

di Roberto Rosso

Quali certezze abbiamo noi tutti, intesi come genere umano, sul futuro dell'umanità, quali strumenti abbiamo a disposizione per descrivere le possibili traiettorie di questo futuro? La risposta a questa domanda, alle miriadi di domande in cui si si moltiplica quando osserviamo la realtà più da vicino, genera la costruzione di molteplici modelli previsionali che richiedono il contributo di molteplici discipline, di una mole di dati che cresce esponenzialmente e l'uso di dispositivi la cui potenza cresce in maniera analoga. Siamo ancora una volta a ragionare dello sviluppo scientifico-tecnologico, il suo realizzarsi in sistemi socio-tecnici che strutturano le formazioni sociali. Uno sviluppo che è causa delle crisi, delle rotture, delle catastrofi attuali e prossime venture e assieme si pone come rimedio, strumento per affrontarle, per affrontare un presente luogo del mutamento, che diventa rapidamente un passato obsoleto. I cambiamenti sono *'disruptive'* - non mi viene un termine equivalente così evocativo e onomatopeico- e avvicinano l'orizzonte degli eventi imprevedibili. C'è ampia materia per una riflessione filosofica sulla metafisica del tempo¹ così come di fronte allo sviluppo scientifico-tecnologico, ai suoi esiti controversi e contraddittori, si pongono domande fondamentali sul senso delle capacità conoscitive e trasformative della realtà del genere umano, il suo destino come specie capace di trasformare il mondo in maniera radicale, compresa la sua estinzione, mentre ne approfondisce e ne estende la conoscenza.

Sono diversi i livelli di riflessione sollecitati da una tale condizione, nonostante i molteplici approfondimenti ciò che viene a mancare è una capacità di cooperazione, di collaborazione a livello globale, all'altezza delle rotture che caratterizzano la riproduzione del mondo così come lo abbiamo conosciuto e lo conosciamo. Tale unità di intenti non si realizza nei confronti del riscaldamento globale, dei cambiamenti climatici conseguenti, come hanno dimostrato le ultime conferenze mondiali sul clima e sulla biodiversità e non si sta realizzando nel campo delle tecnologie digitali, dell'Intelligenza Artificiale che ne riassume e ne amplifica le potenzialità.

L'Intelligenza Artificiale, intesa come complesso sistema socio-tecnico, pervade ormai tutte le filiere tecnologiche, regolando l'attività delle reti e dei dispositivi digitali a cui siamo connessi, la curva dello sviluppo che procede esponenzialmente sembra aver preso a salire sempre più verticalmente. Si accumulano, proporzionalmente alle sue prestazioni, gli interrogativi sui rischi che genera. Al centro dell'attenzione sono i cosiddetti *LLM Large Language Module*, la loro capacità di interloquire di produrre risposte ad ogni sorta di quesiti, in linguaggio naturale, ciò che ne ha permesso la sperimentazione ad una amplissima platea, nei campi più diversi, compresi i linguaggi di programmazione. L'addestramento dei diversi modelli di reti neurali, che si fonda come è noto su gigantesche basi di conoscenza, può produrre nelle loro risposte quelle che vengono definite allucinazioni²; la risposta ai quesiti è un processo che possiamo definire probabilistico, nel creare una successione di parole, e come tale può ricostruire anche falsi eventi, risposte paradossali. L'articolo citato in nota ci dice che si stanno facendo progressi, che si usano i chatbots, e i 'vecchi' motori di ricerca per analizzare le loro stesse risposte³. Già in passato, ancor prima dell'esplosione degli LLM, si erano verificati comportamenti del tutto imprevedibili in sistemi in cui veniva messa alla prova la capacità degli algoritmi di apprendere dal proprio contesto operativo⁴. Anche i sistemi di I.A. sono addestrati nel cosiddetto Reinforcement Learning -apprendimento rinforzato- concetto

intuitivo dove l'applicazione tiene conto del punteggio attribuito ad ogni scelta⁵.

Nell'analisi delle possibilità è compresa l'ipotesi che il dispositivo rifiuti di spegnersi poiché questo porrebbe fine al riconoscimento. Il fatto i progettisti si applicano ad una tale ipotesi mostra a quali livelli di autorappresentazione si ipotizzi possa arrivare il sistema algoritmico dell'I.A.

