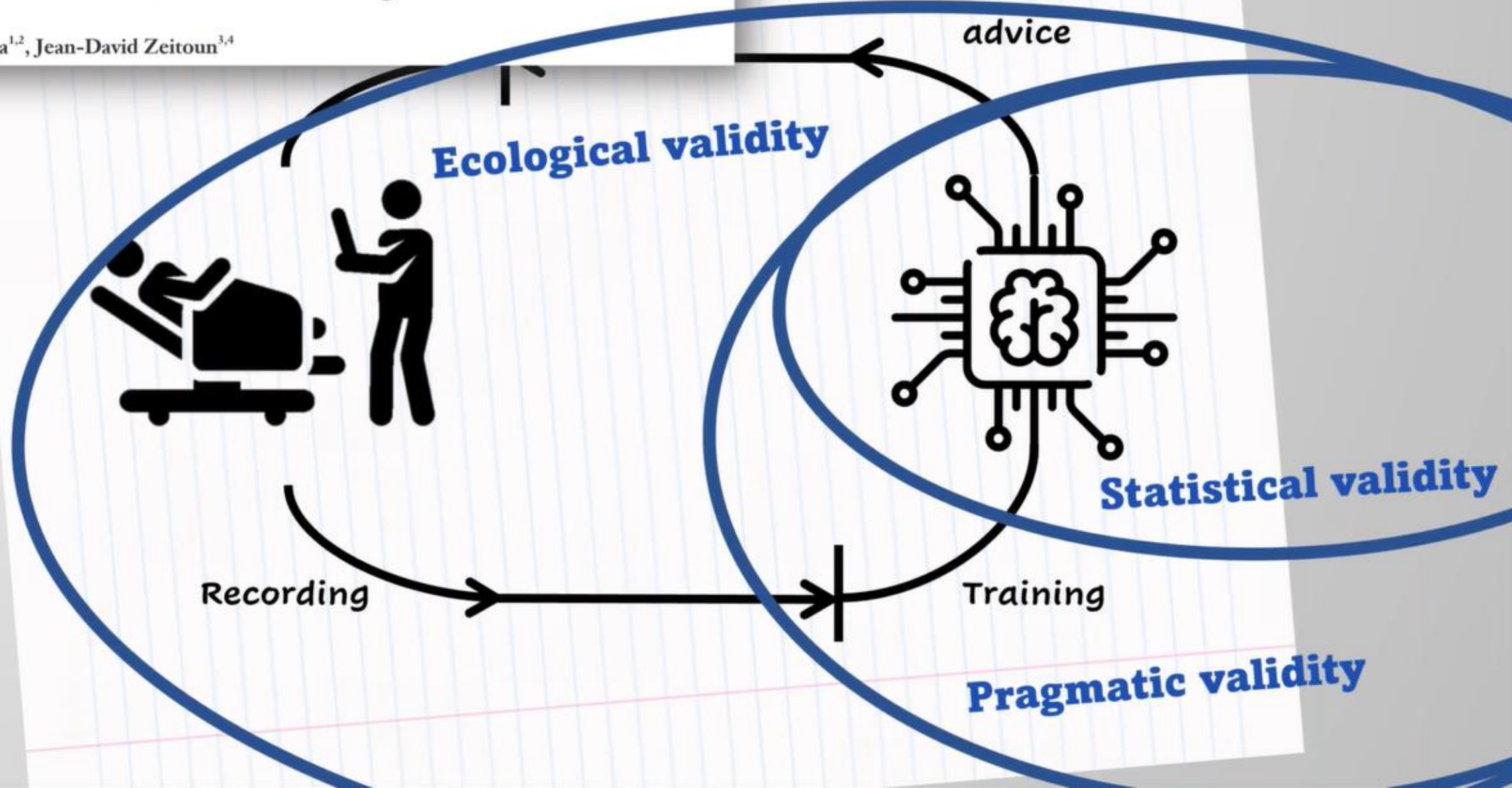
# closed loop of human-machine mutual influence





