

Considerazioni finali

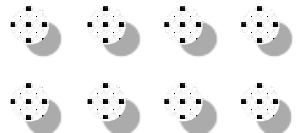
Forum Segnalazioni
Sistemi Intelligenti
22 maggio 2024

Segnalazioni correnti

- Principalmente **AI generativa (LLM)**
- **Esempi di applicazione** (ambito biomedico, sportivo, artistico, linguistico)
- **Rischi** (sostituzione dell'uomo, fake, ...)
- Comparire più volte il termine "**promette**" (questa tecnologia promette di rivoluzionare il settore X nei prossimi anni) – strategia di comunicazione –

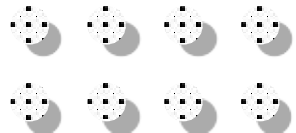


Illustrations by Pixeltrue on [icons8](#)



Qualche dietro le quinte

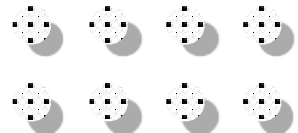
- Es: l'articolo porta alla luce un aspetto che, personalmente non conoscevo, vale a dire l'**enorme consumo di acqua** per raffreddare ed evitare il surriscaldamento dei tanti data center in seguito alle richieste che arrivano a servizi come ChatGPT o Bard ...



VOSTRE SEGNALAZIONI

**Large
Language
Model**

**Deep
Learning**

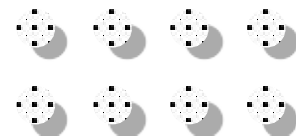


Parte di un sistema più ampio, descritto
solitamente a metà

**Large
Language
Model**



**Deep
Learning**

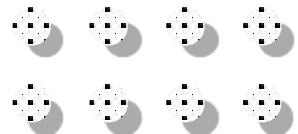


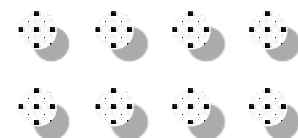
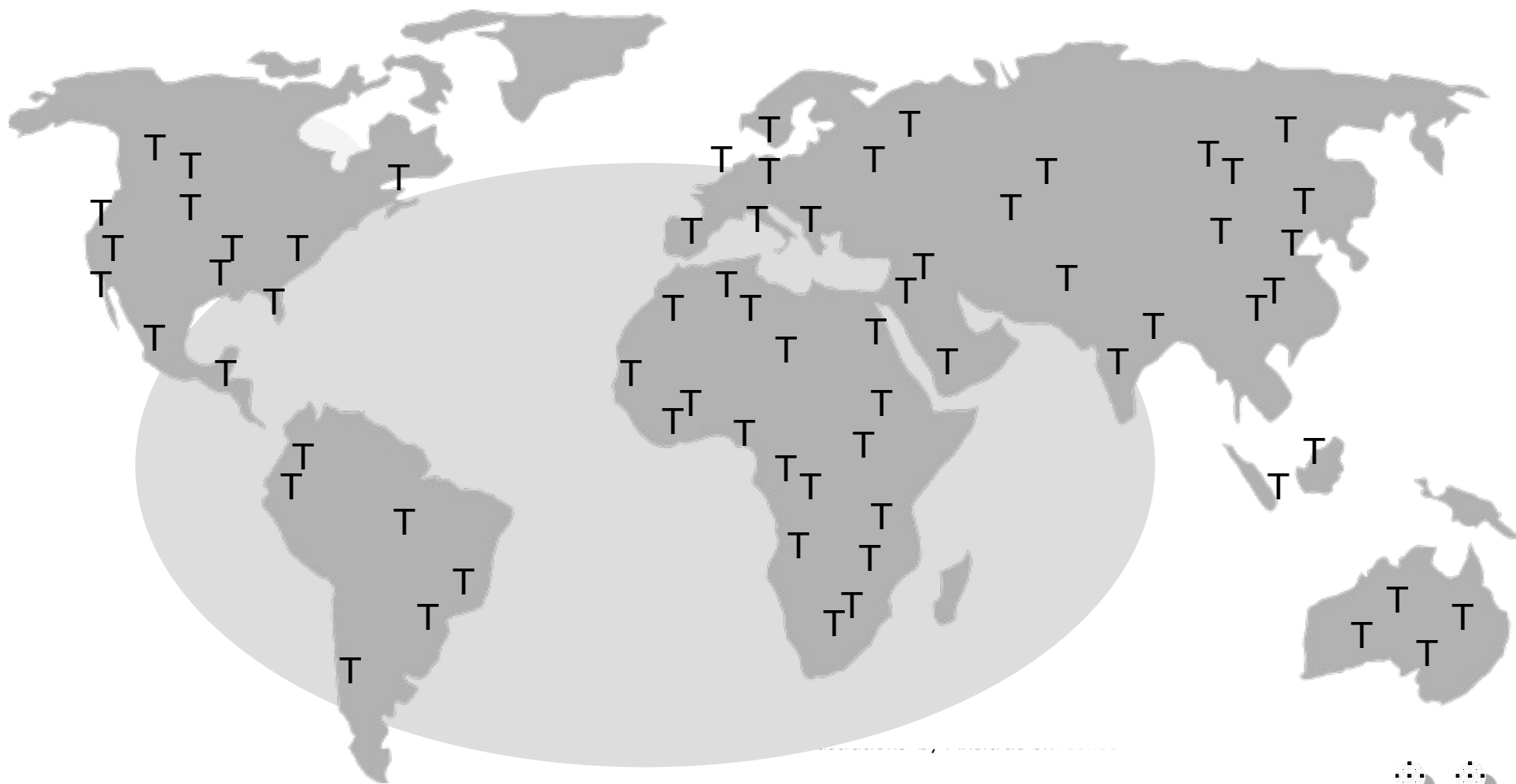
Al volgere del millennio Amazon aveva l'esigenza di identificare ed eliminare migliaia di duplicati dei prodotti mostrati dal sito web. Una persona ne sarebbe stata capace facilmente ma la mole di dati rendeva il compito impossibile

