

L'ECONOMIA DEI DATI NELL'INTELLIGENZA ARTIFICIALE

Una guida per combattere le diseguaglianze industriali
contro ogni esternalità economica e sociale



 seeweb

L'ECONOMIA DEI DATI NELL'INTELLIGENZA ARTIFICIALE

Una guida per combattere le diseguaglianze industriali
contro ogni esternalità economica e sociale



Pubblicato nel mese di Aprile 2024

PREMESSA

Quando abbiamo pensato ad una guida per far crescere l'offerta di servizi e di infrastrutture di intelligenza artificiale nel nostro Paese, sapevamo che il primo grande ostacolo era dissipare lo scetticismo che spesso accompagna chi ritiene questa sfida già persa in partenza. **L'Italia, pur avendo una riconosciuta tradizione di eccellenza scientifica ed una forte presenza di piccole e medie imprese innovative sul territorio, non ha ancora sviluppato un contesto industriale nel campo dell'AI.** Questo fattore si accompagna a ritardi nella definizione di obiettivi nazionali e di conseguenza, nell'allocazione efficiente delle risorse pubbliche che l'Italia si appresta a definire per il settore, rendendo quindi più complessa la possibilità di cooperare con chi sta già investendo nell'AI, primi tra tutti gli Stati Uniti, la Cina ma anche alcuni Paesi in Europa.

Si stima infatti che negli ultimi 10 anni, gli investimenti in AI siano cresciuti di 13 volte e conseguentemente siano nate oltre 41 mila imprese già impegnate nel settore. Di queste, la metà sono negli Stati Uniti (50,6%) mentre – in Europa – emerge la posizione di guida del Regno Unito (11,1%), che ha ampiamente superato la Germania (5,2%) e la Francia (3,6%)¹. E' evidente che c'è un alto livello di concentrazione del mercato in capo ad un ristretto numero di *player* globali localizzati in poche aree del mondo ed una conseguente difficoltà a promuovere la concorrenza man mano che i sistemi di intelligenza artificiale continuano a svilupparsi ed a diventare sempre più pervasivi nell'uso quotidiano.

Non condividiamo però l'approccio di chi rinuncia a competere con i giganti mondiali dell'innovazione. Una strategia nazionale su AI per permettere all'Italia di colmare questo divario, è in corso di definizione da parte del Governo e - nel solco della strategia europea - passerà probabilmente attraverso aiuti e sgravi fiscali per il settore al fine di incentivare la ricerca, lo sviluppo e la diffusione di infrastrutture e di applicazioni nazionali. Ma oltre all'uso da parte di cittadini, aziende ed istituzioni di modelli generativi globali, è fondamentale anche richiamare l'attenzione sulla necessità di sviluppare sistemi AI italiani per assicurare il progresso tecnologico e la sostenibilità del nostro ecosistema all'insegna di una reale indipendenza tecnologica - anche al fine di garantire una *governance* etica - aprendo ad una collaborazione tra tutti gli attori pubblici e privati coinvolti.

Non appare a tal fine necessario coltivare l'illusione di un solo "campione nazionale"² da finanziare per poi lasciarlo da solo a confrontarsi con gli scenari economici globali. In un contesto come quello italiano, meglio si attagliano forme di sostegno distribuite a più imprese, specialmente tra chi opera già nella filiera dell'AI, cercando di sfruttare appieno il potenziale acquisito fino ad oggi in termini di potenza di calcolo, di competenze specializzate e di infrastrutture cloud, lasciando che sia il mercato a selezionare i modelli migliori.

Ecco perché con questo studio vogliamo portare evidenze di natura competitiva, economica ed industriale come momento necessario di analisi di un fenomeno che se non opportunamente governato, rischia di travolgere la nostra economia. Nel discutere di come

¹ Cfr. Artificial Intelligence Index Report 2023, Stanford University https://aiindex.stanford.edu/wp-content/uploads/2023/04/HAI_AI-Index-Report_2023.pdf

² Cfr. Dario Denni - La sentenza sul Polo Strategico Nazionale ed il "Terribile diritto" di Stefano Rodotà: alla ricerca del bene comune <https://www.key4biz.it/la-sentenza-sul-polo-strategico-nazionale-ed-il-terribile-diritto-di-stefano-rodota-alla-ricerca-del-bene-comune/464809/>



l'intelligenza artificiale verrà utilizzata da istituzioni, cittadini ed imprese, saranno di seguito analizzate alcune proposte operative per garantire un rilancio dell'economia digitale italiana che massimizzi gli impatti positivi di questa tecnologia. **Riteniamo a tal fine che sia necessario un coinvolgimento più significativo delle aziende italiane impegnate nello sviluppo di sistemi e di infrastrutture di cloud computing e di intelligenza artificiale**, e che sia garantita la trasparenza nell'allocazione delle risorse impiegate con priorità al benessere dei cittadini e delle imprese del nostro Paese.

Un elemento chiave per spezzare il fenomeno oligopolistico che accompagna attualmente lo sviluppo di soluzioni di intelligenza artificiale, è quello di favorire l'accesso e la condivisione dei dati pubblici su cui si stanno allenando i sistemi di AI. Questi dati sono stati oggetto di un enorme accumulo nel corso degli ultimi 20 anni ad opera dei grandi operatori globali che sono attualmente gli unici a possederli in forma esclusiva ed inaccessibile. La disponibilità e l'utilizzo dei *Very Big Data* è diventato un prerequisito indispensabile per il successo dei servizi di Intelligenza Artificiale, in modo non diverso rispetto a quanto è accaduto in passato in molti altri mercati digitali, come quelli dei *social network*, dei motori di ricerca, della pubblicità online e dell'e-commerce. L'esistenza di un "collo di bottiglia" nel mercato dei dati, e la spontanea – e a volte inconsapevole – cessione dei nostri dati ai giganti globali di servizi cloud, sta limitando la concorrenza nel mercato immediatamente a valle dell'intelligenza artificiale, consentendo alle piattaforme globali di estrarre una quota sproporzionatamente elevata del valore creato nell'ecosistema, tale da permettergli di ampliare la dominanza in tutti gli altri settori adiacenti³.

Nello specifico, crediamo che l'uso di dati di pubblico dominio per creare dei *dataset* di addestramento, dovrebbe portare con sé l'obbligo affinché tali *dataset* siano resi accessibili a tutti. **Si dovrebbe a tal fine dare vigore alla figura dell'intermediario dei dati⁴ previsto dal Data Governance Act, come soggetto che potrebbe gestire queste relazioni di condivisione e garantire che i dati vengano utilizzati secondo i termini negoziati.** Le ricadute concrete e dirette sarebbero determinanti per favorire la competizione, perché si permetterebbe a nuovi soggetti di entrare nel mercato evitando che l'intelligenza artificiale si concentri oltremodo a vantaggio di pochissimi operatori globali. In ultima analisi, questo consentirebbe un controllo pubblico per prevenire raccolte abusive o illegittime e risolverebbe non pochi problemi di conformità alla disciplina europea sulla privacy.

Particolarmente oggi, di fronte ai drammatici effetti della crisi globale che colpisce in termini inflattivi, di approvvigionamento energetico, di approvvigionamento di *chip* e di altre materie prime, **bisogna fare ogni tentativo per definire una politica industriale più incisiva a vantaggio delle aziende italiane operanti nel settore del cloud computing come condizione per la fornitura di infrastrutture e servizi di intelligenza artificiale.** Questo perché gli sviluppatori di AI che non dispongono già di proprie infrastrutture e *data center* si rivolgono generalmente ai fornitori di servizi cloud (cd. "CSP") per l'elaborazione dei dati, ritrovandosi spesso ingabbiati in contratti pluriennali che aumentano la dipendenza tecnologica da fornitori esteri⁵ i quali - forti di offerte, sconti e *voucher* indirizzati a favorire un

³ Recentemente Google e Microsoft hanno integrato le funzionalità del modello di base AI nei loro prodotti già in commercio (i.e. Microsoft Office, Gmail e Google Documents ecc).

⁴ Sulla figura dell'intermediario dei dati e del Data Governance Act si legga <https://digital-strategy.ec.europa.eu/it/policies/data-governance-act-explained>

⁵ Alcuni esempi: 1) Google Cloud: Sconti per impegno di utilizzo (CUD) per Compute Engine <https://cloud.google.com/compute/docs/instances/committed-use-discounts-overview?hl=it> 2) AWS Savings plan <https://docs.aws.amazon.com/savingsplans/latest/userguide/what-is-savings-plans.html>



effetto cd. “*Lock-in*”⁶ - in questo momento stanno dominando il mercato globale. Riteniamo a tal fine che **si dovrebbe limitare l'importo dei cd. “crediti cloud” privati, offerti dai fornitori di servizi cloud ai clienti.** Come si vedrà nel seguito della trattazione, questi *voucher* teoricamente costituiscono un aiuto alle *startup*, ma tipicamente vengono poi utilizzati dagli operatori cloud dominanti per stroncare ogni concorrenza con altri fornitori di servizi cloud creando una dipendenza tecnologica – che poi diventa anche economica - in capo agli utenti, tra cui quelli che vogliono sviluppare sistemi di intelligenza artificiale.

In tal senso, va pure tutta la regolamentazione europea che nasce principalmente per evitare certe asimmetrie tecnologiche, informative e di potere. E' fortemente avvertito in Italia il richiamo alla normativa sull'abuso della dipendenza economica⁷ dalle piattaforme. Così come lo stesso regolamento *Platform to Business*, il *Data Act*, il *DMA*, il *DSA* sono tutti atti che non nascono perché c'è una “frenesia regolamentare a Bruxelles”⁸, come si vuol far credere, ma perché evidentemente il tema dimensionale e gli abusi che lo accompagnano, hanno davvero lasciato il segno.

Ci auguriamo, dunque, che questa guida sia d'aiuto affinché l'Italia possa avere un ruolo non marginale nello sviluppo dell'intelligenza artificiale. Per questo, se si vuole ridurre il divario dell'Italia rispetto al resto d'Europa, è più che mai **necessario che il programma di *public procurement* di beni e servizi digitali – ivi compresi i servizi cloud necessari per il settore pubblico nazionale e locale - vada a sostenere le imprese italiane ad alta tecnologia favorendone la crescita, lo sviluppo di competenze e l'assorbimento dei talenti.** Crediamo che questo possa accadere solo se sapremo indirizzare la commessa pubblica in modo da far sviluppare le aziende locali che sono impegnate negli investimenti di infrastrutture e servizi di intelligenza artificiale e di cloud computing. In questo modo si riuscirebbe ad accelerare la produzione interna di soluzioni innovative per dare sostanza alla transizione digitale in corso, ma anche al rilancio economico del Paese.

E' evidente che per raggiungere i risultati proposti, occorrerà adottare alcune misure correttive di alto valore sistemico⁹ perché solamente garantendo uno sviluppo industriale ed una adeguata formazione, si potranno minimizzare le esternalità negative che portano anche fenomeni di “brain drain”, ovvero di fuga dei cervelli e minimizzare gli impatti sul mondo del lavoro. Queste misure riguardano principalmente 1) la regolazione dell'accesso e del controllo dei *dati di input* utilizzati per allenare i sistemi di intelligenza artificiale, 2) la prevenzione delle distorsioni e delle concentrazioni del mercato da parte delle piattaforme dominanti extraeuropee 3) ed il sostegno alle aziende italiane della filiera dell'AI non solo incentivandone l'adozione ma soprattutto curandone l'intera filiera del valore partendo dalle infrastrutture digitali. Nel prosieguo, illustreremo nel dettaglio ciascuno di questi aspetti e le loro implicazioni, sperando di stimolare un dibattito costruttivo e propositivo sul futuro dell'AI in Italia.

⁶ Il Lock in tecnologico è dovuto principalmente al vantaggio iniziale delle soluzioni di cloud pubblico che prevedono voucher e sconti significativi rispetto ai modelli di data center tradizionali ma con il rischio di diventare dipendenti dai servizi solo di determinati fornitori nel momento in cui il passaggio da una tecnologia o da un fornitore a un altro risulta ostacolato dal provider, oppure difficile, dispendioso in termini di tempo e sproporzionatamente costoso.

⁷ Sul caso Antitrust italiano di abuso di dipendenza economica tra Meta e SIAE per interruzione delle trattative si veda <https://dirittoaldigitale.com/2023/12/06/siae-meta-agcm/>

⁸ Bruxelles Effect: https://en.wikipedia.org/wiki/Brussels_effect

⁹ Le 10 misure proposte in questo studio si trovano nella pagina finale del presente documento.



1] DATA DRIVEN ECONOMY: AI AFFAMATA DI DATI

“I dati sono diventati il fattore economico più importante dei nostri tempi. Il valore di mercato delle aziende più preziose del mondo si basa sui dati. L'economia dei dati è diventata un tema di tecnologia, commercio e politica estera e di sicurezza nei grandi Stati membri dell'Unione europea” (Cfr. Timo Harakka - “Data capitalism in a world of crises”)¹⁰.

I DATI DI ALLENAMENTO DEI SISTEMI DI INTELLIGENZA ARTIFICIALE

La digitalizzazione di beni e servizi ha cambiato profondamente il nostro sistema economico. L'economia dei dati abbraccia contemporaneamente molti settori, dalla tecnologia al commercio, dalla politica estera alla sicurezza degli Stati. Per questo, i *Big Data* si sono rivelati fondamentali per innovare e personalizzare i servizi offerti online. Ma a distanza di anni, il mercato dei dati resta ancora dominato dalle *BigTech* che hanno raccolto storicamente grandi quantità di dati, traendo un enorme vantaggio competitivo e determinando così il controllo del mercato dei servizi online (prima) e dell'intelligenza artificiale (adesso) in modo da creare una complementarità tra prodotti correlati. Di conseguenza, le aziende che producono i modelli di fondazione oggi più conosciuti, sono le stesse piattaforme globali che hanno già iniziato a raggruppare l'uso del loro modello generativo con i vari servizi che offrono sul mercato (es. integrazione *Copilot* nei prodotti Microsoft o di *Bard* nell'ecosistema Google).

Nel campo dell'intelligenza artificiale, i “*dati di input*” sono quei dati utilizzati per l'apprendimento automatico e si innestano nel processo decisionale dei sistemi di AI rendendoli autonomi e capaci di produrre risultati che si avvicinano a come gli esseri umani pensano, parlano e scrivono. Come riporta il sito ufficiale di Chatgpt¹¹, i sistemi di intelligenza artificiale si allenano con (1) informazioni disponibili pubblicamente su Internet, (2) con informazioni concesse in licenza da terze parti e (3) con informazioni fornite da utenti o da formatori umani. Già oggi è possibile addestrare i modelli AI utilizzando esclusivamente set di dati disponibili al pubblico¹², senza ricorrere a set di dati proprietari e/o inaccessibili. Ma quello che sta accadendo è che **i principali player globali stanno limitando l'accesso ai dataset raccolti negli anni, e questo sta portando come conseguenza diretta, ad un numero chiuso di operatori che li possono sfruttare pienamente.**

Questo spiega bene perché l'Intelligenza Artificiale si sta concentrando maggiormente negli USA e molto meno in Europa dove c'è molta regolamentazione e poca innovazione. Soprattutto spiega l'accelerazione incredibile di eventi legati all'intelligenza artificiale, in un tempo che in apparenza sembra molto ristretto. In un rapporto sul potenziale economico dell'intelligenza artificiale generativa scrive McKinsey¹³ “*ChatGPT è stato rilasciato nel novembre 2022. Quattro mesi dopo, OpenAI ha rilasciato un nuovo modello linguistico di grandi dimensioni, o LLM, chiamato GPT-4 con funzionalità notevolmente migliorate. Allo*

¹⁰ Cfr. Timo Harakka - Data capitalism in a world of crises - <https://timoharakka.fi/wp-content/uploads/2023/06/Timo-Harakka-Data-capitalism-in-the-world-of-crises.pdf>

¹¹ How ChatGPT and Our Language Models Are Developed - <https://help.openai.com/en/articles/7842364-how-chatgpt-and-our-language-models-are-developed>

¹² Cfr. LLaMA: Open and Efficient Foundation Language Models <https://arxiv.org/pdf/2302.13971.pdf>

¹³ Cfr McKinsey “The economic potential of generative AI: The next productivity frontier” <https://www.mckinsey.com/capabilities/mckinsey-digital/our-insights/the-economic-potential-of-generative-ai-the-next-productivity-frontier#introduction>



stesso modo, a maggio 2023, l'intelligenza artificiale generativa di Anthropic, Claude, è stata in grado di elaborare 100.000 token di testo, pari a circa 75.000 parole in un minuto – la lunghezza di un romanzo medio – rispetto a circa 9.000 token quando è stata introdotta nel marzo 2023”.

