



**UNIVERSITÀ
DEL SALENTO**

Il Multiverso della governance dei dati digitali nella sperimentazione clinica: dati sintetici open access?

Marco Mancarella

Professore Associato di Informatica giuridica – UniSalento
Direttivo Centro Data Science - UniSalento
Direttivo Centro di ricerca Interuniversitario di Informatica Giuridica
Direttivo Associazione Nazionale Docenti di Informatica Giuridica
Componente @LawLab – LUISS Guido Carli di Roma
Fondatore LiquidLaw srl – Azienda spinoff UniSalento
Avvocato esperto in Diritto delle nuove tecnologie



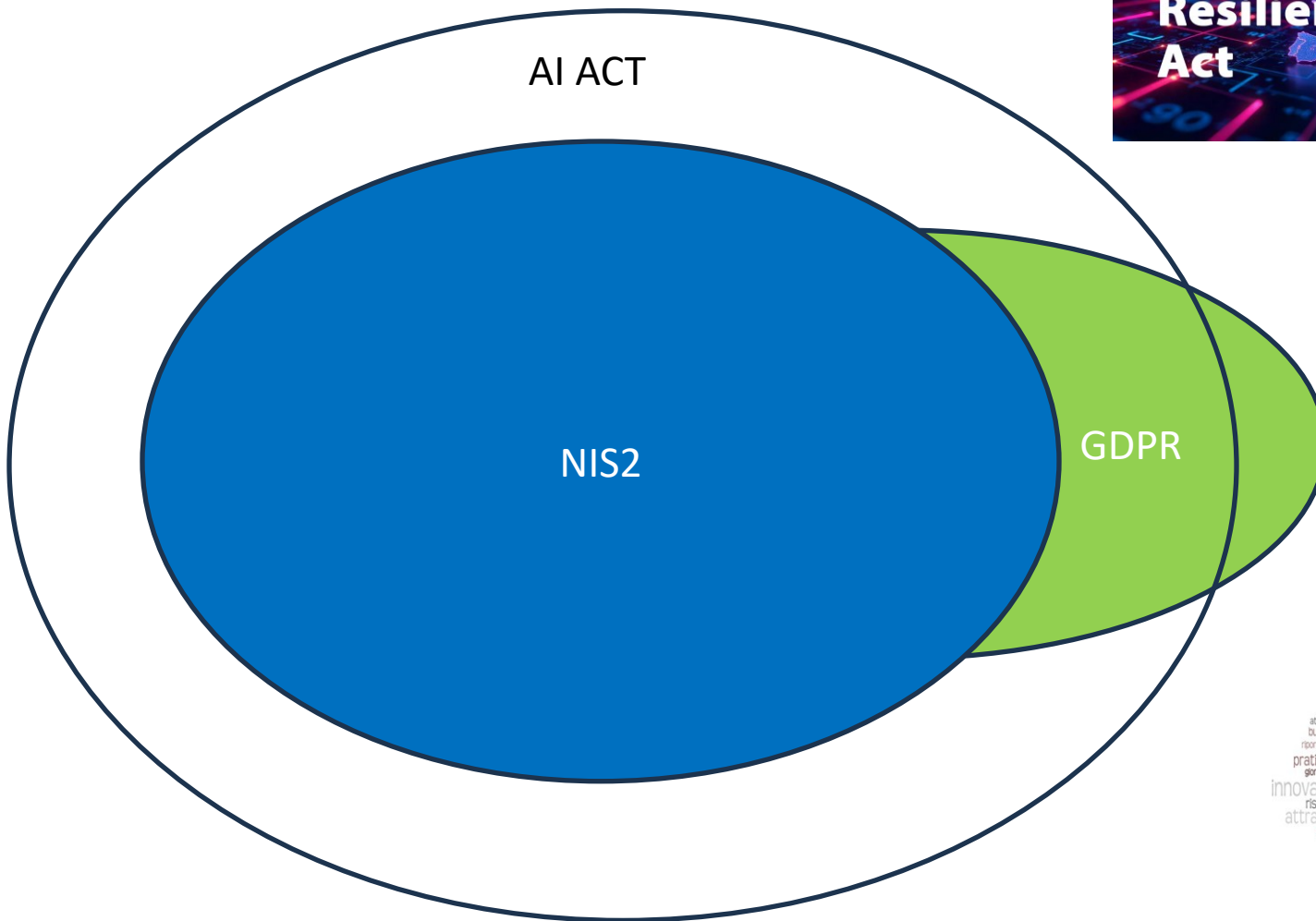
spin-off
UniSalento
2018-22







ALCUNI STRATI DEL MULTIVERSO NORMATIVO...OGGI





CRITICITA'



- **GDPR: DPO**
- **Rischio NIS2: PDC, Referente CSIRT, CISO...**
- **AI ACT: RTD?**





