

Privacy & I.A. UTOPIA?

Privacy Day Forum
Pisa | 25 maggio 2023

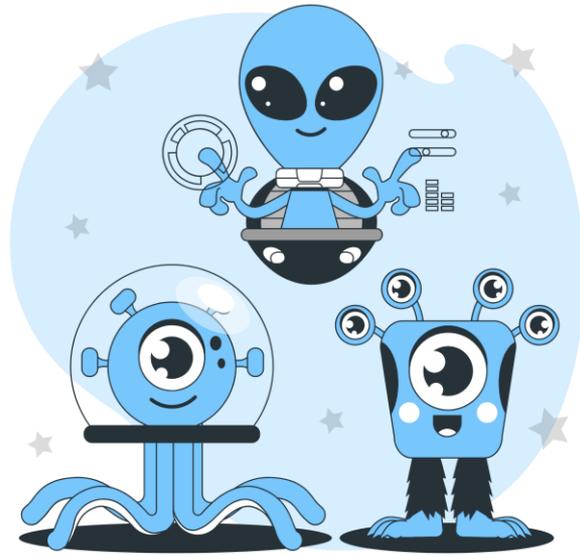
Diego Maranini

UTOPIA

the software

Dilaxia S.p.A.





Customer Experience

Aiuto i clienti a sfruttare al massimo i nostri servizi e
a tradurre le loro esigenze in righe di codice per il
team di sviluppo.

...

Contenuti

01

Tipologie di I.A.

...

02

BIAS ed effetti

...

03

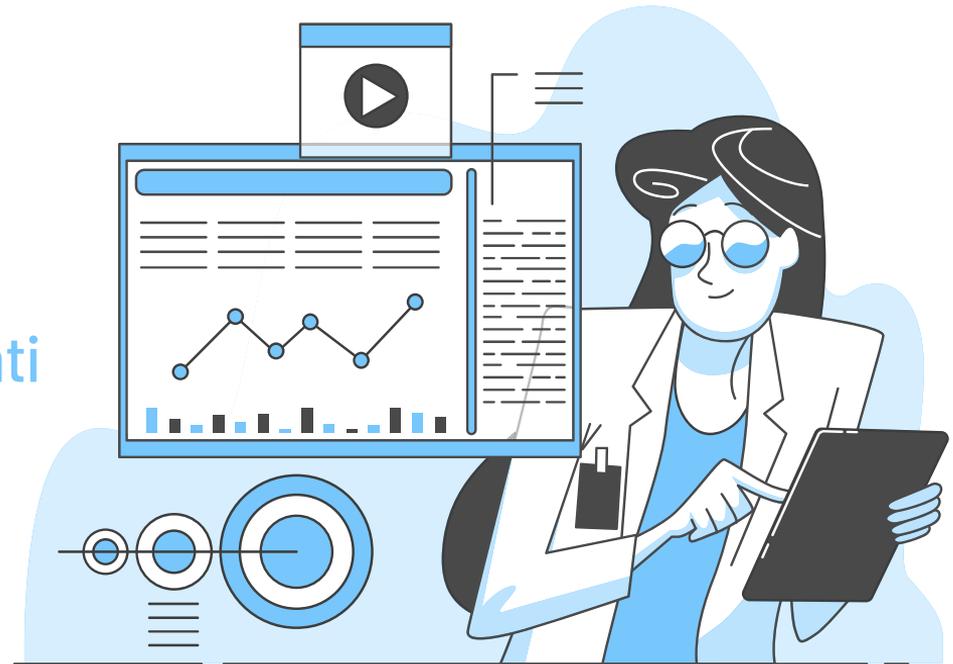
Privacy e adempimenti

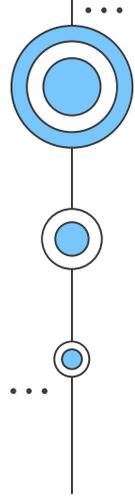
...

04

I.A. per la Privacy

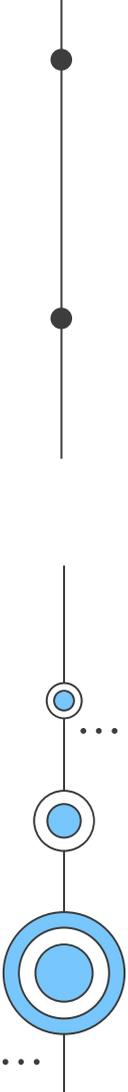
...





01

Intelligenza



Tipologie di “intelligenza”