La disponibilità e l'utilizzo dei *Big Data* è diventato oggi un prerequisito fondamentale per il successo dei servizi di Intelligenza Artificiale, in modo non dissimile a quanto era già accaduto nei mercati dei *social network*, dei motori di ricerca, della pubblicità online e dell'e-commerce. Stiamo vivendo in un'economia *data-driven* con forti barriere all'ingresso¹⁴, alimentata dall'uso di *Big Data* e tecnologie proprietarie polarizzanti, regolate con norme continentali difformi, a fronte di pochi soggetti che dispongono di enormi capitali privati con i quali catturano ulteriori competenze scientifiche e *startup* innovative. Di conseguenza, anche **le acquisizioni di business basati sui dati sono in grande aumento, proprio perché attraverso le concentrazioni si implementano nuovi dataset specializzati che vanno ad affinare, confrontare e rafforzare le informazioni già raccolte.** Il valore che in questo modo va estraendo l'industria dell'AI è solamente un *quid tertium* rispetto a quanto ha già trovato pubblicamente sul web o raccolto dagli utenti dei servizi online.

INTESE RESTRITTIVE E ACCORDI DI ESCLUSIVA TRA CLOUD E AI

E' così che “l'economia dei dati” in pochi anni è riuscita a superare “l'economia della conoscenza” e ci ha catapultati direttamente sulla prossima rivoluzione industriale che sarà quella accelerata dall'intelligenza artificiale. Il valore di mercato delle aziende più preziose del mondo si basa dunque sui dati che sono il fattore economico più importante dei nostri tempi. Se dovessimo limitarci alle sole prospettive di sviluppo dell'IA generativa, si stima una dimensione tra i 2,6 ed i 4,4 trilioni di dollari all'anno, tra il 2,6 ed il 4,4% del PIL mondiale¹⁵.

Ricordiamo che in pochissimo tempo, la valutazione di OpenAI è arrivata a 80 miliardi di dollari¹⁶ e c'è un interesse della Commissione Europea ad aprire un'indagine¹⁷ per esaminare alcuni degli accordi conclusi tra grandi attori del mercato digitale e gli sviluppatori di AI generativa, quindi probabilmente anche sul ruolo di Microsoft¹⁸ in tutta la vicenda societaria e finanziaria di OpenAI, tra cui quella che ha portato all'allontanamento temporaneo di Sam Altman in modalità del tutto inedite e che non trovano precedenti specifici nemmeno sotto il profilo del suo reintegro¹⁹. **L'investimento di Microsoft in OpenAI ha dato a Microsoft - tra le altre cose - il diritto di essere il fornitore esclusivo di servizi cloud per OpenAI.** Recentemente Microsoft ha deciso di investire anche in Mistral AI, una start-up francese fondata meno di un anno fa. OpenAI ha un modello di licenza più chiuso e un investimento finanziario significativo da parte di Microsoft (stimato in circa 13miliardi di dollari), Il suo modello GPT-4 è considerato una “scatola nera”, in cui i dati e il codice utilizzati per costruire il modello non sono accessibili a terzi. All'opposto, Mistral AI adotta un modello *open source* ed una partnership che amplia ulteriormente la

¹⁴ Le barriere all'ingresso riducono la concorrenza a svantaggio dei consumatori, perchè aumentano il costo di ingresso nel mercato di nuovi operatori che per entrare nel mercato devono investire di più in una parte della catena del valore.

¹⁵ dati McKinsey, ripresi da Oecd, settembre 2023.

¹⁶ Cfr. OpenAI nach Deal wohl mit 80 Milliarden Dollar bewertet <https://www.handelsblatt.com/technik/it-internet/bewertung-openai-nach-deal-wohl-mit-80-milliarden-dollar-bewertet/100015708.html>

¹⁷ Commission launches calls for contributions on competition in virtual worlds and generative AI https://ec.europa.eu/commission/presscorner/detail/en/IP_24_85

¹⁸ Cfr. Euractiv - “EU Commission to examine Microsoft-OpenAI partnership”

<https://www.euractiv.com/section/competition/news/eu-commission-to-examine-microsoft-openai-partnership/>

¹⁹ Cfr. “Sam Altman on being fired and rehired by OpenAI” <https://www.theverge.com/2023/11/29/23982046/sam-altman-interview-openai-ceo-rehired>



dominanza di Microsoft (di ulteriori 2 miliardi di dollari) nel settore dell'intelligenza artificiale generativa²⁰.

Più in generale, l'integrazione verticale²¹ di alcuni attori digitali già presenti sul mercato con il loro ecosistema di servizi online, che intervengono con partecipazioni economiche rilevanti in sistemi che sviluppano intelligenza artificiale, sta conferendo a queste imprese un potere tale da escludere i potenziali concorrenti o l'ingresso sul mercato di altre imprese attive nello stesso settore. Vediamo come.

MERGER CONTROL: FUSIONI E ACQUISIZIONI BASATE SUI DATI

I sistemi di intelligenza artificiale generativa hanno bisogno continuamente di nuovi dati "di addestramento" per imparare e questo comporta significativi impatti sulla privacy delle persone, sulla protezione dei consumatori e sulla salvaguardia della concorrenza. Da alcuni anni, anche le Autorità antitrust europee e statunitensi hanno posto in evidenza gli effetti sul mercato di un'economia basata sui dati e stanno ripensando le politiche della concorrenza²² come ad esempio le norme che regolano le fusioni²³ tra le aziende attive nell'economia digitale.

Questo perché le *BigTech* stanno procedendo con acquisizioni nei mercati ancillari - verticali ed orizzontali - creando così concentrazioni rilevanti ed aumentando a dismisura la dominanza nei settori in cui già dispongono di enormi set di dati di addestramento: dati predittivi, dati personali, dati aziendali, dati finanziari, dati di posizionamento, dati generici. Per questo motivo - come analizzeremo più avanti - le autorità antitrust dovrebbero prestare molta attenzione alle acquisizioni effettuate da società dominanti nel cloud e nell'AI, in particolare delle *startup* che potrebbero competere con loro.

L'evoluzione delle regole europee e la loro applicazione per la protezione dei dati personali ed il diritto della concorrenza applicato ai mercati digitali, hanno accumulato un discreto ritardo nel regolare questi fenomeni. E' chiaro che l'accumulo di dati attraverso metodi di raccolta e di acquisizione automatizzata, spesso non sono aderenti alla disciplina sulla privacy, sta rafforzando posizioni di dominanza di mercato in capo a pochi soggetti extraeuropei che sono gli stessi che già oggi possono fornire servizi "gratuiti" (o a pagamento) di intelligenza artificiale.

In questo contesto, **Big Data e Big Analytics svolgeranno sempre di più un ruolo fondamentale nel processo decisionale di molte aziende** che già oggi stanno cercando di acquisire velocemente un vantaggio competitivo in termini di maggiore disponibilità di dati rispetto ai rivali, anche al fine di poter fornire servizi migliori e più performanti. Lo sviluppo dell'AI generativa richiede una forte concentrazione di elementi essenziali tra cui una grande potenza di calcolo (*datacenter*, microprocessori GPU ecc), un'enorme quantità di dati ed un pool di competenze tecniche di notevole specializzazione. Inoltre, i sistemi di AI devono disporre di un'architettura cloud sofisticata ed immediatamente disponibile perché è questa la preconditione tecnologica di ogni avanzamento applicativo nell'intelligenza artificiale.

²⁰ Sul punto si legga: "Microsoft, dopo OpenAI c'è Mistral: prosegue la corsa all'intelligenza artificiale" https://www.ilsole24ore.com/art/microsoft-openai-c-e-mistral-prosegue-corsa-all-intelligenza-artificiale-AFJUCXrC?TB_iframe=true&refresh_ce=1

²¹ L'integrazione verticale si riferisce alla combinazione di diverse fasi della produzione sotto un'unica azienda o gruppo.

²² Cfr. Maurice E. Stucke and Allen P. Grunes, *Big Data and Competition Policy* (Oxford: Oxford University Press, 2016) https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2849074

²³ Cfr. UE Mergers Notices and Guidelines https://competition-policy.ec.europa.eu/mergers/legislation/notices-and-guidelines_en



LA FILIERA DELL'INTELLIGENZA ARTIFICIALE: IL RUOLO DEL CLOUD

Come recentemente ha scritto il direttore de "La Repubblica" Maurizio Molinari²⁴: *"È chiaro che le app sono il campo più competitivo al momento, ma per regolare l'IA bisogna impedire che le grandi aziende tecnologiche dominino l'intera filiera verticale dei livelli della catena produttiva più avanzata della storia umana"*.

Quando si parla di *Big Data*, spesso sottolineiamo l'aspetto della quantità e sottovalutiamo quello della qualità dei dati, ovvero la loro accuratezza e pertinenza. La qualità dei dati e le dimensioni del *set di addestramento* influiscono entrambi sui risultati di un sistema di intelligenza artificiale. Questo ci porta ad una nuova consapevolezza: i dati di *input* di un sistema di intelligenza artificiale sono importanti quanto i dati di *output*²⁵. I risultati dell'elaborazione di enormi volumi di dati non sono quelli che ci si può aspettare razionalmente, anche perché **spesso i sistemi di intelligenza artificiale restituiscono risultati imprevedibili, sorprendenti e persino inspiegabili a chi ci lavora**. A questo proposito, si è affermato²⁶ che *"lo sviluppo dei Big Data può avere ripercussioni su libertà e diritti fondamentali delle persone se non accompagnato da garanzie e da principi etici che li rendano compatibili con i valori delle società democratiche. [...] In generale, una Big Data strategy deve essere predisposta nella piena consapevolezza dei limiti, fissati anche dalle norme in materia di protezione dei dati personali, alle analisi conducibili e ai risultati ottenibili, per evitare di ledere diritti e libertà fondamentali delle persone"*.

Pertanto, i grandi attori digitali che già controllano gli elementi essenziali per formare sistemi AI - oppure sono dominanti nei mercati adiacenti come ad esempio quello del cloud computing - stanno via via consolidando il loro attuale potere di mercato a monte²⁷ della catena del valore, per andare successivamente a svilupparsi a valle, ovvero nel settore dell'intelligenza artificiale che è in forte espansione. **Come è noto, il mercato del cloud appare estremamente concentrato attorno ad un ristretto gruppo di operatori globali che - in questo momento - coincidono con i pionieri dell'intelligenza artificiale**. Soddisfatta la richiesta di potenza computazionale, si rende poi necessaria la disponibilità di grandi quantità di dati (personali e non personali) per addestrare i modelli di base. E per questa ragione, i grandi *player* non hanno intenzione di condividere l'accesso ai dati indicizzati sui loro motori di ricerca o accumulati attraverso i loro servizi online, ma tendono a riservarli per proprio uso, oppure a concludere accordi di esclusività con aziende partner emergenti, nell'ottica di mantenere il perimetro tecnologico ed informativo anche delle più promettenti *startup* innovative.

RISCHI E PERICOLI DELL'INNOVAZIONE SENZA CONSENSO

Tutto ciò è avvenuto perché l'innovazione "senza consenso", quella *disruptive*, si incontra e si scontra con i principi cardine delle democrazie evolute. Scarica sulla società tutte le externalità negative che in questo studio potremmo solo enumerare vista la loro grande

²⁴ Cfr. Maurizio Molinari – La Repubblica - "Evitare l'oligopolio nell'Intelligenza artificiale"

https://www.repubblica.it/editoriali/2024/01/20/news/intelligenza_artificiale_evitare_oligopolio-421929452/

²⁵ Si veda principio "Garbage in, garbage out" https://it.wikipedia.org/wiki/Garbage_in,_garbage_out

²⁶ Cfr. Cosimo Comella, Origine dei "Big Data" <https://www.sicurezza nazionale.gov.it/sisr.nsf/wp-content/uploads/2017/07/gnosis-2-2017-Comella.pdf>

²⁷ L'integrazione verticale a monte si riferisce all'integrazione con i fornitori di potenza computazionale (i.e chip maker, datacenter ecc.) e dati (utenti, merger, ecc). L'integrazione verticale a valle si riferisce all'integrazione con produttori di modelli perfezionati o con i distributori o con i produttori di beni e servizi finali.



portata in termini etici, giuridici, economici e sociali, stanno perpetuando i pregiudizi, le ingiustizie e le discriminazioni presenti in ogni società.

La motivazione spesso addotta - ma che non scusa e non giustifica - è che l'innovazione apporterebbe 1) vantaggi all'utenza 2) prezzi più bassi o gratuiti 3) nuovi servizi 4) grande vantaggio per l'umanità. Questo ragionamento apodittico trascura per partigianeria tutti quelli che sono i rischi - tra cui quelli legati alle pratiche di trattamento illecito dei dati - ivi compreso il *data scraping* che allena i sistemi AI che affronteremo nel prosieguo della trattazione.

Come fare, quindi, per impedire che le grandi aziende tecnologiche dominino l'intera filiera verticale dei livelli della catena produttiva dell'intelligenza artificiale? Quando guardiamo a questo campo, la maggior parte dei principali set di dati sono estremamente concentrati in pochi *player* del mondo e ne vediamo chiaramente il risultato.

Occorre dunque agire sull'efficienza allocativa della risorsa principale dell'AI che è rappresentata dai *Big Data* su cui si allenano i sistemi di intelligenza artificiale, perché in un mercato così concentrato e chiuso sia a valle che a monte, si può agire solo aumentando il numero di soggetti economici per poter migliorare la concorrenza ed il benessere dei consumatori.

Garantire la concorrenza nel campo dell'intelligenza artificiale preserverà anche la democrazia, ci dicono gli esperti²⁸. Il mercato da solo ha mostrato di non essere in grado di correggere queste asimmetrie e per questo i grandi *player* continuano a dominare, quasi indisturbati, nello scenario globale, estendendo la loro dominanza sui mercati adiacenti come quelli aperti dall'intelligenza artificiale. In questo modo, l'AI non resterà a lungo un prodotto autonomo sul mercato, ma si candida a diventare una *feature*, una funzionalità aggiuntiva di sistemi e prodotti già in uso sul mercato digitale.

