

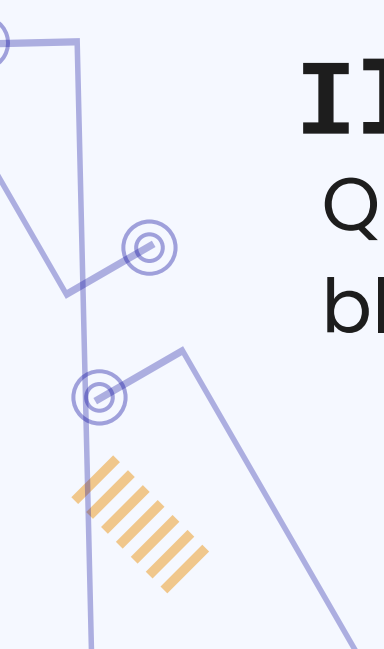
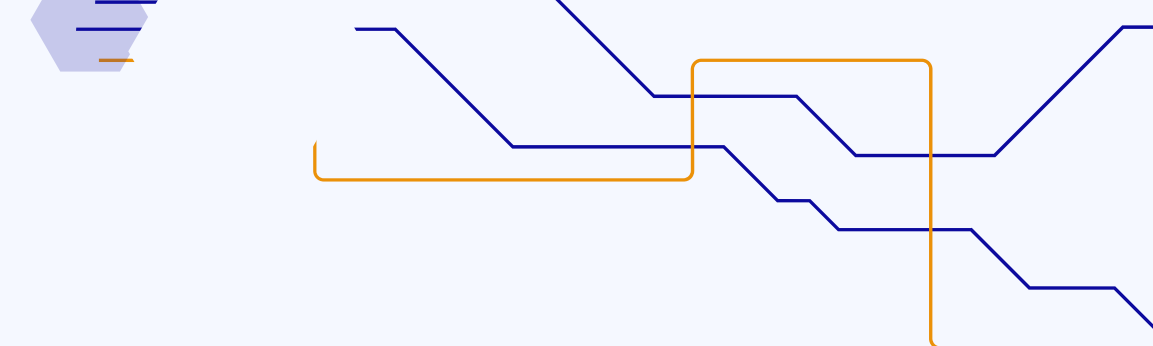
Open-washing nell'era dell'AI

L'illusione dell'Intelligenza Artificiale aperta



FERMOLUG
www.fermolug.org

LICENZA CC BY-NC-SA 4.0



Indice

01

Introduzione

Cos'è l'open-washing e perché inganna la community?