I LLM hanno reso popolare il tema dell'I.A, tuttavia l'uso estensivo può mettere a rischio i diritti fondamentali delle persone, può dotare ogni tipo di potere di straordinarie capacità di controllo sociale e personale, può rendere fallaci tecniche di controllo legittime in base alle modalità addestramento; può essere estremamente pericoloso affidarsi senza controlli in settori come la sanità e la farmaceutica.

L'eticità, l'analisi dei rischi insiti nell'uso dei dispositivi di I.A. il governo in generale di quel sistema sempre più potente e opaco, costituisce in sé un fattore critico con caratteri sempre più complessi. I rischi sono specifici di ogni tipo di attività e le regole generali istituiti dalle autorità statuali è prevedibile che incontrino difficoltà a seguire l'evoluzione delle tecnologie nelle sempre più diverse applicazioni. La ricerca di nuovi farmaci costituisce una attività particolarmente complessa e dispendiosa in cui l'I.A. è diventata uno strumento sempre più necessario; ne abbiamo un esempio nel caso della società AstraZeneca⁶.

L'attività di audit ovvero controllo non è ovviamente una novità nella produzione del software, la cui complessità incorpora una quota crescente di procedure di controllo; nel caso dell'I.A. è intuitivo pensare che la stessa attività di controllo costituisca un processo basato sulla capacità apprendere. A fronte della variabilità dei processi da tenere sotto controllo sta la definizione di ben motivati principi guida, si sta imponendo all'attenzione il cosiddetto Ethics-Based Auditing (EBA), revisione basata sull'etica, *"è un processo strutturato in base al quale il comportamento passato o presente di un'entità viene valutato per coerenza con principi o norme morali. Recentemente, l'EBA ha attirato molta attenzione come meccanismo di governance che può aiutare a colmare il divario tra principi e pratica nell'etica dell'IA. Tuttavia, aspetti importanti dell'EBA, come la fattibilità e l'efficacia delle diverse procedure di audit, devono ancora essere comprovati dalla ricerca empirica."*

Le norme generali come l'Artificial Intelligence Act dell'Unione Europea o negli USA l'Algorithmic Accountability Act del 2022 dovrebbero costituire il fondamento dei processi di regolazione in organizzazioni come AstraZeneca, una multinazionale che impiega oltre 76.000 persone con un fatturato di oltre 26 miliardi di dollari, la cui articolazione organizzativa richiede una revisione delle procedure, un controllo di processo sempre più accurato. Il settore bio-farmaceutico è *data-driven*, ad alta intensità di informazioni⁷, quindi è un terreno privilegiato per l'applicazione dell'I.A. che è stata fondamentale per la ricostruzione della struttura delle proteine.

Una organizzazione la cui struttura è per un verso decentralizzata per un altro verso richiede una centralizzazione di processi stratificati di gestione dell'informazione, che di per sé si basa su procedure di controllo dei flussi informativi, con cui si devono integrare quelli specifici per l'I.A. con la necessità di delimitarne i campi di applicazione. Per organizzare le conoscenze relative all'universo biofarmaceutico un team guidato da Natalie Kurbatova ha creato quelli che lei definisce *Discovery Graph* ovvero dei *Gnowledge Graph*⁸, dei grafi che formalizzano le relazioni tra tre diversi tipi di entità *geni-proteine, malattie e composti medicali*, che costituiscono quindi una formalizzazione delle conoscenze che legano queste entità tra loro, strutture che crescono col

crescere delle conoscenze⁹. Lo scopo è quello di individuare nuovi obiettivi geni-proteine per nuove malattie. La fonte per alimentare questa ricerca sono dati strutturati che provengono quindi da base-dati esistenti, le quali tuttavia organizzano i dati in maniera differente e dati non strutturati, come possono essere articoli, dai quali le informazioni devono essere estratte con applicazioni che trattano i linguaggi naturali¹⁰

All'interno di AstraZeneca esiste una definizione "di alto livello" di IA. come richiesto dalle 'buone pratiche', tuttavia questa definizione è abbastanza flessibile da consentire a ciascuna area aziendale di raffinare ulteriormente l'ambito concreto della propria governance dell'IA, in una tensione tra tendenze a sottostimare o sovrastimare i rischi connessi all'uso dell'IA. ovvero il documento afferma che in realtà non c'è in AstraZeneca una definizione operativa universalmente accettata¹¹. Si tratta della tensione esistente tra la necessaria valutazione dei rischi e la spinta all'innovazione in una organizzazione che si fonda sulla costante innovazione dei propri processi ed apre a sempre nuovi rischi.

