



HAL
open science

Preparare, verificare, imitare: perché il lavoro umano è necessario alla produzione di intelligenze artificiali

Antonio A. Casilli

► **To cite this version:**

Antonio A. Casilli. Preparare, verificare, imitare: perché il lavoro umano è necessario alla produzione di intelligenze artificiali. L'enigma del valore. Il digital labor e la rivoluzione tecnologica (Atti del convegno organizzato da Effimera, 1° giugno 2019, Milano, Casa della Cultura), Effimera, pp.25-41, 2020. halshs-02882795

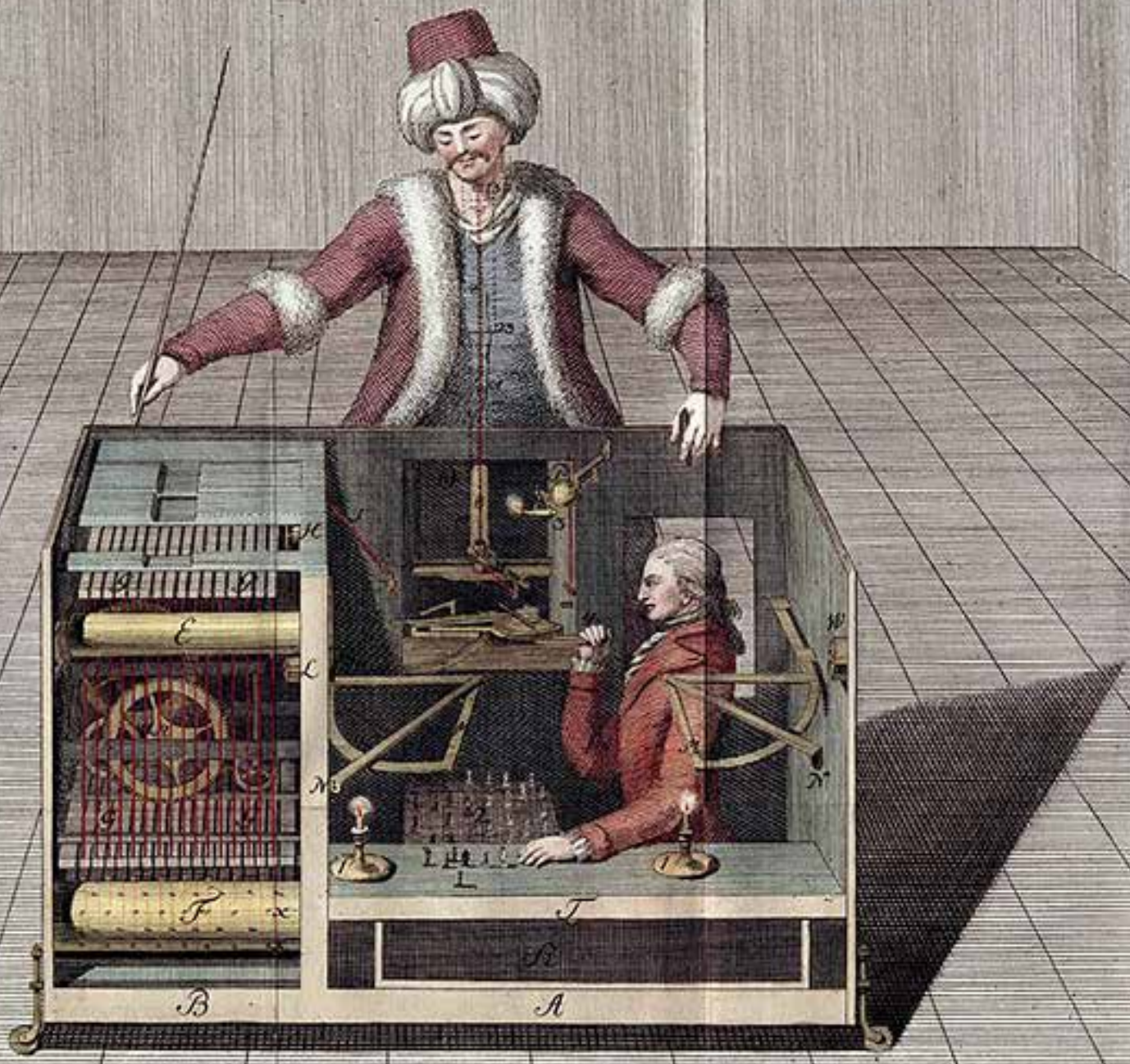
HAL Id: halshs-02882795

<https://shs.hal.science/halshs-02882795v1>

Submitted on 6 Jul 2020

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



L'ENIGMA DEL VALORE

IL DIGITAL LABOUR E LA NUOVA
RIVOLUZIONE TECNOLOGICA

Atti del convegno organizzato da Effimera,
1° giugno 2019, Milano, Casa della Cultura

EFFIMERA
CRITICA E SOVVERSIONI DEL PRESENTE

Antonio A. Casilli

Addestrare, verificare, imitare: perché il lavoro umano è necessario alla produzione dell'intelligenza artificiale

Le tematiche che saranno affrontate in questo intervento sono principalmente tratte da tre testi. Il primo è il mio saggio *Schiavi del clic* (2020)¹⁶. Il secondo invece è un rapporto basato sullo studio *DiPLab* (Digital Platform Labor), condotto nel 2017-2019 con il mio gruppo di ricercatori del CNRS e dell'università Telecom Paris, e pubblicato con il titolo "Il Micro Lavoro in Francia" ("Le Micro-Travail en France", 2019)¹⁷. L'ultimo è un articolo pubblicato nella rivista *Big Data & Society*¹⁸. In quanto segue proporrò spunti teorici tratti dal libro ed elementi empirici tratti dal rapporto e dall'articolo.

L'automazione intelligente non è labor-saving

Il punto di partenza è, sostanzialmente, la necessità di distaccarci dalla *doxa* attuale sull'impatto dell'Intelligenza Artificiale sul lavoro, attorno alla quale si concentra la maggior parte della produzione scientifica degli ultimi anni in economia, in sociologia e in altre discipline—e di cui troviamo echi anche nel dibattito pubblico e politico. Secondo questi approcci, staremmo oggi assistendo a un'ondata di tecnologie intelligenti e di processi automatici che, in quanto *labor-saving*, provocherebbe la distruzione di percentuali non trascurabili di impieghi. Le stime possono variare. Si va dal 47% dei posti di lavoro perduti per colpa dell'automazione a orizzonte 2030 (contenuto nell'ormai famigerato studio pubblicato in un white paper di due ricercatori di Oxford del 2013)¹⁹ a stime molto più prudenti. Uno studio dell'OCSE su diciassette paesi tra il 1993 e il 2007 mostra l'assenza di effetti significativi dell'introduzione di

¹⁶ Antonio A. Casilli, *Schiavi del clic. Perché lavoriamo tutti per il nuovo capitalismo*, Milano, Feltrinelli, 2020 [trad. it. Raffaele Alberto Ventura di Id. *En attendant les robots. Enquête sur le travail du clic*. Parigi, Editions du Seuil, 2019].