²⁸ Cfr. "Ensuring competition in AI will also preserve democracy, experts say" - <https://www.euractiv.com/section/competition/news/ensuring-competition-in-ai-will-also-preserve-democracy-experts-say/>



2] IL PERIMETRO SOGGETTIVO DEL FENOMENO OLIGOPOLISTICO LEGATO ALL'AI

L'ultimo anno ha mostrato con grande evidenza la rapida evoluzione di sistemi AI promossi quasi esclusivamente da soggetti giuridici globali extraeuropei, ben noti alle Autorità di tutto il mondo, ed accomunati tutti da una serie di caratteristiche oligopolistiche che hanno permesso finora di sfruttare l'enorme vantaggio competitivo accumulato negli anni.

ELEMENTI CHE DETERMINANO LA CONCENTRAZIONE DEL MERCATO

1) capitali in misura maggiore dei *venture capital*²⁹, liquidi e pronti ad investire nella prossima *startup-AI* che in pochissimo tempo riesce ad acquisire finanziamenti multi miliardari con cui fare sviluppo di nuovi sistemi LLM, andando anche a conquistare le migliori risorse umane³⁰ disponibili sul mercato, a volte sottraendole a livello locale ad altri soggetti economici, compresi quelli pubblici ovvero di coloro che dovrebbero esercitare il controllo;

2) data center enormi, numerosi, distribuiti a livello globale, dotati dei migliori *chip avanzati* che a volte si autoproducono³¹ o di cui fanno incetta³² selezionando i migliori sul mercato³³ contribuendo pro-quota al cd. *Shortage* che è un fenomeno sicuramente legato alla concentrazione della produzione in poche aree mondiali ma non solo, in quanto interessa anche la finalità della produzione (i.e. effetti rete diretti ed indiretti);

3) di *very large big data* accumulati negli ultimi 20 anni attraverso servizi verticali offerti in forma quasi-gratuita online per conquistare il cliente in cambio del tracciamento della sua vita online e offline, spesso attraverso un consenso che è stato coartato in quanto non offre reali opzioni, e recentemente con un sistema "*pay or consent*³⁴" di dubbia legalità in quanto sembra non consentire un diritto di rifiuto, né un mezzo chiaro, conveniente e non faticoso per interagire con i servizi digitali senza dover rinunciare alla cessione delle informazioni personali;

4) di un ulteriore perfezionamento dei dati raccolti con nuove acquisizioni delle *startup* più promettenti, potenziali *competitor*, un fatto che ha sicuramente rilevanza antitrust ma anche di tutela dei dati personali acquisiti con i nuovi *merger* che vanno a perfezionare - attraverso diverse tecniche di trattamento - i *data lake* chiusi che sono già in loro possesso e sono inaccessibili;

5) ed espandere la dominanza su altri servizi verticali e orizzontali che in questo modo hanno dato vita prima a intere piattaforme e subito dopo a quegli ecosistemi che oggi

²⁹ Big Tech is spending more than VC firms on AI startups - Microsoft, Google, and Amazon have crowded out traditional Silicon Valley investors. <https://arstechnica.com/ai/2023/12/big-tech-is-spending-more-than-vc-firms-on-ai-startups/>

³⁰ Spesso gli esperti AI sono assunti con clausole di non-concorrenza. Sul punto si veda la proposta dell'antitrust americana "FTC Proposes Rule to Ban Noncompete Clauses, Which Hurt Workers and Harm Competition" <https://www.ftc.gov/news-events/news/press-releases/2023/01/ftcproposes-rule-ban-noncompete-clauses-which-hurt-workers-harm-competition>

³¹ Es. Graphcore IPU (Intelligence Processing Unit), Amazon Inferentia, Microsoft Athena e Google TPU

³² Cfr. FT - "Intel to manufacture chips for Microsoft as AI drives demand" <https://www.ft.com/content/be22fa9d-d5ff-415d-b7d0-1b756751e344>

³³ NVIDIA è attualmente il principale fornitore di GPU utilizzate per scopi di intelligenza artificiale - NVIDIA H100 Tensor Core GPU <https://www.nvidia.com/en-us/data-center/h100/>

³⁴ Per approfondimenti si veda "Noyb files GDPR complaint against Meta over "Pay or Okay" <https://noyb.eu/en/noyb-files-gdpr-complaint-against-meta-over-pay-or-okay>



rivaleggiano contendendosi l'innovazione utilizzata come arma da opporre a chiunque contesti il metodo illecito utilizzato per acquisire la dominanza ora spiegata.

Dato il ruolo significativo dell'innovazione nel promuovere la crescita economica e il benessere generale, l'applicazione delle norme antitrust ha finora cercato di promuovere l'innovazione anche a scapito della concorrenza sul mercato. Il presupposto è che le forze di mercato e la domanda dei consumatori costringano le aziende ad investire di più in ricerca e sviluppo e la società ne trarrebbe vantaggio dall'aumento dell'efficienza e dal conseguente abbassamento dei prezzi. L'OCSE, ad esempio, ha definito l'innovazione come un *“prodotto (bene o servizio) nuovo e significativamente migliorato”*³⁵.

Vedremo nel prosieguo dell'analisi quante e quali conseguenze sociali ed economiche porta l'innovazione senza consenso cd. *“disruptive innovation”*.

Tutto quanto sopra espresso e sintetizzato rappresenta il quadro soggettivo che definisce a maggior ragione il perimetro oggettivo di una serie di interventi delle Autorità che ad ogni titolo, non solo in Italia ed in Europa, si sono interessate di numerose vicende patologiche legate in maniera concatenata, ad infrazioni privacy, cybersecurity, antitrust, tutela consumatore, tutela del diritto d'autore e non da ultimo di concorrenza fiscale.

LA CONTESA DEI TALENTI AI: IL RUOLO DELL'UNIVERSITÀ

Tuttavia sta aumentando il divario e le esternalità negative che si abbattano sulla società³⁶.

Si tenga conto che i soggetti sopra descritti hanno acquisito una formidabile tecnica di ingegneria giuridica avendo a disposizione centinaia di giuristi, professori, avvocati, economisti capaci di opporsi con successo ad ogni provvedimento su ogni mercato geografico e tecnologico, e sono in grado di disporre del più grande numero di “rappresentanti di interessi” (i.e. lobbisti e associazioni di interesse anche costituite ad hoc) particolarmente efficaci e che talvolta vedono il coinvolgimento anche delle istituzioni pubbliche quali **le Università che rappresentano il crogiolo dove si allenano i talenti richiesti dal mercato. Talenti che spesso vengono “forgiati” per lavorare unicamente sui sistemi proprietari oggi dominanti, tutti extra-europei**, di fatto escludendo una conoscenza di tecniche ulteriori e diverse³⁷.

Riteniamo a tal fine che il ruolo delle Università sia centrale nella formazione per il mercato ma occorre principalmente considerare che attualmente l'Italia vive la rivoluzione innescata dalla tecnologia dell'AI in maniera quasi passiva, come se dovessimo necessariamente identificare le soluzioni estere come un modello univoco a cui indirizzare le capacità dei soggetti che si avviano ad intervenire nei diversi settori produttivi della nostra economia.

Una serie di fattori comunemente osservati porta a queste considerazioni:

- il progressivo depauperamento degli iscritti ai corsi in materie STEM può essere considerato in ragione del calo demografico e di insufficienze nel corretto indirizzamento ai corsi di laurea;

³⁵ Cfr. Oslo Manual: Guidelines for Collecting and Interpreting Innovation Data, pag.46
<http://dx.doi.org/10.1787/9789264013100-en>

³⁶ Cfr. Chi controlla l'intelligenza artificiale controlla il futuro <https://www.redhotcyber.com/post/chi-controlla-lintelligenza-artificiale-controlla-il-futuro-e-ora-di-scegliere-da-che-parte-stare>

³⁷ La formazione dei talenti su AI sui modelli degli operatori dominanti crea due ordini di problemi: il costo di re-skilling su modelli diversi ed il costo per acquisire tali talenti che tipicamente vengono sottratti ad altre aziende del settore.



- ISTAT stima che 337.000 giovani laureati hanno lasciato l'Italia tra il 2012 ed il 2022 e questo è un dato estremamente rilevante per gli impatti specifici sul settore dell'innovazione;
- il tema salariale porta con sé probabili ricadute anche sul fatto che molti brillanti laureati sono attratti da offerte di aziende multinazionali e quindi lavorano dall'Italia ma per imprese che operano dall'estero in più mercati;
- l'assenza di un contesto industriale dedicato all'intelligenza artificiale comporta inoltre la scarsa attrattiva dell'Italia rispetto ad eccellenze che potrebbero arrivare nel nostro Paese arricchendo l'offerta di personale specializzato qualificato da altri Paesi.

Occorrerà quindi agire su diversi livelli di intervento per poter superare queste criticità ed intendiamo sottolineare il ruolo della commessa pubblica di servizi indirizzati ad aziende italiane della filiera dell'intelligenza artificiale perché facendo crescere le economie locali specializzate nelle nuove tecnologie si innesca un effetto immediatamente attrattivo anche per i talenti che si formano nelle nostre università.



3] LA STRATEGIA EUROPEA SUI DATI E LA RAPIDA OBSOLESCENZA DELLE REGOLE

L'Unione europea ha da tempo approvato un programma politico per il decennio digitale, in base al quale ogni Stato membro ha sviluppato la propria strategia digitale (cd. “*Digital Compass*”) per cercare di raggiungere gli obiettivi proposti entro il 2030. Ma nonostante sia riconosciuto da tutti il valore dei dati, non abbiamo ancora una vera strategia sui dati che possa favorire il mercato interno e - di conseguenza - facciamo fatica ad inseguire il mondo che cambia. Il sistema di regole che abbiamo pensato finora – anche nelle speculazioni più recenti - potrebbe non essere in grado di cogliere la specificità del fenomeno AI su cui è nato da poco un dibattito globale tra etica e diritto. Molte delle problematiche finora sollevate - consenso, equità, proprietà intellettuale e simili - non hanno trovato una risposta convincente da parte dei Governi europei.

ELEMENTI DI RITARDO NEL PROCEDIMENTO REGOLAMENTARE

Si calcola che il solo AI ACT necessiterà di 20 atti delegati implementativi e almeno 6 linee guida per una corretta attuazione. Un tempo non compatibile con lo sviluppo e la commercializzazione dei sistemi di AI ed al tempo stesso, una finestra di opportunità temporale che crea un enorme vantaggio per i colossi industriali che oggi guidano l'intelligenza artificiale e che non staranno ad aspettare. E al contempo, l'incertezza regolamentare sta rallentando gli investimenti europei che attendono di essere impattati per primi dalle norme, in quanto operanti all'interno della giurisdizione eurocomunitaria.

Come scrive il Prof. G.M. Riccio “*Nell’era della self-responsability e delle autovalutazioni, dov’è la certezza del diritto? Nel GDPR, su 99 articoli e 373 Considerando, la parola assessment ricorre 47 volte: sembra un numero enorme, ma poi ci si imbatte nella proposta di Regolamento sull’AI dove, su 85 articoli e 89 Considerando, ricorre invece 152 volte e nel Cyber Resilience Act, su 57 articoli e 71 Considerando, addirittura 212 volte! Quindi, ci troviamo al cospetto di un continuo delegare l’individuazione del contenuto degli obblighi ai destinatari degli stessi, spesso ricorrendo a norme extragiuridiche (standard, protocolli, ecc.)*”³⁸.

E’ riconosciuto da tutti che l’Unione Europea ha dimostrato una particolare capacità nel definire prima di altri, le regole per i servizi digitali e vanta una primazia anche sulle norme condivise per l’intelligenza artificiale. Solo per citare qualche esempio, è entrato in vigore nel 2016 il Regolamento Generale sulla Protezione dei Dati. Più recentemente il Digital Services Act, il Digital Markets Act, il Data Act. Ultimo, ma non per importanza, il Data Governance Act che stabilisce un quadro generale per la condivisione ed il riutilizzo dei dati. Ed è in procinto di approvazione il AI Act.

Tuttavia gli obblighi regolamentari europei che si applicano ai grandi ecosistemi digitali stanno portando ad una serie di compromessi tra l’obbligo di apertura delle piattaforme da un lato, e la sicurezza, la privacy, l’innovazione, l’integrità del servizio dall’altro. Come scrive

³⁸ Cfr. Giovanni Maria Riccio - “Regole tecniche e regole giuridiche nella disciplina dei dati personali e dell’AI” - <https://dirittodiinternet.it/wp-content/uploads/2024/01/Riccio-Rel.pdf>



correttamente CERRE³⁹ “Questi compromessi sono riconosciuti nel DMA e sono parte del relativo bilanciamento che dovrà essere effettuato nel processo di attuazione dalle autorità regolamentari quando adotteranno le loro misure di conformità, poi dalla Commissione e dai tribunali nazionali quando valuteranno tali misure e, infine, saranno giudicati dalla Corte di Giustizia dell'UE che produrrà l'interpretazione definitiva del DMA”.

LA TENACE OPPOSIZIONE DEI SOGGETTI REGOLATI NEL MERCATO

Uno dei motivi dell'inefficacia delle attuali norme sulla concorrenza nei mercati digitali è che non creano un vero effetto deterrente nei confronti degli autori delle violazioni.

Si fa fatica perfino a introdurre dei “remedies” efficaci che possano ripristinare le condizioni di concorrenza a fronte delle continue violazioni che sono state ormai tipizzate dal Digital Market Act⁴⁰.

Sofferamoci per un attimo sui concetti stessi di “gatekeeper”, “ecosistema” o di “effetti di rete” che di recente vengono utilizzati nel dibattito sulle politiche europee della concorrenza. Questi termini sono tutti di derivazione economica e stanno aiutando certamente alla definizione degli obblighi in capo alle grandi piattaforme digitali, che poi sono quelle che tipicamente pongono in essere comportamenti escludenti. Già adesso osserviamo che la regolazione del digitale sta incontrando forti resistenze da parte dei soggetti obbligati, soprattutto intorno al *quantum*⁴¹ ed al *qualis*⁴².

Fin dai primi momenti, infatti, è apparsa chiara la volontà dei primi “Gatekeeper” designati, di avviare percorsi giudiziari per resistere alle norme europee. Inevitabilmente è andata sviluppandosi anche una dottrina opposta⁴³, che in nome dell'innovazione senza-consenso argomenta su come divincolarsi dalla fastidiosa notifica di Bruxelles.

CORRETTIVI AL DIGITAL MARKET ACT: I SERVIZI CLOUD

Ma c'è di più. Il Cloud computing rappresenta la risorsa tecnologica che è la preconditione dell'offerta di servizi di intelligenza artificiale. Infatti tutto il settore dell'AI si sta rapidamente strutturando attorno alle grandi aziende digitali già esistenti in quanto fortemente presenti nei mercati adiacenti, primo tra tutti quello della fornitura di servizi cloud. **La capacità del Digital Market Act di prevenire la monopolizzazione dell'intelligenza artificiale è ostacolata proprio dal fatto che né il cloud computing né i modelli di AI rientrano nel perimetro del regolamento.** E questo è accaduto nonostante il cloud computing fosse stato chiaramente elencato come “servizio di piattaforma principale” ai sensi del DMA.