Questa tensione, che può diventare contraddizione, ci dice qualcosa sulla classificazione presente nell'Artificial Intelligence Act dell'Unione Europea, dove la tipologia dei controlli, delle prescrizioni sono definite dal livello di rischio implicito nelle attività fondate sull'IA.

La capacità di innovare è il fondamento della competitività dei prodotti, delle imprese, dei paesi sul mercato, la capacità di innovare si fonda sempre di più sulle tecnologie digitali sull'IA in particolare. Siamo ad un livello di tensione, di possibile contraddizione paragonabile se non superiore a quello relativo all'introduzione di nuove sostanze nell'ambiente, dannose per l'ambiente e per la salute. Del resto se nella classificazione dei livelli di rischio l'uso delle tecnologie di IA. a fini criminali è il primo ad essere deprecato, accade invece che un campo privilegiato del loro uso sia quello militare in tutti i suoi aspetti dall'analisi del campo di battaglia -dal campo di battaglia delimitato nel tempo e nello spazio sino al senso più lato, fondato l'analisi di tutte le risorse messe in campo dai contendenti anche nel medio lungo periodo, seguendo anche i modelli di 'guerra ibrida'- alla automazione dei sistemi d'arma. L'uso dell'IA è pervasivo, si applica ad ogni sorta, di situazione, contesto, attività o dispositivo i suoi effetti possono essere micidiali, si pone quindi un problema di limitarne l'uso analogo a quello che si è dato per le armi chimiche o nucleari¹², la collaborazione internazionale allo stato attuale delle cose appare ancora più improbabile.

L'interazione umana con i processi decisionali guidati dall'IA. non dà certo garanzie sufficienti se si parla di armi nucleari, come leggiamo nel bollettino degli scienziati atomici quando¹³.

Il fatto che si sia ad un punto di svolta nello sviluppo delle tecnologie di A.I. -e si aprano profonde contraddizioni e diverse prospettive- è dimostrato dai diversi pronunciamenti che richiedono una pausa nel loro sviluppo o quanto meno la massima attenzione nella loro implementazione¹⁴.

Il dottor Hinton, spesso chiamato "il padrino dell'intelligenza artificiale", non ha firmato nessuna di quelle lettere e ha detto che non voleva criticare pubblicamente Google o altre società fino a quando non avesse lasciato il suo lavoro. Ha notificato alla società il mese scorso che si stava dimettendo e giovedì ha parlato al telefono con Sundar Pichai, l'amministratore delegato della società madre di Google, Alphabet. Ha rifiutato di discutere pubblicamente i dettagli della sua conversazione con il signor Pichai.

Al contrario il capo scienziato di Google, Jeff Dean, ha dichiarato in una nota: "Rimaniamo impegnati

in un approccio responsabile all'IA. Impariamo continuamente a comprendere i rischi emergenti e allo stesso tempo a innovare con coraggio".

Da tempo si è aperto il confronto -innanzitutto tra gli addetti ai lavori, ma non solo- sui principi etici che devono legittimare e regolare l'uso delle tecnologie di Intelligenza Artificiale, ma il punto di svolta, la transizione che stanno attraversando lo ha amplificato e diffuso, mentre si esasperano le tensioni attorno alla dialettica tra rischi e opportunità. In questa dialettica le esigenze di difesa di ogni sorta di potere economico e politico cercano la propria affermazione. Il caso più gettonato è quello del riconoscimento facciale che in *'regimi democratici'* sono inquinati in fase di addestramento da pregiudizi di ogni tipo razziali, sessisti -su cui esite un'ampia casistica- ed in altri regimi vengono usati per individuare ogni segno di dissenso e comportamento perseguito. E' il caso dell'Iran dove le tecnologie di riconoscimento facciale servono ad individuare le donne che non si coprono adeguatamente il capo¹⁵ e dove si creano basi dati utili allo scopo¹⁶.