03

Il Mito dei Pesi Aperti

Quando "open weight" blocca l'innovazione

02

Il Paradosso del Profitto

Come le big tech monetizzano progetti open source

04

Stato dell'Arte

Il divario reale tra modelli veramente aperti e closed source



01

Introduzione

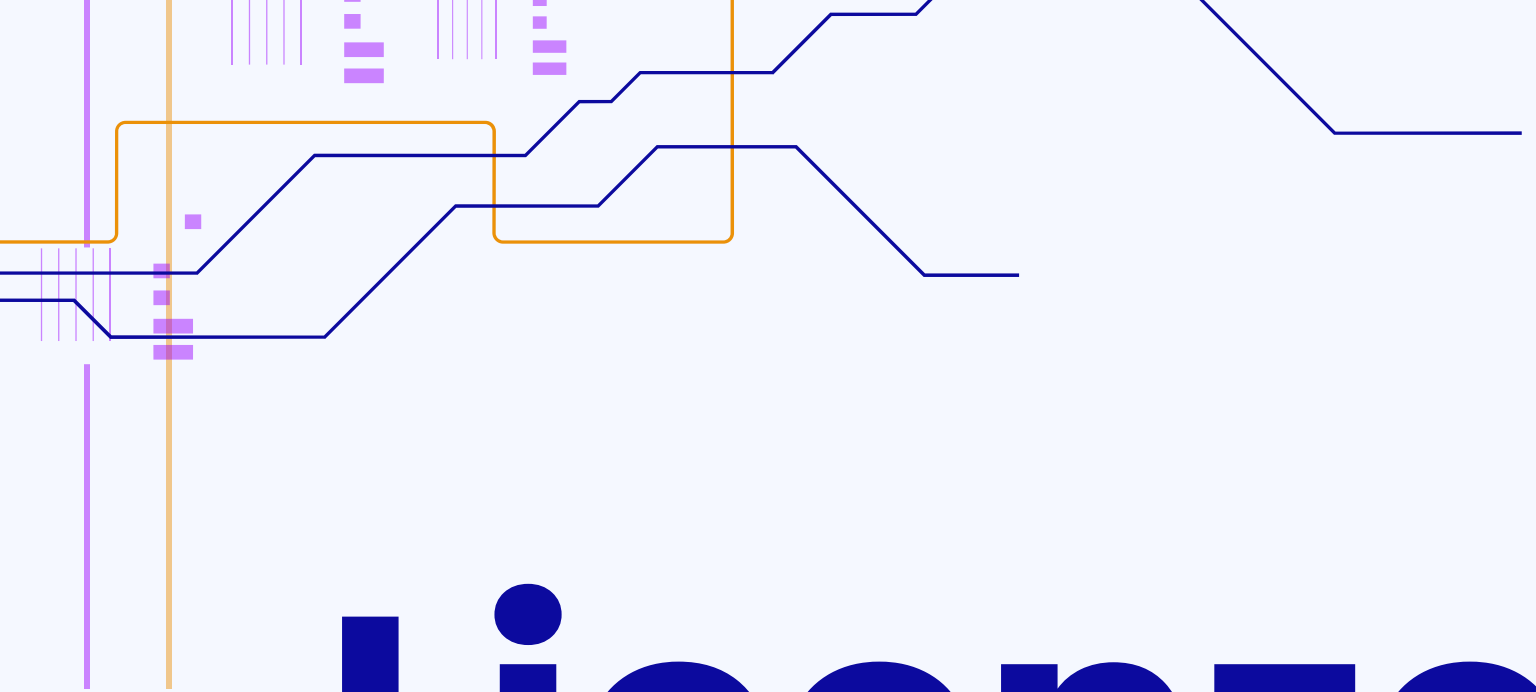
Cos'è l'open-washing e perché inganna la community?



Introduzione

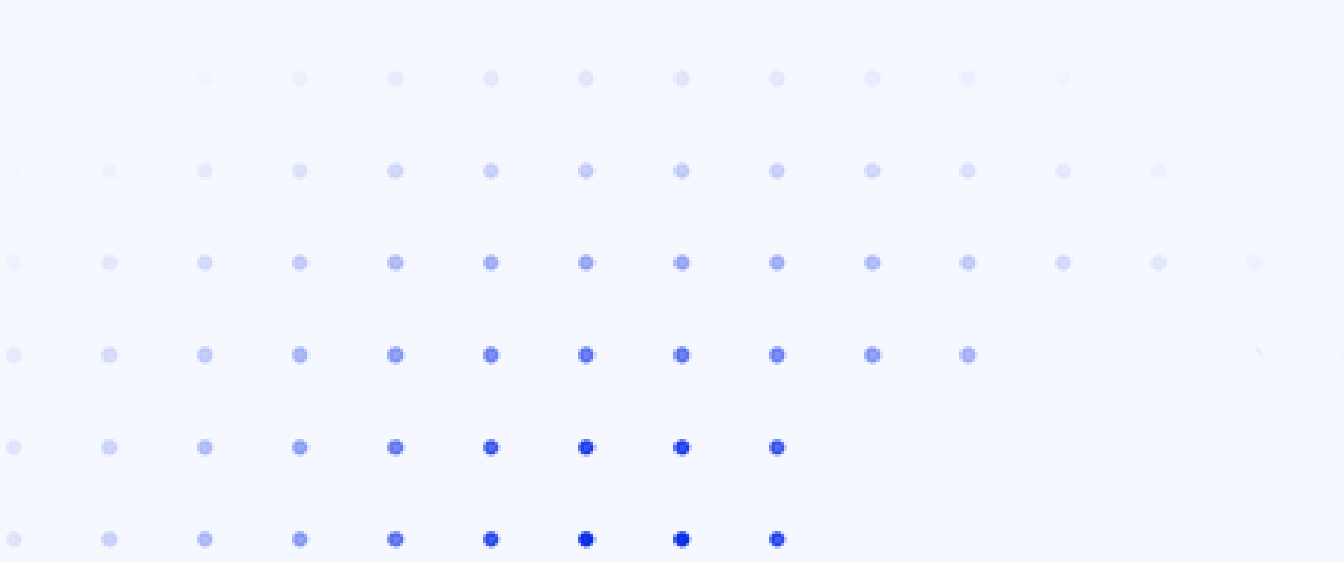
“L’open-washing nell’IA indica quando le aziende **esagerano il loro impegno verso l’apertura**, pur **mantenendo parti cruciali proprietarie**. Non è una novità: prima c’erano il cloud-washing, l’AI-washing, e ora questo. Le agenzie di marketing vogliono che “open” le faccia sembrare aziende virtuose che salvano cuccioli di foca dalle maree nere.”