Tuttavia, finora nessun servizio cloud è stato incluso nel primo gruppo di designazioni fatte dalla Commissione Europea lo scorso settembre 2023. Di conseguenza manca un presidio forte della Commissione sulle occasioni che a monte, ovvero sul cloud, possono determinare

³⁹ Cfr. CERRE - “Implementing the DMA: Substantive and Procedural Principles” - <https://cerre.eu/wp-content/uploads/2024/01/CERRE-BOOK-IMPLEMENTINGDMA.pdf>

⁴⁰ Cfr. Anush Ganesh, “Effective remedies in digital market abuse of dominance cases” <https://cdn.sanity.io/files/hr4v9eo1/production/f8dca7c69e166ceff714ab3b65641e5dbc591510.pdf>

⁴¹ Cfr. EU seeks to investigate Apple over cutting off web apps <https://www.ft.com/content/d2f7328c-5851-4f16-8f8d-93f0098b6adc>

⁴² Cfr. “Meta and TikTok sue European Commission over Digital Services Act fee” <https://www.euractiv.com/section/platforms/news/meta-and-tiktok-sue-european-commission-over-digital-services-act-fee/>

⁴³ Cfr. UEA - “Rebutting gatekeeper presumption” - <https://cdn.sanity.io/files/hr4v9eo1/production/266b7521dfadee151ed98fc2865fb01abe2e0ade.pdf>



una chiusura del mercato a valle, ovvero quello dell'AI. Per forza di cose, nemmeno i modelli di fondazione AI sono rientrati nel perimetro di regole europee vigenti, proprio perché non furono nemmeno menzionati nel DMA, non fosse altro perché sfuggirono temporalmente alla consapevolezza stessa del Legislatore europeo all'epoca dei negoziati sul regolamento.

Come ha sottolineato il ceo di Leonardo, Roberto Cingolani⁴⁴, intervenendo in audizione alla Commissione Attività produttive della Camera nell'ambito di una indagine conoscitiva sull'intelligenza artificiale: *"Abbiamo una Commissione europea che ha fatto diversi lavori regolatori basati su un approccio giuridico, con il rischio di fare il codice della strada prima di avere l'automobile. Negli Stati Uniti c'è stato un approccio tecnologico forse troppo permissivo e se la sono cavata dicendo l'etica ce la dà la trasparenza"*. Tuttavia, al netto delle differenze, entrambe i sistemi giuridici - quello americano e quello europeo - si accomunano per le forti difficoltà applicative nel settore digitale.

Questo significa che il complesso sistema di regole pensato per i grandi operatori globali, ivi compreso quello di imminente approvazione cd. **AI ACT, non può risolvere a livello applicativo le esternalità riscontrate, senza passare per un percorso lungo, tortuoso e pieno di insidie**. Analizzeremo nel prosieguo dello studio le ricadute concrete sul mercato di quanto affermato finora.

⁴⁴ Indagine conoscitiva Intelligenza artificiale - Audizioni - Roberto Cingolani, amministratore delegato di Leonardo S.p.A
<https://webtv.camera.it/evento/24574>



4] GLI EFFETTI DEL DISALLINEAMENTO TRA REGOLE E I TEMPI DEI PROCEDIMENTI

Scrive su Euractiv Max von Thun - direttore Open Markets Institute⁴⁵ - *“Considerata la velocità e la portata con cui l'intelligenza artificiale si sta sviluppando e diffondendo, attendere l'applicazione della legge sull'intelligenza artificiale per agire non è un'opzione per l'Europa. Una manciata di giganti della tecnologia sta già sfruttando il proprio dominio esistente su infrastrutture informatiche, dati e talenti per prendere il controllo dell'ecosistema emergente dell'intelligenza artificiale, compresa la cooptazione di potenziali sfidanti attraverso partnership e investimenti esclusivi”.*

L'EFFICACIA DELLE NORME TRA PROCEDIMENTI E SANZIONI

Anche quando le azioni delle Autorità coinvolte (i.e. Privacy, Antitrust, nazionali, europee ecc) vengono attuate, raramente le regole attuali possono offrire un rimedio tempestivo per sanare le condotte illecite dei giganti della tecnologia. Questo perché ad eccezione delle grandi acquisizioni/concentrazioni che hanno un percorso di pre-notifica a Bruxelles, **l'applicazione delle norme europee finora è avvenuta ex-post, quindi in maniera successiva e reattiva rispetto ai fatti.** A tal proposito, i governi potrebbero dotare le autorità di regolamentazione antitrust di maggiori poteri di intervento, soprattutto per fermare le acquisizioni che possono ridurre significativamente la concorrenza anche in termini prospettici e di ricadute su altri mercati adiacenti. Le misure ex post nei mercati digitali sono spesso limitate a *remedies* (o affidate ad impegni vincolanti) e sovente arrivano troppo tardive rispetto all'evoluzione del mercato, perché nel frattempo un'importante proprietà intellettuale potrebbe venire trasferita all'acquirente anche prima che una fusione venga perfezionata o annullata.

Più semplicemente, quando un'azienda vuole procedere ad una acquisizione, oppure adotta una condotta illecita e la concorrenza sul mercato viene ridotta, l'Autorità nazionale di regolazione si attiva e apre un'indagine - d'ufficio o sulla base di una segnalazione - a cui fa seguito l'attività istruttoria, a valle della quale e se esistono prove sufficienti, contesta all'azienda il comportamento passato. Nei casi più complessi, questo procedimento richiede spesso anni di indagini, rilievi, contestazioni formali, processi decisionali, contenziosi giurisdizionali ed i relativi ricorsi. Tutto ciò può continuare all'infinito, fino a quando un tribunale di ultima istanza non ordini all'azienda la cessazione della condotta o l'annullamento dell'acquisizione. Finanche il procedimento sanzionatorio è appellabile e ripete il percorso del procedimento principale.

Dunque se negli anni l'autorità giudiziaria coinvolta a tutti i livelli, ha svolto un ruolo importante per correggere le distorsioni sul mercato oppure una regolamentazione non correttamente applicata a livello nazionale, è importante rilevare che progressivamente non si è riusciti a garantire in alcun modo il ripristino delle condizioni competitive, perché anche quando si raggiunge una sentenza di accoglimento, il provvedimento dovrà tornare all'Autorità che lo ha emesso per un riesame e sarà soggetto a sua volta ad una nuova consultazione pubblica che darà luogo ad una nuova delibera, che sarà anche essa

⁴⁵ Cfr. Euractiv “EU does not need to wait for the AI Act to act” <https://www.euractiv.com/section/artificial-intelligence/opinion/eu-does-not-need-to-wait-for-the-ai-act-to-act>



impugnabile. **Il fattore tempo è dunque un elemento di rischio associato al rischio regolatorio quando si somma al rischio giudiziario che è anche esso un rischio economico per tutto il comparto.**

IL WEBSCRAPING PER ALLENARE SISTEMI DI AI: IL CASO CHATGPT

La Commissione Europea, ad esempio, per regolare l'intelligenza artificiale ha proposto un approccio basato sul rischio, con quattro livelli di gradualità per i sistemi di AI, nonché un'identificazione di rischi specifici per i modelli di uso generale. Tra essi, viene considerato "rischio inaccettabile" - tra gli altri - lo *"Scraping non mirato di Internet o CCTV per immagini facciali per creare o espandere database"*⁴⁶.

Non stupirà che proprio il cd. *"webscraping"* - una pratica in uso da anni - è venuto immediatamente in evidenza oggi, non solo in Italia⁴⁷ ma in maniera precipua per l'intelligenza artificiale, a causa dell'utilizzo di questa tecnica anche per addestrare sistemi di AI. E' chiaro che la raccolta a strascico di dati personali pone diversi ordini di criticità perché in assenza di idonea base giuridica, ogni trattamento è da considerarsi per sua natura illecito. Queste considerazioni riportano a diversi tipologie di adempimenti previsti dal GDPR⁴⁸. Tutte le volte che la protezione dei dati personali si scontra con la AI generativa: 1) occorrerà avere una base giuridica adeguata per la formazione; 2) una base giuridica adeguata per il trattamento delle richieste; 3) ottemperare i requisiti di informazione; 4) attuare un'inversione del modello, proteggere la fuga di dati e garantire il diritto alla cancellazione; 5) garantire la tutela dei minori; 6) e infine procedere alla limitazione dello scopo. **Si tenga conto inoltre che l'intelligenza artificiale non viene alimentata solo di dati personali ma anche di dati-non-personali compresi quelli aziendali che pongono un diverso ordine di problemi, per lo più economici e competitivi, ma non per questo di minore rilevanza.** E che in ultima istanza, i dati-non-personali potrebbero ricondurre attraverso un trattamento di intelligenza artificiale - a dati personali.

L'Autorità italiana Garante per la protezione dei dati personali ha recentemente emesso una notifica a OpenAI⁴⁹, affermando che ChatGPT, il *chatbot* AI più conosciuto al mondo, avrebbe infranto alcune norme sulla protezione dei dati dell'Unione Europea. A valle di indagini avviate lo scorso aprile 2023⁵⁰, l'Autorità italiana aveva già imposto un divieto temporaneo a ChatGPT a causa di molteplici infrazioni, tra cui la raccolta illegale di dati personali e l'assenza di meccanismi per verificare l'età dei minori. Soprattutto il Garante italiano ha sottolineato che OpenAI non avrebbe informato adeguatamente gli utenti della raccolta dei loro dati personali, violando così aspetti cruciali del Regolamento generale sulla protezione dei dati (GDPR). **A distanza di un anno, sappiamo che il procedimento di contestazione sta proseguendo il suo iter, ma non abbiamo evidenze certe che i trattamenti contestati a ChatGPT siano effettivamente cessati.** *"Il problema da risolvere - scrive il Prof. Andrea Monti⁵¹ - è quindi se lo scraping e l'ulteriore trattamento dei dati da*

⁴⁶ Cfr. European Commission, AI Q&A - https://ec.europa.eu/commission/presscorner/detail/en/QANDA_21_16839

⁴⁷ Generative AI first call for evidence: The lawful basis for web scraping to train generative AI models <https://ico.org.uk/about-the-ico/what-we-do/our-work-on-artificial-intelligence/generative-ai-first-call-for-evidence/>

⁴⁸ Cfr. per una lista di controllo sulla protezione dei dati per l'IA è utile il documento BayLDA https://www.lda.bayern.de/media/ki_checkliste.pdf

⁴⁹ Cfr. ChatGPT: Garante privacy, notificato a OpenAI l'atto di contestazione per le violazioni alla normativa privacy <https://www.garanteprivacy.it/home/docweb/-/docweb-display/docweb/9978020>

⁵⁰ ChatGPT: Garante privacy, limitazione provvisoria sospesa se OpenAI adotterà le misure richieste. L'Autorità ha dato tempo alla società fino al 30 aprile per mettersi in regola <https://www.garanteprivacy.it/web/guest/home/docweb/-/docweb-display/docweb/9874751>

⁵¹ GDPR e OpenAI: quanto sono fondate le accuse del Garante? https://www.repubblica.it/tecnologia/blog/strategikon/2024/01/31/news/gdpr_e_openai_atto_secondo-422019996/



parte di OpenAI ledano, almeno in termini potenziali, i diritti fondamentali dei cittadini degli Stati membri dell'UE e, cosa più importante, quelli garantiti dalla nostra Costituzione”.

E' chiaro quindi che i sistemi AI più conosciuti al mondo, non hanno mostrato finora particolare attenzione all'acquisizione del consenso, alla revoca del consenso, o finanche alla negazione del consenso (le sanzioni ed i numerosi richiami delle Autorità ne sono prova). Il tema del consenso appare quindi sostanzialmente ignorato in certi contesti o addirittura viene trattato come un succedaneo dell'innovazione, come sopra ampiamente argomentato.

Una vicenda, questa, che ha nuovamente interessato il nostro Garante per la Protezione dei Dati Personali che sul tema del webscraping da parte di sistemi di intelligenza artificiale, ha avviato un'indagine conoscitiva⁵². Nell'attesa degli esiti, le esternalità negative dirette sulla società causate da un'innovazione-senza-consenso (come quella che raccoglie i dati personali a strascico pescando in rete immagini e fotografie per allenare sistemi di intelligenza artificiale) sono evidenti già adesso, ed ampiamente indagate da tutte le Autorità europee che faticano a inseguire questi fenomeni non solo per via della loro opacità⁵³ ma anche della loro complessità. **In una recente intervista il Presidente dell'Autorità Garante della Privacy Prof. Pasquale Stanzone ha dichiarato⁵⁴ che “sacrificare diritti e libertà sull'altare del mercato è incompatibile con i nostri principi costituzionali. L'articolo 41 della Costituzione stabilisce che l'iniziativa economica privata non può svolgersi in modo da recare danno alla libertà e alla dignità umane. Un mercato basato su forme di oppressione non sarebbe sostenibile, non funzionerebbe. Non possiamo accettare queste forme di capitalismo estrattivo”.**

La domanda che alcuni studiosi iniziano a porsi è se il sistema basato sul “consenso” sia ancora attuale nello scenario alimentato dall'intelligenza artificiale. Secondo l'Università di Stanford⁵⁵ *“La sfida che avanza in un mondo con una maggiore domanda di dati è come mitigare la raccolta di dati in eccesso senza aggiungere troppi attriti con richieste di consenso eccessive. I servizi digitali hanno bisogno dei dati dei consumatori per funzionare e non tutte queste richieste sono eccessive”.* **L'attuale impianto regolamentare sta creando nuove disuguaglianze industriali tra chi è veramente obbligato perché è un'impresa locale dimensionalmente e giuridicamente più raggiungibile, e chi sfugge perché si trova ad operare in un'altra giurisdizione o ha una capacità di resistere alle sanzioni superiore all'efficacia delle autorità che controllano il mercato.**

Peraltro, per trattare ulteriormente i dati personali per una finalità diversa e non-incompatibile con quella per cui essi sono stati raccolti - il regolamento sulla privacy prevede un rigido quadro di garanzie, tra cui quella di fornire adeguate informazioni in merito a tale diversa finalità e ogni ulteriore informazione pertinente. Cosa che spesso, di fatto, non avviene.

⁵² Intelligenza artificiale: Garante privacy apre un'indagine sulla raccolta di dati personali on line per addestrare gli algoritmi. L'iniziativa è volta a verificare l'adozione di misure di sicurezza da parte di siti pubblici e privati <https://www.garanteprivacy.it/home/docweb/-/docweb-display/docweb/9952078>

⁵³ The Foundation Model Transparency Index (Stanford) <https://crfm.stanford.edu/fmti/>

⁵⁴ ChatGpt, il Garante della privacy: "No al mercato senza freni. Noi terza via tra Usa e Cina" - <https://www.garanteprivacy.it/home/docweb/-/docweb-display/docweb/9879276>

⁵⁵ Cfr. Jennifer King, Caroline Meinhardt - "Rethinking Privacy in the AI Era: Policy Provocations for a Data-Centric World" <https://hai.stanford.edu/white-paper-rethinking-privacy-ai-era-policy-provocations-data-centric-world#:~:text=February%2022%2C%202024-.Rethinking%20Privacy%20in%20the%20AI%20Era%3A%20Policy.for%20a%20Data%2DCentric%20World&text=This%20white%20paper%20explores%20the.harms%20in%20an%20AI%20era.>



5] COMBATTERE LE DISEGUAGLIANZE INDUSTRIALI E RISOLVERE LE ESTERNALITÀ SULLA SOCIETÀ

A questo punto della trattazione, ci appare necessario dare sostanza al tema delle esternalità negative dell'intelligenza artificiale, specialmente quelle che si abbattano sulla società civile. Questo perché i grandi operatori globali hanno di fatto ignorato o posto in seria discussione "i diritti sui dati" prima ancora che i dati stessi, fagocitandoli e rielaborandoli fin dal loro ingresso. Essi cioè vengono "masticati" e trasformati per renderli edibili ai sistemi di intelligenza artificiale.