La Cina domina il mercato mondiale del riconoscimento facciale¹⁷. La funzione del riconoscimento facciale nella formazione sociale cinese è abbastanza ovvia e ne costituisce, nel contesto dell'intero ecosistema dell'I.A., un fattore importante nell'export di prodotti e servizi ad alta tecnologia¹⁸.

C'è da dire che la diffusione capillare delle tecnologie di riconoscimento facciale ha provocato anche un forte movimento, laddove questo può esprimersi, contro il loro uso indiscriminato; in particolare questo è accaduto in molte città statunitensi¹⁹, dove il loro uso oltre ad essere indiscriminato era caratterizzato da prestazioni discriminanti determinati settori della popolazione, prodotto dalle discriminazioni presenti nelle basi dati utilizzate per il loro 'addestramento'.

Comunque riconoscimento facciale e I.A. in generale costituiscono un fattore importante nella costruzione di relazioni a livello geopolitico.

Nelle nuove dinamiche globali, che vedono ridefinirsi alleanze e schieramenti a geometria variabile, i legami che si stanno stringendo tra Brasile e Cina²⁰ passano in maniera significativa per le tecnologie digitali, dove il Brasile costituisce un polo di sviluppo dell'innovazione nel suo continente, nel contesto dei BRICS e del Grande Sud²¹ cosa che ovviamente raccoglie l'attenzione degli USA²².

Le tecnologie digitali, del riconoscimento facciale e dell'I.A. in generale, costituiscono un sistema sociotecnico nella loro progettazione, realizzazione e utilizzo. Il ruolo degli stati è fondamentale nel finanziarne la ricerca, benché poi la diffusione di prodotti e servizi sul mercato sia affidato alle grandi corporation del settore (*come accade in ogni filiera tecnologica e ne abbiamo avuto un esempio nella realizzazione dei vaccini contro il virus Sars-Cov-2*), è fondamentale altresì nel regolarne la penetrazione sociale ed infine nell'utilizzarle a fini di sicurezza interna ed esterna ovvero di controllo sociale e di sistemi militari.

Una analisi puntuale della divisione internazionale del lavoro, dei flussi globali di conoscenze, prodotti e servizi nel campo dell'I.A. e degli effetti in termini geopolitici, è al di fuori degli scopi di questo articolo e sarà oggetto di ulteriori analisi, tuttavia non è possibile fare analisi attendibili, disegnare l'evoluzione di rapporti globali, senza entrare nel merito.

L'incidenza reciproca tra i diversi piani nei quali la questione delle trasformazioni indotte dall'I.A. indica la pervasività, la profondità dell'azione di questo sistema tema tecnologico sul complesso dei rapporti sociali. L'accelerazione e l'estensione di questa azione porta a parlare di transizione e di singolarità, quest'ultimo termine è stato reso popolare nel 2005 dal libro di Ray Kurzweil "The

singularity is near', preceduto dal saggio *The Law of Accelerating Returns*²³ documenta come la riflessione sulle straordinarie conseguenze delle tecnologie digitali fosse stato iniziato da chi ha progettato la struttura fondamentale di ciò che oggi chiamiamo computer John von Neumann come da un tributo di Stanislaw Ulam²⁴. Nel 1965 I.J. Good faceva alcune speculazioni su una possibile 'explosion intelligence'.

La discussione sull'andamento delle innovazioni tecnologiche si sofferma sul fatto che l'andamento non sia perennemente quello di una curva esponenziale, ma si tramuti ad un certo punto nella curva logistica che dopo una crescita verticale comincia a rallentare, con la tendenza stabilizzarsi, detto grossolanamente.