La realizzazione di un software fallì

Il manager Venky Harinarayan ideò uno schema organizzativo ibrido uomo/macchina, in cui un task viene spezzettato in tante piccole parti distribuite a tante persone diverse

Nascono il concetto di **HIT** (human intelligence task) e la piattaforma digitale **Amazon Mechanical Turk**





Amazon mechanical
turk

Digital labour platforms and the future of work

Towards decent work in the online world

J. Berg, M. Furrer, E. Harmon, U. Rani, M.S. Silberman
International Labour Office • Geneva 2018

... Microtask platforms initially emerged due to a need for uniquely human intelligence.
... microtask platforms such as AMT were “born out of the failures of artificial intelligence to meet the needs of internet companies seeking to expand the domain of the data they could store, classify and serve up on line” (p. 225).

Artificial intelligence could not “classify the nuances of the images, sounds and texts that filled Web 2.0” (ibid.), thus people, rather than algorithms, were needed to fill the gap.

In consequence, humans have become integral to the provision of services commonly marketed or described as “artificial intelligence” (AI)

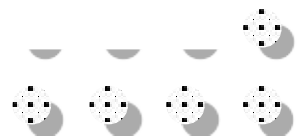
...



... Ironically, AI as it exists in current industry practice has much less to do with its original intellectual and practical aims as a field, which was to create entirely automated systems that could solve problems that previously only humans could solve (e.g. whether an image satisfies a certain set of criteria)

Rather, AI tends to reorganize those problems so they can be **(a) partially automated** and **(b) outsourced to a low-cost, flexible, algorithmically managed workforce**

Paradoxically, some microtask platforms specialize in providing human-labelled data sets that are used to train machine-learning algorithms in the hope that they will be able to function independently of human input in the future



**Raccolta e
Pulizia dei Dati**

Data Splitting

Separazione di training e test set

Model Set-Up

Identificazione dell'architettura

Training

Checking

Integrazione

Integrazione nel sistema di utilizzo



Raccolta e Pulizia dei Dati

Crowdworking tramite piattaforma
digitale di microtask

Data Splitting

Separazione di training e test set

Model Set-Up

Identificazione dell'architettura

Training

Checking

Integrazione

Integrazione nel sistema di utilizzo



AI Training Data – Quality Data for Your Algorithm

AI training data is the basis for building and improving AI models. Your algorithms need human interaction if you want them to provide human-like results. Our AI training data services focus on computer vision and conversational AI. Find out more and buy quality AI training data.



✉ Get in touch with us!

+1 (812) 878-6534

+49 201 30571233

Le aziende usano la piattaforma per trovare e impiegare microworker su compiti di interesse

AI Training Datasets by clickworker

With over 6 million Clickworkers, we can help you get more out of your algorithms by **generating**, **labeling** and **validating** unique AI datasets, specifically tailored to your needs. We can also provide you with a solution to quickly analyze the results of your AI's output.

E S E M P I O

Vcode Generation, Autosubmission

VCode generation
Auto Submission to MW



When you create a campaign in Microworkers-TTV, you don't always have to start from scratch. All you need to do is pick from any of our ready-to-use templates. These templates are built all you need to do is use our Clone & Edit option, personalize your chosen template, then go ahead with adding your project requirements.

All of our templates are fully customizable. You wouldn't need to work hard on putting one together - it's all done for you! Got a project in mind? Let's get started right away!



Image Categorization (Product Properties) (Categorization)

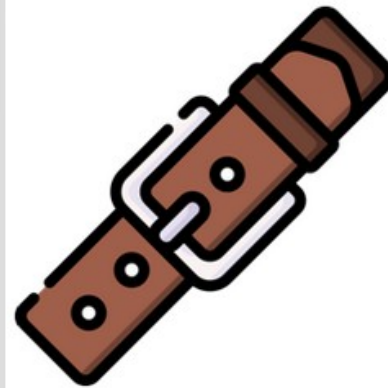
Using a series of images and classification, this template helps improve online product shopping and/or searching.

Task Instructions:

Look at the product image and provide the following

- Most suitable category
- Is this product is suited for a child or a grown up?
- Is this product is suited for male or female?
- Define product style (casual/formal/sport)

microWorkers
sample



microWorkers
sample



Fill in

Evaluation of Surprising Products (Sentiment Analysis)

Emotions can be expressed through images. Do you want to know the feelings of the respondents when they see your images? Use this template to find out.

Task Instructions:

- Here you find the pictures of 10 products, labelled with letters from **(a)** to **(j)**, that are supposed to be surprising
- Without considering your judgement about the aesthetics of the product, which is not interesting to the purpose of the present survey, **select three (3)** of them, which you judge the most surprising

Note:

For any of them answer the following questions, by specifying the **letter associated** with the **reference picture**

Questionnaire:



(a)



(b)



(c)



(d)



Tweet Emotion (Sentiment Analysis)

Tweets may have different impacts on readers. Use this template to annotate emotions from your given tweet link.

Task Instructions:

Read the tweet and label the content **according to the emotion** it discloses.

" Template Form Builder From JSON Schema blog.microworkers.com/template-form-builder-from-json-schema/ #template #JSON #tcampaign #JSONdata #JSONschema #TTV "

Emotion:

<input type="radio"/> Neutral	<input type="radio"/> Boredom	<input type="radio"/> Relief	<input type="radio"/> Surprise	<input type="radio"/> Fun	<input type="radio"/> Happiness	<input type="radio"/> Love	<input type="radio"/> Sadness	<input type="radio"/> Enthusiasm	<input type="radio"/> Worry	<input type="radio"/> Anger	<input type="radio"/> Hate
-------------------------------	-------------------------------	------------------------------	--------------------------------	---------------------------	---------------------------------	----------------------------	-------------------------------	----------------------------------	-----------------------------	-----------------------------	----------------------------

Determining Image Color (Content Moderation)

Lively colors can be deceiving. Identify and choose dominant colors from images with this template. Contains selections of colors where workers have to pick the most

Task Instructions:

Please help us to determine the color of the image.

- The product shown can have multiple colors. Please make selection on all the dominant colors that apply
- If there is other color(s) present on the image, please make sure to check "**Other**" and write such in text box below



- ☐ White
- ☐ Black
- ☐ Red
- ☐ Yellow
- ☐ Pink
- ☐ Sky Blue
- ☐ Other (please specify)

Other Color


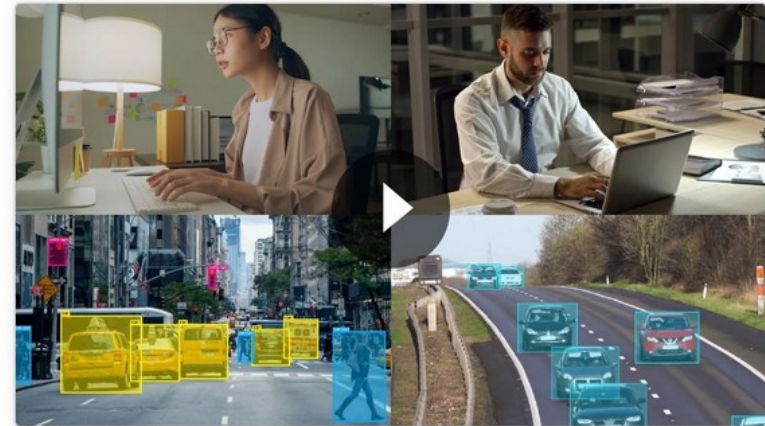
Pronto a trovare i dati che ti mancavano?