Questi prodotti e servizi si alimentano continuamente di dati che rappresentano la vera contropartita di valore rispetto alla percezione che comunemente si attribuisce all'innovazione tecnologica. **Perfino le pratiche di *scraping* più invasive, utilizzate per alimentare sistemi AI, vengono sovente legittimate dalla tesi ormai "consumata" per cui l'innovazione non si può fermare perché migliora la qualità della vita dei cittadini e dei consumatori.** Non è credibile, infatti, che gli enormi set di dati accumulati dai colossi globali oggi leader della AI siano stati rispettosi nel raccogliere il consenso individuale.

E' chiaro quindi che si deve fare di più per efficientare questo percorso o quantomeno provare ad identificare un interesse valido e lecito ma che sia totalmente diverso da quello puramente commerciale. Soprattutto occorrerà dimostrare che i rischi per gli individui siano sempre tenuti in considerazione e significativamente mitigati "*by design*", cosa che ad oggi non sembra che stia accadendo.

C'è bisogno che anche le Autorità di settore, sviluppino nuove competenze informatiche che spaziano dalla conoscenza di tecniche di *data mining*, elaborazione del linguaggio naturale, analisi del testo, *web scraping*, visualizzazione dei dati e apprendimento automatico, nonché competenze nelle principali aree tematiche giuridiche, economiche e statistiche che si applicano all'intelligenza artificiale.

ETICA E DIRITTO: L'INSEGNAMENTO DEL PROF. STEFANO RODOTÀ

Un nuovo punto di vista, finora non sufficientemente esplorato, ruota intorno al tema delle disuguaglianze industriali acuite dall'avvento dell'AI. Scrive il Prof. Gianfranco Viesti nel suo libro "Riuscirà il Pnrr a rilanciare l'Italia?"⁵⁶ - "*Un paese non si rilancia con una lista di riforme e di investimenti scritta da tecnici, anche se di grande qualità; ma solo attraverso una visione politica del suo futuro*". Il Pnrr, secondo Viesti, può rappresentare un'utilissima inversione di rotta rispetto alle politiche precedenti, ma da solo difficilmente potrà cambiare l'Italia, se non affronta i principali nodi politici alla radice delle sue difficoltà e delle sue disuguaglianze. Ciò si collega anche a questioni di equità e di concorrenza dinamica al fine della valutazione degli impatti di lungo periodo che i sistemi di intelligenza artificiale avranno sulla società.

Torna di estrema attualità il discorso al Parlamento del **Prof. Stefano Rodotà che già nel 1999 si domandava "se tutto ciò che è tecnicamente possibile è anche eticamente ammissibile, socialmente accettabile, giuridicamente lecito"**. Queste parole

⁵⁶ Cfr. Gianfranco Viesti "Riuscirà il Pnrr a rilanciare l'Italia?" - 2023 Donzelli Editore



dell'indimenticato accademico ci interrogano ancora oggi sul futuro che stiamo costruendo e meritano doverosamente una risposta da parte dei *policy maker*.

Combattere le disuguaglianze industriali ci appare il metodo migliore per correggere, a valle, le disuguaglianze sociali generate dall'intelligenza artificiale, che stanno avendo il maggiore impatto sul mondo moderno. Invertendo la prospettiva si scopre la debolezza delle basi di un ragionamento che – almeno finora - davamo per scontate. Il perseguimento dell'efficienza da parte delle imprese è sicuramente un requisito che può portare una maggiore concorrenza a vantaggio dei consumatori. Ma **l'obbiettivo adesso non è semplicemente promuovere la concorrenza tra sistemi di AI, quanto piuttosto permettere a nuovi soggetti – auspicabilmente italiani - di entrare in questo mercato e salvaguardare i diritti dei consumatori, dei cittadini, delle aziende, a tutti questi livelli.**

Molti problemi affondano le loro radici nel processo di raccolta dei dati: questioni come il consenso, la privacy, l'equità e la proprietà intellettuale, la cybersicurezza, stanno sollevando interrogativi⁵⁷ di non poco momento, portando a nuove disuguaglianze che originano da un'innovazione senza-consenso, ovvero quella che stiamo subendo e che facciamo fatica a comprendere. A ben vedere non ci troviamo di fronte ad una novità assoluta e nemmeno di fronte ad una pratica in uso solo per i sistemi AI, ma anzi stiamo analizzando la modalità principale con cui si sono imposti negli ultimi 20 anni, i sistemi di *search*, *e-commerce*, geolocalizzazione, riconoscimento facciale, marketing e pubblicità online e tanto altro che oggi domina il mercato e sta decretando lo straordinario successo delle piattaforme di AI finora annunciate.

CESSIONE DEI DATI PERSONALI: IL RUOLO DEI PRIVATI

Tuttavia c'è un altro aspetto di questa vicenda che resta piuttosto sottovalutato ed è quello che passa attraverso la partecipazione attiva dei cittadini e delle parti interessate nella cessione dei dati personali. **Sono infatti i privati cittadini quelli che sembrano cedere più volentieri - e in modo spesso inconsapevole - i propri dati a sistemi esteri di AI ma che al contempo mostrano resistenze a cedere gli stessi dati alla Pubblica Amministrazione o ad altri sistemi locali.**

Scrivono Roberto Viola e Luca De Biase nel libro Codice del Futuro⁵⁸: *“La popolazione tende a fidarsi delle grandi aziende di internet, regalando loro una quantità di informazioni personali tale da costruire profili che descrivono nei minimi dettagli gli individui. Tende a fidarsi anche della Cina e del suo Stato ipervigilante, concedendo a grandi organizzazioni cinesi una quantità di dati gigantesca. Eppure, molti pensano di non volersi fidare degli Stati europei e dell'Europa stessa, istituzioni che costituzionalmente sono molto più orientate a garantire i diritti delle organizzazioni cui si è accennato sopra”.*

Giova a questo punto ricordare che durante la pandemia due operatori oggi leader nell'intelligenza artificiale, offrivano pubblicamente i dati di tracciamento del contagio più aggiornati e ben prima rispetto a quelli ufficiali distribuiti dalle Autorità. Il rischio – anche a

⁵⁷ Sul punto si veda “Generative AI in EU Law: Liability, Privacy, Intellectual Property, and Cybersecurity” - Claudio Novelli - Federico Casolari - Philipp Hacker - Giorgio Spedicato - Luciano Floridi <https://media.licdn.com/dms/document/media/D4D1FAQHtySsprVhUtQ/feedshare-document-pdf-analyzed/0/1708441757476?e=1709164800&v=beta&t=Imd-W9uLnt4DQY75w7Jn1Eu2pcScjVo7IPf1N7VB6nE>

⁵⁸ Cfr. Roberto Viola, Luca De Biase Il codice del futuro - <https://ecommerce.ilsolo24ore.com/shopping24/il-codice-del-futuro.html>



fronte di un'utenza che sembra non avere ancora compiutamente acquisito il valore dei propri dati personali – è che si cedano dati personali più volentieri ed in maniera più agevole e concedente agli operatori globali, ed in maniera più complessa ai soggetti nazionali che – essendo più raggiungibili e meno efficienti perché non godono di economie di scala e di scopo – generalmente sono gravati da obblighi che li portano ad un rispetto letterale delle norme e quindi l'adesione ai loro servizi è più capziosa e complessa rispetto a un “click to join”, “log in with”, “tap here”, “consent” su più servizi di una piattaforma che fa parte di un ecosistema che cattura l'utente. L'integrazione dei modelli di base di AI nelle piattaforme globali oggi dominanti - ed in alcuni casi la loro trasformazione in ecosistemi proprietari di servizi e prodotti AI⁵⁹ - rappresenta la sfida competitiva principale che dovremo affrontare.

LA CENTRALITÀ DELL'UOMO NEL DIBATTITO SU AI

Nel preambolo della Dichiarazione europea sui diritti e i principi digitali per il decennio digitale, si legge «*Il nostro obiettivo è di promuovere un modello europeo per la trasformazione digitale, che metta al centro le persone, sia basato sui valori europei e sui diritti fondamentali dell'UE, riaffermi i diritti umani universali e apporti benefici a tutte le persone, alle imprese e alla società nel suo complesso*».

La protezione dei dati è una conseguenza della protezione di diritti individuali e come correttamente ha osservato **il Prof. Andrea Monti stiamo assistendo ad una perdita di questi diritti: “Nel momento in cui rinunciamo ai nostri diritti e li consideriamo negoziabili è questo il vero oro che abbiamo messo in mano alle Bigtech e non sono i dati⁶⁰”**. (...) *La profilazione individualizzata a livello nominativo non serve più perché quello che conta è avere una quantità di dati, tale che si possano addestrare i sistemi di AI in modo che essa possa operare senza dovere ogni volta avere accesso ad una base di dati⁶¹”*.

Il dato singolo *ex se* – ivi incluso il dato personale - mostra oggi di avere una sua singolare “irrilevanza” se non quando trattato e inserito in un contesto trasformativo che ne fa nascere il valore. E' così che sono nati anche i cd. “Dati sintetici o artificiali”, che pongono certamente un altro ordine di problemi ma che pur sempre rappresentano – nell'ambito di questa analisi - la forza irresistibile e *disruptive* di cui solo pochi operatori globali sono capaci.

Tuttavia, ancora oggi poche persone si rendono conto di quali dati personali vengono ceduti aderendo ad un servizio online ed eventualmente a quale uso hanno prestato il consenso⁶². Né sanno valutare se quel consenso è stato acquisito correttamente od estorto senza una vera alternativa. Questo perché l'approccio di notifica e consenso per la privacy impone un onere significativo a carico degli individui non solo nell'esercizio di questi diritti, ma anche nel comprendere il motivo per cui ciò si dovrebbe fare.

Ad esempio, ancora oggi molti utenti di Internet faticano a capire cosa siano i *cookie* e come la loro raccolta possa compromettere la loro privacy. E tutto questo sta già accadendo nel modo più opaco ai danni di cittadini, aziende e istituzioni.

Si nota infatti che non esistono requisiti di trasparenza specifici affinché le aziende riportino con esattezza la provenienza dei propri dati di formazione con cui allenano sistemi di AI, né

⁵⁹ Es. Apple vision pro nell'ecosistema Apple, Meta Quest e metaverso Facebook, Starlink satelliti guidati da AI in orbita bassa ecc.

⁶⁰ Cfr. Protecting Personal Information - A.Monti, R. Wacks, Hart Publishing 2019 - video minuto 1:02:03 <https://youtu.be/oKgOB86KFpY>

⁶¹ Cfr. Protecting Personal Information - A.Monti, R. Wacks, Hart Publishing 2019 - video minuto 54:48 <https://youtu.be/oKgOB86KFpY?t=3286>

⁶² How Americans View Data Privacy <https://www.pewresearch.org/internet/2023/10/18/how-americans-view-data-privacy/>



possono documentare fattualmente se essi siano conformi ai principi di minimizzazione dei dati e di specificazione dello scopo per cui essi sono stati raccolti o per affrontare problemi legati alla qualità dei dati e le relative conseguenze in termini di "allucinazioni". Temi anche questi estremamente delicati di cui non potremo occuparci in questa trattazione ma che meritano un grande rilievo.

MODELLI OPACHI: TECNICHE DI ACQUISIZIONE DEL CONSENSO

Secondo il Prof. Ignacio Cofone autore del libro *The Privacy Fallacy*⁶³ **"Ignoriamo le impostazioni sulla privacy e i termini e le condizioni non perché siamo pigri, ma piuttosto perché i progetti dei fornitori di servizi ci spingono verso scelte svantaggiose. (...) I progettisti modellano abitualmente architetture di scelta per indurre le persone ad accettare termini che avvantaggiano le aziende. (...) Inquadrare le impostazioni relative alla scelta della privacy in modi convenienti, ad esempio etichettandole come "impostazioni dell'app" anziché "impostazioni sulla privacy", aumenta la probabilità che le persone accettino le pratiche relative ai dati. Inoltre, i design non user-friendly ostacolano i controlli sulla privacy percepiti"**.

Ma la questione a destare più allarme è che tutto il processo algoritmico resta opaco anche a chi è chiamato a controllare e governare questi fenomeni distorsivi che poggiano su tre asset critici e difficilmente replicabili quali: A) sistemi non trasparenti, B) algoritmi che cambiano nel tempo, C) enormi *dataset* privati spontaneamente conferiti e che attraverso un "processo di privatizzazione" diventano non accessibili e quindi possono arrivare a mettere a rischio la tenuta delle democrazie. E' chiaro che se affidiamo agli algoritmi delle decisioni importanti, dovremmo anche sapere in maniera trasparente come funzionano. Ma tutto questo non sta accadendo.

Questo perchè il sapere algoritmico non ha contezza di se e del tempo, mentre l'uomo si⁶⁴. Anche il processo di *machine learning* è un reale consolidato⁶⁵ in miliardi di ipotesi simili, tutte possibili e diverse. Ma l'intelligenza artificiale generativa - con il suo carico di predizioni fondate su meccanismi neurali ispirati alla struttura del cervello umano - suggerisce un'ipotesi di senso⁶⁶ ulteriore rispetto ad un senso già trovato - che sappiamo sfuggire persino ai soggetti che oggi stanno sviluppando sistemi di intelligenza artificiale. Scrive a tal proposito il prof. Stefano Epifani⁶⁷: *"Impattando sui motori di scelta, la trasformazione digitale, dunque, non si limita ad agire a livello di processo, ma agisce letteralmente a livello di senso: ha il potere, infatti, di cambiare il senso delle cose. Una vera e propria rivoluzione di senso che basandosi sulla mutata percezione del concetto di valore (applicato di volta in volta ai diversi aspetti: valore delle relazioni interpersonali, valore delle istituzioni, valore dei beni e dei servizi con i quali entriamo in contatto, valore dell'informazione, finanche valore delle persone) produce un cambiamento profondo in ogni aspetto della vita"*.

Tutto ciò meriterebbe un approfondimento specifico ed un dibattito integrato tra filosofia, psicologia ed etica - che non possiamo fare in questa sede - ma che ci torna utile come elemento di analisi delle disuguaglianze industriali e delle conseguenze negative sulla

⁶³ Cfr. Ignacio Cofone - *The Privacy Fallacy - Harm and Power in the Information Economy*
<https://www.cambridge.org/core/books/privacy-fallacy/547578F2A1AE0C40963105CE066B412E>

⁶⁴ Cfr. Martin Heidegger, "Essere e Tempo"

⁶⁵ Cfr. J. Lacan "1-Seminario, Reale, simbolico, immaginario"

⁶⁶ Secondo alcuni studiosi, la tecnologia AI potrebbe non essere altro che un "pappagallo stocastico" che imita il linguaggio umano senza alcun collegamento con il significato. Sul punto si veda "On the Dangers of Stochastic Parrots: Can Language Models Be Too Big?" <https://dl.acm.org/doi/10.1145/3442188.3445922>

⁶⁷ Cfr. Stefano Epifani "Sostenibilità Digitale: Perché la sostenibilità non può fare a meno della trasformazione digitale"



società. E soprattutto per rimettere al centro l'uomo nel dibattito sull'intelligenza artificiale. Come recentemente hanno scritto Henry A. Kissinger, Eric Schmidt, Daniel Huttenlocher e Aldo Piccato⁶⁸ *“Nessun esperto, quale che sia il suo settore di studi, è in grado di comprendere da solo un futuro in cui le macchine imparano e utilizzano una logica che oltrepassa l'attuale portata della ragione umana. Le società, quindi, devono collaborare non soltanto per capire, ma anche per adattarsi”*.