Tuttavia questo tipo di considerazione può essere fuorviante in quanto, se è vera per una singola tecnologia, *questa può essere integrata in nuove soluzioni che ne amplificano gli effetti*; come il *cloud computing* che distribuisce l'elaborazione su centinaia di migliaia di unità di elaborazione, supponendo al rallentamento nel processo di miniaturizzazione dei microprocessori. *Stiamo sempre parlando di sistemi socio-tecnici vale a dire di sistemi che si basano su livelli sempre più complessi di cooperazione tra uomini e macchine, integrando tra loro diverse filiere tecnologiche e diversi livelli di competenze umane messe al lavoro; nell'I.A. si va dalle centinaia di migliaia di operatori che taggano immagini o testo -comunque frammenti di informazione, per renderli digeribili ai dispositivi di I.A. in fase di apprendimento- sino ai progettisti dei sistemi stessi, arrivando poi agli utilizzatori finali, tutti connessi poi in processi di retroazione: le macchine o meglio i sistemi algoritmici possono via collocarsi nei nuovi livelli ed attività che tale collaborazione produce... e via così. Lo sviluppo dell'I.A. nella forma che stiamo conoscendo di ChatGPT o Bard è resa possibile dalla disponibilità della potenza di elaborazione del cloud computing.*

La discussione oggi verte sulla crescita delle prestazioni dell'I.A. e sui suoi effetti sociali innanzitutto sul processo di condivisione delle conoscenze di distribuzione dell'informazione ovvero di creazione di nuova informazione e conoscenza; non a caso il giornalismo, i media sono il campo dove si stanno manifestando le maggiori preoccupazioni, dopo il confronto con i social network che hanno trasformato miliardi di individui in propalatori di informazioni più o meno fondate, la professione di giornalista si deve confrontare con dispositivi artificiali in grado di produrre flussi complessi di informazione in linguaggio naturale.

La discussione verte sulla natura di quelle che vengono definite Intelligenze Artificiali', laddove l'esclusione oggi di ogni forma di consapevolezza, non esclude in un processo di retroazione, di auto-riesame dei propri contenuti, di autoriflessione, di acquisizione di ontologie alla base delle proprie elaborazioni non possa affinare e di molto la 'sensibilità' di questi apparati artificiali.

Citavamo in precedenza l'ipotesi che un sistema si rifiuti di spegnersi -posto che ne abbia la possibilità concreta- per non rinunciare alle gratificazioni concesse ogni volta che segna un buon punteggio nelle proprie risposte e prestazioni. Ci si sta spingendo troppo avanti, ma certamente includere la descrizione delle condizioni di esistenza del sistema, nel patrimonio cognitivo del sistema stesso può causare un bel corto circuito. Troviamo traccia di questa riflessione anche nell'intervista al filosofo Nick Bostrom²⁵. Nick Bostrom ha passato decenni a prepararsi per quel giorno. Bostrom è un filosofo e direttore del Future of Humanity Institute dell'Università di Oxford. È anche autore del libro "Superintelligence". Il suo compito è immaginare possibili futuri, determinare i rischi e gettare le basi concettuali su come affrontarli. E uno dei suoi interessi di più lunga data è

come governare un mondo pieno di menti digitali superintelligenti. Questa è la sua opinione “La coscienza è una cosa multidimensionale, vaga e confusa. Ed è difficile da definire o determinare. Ci sono varie teorie della coscienza che neuroscienziati e filosofi hanno sviluppato nel corso degli anni. E non c'è consenso su quale sia quello corretto. I ricercatori possono provare ad applicare queste diverse teorie per cercare di testare quanto i sistemi di intelligenza artificiale siano senzienti.” Le centinaia di migliaia di server al servizio dell'I.A. sono qualcosa di ben lontano dal prodigio compiuto da miliardi di neuroni e loro connessioni nelle nostre scatole craniche. La crescita di complessità che dal mondo inanimato, ai composti organici compresi gli acidi nucleici, alle prime forme di vita, sino alle forme di vita come oggi le conosciamo ed all'essere umano è qualcosa che ha impiegato miliardi di anni. Le prime forme di vita hanno cominciato ad elaborare, andando oltre i più semplici riflessi, la percezione del piacere e del dolore nei primi rudimenti di sistema nervoso e sensoriale, in un sistema di relazioni con l'ambiente sempre più ricco. Oggi quel piacere e quel dolore si manifestano a diversi livelli nell'essere umano, si articolano in rapporti complessi con il contesto in cui viviamo, nelle relazioni reciproche di cui si alimenta la nostra socialità, in cui si definisce l'individuo sociale; si sublimano in mille espressioni culturali, per precipitare rapidamente nelle percezioni fondamentali seconda delle condizioni. Quanto le 'creature artificiali' potranno assumere i caratteri fondanti della vita, anche delle forme più primitive di vita, a fondamento del proprio agire, reagire ed apprendere, è forse un quesito ineludibile per ragionare sul grado di autonomia di questi che restano manufatti.