¹⁷ Antonio A. Casilli, Paola Tubaro, Clément Le Ludec, Marion Coville, Maxime Besenval, Touhfat Mouhtare, Elinor Wahal, "Le Micro-travail en France. Derrière l'automatisation de nouvelles précarités au travail?", rapporto del progetto DiPLab (Digital Platform Labor), 2019 <<http://diplab.eu>>.

¹⁸ Paola Tubaro, Antonio A. Casilli e Marion Coville, "The trainer, the verifier, the imitator: Three ways in which human platform workers support AI", *Big Data & Society*, Gennaio 2020.

¹⁹ Carl Benedikt Frey & Michael A. Osborne, "The future of employment: How susceptible are jobs to computerisation?", *Technological Forecasting and Social Change*, Vol. 114, 2017, p. 254-280.

robot industriali sull'occupazione globale in termini di ore lavorate.²⁰ Anche senza prendere per buona la retorica industriale secondo cui la digitalizzazione e la robotizzazione stimolerebbero l'occupazione, basta confrontare gli indicatori sul livello di automazione e i tassi di disoccupazione nei paesi industrializzati del G20 nella prima metà del decennio 2010 per notare che quelli con il più alto tasso di automazione (Corea del Sud, Giappone, Germania...) hanno anche dei tassi di disoccupazione più bassi.²¹

Spunti di spiegazione ci sono forniti da economisti come David H. Autor, del MIT, il quale alla domanda “come mai di fronte a questa esplosione dell'automazione ci sono ancora tanti posti di lavoro?” risponde mostrando come, per due secoli, il rapporto tra occupazione e popolazione sia cresciuto continuamente, senza che il livello globale della disoccupazione aumentasse stabilmente. Secondo Autor, esiste un rapporto di complementarità profonda tra il gesto produttivo umano e il funzionamento delle macchine. La dialettica automazione/lavoro, sebbene non priva di tensioni, determina un aumento della domanda di lavoro.²² Un esempio lampante di questa complementarità si trova nel settore bancario, dove l'introduzione degli sportelli automatici nel periodo 1980-2010 ha portato a una riqualificazione, e non a una soppressione, di alcune categorie di dipendenti. La presenza dei bancomat ha, certo, stimolato la moltiplicazione delle agenzie, che sono aumentate del 43% dal 1990²³. Tuttavia, il numero dei cassieri nelle banche non è diminuito ma si è stabilizzato, grazie all'espansione del settore.

Che cosa è il *digital labor*?

Ciò che mi preme veramente qui è allontanarmi da questo tipo di approccio, che da molti anni non porta a risultati concludenti, per spostare l'attenzione sulle modalità effettive di *produzione* dell'IA, perché attardarsi solo sui presunti effetti della sua *diffusione* significa essere tributari di una retorica marketing, che ci presenta l'automazione come efficace e inoppugnabile. E invece dobbiamo prima di tutto chiederci che cosa sono le tecnologie intelligenti, se sono davvero quello che dicono di

²⁰ Georg Graetz e Guy Michaels, “Robots at Work”, IZA Discussion Papers, no 8938, Institute for the Study of Labor (IZA), 2015 <<http://EconPapers.repec.org/RePEc:iza:izadps:dp8938>>.

²¹ International Federation of Robotics, World Robotics Industrial Robots, rapporto 2016, <<https://ifr.org/worldrobotics/>>.

²² David H. Autor, “Why are there still so many jobs? The history and future of workplace automation”, *Journal of Economic Perspectives*, vol. 29, no 3, 2015, p. 3-30 (5).

²³ James Bessen, “Toil and technology”, *Finance and Development*, vol. 52, no 1, p. 16-19, 2015.

essere. E se lo sono, come sono prodotte. Così facendo, ci rendiamo subito conto che l'IA oggi è basata soprattutto sul *machine learning*, ovvero quel ramo dell'apprendimento statistico che, partendo da dati concreti, riesce a dedurre regolarità in maniera adattativa e iterativa. Per poter mettere in atto delle soluzioni di *machine learning*, c'è bisogno di creare, preparare e verificare dati.

E questo è il nodo del *digital labor*, nodo che si stabilisce all'interno della nostra produzione attuale di IA. L'automazione intelligente si basa sui dati, e, in assenza di questi ultimi, non ci sono robots che comunicano, non ci sono algoritmi che imparano, non ci sono soluzioni che automatizzano. Sono l'alpha e l'omega di questo tipo di automatizzazione. Ma i dati non cadono dal cielo. Non sono "dati", nel senso etimologico del termine. Sono piuttosto "prodotti" e la loro produzione è appunto delegata a collettività lavorative umane che effettuano ciò che noi chiamiamo lavoro digitale. La definizione che ne davo in un mio testo del 2015 ("Noi chiamiamo digital labor la riduzione dei nostri legami digitali a un momento del rapporto di produzione, la sussunzione del sociale nel commerciale nel contesto delle nostre pratiche tecnologiche.")²⁴, merita oggi di essere un po' rimaneggiata. Del digital labor do oramai una definizione più fattuale, ancorata al contesto empirico delle piattaforme digitali.

Sulle piattaforme digitali che la estraggono e la mettono a profitto, questa attività si suddivide in tre grandi famiglie (fig. 1). La prima è sicuramente la più conosciuta in Italia: il *lavoro on demand*. Si tratta di un insieme di mansioni svolte a livello locale, molto spesso come servizio alla persona, per il trasporto, la manutenzione, la logistica. Questo primo tipo di digital labor va sostanzialmente dall'occupazione dei conducenti di Uber a quella dei ciclo-fattorini di Deliveroo, al lavoro necessario per vendere beni e servizi su marketplace digitali e siti di e-commerce, da Etsy a AirBnB. Questa famiglia di attività, estremamente visibile e localizzata in contesti urbani, è inoltre *ostensiva*: effettivamente noi riconosciamo senza mediazioni che le mansioni svolte di una persona che guida per Uber sono del lavoro. Quello che non vediamo, la parte non-ostensiva della sua attività, sono in realtà i processi di produzione di dati che si addossano a questa parte visibile. Un lavoratore di Uber, in realtà, passa meno della metà delle sue giornate al volante. Il resto del suo tempo, lo passa sulla sua *app*, a produrre dati. Idem per i rider, e per tutti coloro che usano piattaforme di e-commerce.