ANI

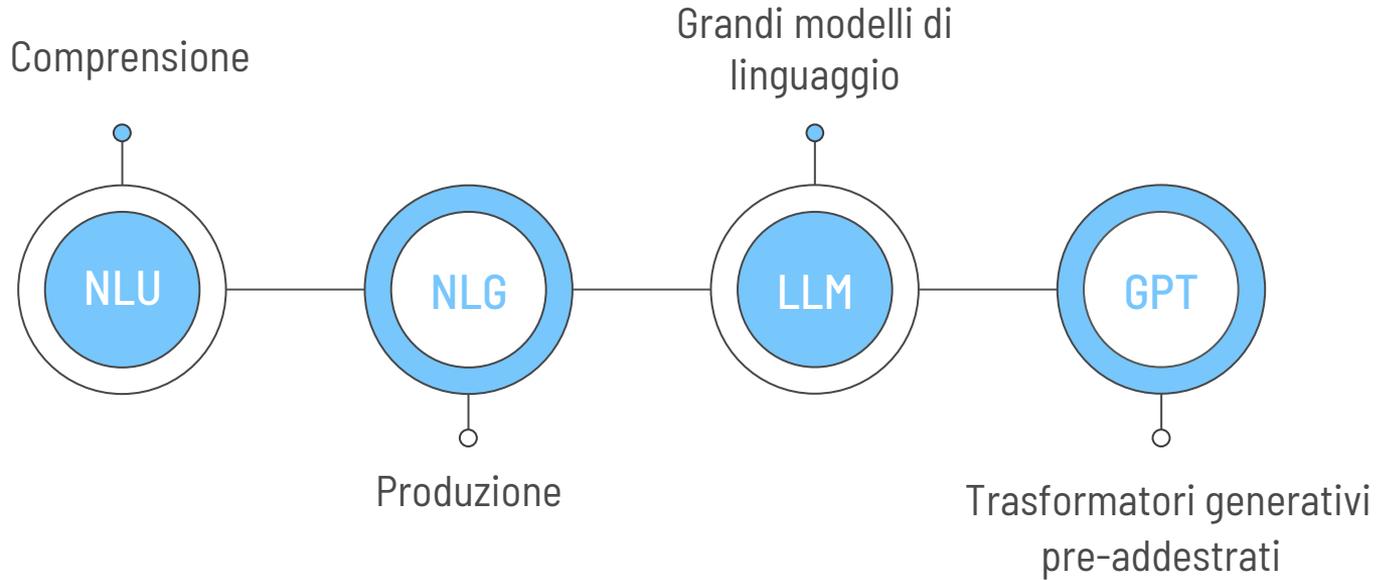


AGI



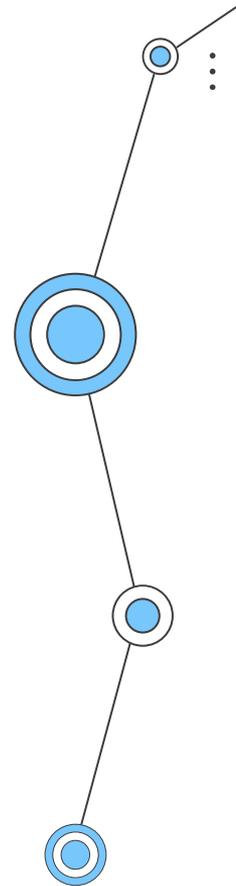
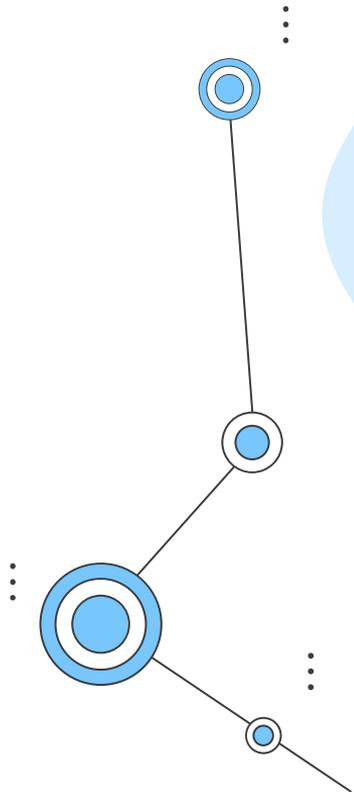
ASI

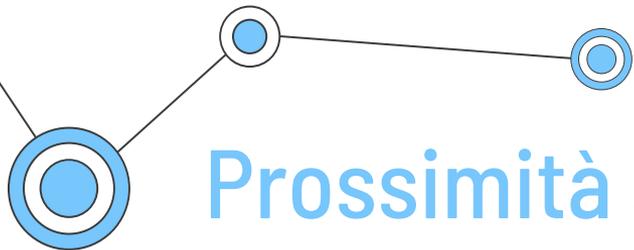
Nulla di più vicino





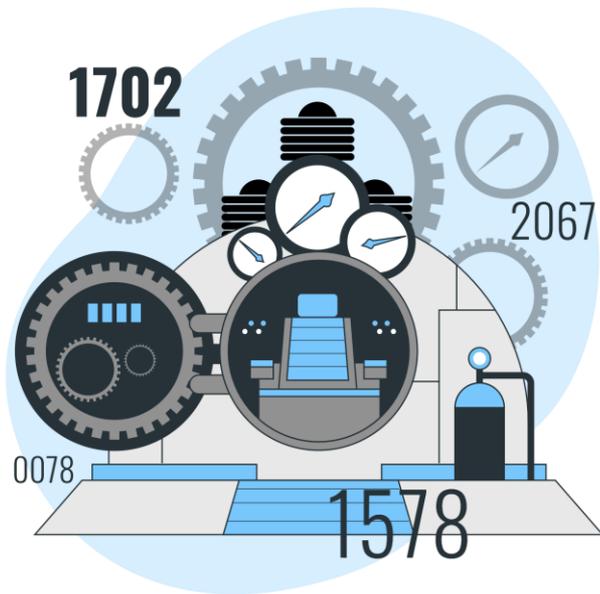
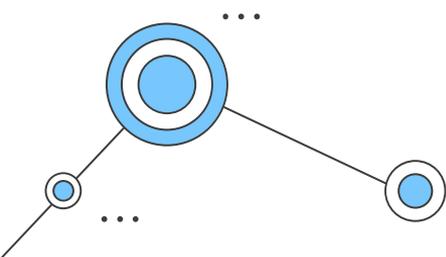
1 trillione





Prossimità statistica

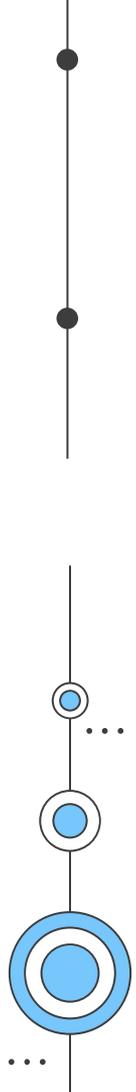
Le parole scelte dalla prossimità statistica però potrebbero **non corrispondere** a situazioni reali, ciò che gli scienziati dei dati chiamano il «*problema della messa a terra*».

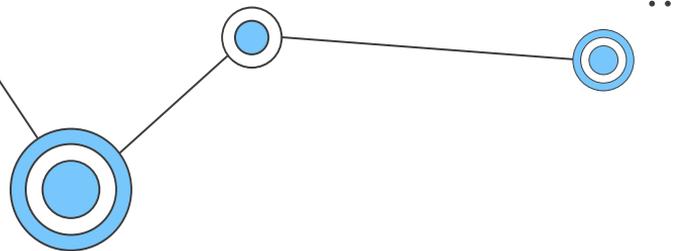




02

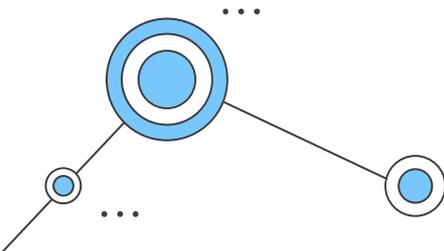
Effetti

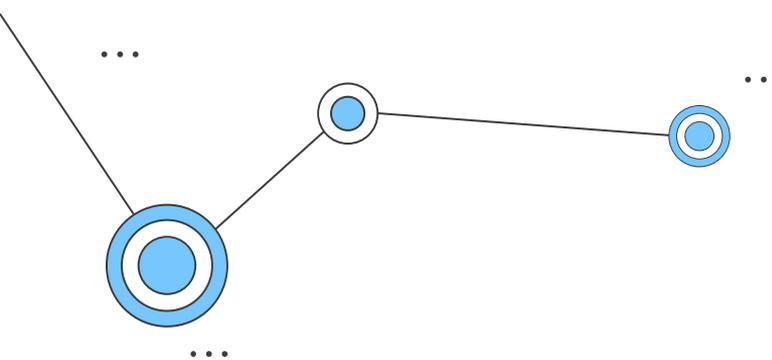




Un algoritmo di intelligenza artificiale basato su un approccio *data-driven* non predice il futuro, ma codifica il passato.