Source: <https://www.infoworld.com/article/3615678/open-washing-and-the-illusion-of-ai-openness.html>



Licenze a confronto

Quali sono i tipi di modelli che possiamo trovarci davanti?



Licenze a confronto

Open Source

Dataset	✓
Pesi modello	✓
Codice	✓

Open Weights

Dataset	✗
Pesi modello	✓
Codice	✗

Closed / Modello proprietario

Dataset	✗
Pesi modello	✗
Codice	✗



Licenze a confronto

Open Source

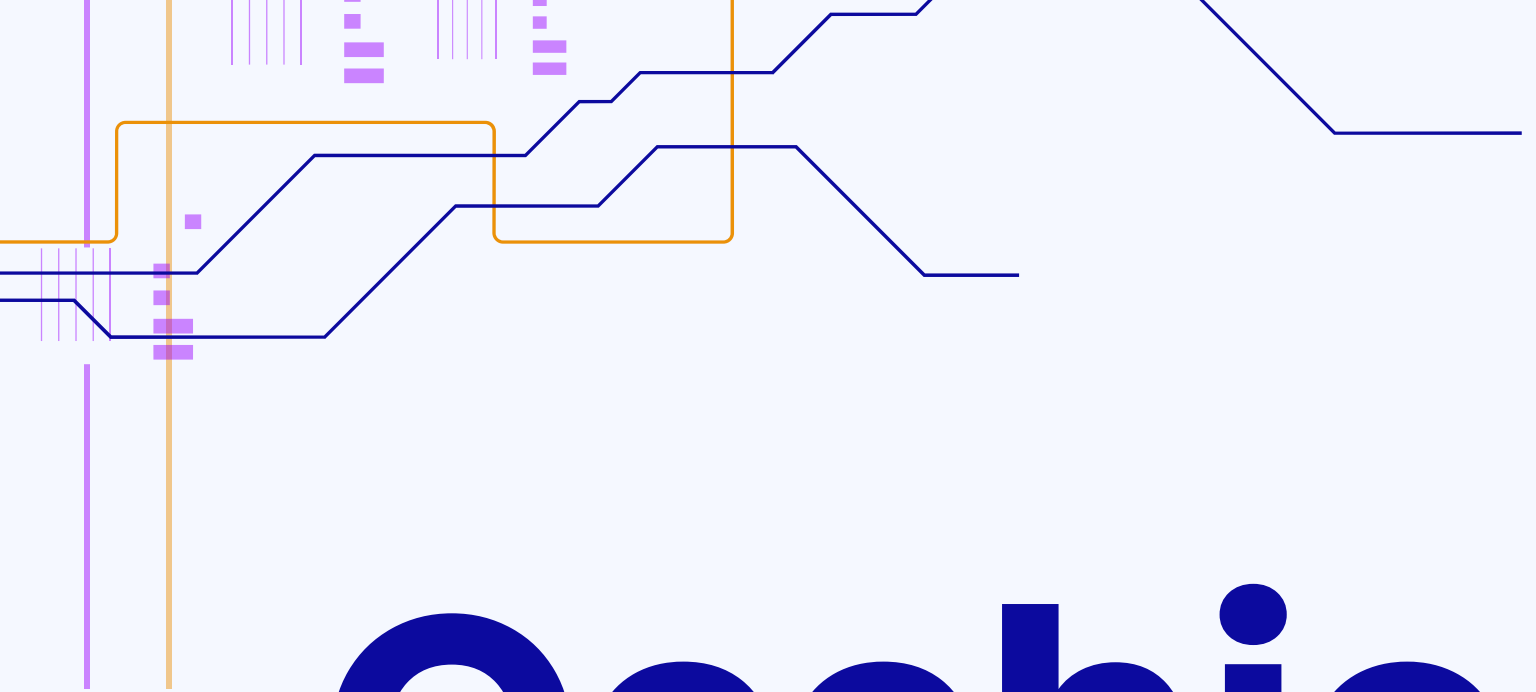
Dataset	✓
Pesi modello	✓
Codice	✓
Inferenza	✓
Modifica/fine-tuning	✓
Replicazione	✓

