

Prospettive sulla divisione del lavoro di computazione

Origini del microwork e produzione
dell'intelligenza artificiale¹

Lorenzo De Lellis

Origini del microwork

Il termine microwork si riferisce a un'attività lavorativa svolta all'interno di piattaforme online e finalizzata, principalmente, al *training* e al *fine-tuning* dell'intelligenza artificiale (IA). Questo lavoro comprende una serie di task differenti che possono, in linea di massima, essere ricondotte a due categorie principali: (i) l'etichettatura o la produzione di dati da inserire all'interno dei dataset utilizzati per addestrare le reti neurali e (ii) la verifica degli output

1. Finanziato dall'Unione europea nell'ambito del progetto ERC AIMODELS (GA n. 101088645). Le opinioni espresse appartengono tuttavia al solo autore e non riflettono necessariamente le opinioni dell'Unione europea o dell'Agenzia esecutiva europea per la ricerca (REA). Né l'Unione europea né l'amministrazione erogatrice possono esserne ritenute responsabili.

prodotti da un modello, come la corretta inclusione di specifiche informazioni all'interno di un testo generato da un *large language model* (LLM).

La prima piattaforma di microwork è stata Mechanical Turk, creata da Amazon nel 2003. Al suo interno, in modo analogo a quanto avviene nei magazzini dell'azienda, il lavoro è stato fin dall'inizio organizzato tramite procedure standardizzate, descritte da istruzioni dettagliate funzionali a regolare ogni fase del processo produttivo, con l'obiettivo di uniformare e semplificare le operazioni e garantire la massima flessibilità e sostituibilità dei lavoratori. Questa configurazione modulare, che tuttora caratterizza la maggior parte delle piattaforme di microwork, consente una significativa riduzione dei costi, ottenuta non soltanto grazie alla delocalizzazione su scala globale del processo produttivo, ma anche attraverso la sua frammentazione in operazioni ripetitive e facilmente replicabili. Ne emerge un'organizzazione del lavoro in cui il contributo umano si configura prevalentemente come una sequenza di processi ripetitivi di classificazione delle informazioni, che costituiscono la struttura portante dei moderni sistemi di IA.

In un celebre intervento al MIT nel 2006, Jeff Bezos evidenziava già la funzione centrale del microwork come strumento di addestramento dell'IA, sottolineando la necessità di incorporare la cognizione umana all'interno delle sequenze di esecuzione del codice per migliorarne l'accuratezza². Questo processo di integrazione delle decisioni umane nei cicli di addestramento della macchina, generalmente indicato con l'espressione «human-in-the-loop», costituisce la funzione principale del microwork, il cui obiettivo primario è oggi quello di allineare gli output dei modelli di IA.

2. J. Pontin, *Opening Keynote and Keynote Interview with Jeff Bezos*, «MIT TechTV», techtv.mit.edu/videos/16180-opening-keynote-and-keynote-interview-with-jeff-bezos, letto il 29/09/2025.

Legame tra microwork e IA

Anche se la stretta relazione tra microwork e IA era già inscritta nei principi fondativi di Mechanical Turk, uno specifico evento contribuì a consolidare ulteriormente questo legame: la creazione di ImageNet. Fei-Fei Li, allora ricercatrice informatica all'università di Princeton, avviò nel 2007 il progetto per costruire un ampio dataset per addestrare modelli di *computer vision* capaci di riconoscere autonomamente oggetti presenti in immagini digitali. Consapevole della necessità di disporre di una base di dati vasta e strutturata per addestrare sistemi di IA efficaci, Li decise di ricorrere a Mechanical Turk per l'etichettatura delle immagini – una soluzione che permetteva di coniugare basso costo del lavoro, velocità e scalabilità. ImageNet rappresenta, dunque, il primo esempio di un ampio dataset per il *training* dell'IA basato su dati prodotti dagli utenti di Internet – in quanto composto da immagini estratte da motori di ricerca come Google Images e da social network come Flickr – e successivamente annotato da lavoratori di piattaforma.

L'utilizzo di Mechanical Turk per operazioni di addestramento dell'IA trovò quindi in ImageNet il primo grande progetto capace di dimostrare il potenziale delle piattaforme di microwork. ImageNet non rappresentò soltanto un'innovazione tecnica, ma anche l'avvio di una trasformazione del modo in cui viene organizzato il processo di produzione dell'IA: il lavoro necessario al suo addestramento è stato infatti, a partire dalla creazione del dataset, sempre più esternalizzato, frammentato e distribuito su scala globale. Il modello modulare del microwork, fondato su standardizzazione e replicabilità, non soltanto riduce i costi, ma accelera anche significativamente i tempi di sviluppo della tecnologia, e la forza lavoro a basso costo, scalabile e flessibile delle numerose piattaforme nate sul modello di Mechanical Turk è oggi divenuta una componente strutturale del processo di creazione dell'IA. Questa tecnologia richiede, infatti, una grande quantità di lavoro umano per produrre output sensati, coerenti e affidabili. Il microwork serve quindi, essenzialmente, a fornire un numero elevato di esempi per consentire alla macchina di replicare correttamente le scelte umane,

anche in relazione a oggetti e contesti non precedentemente noti. Ciò dimostra che l'IA contemporanea non si fonda esclusivamente sull'appropriazione dei dati degli utenti di Internet, ma dipende anche dal lavoro umano necessario alla loro organizzazione, che costituisce una componente strutturale – e non residuale o transitoria, come talvolta si pensa – del suo processo produttivo.

