

Corso di laurea in Governo, Amministrazione e Politica con
Major in Politica e Comunicazione

Cattedra Metodi e Tecniche del Giornalismo Economico-Politico

Il futuro dell'informazione: l'utilizzo dell'intelligenza
artificiale nel giornalismo

Prof. Fabio Carducci Artenisio

RELATORE

Prof. Cristina Alaimo

CORRELATORE

Alessio Barca

Matricola 646082

CANDIDATO

Indice

Introduzione	3
CAPITOLO 1 Concetti, definizioni e storia dell'intelligenza artificiale nel giornalismo	4
1.1 <i>Concetti e definizioni dell'intelligenza artificiale</i>	4
1.1.1 Concetti e definizioni generali.....	4
1.1.2 Tipologie di Intelligenza Artificiale.....	7
1.2 <i>La nascita dell'Intelligenza Artificiale come campo scientifico</i>	10
1.2.1 Dall'antica Grecia alla Conferenza di Dartmouth.....	10
1.2.2 La Conferenza di Dartmouth e gli sviluppi successivi.....	13
1.3 <i>L'Intelligenza Artificiale nel giornalismo</i>	15
1.3.1 I termini appropriati.....	16
1.3.2 Tipologie di Intelligenze Artificiali utilizzate nel giornalismo	17
1.4 <i>Storia delle Intelligenza Artificiale nel giornalismo</i>	19
1.4.1 Articoli automatici di previsioni meteo	19
1.4.2 Stats Monkey	21
CAPITOLO 2 L' intelligenza artificiale nel giornalismo: utilizzi ed esempi virtuosi	25
2.1 <i>Le intelligenze artificiali per il news gathering</i>	26
2.1.1 La Computer Vision	26
2.1.1.1 L'utilizzo della Computer Vision al New York Times	31
2.1.2 Ricerca di informazioni e previsione di tendenze	33
2.1.2.1 Asimov	34
2.2 <i>Le intelligenze artificiali per la news production</i>	35
2.2.1 Software audio e video	35
2.2.1.1 AI Anchor del Sole24Ore	36
2.2.2 La scrittura automatica	37
2.2.2.1 GPT-3	38
2.2.3 Deep Fake	40
2.2.3.1 L'Anchor Man digitale di Xinhua.....	41
2.3 <i>Le intelligenze artificiali per la News Distribution</i>	43
2.3.1 Gli Alert e la personalizzazione dei contenuti.....	44
2.3.1.1 Voitto: l'assistente per le notizie che propone alternative	46

CAPITOLO 3 Legislazione e questione etiche	47
3.1 <i>La legislazione italiana sul giornalismo</i>	50
3.1.1 Dal cartaceo al digitale	50
3.1.2 Dal digitale alle intelligenze artificiali	51
3.1.3 La responsabilità dell'operato dell'IA	54
3.2 <i>Questioni etiche</i>	56
3.2.1 La trasparenza	56
3.2.1.1 La conoscibilità dei dati	56
3.2.1.2 La trasparenza dell'addestramento	58
3.2.2 L'impatto ambientale delle intelligenze artificiali	60
CAPITOLO 4 Il cambiamento della professione giornalistica e la sua percezione	62
4.1 <i>Deontologia</i>	62
4.1.1 Le fake news	63
4.1.1.1 La psicologia delle fake news	63
4.1.1.2 Scienza e fake news	65
4.1.1.3 Il regime giuridico delle fake news	66
4.1.2 Le intelligenze artificiali per le fake news	67
4.2 <i>L'impatto delle intelligenze artificiali sul settore giornalistico</i>	73
4.2.1 L'impatto occupazionale delle intelligenze artificiali nelle redazioni	74
4.2.1.1 La necessità di un nuovo ruolo per il giornalista	77
4.2.2 La sostenibilità economica delle redazioni	78
4.3 <i>La percezione delle IA nel giornalismo</i>	81
4.3.1 La percezione della personalizzazione delle notizie	83
Conclusione	85
Bibliografia	86
Sitografia	92
Riassunto	96

Introduzione

“Il futuro dell'informazione: l'utilizzo dell'intelligenza artificiale nel giornalismo” è una tesi magistrale che vuole indagare gli scenari attuali del mondo del giornalismo, analizzando nello specifico lo stato attuale delle intelligenze artificiali in questo ambito e gli aspetti ad esse correlati, proiettandosi verso il futuro.

Ho scelto questo argomento per via del mio spiccato interesse per l'informatica e per il mondo dell'informazione. Nell'esperienza da rappresentante di istituto al liceo, la mansione che mi piaceva svolgere di più era quella di comunicare, sia fisicamente che digitalmente, con studenti e professori, modificando la forma e il registro linguistico a seconda dei casi. Già la mia tesi triennale ha riguardato argomenti simili, nello specifico i criteri di usabilità per la realizzazione di siti web. Mi ha sempre affascinato l'analisi della percezione degli utenti, con l'intento di perseguire allo stesso tempo interessi economici e utilità per il proprio pubblico di riferimento. Nella scrittura della mia tesi triennale l'obiettivo era quello di dimostrare che perseguire l'usabilità, cioè l'adattamento della struttura di un sito web alle esigenze dell'utente in modo da rendere più fluida possibile la sua navigazione all'interno di esso, andava di pari passo con gli obiettivi di engagement che inevitabilmente un sito si pone. Altro obiettivo era quello di dimostrare che il rispetto dei criteri stabiliti dagli algoritmi di Google, più che essere dei vincoli, sono delle opportunità, in quanto impostati in modo da rispettare sia le leggi che le esigenze degli utenti.