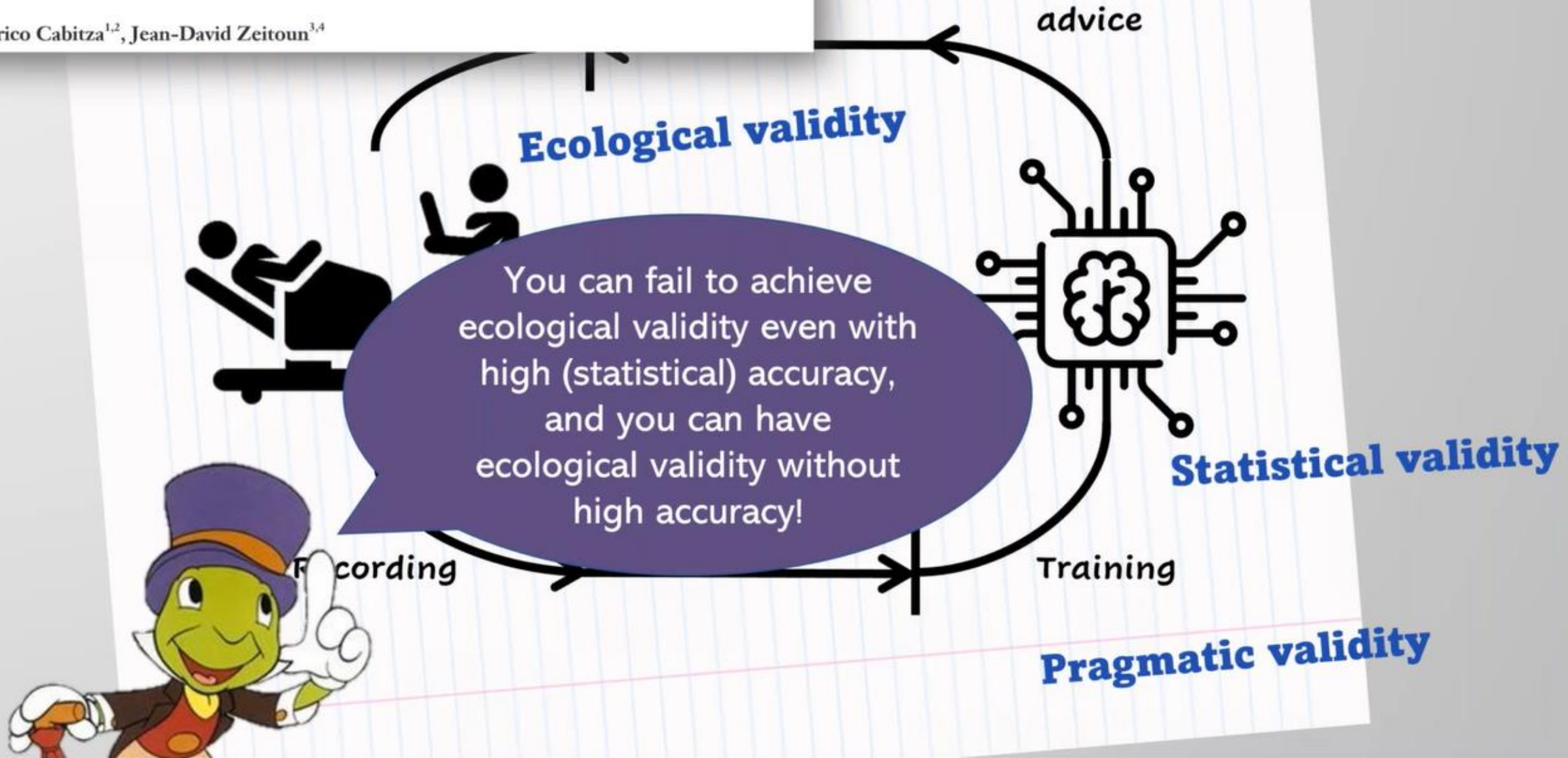
## The proof of the pudding: in praise of a culture of real-world validation for medical artificial intelligence

Federico Cabitza<sup>1,2</sup>, Jean-David Zeitoun<sup>3,4</sup>



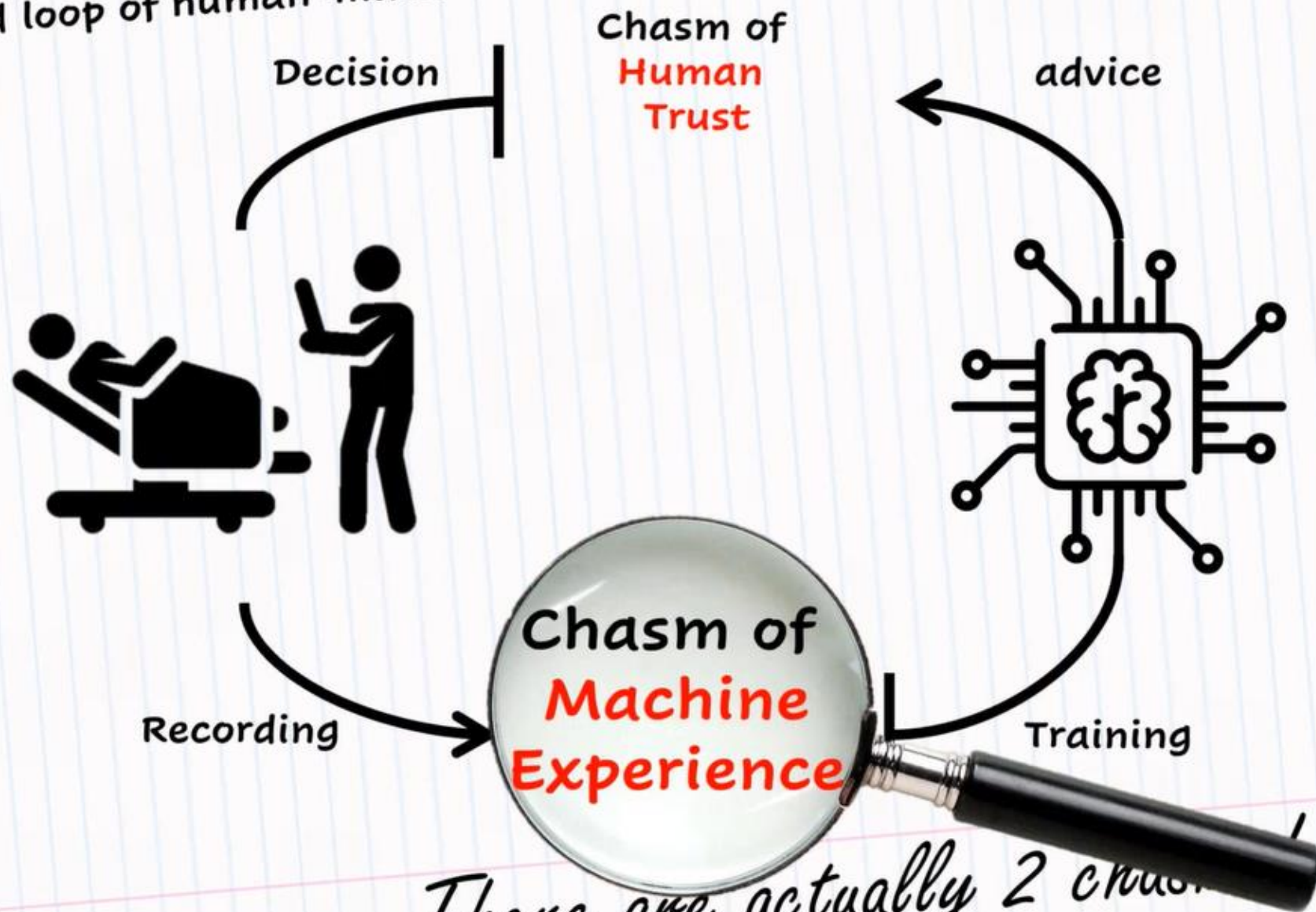
# The proof of the pudding: in praise of a culture of real-world validation for medical artificial intelligence