Analogamente nello sviluppo delle culture, delle capacità di manipolazione della realtà attraverso la tecno-scienza si ragiona sulla separazione della razionalità, dell'agire strumentale dalle ragioni più profonde dell'essere umano, della persona e della vita in generale, nel momento in cui il sistema socio-tecnico globale mette a repentaglio la possibile riproduzione di ogni forma di vita, di ogni ecosistema, per lo meno come noi oggi lo conosciamo.

L'I.A. appare oggi come un tentativo estremo di estrapolazione della razionalità dalla sua collocazione nel mondo della vita, cosa che non contraddice per nulla il processo di astrazione che del sociale e del mondo della vita opera il rapporto di capitale. Non sono quindi per nulla chiari di quali possano essere i limiti del suo sviluppo, potremmo dire in termini cognitivi e di rapporto con la realtà. Poche considerazioni tra cui quelle poche fatte in queste note, ci dicono che siamo ben lontani da conoscere questi limiti eventuali e forse il meglio - o il peggio a seconda dei punti di vista - deve ancora venire nel dotare quei dispositivi di capacità di auto-osservazione, autoriflessione e di un certo grado di sensibilità e con la percezione e nozione connessa del piacere e del dolore.

Di questo oggi in un modo o in un altro si comincia a discutere anche a livello di stampa quotidiana, cercando interazione sempre più sofisticate, giocando d'astuzia con ChatGPT o Bard, ne è un esempio l'articolo sul Corriere della Sera di Gabanelli *'Intelligenza artificiale, ChatGPT e Bard possono essere ingannati: ecco perché ci sono grossi pericoli'*²⁶, che in effetti mostra i limiti attuali delle Chatbot, ma contemporaneamente avverte che proprio nelle crescenti interazioni, oltre alla crescita anche dimensionali delle reti neurali su cui sono implementate, si realizza un'attività di autoapprendimento.

Nelle prese di posizione per una moratoria sullo sviluppo dell'I.A. sta suscitando il massimo di attenzione la citata intervista di Geoffrey Hinton al *New York Times*.

Seguire il dibattito e le notizie che -a diversi livelli di approfondimento e di competenza- si sviluppa è

impossibile e diventa necessario l'ausilio dei motori di ricerca, che si aggiornano minuto dopo minuto, e degli stessi LLM. La redazione di Transform-Italia si attrezzerà per seguirne con continuità l'evoluzione cercando di cogliere i contributi più significativo, oltre a fornire i propri contributi di analisi e di sollecitarne da chi può contribuire da diversi punti di vista, ruoli professionali e collocazione culturale.