E' così che ci ritroviamo, quasi impercettibilmente, in una seconda categoria di digital labor, chiamata *micro-lavoro*, realizzato da persone ingaggiate in forme

²⁴ Dominique Cardon e Antonio A. Casilli, *Qu'est-ce que le digital labor?* Bry-sur-Marne, Éditions de l'INA, 2015

estremamente parcellizzate, per qualche minuto appena, per fare del “data entry”, o delle ricerche internet, o del filtraggio di contenuti, o ancora per rispondere a messaggi di assistenza tecnica. In breve, sono miriadi di micro-mansioni di back-office effettuate da cottimisti e pagate somme irrisorie: da meno di un centesimo a qualche euro. I lavoratori europei o statunitensi vedono in questo tipo di lavoro una occasione per un reddito integrativo, o per sostenersi in situazioni di povertà e precarietà estreme. Invece per i residenti in paesi emergenti o in via di sviluppo, dove il salario medio può essere di appena venti o trenta dollari al mese, micro-lavorare diventa un’occupazione interessante, per quanto non inquadrata in un impiego formale, poiché le piattaforme che mettono in contatto aziende acquirenti e micro-lavoratori non riconoscono questi ultimi come loro dipendenti.

Considero il micro-lavoro il vero perno tra il lavoro visibile, e ciò nonostante piattaformaizzato, delle persone che producono on demand, e tutto l’universo, estremamente controverso e difficile da definire, del *lavoro dell’utente* dei social media o delle grandi piattaforme alla Google. Quest’ultimo rappresenta la terza famiglia. Questo schema concettuale ci permette anche di vedere come questi tre tipi di digital labor si situino su un continuum tra lavoro *sottopagato* (quello dei rider o degli autisti di Uber), il lavoro *micro-remunerato* (quello dei Turkers di Amazon) e quello *non pagato* (degli utenti di Facebook o Google) che alcuni, ancora adesso, si ostinano erroneamente a definire “gratuito”.

Quando vent’anni fa Tiziana Terranova scriveva che questo era *free labor*²⁵, la situazione era in realtà molto diversa: il web 1.0 non era ancora stato sottoposto alle dinamiche di predazione capitalista delle grandi piattaforme sociali e dei behemot dei dati, e persistevano “isole nella Rete” basate su un misto di volontariato e spirito do-it-yourself. Sinceramente ci vuole oggi uno sforzo politico, ed anche non poca cattiva fede, per continuare a caratterizzare il lavoro dell’utente estratto e recuperato da oligopoli digitali come del lavoro “gratuito”. Il dubbio più serio è provocato dalla prossimità economica tra questo tipo di digital labor non remunerato et quello delle persone che condividono gli spazi con gli utenti organici dei social, ovvero moderatori, *click farmers*, produttori di contenuto a cottimo. Tutta una fauna di proletari del web sociale, generalmente micro-pagati, ovvero retribuiti appena qualche centesimo o qualche euro per realizzare delle micro-prestazioni che durano qualche secondo, o qualche minuto. Per i lavoratori del clic di paesi del sud globale come l’Indonesia o il Bangladesh, queste micro-prestazioni possono essere pagate fino

²⁵ Tiziana Terranova, “Free labor: Producing culture for the digital economy”, *Social Text*, vol. 18, n. 2, 2000, p. 33-58 (33).

0,00001 centesimi di dollaro ciascuna. Ora, la domanda bruciante è: che differenza c'è fra un lavoro pagato 0,00001 centesimi e un lavoro pagato 0 centesimi, se non l'hybris degli utenti di piattaforme del nord globale che si vivono come volontari entusiasti del web sociale e non come lavoratori nel bisogno?

Fig. 1 - I tre tipi di digital Labor. Fonte: elaborazione dell'autore (2019).



Che cosa è, allora, il digital labor? Alla luce di questa tipologia, la mia risposta di oggi alla domanda che mi si rivolgeva nel 2015, è la seguente: una galassia di attività lavorative *parcellizzate e datificate, ovvero ridotte a micro-prestazioni orientate verso la produzione e lo sfruttamento di dati.*

Il valore in un regime di lavoro digitale

Questa definizione del digital labor dà adito ad una composizione specifica del valore. Attualmente possiamo vedere emergere tre tipi di valore-lavoro "generato dagli utenti" e estratto dalle piattaforme digitali: il *valore di qualificazione* (gli utenti organizzano l'informazione lasciando commenti o dando voti su beni, servizi e/o su altri utenti della piattaforma), che permette il funzionamento regolare delle piattaforme; il *valore di monetizzazione* (il prelievo di commissioni o la cessione lucrativa di dati forniti da attori ad altri attori), che fornisce liquidità a breve termine; il

valore di automazione (l'utilizzo di dati e contenuti degli utenti per addestrare le intelligenze artificiali), che è da considerare un investimento a più lungo termine.

Per inquadrare questa visione, dobbiamo innanzitutto superare la distinzione tradizionale tra valore creato nel processo di produzione e valore scambiato nella sfera della circolazione dei beni e dei servizi, ovvero sul mercato. Nella misura in cui le piattaforme sono degli ibridi aziende-mercati, questa distinzione, caratteristica sia degli approcci marxisti che della cosiddetta economia delle convenzioni, è stata messa in discussione da diversi autori. In particolare, il sociologo Michel Callon ha sottolineato che la costituzione del valore di un prodotto non dipende soltanto dal processo di fabbricazione ma anche dalla logica del valore di scambio che determina il trasferimento della proprietà.²⁶ Questo è tanto più vero sulle piattaforme digitali, dove i dati, i contenuti e i servizi sono fabbricati dagli utenti in vista della loro stessa circolazione. Ma questa circolazione “non è altro che una fase dell'intero processo, mai concluso, di qualificazione e riqualificazione dei beni, dal quale non può essere distinta”.²⁷

Nel contesto del settore digitale, questa logica si manifesta nell'attività stessa degli utenti che attraverso i loro scambi e le loro interazioni forniscono un contributo fondamentale per testare gli strumenti messi a disposizione dalle piattaforme e scoprire nuove possibili applicazioni per le soluzioni tecniche che utilizzano. Non bisogna immaginare questo lavoro come qualcosa che pertiene agli esperti: tutti gli utenti di Gmail hanno contribuito a migliorare, attraverso l'uso, il servizio di posta elettronica, restato in fase Beta per 5 anni; il gioco online Minecraft si basa sulla creazione di nuove situazioni e spazi da parte degli stessi giocatori; le interfacce delle principali piattaforme generaliste (Facebook, Instagram, Twitter) continuano a fare test proponendo ai loro membri diverse varianti della stessa interfaccia. Gli utenti hanno inoltre un ruolo cruciale nel produrre valutazioni (attraverso commenti, critiche, resoconti, eccetera) e metriche di valutazione (voti, like, stelline, eccetera) che costituiscono informazione e permettono di estrarre strati di dati supplementari su Airbnb, Upwork, Uber, Amazon... In questo modo contribuiscono a un processo ciclico di qualificazione-riqualificazione di ogni oggetto, che acquisisce valore a ogni stadio della sua metamorfosi.