In questo lavoro ho cercato di mantenere lo stesso tipo di approccio. L'obiettivo è stato quello di capire come poter sfruttare le tecnologie per la ripresa economica del settore giornalistico, provando contemporaneamente ad offrire contenuti innovativi e di qualità ai lettori. Se non per le nozioni di programmazione, era a me sconosciuto il mondo delle intelligenze artificiali.

La prima parte dell'analisi indagherà, infatti, i concetti e le terminologie principali delle intelligenze artificiali, sia in generale che nello specifico, per il mondo del giornalismo, con approfondimenti storici che ci faranno capire come si è evoluto l'ambito. Cercando poi di rendere il più comprensibili possibile concetti tecnici, si analizzeranno le tipologie di software utilizzate nel giornalismo, riportando esempi virtuosi.

Tutta questa analisi non potrà prescindere dalle leggi e dai principi etici di un rapporto, quello uomo-macchina, che è sempre stato controverso e che oggi assume un'accezione quasi conflittuale. Responsabilità, trasparenza e impatto ambientale sono i principali temi da porre per un utilizzo consapevole e appropriato delle tecnologie.

Infine, siamo andati al nocciolo della questione iniziale: cosa può fare oggi un giornalista per la democrazia? Le intelligenze artificiali gli ruberanno il lavoro o possono essere uno strumento veramente utile alla causa? Come percepiscono i lettori i contenuti di un'intelligenza artificiale o di un articolo scritto da un giornalista in “collaborazione” con un IA?

Si è metodologicamente usufruito sia di paper scientifici che di articoli sul web, da una parte per indagare al meglio e in maniera approfondita le questioni poste, dall'altra per essere aggiornati con i più recenti sviluppi e comprendere il tono della discussione mediatica sull'argomento.

CAPITOLO 1 Concetti, definizioni e storia dell'intelligenza artificiale nel giornalismo

Cos'è l'Intelligenza Artificiale? Quand'è nata? Come la intendiamo oggi? Quando e come è entrata a far parte delle redazioni?

Sono questi i quesiti cui tenteremo di rispondere in questo capitolo.

1.1 Concetti e definizioni dell'intelligenza artificiale

1.1.1 Concetti e definizioni generali

Il termine Intelligenza Artificiale è generalmente inteso come “l'abilità di una macchina di mostrare capacità umane quali il ragionamento, l'apprendimento, la pianificazione e la creatività”¹. Il vocabolo è diventato d'uso comune negli ultimi decenni ed è stato coniato come vero e proprio termine indicante un ambito scientifico solo nel 1956, durante la Conferenza di Dartmouth².

La moderna concezione di Intelligenza Artificiale, però, è più complessa di quanto si possa pensare. In informatica si intende un “insieme di studi e tecniche che tendono alla realizzazione di macchine, specialmente calcolatori elettronici, in grado di risolvere problemi e di riprodurre attività proprie dell'intelligenza umana”³.

Il fatto che riproducano e mostrino capacità simili e non identiche all'uomo rende tale concetto più complesso: si può parlare di intelligenza artificiale forte, cioè che riproduce in tutto e per tutto azioni tipiche dell'uomo, e intelligenza artificiale debole, cioè un'intelligenza che riproduce solo in parte compiti umani e che tipicamente viene affiancata da un vero e proprio essere umano nello svolgimento della propria funzione. È facilmente intuibile che, se non per esempi cinematografici o di fantascienza, l'intelligenza artificiale forte non esista, o quantomeno non ancora. Non è però solo su questa differenza che la nostra definizione diviene più complessa. Talvolta questo termine viene accostato all'idea di robot ma è in realtà diversa da essa. Il termine robot ha un'accezione più inerente allo spazio fisico e azioni pratiche all'interno di esso, mentre l'intelligenza artificiale, pur essendo usato come termine anche nell'ambito della robotica, riguarda per lo più sistemi e programmi capaci di assolvere azioni accostabili a processi mentali. Non sono però solo il termine robot e l'ambito della robotica in generale a essere confusi con l'intelligenza artificiale. Sicuramente avrete sentito parlare di *Machine Learning*, *Deep Learning* e *Natural Language Processing*. Spesso questi termini vengono scambiati, inglobati, sovrapposti tra loro, talvolta in maniera opportuna, quasi sempre in maniera indebita.

¹Parlamento Europeo, *Che cos'è l'intelligenza artificiale e come viene usata*, Sezione Attualità, 29 marzo 2021

²Il Dartmouth Summer Research Project on Artificial intelligence, anche conosciuto come “Conferenza di Dartmouth” è il primo convegno di studiosi e ricercatori sull'Intelligenza Artificiale, comunemente riconosciuto come il punto di partenza dell'Intelligenza Artificiale in quanto vero e proprio campo scientifico (si approfondirà ulteriormente più avanti)

³T. De Mauro, *Grande dizionario italiano dell'uso*, Torino, 2000

L'ostacolo da tenere bene a mente è che anche se esistono suddivisioni e concettualizzazioni condivise, non lo sono mai in tutto e per tutto e risulta impossibile incasellare, categorizzare e gerarchizzare in maniera precisa tali concetti per tre ordini di motivi:

- 1) Le tecnologie sono spesso convergenti tra loro e assolvono più funzioni insieme e quindi anche i termini che le identificano sono spesso sovrapponibili di per sé (si pensi al concetto di mobile device, smartphone o cellulare)
- 2) Definizioni tecniche devono spesso fronteggiare il problema della fruibilità e comprensione da parte di tutti e spesso nelle semplificazioni subiscono deformazioni
- 3) L'ambito preso in esame è ancora in fieri e quindi anche i termini utilizzati possono essere soggetti a fraintendimenti e cambiamenti

Nonostante ciò, risulta comunque utile cercare di esemplificare il più possibile i termini in questione, oltre che a darne una gerarchia indicativa e fornire più elementi possibili al lettore al fine di una comprensione a 360° degli ambiti dell'Intelligenza Artificiale: una sorta di glossario approfondito e contestualizzato.

Il primo passo per una concettualizzazione consona e utile alla nostra causa è sapere che il termine "Intelligenza Artificiale" è un termine ombrello, omnicomprensivo e comunemente usato per racchiudere tutti gli altri. La sigla abbreviata IA in italiano sta ad indicare "Intelligenza Artificiale" mentre in inglese "*Intelligent Agent*" cioè "Agente Intelligente", che ha un'accezione più specifica⁴. La sigla AI invece indica il contrario: "Agente Intelligente" in italiano e "*Artificial Intelligence*" in inglese. Ovviamente queste sigle sono così interpretate nell'ambito informatico, mediatico e della robotica, ma in altri ambiti potrebbero avere diverse traduzioni, come Inseminazione Artificiale in ambito medico (*Artificial Insemination* in inglese) o Amnesty International in ambito sociopolitico. Usate però in un determinato contesto sono meno soggette a misinterpretazioni e verrà quindi usata in questa sede la sigla IA. Verranno inserite tra parentesi le sigle solitamente utilizzate anche per altri termini.

Il *Machine Learning* (ML) è una branca dell'Intelligenza Artificiale. Il termine è utilizzato per identificare quelle tecnologie che auto-apprendono: in base a come sono state impostate sono in grado, di volta in volta, di generalizzare in maniera sempre più precisa, senza bisogno di modificare gli input. La definizione forse più accurata che possiamo dare di questo concetto è quella del vincitore del Turing Award nonché padre del *deep learning* Yoshua Bengio: << L'obiettivo è fornire conoscenza ai computer attraverso dati, osservazioni e interazioni con il mondo e attraverso questa conoscenza far sì che i computer possano generalizzare verso nuove impostazioni >>⁵.

⁴È definito agente intelligente qualunque strumento riesca a percepire l'ambiente esterno e prenda decisioni autonome (anche un termostato è un Agente Intelligente). Nonostante possa essere comunque considerato come un'IA si pone a metà tra essa e il concetto di robot

⁵Daniel Fagella, *What is Machine Learning?*, Emerj The AI Research and Advisory Company, 26 febbraio 2020

Artificial Neural Network (ANN) sta ad indicare le reti neurali, sorta di reti neuronali artificiali, che sono una sottobranca del *Machine Learning*, con una complessità maggiore rispetto a software o algoritmi elementari di auto-apprendimento, ma non così sviluppati come il *Deep Learning*.

Il *Deep Learning* è anch'esso quindi una sotto branca del *Machine Learning*, forse la più innovativa e quella con più prospettive future, che sfrutta reti neurali capaci di apprendere senza supervisione, cambiamenti di input e senza la necessità di dati strutturati.

Queste sono forse le quattro aree più importanti dell'IA, comprendenti a loro volta molte pratiche, tipi di software e algoritmi. Lo schema esemplificativo di tutto ciò che comprende l'Intelligenza Artificiale è raffigurato nella figura 1. Le differenze tra *Machine Learning*, *Artificial Neural Network* e *Deep Learning* sono raffigurate nelle figure 2 e 3 sottostanti.

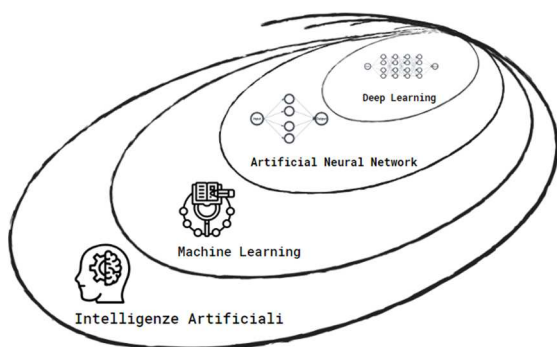


Fig. 1 Categorizzazione Ambiti Intelligenze Artificiali

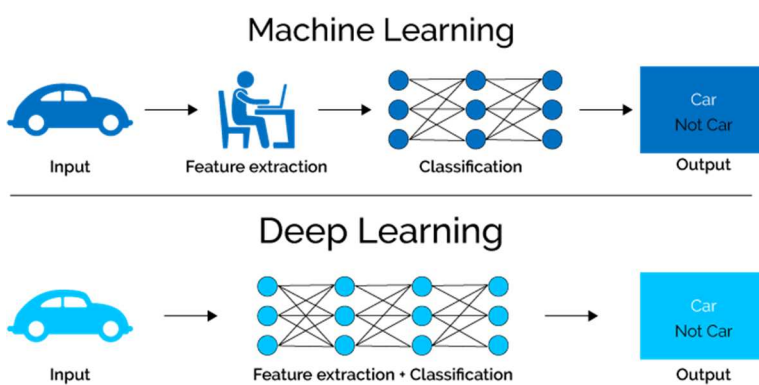


Fig. 2 Caratteristiche e Differenze tra *Machine Learning* e *Deep Learning*
(Credits: Informatica e Ingegneria Online, Vito La Vecchia)