– Gianmarco Genalti





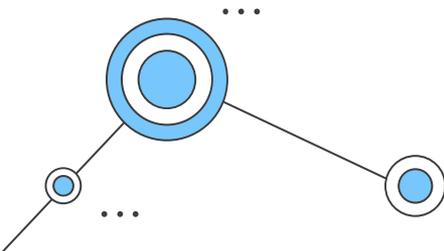
ChatGPT **PLUS**

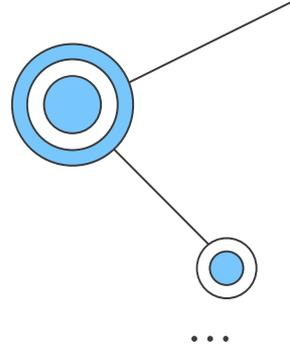
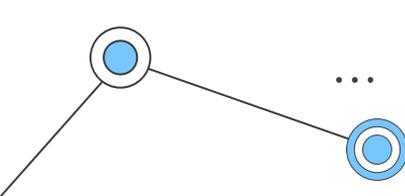
GPT-4 currently has a cap of 25 messages every 3 hours.

Send a message.



ChatGPT may produce inaccurate information about people, places, or facts. [ChatGPT May 12 Version](#)





Comprendere i BIAS

01

Campione

Il set di dati utilizzato è parziale o errato.

02

Presentazione

Quando un modello favorisce determinati risultati rispetto ad altri.

03

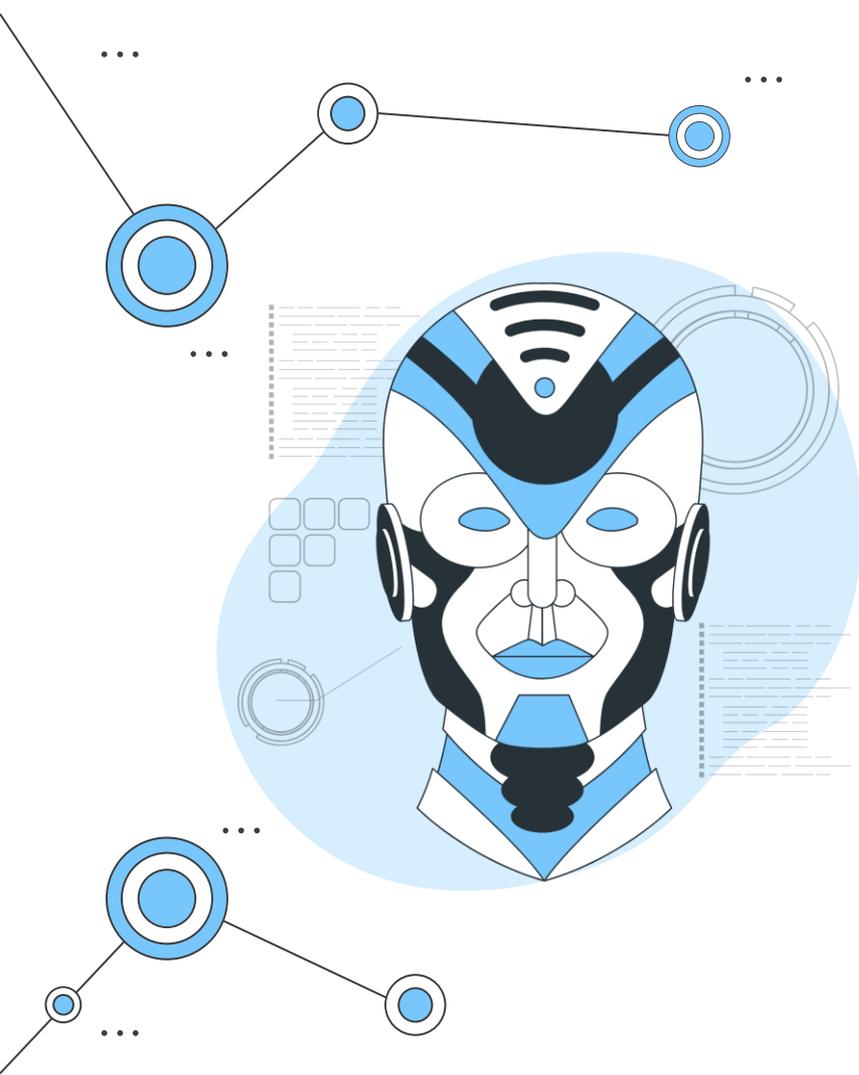
Pregiudizi

Bias provenienti spesso dagli umani, anche involontariamente.

04

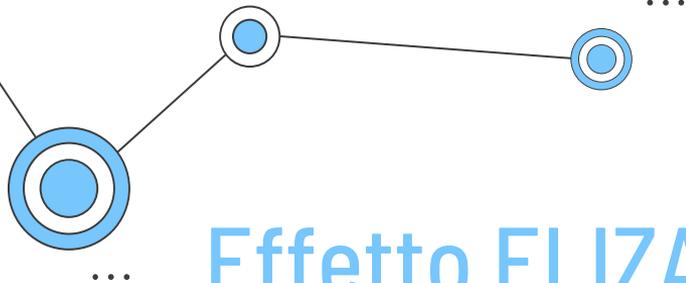
Discriminazioni

Il set di dati usato per l'addestramento non è rappresentativo.



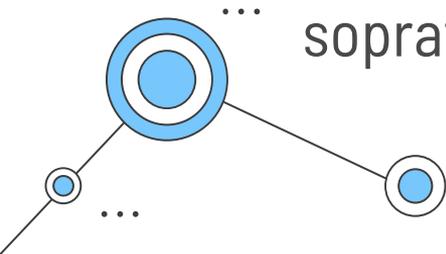
Le macchine non capiscono i BIAS. Le ipotesi errate spesso formulate durante la selezione dei set di dati e degli algoritmi, consapevolmente o inconsapevolmente, portano a BIAS dell'algoritmo.