[Contatti](#)

Servizi di raccolta dati completamente gestiti

Poiché i dati sono della massima importanza per il successo di ogni organizzazione, si stima che in media i team di intelligenza artificiale trascorrono l'80% del loro tempo a preparare i dati per i modelli di intelligenza artificiale. Questa preparazione dei dati di solito include più passaggi come:

- Identificare i dati richiesti
- Identificare la disponibilità dei dati
- Profilazione dei dati
- Recupero dei dati
- Integrazione dei dati
- Pulizia dei dati
- Preparazione dei dati



Il team Shaip, aiutato dal nostro strumento di raccolta dati proprietario (app mobile disponibile per Android e iOS), gestisce una forza lavoro globale di raccoglitori di dati per raccogliere dati di formazione per i tuoi progetti AI e ML. Attingendo da un'ampia varietà di fasce d'età, dati demografici e background educativi, possiamo aiutarti a raccogliere grandi volumi di set di dati di apprendimento automatico per soddisfare le iniziative di intelligenza artificiale più esigenti. Shaip ti assiste durante tutto il processo di raccolta dei dati e ti consente di concentrarti sul risultato e guidare il tuo progetto di intelligenza artificiale in una direzione: **INOLTRE.**

Catalogo e licenza di dati medici pronti all'uso:

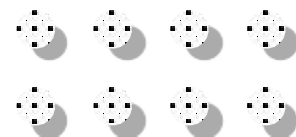
- 5 milioni di record e file audio medici in 31 specialità
- 2 milioni di immagini mediche in radiologia e altre specialità (MRI, TC, USG, XR)
- Oltre 30 documenti di testo clinico con entità a valore aggiunto e annotazione di relazioni

Catalogo dati immagine e video e licenze:

- Raccolta di immagini di alimenti/documenti
- Raccolta di video sulla sicurezza domestica
- Raccolta di immagini/video facciali
- Fatture, PO, Ricevute Raccolta documenti per OCR
- Raccolta di immagini per il rilevamento dei danni ai veicoli
- Raccolta di immagini della targa del veicolo
- Collezione di immagini per interni auto
- Raccolta di immagini con automobilista a fuoco
- Raccolta di immagini legate alla moda

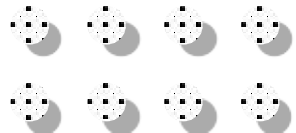
Catalogo e licenza dei dati vocali pronti all'uso:

- Oltre 40 ore di dati vocali (50+ lingue/100+ dialetti)
- Oltre 55 argomenti trattati
- Frequenza di campionamento - 8/16/44/48 kHz
- Tipo di audio: parole spontanee, sceneggiate, monologiche, di risveglio
- Set di dati audio completamente trascritti in più lingue per conversazioni uomo-uomo, uomo-bot, conversazione call center uomo-agente, monologhi, discorsi, podcast, ecc.



Microlavoro, implicazioni sociali?

- ❁ Micropaga (mediamente 2\$ / ora, comunque minore del minimo orario consentito per legge in molti paesi)
- ❁ Algoritmi di gestione dell'offerta di micro-work a microworker non trasparenti
- ❁ Mancanza di supporto nelle contestazioni (microworker non pagato dall'azienda committente)
- ❁ Strumentazione: device propri
- ❁ Diritti dei lavoratori?
- ❁ Malattia?
- ❁ Pensione?



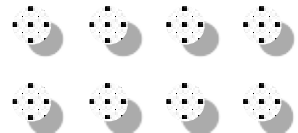
Microlavoro, implicazioni sociali?

- ✿ Lavoro saltuario compatibile con necessità (temporanea) di arrotondamento
- ✿ Fonte di guadagno in periodi di perdita di lavoro
- ✿ Autonomia nella scelta dei tempi e delle quantità di task
- ✿ Flessibilità per chi deve dedicarsi all'assistenza (bambini, anziani, ...)



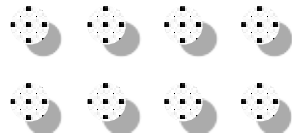
Fair Crowd Work

- ✿ <http://faircrowd.work/>
- ✿ raccoglie informazioni su crowd work, app-based work e simili dalla prospettiva di lavoratori e associazioni di lavoratori
- ✿ Offre valutazioni delle condizioni di lavoro sulle varie piattaforme online basate su survey



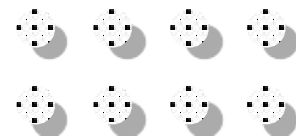
Qualche lettura

- ✿ Untold History of AI: How Amazon's Mechanical Turkers Got Squeezed Inside the Machine,
<https://spectrum.ieee.org/untold-history-of-ai-mechanical-turk-revisited-tkkt>
- ✿ Digital Labour Platforms and the future of Work
https://www.researchgate.net/publication/328043472_Digital_labour_platforms_and_the_future_of_work_Towards_decent_work_in_the_online_world
- ✿ <https://it.shaip.com/blog/a-guide-large-language-model-llm/>



Scorso anno: applicazioni religiose

- Tarteel AI
- Prega.org



Scorso anno: applicazioni religiose

- Tarteel AI: **esiste ancora**
- Prega.org: **sospeso**
(<https://www.parmateneo.it/?p=90471>)