Divisione del lavoro di computazione e conoscenza collettiva

Per comprendere più a fondo i principi fondativi del microwork è utile tornare, per un momento, ai presupposti alla base della creazione di Mechanical Turk. L'idea della piattaforma si deve a Venky Harinarayan, Anand Rajaraman e Anand Ranganathan, tre ingegneri che hanno brevettato per Amazon il modello di organizzazione produttiva tipico di queste piattaforme. Nel brevetto, la loro invenzione è descritta come «[a] hybrid machine/human computing arrangement which advantageously involves humans to assist a computer to solve particular tasks, allowing the computer to solve the tasks more efficiently»³. In sostanza, viene proposta una forma di divisione del lavoro in cui le necessità delle macchine generano una serie di *microtask* che devono essere svolte da esseri umani.

Si può andare indietro di circa due secoli da questo brevetto nella ricerca delle origini concettuali di questa forma di divisione del lavoro. Lorraine Daston ha mostrato chiaramente che Gaspard de Prony, matematico francese incaricato di calcolare un'ampia serie di tavole logaritmiche da utilizzare in ambito catastale, astronomico e navale, si ispirava apertamente al concetto di divisione del lavoro formulato da Adam Smith, schematizzandolo all'interno di una piramide che rappresentava la specifica struttura gerarchica necessaria per svolgere il compito assegnatogli dal governo fran-

3. V. Harinarayan, A. Rajaraman e A. Ranganathan, *Hybrid machine/human computing arrangement*, United States Patent: 7197459, 2001.

cese⁴. Al vertice della piramide troviamo un'élite di matematici dotati di superiori capacità intellettive, al centro i tecnici incaricati di tradurre le idee dei vertici in una serie di operazioni elementari facilmente eseguibili – semplici addizioni e sottrazioni – e alla base i lavoratori addetti al calcolo di queste operazioni, interpretabili a tutti gli effetti come *subtask* del processo di computazione complessivo. Secondo Charles Babbage, da molti considerato il padre della computazione moderna e fortemente influenzato dalle idee di De Prony, i lavoratori incaricati di svolgere queste operazioni potevano essere considerati come entità puramente meccaniche nell'esecuzione dei loro compiti.

Circa un secolo e mezzo dopo, Herbert Simon, figura cruciale nella storia dell'IA, affermava in una lezione alla Operations Research Society che il vero inventore della computazione moderna è stato, in realtà, Adam Smith, la cui idea di divisione del lavoro venne successivamente incorporata nelle strutture delle macchine a partire dagli sviluppi concettuali apportati da De Prony e Babbage. Nel corso della lezione, Simon precisava inoltre come, più in generale, lo scopo fondamentale della computazione – e dunque anche dell'IA, che ne rappresenta in un certo senso la massima espressione – sia quello di applicare sistematicamente soluzioni “intelligenti”, cioè efficienti, ai processi amministrativi delle aziende e degli apparati burocratici. Questo approccio trovava, secondo Simon, la sua prima compiuta espressione nel lavoro di Frederick Winslow Taylor⁵. In altre parole, Simon proponeva già negli anni Cinquanta l'idea che l'IA si fondi essenzialmente sulla divisione del lavoro e sul management scientifico di matrice taylorista. Simon fu in effetti, oltre che uno dei pionieri dell'IA, soprattutto un importante economista, attivo in particolare nell'analisi dell'ottimizzazione dei processi decisionali e aziendali attraverso il miglioramento dei si-

4. L. Daston, *Calculation and the division of labor, 1750-1950*, «Bulletin of the German Historical Institute», 62 (Spring), 2017, pp. 9-30.

5. H. A. Simon e A. Newell, *Heuristic Problem Solving: The Next Advance in Operations Research*, «Operations Research», 6 (1), 1958, pp. 1-10.

stemi di elaborazione dell'informazione. In sintesi, si può affermare che fu colui che, per primo e più chiaramente di altri, comprese sia il potenziale dell'IA per l'efficientamento produttivo che il suo legame con i principi dell'organizzazione del lavoro, dai quali derivano sia la struttura tecnica che concettuale di questa tecnologia.

Naturalmente, dal discorso di Simon alla Operations Research Society alla nascita di Mechanical Turk sono cambiate molte cose nell'IA. Tuttavia, il lascito delle sue idee seminali può essere ancora rintracciato nel brevetto di Harinarayan, Rajaraman e Ranganathan, nel quale il processo di computazione non è più concepito semplicemente come una sequenza di operazioni interne ai dispositivi, ma piuttosto come il prodotto di un network ibrido umano-macchina, ossia come un complesso socio-tecnico fondato sulla divisione del lavoro di computazione e sulla sua organizzazione secondo i principi del management scientifico. La scomposizione del lavoro di computazione risulta qui finalizzata all'addestramento di macchine capaci di riprodurre su larga scala la conoscenza collettiva estratta dai dati e sistematizzata attraverso il lavoro umano. Questa duplice dimensione, analogica e generalizzante, rende l'IA un oggetto particolare, diverso da precedenti forme di strutturazione della conoscenza collettiva – quali, ad esempio, le tavole logaritmiche di De Prony – ma pur sempre costruito catturando la dimensione sociale della conoscenza attraverso l'organizzazione del lavoro produttivo. Si può affermare, in conclusione, che la vera intelligenza dell'IA risieda oggi nella sua capacità di organizzare la società sul piano materiale ed epistemico attraverso strutture produttive globalmente distribuite, piuttosto che nel produrre ragionamenti e astrazioni simili a quelli umani.