VERSO UN APPROCCIO MULTIDISCIPLINARE

Se i ricercatori e le Autorità preposte avessero accesso a una serie di dati che rivelano il funzionamento interno dei sistemi di AI, potrebbero quantificarne gli effetti affinché possano porvi rimedio. Questioni come la diffusione della disinformazione online, dell'estremismo, la salute mentale dei minori, o i fenomeni di polarizzazione ideologica, verrebbero comprese nei profondi meccanismi che le determinano a livello algoritmico. Ciò si tradurrebbe in controlli e probabili sanzioni ed è per questo che gli operatori globali che guidano l'AI oppongono una significativa resistenza limitandosi ad impegni, da parte delle aziende stesse, a sviluppare nuovi sistemi di auto-controllo. Ma non può essere sufficiente.

Come recentemente ha affermato il Presidente del Comitato editoria e AI, Prof. Padre Paolo Benanti⁶⁹ *“Il tema è sfidante, non si risolve con una sola prospettiva, bisogna vivere la complessità e creare un sistema che possa garantire benessere alla maggior parte delle persone”*. Per far fronte a tutto questo, non basta quindi la prospettiva del giurista, dell'economista, del filosofo e dell'eticista. Occorre un concorso sinergico di più discipline per indagare meglio le ragioni di questa rivoluzione tecnologica, partendo dal rapporto tra uomo ed intelligenza artificiale per promuovere sistemi affidabili ed etici lungo l'intero percorso del loro sviluppo, della loro diffusione e del loro utilizzo. E soprattutto, è importante che essi siano in linea con i valori dell'UE.

L'algoritmo ad esempio, può essere etico o non etico, ma resta agnostico anche quando è addestrato ad una specifica finalità. E sicuramente è opaco anche a chi dovrebbe esercitare un controllo su di esso. I *dataset* di addestramento cambiano anche essi continuamente, determinando un *output* diverso anche a seconda del modello che li addestra e che presto farà a meno di attingere ogni volta dallo stesso *repository* di dati, perché dopo che li ha “digeriti” e li ha inglobati, si trova nel bisogno di avere ulteriori dati, nuovi e diversi, per un ulteriore confronto ed un nuovo sviluppo.

RESPONSABILITÀ DEGLI ALGORITMI: RICONOSCIMENTO E DIRITTO

L'anticipazione del futuro da parte di un'intelligenza artificiale ha cambiato gran parte delle nostre convinzioni anche in materia di responsabilità. Siamo inevitabilmente travolti da anglicismi anche nel linguaggio giuridico. *“Intelligenza artificiale - Cos'è davvero, come funziona, che effetti avrà”⁷⁰* è un libro adatto per cogliere le principali differenze tra 1) *accountability* (morale) 2) *liability* (giuridica) 3) *responsability* (causale). Per scriverlo, un noto

⁶⁸ Cfr. *“L'era dell'intelligenza artificiale Il futuro dell'identità umana”* - Henry A. Kissinger, Eric Schmidt, Daniel Huttenlocher, Aldo Piccato

⁶⁹ Cfr. La Repubblica - Paolo Benanti, presidente della Commissione di Palazzo Chigi: “Regole per rendere subito riconoscibile un prodotto dell'Intelligenza artificiale” https://firenze.repubblica.it/cronaca/2024/01/24/news/padre_benanti_intelligenza_artificiale-421967048/?ref=RHLF-BG-P7-S3-T1

⁷⁰ Cfr. Stefano Quintarelli, Claudia Giulia Ferrauto, Francesco Corea, Fabio Fossa, Andrea Loreggia, Salvatore Sapienza, Claudia Giulia Ferrauto - *“Intelligenza artificiale Cos'è davvero, come funziona, che effetti avrà”*



informatico, imprenditore e tecnologo italiano, Stefano Quintarelli ha riunito le competenze di sei diversi esperti del settore partecipando anche al primo "AI High Level Expert Group" della Commissione Europea con il fine di valutare gli impatti di questa nuova tecnologia sulle nostre vite.

In tutto questo, non dobbiamo cadere nell'inganno terminologico a cui può indurci l'uso della parola "intelligenza" come forma di "antropomorfizzazione" dell'AI⁷¹. Sappiamo ad esempio che in un rapporto bilaterale uomo-macchina, la capacità di agire dell'uomo non è mai stata riconosciuta dalla macchina che fino a ieri si è limitata ad accettare passivamente ed eseguire il comando dell'uomo, in maniera agnostica e formalmente senza riconoscerlo.

Eppure è la stessa macchina che già oggi può facilmente identificarci con i dati biometrici e differenziare così l'output prodotto. Un bel passo avanti, avvenuto peraltro in pochissimo tempo e solo marginalmente arginato dalla regolamentazione. E' così che l'enorme mole di dati personali catturati negli anni dai fornitori di servizi online, soprattutto quelli che si sono imposti con il loro sapere tecnologico a livello globale, ha capovolto la convinzione che le macchine siano dei semplici automi governati da un algoritmo. Non è più così.

Senza autodeterminazione la scelta operata dall'intelligenza artificiale porta semplicemente all'assunzione di una responsabilità in capo ad un altro soggetto che peraltro resta ancora da identificare in termini tecnici e giuridici ma che probabilmente è il produttore o chi ne trae profitto⁷². E su questo tema, quello della responsabilità, si è aperto un ampio dibattito pubblico oltre che un inevitabile ed asprissimo confronto giudiziario. Una parte della dottrina propone una pluralità di soggetti giuridici individuabili come responsabili all'interno in un perimetro chiuso di illeciti informatici, peraltro quasi tutti già noti e codificati. Un'altra parte della dottrina, invece, propone un nuovo modello di responsabilità mediata che passa attraverso la creazione di un fondo di garanzia a ciò destinato, anche per far fronte all'impossibilità che abbiamo oggi di individuare un modello assicurativo che si attagli alle fattispecie in esame. Del resto, quale impresa assicurativa potrebbe farsi carico di un rischio non definito?

Se è improbabile un controllo umano diretto sull'intelligenza artificiale e se si vuole risolvere la questione della responsabilità e degli impatti sociali e geopolitici che genera, si dovrebbe agire rompendo le disuguaglianze industriali e ricreando un *"level playing field"* tra i diversi attori in gioco, partendo proprio dalla rottura dell'oligopolio dei dati in capo a pochi *stakeholder*.

⁷¹ Cfr. - Giusella Finocchiaro "Intelligenza artificiale: quali regole?" - il Mulino 2024

⁷² Sul tema si veda il caso canadese: *Moffatt v. Air Canada*, 2024 Bccr 149, del 14 febbraio 2024, secondo l'analisi della Prof.ssa Giusella Finocchiaro si potrebbe riassumere limitandosi a citare il noto brocardo latino: «*cuius commoda, eius et incommoda*» (chi trae vantaggi da una situazione, deve sopportarne anche gli svantaggi).



6] L'ACCESSO AI BIG DATA COME PRESUPPOSTO DELLA COMPETIZIONE E DEMOCRATIZZAZIONE DELL'AI

La barriera all'ingresso principale che ostacola la crescita del settore AI nel nostro Paese è rappresentata principalmente dall'indisponibilità di grandi *dataset di training*. Questi rappresentano delle "barriere artificiali" che derivano, in parte, dalla predazione deliberata dei dati avvenuta negli ultimi 20 anni ed operata a livello globale dai grandi fornitori di servizi online. Gli attori che già possiedono grandi set di dati hanno un vantaggio significativo rispetto agli altri sviluppatori di sistemi AI che non dispongono di riserve di dati o di capacità di calcolo, e devono raccogliergli in modo conforme alle norme, acquistarli o concederli in licenza.

Ciò comporta un forte ostacolo alla nascita di nuovi sistemi di intelligenza artificiale in Italia, e crea numerosi intralci applicativi alla regolamentazione sull'uso dei dati (personali, non personali, aziendali, critici, strategici, ordinari ecc) al netto del fatto che un nuovo quadro regolamentare arriverà tardivo ed in maniera non uniforme tra Europa e Stati Uniti⁷³.

ECONOMIE DI SCALA E DI SCOPO - LE ESTERNALITÀ NEGATIVE

Come è stato correttamente osservato dalle **Professoressa Bonassi e Iossa**⁷⁴, **“La disponibilità di dati in capo ai grandi operatori digitali sottende un trade-off in termini di effetti sui consumatori. Da un lato, si può assistere ad un miglioramento dell'efficienza nell'offerta di prodotti/servizi che a sua volta può riverberarsi positivamente sulla loro qualità a beneficio dei consumatori. Dall'altro lato, si rafforza il loro vantaggio competitivo, nella misura in cui possono sfruttare i dati cui hanno accesso per adottare comportamenti anticoncorrenziali idonei a produrre un impatto negativo sul mercato e sui consumatori”**.

I dati sono preziosi perché permettono ad esempio di profilare le abitudini di milioni di consumatori e quindi interessano a diversi settori economici, come ad esempio i mercati della pubblicità e del commercio online, ma non solo. E' chiaro che se tutti avessero grandi quantità di dati ed i mezzi per analizzarli, le aziende oggi leader dell'intelligenza artificiale perderebbero la loro supremazia che non è data soltanto un "sapere tecnologico" che potrebbe essere sviluppato in molti altri Paesi (tra cui l'Italia), quanto piuttosto dal monopolio verticale sui *Big Data* privatizzati che sono stati accumulati negli anni.

A tal proposito, ricordiamo che la Dichiarazione europea sui diritti e i principi digitali per il decennio digitale recita: «*Ogni persona dovrebbe avere la possibilità di competere lealmente e innovare nell'ambiente digitale. Tutto ciò dovrebbe apportare benefici anche alle imprese, comprese le PMI*». Dunque non si discute adesso di tutti gli altri ostacoli all'ingresso sul mercato dell'intelligenza artificiale, che possono derivare ad esempio dagli effetti rete⁷⁵ o dalle economie di scala⁷⁶ e di scopo⁷⁷ già rafforzate a livello globale dagli operatori

⁷³ Secondo alcuni studiosi è difficile competere con soggetti che dispongono già di dati liberi e non regolamentati, in particolare quando esiste concorrenza tra giurisdizioni legali.

⁷⁴ Cfr. C. Bonassi, E. Iossa - "Il ruolo dei dati nella tutela della concorrenza e dei consumatori: l'esperienza dell'AGCM" - <https://www.linkedin.com/feed/update/urn:li:activity:7150896077901049858/>

⁷⁵ Si ha tipicamente un effetto rete nell'intelligenza artificiale nel momento in cui i dati generati dalle interazioni con il modello generativo, vengono raccolti ed utilizzati per il miglioramento del modello stesso, diventando a loro volta dati di addestramento.

⁷⁶ Le economie di scala nell'AI sono certificate dalla rapida crescita del mercato dei modelli di fondazione e la loro espansione su scala globale in un tempo ristretto e la loro concentrazione geografica e di mercato.



dominanti. Anche perché in quel caso ci troveremmo di fronte a barriere potenzialmente superabili con la regolamentazione attuale, in quanto riflettono forme di diversa efficienza degli operatori sui differenti mercati in cui operano.

IL MERCATO DEI DATI: LE ASIMMETRIE E LE CONCENTRAZIONI

Ma il tema centrale è che le aziende globali leader dell'AI non intendono condividere gli enormi set di dati che hanno accumulato, proprio perché quello è il loro *asset* principale ed è anche il più redditizio, e quindi da proteggere. Come ha osservato la commissaria europea alla concorrenza Margrethe Vestager⁷⁸, *“Più dati puoi raccogliere, più sai, migliore è il prodotto che poi fornisci, ma anche più potente sarai nei confronti degli altri”*. Le imprese italiane interessate a sviluppare sistemi di intelligenza artificiale in Italia possono crescere allo stesso modo in assenza della disponibilità di Big Data pubblici?

Come recentemente ha scritto il **Prof. Viktor Mayer-Schönberger**⁷⁹: ***“It’s simply foolish to let a few data-rich corporations limit the value and insights society can gain from data.” (...)*** ***“Data monopolies are theft of progress. Data usage is service to the common good”***.

In questo momento, solo le *BigTech* dispongono di enormi *set di dati di addestramento*, dati predittivi, dati personali, dati aziendali, dati finanziari, dati di posizionamento, dati generici. Il valore che va estraendo l'industria dell'AI oggi, è un *quid tertium* rispetto a quanto ha già trovato e raccolto negli anni precedenti, facendo *scraping* sul web, trattando i dati dei clienti dei servizi online o attraverso le acquisizioni di promettenti *startup* del settore. Si calcola che dal 2017 ad oggi, le quattro grandi piattaforme digitali globali americane (i.e. Microsoft, Apple, Meta e Alphabet) hanno investito miliardi di dollari per acquisizioni di imprese di AI: finora sono 59, con Apple che distanzia tutte le altre *BigTech* con ben 21 acquisizioni, rispetto alle 12 di Microsoft⁸⁰. Non si contano invece le *partnership*, i paternariati strategici, gli investimenti in capitale o la cessione gratuita di servizi cloud fatta a centri di ricerca specializzati in intelligenza artificiale.

In un recente articolo apparso sul Financial Times⁸¹ in tema di concentrazioni, si evidenziano differenze tra il sistema antitrust americano e quello europeo *“(...) data la naturale capacità delle piattaforme di sfruttare gli effetti di rete per schiacciare potenziali concorrenti fin dalla nascita – e per imporre elevati costi di cambiamento ai consumatori, la FTC prevede di esaminare non solo i prezzi o anche le dimensioni del mercato, ma anche se le partnership AI tra aziende comportano forme di esclusività o di accesso preferenziale ai dati”*. **Solo favorendo l'accesso a questi enormi dataset, si potrebbe ripristinare un “level playing field” e favorire la nascita e la crescita di altre imprese capaci di competere sul mercato globale dell'intelligenza artificiale, contribuendo al benessere sociale ed economico del paese.**

⁷⁷ Le economie di scopo sul lato della domanda di AI, derivano dal raggruppamento di diversi servizi in modalità “one stop shop” in modo da creare una complementarità tra prodotti correlati (es. Copilot su prodotti Microsoft, integrazione Bard nell'ecosistema Google ecc.).

⁷⁸ Cfr. MLex Special Report, 22 January 2015 ‘Vestager Interview’

⁷⁹ Cfr. Viktor Mayer-Schönberger - Access Rules - Freeing Data from Big Tech for a Better Future <https://www.ucpress.edu/book/9780520387737/access-rules>

⁸⁰ Cfr. Financial Times, 25 gennaio 2024

⁸¹ Cfr. FT The great US-Europe antitrust divide - There are fundamental differences in how the two see corporate concentration <https://www.ft.com/content/065a2f93-dc1e-410c-ba9d-73c930cedc14>



L'università di Stanford⁸² ha recentemente promosso un approccio molto simile alla catena di approvvigionamento che potrebbe regolare meglio la fornitura e la *governance* dei dati di addestramento per l'AI, apportando maggiore trasparenza all'intero ciclo di vita dello sviluppo dei *dataset* e mettendo in risalto la necessità di una gestione responsabile dei dati, dalla creazione alla cancellazione. Questa strategia richiederebbe sicuramente una forte volontà politica ed una regolamentazione adeguata che tenga conto di tutto quanto sopra espresso. Nel prossimo capitolo analizzeremo alcune proposte di politica industriale che tengano conto dell'interesse nazionale nello sviluppo dell'intelligenza artificiale.

