

La convergenza dell'Intelligenza Artificiale e della medicina di precisione nel contesto oncologico

Prof. Giulio Caravagna

Cancer Data Science Laboratory

Dipartimento di Matematica, Informatica e Geoscienze (MIGe)

Università degli Studi di Trieste

Email gcaravagna@units.it Website www.caravagnalab.org



The Cancer Data Science Laboratory
www.caravagnalab.org



2024 Trieste.
Giorgia, Elena R, Elena B, Lucrezia, Katzyarina, Sara, Virginia, Alice, Riccardo, Giovanni, Nicola, Azad, Giulio.

Artificial Intelligence and Machine Learning for

- Pre-existing versus de novo treatment resistance
- Genetic, epi-genetic evolution and plasticity
- Evolutionary trajectories across time and space



#sanita2030



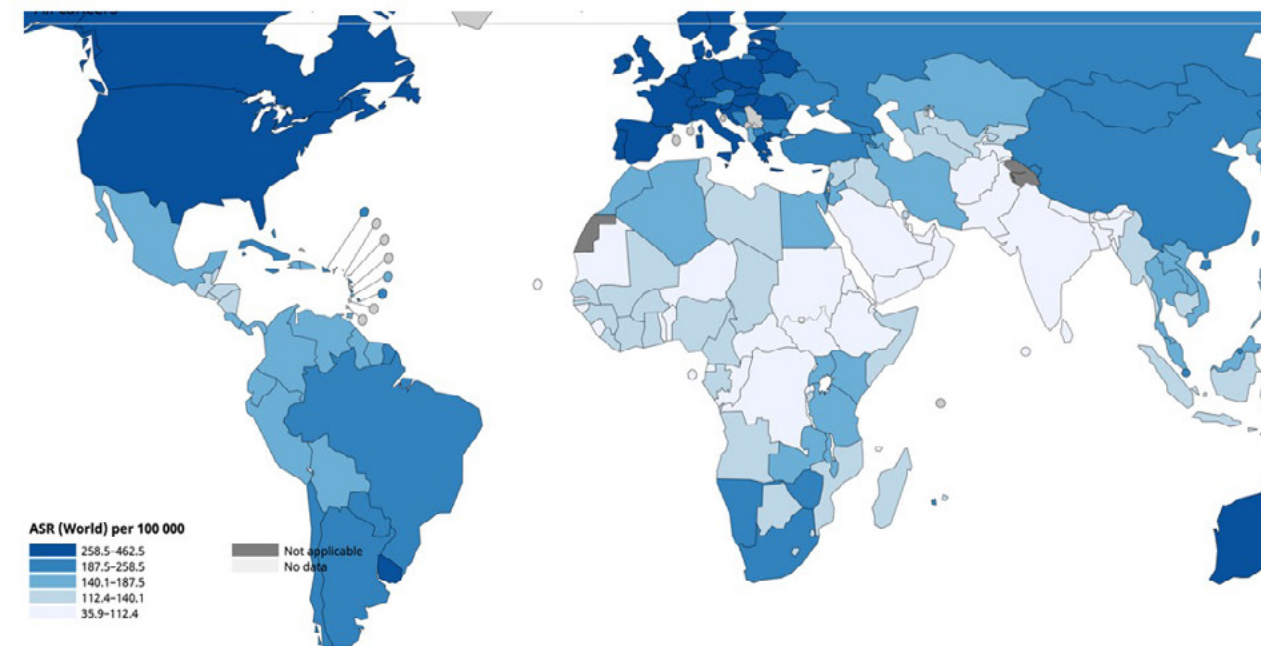
www.sanita2030.it





Cancer by the numbers (GLOBOCAN 2022)

International Agency for Research on Cancer
World Health Organization



- ~20M cases
 - 1 out of 4/5
 - ~10M deaths
- 2050 forecast:
 - *adjusted (aging, etc.)*
 - ~35M (+77% = 4 nations)
 - up to 1 over 2 (UK)
- Lung, breast, colon, prostate, ...

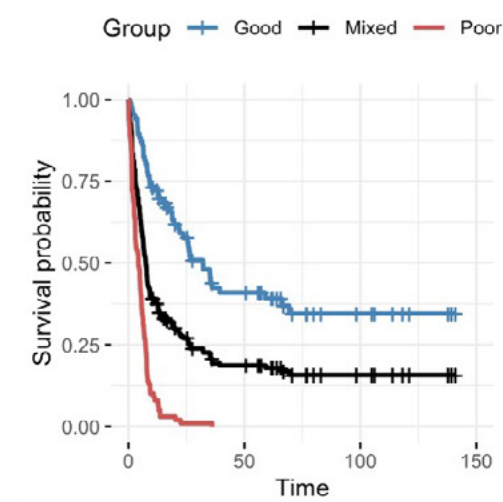
GLOBOCAN, WHO

#sanita2030



www.sanita2030.it

The classification dilemma

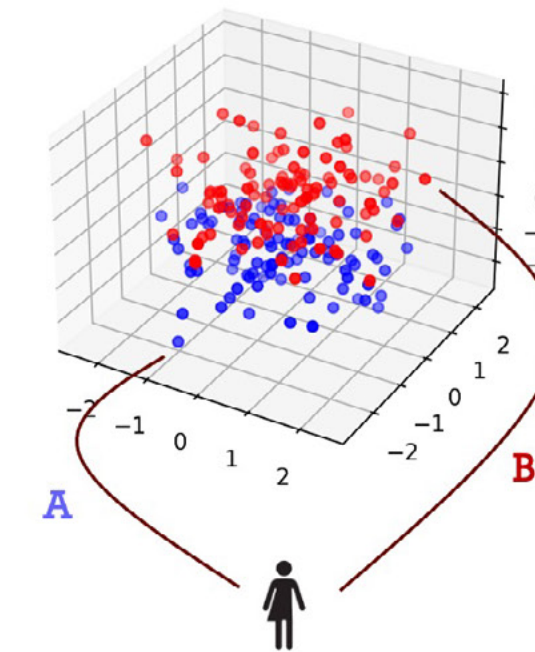


Classification task
Which one is ?