– Anu Ganesan

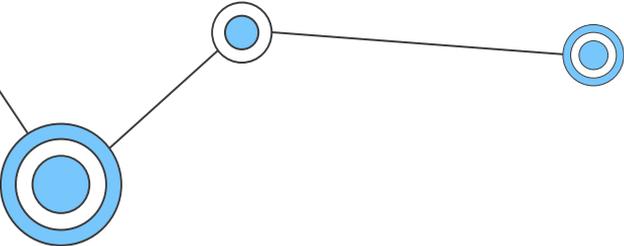


Effetto ELIZA (falsa intimità)

Le persone **proiettano qualità umane** come emozioni e comprensione sui modelli di intelligenza artificiale, inducendole a pensare che questi siano affidabili oracoli capaci di prendere decisioni. Questo può essere fuorviante e persino pericoloso, soprattutto in caso di persone vulnerabili.



...



ChatGPT è uno "stagista onnisciente", desideroso di compiacere, che a volte mente".

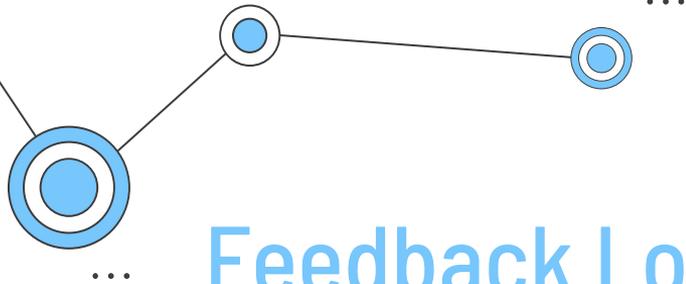
– Ethan Mollick





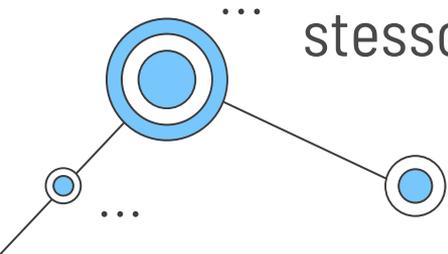
**Siamo a bordo delle
montagne russe statistiche
del significato.**

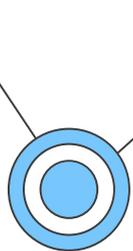
Leif Weatherby



Feedback Loops

La proposta fatta dall'algoritmo condiziona l'utente e non gli permette di valutare le alternative. Il dato raccolto però sarà condizionato da ciò, portandolo a un loop dove i nuovi dati che osserverà saranno frutto del suo stesso comportamento.





Research ▾

Product ▾

Developers ▾

Safety

Company ▾

May 18, 2023

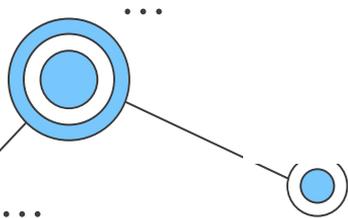
[Product, Announcements](#)

Since the release of ChatGPT, we've heard from users that they love using ChatGPT on the go. Today, we're launching the ChatGPT app for iOS.

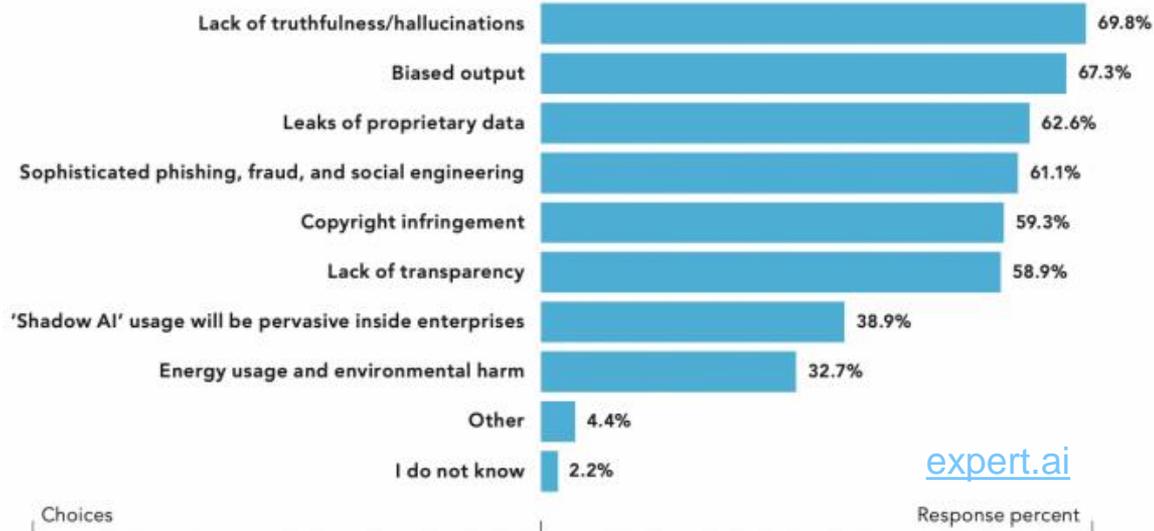
The ChatGPT app is free to use and syncs your history across devices. It also integrates [Whisper](#), our open-source speech-recognition system, enabling voice input. [ChatGPT Plus subscribers](#) get exclusive access to [GPT-4's capabilities](#), early access to features and faster response times, all on iOS.

Discover the versatility of ChatGPT:

- Instant answers: Get precise information without sifting through ads or multiple results.