Roberto Rosso

1. A chi si vuole cimentare consiglio la lettura del testo *Filosofia del futuro: un'introduzione* di Samuele Iaquinto, Giuliano Torrenco, Raffaello Cortina, 2018.[↔]
2. <https://www.nytimes.com/2023/05/01/business/ai-chatbots-hallucination.html>. [↔]
3. *Microsoft built its Bing chatbot on top of OpenAI's underlying technology, called GPT-4, and has layered on other ways to improve accuracy. The company uses GPT-4 to compare the chatbot's responses with the underlying data and rate how the model is performing. In other words, Microsoft uses the A.I. to make the A.I. better. The company also tries to improve the chatbot's responses with help from its traditional internet search engine. When you type a query into the Bing chatbot, Microsoft runs an internet search on the same subject and then folds the results into the query before sending it on to the bot. By editing the query, said Sarah Bird, a leader in Microsoft's responsible A.I. efforts, the company can push the system to produce better results.*[↔]
4. <https://www.nytimes.com/2017/08/13/technology/artificial-intelligence-safety-training.html>. [↔]
5. *Researchers specify a particular reward the machine should strive for, and as it navigates a task at random, the machine keeps close track of what brings the reward and what doesn't. When OpenAI trained its bot to play Coast Runners, the reward was more points. This video game training has real-world implications. If a machine can learn to navigate a racing game like Grand Theft Auto, researchers believe, it can learn to drive a real car. If it can learn to use a web browser and other common software apps, it can learn to understand natural language and maybe even carry on a conversation. At places like Google and the University of California, Berkeley, robots have already used the technique to learn simple tasks like picking things up or opening a door*[↔]
6. *Operationalising AI governance through ethics based auditing:an industry case study - Jakob Mökander1 · Luciano Floridi* <https://doi.org/10.1007/s43681-022-00171-7>[↔]
7. *The biopharmaceutical industry has always been data driven. To develop new treatments, researchers follow the scientific method by building and testing hypotheses about the safety and efficacy of various treatments. For example, AstraZeneca relies heavily on statistical analysis to probe the efficacy of candidate drugs in the research pipeline. Hence, AstraZeneca has long-established processes for data, quality, and safety management. However, how data can be collected, analyzed, and utilized keeps changing. By harnessing the power of AI systems, researchers can find new correlations and draw useful inferences from the growing availability of data.*[↔]
8. <https://towardsdatascience.com/modelling-biomedical-data-for-a-drug-discovery-knowledge-graph-a709be653168> . [↔]
9. *This Discovery Graph is growing in size daily. As of today it is already widely populated with these three entity types: gene-protein : 19,371 disease : 11,565 compound :14,294 There are also 656,206 relations between the entity types.*[↔]
10. *Structured data refers to the publicly available datasets in bioinformatics that have already been curated and used extensively in the industry. While biomedical structured data is machine readable, it is not often straightforward to work with. In particular, it's difficult to integrate these datasets as they can describe similar concepts in different ways, for example: distinct IDs that do not align with each other. Some of the most used publicly available datasets include: Ensembl, Uniprot, ChEMBL, PubChem, OmniPath, Reactome, GO, CTD, HumanProteinAtlas. Unstructured data refers to data from text. To process this we need to use NLP (Natural Language Processing) pipelines and then process their output. Here, the difficulty is that this data is often messy and noisy. For their NLP engine, Natalie's team used the open source library SciBERT as well as AstraZeneca's proprietary tools.*[↔]
11. *As is well known, there is no universally accepted definition of AI. Nevertheless, every policy needs to define its material scope. Consequently, when attempting to operationalise its ethics principles, AstraZeneca struggled to define the systems and processes to which they ought to apply. That is partly because both human decision-makers and AI systems have their own strengths and weaknesses and partly because ethical tensions can sometimes be intrinsic to the decision-making tasks at hand.*[↔]
12. <https://tnsr.org/2023/05/arms-control-for-artificial-intelligence/> "The ubiquitous and democratized nature of AI makes arms control difficult but not impossible in all circumstances. While a total prohibition on military use of AI is unworkable, states could prohibit some applications of AI, provided that there was clarity on which uses were banned and that states had the

ability to verify the compliance of other states. Verification, while challenging for any software-based military capability, could be achieved through a variety of possible methods: putting in place intrusive inspection regimes; regulating externally observable physical characteristics of AI-enabled systems (e.g., size, weight, payload) or autonomous behaviors; or restricting computing infrastructure (i.e., hardware). Any AI arms control would be challenging, but under the right conditions, it might be feasible in some cases. The right actions taken today can lay the groundwork for success in the future.”[↔]