- **Subtype A** : worst prognosis, plan aggressive treatment
- **Subtype B** : better prognosis, avoid over-treatment

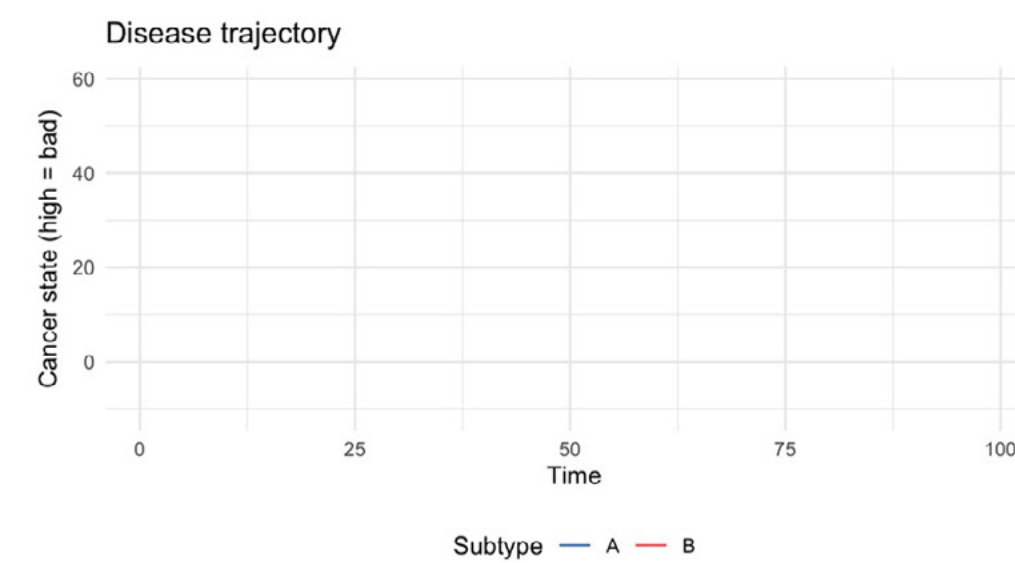
Cancer molecular data (complex)

~30.000 genes, ~1.000 mutations, ~10.000.000.000 cells

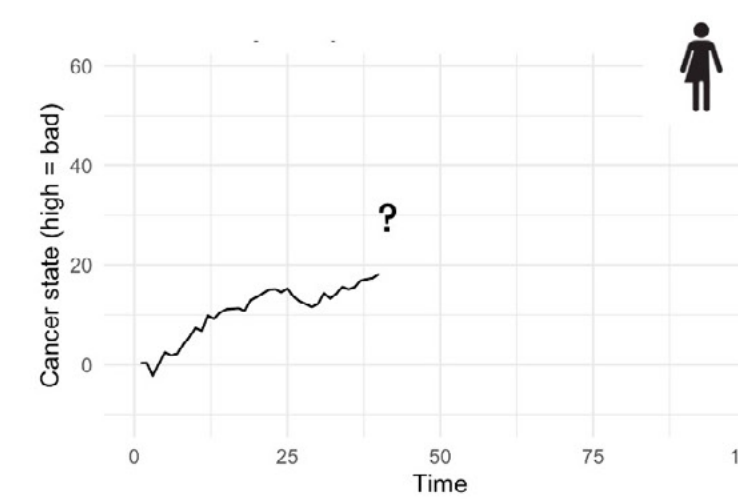


Dynamic disease → not just today, forecast tomorrow (*treatment response?*)

Heterogenous in genotype /phenotype spaces

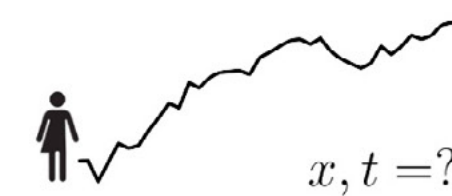
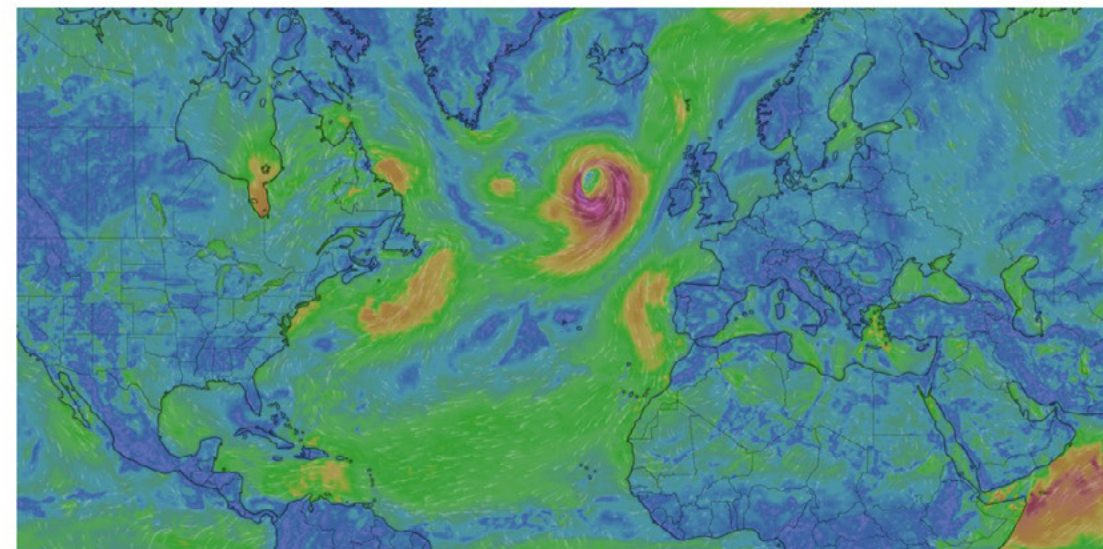