ESG Risks with Generative AI and LLMs



"What environmental, social, and governance risks are created by the adoption of generative AI and Large Language Models? (Select all that apply)"

I.A. responsabile



Completezza

Nessun modello si adatta a tutti gli scenari



Allenamento

L'attività di training deve essere governata



Uomo nel mezzo

Ricerca costante dei BIAS



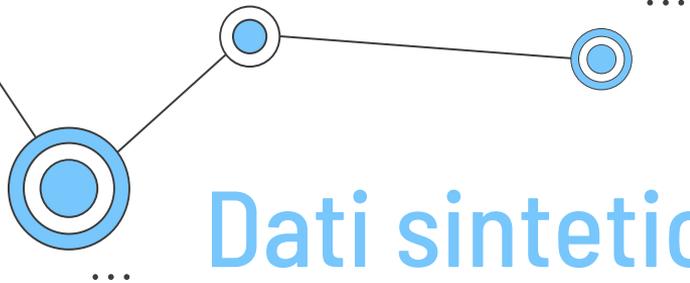
Dati reali

Testare con i dati del mondo reale



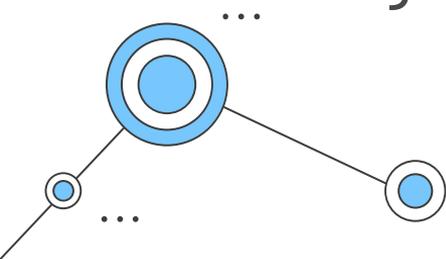
Monitoraggio

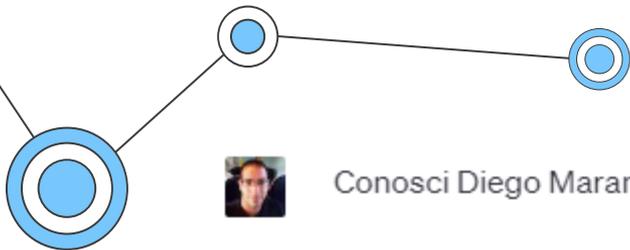
Controllo degli output e delle variabili



Dati sintetici

I dati sintetici sono dati generati artificialmente attraverso l'utilizzo di algoritmi e tecniche di sintesi, creati con lo scopo di simulare dati reali e allenare l'algoritmo.

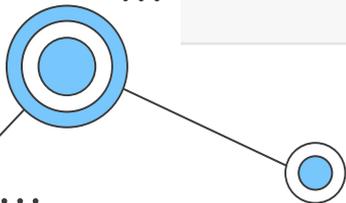


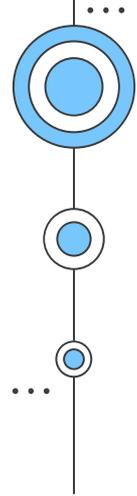


Conosci Diego Maranini?



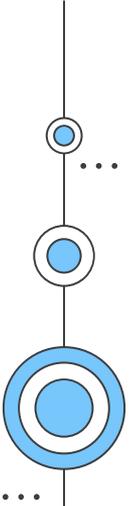
Mi dispiace, ma non ho informazioni su un individuo chiamato Diego Maranini nel mio database. La mia formazione si basa su un dataset che includeva informazioni pubblicamente disponibili fino a settembre 2021. Se Diego Maranini è diventato noto dopo questa data o se le sue informazioni non erano pubblicamente disponibili, non avrò dettagli da fornirti. Per favore, considera che sono un modello AI progettato per rispettare la privacy degli individui e non ho accesso a informazioni personali a meno che non siano state condivise con me durante la nostra conversazione.

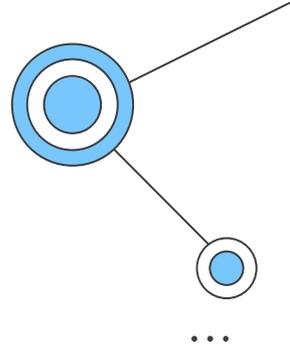
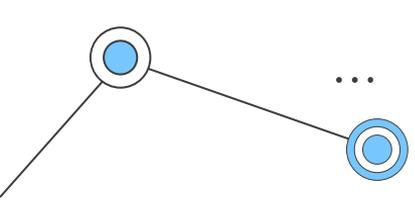




03

Protezione





Privacy

01

Cambio delle
finalità

02

Necessità

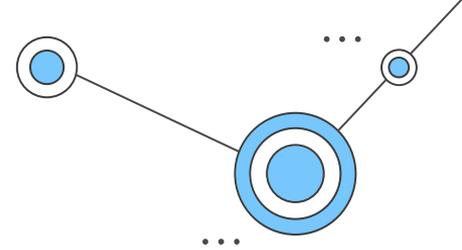
03

Limitazione

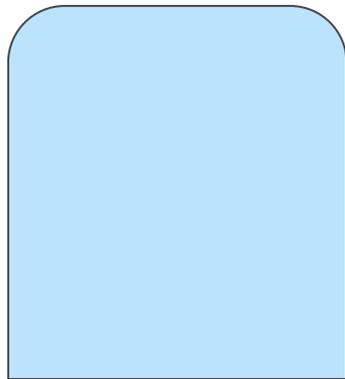
04

Ruoli e
attribuzioni

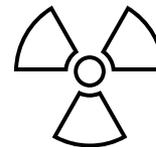
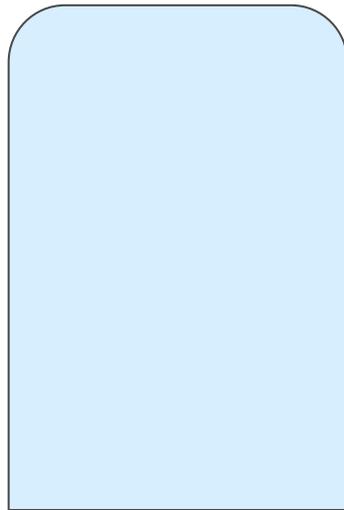
Regolamento I.A.