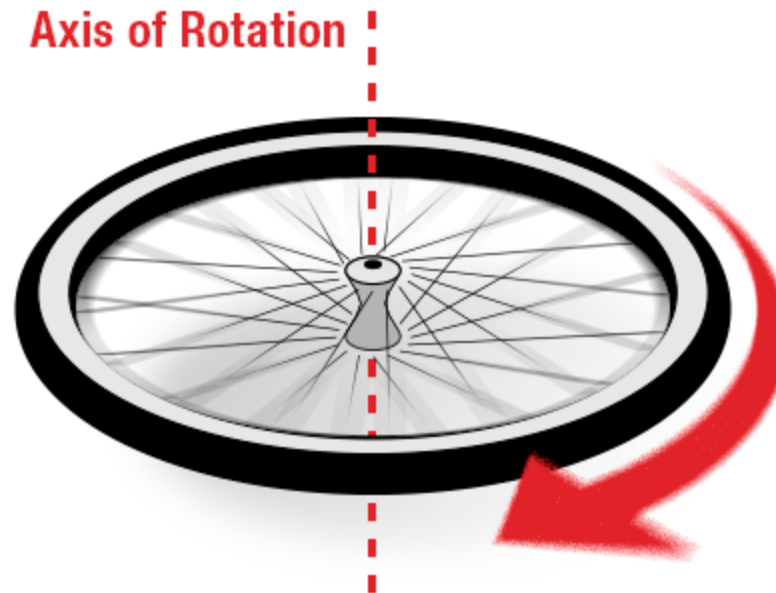
SENSAZIONI





Da dove partiamo?

«dignità umana»







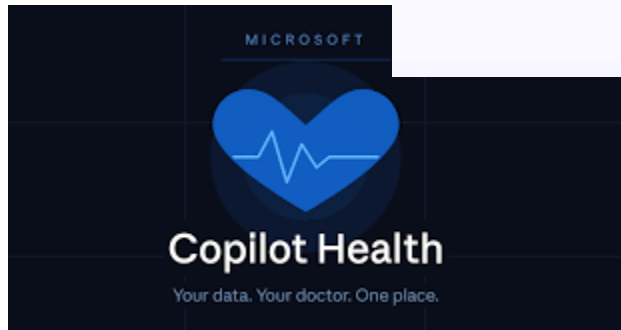
report "**Generative AI and synthetic data for clinical application**", presentato pochi giorni fa da **NTT DATA**, leader mondiale nei servizi tecnologici e business digitale, e **Train**, spin-off dell'IRCCS Istituto Clinico Humanitas specializzato nello sviluppo dell'AI generativa

entro il 2035, il 60% dei bisogni clinici della popolazione saranno quesiti clinici complessi, molto più di adesso e quasi ogni giorno si scoprono malattie rare o si trovano fenotipi diversi all'interno della stessa patologia





06 MARCH 2026
**AWS Launches
 Amazon Connect
 Health to Expand AI
 Agent Capabilities in
 Healthcare**







E se puntissimo sui dati sintetici?

Il dato sintetico **non è il dato anonimizzato** (il dato al quale si tolgono le informazioni sensibili e si lasciano solo quelle sanitarie, ma il dato rimane unico e riferibile a un solo paziente), **non è nemmeno il gemello digitale** (una rappresentazione virtuale di un sistema fisico, progettata per riflettere accuratamente le caratteristiche e il comportamento dell'entità reale). Il dato sintetico è **un dato artificiale, generato attraverso modelli matematici, che imita i dati reali**



vantaggi sostanziali:

- possono essere **generati in grandi quantità secondo necessità**, fornendo una preziosa fonte per analisi e addestramento di modelli di machine learning
- garantiscono la **protezione della privacy (a valle)** non contenendo informazioni personali
- possono essere **progettati specificamente per ridurre i bias** presenti nei dati reali **(evitando di ereditarli)**
- **ma trattamento sempre informato (realmente informato)!**



AI Act riconosce i **dati sintetici** come **risorsa strategica** proprio per consentire la conformità in tema di privacy e governance

Miglioramento e mitigazione dei rischi: I dati sintetici sono riconosciuti come strumenti per migliorare la qualità dei dataset di addestramento, correggere distorsioni (*bias*) e mitigare i rischi legati al trattamento dei dati personali

Alternative ai dati reali: Nei contesti di sistemi di IA ad *alto rischio*, *l'art. 10 dell'AI Act consente e incentiva l'uso di dati sintetici o anonimizzati per l'addestramento e il collaudo*, riducendo la necessità di utilizzare dati personali sensibili

Compliance e Privacy: L'uso di dati sintetici facilita la conformità al GDPR

Sviluppo nei Sandbox: L'AI Act promuove l'utilizzo di dati sintetici all'interno dei *Regulatory Sandbox* (ambienti di sperimentazione normativa) per permettere alle aziende di testare sistemi di IA in un ambiente sicuro e regolamentato



E in Italia?