Open Weights

Dataset	✗
Pesi modello	✓
Codice	✗
Inferenza	✓
Modifica/fine-tuning	✓
Replicazione	✗

Closed / Modello proprietario

Dataset	✗
Pesi modello	✗
Codice	✗
Inferenza	✓
Modifica/fine-tuning	🚦
Replicazione	✗



Occhio alla Licenza

E' davvero sufficiente una licenza MIT per dire che un modello
AI è Open Source?



Da sviluppatori:

- Non dobbiamo limitarci a leggere il nome della licenza, è importante approfondire per essere sicuri a cosa andiamo incontro
- La differenza principale tra il Software classico e l'AI è il costo di produzione, è piuttosto comune trovare licenze come questa

Modified MIT License

Copyright (c) 2025 Moonshot AI

Permission is hereby granted, free of charge, to any person obtaining a copy of this software **and** associated documentation files (the "Software"), to deal in the Software without restriction, including without limitation the rights to use, copy, modify, merge, publish, **distribute, sublicense, and/or** sell copies of the Software, **and** to permit persons to whom the Software is furnished to do so, **subject** to the following conditions:

The above copyright notice **and** this permission notice **shall be** included in all copies **or substantial** portions of the Software.

THE SOFTWARE IS PROVIDED "AS IS", WITHOUT WARRANTY OF ANY KIND, EXPRESS **OR** IMPLIED, INCLUDING **BUT** NOT LIMITED TO THE WARRANTIES OF MERCHANTABILITY, FITNESS FOR A PARTICULAR PURPOSE **AND** NONINFRINGEMENT. IN NO EVENT **SHALL** THE AUTHORS **OR** COPYRIGHT HOLDERS **BE** LIABLE FOR ANY CLAIM, DAMAGES **OR** OTHER LIABILITY, WHETHER IN AN ACTION OF CONTRACT, TORT **OR** OTHERWISE, ARISING FROM, OUT OF **OR** IN CONNECTION WITH THE SOFTWARE **OR** THE USE **OR** OTHER DEALINGS IN THE SOFTWARE.