Federico Cabitza<sup>1,2</sup>, Jean-David Zeitoun<sup>3,4</sup>





closed loop of human-machine mutual influence

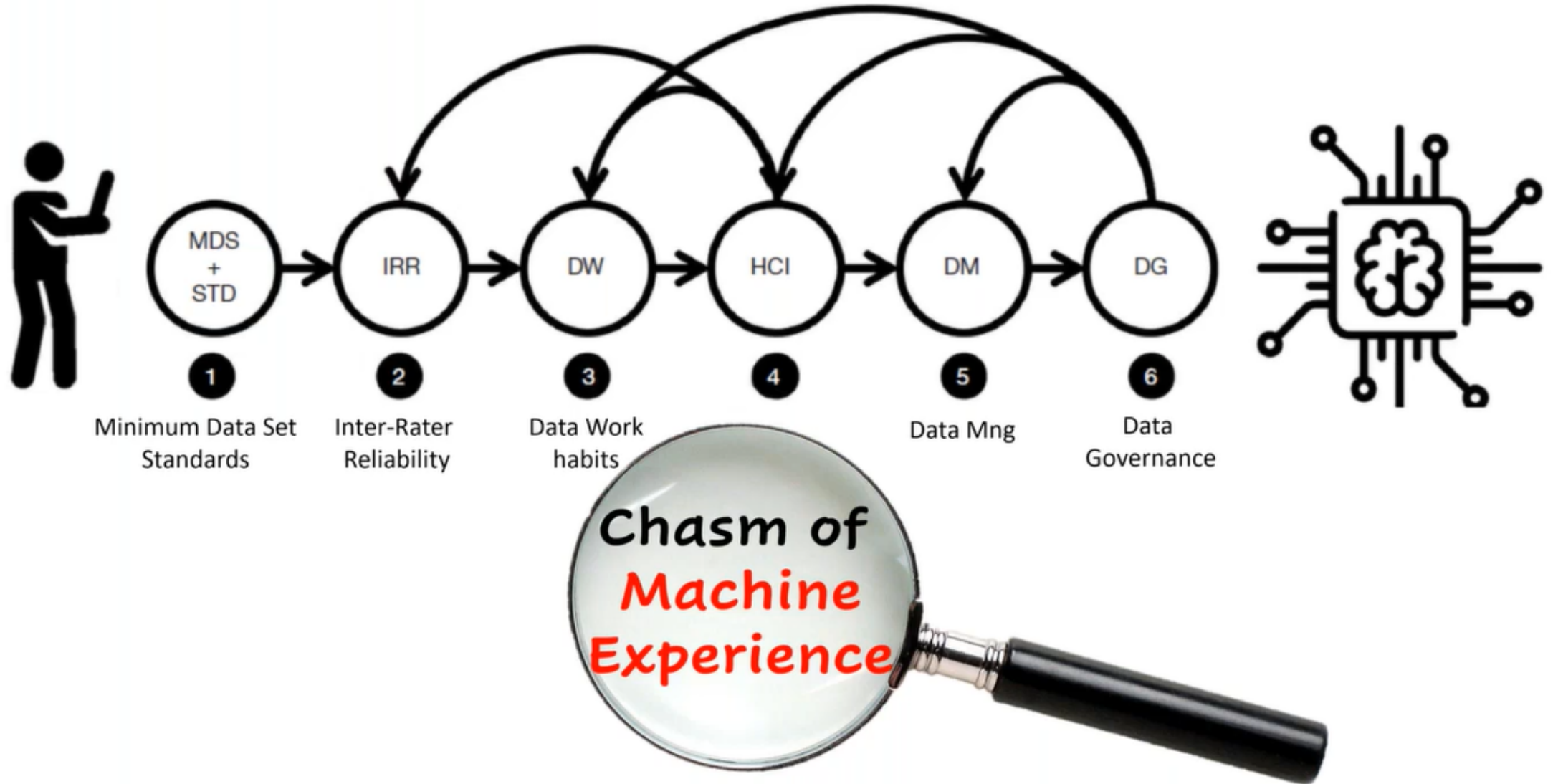


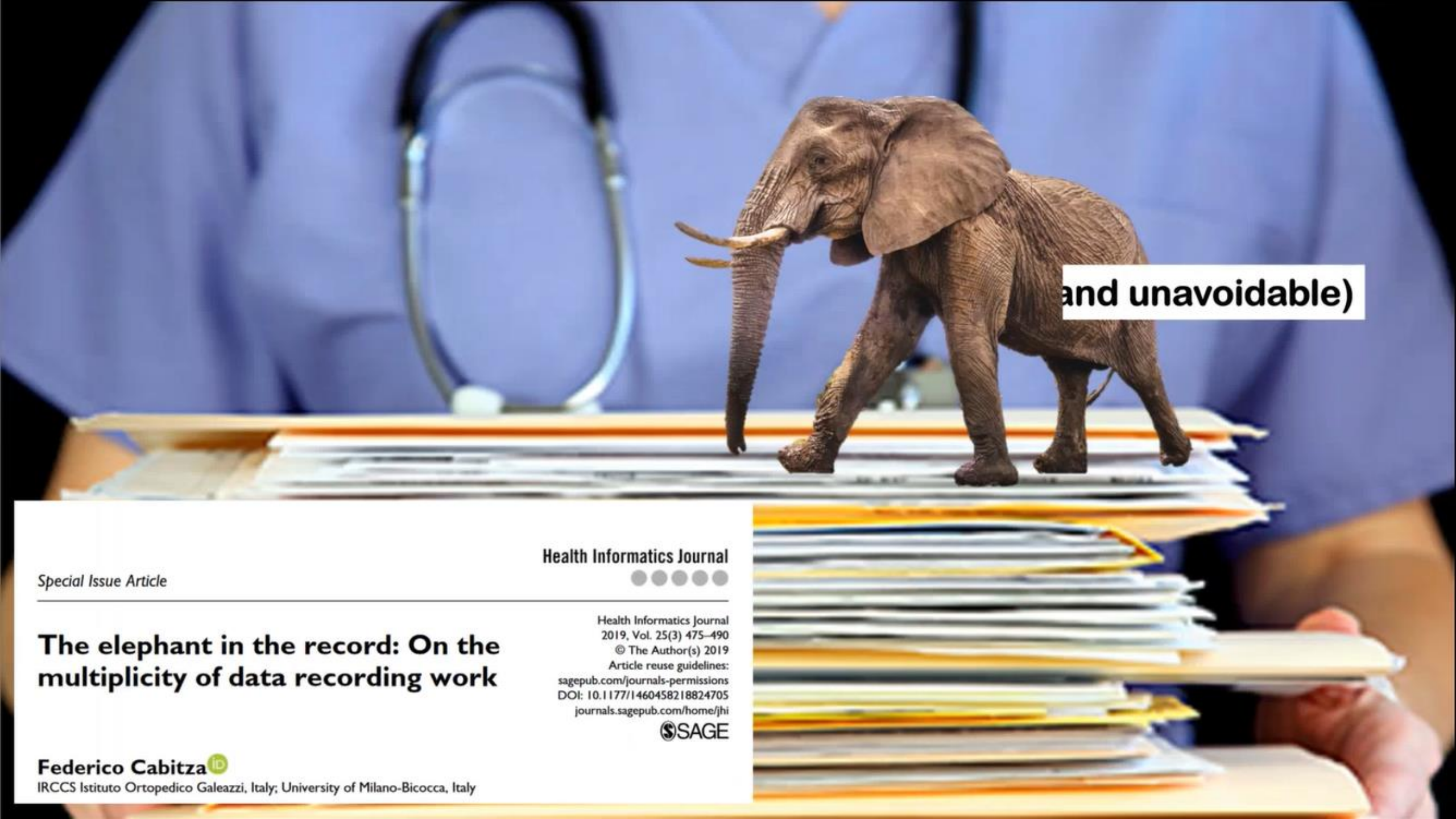
There are actually 2 chasms.



## Bridging the “last mile” gap between AI implementation and operation: “data awareness” that matters

Federico Cabitza, Andrea Campagner, Clara Balsano





and unavoidable)

*Special Issue Article*

## The elephant in the record: On the multiplicity of data recording work

**Federico Cabitza** 

IRCCS Istituto Ortopedico Galeazzi, Italy; University of Milano-Bicocca, Italy

**Health Informatics Journal**



Health Informatics Journal

2019, Vol. 25(3) 475–490

© The Author(s) 2019

Article reuse guidelines:

[sagepub.com/journals-permissions](http://sagepub.com/journals-permissions)

DOI: 10.1177/1460458218824705

[journals.sagepub.com/home/jhi](http://journals.sagepub.com/home/jhi)