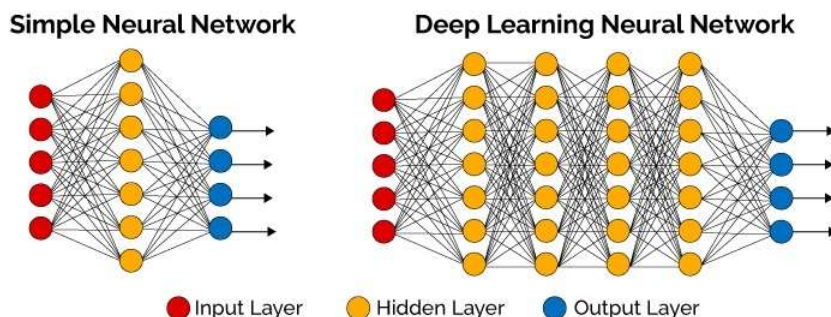


Fig.3 Differenza sostanziale tra *Artificial Neural Network* e *Deep Learning*
(Credits: <https://goo.gl/b84z5Q>)

1.1.2 Tipologie di Intelligenza Artificiale

La differenza tra software e algoritmo è molto banale: un software (o un programma) può essere definito in maniera esemplificativa come un insieme di algoritmi. Un algoritmo, in informatica, è “una sequenza finita di operazioni elementari, eseguibili facilmente da un elaboratore che, a partire da un insieme di dati (input) produce un altro insieme di dati (output) che soddisfano un preassegnato insieme di requisiti”⁶. La maggior parte delle tecnologie che menzioneremo o analizzeremo saranno software, ma non mancano anche esempi di singoli algoritmi che vengono utilizzati per determinati scopi. Gli algoritmi si dividono poi in tre categorie principali: ordinamento, classificazione e associazione.

Tutto ciò che rientra sotto “l’ombrello” dell’IA, ma non nel *Machine Learning*, viene categorizzato come “*reasoning system*” (sistema ragionante) che si differenzia dai “*learning systems*” che, come detto, sono sistemi di autoapprendimento. Questa categoria viene definita anche come “IA classica” e comprende tutto ciò che è programmato, preimpostato, e non necessita di modifiche ulteriori. Sono esempi di queste tecnologie:

- Logica fuzzy (anche detta logica sfocata): comprende tutte quelle tecnologie che utilizzano variabili non binarie ma graduali, cioè che oltre a poter assumere valori 0 e 1 possono assumere valori intermedi tra di essi
- Sistemi esperti: comprende tutte quelle tecnologie che simulano il *decision-making* di un umano per lo più attraverso blocchi di conoscenza impostati con strutture simili a “se succede questo faccio questo” combinati tra loro
- Motori di ricerca: sono tutti quei sistemi di ricerca automatici che attraverso determinati criteri assegnano punteggi a determinate categorie di contenuti (come, ad esempio, le pagine web) dando come risultato una sorta di classifica di tali contenuti. Si basano sulla ricerca semantica
- Riduzione dimensionale: comprende tutte quelle tecnologie che permettono di ridurre file o dati a strutture esigue mantenendo le proprietà principali. Risulta utile quando si devono analizzare grosse quantità di dati o dati complessi non tracciabili computazionalmente
- Programmi logici induttivi (ILP): comprende tutti quei programmi che formulano ipotesi sulla base di un set di dati, esempi positivi ed esempi negativi
- Agenti Intelligenti (AI): tutte quelle strumentazioni che percepiscono l’ambiente circostante e possono prendere decisioni al fine di raggiungere un determinato obiettivo
- Ragionamento basato su istanze: comprende tutte quelle tecnologie che partendo da un set di frasi preimpostate riescono a svolgere funzioni come risposta o contestualizzazione

⁶Enciclopedia Treccani online, *Definizione di algoritmo*

Il *Natural Language Processing* (NLP) e i *Recommender Systems* sono sistemi che potrebbero anche rientrare nelle IA classiche, ma che in realtà sono usualmente settati per auto apprendere e che rientrano quindi nel *Machine Learning*. È un classico esempio di NLP l'assistente automatico online, cioè una chatbot online che spesso si trova sui siti e-commerce e che risponde ai quesiti degli utenti. Esso, quando rientra nelle tecnologie ML, impara ad ogni conversazione per poter essere sempre più funzionale ai dubbi degli utenti e sembrare sempre più simile a loro. Un classico esempio di *Recommender Systems* è quello che consiglia i film o le serie TV sulle piattaforme online o gli algoritmi di Spotify, Youtube e dei Social Network in generale che ci consigliano, sulla base dei nostri gusti, contenuti che potrebbero interessarci.