⁸² Cfr. Jennifer King, Caroline Meinhardt - "Rethinking Privacy in the AI Era: Policy Provocations for a Data-Centric World"
<https://hai.stanford.edu/white-paper-rethinking-privacy-ai-era-policy-provocations-data-centric-world#:~:text=February%2022%2C%202024-.Rethinking%20Privacy%20in%20the%20AI%20Era%3A%20Policy.for%20a%20Data%2DCentric%20World&text=This%20white%20paper%20explores%20the.harms%20in%20an%20AI%20era.>



7] LA SOVRANITÀ DIGITALE E IL RUOLO DELLA COMMESSA PUBBLICA ITALIANA SU AI

Per garantire all'Italia uno sviluppo di sistemi di intelligenza artificiale che possa ambire ad essere competitivo nello scenario globale, occorre un cambiamento nell'approccio e negli strumenti di politica industriale finora adottati.

LA COMMESSA PUBBLICA PER LO SVILUPPO DELLE COMPETENZE

Per ridurre il divario dell'Italia rispetto al resto d'Europa e rispetto ai *competitor* americani, è **oggi più che mai necessario che il nostro programma di *public procurement* di beni e servizi indirizzati per il settore pubblico, vada a sostenere le imprese italiane ad alta tecnologia in modo da accelerare la produzione locale delle diverse soluzioni innovative per l'erogazione dei nostri servizi pubblici.** Questo perché gli investimenti pubblici a livello nazionale possono andare a sollecitare nuovi investimenti pubblici a livello locale, ed a seguire, anche quelli dei privati creando un volano positivo di crescita della nostra economia interna.

Per far questo, è fondamentale che le decisioni sulla transizione digitale della pubblica amministrazione, abbiano prioritaria contezza dei vari soggetti italiani interessati a fornire infrastrutture e servizi digitali alla PA in modo che si possano favorire investimenti che portino valore aggiunto al nostro Paese anche in termini di competenze specializzate, che invece oggi si vedono costrette - una volta formate nelle nostre università - a doversi spostare all'estero. Purtroppo la forza lavoro specializzata nell'intelligenza artificiale e nelle materie STEM, viene formata dalle nostre università ma poi finisce inesorabilmente per essere attratta dalle aziende tecnologiche più note a livello globale.

E' possibile che molti di questi ragionamenti siano ormai passati in subordine rispetto ad urgenze contingenti, anche se per molti mesi si è parlato di sovranità digitale⁸³. **Di fronte ai nuovi scenari economici e macroeconomici su larga scala, il nostro interesse resta però quello di garantire una risposta di servizi locali che possano concorrere a fornire una vera alternativa alle Bigtech**, con servizi di qualità a prezzi competitivi, grazie alla capacità di investire in sviluppo, innovazione e infrastruttura. Ogni volta che installiamo un'applicazione AI, scegliamo un servizio cloud o acquistiamo un dispositivo, stiamo prendendo una decisione che influenzerà la nostra indipendenza tecnologica. Dobbiamo essere consapevoli delle implicazioni di queste scelte: chi controlla i nostri dati? Quali standard tecnologici stiamo adottando? **Se la sovranità digitale è un desiderio strategico a livello di nazione, l'indipendenza tecnologica è la sua manifestazione pratica a livello individuale e collettivo. Pertanto, dovremmo considerare attentamente le**

⁸³ Cfr. "Digital sovereignty for Europe"

[https://www.europarl.europa.eu/RegData/etudes/BRIE/2020/651992/EPRS_BRI\(2020\)651992_EN.pdf](https://www.europarl.europa.eu/RegData/etudes/BRIE/2020/651992/EPRS_BRI(2020)651992_EN.pdf)



tecnologie che adottiamo e promuovere una cultura di consapevolezza e autonomia digitale.

Seguendo l'esempio di altri Stati europei, anche l'Italia può introdurre misure selettive per garantire alle PMI dell'innovazione attive nel AI di poter ricevere adeguate porzioni di proventi dalle vendite di servizi digitali, per continuare a crescere ed investire grazie all'importante volano costituito dalla commessa pubblica. **Occorre di conseguenza, anche superare il peso della burocrazia che penalizza in misura non proporzionale le aziende più piccole, rispetto ai grandi player globali.** Una grande occasione ci è offerta oggi dal PNRR. Scrive il Prof. Gianfranco Viesti⁸⁴ *"I Piani rappresentano una risposta allo shock Covid e quindi vanno realizzati in un arco di tempo molto ristretto: gli interventi vanno necessariamente conclusi tutti entro il 2026 (...) il Piano manca di una visione di politica industriale: sembra partire dall'idea che basti potenziare la domanda per ottenere un sufficiente sviluppo dell'offerta interna"*.

L'impegno pubblico ad accelerare l'economia italiana è dunque intrinseco alle finalità stesse del PNRR. Per ogni euro investito in Italia nella transizione digitale, si dovrebbe andare a produrre uno *spill over effect* su tutti gli altri settori individuati nel Piano, creando ulteriori ricadute positive su tutta l'economia italiana. Se si vuole ridurre il divario dell'Italia sull'intelligenza artificiale rispetto al resto d'Europa e del mondo, **è più che mai necessario quindi che il programma di public procurement di beni e servizi necessari per la digitalizzazione del settore pubblico nazionale e locale, vada a sostenere le imprese italiane ad alta tecnologia**, in modo da accelerare la produzione interna di soluzioni di intelligenza artificiale che diano sostanza alla transizione, ma anche al rilancio economico del Paese. Una priorità che potrebbe essere facilmente individuata dal Governo è sicuramente quella di incentivare la nascita di nuove imprese nel campo dell'intelligenza artificiale ma prima ancora si dovrebbe provvedere allo sviluppo ed alla crescita di quelle realtà italiane che sono già riconosciute nella filiera dell'AI. Tutto questo oggi non sta accadendo, in parte anche per la cd. *"cattura del regolatore"*⁸⁵.

Per interrompere questo andamento negativo non bastano più affermazioni di principio che hanno dimostrato limiti piuttosto severi nel momento in cui la sovranità digitale è stata prima problematizzata, poi ridotta ad una garanzia di mantenere la produzione di apparati o i dati all'interno di confini territoriali precisi. Ma è davvero troppo poco ed ancora oggi, **assistiamo ad una resa del decisore pubblico di fronte al dominio tecnologico di fornitori esteri, anche quando non sembra affatto necessario perché sono già presenti sul mercato valide alternative italiane di prodotti e servizi della società dell'informazione.**

MISURE DI CONTRASTO AL LOCK-IN TECNOLOGICO

Alcune delle più grandi realtà economiche pubbliche e private che hanno già intrapreso e concluso il percorso di migrazione al cloud si trovano oggi una nuova voce di costo nel

⁸⁴ Cfr. Gianfranco Viesti *Riuscirà il Pnrr a rilanciare l'Italia?* - 2023 Donzelli Editore

⁸⁵ Cfr. Wikipedia, "Regulatory capture" https://it.wikipedia.org/wiki/Regulatory_capture



bilancio. Questo perché *il marketing di BigTech*⁸⁶, fatto di *voucher* gratuiti pubblici⁸⁷ e privati⁸⁸ per la migrazione in cloud, ha aperto i cancelli a una serie di risorse tecnologiche e di funzionalità di cui spesso le aziende e le pubbliche amministrazioni non hanno ancora bisogno. E quei cancelli si chiudono velocemente lasciando i clienti sia pubblici che privati, intrappolati in fastidiosi *effetti cd. di "lock-in"* da cui è difficile liberarsi⁸⁹. **Il lock-in tecnologico è dovuto principalmente al vantaggio iniziale delle soluzioni di cloud pubblico che prevedono voucher e sconti significativi rispetto ai modelli di data center tradizionali.** Ma nonostante la maggiore flessibilità disponibile nel cloud, esiste il rischio di diventare dipendenti dai prodotti e servizi solo di determinati fornitori (spesso extraeuropei) e viene in evidenza nel momento in cui il passaggio da una tecnologia ad un'altra o da un fornitore a un altro risulta ostacolato dal provider, oppure difficile, dispendioso in termini di tempo e sproporzionatamente costoso.

Ad esempio durante il lockdown, Google ha reso Classroom gratuito per le scuole italiane, contribuendo a garantire la continuità dell'istruzione ma successivamente, è diventato un servizio a pagamento. L'uso di Classroom ha evidenziato l'importanza dell'indipendenza tecnologica. Le scuole e gli studenti dovrebbero valutare attentamente le loro scelte tecnologiche anche in funzione della privacy. Lo Stato da parte sua, dovrebbe indirizzarsi a servizi italiani di didattica a distanza, visto anche il successo e le casistiche ormai consolidate nell'Istruzione. L'ex Garante Privacy Antonello Soro, incoraggiava l'Italia a dotarsi di una *"piattaforma pubblica gestione dati. Non possiamo continuare ad appoggiarci a strutture cinesi o americane"*⁹⁰. (...) *"forse è più prudente utilizzare il registro elettronico, che pure non è privo di problemi: fra il registro elettronico e la piattaforma di una multinazionale di cui non si sa nulla", raccomandava Soro, "è meglio nel presente dare indicazioni perché le scuole ricorran tutte le volte che è possibile al primo"*.

A tal proposito, la Francia⁹¹ ha recentemente introdotto una regolamentazione dei voucher commerciali dei servizi cloud, che nascono proprio per attrarre le *startup* e farle rimanere incollate dentro un castello di funzioni gratuite in sistemi che spesso non si usano e che successivamente si ritrovano a dover pagare. E' noto infatti, che gli sviluppatori di AI che non dispongono già di proprie infrastrutture e *data center* si rivolgono generalmente ai fornitori di servizi cloud (cd. "CSP") per l'elaborazione dei dati, ritrovandosi spesso ingabbiati in contratti pluriennali che aumentano la dipendenza tecnologica da fornitori esteri⁹² - gli stessi che stanno dominando il mercato globale del cloud e dell'intelligenza artificiale.

⁸⁶ AWS Public Sector Symposium - <https://aws.amazon.com/it/events/summits/public-sector-symposium-rome/>

⁸⁷ Dal PNRR altri 50 milioni per migrare in cloud dati e servizi dei Comuni <https://innovazione.gov.it/notizie/articoli/dal-pnrr-altri-50-milioni-per-migrare-in-cloud-dati-e-servizi-dei-comuni/>

⁸⁸ "AWS aiuta le startup a dare vita alle loro idee tramite AWS Activate". Se fai parte di AWS Activate, potresti avere diritto a ricevere fino a 100.000 dollari di crediti per compensare la fattura AWS. <https://aws.amazon.com/startups/credits?lang=it>

⁸⁹ Sulle tecniche anti-Lock-in si veda GOV.UK "Managing technical lock-in in the cloud" <https://www.gov.uk/guidance/managing-technical-lock-in-in-the-cloud>

⁹⁰ Cfr. Key4biz - "Scuola e sovranità digitale. Il ministro francese boccia Google e Microsoft. E in Italia?" - <https://www.key4biz.it/scuola-e-sovranita-digitale-il-ministro-francese-boccia-google-e-microsoft-e-in-italia/430956/>

⁹¹ Cfr - Senato francese e voucher cloud - <https://www.senat.fr/tableau-historique/pjl22-593.html>
 Il. - "Un fornitore di servizi di cloud computing può concedere un credito di cloud computing a una persona che svolge attività di produzione, distribuzione o servizio solo per un periodo limitato che non può superare un anno, anche se la concessione di tale credito viene rinnovata. "La concessione di un bene di cloud computing non può essere accompagnata da una condizione di esclusività di alcun tipo da parte del beneficiario nei confronti del fornitore di tale bene".

⁹² Alcuni esempi: 1) Google Cloud: Sconti per impegno di utilizzo (CUD) per Compute Engine <https://cloud.google.com/compute/docs/instances/committed-use-discounts-overview?hl=it> 2) AWS Savings plan <https://docs.aws.amazon.com/savingsplans/latest/userguide/what-is-savings-plans.html>



Un fornitore di servizi cloud locale non può permettersi di mettere sul mercato dei buoni gratuiti da centinaia di migliaia di euro (o dollari) in servizi cloud da scontare in fattura, e quindi inizialmente potrebbe sembrare più costoso rispetto a chi invece offre crediti gratuiti da scontare in fattura per catturare l'utente. E' un tema, questo, su cui le Autorità Antitrust degli Stati Uniti⁹³, Francia⁹⁴ e del Regno Unito⁹⁵ hanno aperto un focus specifico. Riteniamo a tal fine che anche in Italia si dovrebbe limitare l'importo dei "crediti cloud" che possono essere offerti dai fornitori di servizi di cloud computing ai clienti. Questi *voucher*, teoricamente costituiscono un aiuto alle *startup*, ma tipicamente vengono poi utilizzati dagli operatori cloud dominanti per stroncare ogni concorrenza con altri fornitori di servizi cloud ma soprattutto creano una dipendenza tecnologica per i clienti, spesso costretti a rapporti di esclusiva⁹⁶. Ciò rappresenta una forte minaccia per la nostra sovranità digitale.

In particolare, l'utilizzo dei "voucher cloud" offerti dai privati può rendere le nostre PMI più vulnerabili a cambiamenti nei prezzi o nelle condizioni di servizio dei fornitori cloud extraeuropei limitando la capacità delle PMI di innovare e sviluppare le proprie applicazioni di intelligenza artificiale. Ecco perché oltre a un livello massimo di credito gratuito erogabile, si dovrebbe anche seguire l'esempio francese ed imporre un limite temporale massimo di un anno per erogare eventuali crediti cloud. Crediti che comunque dovrebbero essere soggetti a tassazione⁹⁷ perché il fatto che una parte del corrispettivo percepito per la cessione del servizio non sia stata materialmente versata dall'utente finale stesso, ma sia stata messa a sua disposizione attraverso un *voucher*, non assume alcuna rilevanza ai fini della determinazione della base imponibile del soggetto che eroga il servizio, almeno questa è la regola che vige per le imprese che operano in Italia.

IL CLOUD NAZIONALE, LA SOVRANITÀ DIGITALE E LA SICUREZZA

L'Italia dovrebbe quindi puntare a essere molto forte sul cloud e sulla *cybersecurity* a livello nazionale mentre sulle altre priorità tecnologiche dovrebbe puntare ad alleanze in Europa.

Come ha correttamente affermato il CEO di Leonardo, Roberto Cingolani⁹⁸, rispondendo a una domanda sulle priorità per l'Italia a livello tecnologico nel corso dell'audizione alla Commissione attività produttive della Camera: *"Punterei ad un cloud molto forte a livello nazionale: c'è un tentativo adesso ma siamo un po' lenti, per cui già se avessimo buona*

⁹³ Cfr. FTC - "Cloud Computing RFI: What we heard and learned" - https://www.ftc.gov/policy/advocacy-research/tech-at-ftc/2023/11/cloud-computing-rfi-what-we-heard-learned?utm_campaign=new_ftc_blog_post_summari&utm_content=1700157150&utm_medium=social&utm_source=twitter

⁹⁴ Cfr. "Cloud computing: The Autorité de la concurrence issues an opinion on certain provisions of the draft law to secure and regulate the digital space" <https://www.autoritedelaconcurrence.fr/en/press-release/cloud-computing-autorite-de-la-concurrence-issues-opinion-certain-provisions-draft>

⁹⁵ Ofcom - Cloud services market study https://www.ofcom.org.uk/_data/assets/pdf_file/0029/256457/cloud-services-market-study-interim-report.pdf

⁹⁶ Ad esempio Google Cloud è il fornitore esclusivo di servizi cloud per Anthropic come Microsoft per OpenAI. <https://www.anthropic.com/news/anthropic-partners-with-google-cloud>

⁹⁷ Si legga sul punto "Il trattamento Iva delle gift card" a cura di Ennio Vial <https://clienti.euroconference.it/documenti/sProvvedim/20111219sD/20111219A05.pdf>

⁹⁸ Indagine conoscitiva Intelligenza artificiale - Audizioni - Roberto Cingolani, amministratore delegato di Leonardo S.p.A <https://webtv.camera.it/evento/24574>



competenza di calcolo e informazioni dei cittadini protette in un cloud mi sentirei in un Paese più sicuro". Lo scenario italiano deve essere pronto a cogliere la sfida del cloud pubblico con maggiore convinzione, facendo leva sulle competenze italiane per la progettazione, l'implementazione e l'ottimizzazione delle soluzioni cloud, che sono in continua evoluzione e che richiedono una formazione costante.