²⁶ Michel Callon, “Postface. La formulation marchande des biens”, in François Vatin (a cura di), *Évaluer et valoriser. Une sociologie économique de la mesure*, Toulouse, Presses universitaires du Mirail, 2009, p. 247-269.

²⁷ *Ibid.*, p. 251

Malgrado l'insistenza sulla produzione di dati e di metriche, la genesi del valore non viene operata attraverso un processo di trasformazione della qualità in quantità. Gli utenti non si limitano a fornire feedback numerici sulla loro esperienza, ma operano un “lavoro di qualificazione” valorizzato dalla piattaforma. La principale attività dell'utente di un servizio digitale, ricorda altrove Michel Callon, “sarà di qualificare (ovvero classificare, valutare e giudicare) i prodotti che gli vengono messi a disposizione, confrontandoli o mettendoli in relazione”.²⁸ Questo lavoro di qualificazione si realizza grazie a una pianificazione dell'esperienza-utente che lo mette in condizione di accedere alle informazioni e di collegarle tra loro, così producendo del valore.

Altre attività degli utenti delle piattaforme servono a principalmente a estrarre valore di *monetizzazione*. Anche se non-poche piattaforme, soprattutto nel digital Labor on demand o nel micro-lavoro, captano questo valore attraverso complessi sistemi di commissioni e architetture di prezzi, la commercializzazione dei servizi, contenuti e dati monetizzabili costituisce un mercato miliardario orientato soprattutto al targeting pubblicitario. Un sistema di monetizzazione di questo tipo può essere illustrato dal noto programma Facebook Partners²⁹ che permette alla piattaforma di siglare accordi commerciali con grandi *data brokers*. Si tratta di aziende commerciali che aggregano e incrociano informazioni su centinaia di milioni di cittadini a partire da media, agenzie pubblicitarie, amministrazioni pubbliche, archivi di *open data* e logicamente anche da dati raccolti dalle aziende su Internet. Incrociando i proprio set di dati con quelli di partner specializzati nella monetizzazione di dati personali (Acxiom, Epsilon, Datalogix e altre) Facebook contribuisce a profilare le abitudini di 500 milioni d'internauti in tutto il mondo. Anche prima che grandi scandali come quello di Cambridge Analytica rendessero noti i rischi associati all'accumulazione dei dati dei social, questi abbinamenti di database si dimostravano particolarmente preoccupanti. I data brokers aggregano informazioni sulla salute, le opinioni politiche, gli orientamenti sessuali o le credenze religiose dei cittadini — e questo in barba agli appelli alla regolazione e alla responsabilità di queste organizzazioni davanti alla Commissione federale del commercio degli Stati Uniti.³⁰

²⁸ Michel Callon, Cécile Méadel e Vololona Rabeharisoa, “L'économie des qualités”, *Polittix*, vol. 13, no 52, 2000, p. 211-239.

²⁹ “Solutions Explorer”, *Facebook Business*, 2018, <https://www.facebook.com/business/solutions-explorer/campaign_management/>.

³⁰ Edith Ramirez, Julie Brill, Maureen K. Ohlhausen, Joshua D. Wright, e Terrell McSweeney, “Data brokers : A call for transparency and accountability”, rapporto della Federal Trade Commission, 2014, <<https://www.ftc.gov/system/files/documents/reports/data-brokers-call-transparency-accountability-report-federal-trade-commission-may-2014/140527-databrokerreport.pdf>>.

La *qualificazione* e la *monetizzazione* non esauriscono tutte le forme di estrazione del valore da parte delle piattaforme. Le informazioni prodotte dagli utenti sono utilizzate per addestrare gli algoritmi o costituire dei database necessari ad ambiziosi progetti di *deep learning*. Come lo spiega l'informatico Yann LeCun, direttore del FAIR (Facebook Artificial Intelligence Research) e pioniere di questa branca dell'IA, i progressi nel settore dell'apprendimento automatico o delle reti neurali artificiali non sono dovuti a scoperte o miglioramenti nel metodo occorsi negli ultimi anni, ma alla disponibilità massiccia di centinaia di milioni di *esempi* d'immagini, di testi, di suoni in milioni di categorie.³¹ Questa profusione di contenuti e di dati è necessaria perché, a differenza degli umani che sono capaci di apprendere a partire da pochissimi esempi, le macchine imparano moltiplicando considerevolmente le osservazioni.³²

Gli esempi prodotti dagli utenti tuttavia non sono altro che il primo passo dell'automazione. Per funzionare, i metodi di apprendimento automatico hanno bisogno di altri interventi da parte di umani che annotano, arricchiscono e organizzano i contenuti e i dati. Per questo alcune piattaforme incitano i loro utenti ad aggiungere dei tag sui loro contenuti. Ad esempio, quando su Instagram un utente attribuisce a una foto delle parole-chiave come “spiaggia”, “Ibiza”, “ombrellone”, “sdraio”, “estate”, “2017”, “dallamiastanza”, “sole”, “relax” sta descrivendo con grande precisione l'immagine — a tal punto che un software di riconoscimento potrebbe usare queste informazioni per capire il contenuto, il momento e il luogo esatto della fotografia, oltre che i sentimenti ad essa associati. Facebook, come Instagram o Oculus VR, ormai usano queste immagini etichettate su larga scala per riconoscere i prodotti nelle foto e personalizzare le offerte commerciali.³³

Ma le altre piattaforme di digital labor sono anch'esse capaci di estrarre valore d'automazione. Su Uber, tanto i conducenti che i passeggeri sono dei produttori di dati e dei beta tester per il suo ambizioso programma di veicoli autonomi. Poiché qualsiasi persona che utilizza la app di Uber è oggetto di registrazioni sistematiche, tanto le informazioni dei passeggeri che quelle dei conducenti e dei veicoli stessi sono cruciali nel processo di automazione. In fondo queste automobili non sono altro che oggetti su

³¹ Yann LeCun, “Deep learning”, comunicazione al convegno USI 2015, Paris, <https://youtu.be/RgUcQceqC_Y>.