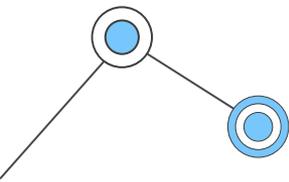
Rischio alto

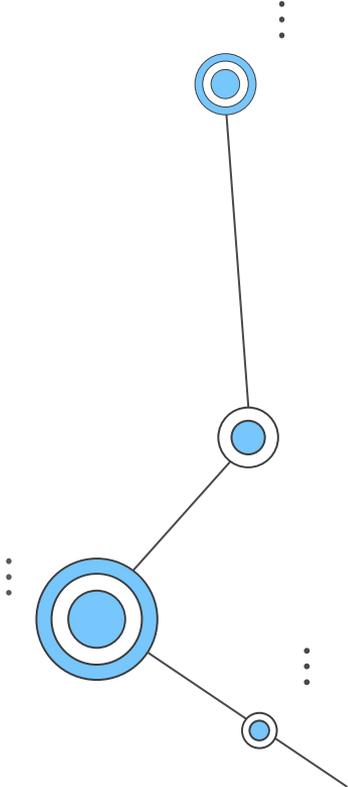


Rischiosissime



Rischio basso

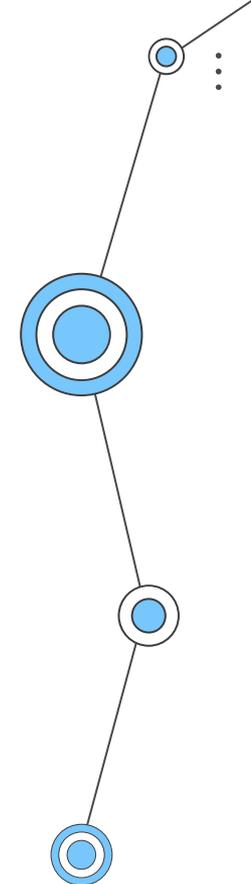




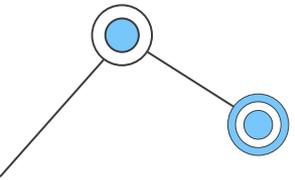
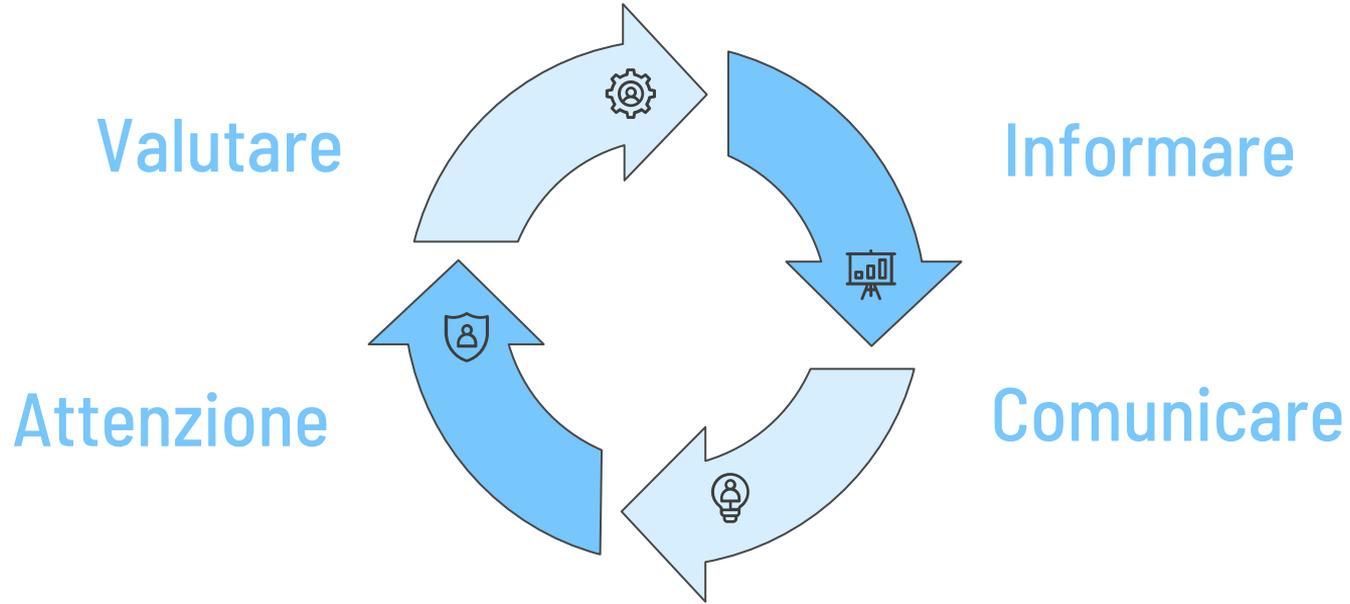
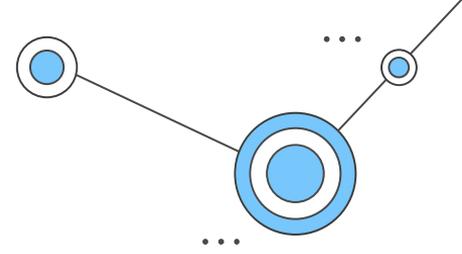
Articolo 22 GDPR

Considerando 71

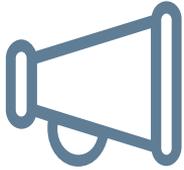
D.Lgs. 104/2022



Decreto trasparenza



I fondamentali



Finalità



Base
giuridica



Misure

Base giuridica (Art. 6 GDPR)

(a)

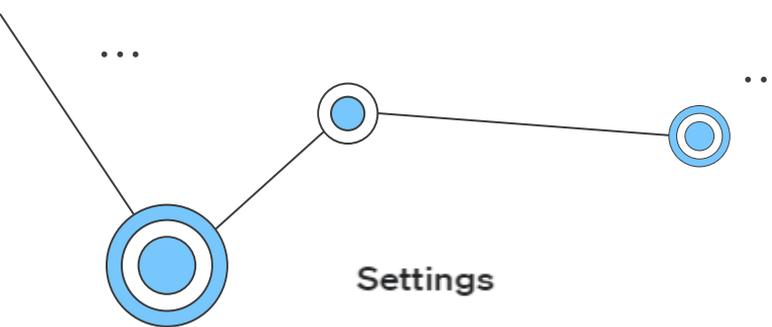
Utenti

Consenso con diritto di revoca, strumenti per la rettifica e dove non possibile, la cancellazione.

(f)

Non utenti

Legittimo interesse con di diritto di opposizione e analoghi strumenti degli utenti per rettifica e cancellazione.



Settings

✕

 General

 Beta features

 **Data controls**

Chat History & Training

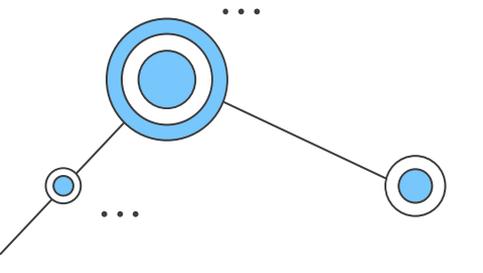
Save new web chats to your history and allow them to be used to improve ChatGPT via model training. Unsaved chats will be deleted from our systems within 30 days. [Learn more](#)

Export data

Export

Delete account

Delete



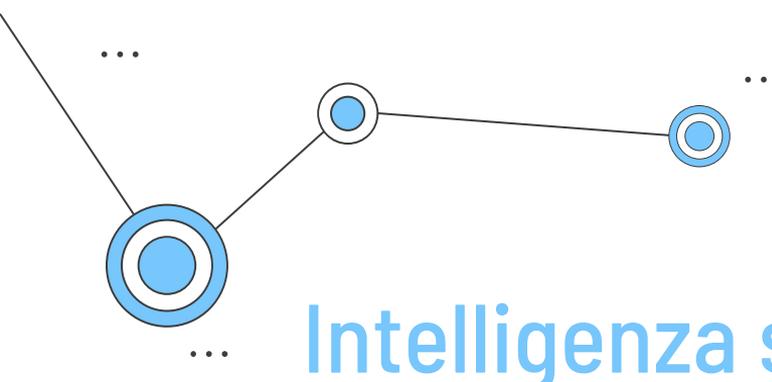
Fin dalla progettazione

Spiegabile

L'intelligenza c.d.
"simbolica, basata su
regole, vocabolari,
tassonomie e grafici
della conoscenza.