Our only modification part is that, **if the Software (or any derivative works thereof) is used for any of your commercial products or services that have more than 100 million monthly active users, or more than 20 million US dollars (or equivalent in other currencies) in monthly revenue, you shall prominently display "Kimi K2" on the user interface of such product or service.**

Alcuni esempi di Open-washing



GPT-OSS

- Nome ingannevole OSS
- Mancano dati e codice
- Replicazione impossibile



Llama 4

- Restrizioni geografiche Europa
- Limiti uso commerciale
- Viola principi OSI

Come le Big Tech monetizzano l'Open Source?

Strategia del Complemento

AI gratuita genera contenuti sui social.

contenuti > engagement > ads



Vendor Lock-in Cloud

Modello gratuito, infrastruttura a pagamento.
Cloud hosting, API premium, servizi enterprise e supporto tecnico generano i profitti reali.



R&D Esternalizzato

Migliaia di sviluppatori sperimentano gratis.

Le aziende osservano cosa funziona e acquisiscono solo le soluzioni vincenti già validate.



Perché dovremmo prediligere l'uso di modelli veramente Open Source?

OpenAI recently released their open-weights model. Here we'll discuss how that inevitably leaks some information about their model training stack, and, on the way, show that GPT-5 was trained on phrases from adult websites.

<https://fi-le.net/oss/>

In a joint study with the UK AI Security Institute and the Alan Turing Institute, we found that as few as 250 malicious documents can produce a "backdoor" vulnerability in a large language model—

<https://www.anthropic.com/research/small-samples-poison>

Progetti veramente Open Source



OLMo

- Dataset completo pubblico (Dolma)
- Codice training open source
- Licenza Apache 2.0

BigScience



BLOOM

- Progetto collaborativo trasparente
- Dataset ROOTS documentato
- Licenza veramente libera