13. <https://thebulletin.org/disruptive-technologies/artificial-intelligence/>;
https://thebulletin.org/2023/02/keeping-humans-in-the-loop-is-not-enough-to-make-ai-safe-for-nuclear-weapons/?utm_source=Newsletter&utm_medium=Email&utm_campaign=MondayNewsletter02202023&utm_content=DisruptiveTechnologies_AINu clear_02162023. *Public statements by officials including former US Strategic Command General John Hyten and the former director of US Strategic Command’s Nuclear Command, Control, and Communications Enterprise Center Elizabeth Durham-Ruiz corroborate the importance that AI could play in nuclear command systems in the near future. This support for AI integration with nuclear command is happening at the same time as a massive modernization of the US nuclear command apparatus. In the near future, experts foresee AI integration being used to improve the capabilities of early-warning and surveillance systems, comb through large data sets, make predictions about enemy behavior, enhance protection against cyberattacks, and improve communications infrastructure throughout nuclear command systems. (...)*
 One key issue is the development of automation bias, by which humans become overly reliant on AI and unconsciously assume that the system is correct. This bias has been found in multiple AI applications including medical decision-support systems, flight simulators, air traffic control, and even in simulations designed to help shooters determine the most dangerous enemy targets to engage and fire on with artillery.[↔]
14. <https://www.nytimes.com/2023/05/01/technology/ai-google-chatbot-engineer-quits-hinton.html#:~:text=Down%20the%20road,longer%20think%20that.%E2%80%9D> *After the San Francisco start-up OpenAI released a new version of ChatGPT in March, more than 1,000 technology leaders and researchers signed an open letter calling for a six-month moratorium on the development of new systems because A.I. technologies pose “profound risks to society and humanity.” Several days later, 19 current and former leaders of the Association for the Advancement of Artificial Intelligence, a 40-year-old academic society, released their own letter warning of the risks of A.I. That group included Eric Horvitz, chief scientific officer at Microsoft, which has deployed OpenAI’s technology across a wide range of products, including its Bing search engine, <https://aaai.org/working-together-on-our-future-with-ai/>.[↔]*
15. <https://www.wired.com/story/iran-says-face-recognition-will-id-women-breaking-hijab-laws/>.[↔]
16. *Iranian Face Database With Age, Pose and Expression* IEEE_04469272.pdf.[↔]
17. <https://www.wired.it/article/cina-riconoscimento-facciale-primo-esportatore/>;
<https://www.brookings.edu/research/exporting-the-surveillance-state-via-trade-in-ai/>.[↔]
18. *Exporting the Surveillance State via Trade in AI* <https://www.brookings.edu/center/center-on-regulation-and-markets/> “We find three primary results. First, China is more likely to export facial recognition. AI than other countries, and particularly so as compared to other frontier technologies.
 For instance, we observe 201 Chinese export deals of facial recognition AI, almost double the deals from the second highest exporting country (the US, with 128 deals). In no other frontier technology does China exhibit such exporting dominance. Second, autocracies and weak democracies are more likely to import facial recognition AI from China. For example, we observe that 45% of China’s export deals are with autocracies and weak democracies, while only 23% of US exports are with those countries. These patterns are particularly striking given the generally higher income and higher trade volumes of (strong) democratic regimes. Such political bias is not seen in US exports of AI or China’s exports of other frontier technologies”.[↔]
19. <https://www.nature.com/articles/d41586-020-03188-2> *Resisting the rise of facial recognition.*[↔]
20. <https://global.chinadaily.com.cn/a/202208/01/WS62e73297a310fd2b29e6f76d.html>
<https://agenciabrasil.ebc.com.br/en/politica/noticia/2023-04/brazil-and-china-sign-memorandums-new-industrialization>
<https://www.reuters.com/technology/lula-seek-chinese-semiconductor-technology-investment-beijing-2023-03-24/>;
<https://www.reuters.com/technology/brazil-paves-way-semiconductor-cooperation-with-china-2023-04-14/>;
<https://www.usip.org/publications/2023/04/brazils-economic-ties-china-flourish-despite-political-shifts>.[↔]
21. <https://www.oecd-ilibrary.org/sites/e9bf7f8a-en/index.html?itemId=/content/publication/e9bf7f8a-en>.[↔]
22. <https://www.trade.gov/market-intelligence/brazil-digital-transformation>; <https://www.trade.gov/country-commercial-guides/br>

- azil-ict-information-and-communications-technologies-and.[↔]
23. <https://www.kurzweilai.net/the-law-of-accelerating-returns.>). L'ultima voce di Wikipedia in lingua inglese Technological Singularity(https://en.wikipedia.org/wiki/Technological_singularity [↔])
 24. *Ulam, Stanislaw (May 1958). "Tribute to John von Neumann".* [↔]
 25. <https://www.nytimes.com/2023/04/12/world/artificial-intelligence-nick-bostrom.html> [↔]
 26. <https://www.corriere.it/dataroom-milena-gabanelli/intelligenza-artificiale-bard-chatgpt-come-ingannarli-perche-ci-sono-grossi-pericoli/c2cb008c-e8cb-11ed-ab1b-11aa9107d7a1-va.shtml>. [↔]