³² Yann LeCun, “Qu'est-ce que l'intelligence artificielle?”, chaire Recherches sur l'intelligence artificielle, informatique et sciences numériques (2015-2016), Collège de France, <https://www.college-de-france.fr/media/yann-lecun/UPL4485925235409209505_Intelligence_ArtificielleY.LeCun.pdf>.

³³ Jérôme Marin, “Instagram se convertit au e-commerce”, *Le Monde*, 7 novembre 2016, <http://www.lemonde.fr/economie/article/2016/11/07/instagram-se-convertit-au-e-commerce_5026620_3234.html>.

ruote che raccolgono dati e li comunicano ad altre automobili, nonché ai server delle società di trasporto³⁴.

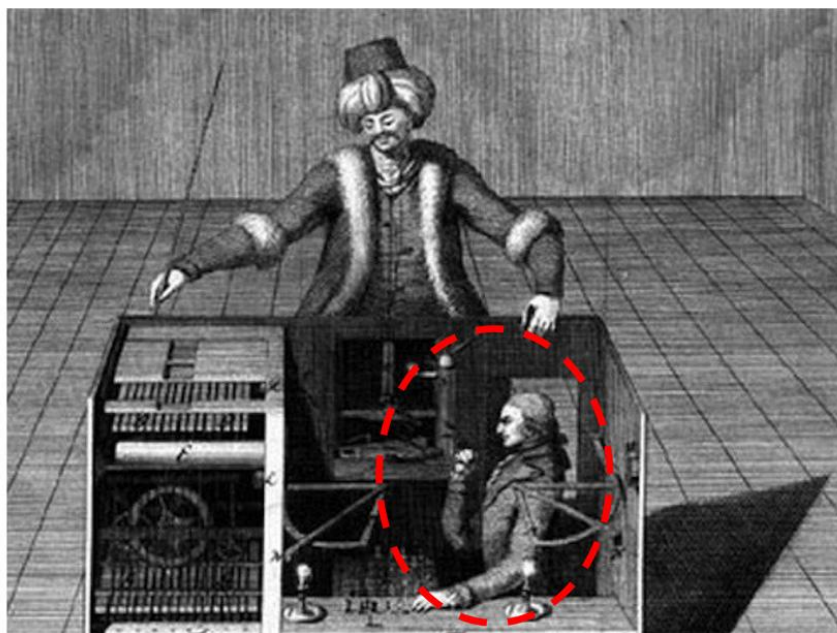
E ovviamente, una parte essenziale del modello economico delle piattaforme di micro-lavoro risiede inoltre nella creazione di ecosistemi che permettono di insegnare alle macchine a realizzare certe operazioni. Se gli utenti dei social e dell'economia on demand producono suoni, immagini o testi che riempiono i data center dei giganti del digitale, queste masse enormi di dati non-lavorati hanno bisogno di essere etichettate, normalizzate, spesso filtrate per eliminare errori e doppioni. Il lavoro di ripulitura dei dati è spesso la fase più delicata dell'apprendimento automatico - e quella che richiede più manodopera. Per sviluppare un sistema di traduzione automatica, per esempio, ci sarà bisogno di una massa di casi di conversazione più larga possibile nelle due lingue selezionate, nonché di un lavoro molto preciso di annotazione di ogni parola per risolvere i casi di polisemia e identificare ogni espressione idiomatica. È proprio qui che il microlavoro appare come un eccellente modo per ridurre i costi e il tempo dell'operazione. Dopo esser stati processati, i dati non saranno più grezzi, bensì dati di qualità sui quali i modelli matematici potranno addestrarsi.

Perché non ci sbarazzeremo mai del turco meccanico

Così presentata, l'utilità di questo lavoro umano di addestramento degli algoritmi sembra abbastanza astratta. Bisogna allora ancorarlo ad alcuni punti di riferimento storici e empirici, che ci permetteranno anche di chiarire di che tipo di automazione intelligente stiamo parlando: non quella dei robots industriali e delle intelligenze artificiali "generaliste", ma le più modeste soluzioni algoritmiche disponibili nei nostri oggetti connessi di consumo di massa. Tra le piattaforme del digital labor (e più in particolare del micro-lavoro, sul quale mi concentrerò) spicca il già citato Mechanical Turk, creato da Amazon nel 2005, sulla base di un brevetto che risale addirittura al 2001. Una piattaforma che ha oggi due decenni di vita, ispirata a sua volta da una metafora vecchia di due secoli e mezzo. Il "turco giocatore di scacchi" era un automa introdotto nel 1769 alla corte d'Austria dall'ingegnere Von Kempelen. All'imperatrice Maria Teresa venne presentato un automa antropomorfo capace di simulare i processi cognitivi umani di un giocatore di scacchi. Si trattava né più né meno della prima intelligenza artificiale.

³⁴ Paola Tubaro e Antonio A. Casilli (2019), "Micro-work, artificial intelligence and the automotive industry", *Journal of Industrial and Business Economics*, vol. 46, n. 3, p. 333-345.

Fig. 2 - Schema del “giocatore di scacchi” di Von Kempelen.



Fonte: Karl Gottlieb von Windisch, *Briefe über den Schachspieler Des Herrn Von Kempelen*, 1784.

Ma, allora come oggi, questa IA aveva ben poco di artificiale. Perché all’interno dell'automa si nascondeva in realtà un essere umano (fig. 2) . Le ipotesi sull’identità del suddetto essere umano sono molteplici e tutte problematiche: secondo alcune voci si sarebbe trattato di un militare che aveva perduto le gambe in battaglia, altri parlavano di un bambino molto magro che riusciva quindi con le mani a far muovere l'automa. Edgar Allan Poe, in un celebre saggio del 1836 parlava di un “italiano al servizio [di von Kempelen]”, “di taglia media” e “dalle spalle singolarmente curve”³⁵. Un secolo più tardi, Walter Benjamin parlerà di un “nano gobbo”.³⁶ Per stravaganti che fossero queste illazioni, in nessun caso era avanzata l’ipotesi che questo operatore umano fosse un gran maestro degli scacchi, un genio o un intelletto al di sopra del comune.

Questa, è forse la cosa più interessante (e più sfacciatamente cinica) del progetto Amazon Mechanical Turk. Con questo nome, Jeff Bezos sembra alludere al fatto che la sua piattaforma opera la stessa mistificazione che quella del turco meccanico del ‘700. Ma all’interno dei suoi processi automatici non nasconde una sola persona, bensì centinaia di migliaia di operai del clic, di ‘nani gobbi’, di lavoratori non specializzati e

³⁵ Edgar Allan Poe, *Il giocatore di scacchi di Maelzel*, SE, Milano, 2009, p. 17.