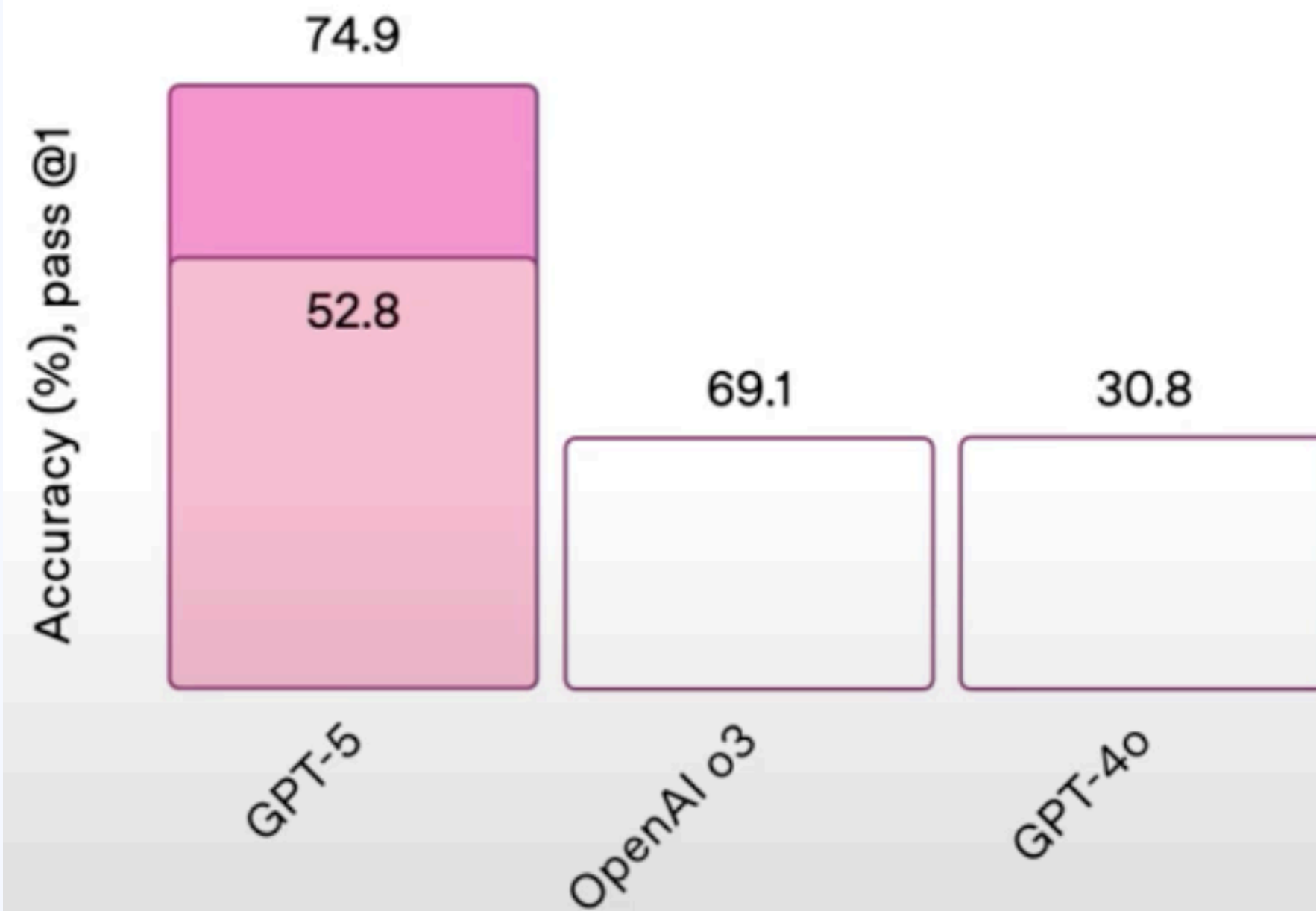
Performance a confronto



SWE-bench Verified

Software engineering

Without thinking With thinking



Approcci novelli

SAMSUNG

TRM

- Introduce una nuova architettura.
- Modello da 7M di parametri che **batte GPT-5 e DeepSeek 3.1 in alcuni benchmark**

github.com/SamsungSAILMontreal/TinyRecursiveModels

Ling-V2

- **Performance Stato dell'Arte**
- Estremamente efficiente
- Supporta reasoning
- **Addestrato su 20T di token**
(DeepSeek R1 è addestrato su 14,8T di token invece)

github.com/inclusionAI/Ling-V2



Approcci novelli



Reasoning with Sampling

- Sampler per migliorare l'accuratezza di risposte dei modelli LLM senza la fase di Reinforcement Learning
- Possiamo migliorare le performance di un LLM già addestrato del +20% (usando Monte Carlo Markov Chain)

<https://arxiv.org/pdf/2510.14901>

Supportare l'Open Source a casa



open|code

<https://opencode.ai>

+



GLM-4.6, Z.ai

Per riassumere

- Da sviluppatori e addetti ai lavori dobbiamo saper riconoscere le diverse licenze per evitare di violarle
- Come aziende è importante sapere che Open Weight non vuol dire gratis
- Da utenti è spesso possibile usare modelli Open Weight con performance simili ai modelli proprietari ma costi ridotti



Alcune fonti di apprendimento

- **Kaggle:** datasets, sfide, modelli
 - **HuggingFace:** datasets, corsi, modelli
 - **Google Colab:** esperimenti, fine-tuning, training
 - **scikit-learn:** guida e tools per analisi dati e machine learning
 - **Google Skills:** impara skills AI da zero con corsi gratuiti
- 
- 

Grazie dell'attenzione

```
profile.js

class Developer {
  constructor() {
    this.name = "Aldo Giambuzzi";
    this.age = 20;
    this.location = "Camerino, Italy";
    this.passions = ["Programming", "Artificial Intelligence",
    "Web Development", "Guitar", "Calisthenics"];
  }
}

const aldo = new Developer();
await aldo.createAmazingThings();|
```

 **Aldo Giambuzzi**



FERMOLUG
www.fermolug.org