Rientrano nel *Machine Learning*:

- Alberi decisionali: comprendono tutte quelle tecnologie che si basano su strutture ad albero che valutando decisioni consequenziali prevedono la strategia migliore
- *Hidden Markov Models* (HMM): comprende tutte quelle tecnologie che si basano su un processo in cui il risultato (*outcome*) è condizionato da una variabile “nascosta” (*hidden*) di cui si conoscono solo le modalità con cui essa influenza il risultato. Sono utilizzate in bioinformatica, economia, finanza e numerosi altri campi
- Inferenza Bayesiana: comprende tutte quelle tecnologie che si basano su modelli di inferenza statistica che calcolano la probabilità sulla base di due variabili principali: la variabile antecedente e la “funzione di verosomiglianza” (modello statistico derivato dai dati osservati)
- Algoritmi di Regressione: comprende tutte quelle tecnologie che utilizzano la regressione come modello
- *Support Vector Machine*: comprende tutte quelle categorie che si basano su modelli derivanti da un set di esempi, suddivisi in due categorie con possibili risultati intermedi tra le due
- Algoritmi evolutivi: comprende tutte quelle categorie basate sull'evoluzione di una popolazione predetta attraverso una sorta di funzione dell'evoluzione biologica (a sua volta basata su meccanismi di riproduzione, selezione, ricombinazione e selezione)

Fanno sempre parte di questa categoria, ma non hanno bisogno di supervisione:

- Analisi dei componenti principali: comprende tutte quelle tecnologie che si basano su tecniche di analisi di grandi moli di dati, per riuscire a destrutturarli e analizzarli al meglio. Rientra nelle tecnologie di riduzione dimensionale ma è ovviamente più specifica e complessa
- *Clustering*: comprende tutte quelle tecnologie che utilizzano tecniche di raggruppamento. Sarebbe una categoria generale ma ultimamente vengono utilizzate tecnologie molto complesse di clustering soprattutto sui Big Data.
- Regole associazionali: comprende tutte quelle tecnologie che identificano relazioni interessanti all'interno di grandi moli di dati in maniera automatica

Nelle tecnologie di *Artificial Neural Network*, quasi sfocianti nella sottocategoria del *Deep Learning*, rientrano:

- *Recurrent Neural Network* (RNN): comprende tutte quelle tecnologie in cui gli output influenzano a loro volta gli input e riformulano il modello, migliorandolo.
- *Convolutional Neural Network* (CNN): comprende tutte quelle tecnologie che utilizzano strati di perceptron che riescono ad analizzare elementi visuali come immagini e video. Sono spesso utilizzate per analisi, *clustering* e selezione di immagini.

Nelle tecnologie di *Deep Learning* vero e proprio rientrano:

- *Autoencoder*: comprende tutte quelle tecnologie che modificano i propri input in maniera autonoma per essere più efficienti e risparmiare sull'utilizzo di dati trascurabili. Sono tecnologie diverse dagli RNN perché riformulano la formattazione e i codici del software/ algoritmo nello specifico
- *Multilayer perceptrons* (MLP): sono ANN con almeno tre strati di *hidden layers* e i quali neuroni, eccetto per quelli di input, usano funzioni non lineari di attivazione (simili alla seconda rappresentazione in figura 3)

Le categorie presentate sono molte e come detto molte di esse comprendono le altre, si intersecano tra di loro o variano per piccolissime differenze tecniche. Sono comunque utili alla comprensione dell'ambito scientifico delle IA, anche se presentate in linea generale e suddivise secondo criteri di specificità e complessità⁷ (altri ricercatori le suddividono secondo criteri diversi). Nel momento in cui analizzeremo le Intelligenze Artificiali nel mondo del giornalismo sapremo indicativamente di quali tipi di tecnologie stiamo parlando e quali sono le differenze principali e sostanziali tra un tipo di tecnologia e un altro.

Tralasciando gli aspetti tecnici, possiamo suddividere le IA in tre macrocomponenti⁸:

- *Interfaccia con l'utente*: sono le modalità di comunicazione e trasmissione all'utente delle informazioni elaborate, in un formato comprensibile all'utente
- *Informazioni di base*: è la componente che racchiude tutti i fatti e le informazioni potenzialmente utili su un esatto ordine di problemi. Li rende anche accessibili e fruibili da parte dell'interfaccia motore
- *Interfaccia motore*: il motore di interfaccia è il programma che individua le informazioni adeguate nella base di informazioni, e inferisce nuove informazioni applicando strategie di elaborazione e analisi coerenti.

Per andare ancora più a fondo e capire a pieno la natura delle Intelligenze Artificiali è necessario tornare indietro di millenni, andando a capire cosa ha motivato la creazione di strumenti di questo genere e come siamo arrivati all'odierno *status quo*.

⁷Più si va a fondo in una sottocategoria più la nostra categoria ha compreso tecnologie più complesse

⁸Oliver Jeffery, *Explain Artificial Intelligence an history of Artificial intelligence*, University of Alberta, novembre 2022

1.2 La nascita dell'Intelligenza Artificiale come campo scientifico

1.2.1 Dall'antica Grecia alla Conferenza di Dartmouth

Il termine Intelligenza Artificiale, come detto, è diventato d'uso comune negli ultimi decenni ed è stato coniato come vero e proprio termine indicante un ambito scientifico solo nel 1956, durante la Conferenza di Dartmouth⁹.