# Development, evaluation, and validation of machine learning models for COVID-19 detection based on routine blood tests

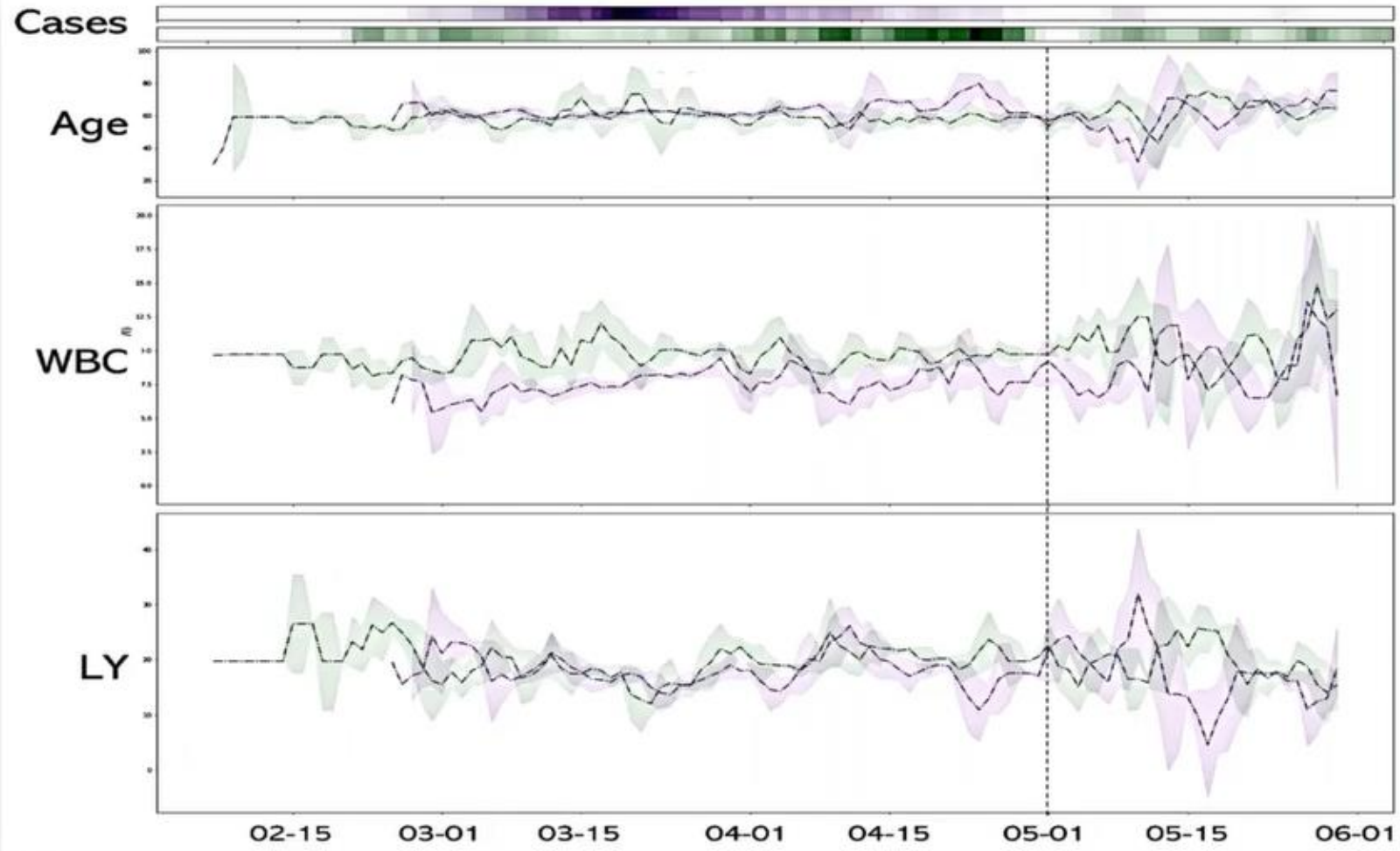


Clinical Chemistry  
and Laboratory  
Medicine (CCLM)  
Volume 59 Issue 2

ACCURACY DECREASE

CONCEPT DRIFT

TIME



2<sup>nd</sup> elephant:

Current accuracy  
is a fictional construct.





2<sup>nd</sup> elephant:

Current accuracy  
is a fictional construct.

Open Access Article

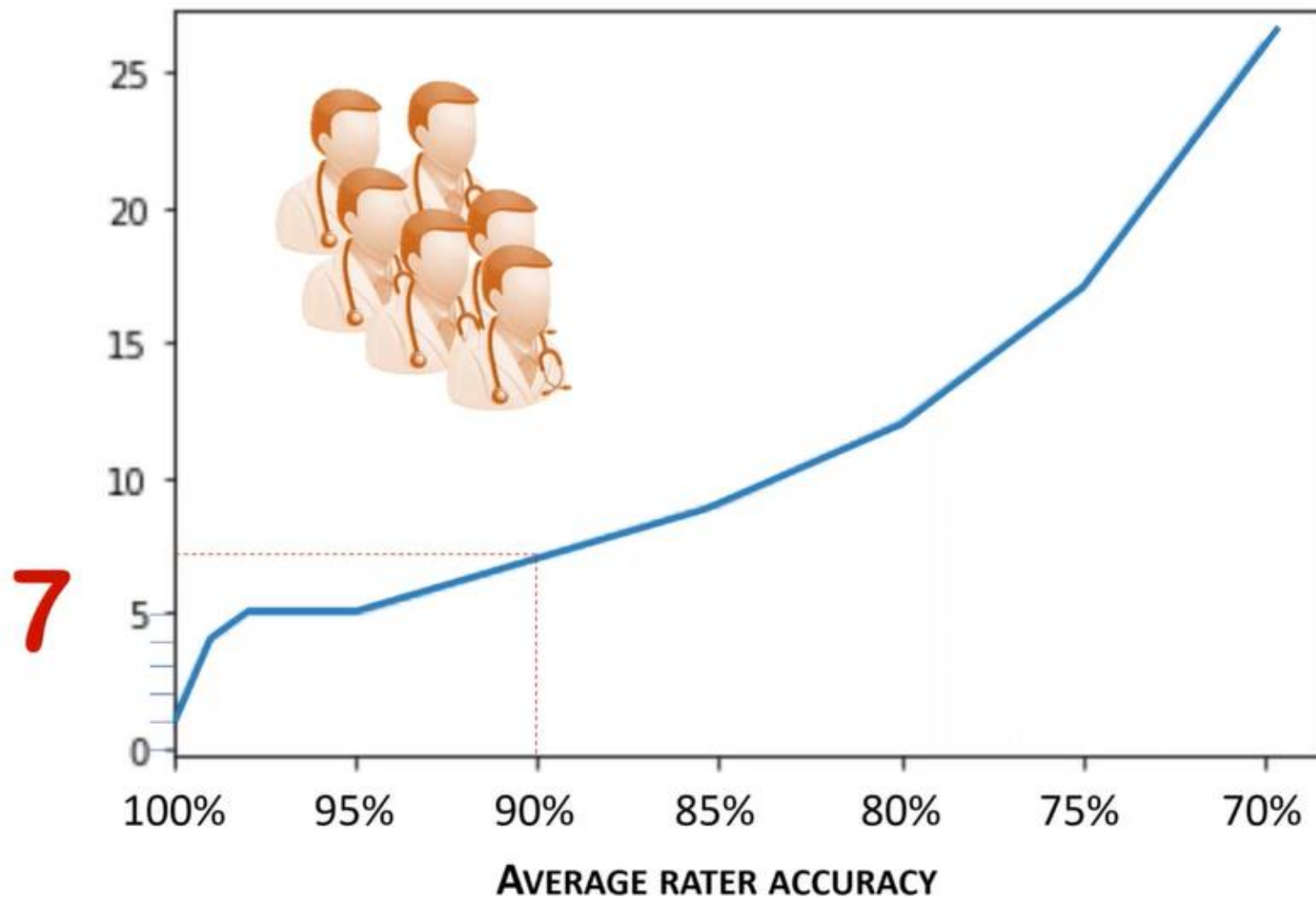
## The Elephant in the Machine: Proposing a New Metric of Data Reliability and its Application to a Medical Case to Assess Classification Reliability

by  Federico Cabitza <sup>1,\*</sup>  ,  Andrea Campagner <sup>1</sup> ,  Domenico Albano <sup>2,3</sup>  ,  Alberto Aliprandi <sup>4</sup> ,  
 Alberto Bruno <sup>3</sup>  ,  Vito Chianca <sup>2</sup> ,  Angelo Corazza <sup>2</sup> ,  Francesco Di Pietto <sup>5</sup> ,  
 Angelo Gambino <sup>2</sup> ,  Salvatore Gitto <sup>6</sup>  ,  Carmelo Messina <sup>2,6</sup> ,  Davide Orlandi <sup>7</sup>  ,  
 Luigi Pedone <sup>2</sup> ,  Marcello Zappia <sup>8,9</sup>   and  Luca Maria Sconfienza <sup>2,6</sup>  

<sup>1</sup> Department of Informatics, Systemics and Communication (DISCo), University of Milano-Bicocca, 20126 Milano, Italy

<sup>2</sup> IRCCS Istituto Ortopedico Galeazzi, 20161 Milano, Italy

## NUMBER OF RATERS TO INVOLVE TO GET A LABELLING ERROR UNDER 5%



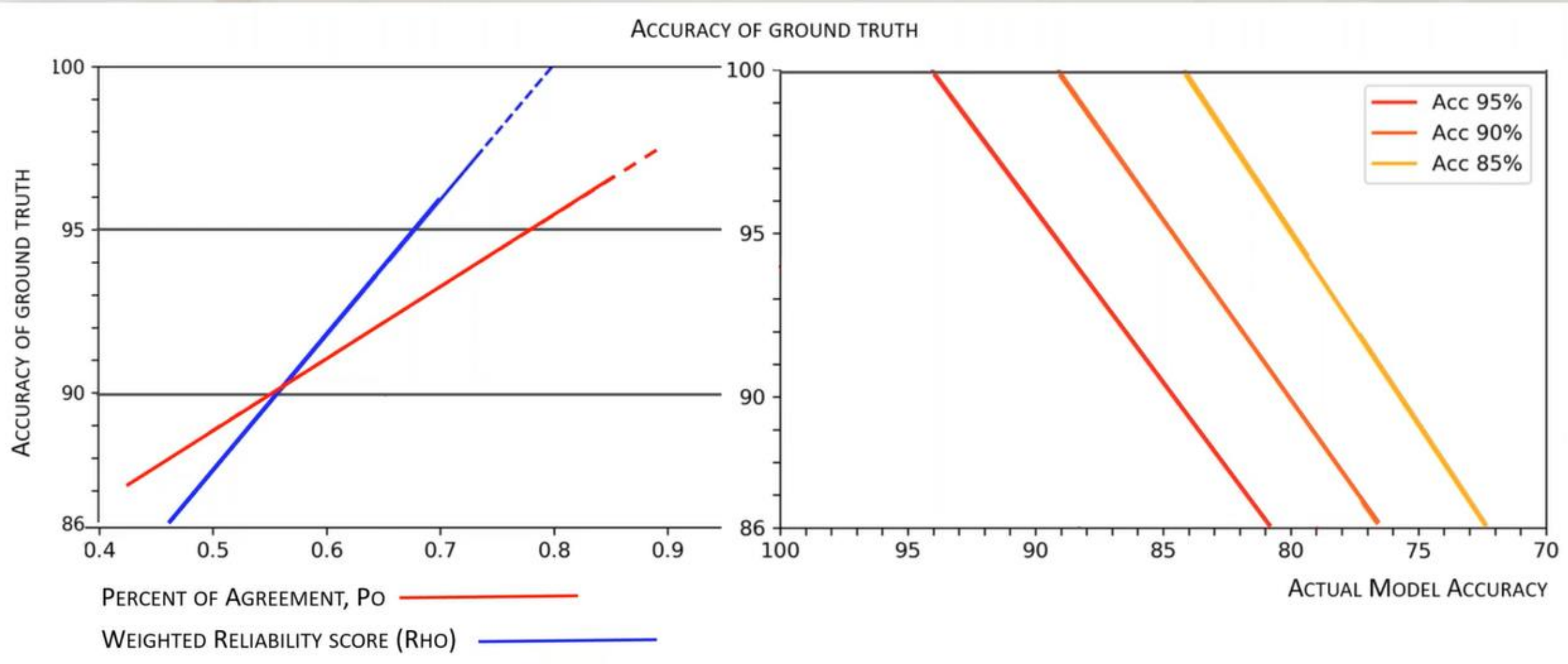
2<sup>nd</sup> elephant:

7

Current accuracy  
is a fictional construct.



# A nomogram



We proposed a new metric

76

74

72

7M

68

66

64

62

60

58

56

the weighted Utility (wU)

$$wU(\tau, \mathbf{r}, S, h) = \frac{1}{\mathbf{r}(Pos)} \sum_{x_i: y_i=1} r(x_i) \cdot \sigma_Y(\mathbf{h}(x_i) | \tau(x_i)) - \frac{1}{\mathbf{r}(Pos)} \sum_{x_i: y_i=0} r(x_i) \cdot \frac{\tau(x_i)}{1 - \tau(x_i)} \cdot \sigma_Y(\mathbf{h}(x_i) | \tau(x_i)).$$



76

74

72

7M

68



$$wU(\tau, \mathbf{r}, S, h) = \frac{1}{\mathbf{r}(Pos)} \sum_{x_i: y_i=1} r(x_i) \cdot \frac{\tau(x_i)}{1 - \tau(x_i)} \cdot \sigma_Y(\mathbf{h}(x_i) | \tau(x_i)).$$



Intuitively, a decision support is useful if the number of times it is right in detecting a health problem is higher than the number of times it is wrong so.



$$wU(\tau, \mathbf{r}, S, h) = \frac{1}{r(Pos)} \sum_{x_i: y_i=1} r(x_i) \cdot \sigma_Y(h(x_i)|\tau(x_i)) - \frac{1}{r(Pos)} \sum_{x_i: y_i=0} r(x_i) \cdot \frac{\tau(x_i)}{1 - \tau(x_i)} \cdot \sigma_Y(h(x_i)|\tau(x_i)).$$

True positive rate

False positive rate





And if:

- 1 It is optimized to avoid the most impactful kind of error (at class level)
- 2 It helps you when you need it most (i.e., most difficult/rarest cases)
- 3 It doesn't take guesses.

Intuitively, a decision support is useful if the number of times it is right in detecting a health problem is higher than the number of times it is wrong so.



$$wU(\tau, r, S, h) = \frac{1}{r(Pos)} \sum_{x_i: y_i=1} \overset{2}{\sigma} \overset{3}{c}(\tau(x_i)) - \frac{1}{r(Pos)} \sum_{x_i: y_i=0} \overset{2}{\sigma} \cdot \frac{1}{1} \cdot \overset{3}{c}(\tau(x_i)).$$

True positive rate

False positive rate

## The IJMEDI checklist for assessment of medical AI

Requirement	Authors			Reviewers		
	NA	No	Yes	OK	mR	MR
Problem Understanding						
1. Is the study population described, also in terms of inclusion/exclusion criteria (e.g., patients older than 18 tested for COVID-19; all inpatients hospitalized for 24 or more hours)? §	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
2. Is the study design described? (e.g., retrospective, prospective, cross-sectional [1], observational, randomized control trial [2]) §	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
3. Is the study setting described? (e.g., teaching tertiary hospital; primary care ambulatory, nursing home, medical laboratory, R&D laboratory) §	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
4. Is the source of data described? (e.g., electronic specialty registry; laboratory information system; electronic health record; picture archiving and communication system) §	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
5. Is the medical task reported? (e.g., diagnostic detection, diagnostic characterization, diagnostic staging, prognosis (on which endpoint), event prediction, risk stratification, anatomical structure segmentation, treatment selection and planning, monitoring) §	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
6. Is the data collection process described, also in terms of setting-specific data collection strategies (e.g. whether body temperatures are measured only in the morning; whether some blood tests are performed only in light of a specific diagnostic hypothesis)? Any consideration about data quality is appreciated, e.g., in regard to completeness, plausibility, and robustness with respect to upcoding or downcoding practices	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Data Understanding						
7. Are the subject demographics described in terms of	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
1. average age (mean or median);	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
2. age variability (standard deviation (SD) or inter-quartile range (IQR));	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
3. gender breakdown (e.g., 55% female, 44% male, 1% not reported); §	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
4. main comorbidities;	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
5. ethnic group (e.g., Native American, Asian, South East Asian, African, African American, Hispanic, Native Hawaiian or Other Pacific Islander, European or American White)	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
4. Rho [1]	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
5. Labelling technique [1]	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>



International Journal of Medical Informatics

Available online 2 June 2021, 104510

In Press, Journal Pre-proof



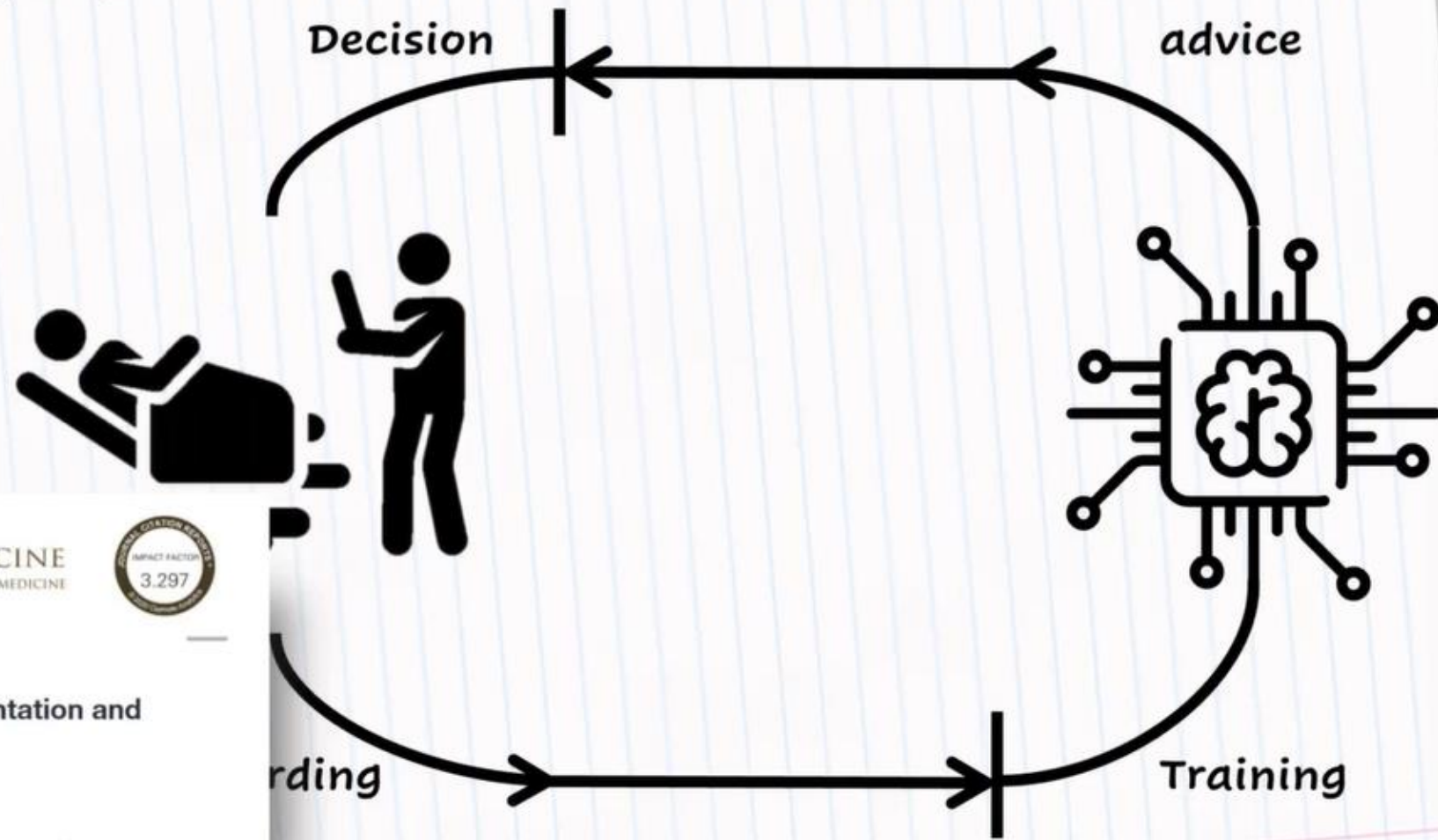
## The need to separate the wheat from the chaff in medical informatics

Federico Cabitza, Andrea Campagner

To discover more, please refer to the new IJMEDI checklist for assessment of medical AI



# closed loop of human-machine mutual influence



**ATM** ANNALS OF TRANSLATIONAL MEDICINE  
AN OPEN ACCESS JOURNAL COVERING ALL SUBSPECIALTIES OF TRANSLATIONAL MEDICINE



Review Article

## Bridging the “last mile” gap between AI implementation and operation: “data awareness” that matters

Federico Cabitza<sup>1</sup>, Andrea Campagner<sup>2</sup>, Clara Balsano<sup>3,4</sup>

<sup>1</sup>Dipartimento di Informatica, Sistemistica e Comunicazione, Università degli Studi di Milano-Bicocca, Milano, Italy; <sup>2</sup>IRCCS Istituto Ortopedico Galeazzi, Milano, Italy; <sup>3</sup>Dipartimento di Medicina Clinica, Sanità Pubblica, Scienze della Vita e dell’Ambiente, Università degli Studi dell’Aquila, L’Aquila, Italy; <sup>4</sup>Francesco Balsano Foundation, Via Giovanni Battista Martini 6, Rome, Italy

**Contributions:** (I) Conception and design: F Cabitza, C Balsano; (II) Administrative support: A Campagner; (III) Provision of study materials or patients: None; (IV) Collection and assembly of data: F Cabitza, A Campagner; (V) Data analysis and interpretation: F Cabitza; (VI) Manuscript writing: All authors; (VII) Final approval of manuscript: All authors.

**Correspondence to:** Federico Cabitza, PhD, Dipartimento di Informatica, Sistemistica e Comunicazione, Università degli Studi di Milano-Bicocca, Milano, Italy. Email: federico.cabitza@unimib.it

L'importanza di investire in ponti e... costruttori di ponti!







**GRAZIE!**



federico.cabitzza @ unimib.it



@cabitzaf