³⁶ Walter Benjamin, *Sul concetto di storia*, Torino, Einaudi 1997.

spesso esclusi dal mercato del lavoro. Chi sono questi “Turker”? Si tratta sostanzialmente, di persone che non hanno qualifiche né competenze specifiche. Non sono ingegneri, informatici, né esperti di *machine learning*. Si tratta di semplici utenti che vogliono guadagnare un complemento di reddito connettendosi ad una applicazione dove realizzano delle micro-mansioni (*micro-tasks*), ripetitive ed estremamente semplici: trascrivere decine di scontrini, organizzare playlists, annotare immagini, ecc.

Tutto questo serve ad addestrare degli algoritmi—di riconoscimento di forme o di testo, di interpretazione di lingue, di scelta di percorsi GPS, per esempio. Anche le autovetture autonome hanno bisogno di micro-lavoratori che, a partire dalle immagini registrate dal veicolo, distinguono un pedone, interpretano un cartello stradale, spiegano cosa fare davanti a un semaforo. Sono proprio i micro-lavoratori che, tutto il giorno, guardano delle immagini registrate da automobili e le classificano.

Amazon non è l’unica azienda a essersi dotata di una piattaforma di micro-lavoro. Molte delle grandi società del settore tech propongono oggi lo stesso tipo di servizio, di solito per migliorare la propria offerta. Microsoft, per esempio, ha creato un clone di Mechanical Turk che si chiama UHRS per parametrare i risultati di Bing. Apple si serve della piattaforma di micro-lavoro TryRating per ottimizzare la performance della sua app “Map”, e Alphabet ha un’altra applicazione dello stesso tipo, RaterHub, per calibrare Google Search.

Queste piattaforme sono necessarie per produrre l’IA e gli algoritmi di apprendimento automatico. La loro necessità si è imposta alla nostra comprensione come una costante e come un elemento ineliminabile alla fine della nostra ricerca a tutto campo sul micro-lavoro in Francia pubblicata nel 2019³⁷. Il digital labor umano alimenta l’IA in tre maniere diverse. La prima, già ampiamente documentata, consiste nella generazione dei dati e nell’*addestramento* dei modelli matematici. Ma il lavoro umano piattaformaizzato è anche utilizzato per *verificare* la pertinenza dei risultati dell’IA, nonché per *simularne* l’esistenza stessa³⁸.

Un esempio della prima funzione del digital labor ci è fornito dall’intervista approfondita realizzata con il fondatore di una start-up che aiuta gli sviluppatori di applicazioni intelligenti per oggetti connessi (smartphone, altoparlanti wireless, ecc). Se un software developer volesse creare una applicazione che permette al suo utilizzatore di pronunciare il titolo di una canzone a caso e di lanciare la versione

³⁷ AA. VV., “Le Micro-travail en France. Derrière l’automatisation de nouvelles précarités au travail?”, cit.

³⁸ Paola Tubaro, Antonio A. Casilli e Marion Coville, “The trainer, the verifier, the imitator”, cit.

karaoke della stessa, avrebbe bisogno di milioni di esempi di titoli di canzoni per costituire la sua base dati di addestramento. La startup mette allora a disposizione degli sviluppatori un programma che in apparenza è interamente automatico, con tanto di pulsante “*generate data*” ben visibile che permette di creare ex novo centinaia di migliaia di esempio. Ma ciò è possibile perché il pulsante si connette in realtà a un servizio a pagamento che fornisce dei “turchi meccanici”, ovvero sia di micro-lavoratori. Lo sviluppatore può scegliere di acquistare dati per 100, 250, 500 euro, ma nel back-office, al micro-lavoratore entrano in tasca solo i pochi centesimi corrispondenti ai secondi di tempo necessari per pronunciare uno o più titoli di canzoni³⁹.

Invisibilizzati, i micro-lavoratori realizzano le operazioni necessarie per fornire esempi a partire dai quali le applicazioni intelligenti apprendono. Ciononostante, il loro non è un lavoro di “professori” degli algoritmi, dotati di responsabilità pedagogiche, di autonomia e di una visione chiara delle mansioni che la macchina svolge. Al contrario, si tratta di attività opache, che distaccano i lavoratori dalle loro realizzazioni. Costituiscono chiari casi di alienazione e introducono rischi psico-sociali specifici. Una micro-lavoratrice intervistata nel contesto della nostra inchiesta descriveva la sua ultima mansione (preparare immagini per una applicazione di consigli dietetici) come un compito tanto misterioso quanto inutile: “Disegnare un quadrato intorno a un pomodoro... ci danno l’immagine di un piatto e ci dicono: ‘disegnate un quadrato intorno a un pomodoro’. Non sappiamo perché. Tutti sanno cos’è un pomodoro, spero... poi se c’è, penso debba essere utile a qualcuno per qualcosa, ma... non so perché”⁴⁰.

Alienazione, dunque: una persona distaccata dal senso della sua attività, che in questo caso specifico consisteva nell’etichettare immagini per una applicazione di consigli dietetici personalizzati: scatto una foto del mio piatto con lo smartphone e un modello matematico calcola calorie e coerenza con il mio fabbisogno nutritivo. Questo lavoro necessario di annotazione di grandi masse di dati viene frammentato fino a proporre ai micro-lavoratori diverse mansioni per ciascuna immagine: qualcuno si occuperà allora di identificare i pomodori, un altro l’insalata e così via.

Davanti a una procedura tanto rigorosa di preparazione di database di addestramento per le IA, la reazione potrebbe essere di aderire al discorso ideologico degli informatici e degli investitori della tech, che considera il *digital labor* come un

³⁹ AA. VV., “Le Micro-travail en France. Derrière l’automatisation de nouvelles précarités au travail?”, cit., p. 57.

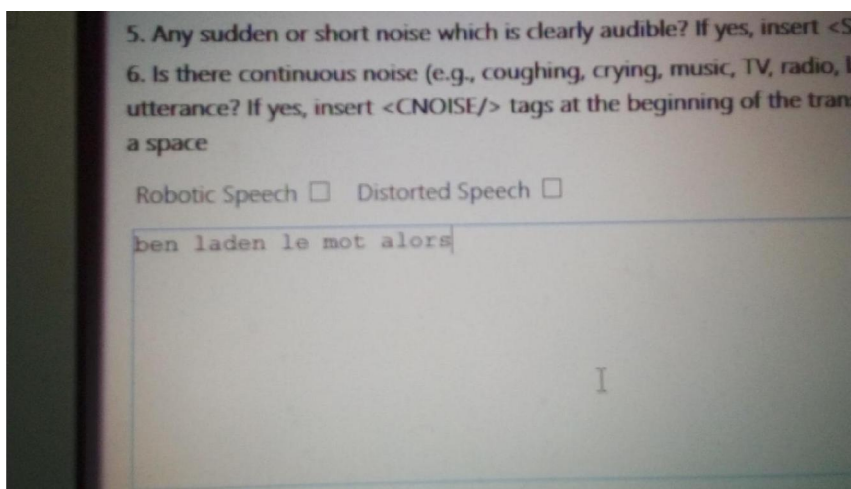
⁴⁰ *Ibid.*, p. 35.

fenomeno transitorio, destinato a scomparire una volta che tutte le IA del mondo saranno state addestrate, che tutti i modelli matematici alla base di queste tecnologie avranno appreso a partire da tutti i dati necessari. In quel momento non ci sarà più bisogno del lavoro umano, e si raggiungerà quella che viene spesso designata come “piena automazione”, che coinciderebbe davvero con la perdita di percentuali enormi di impieghi umani.

Ma la fede ingenua nella possibilità di addestrare una volta per tutte le IA si scontra con la realtà produttiva e commerciale di queste tecnologie. Anche quando esse esistono e funzionano, sono messe su mercati dove devono conquistare sempre nuovi segmenti di mercato e competere con altre applicazioni che presentano più features. Questo implica il bisogno costante di riaddestrarle, non solo perché imparino nuove lingue, nuove funzionalità, ma anche e soprattutto perché non commettano errori nel momento in cui si confrontano a una nuova operazione. Bisogna allora mobilitare un altro esercito di micro-lavoratori che, prima di tutto, verifichino e controllino i risultati delle IA. Se coloro che li addestravano potevano ancora essere considerati come dei collaudatori di prototipi, i verificatori sono coloro che si occupano della manutenzione di sistemi in attività.

Un'altra micro-lavoratrice da noi intervistata, svolgeva ad esempio un lavoro alquanto peculiare (e problematico) di correzione dei risultati prodotti da un assistente virtuale. Le micro-mansioni alle quali era adibita consistevano in piccoli estratti audio che duravano appena da 5 a 15 secondi. L'IA aveva già tentato di trascrivere le conversazioni, e alla micro-lavoratrice incombeva il compito di verificare che il trascritto corrispondesse a quanto contenuto nei file audio. La remunerazione era tutt'altro che generosa e i ritmi infernali: questa persona riceveva più di 100 task all'ora, una massa di lavoro che non poteva ovviamente svolgere alla cadenza stabilita dalla piattaforma. Oltre ai problemi di violazione della vita privata dei consumatori che questa attività può comportare (gli assistenti vocali hanno la fastidiosa abitudine di attivarsi in momenti inopportuni, catturando conversazioni intime, informazioni sensibili, segreti professionali), resta il fatto che la qualità delle trascrizioni trattate era piuttosto scadente. E questo può avere conseguenze negative sugli utilizzatori.

Fig. 3 - Foto di una schermata contenente una micro-mansione di verifica di trascrizioni automatiche. Fonte: inchiesta DiPLab.



Fonte: inchiesta DiPLab.

Una delle micro-mansioni trattate, per esempio, aveva prodotto in francese la trascrizione “*Ben Laden le mot alors*” (“Ben Laden la parola allora”). Frase sibillina e, nel contesto delle leggi ipersecuritarie della Francia, potenzialmente da segnalare alle autorità competenti. Tuttavia, dopo l’ascolto della micro-lavoratrice madrelingua la trascrizione è stata corretta in un molto più prosaico “*bien, donne-le moi alors!*” (“beh scusa, dammelo!”) (fig. 3), pronunciato forse da qualcuno che stava chiedendo ad un’altra persona di passargli il sale...⁴¹

Anche per funzionalità a basso impatto come scegliere una canzone o farsi suggerire un buon ristorante, l’implicazione continua dei micro-lavoratori che verificano i risultati dell’IA sembra tutt’altro che transitoria. Questo è ancora più vero per le funzionalità critiche: non lasceremmo senza verifica (anche perché è vietato dall’articolo 22 del GDPR) un sistema automatico che guidi un’automobile, scelga una terapia o assegni un mutuo. Ciò basterebbe ad ammettere che ci sono limiti naturali e sociali all’automazione completa. Ma un altro limite, che ancora una volta suggerisce

⁴¹ *Ibid.*, p. 58.

la perpetuazione del digital labor, è di natura economica: a volte sviluppare un'IA interamente funzionante costa molto di più che ingaggiare migliaia di micro-lavoratori. La terza ed ultima attività necessaria per la fabbricazione di IA, è quella che chiamiamo *A.I. impersonation*, ovvero l'imitazione di intelligenze artificiali.

L'esempio chiave è il caso che riporto in apertura del mio libro *Schiavi del Clic*: un giovane stagista in una startup che dice produrre raccomandazioni algoritmiche per prodotti di lusso per una clientela VIP, scopre in realtà che il lavoro è realizzato a mano da persone reclutate in Madagascar le quali, dopo aver raccolto informazioni personali sui profili social dei clienti potenziali, propongono beni e servizi. Fintanto che il lavoro di una miriade di operai del clic sarà meno caro di quello di una squadra di informatici specializzati nello sviluppo di soluzioni realmente automatiche, la startup non avrà nessuna ragione economica di sviluppare l'IA che dichiara di possedere già. “L'ideale sarebbe di metterla in cantiere”, riconosce uno dei suoi fondatori, “ma a questo stadio le richieste dei nostri clienti sono così numerose che dobbiamo concentrare i nostri sforzi sulla piattaforma esistente, per renderla più efficiente e redditizia.”⁴²

Comprensibilmente, lo stagista da noi intervistato interpretava questa situazione come un caso di pubblicità ingannevole. Ma la realtà è più complessa: anche le piattaforme ufficiali come Amazon Mechanical Turk sono basate su un principio di simulazione dell'intelligenza artificiale. In occasione del suo lancio nel 2006 al MIT, Jeff Bezos stesso presentava un caso in cui, al posto di sviluppare un'applicazione che genera raccomandazioni di attività ricreative a New York, il codice permettesse di pubblicare delle micro-mansioni per chiedere ai Turker di suggerire luoghi di uscita e ristoranti in tempo reale⁴³. Altre aziende chiamano questa maniera di funzionare “apprendimento in tempo reale” (*real time learning*). Infatti, non siamo molto lontani dalla situazione di addestramento per l'applicazione di karaoke citata qualche pagina fa, solo che in questo caso lo scopo non è di creare, a termine, una vera IA. Lo scopo è di fornire un servizio all'utilizzatore, con ogni mezzo necessario e al prezzo più basso.

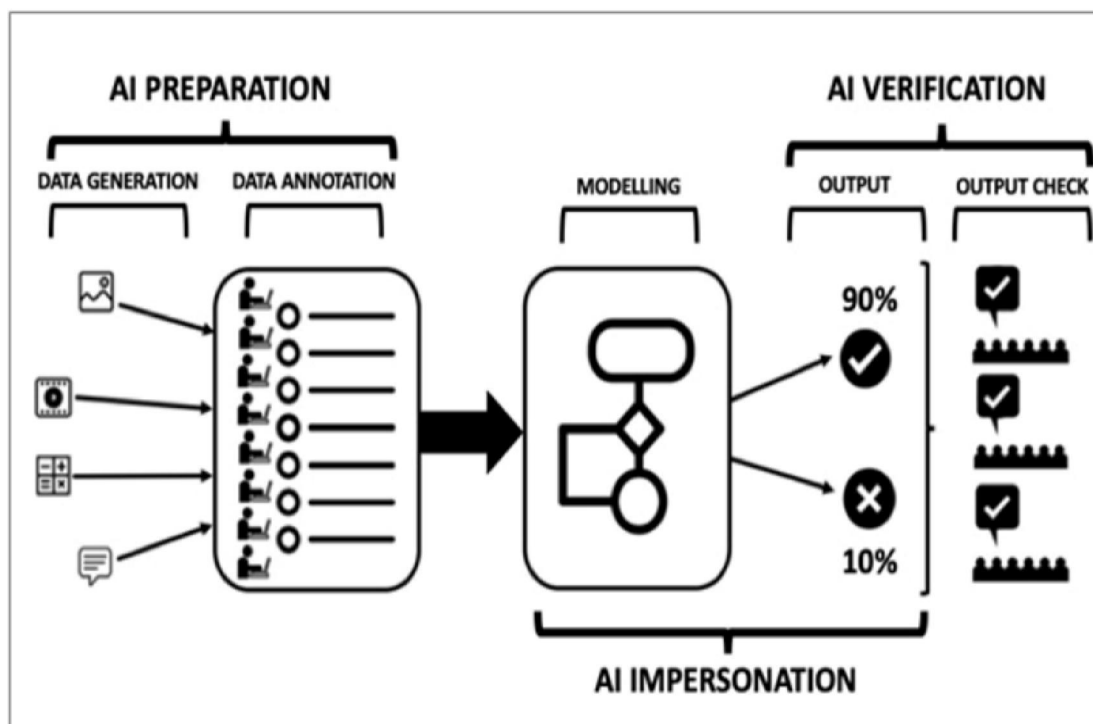
Lo schema in fig. 4 riassume quanto detto finora. Secondo i nostri risultati, l'automazione intelligente richiede una quantità considerevole di digital labor umano, e non solo in fase di preparazione e addestramento. La sua utilità anche a dei fini di verifica e di imitazione è confermata⁴⁴.

⁴² A. A. Casilli, *Schiavi del clic*, cit., p. 19.

⁴³ *Ibid.*, p. 116-119.

⁴⁴ Paola Tubaro, Antonio A. Casilli e Marion Coville, “The trainer, the verifier, the imitator”, cit.

Fig. 4 - Preparazione, verifica, imitazione: le tre funzioni del digital labor necessarie per lo sviluppo di soluzioni di IA.



Fonte: Tubaro, Casilli, Coville (2020).

Soprattutto l'imitazione può sembrare una soluzione complessa e estremamente costosa, ma fintanto che i micro-lavoratori sono pagati qualche centesimo e sottoposti a una concorrenza reciproca senza controllo, rimarrà sempre più conveniente dal punto di vista economico di produrre della "intelligenza artificiale artificiale" (la frase è sempre di Jeff Bezos), invece che costruire un modello matematico, creare un database, aggiornare il processo, correggere i risultati, ecc. E per abbassare i prezzi di questo lavoro, strumenti classici come l'offshoring o il dumping sociali sono utilizzati senza ritegno. Malgrado le piattaforme che producono questi servizi li presentino come "*Data with a human touch*" e mettano in avanti nella loro comunicazione aziendale modelli biondi e persone che sembrano residenti di paesi scandinavi, i loro

effettivi sono per la maggior parte situati in paesi emergenti o in via di sviluppo⁴⁵. Di fronte alle decine di milioni di lavoratori dichiarati in Bangladesh, Filippine, Sudafrica, Venezuela da piattaforme globali come Upwork, Zhubajie, FigureEight, Appen, Microworkers, impallidiscono i 50.000 Turker che micro-lavorano per Jeff Bezos o i 260.000 lavoratori del clic da noi stessi stimati in Francia, i quali appartengono spesso a classi popolari o sono membri di famiglie che vivono sotto la soglia di povertà⁴⁶. È un mercato globale dove le remunerazioni sono negoziate al ribasso.

Un altro imprenditore, intervistato nel contesto dell'inchiesta DiPLab, dichiara: “quello di cui la gente non si rende conto è che la stragrande maggioranza delle startup, soprattutto quelle business-to-business, che conosciamo, sono *human-based*, sono basate su degli esseri umani... Ma io capisco che per loro, è una scommessa sul futuro. Devono creare i propri dati e poi, su quella base, sviluppare dei modelli di machine learning, nella speranza che un giorno il processo sia del tutto automatizzato”⁴⁷.

È come in quell'adagio americano: “*Fake it until you make it*”, fai finta finché non ci riesci. Ma in che momento possiamo dire che l'intelligenza artificiale *c'è riuscita*? Buona parte di quello che viene presentato oggi nei media come automatizzazione dei processi produttivi (*business process automation*), è in realtà esternalizzazione di questi stessi processi produttivi (*business process outsourcing*). Invisibilizzando il digital labor, le imprese tecnologiche contemporanee stanno incoraggiando una vasta impresa storica di delegazione a paesi terzi o a popolazioni fragilizzate dei paesi del Nord globale, processi che hanno ben poco di automatizzato e che contengono una parte cospicua e ineliminabile di lavoro vivo.

⁴⁵ Mark Graham, Isis Hjorth e Vili Lehdonvirta, “Digital labour and development: impacts of global digital labour platforms and the gig economy on worker livelihoods”, *Transfer: European Review of Labour and Research*, vol. 23, n. 2, 2017, pp. 135-162.

⁴⁶ AA. VV., *Le micro-travail en France*, cit., p. 29, 69-72.

⁴⁷ *Ibid.*, p. 59.