Il concetto di intelligenza artificiale in senso lato però, ha origine nell'antichità: dai servi meccanici di Efesto¹⁰, alle macchine di Erone d'Alessandria¹¹, sino a Talos¹², il gigante di bronzo guardiano di Creta, l'idea che qualcosa di altro, di diverso dall'uomo, potesse assumere tratti simili ad esso ha sempre caratterizzato la mitologia e la storia umana. Spostandoci più avanti negli anni continuiamo a trovare esempi di storie riguardanti intelligenze autonome, come la colomba volante di legno che “da sola si spostava di ramo in ramo” di Archita da Taranto nel IV Secolo A. C.¹³, o la testa di bronzo del “papa mago” Silvestro II, capace di rispondere sì e no a semplici domande¹⁴. Tralasciando discussioni sulla veridicità e sulla coerenza dell'accostamento di tali esempi al concetto di Intelligenza Artificiale, ci è utile sottolineare come la congettura su un qualcosa di assimilabile alla ragione, diverso dall'essere umano, sia sempre esistita.

A livello teorico essa pone le sue radici nella teoria filosofica del meccanicismo¹⁵ e nel determinismo¹⁶, dottrine che definiscono il mondo e l'essere umano come una serie di meccanismi causa-effetto individuabili,

⁹Il Dartmouth Summer Research Project on Artificial intelligence, anche conosciuto come “Conferenza di Dartmouth” è il primo convegno di studiosi e ricercatori sull'Intelligenza Artificiale, comunemente riconosciuto come il punto di partenza dell'Intelligenza Artificiale come vero e proprio campo scientifico (si approfondirà ulteriormente più avanti)

¹⁰Macchine semoventi in grado di aiutare Efesto, dio del fuoco, delle fucine, dell'ingegneria, della scultura e della metallurgia, nella sua fucina nelle viscere dell'Etna

¹¹Unico congegno automatico, simile ad un teatrino, descritto nel dettaglio in “Automata”, tradotto dal greco nel 1589 da Girolamo Porro

¹²Costruito dallo stesso Efesto e incaricato da Minosse di sorvegliare l'isola di Creta, era un guardiano di bronzo invulnerabile

¹³*“In particolare il pensatore tarantino avrebbe elaborato un principio di armonia di scala, utilissimo per il dimensionamento e la verifica di alcuni componenti meccanici, rendendosi quindi protagonista di una vera e propria rivoluzione metodologica senza precedenti, consistente in un approccio sistematico alla meccanica mediante l'utilizzo di principi matematici e anticipando quindi di un secolo le indagini altrettanto sorprendenti e raffinate che gli ingegneri alessandrini eseguiranno in epoca ellenistica” (La colomba di Archita, Antonio Tagliente, 2011, Scorpione editrice)*

¹⁴*“Possedeva nel suo palazzo, una testa di bronzo che rispondeva sì o “no” alle domande che egli le rivolgeva sulla politica e sulla situazione della cristianità. Secondo Silvestro II questo procedimento era molto semplice e corrispondeva al calcolo con due cifre. Possibile che Gerberto potesse costruire una macchina analitica in codice binario? Dai cronisti dell'epoca apprendiamo che egli inventò un globo celeste in cui tutti gli astri avevano proprie orbite e propri movimenti e compivano in tempi proporzionati le proprie rivoluzioni. Si tratta, evidentemente, di un planetario: un'invenzione che fu ripresa solo nel XIX secolo. Inventò anche un orologio meccanico e un organo a vapore. Di tutte queste invenzioni e di altre ancora non abbiamo che notizie imprecise, poiché alla sua morte tutto fu distrutto”, (Patrologia Latina, Jacques Paul Migne, XIX sec.)*

¹⁵In generale, concezione di tipo materialistico che tende a spiegare le proprietà degli oggetti e dei processi del mondo fisico in termini esclusivamente meccanici, cioè sulla base di concetti connessi con la materia e il movimento, giungendo talora a interpretare ogni aspetto della realtà, sia naturale sia umana (storica, individuale, sociale), come il prodotto di una causalità deterministica: negando quindi la presenza di una finalità superiore o immanente (in tal senso, il termine è generalmente contrapposto a finalismo, teleologismo), ovvero riducendo la complessità e l'evoluzione del mondo reale a relazioni lineari di causa-effetto che, regolate da leggi immutabili, escludono la considerazione di dinamiche di tipo dialettico o di processi retroattivi. (Treccani)

¹⁶Concezione secondo la quale gli accadimenti della realtà metafisica, fisica o morale sono reciprocamente connessi in modo necessario e invariabile. In particolare il d. riguarda il rapporto di necessità tra causa ed effetto, tra legge naturale e fenomeno, per cui, data una causa o una legge, non può che prodursi in modo necessario e univoco quell'effetto o quel fenomeno specifico. In base a ciò l'Universo non presenta alcun evento contingente sotto forma di variazione spontanea o casuale, né persegue alcuna libera finalità (Treccani)

conoscibili e riproducibili. A tali scuole di pensiero si rifacevano Cartesio ed Hobbes, che potremmo definire in questa sede come i padri della concezione meccanica del pensiero umano. Cartesio nel 1637, nel suo “Discorso sul metodo”, ha individuato due caratteristiche che distinguono l’uomo dagli animali e dalle macchine: l’uso del linguaggio e la ragione e allo stesso tempo si domandava se prima o poi le macchine sarebbero arrivate a riprodurre, in linea teorica e mai totalmente congruente, anche queste due caratteristiche¹⁷; Hobbes è arrivato ad interpretare le dinamiche interne del corpo umano come una macchina, domandandosi se fossero replicabili o meno la nostra struttura e la nostra intelligenza¹⁸.