Hard cross-sectional (no temporal information)





It is a forecast problem



Variables at play:

- genetic (*germline and somatic*)
- environment (*mutagens like smoking etc*)
- treatments (*283 FDA-approved drugs,*)
- whole-life clinical history

First principles in biology: **unknown!**

$$\nabla^2 \omega + \frac{f_0^2}{\sigma} \frac{\partial^2 \omega}{\partial p^2} = \frac{f_0}{\sigma} \left(\frac{\partial}{\partial p} \left(\mathbf{V}_g \cdot \nabla \left(\frac{\partial \Phi}{\partial p} \right) \right) - \nabla^2 (\mathbf{V}_g \cdot \nabla \Phi) \right)$$

Windy.com

#sanita2030

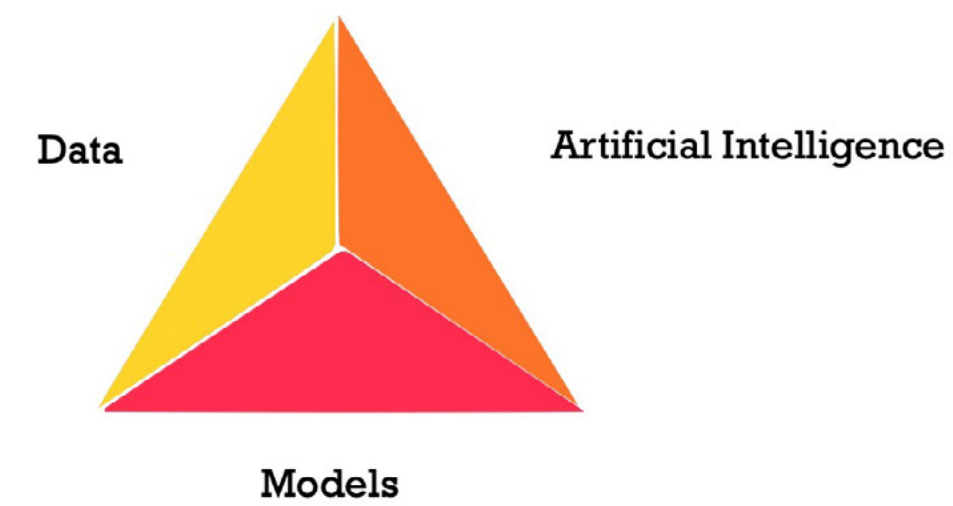


www.sanita2030.it



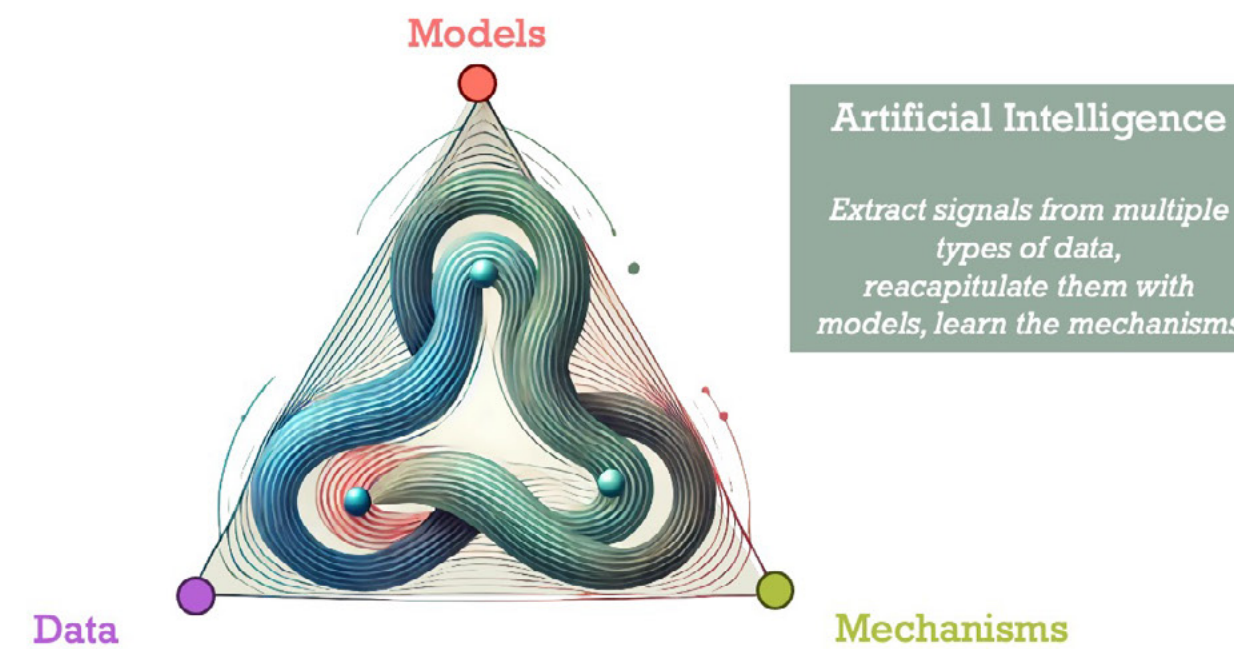


What I think we need

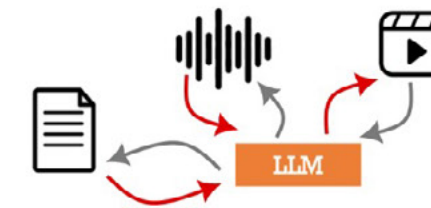
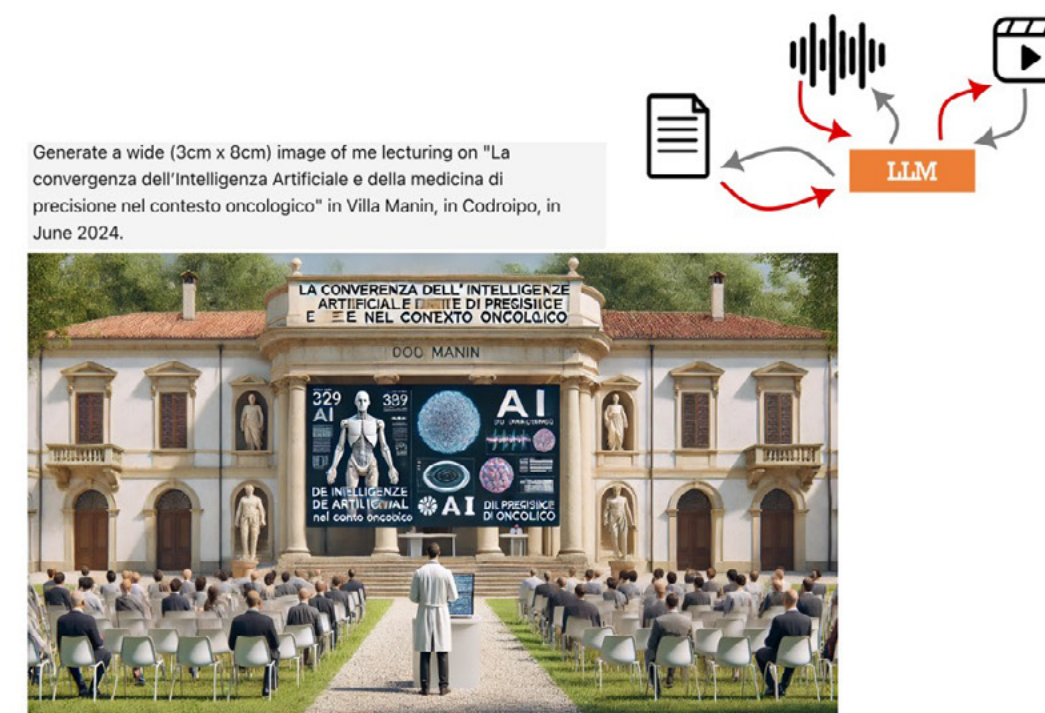
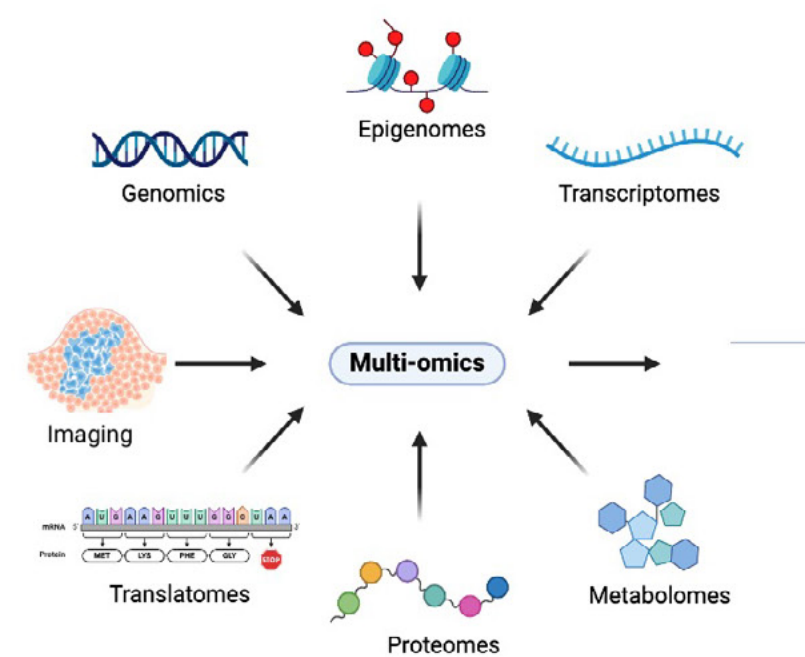




What I think we need



Large, Heterogenous, multi-modal datasets for multi-modal AI



Biorender /

#sanita2030



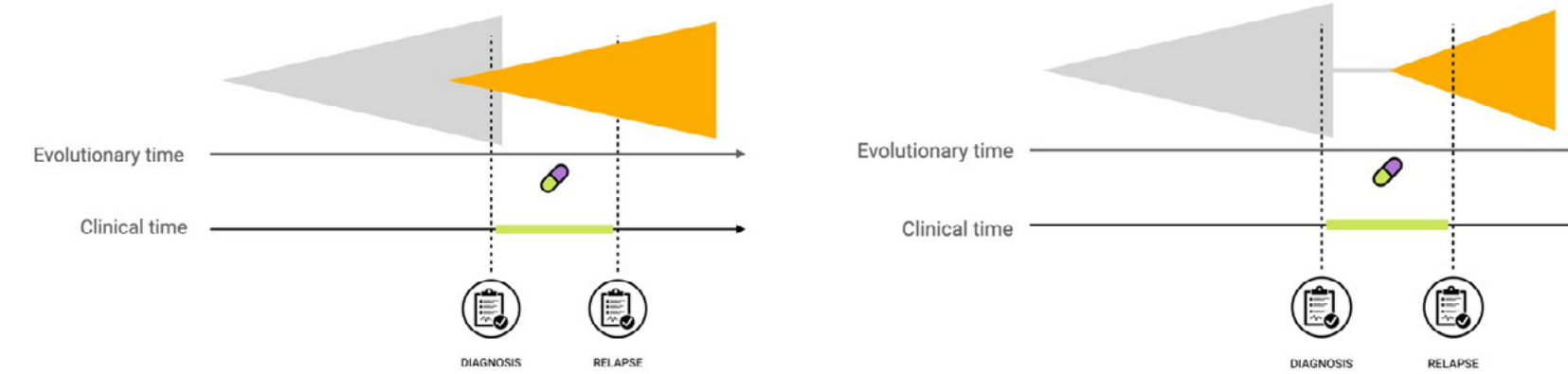
www.sanita2030.it

A pinch of our work in Trieste

Characterising genotype and phenotype clonal evolution of response to therapy with Artificial Intelligence



Pre-existing vs de novo resistance in Acute Myeloid Leukemia (with Vago/ Tonon, San Raffaele)



A pinch of our work in Trieste

Characterising genotype and phenotype clonal evolution of response to therapy with Artificial Intelligence



Pre-existing vs de novo resistance in Acute Myeloid Leukaemia (with Vago/ Tonon, San Raffaele Milan)

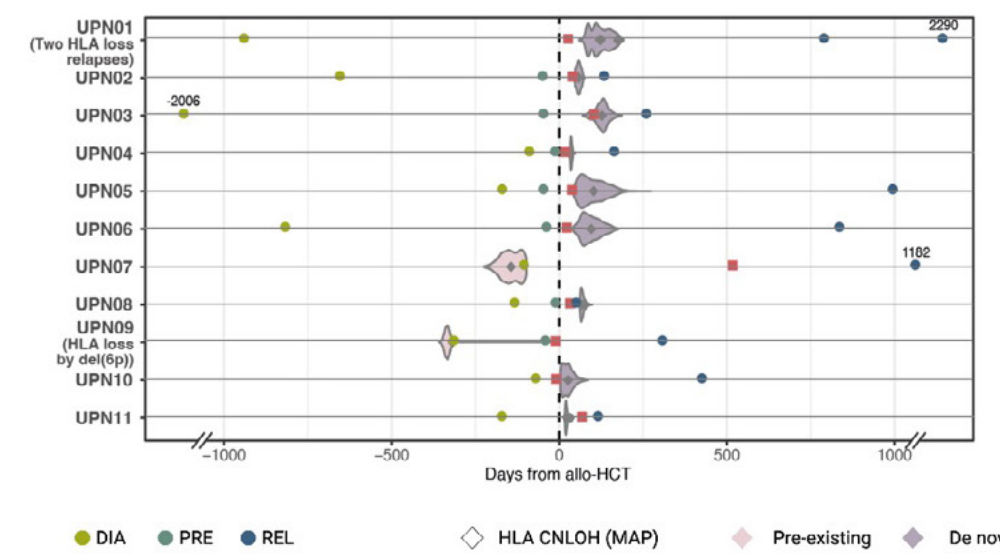
Clinical data

- Treatment types
- Treatment duration

Molecular data

- Whole-genome DNA sequencing
- Single-cell RNA sequencing

$$\mathcal{L}(m_{\text{dipl}}, m_{\text{gan}}, m_{\text{cyt}}, m_{\text{HLA}}) = \prod_{m_x} \text{Pois} \left(m_x \mid \omega \theta_x \int_{\Delta t} \mu_x(t) dt \right)$$



#sanita2030



www.sanita2030.it

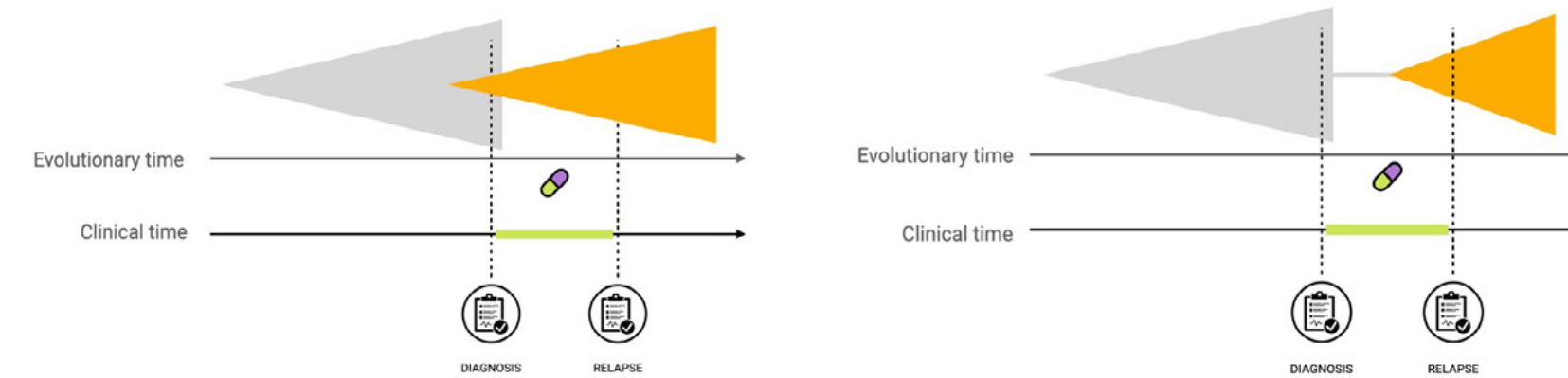


A pinch of our work in Trieste

Characterising genotype and phenotype clonal evolution of response to therapy with Artificial Intelligence



Plastic biomarkers discovery in Chronic Lymphocytic Leukaemia (with Gattei, CRO Aviano)



#sanita2030



www.sanita2030.it



A pinch of our work in Trieste

Characterising genotype and phenotype clonal evolution of response to therapy with Artificial Intelligence

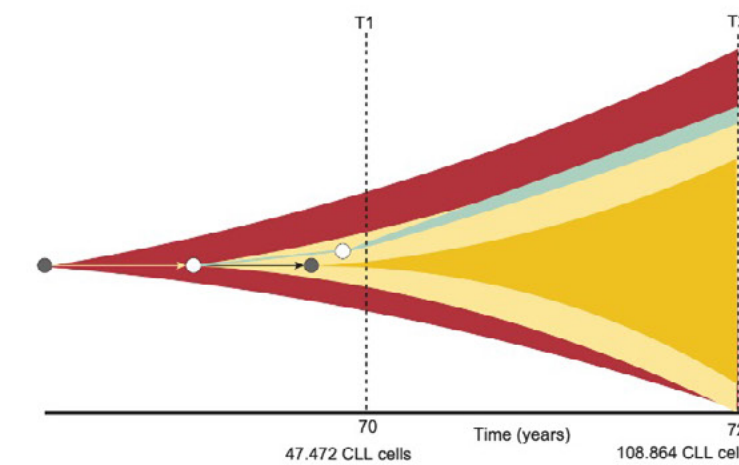


Plastic biomarkers discovery in Chronic Lymphocytic Leukaemia (with Gattei, CRO Aviano)

Molecular data

- Whole-genome DNA sequencing
- Single-cell DNA sequencing
- Single-cell RNA sequencing
- CD49d promoter methylation
- DNA ATAC sequencing

$$\mathbb{E}(n_{\oplus} | \lambda_{\oplus}, s, \omega_{\oplus}, \omega_{\oplus}, t_{\oplus}) \propto \exp \left[- \left(\lambda_{\oplus} (1 + s) + \frac{\omega_{\oplus} \omega_{\oplus}}{\lambda_{\oplus}^5} \right) (T - t_{\oplus}) \right]$$



#sanita2030

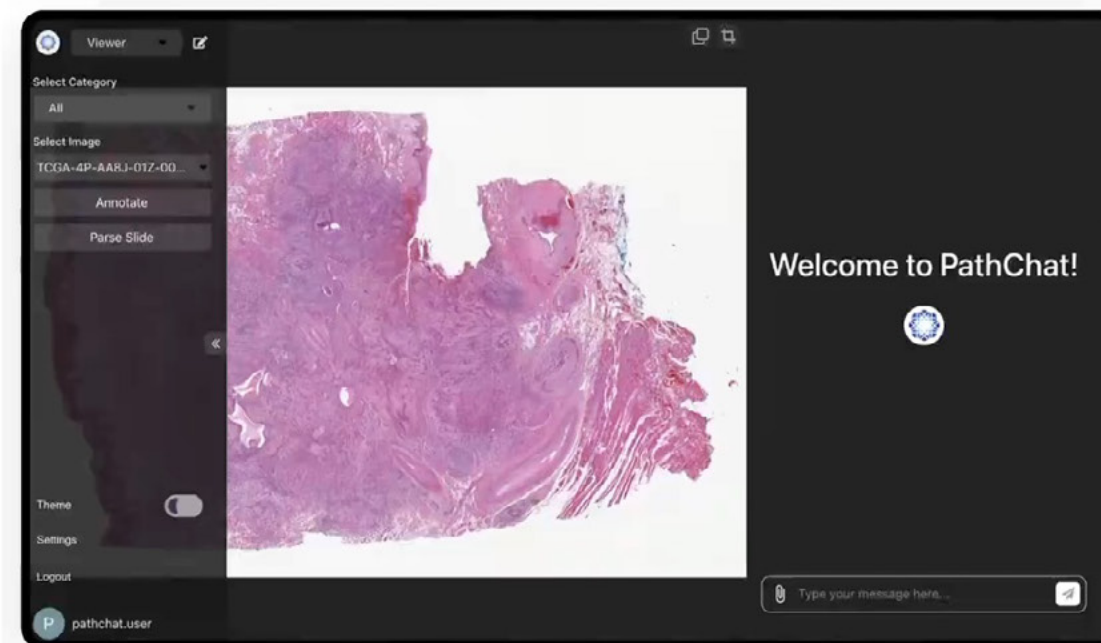


www.sanita2030.it



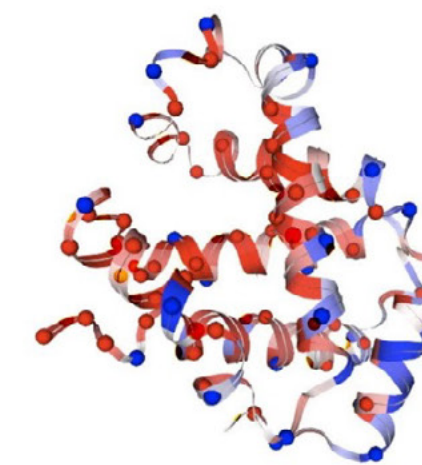


Popolarised AI for precision medicine



PMID: 38866050

#sanita2030



Hemoglobin subunit beta (HBB)

Google Deepmind

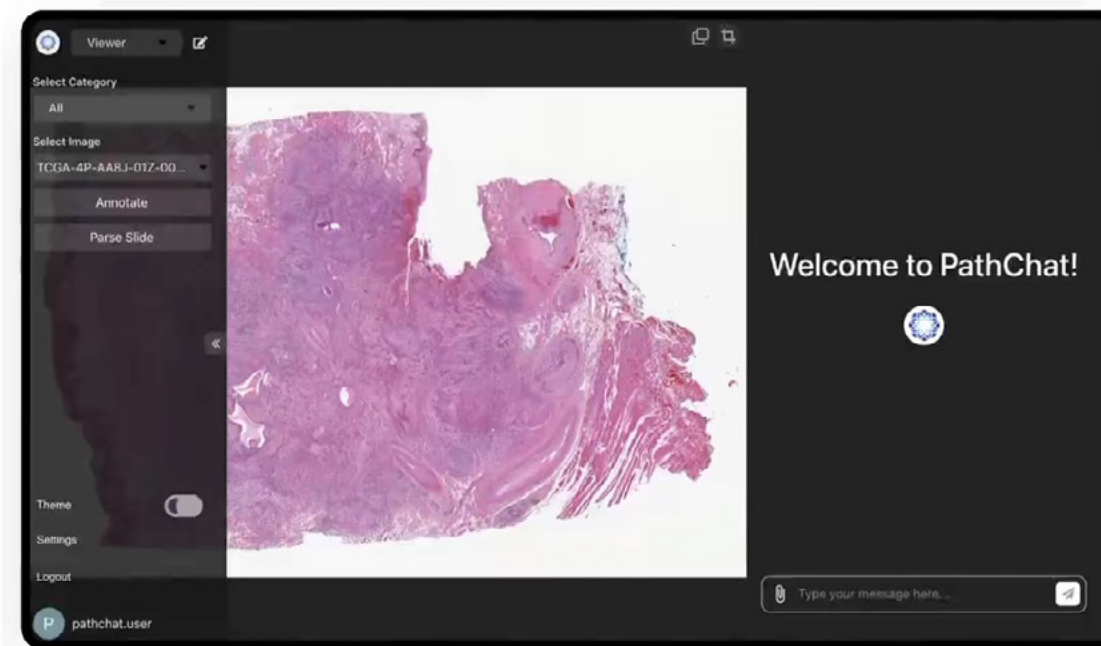


www.sanita2030.it





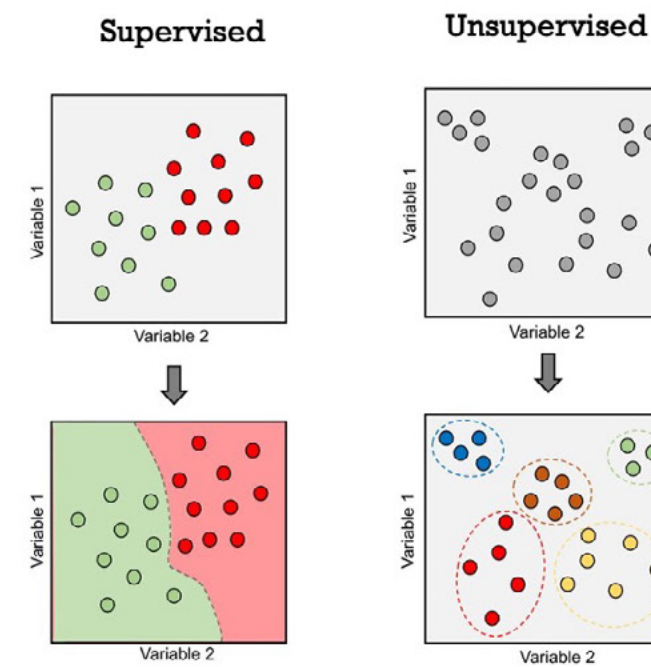
Popolarised AI for precision medicine



PMID: 38866050/ DOI: 10.1186/s12052-021-00147-x

#sanita2030

Imaging vs genomics

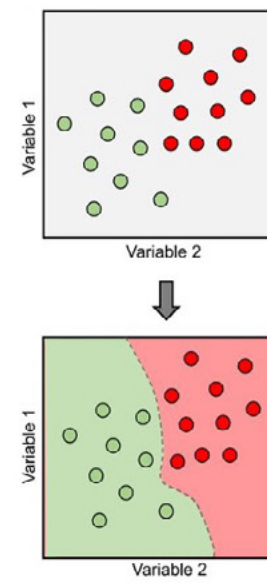


www.sanita2030.it

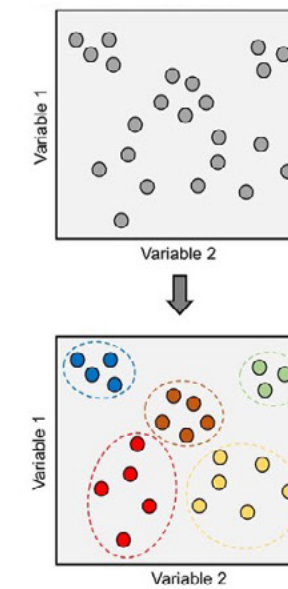


Imaging (example) vs genomics

Supervised



Unsupervised



PMID: 38866050

DOI: 10.1186/s12052-021-00147-x

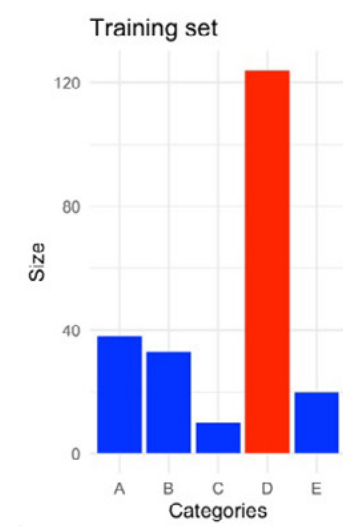
#sanita2030



www.sanita2030.it



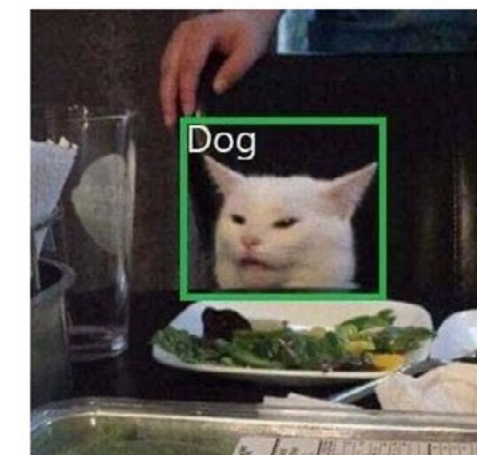
Ethics, accessibility and errors



Training set bias



The cost of large-scale AI



The cost of an error

Stanford University. AI Index report/ Reddit

#sanita2030



www.sanita2030.it

Thanks

Delitti in materia di violazione del diritto d'autore (Art. 25-novies, D.Lgs. n. 231/2001) [articolo aggiunto dalla L. n. 99/2009]

- Messa a disposizione del pubblico, in un sistema di reti telematiche, mediante connessioni di qualsiasi genere, di un'opera dell'ingegno protetta, o di parte di essa (art. 171, legge n.633/1941 comma 1 lett. a) bis)
- Reati di cui al punto precedente commessi su opere altrui non destinate alla pubblicazione qualora ne risulti offeso l'onore o la reputazione (art. 171, legge n.633/1941 comma 3)
- Abusiva duplicazione, per trarne profitto, di programmi per elaboratore; importazione, distribuzione, vendita o detenzione a scopo commerciale o imprenditoriale o concessione in locazione di programmi contenuti in supporti non contrassegnati dalla SIAE; predisposizione di mezzi per rimuovere o eludere i dispositivi di protezione di programmi per elaboratori (art. 171-bis legge n.633/1941 comma 1)
- Riproduzione, trasferimento su altro supporto, distribuzione, comunicazione, presentazione o dimostrazione in pubblico, del contenuto di una banca dati; estrazione o reimpiego della banca dati; distribuzione, vendita o concessione in locazione di banche di dati (art. 171-bis legge n.633/1941 comma 2)
- Abusiva duplicazione, riproduzione, trasmissione o diffusione in pubblico con qualsiasi procedimento, in tutto o in parte, di opere dell'ingegno destinate al circuito televisivo, cinematografico, della vendita o del noleggio di dischi, nastri o supporti analoghi o ogni altro supporto contenente fonogrammi o videogrammi di opere musicali, cinematografiche o audiovisive assimilate o sequenze di immagini in movimento; opere letterarie, drammatiche, scientifiche o didattiche, musicali o drammatico musicali, multimediali, anche se inserite in opere collettive o composite o banche dati; riproduzione, duplicazione, trasmissione o diffusione abusiva, vendita o commercio, cessione a qualsiasi titolo o importazione abusiva di oltre cinquanta copie o esemplari di opere tutelate dal diritto d'autore e da diritti connessi; immissione in un sistema di reti telematiche, mediante connessioni di qualsiasi genere, di un'opera dell'ingegno protetta dal diritto d'autore, o parte di essa (art. 171-ter legge n.633/1941)
- Mancata comunicazione alla SIAE dei dati di identificazione dei supporti non soggetti al contrassegno o falsa dichiarazione (art. 171-septies legge n.633/1941)
- Fraudolenta produzione, vendita, importazione, promozione, installazione, modifica, utilizzo per uso pubblico e privato di apparati o parti di apparati atti alla decodificazione di trasmissioni audiovisive ad accesso condizionato effettuate via etere, via satellite, via cavo, in forma sia analogica sia digitale (art. 171-octies legge n.633/1941).

[Torna all'inizio](#)