Legge 132/2025

Art. 8 - Ricerca e sperimentazione scientifica nella realizzazione di sistemi di intelligenza artificiale in ambito sanitario:

individua un perimetro di liceità specifico per i trattamenti di dati, anche appartenenti alle **categorie particolari** di cui all'articolo 9 del Regolamento (UE) 2016/679, **finalizzati alla ricerca e alla sperimentazione scientifica nella realizzazione di sistemi di intelligenza artificiale per scopi terapeutici e farmacologici:**

dichiara di **rilevante interesse pubblico** tali trattamenti ove necessari alla costituzione e all'utilizzazione di banche dati e di modelli di base, con ciò approntando la **base giuridica nazionale**



Legge 132/2025

Art. 8 - Ricerca e sperimentazione scientifica nella realizzazione di sistemi di intelligenza artificiale in ambito sanitario:

I trattamenti devono essere svolti da **enti pubblici e privati senza scopo di lucro**, dagli Istituti di ricovero e cura a carattere scientifico, ovvero da soggetti privati operanti nel settore sanitario nell'ambito di progetti di ricerca cui partecipino soggetti pubblici o privati **no profit** o **IRCCS**

circuito virtuoso di cooperazione pubblico-privato ancorato a finalità scientifiche

esclusione finalità meramente commerciali



Legge 132/2025

Art. 8 - Ricerca e sperimentazione scientifica nella realizzazione di sistemi di intelligenza artificiale in ambito sanitario:

È **sempre consentito**, previa **informativa all'interessato** (sito web del titolare), il trattamento volto:

- all'anonimizzazione
- alla pseudonimizzazione
- alla **sintetizzazione dei dati sanitari e delle ulteriori categorie particolari**

sia per gli scopi di **ricerca** sia per finalità di programmazione, gestione, controllo e valutazione dell'assistenza sanitaria



Legge 132/2025

Art. 8 - Ricerca e sperimentazione scientifica nella realizzazione di sistemi di intelligenza artificiale in ambito sanitario:

possibile adozione di **linee guida per le procedure di anonimizzazione e la creazione di dati sintetici**, **demandate all'Agenzia nazionale per i servizi sanitari regionali, previo parere del Garante**, nel rispetto degli standard internazionali e dello stato dell'arte

la legge introduce un meccanismo di **comunicazione preventiva al Garante** per la protezione dei dati personali per i trattamenti (possono iniziare decorsi trenta giorni dalla comunicazione, salvo blocco da parte dell'Autorità)



Legge 132/2025

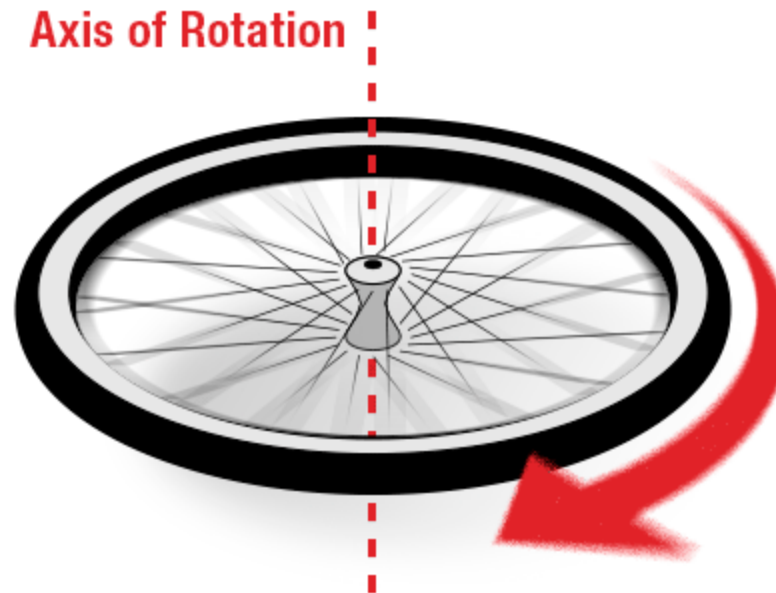
Art. 9 - Disposizioni in materia di trattamento di dati personali:

la norma prevede che, entro centoventi giorni dall'entrata in vigore della legge, il **Ministro della salute**, sentiti il Garante per la protezione dei dati personali, gli enti di ricerca, le strutture sanitarie, le autorità competenti e gli operatori del settore, adotti un **decreto** finalizzato a definire le **modalità operative del trattamento dei dati**

con il massimo delle modalità semplificate consentite dal GDPR per finalità di **ricerca e sperimentazione** anche tramite sistemi di intelligenza artificiale e **machine learning**, inclusi la costituzione e l'utilizzo di spazi speciali di sperimentazione a fini di ricerca



«dignità umana»







Thank you

Marco Mancarella

marco.mancarella@unisalento.it



**UNIVERSITÀ
DEL SALENTO**



spin-off
UniSalento
2018-22