Nel XVIII secolo i tentativi di ricreare figure simili all’umano sono stati diversi, tra cui il celebre “Turco”, un automa che avrebbe dovuto simulare un giocatore di scacchi¹⁹ regalato da un inventore ungherese a Maria Teresa d’Austria. Dello stesso periodo scrivono così alcuni studiosi dell’Università di Genova²⁰:

Fino all'Ottocento i costruttori di 'macchine inutili', di tipo zoomorfo e antropomorfo, venivano finanziati solo se i loro automi erano in grado di divertire e stupire. Correttamente, Vittorio Marchis ha introdotto il termine 'machina ludens' per definire questo strano settore ricco di strutture artificiali impiegate nei teatri, nelle processioni e cerimonie religiose, nei palazzi di re e principi, nei salotti e nelle camere da giochi ove si volevano stupire ospiti di riguardo. La macchina era intesa come spettacolo e l'interpretazione è appropriata in epoche in cui lo spettacolo offerto da oggetti inanimati, date le tecniche a disposizione, poteva essere realizzato solo attraverso l'impiego di simulacri articolati (rigorosamente tridimensionali) di uomini e animali [...] Il passaggio dall'automa di intrattenimento al robot in grado di lavorare avvenne all'inizio del Novecento. È difficile individuare i motivi che hanno causato questa svolta. Tuttavia, è innegabile che l'Ottocento, con l'avvento delle grandi ideologie legate al capitalismo, al socialismo e al materialismo storico, aveva posto le premesse per una sensibilizzazione sulle tematiche del lavoro.

Nell’Ottocento cambia quindi l’obiettivo con cui si tenta, spesso invano, di riprodurre qualcosa di simile agli automi. Lo studio su queste tematiche ha iniziato a riguardare di più il mondo “scientifico” in senso stretto, in particolare matematici e fisici e gli studi nello specifico verteranno principalmente sulle teorie del pensiero deduttivo e sulle connessioni neurali, oltre che sulla logica e la probabilità. Oggi diremmo che si stava passando concettualmente da un interesse per i robot ad un interesse per le intelligenze artificiali.

¹⁷Renè Descartes, *Discorso sul Metodo*, 1637. Egli dirà “assegnando nel pensiero un certo ordine anche a quegli oggetti che nella loro natura non stanno in una relazione di antecedenza e conseguenza”

¹⁸Thomas Hobbes, *Leviathan or The Matter, Forme and Power of a Common Wealth Ecclesiastical and Civil*, 1651

¹⁹Sorta di manichino di legno incorporato nella struttura della scacchiera con all’interno una persona di bassa statura che poteva nascondersi alla finta verifica del fatto che non ci fosse nessuno all’interno e che governava le braccia meccaniche, vedendo le mosse che avvenivano sulla scacchiera attraverso dei magneti posti al di sotto

²⁰Vincenzo Tagliasco (*Cattedra Bioingegneria Università degli Studi di Genova*), Annamaria D’Ursi (*Content provider per sistemi audiovisivi*), Riccardo Manzotti (*LIRA-LAB del DIST Università di Genova*), *Macchine che pensano: dalla cibernetica alla coscienza artificiale*, Media Duemila, 2002

Il primo vero e proprio tentativo di spaccettare la logica di pensiero lo ha effettuato George Boole, un matematico britannico, con il suo *“The Laws of Thoughts: The Mathematical Theories of Logic and Probabilities”* del 1854. Boole accostò la logica alla matematica ed ebbe il merito di introdurre il concetto di variabile booleana (o input/output booleani), cioè una variabile binaria (0,1) che assumerà valore 0 in caso di valore negativo e valore 1 in caso di valore positivo²¹. Pochi anni dopo, la teoria di Boole venne ridimensionata da Frege, che ci tenne a precisare che la logica fosse più accostabile all’aritmetica piuttosto che alla matematica in generale. Frege studiò la logica della deduzione e formulò una procedura per mettere alla prova i principi matematici, più specificatamente se essi portassero o meno a determinati risultati e fossero generalizzabili. Per l’inizio del Novecento era ormai comunemente accettato che la maggior parte dei principi matematici potesse essere espresso da un insieme di nozioni primitive accostabili a logicismi²².

Russel e Whitehead scrissero *“Principia Mathematica”*, constatando che l’opera di Frege fosse incompleta e ampliando le nozioni matematiche da lui considerate esaustive. Hilbert e Ackermann ebbero il merito di spostare la questione sui processi decisionali, ponendo le basi concettuali per i moderni algoritmi. Infine, Alan Turing, considerato il padre dell’Intelligenza Artificiale, insieme ad altri studiosi, consapevoli dell’incompletezza dei modelli precedenti, tentò di introdurre un metodo per testare una macchina, provando a capire se essa fosse capace di svolgere qualsiasi calcolo computazionale. I risultati constatarono che ci fossero dei limiti al tempo considerati invalicabili che la logica matematica non avrebbe potuto superare.

Il focus in questo periodo si spostò dalla volontà di superare i limiti alla possibilità di, consapevoli di questa limitatezza, dare risvolto pratico a tutto questo progresso teorico creando effettivamente delle macchine che potessero eseguire calcoli. La prima vera e propria macchina calcolatrice venne in realtà inventata da Wilhelm Schikard nel 1623, una sorta di orologio calcolante, precedendo quella che per anni fu considerata la prima calcolatrice moderna, cioè la cosiddetta “Pascalina” di Pascal del 1642. Entrambe le macchine, pur considerate sensazionali all’epoca, potevano solo svolgere addizioni e sottrazioni per un massimo di cifre determinato. Per quanto riguarda la paternità dell’Intelligenza Artificiale, Charles Babbage aveva già progettato nel 1837 un computer che svolgesse funzioni elementari²³ ma morì prima di ultimarlo. Ada Lovelace è invece spesso riconosciuta come la prima ad aver stilato un programma computeristico²⁴. Nonostante ciò, è unilateralmente riconosciuta ad Alan Turing la paternità della moderna concezione di Intelligenza Artificiale.

Quindi secondo tali teorizzazioni, tutto ciò che è meccanico è intelligenza artificiale? Quand’è che il termine ha assunto al tempo stesso l’accezione e la complessità che oggi lo caratterizzano?

²¹Ci servirà in questa sede capire come una variabile simile possa essere accostata a qualsiasi domanda avente come possibili risposte sì o no

²²Martijn Kuipers (Universidade Luisiada de Lisboa; INOV-INESC Inovação), Ramjee Prasad (CTIF Global Capsule, Dep. BTech, Aathus University), *Journey of Artificial Intelligence*, Published online 29 October 2021, under exclusive licence to Springer Science+Business Media, LLC, Part of Springer Nature 2021

²³La cosiddetta “Macchina analitica”, primo prototipo di computer meccanico sviluppato per eseguire compiti generici

²⁴Collaborò con Babbage e fu la prima ad intuire che i computer potessero andare oltre il mero calcolo numerico

1.2.2 La Conferenza di Dartmouth e gli sviluppi successivi

L'utilizzo più attinente del termine "Intelligenza Artificiale" per come lo intendiamo oggi risale a più di settant'anni fa, quando lo stesso Alan Turing affrontò il problema e propose un esperimento che divenne noto come test di Turing, un tentativo di definire uno standard per una macchina da poter considerare "intelligente"²⁵.

Nello stesso periodo, McCulloch e Walter Pitts teorizzavano il primo modello di rete neurale²⁶, sviluppato poi da Hebb²⁷ qualche anno dopo. All'indomani della Conferenza di Dartmouth, Minsky e Edmonds progettavano il famoso SNARC, primo vero e proprio computer neurale, avente una decina di neuroni artificiali. Nel 1956, in New Hampshire, precisamente al Dartmouth College, fu organizzato quello che verrà ricordato come l'evento fondante dell'Intelligenza Artificiale come campo scientifico vero e proprio. L'organizzatore, John McCarthy, al tempo professore di matematica del college, dichiarò che la conferenza doveva "procedere sulla base della congettura che ogni aspetto dell'apprendimento o di qualsiasi altra caratteristica dell'intelligenza può essere descritto in linea di principio in modo così preciso da poter essere simulato da una macchina"²⁸. Si tenne così uno studio di due mesi da parte di dieci scienziati per riuscire a capire come far sì "che le macchine usassero il linguaggio, formassero astrazioni e concetti, risolvessero tipi di problemi ora riservati agli esseri umani e migliorino la loro capacità di apprendimento"²⁹. Da lì in poi, fino al 1974 circa, si attraversò un periodo d'oro per l'intelligenza artificiale: nuovi programmi vennero avviati, tra cui computer che risolvevano problemi di algebra e imparavano a parlare inglese, accompagnati dall'entusiasmo dell'opinione pubblica. Si sviluppò il perceptrone, una sorta di neurone adibito alla classificazione, e si diede vita alla prima Chatbot mai esistita³⁰. Nel 1974 però, dopo un iniziale entusiasmo sulle prospettive future e sulla possibilità di superare limiti fino ad allora neanche lontanamente raggiunti, il governo britannico e quello statunitense smisero di finanziare le ricerche riguardanti questo ambito. Parallelamente, le potenzialità di tali progetti si scontrarono con una realtà dei fatti che vide le macchine non avere la capacità di raggiungere gli obiettivi sperati. Questo periodo verrà infatti successivamente ricordato come "l'inverno delle IA". Grazie però ad un aumento della qualità dei computer IBM e Apple e ad un ingente investimento del governo giapponese in progetti su computer di quinta generazione, il settore ebbe una breve ripresa, che si arrestò solo a causa di una bolla

²⁵Turing, A.M., *Computing machinery and intelligence*. Mind, 1950, 433-460.

²⁶Sorta di rete di neuroni assimilabili a variabili Booleane che davano un determinato esito. Un esempio calzante per capire concettualmente come funziona una rete neurale è una domanda il cui esito sì/no è condizionato da altre domande con esito sì/no. A determinate combinazioni delle domande che condizionano corrisponderà un determinato esito della domanda principale.

²⁷La legge di Hebb è la base neuropsicologica dell'apprendimento: "se un neurone A è abbastanza vicino ad un neurone B da contribuire ripetutamente e in maniera duratura alla sua eccitazione, allora ha luogo in entrambi i neuroni un processo di crescita o di cambiamento metabolico tale per cui l'efficacia di A nell'eccitare B viene accresciuta".

²⁸*Artificial Intelligence Coined at Dartmouth*, Dartmouth official website

²⁹J. McCarthy, Dartmouth College M. L. Minsky, *A proposal for the Dartmouth Summer Research Project on Artificial Intelligence*, Harvard University N. Rochester, I.B.M. Corporation C.E. