



Finanziato
dall'Unione europea
NextGenerationEU



Ministero
dell'Università
e della Ricerca



Italiadomani
PIANO NAZIONALE
DI RIPRESA E RESILIENZA



UNIVERSITÀ
DI TRENTO

Missione 4 Istruzione e Ricerca

La prova scientifica acquisita
mediante strumenti basati sull'IA:
l'apporto della metrologia

Luca Mari

Università Cattaneo LIUC

lmari@liuc.it

<https://lmari.github.io>



Finanziato
dall'Unione europea
NextGenerationEU



Ministero
dell'Università
e della Ricerca



Italiadomani
PIANO NAZIONALE
DI RIPRESA E RESILIENZA



UNIVERSITÀ
DI TRENTO

La prova scientifica acquisita mediante strumenti basati sull'IA: l'apporto della metrologia

Prima parte: metrologia

Seconda parte: intelligenza artificiale

Terza parte: verso una sintesi?



Finanziato
dall'Unione europea
NextGenerationEU



Ministero
dell'Università
e della Ricerca



Italiadomani
PIANO NAZIONALE
DI RIPRESA E RESILIENZA



UNIVERSITÀ
DI TRENTO

1. Metrologia

Si misura per **produrre informazione**,
ma non ogni attività di produzione di informazione è una misurazione:
cosa caratterizza la misurazione rispetto ad altre attività,
per esempio la raccolta di opinioni di persone, anche esperte?

Non il contenuto dell'informazione

(si può chiedere un'opinione sulla lunghezza di un oggetto,
e quello che si ottiene rimane un'opinione, non una misura)

Quantificare e misurare sono attività diverse



Finanziato
dall'Unione europea
NextGenerationEU



Ministero
dell'Università
e della Ricerca



Italiadomani
PIANO NAZIONALE
DI RIPRESA E RESILIENZA



UNIVERSITÀ
DI TRENTO

La misurazione è caratterizzata per il metodo con cui viene realizzata:

il come si produce informazione rende esplicito

il **grado di affidabilità** (*trustworthiness*) dell'informazione prodotta,

in termini di quanto bene è riferita all'oggetto misurato (**oggettività**)

e quanto bene è interpretabile socialmente (**intersoggettività**)



Finanziato
dall'Unione europea
NextGenerationEU



Ministero
dell'Università
e della Ricerca



Italiadomani
PIANO NAZIONALE
DI RIPRESA E RESILIENZA



UNIVERSITÀ
DI TRENTO

Chi misura non garantisce la verità dell'informazione prodotta,
ma si impegna, al meglio delle sue capacità,
a dichiarare **quanto l'informazione prodotta è affidabile**

«Nel riportare il risultato di una misurazione di una grandezza fisica,
è obbligatorio fornire una qualche indicazione quantitativa della qualità del risultato,
cosicché gli utenti ne possano accertare l'attendibilità.
Senza tale indicazione i risultati di misura non possono essere confrontati né tra di loro,
né con valori di riferimento assegnati da specifiche o norme.
È pertanto necessario che esista una procedura, di agevole comprensione e applicazione,
nonché generalmente accettata, per caratterizzare la qualità del risultato di una
misurazione, vale a dire, per valutarne ed esprimerne l'incertezza.»

UNI CEI 70098-3, Guida all'espressione dell'incertezza di misura, 0.1



Finanziato
dall'Unione europea
NextGenerationEU



Ministero
dell'Università
e della Ricerca



Italiadomani
PIANO NAZIONALE
DI RIPRESA E RESILIENZA



UNIVERSITÀ
DI TRENTO

Una prima sintesi:

ci possono essere cattive misurazioni, così come buone opinioni,
ma **non ci possono essere misurazioni opache** (*black box*)



Finanziato
dall'Unione europea
NextGenerationEU



Ministero
dell'Università
e della Ricerca



Italiadomani
PIANO NAZIONALE
DI RIPRESA E RESILIENZA



UNIVERSITÀ
DI TRENTO

2. Intelligenza artificiale

Un'entità problematica
(è davvero intelligente? pensa? ha una intenzionalità? è cosciente? ...
le opinioni degli esperti sono contrastanti)

Per comprendere quello che sta succedendo intorno all'IA
la domanda è prima di tutto un'altra:
come fa a fare quello che fa?



Finanziato
dall'Unione europea
NextGenerationEU



Ministero
dell'Università
e della Ricerca



Italiadomani
PIANO NAZIONALE
DI RIPRESA E RESILIENZA



UNIVERSITÀ
DI TRENTO

Il comportamento degli attuali sistemi di IA
è il risultato non di programmazione, ma di **addestramento**

Si dice «*Artificial Intelligence*» ma si intende «***Machine Learning***»

... e questo genera un'interessante analogia tra sistemi di misura e sistemi di IA...



Un sistema viene prima di tutto costruito, con una struttura adattiva / parametrica...

strumento di misura

rete neurale artificiale

... poi lo si adatta al contesto di uso...

taratura
mediante campioni

addestramento
mediante training set

... poi se ne accerta e valida il comportamento...

conferma metrologica

test / benchmark

... e infine lo si usa...

misurazione

inferenza



Finanziato
dall'Unione europea
NextGenerationEU



Ministero
dell'Università
e della Ricerca



Italiadomani
PIANO NAZIONALE
DI RIPRESA E RESILIENZA



UNIVERSITÀ
DI TRENTO

Ma l'analogia tra sistemi di misura e sistemi di IA diventa problematica proprio nell'attività di adattamento (taratura / addestramento):

- la taratura degli strumenti di misura si fonda sul Sistema Internazionale delle unità e su un insieme di campioni di misura primari riconosciuti globalmente, che garantiscono **la riferibilità metrologica** delle misure in ogni luogo e in ogni momento
- l'addestramento delle reti neurali artificiali è invece realizzato, almeno per ora, solo localmente e caso per caso, cosa che rende **complesso valutare l'affidabilità del loro comportamento**



Valutare l'affidabilità del comportamento di un sistema di IA: una prospettiva metrologica

- Caso A: sistemi di Machine Learning per classificazione (*p.es. sistemi di raccomandazione*)
Il «buon comportamento» è ben definito ed esplicitamente esemplificato
Ci possono essere problemi di instabilità e di distorsioni
* Strategie di soluzione consolidate (valutazione di accuratezza ecc.)
- Caso B: sistemi di Natural Language Processing context-free (*p.es. sistemi di traduzione*)
Il «buon comportamento» è solo parzialmente definito ed esemplificato
Ci possono essere problemi non solo di instabilità e di distorsioni, ma anche di incertezza di definizione
* Strategie di soluzione sperimentali (valutazione mediante benchmark vari)
- Caso C: sistemi di Natural Language Processing context-sensitive (*p.es. chatbot*)
Il «buon comportamento» è solo assai limitatamente definito ed esemplificato
Ci possono essere problemi non solo di instabilità e di distorsioni, e la definizione è incerta e contestuale
* Strategie di soluzione ancora preliminari (valutazione solo per confronto diretto)



Finanziato
dall'Unione europea
NextGenerationEU



Ministero
dell'Università
e della Ricerca



Italiadomani
PIANO NAZIONALE
DI RIPRESA E RESILIENZA



UNIVERSITÀ
DI TRENTO

3. Verso una sintesi?

«Benché la presente Guida fornisca uno schema generale per valutare l'incertezza, essa non può sostituirsi al pensiero critico, all'onestà intellettuale e alla capacità professionale. La valutazione dell'incertezza non è né un compito di routine né un esercizio puramente matematico, ma dipende dalla conoscenza approfondita della natura del misurando e della misurazione. La qualità e l'utilità dell'incertezza attribuita al risultato di una misurazione dipendono pertanto, in definitiva, dall'approfondimento, dall'analisi critica e dall'integrità morale di chi contribuisce ad assegnarne il valore.»

UNI CEI 70098-3, Guida all'espressione dell'incertezza di misura, 3.4.8

**Lo stesso, e a maggior ragione, vale per la valutazione
della qualità dell'informazione prodotta dai sistemi di IA**



Finanziato
dall'Unione europea
NextGenerationEU



Ministero
dell'Università
e della Ricerca



Italiadomani
PIANO NAZIONALE
DI RIPRESA E RESILIENZA



UNIVERSITÀ
DI TRENTO

I sistemi di IA ci stanno mettendo a disposizione dei superpoteri ma la responsabilità non può che rimanere nostra





Finanziato
dall'Unione europea
NextGenerationEU



Ministero
dell'Università
e della Ricerca



Italiadomani
PIANO NAZIONALE
DI RIPRESA E RESILIENZA



UNIVERSITÀ
DI TRENTO

Grazie per l'attenzione

Luca Mari

lmari@liuc.it

<https://lmari.github.io>