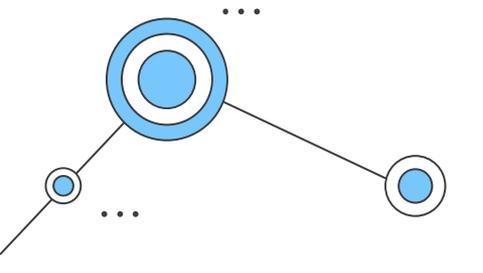
Non spiegabile

Scatola nera in cui
convergono dati ed
algoritmi per la
creazione di un modello
capace di apprendere e
decidere in autonomia.



Intelligenza simbolica

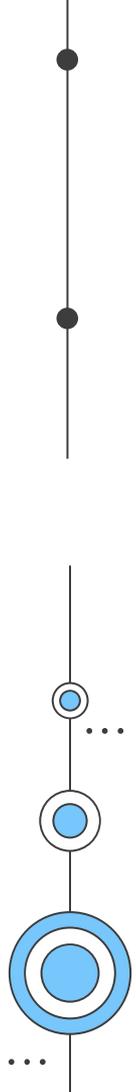
L'intelligenza artificiale simbolica è basata al 100% sulla conoscenza esplicita a tutti i livelli.



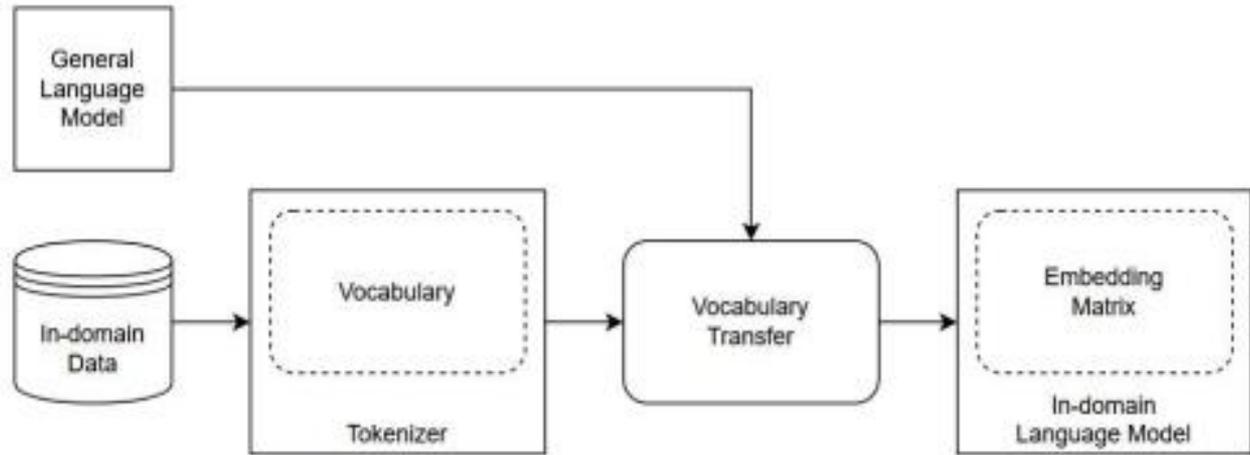


04

Integrazione



Modello ibrido

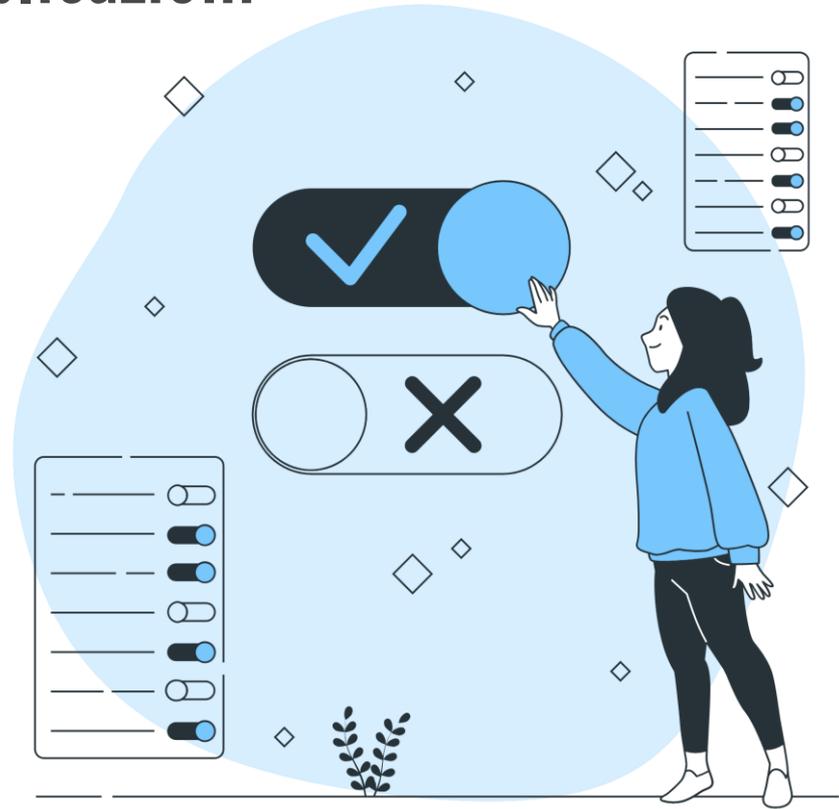


Applicazioni

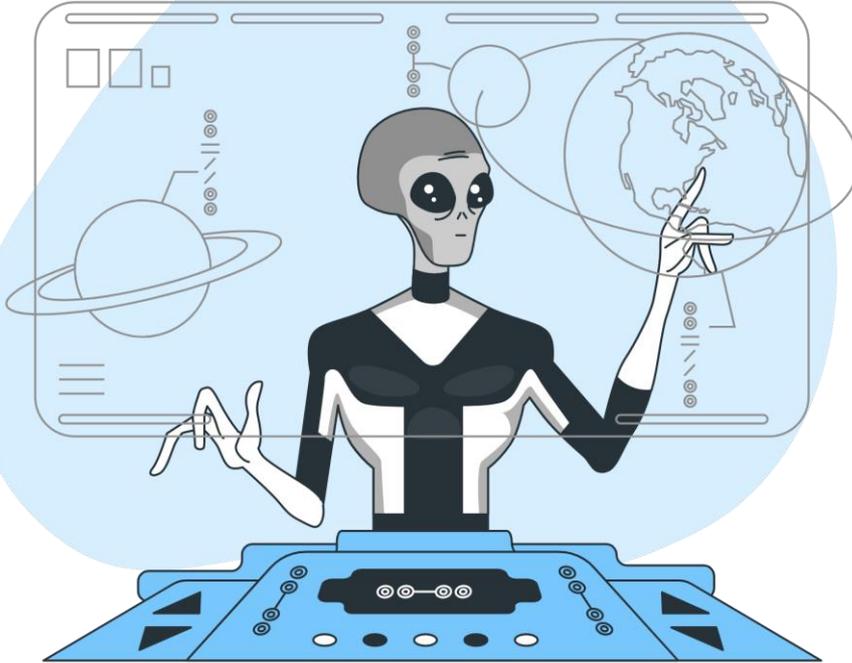
Obbligo
informativo

Esercizio dei
diritti

Registro dei
trattamenti



Applicazioni



[ChatPDF.com](https://chatpdf.com)

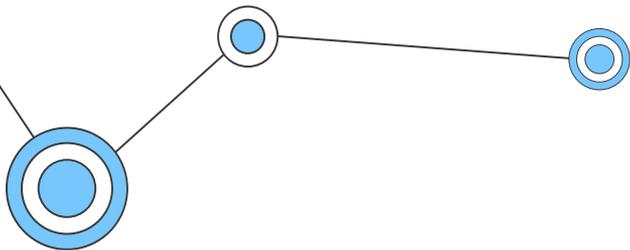
Carica un PDF e chiedi al BOT le informazioni

[Donotpay.com](https://donotpay.com)

Il primo avvocato Robot al mondo

[Spellbook.legal](https://spellbook.legal)

Scrittura di contratti con I.A.



Il documento fornisce una metodologia per valutare la gravità delle violazioni dei dati personali. La gravità viene valutata in base a quattro livelli: basso, medio, alto e molto alto. La metodologia tiene conto di diversi fattori, tra cui il tipo di dati personali coinvolti, il numero di individui interessati, la durata della violazione e le conseguenze per gli individui interessati. Alla fine della valutazione, vengono anche considerati altri criteri rilevanti che non sono stati presi in considerazione nella metodologia e che possono essere segnalati alle autorità competenti se necessario.

Come valuto la gravità?

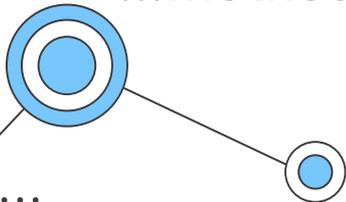
Chat with Data breach severity ...methodology_1.0.pdf

Mi potresti fare degli esempi?

Certamente! Ecco alcuni esempi di violazioni dei dati personali e la loro valutazione della gravità in base alla metodologia descritta nel documento:

1. **Bassa gravità:** un dipendente perde una chiavetta USB contenente dati personali di 10 clienti, ma la chiavetta è protetta da password e crittografia.
2. **Media gravità:** un hacker accede a un database contenente informazioni personali di 5000 utenti, tra cui nomi, indirizzi email e numeri di telefono.
3. **Alta gravità:** un'azienda perde un laptop contenente informazioni personali sensibili di 1000 dipendenti, tra cui numeri di previdenza sociale e informazioni bancarie.
4. **Molto alta gravità:** un attacco informatico compromette il sistema di una banca, consentendo agli hacker di accedere a informazioni personali e finanziarie di milioni di clienti.

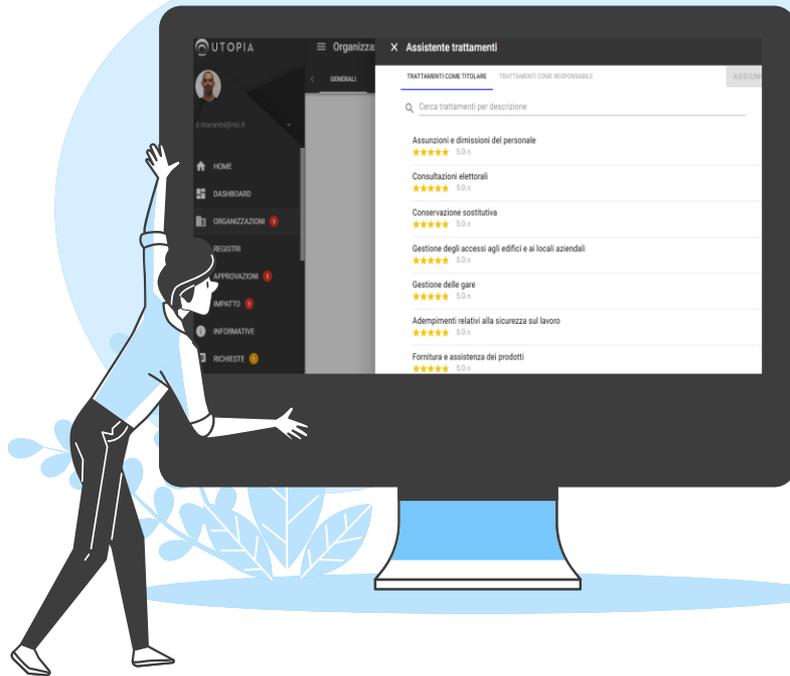
Naturalmente, la valutazione della gravità dipende dalle circostanze specifiche della violazione dei dati personali e potrebbe essere necessario considerare altri fattori oltre a quelli menzionati sopra.



MIA by UTOPIA

Assistente ai trattamenti

Tramite la funzione di «**ranking**» suggerisce i trattamenti più effettuati e grazie ad algoritmi statistici associati a funzioni di NLU simbolica **compila le informazioni più importanti del registro.**



Grazie

Domande?

d.maranini@dilaxia.com

www.utopiathesoftware.com

CREDITS: This presentation template was created by [Slidesgo](#), including icons by [Flaticon](#), infographics & images by [Freepik](#) and illustrations by [Stories](#)