Come è noto, infatti, l'intelligenza artificiale pone anche notevoli rischi di natura geopolitica e securitaria. Sul tema dell'uso legittimo della difesa informatica, c'è una nutrita rassegna di scritti internazionali che hanno analizzato l'ipotesi di un attacco cibernetico ad infrastrutture critiche, come ad esempio una centrale nucleare. E' un'ipotesi che adesso, calata nella realtà quotidiana che stiamo vivendo, assume una concretezza tangibile ed assai preoccupante. In una guerra fredda informatica, ci si muove con semplici finalità deterrenti, volte a prevenire *cyber* attacchi. Ci si muove quindi in un'area grigia che viene prima dell'uso della forza armata. Oggi però è contestuale: siamo in uno scenario che si avvicina pericolosamente alla guerra ibrida. Non è chiaro in che modo si possa evitare che per colpire una nazione-bersaglio si possano evitare effetti indiretti e negativi a nazioni considerate amiche.

Negli ultimi mesi abbiamo assistito a scelte inedite di operatori globali che hanno sfidato gli ordini europei andando quindi a depotenziare la capacità di incidere su scelte strategiche di difesa comune. Resta un rischio inaccettabile il fatto che alcuni colossi globali del cloud e dell'intelligenza artificiale agiscano eseguendo un ordine da "madrepatria" ma in via difforme da quanto ordinato dalle Autorità del Paese estero in cui operano, e questo sta già accadendo per le più varie ragioni di convenienza o di opportunità.

Gli Stati Uniti ad esempio, potranno ordinare ai fornitori di servizi cloud americani di fornire i dati dei clienti stranieri⁹⁹. Sarà obbligatorio verificare l'identità di qualsiasi persona straniera, compresa la modalità di pagamento, l'account IaaS, l'email, il telefono e gli indirizzi IP usati per accedere all'account. Tutto questo, formalmente, per evitare che i sistemi di intelligenza artificiale americani vengano usati per scopi malevoli da altri Stati. L'amministrazione americana inoltre, si sta preparando a utilizzare il *Defense Production Act*¹⁰⁰ per obbligare le aziende tecnologiche a informare il governo quando addestrano un modello di intelligenza artificiale utilizzando una quantità significativa di potenza di calcolo. In passato, alcuni rappresentanti dei fornitori cloud americani hanno affermato di non aver mai ricevuto l'ordine di un giudice americano di accedere ai dati di un cittadino europeo conservati sulla loro infrastruttura. Non sappiamo se sia vero e quante volte sia accaduto. Sappiamo solo che sarebbe potuto accadere in ogni momento e di fatto la situazione si sta aggravando notevolmente a fronte dell'ampliamento del perimetro di cybersicurezza con cui gli Stati Uniti intendono difendere la propria potenza di calcolo alimentata dall'intelligenza artificiale, nel caso in cui sia impiegata da altri Stati esteri.

Quando un Paese come l'Italia affida i dati critici ad operatori esteri in partnership con operatori locali, deve quindi rivalutare meglio che potrebbe perderne il controllo a

⁹⁹ DEPARTMENT OF COMMERCE - 15 CFR Part 7 [Docket No. 240119-0020] Taking Additional Steps To Address the National Emergency With Respect to Significant Malicious Cyber-Enabled Activities <https://aboutgov.com/bct3>

¹⁰⁰ OpenAI and Other Tech Giants Will Have to Warn the US Government When They Start New AI Projects <https://www.wired.com/story/openai-tech-giants-us-government-ai-projects/>



fronte di un ordine di blocco delle forniture e dei servizi, azionato ad esempio da uno Stato estero impegnato a difendersi da un attacco cibernetico. Il sistema di garanzie “contrattuali” nulla potrebbe fare per resistere a questa ipotesi, che non riguarda l’aspetto economico e nemmeno la riservatezza del dato, ma la sicurezza e l’ordine pubblico.

Sempre più spesso si fa strada il dubbio sulla effettiva *governance* dei dati. Ecco perché alcune grandi aziende europee, hanno sentito il bisogno di recuperare¹⁰¹ – almeno in parte - il controllo sulla propria infrastruttura scegliendo di affidarsi ad un cloud provider nazionale, in grado di garantire e proteggere il perimetro giuridico dei dati senza dover ricorrere ad una promessa contrattuale, che oggi è un “*pledge*”¹⁰², domani un “*commitment*”¹⁰³ e dopodomani una sanzione sicura per violazione della privacy o della concorrenza. **Scegliere un fornitore di servizi cloud locale per allenare un sistema di Intelligenza Artificiale piuttosto che una delle Big Tech oltreoceano, può essere la scelta più vantaggiosa, soprattutto per il contesto italiano** che sta vivendo proprio adesso i principali passaggi di questa rivoluzione digitale. Questo perché un operatore locale può accompagnare meglio certi processi e soprattutto in maniera più trasparente e dinamica, permettendo di combinare i benefici del cloud e dell’intelligenza artificiale, con i benefici della prossimità, della personalizzazione e della conformità legale dell’azienda alle regole di settore (es. NIS2, Dora, GDPR, e altro).

¹⁰¹ Cfr. Dario Denni - “NIS2 per le PMI: Il marketing della repatriation amplifica l’importanza di una migrazione al Cloud” <https://www.key4biz.it/nis2-per-le-pmi-il-marketing-della-repatriation-amplifica-limportanza-di-una-migrazione-al-cloud/462238/>

¹⁰² “Pledge” è un impegno contrattuale. Si veda da ultimo AWS Digital Sovereignty Pledge <https://aws.amazon.com/it/blogs/security/aws-digital-sovereignty-pledge-control-without-compromise/>

¹⁰³ “Commitment” è un impegno sovente preso di fronte ad un’Autorità o durante un procedimento di contestazione al fine di evitare una sanzione economica. Da ultimo si veda: Microsoft’s AI Access Principles: Our commitments to promote innovation and competition in the new AI economy <https://blogs.microsoft.com/on-the-issues/2024/02/26/microsoft-ai-access-principles-responsible-mobile-world-congress/#:~:text=The%20standard%20covers%20six%20key,inclusiveness%3B%20transparency%3B%20and%20accountability.>



CONCLUSIONI

La tesi ricorrente in questo *paper* è che per garantire all'Italia uno sviluppo di sistemi di intelligenza artificiale occorre introdurre misure correttive di alto valore sistemico. Tra le misure emerse in questa analisi, assumono maggiore rilevanza quelle che possono avere la capacità di far crescere un mercato competitivo, aperto e sicuro dell'intelligenza artificiale.

Queste misure riguardano principalmente la regolazione dell'accesso e del controllo dei dati di input utilizzati per allenare i sistemi di intelligenza artificiale, la prevenzione delle distorsioni e delle concentrazioni del mercato da parte delle piattaforme dominanti extraeuropee ed il sostegno alle aziende italiane della filiera dell'AI. Da sole, queste misure potrebbero non essere sufficienti per affrontare tutte le questioni che abbiamo sollevato in precedenza, ma riteniamo che costituiscano un importante punto di partenza per andare avanti nella giusta direzione.



L'ECONOMIA DEI DATI NELL'INTELLIGENZA ARTIFICIALE

10 MISURE PER COMBATTERE LE DISEGUAGLIANZE INDUSTRIALI CONTRO OGNI ESTERNALITÀ ECONOMICA E SOCIALE

1. **garantire e regolamentare l'accesso e la trasparenza ai dati di input per allenare sistemi di intelligenza artificiale** in modo che possano essere accessibili alle autorità di controllo e ad altri soggetti che sviluppano soluzioni di intelligenza artificiale e per aprire l'ecosistema dei dati a un insieme di attori più ampio rispetto alle grandi aziende tecnologiche che ne hanno un quasi monopolio;
2. **dare maggiori poteri ex-ante alle Autorità Antitrust per prevenire fenomeni distorsivi ricorrenti** (concentrazioni, acquisizioni di startup AI, self preferencing di servizi, accordi restrittivi sull'uso delle risorse computazionali, contratti di fornitura esclusiva, clausole di lock-in tecnologico ecc);
3. **prevenire la concentrazione del mercato dall'AI attraverso l'applicazione di misure che ostacolano l'estensione della dominanza sui mercati ancillari** (i.e. featuring della AI sui sistemi e prodotti già in uso gratuito oppure offerti con prezzi predatori; utilizzo dei dati raccolti senza consenso o con finalità diverse per allenare sistemi di AI);
4. **vietare l'utilizzo di "crediti cloud" in forma di voucher da scontare nell'acquisto di servizi cloud dalle piattaforme dominanti** (i.e. introdurre limiti temporali e quantitativi all'uso dei voucher);
5. **avviare un'indagine del mercato italiano del cloud computing e dell'intelligenza artificiale** come in Francia e nel Regno Unito per censire le aziende operanti nel settore e far emergere dei possibili temi di natura competitiva;
6. **favorire una commessa pubblica che sia indirizzata a sviluppare aziende italiane della filiera dell'intelligenza artificiale**, compresi i provider di servizi cloud, anche a tutela della sovranità digitale e della cybersicurezza nel trattamento di dati critici e strategici;
7. **individuare un marketplace certificato ACN di servizi di intelligenza artificiale** che segua le caratteristiche di qualificazione per la sicurezza dei servizi cloud per la PA;
8. **introdurre programmi di educazione alle materie STEM che non si limitino ai sistemi dominanti extraeuropei**, ma includano soluzioni italiane, in lingua italiana, di fornitori italiani che pure esistono e competono sul mercato;
9. **garantire l'efficacia dei procedimenti regolamentari e tempistiche in linea con l'evoluzione del mercato** accompagnate da sanzioni proporzionali che rappresentino un deterrente effettivo a fronte delle condotte abusive;
10. **proporre una riduzione delle soglie per la notifica ed un ampliamento dei soggetti e dei servizi ricompresi nella disciplina europea dei mercati digitali**, segnatamente il DMA che non include il cloud ed i modelli fondativi, ed il DSA per i sistemi di intelligenza artificiale in modo armonizzato con AI-Act di prossima pubblicazione.



L'ECONOMIA DEI DATI NELL'INTELLIGENZA ARTIFICIALE

INDICE DEGLI ARGOMENTI

PREMESSA

1] DATA DRIVEN ECONOMY: AI AFFAMATA DI DATI

- I DATI DI ALLENAMENTO DEI SISTEMI DI INTELLIGENZA ARTIFICIALE
- INTESI RESTRITTIVE E ACCORDI DI ESCLUSIVA TRA CLOUD E AI
- MERGER CONTROL: FUSIONI E ACQUISIZIONI BASATE SUI DATI
- LA FILIERA DELL'INTELLIGENZA ARTIFICIALE: IL RUOLO DEL CLOUD
- RISCHI E PERICOLI DELL'INNOVAZIONE SENZA CONSENSO

2] IL PERIMETRO SOGGETTIVO DEL FENOMENO OLIGOPOLISTICO LEGATO ALL'AI

- ELEMENTI CHE DETERMINANO LA CONCENTRAZIONE DEL MERCATO
- LA CONTESSA DEI TALENTI AI: IL RUOLO DELLE UNIVERSITÀ

3] LA STRATEGIA EUROPEA SUI DATI E LA RAPIDA OBSOLESCENZA DELLE REGOLE

- ELEMENTI DI RITARDO NEL PROCEDIMENTO REGOLAMENTARE
- LA TENACE OPPOSIZIONE DEI SOGGETTI REGOLATI NEL MERCATO
- CORRETTIVI AL DIGITAL MARKET ACT: I SERVIZI CLOUD

4] GLI EFFETTI DEL DISALLINEAMENTO TRA REGOLE E I TEMPI DEI PROCEDIMENTI

- L'EFFICACIA DELLE NORME TRA PROCEDIMENTI E SANZIONI
- IL WEBSCRAPING PER ALLENARE SISTEMI DI AI: IL CASO CHATGPT

5] COMBATTERE LE DISUGUAGLIANZE INDUSTRIALI

- ETICA E DIRITTO: L'INSEGNAMENTO DEL PROF. STEFANO RODOTA'
- CESSIONE DEI DATI PERSONALI: IL RUOLO DEI PRIVATI
- LA CENTRALITÀ DELL'UOMO NEL DIBATTITO SU AI
- MODELLI OPACHI: TECNICHE DI ACQUISIZIONE DEL CONSENSO
- VERSO UN APPROCCIO MULTIDISCIPLINARE
- RESPONSABILITÀ DEGLI ALGORITMI: RICONOSCIMENTO E DIRITTO

6] L'ACCESSO AI BIG DATA COME PRESUPPOSTO DELLA COMPETIZIONE E DEMOCRATIZZAZIONE DELL'AI

- ECONOMIE DI SCALA E DI SCOPO - LE ESTERNALITÀ NEGATIVE
- IL MERCATO DEI DATI : LE ASIMMETRIE E LE CONCENTRAZIONI

7] LA SOVRANITA' DIGITALE ED IL RUOLO DELLA COMMESSA PUBBLICA ITALIANA SU AI

- LA COMMESSA PUBBLICA PER LO SVILUPPO DELLE COMPETENZE
- MISURE DI CONTRASTO AL LOCK-IN TECNOLOGICO
- IL CLOUD NAZIONALE, LA SOVRANITA' DIGITALE E LA SICUREZZA

CONCLUSIONI

- 10 MISURE PER COMBATTERE LE DISUGUAGLIANZE INDUSTRIALI
CONTRO OGNI ESTERNALITÀ ECONOMICA E SOCIALE



Seeweb è il Cloud Computing Provider italiano focalizzato sull'assistenza al cliente.

Da 20 anni supportiamo web agency, system integrator, sviluppatori e aziende con servizi di hosting web, soluzioni di posta elettronica, Colocation, streaming e soprattutto Cloud.

La nostra presenza multipla in Italia con più datacenter proprietari nelle sedi di Milano Caldera, Sesto San Giovanni, Frosinone, insieme alla server farm in Svizzera, ci permette di offrire servizi IT ridondati e basati su un'infrastruttura di rete dall'accurata scelta di fornitori e politiche di peering.

La nostra passione è seguire il cliente e accompagnarlo nelle varie fasi di crescita e sviluppo del suo progetto online, garantendogli sempre i massimi standard di performance e sicurezza.



<http://www.seeweb.com>
info@seeweb.com

Sede Legale e Operativa: Corso Lazio, 9/a 03100 Frosinone IT

Filiale nord: Via Caldera, 21 20153 Milano IT p.iva 02043220603

L'ECONOMIA DEI DATI
NELL'INTELLIGENZA ARTIFICIALE

Una guida per combattere le diseguaglianze industriali
contro ogni esternalità economica e sociale



seeweb

Pubblicato nel mese di Aprile 2024

seeweb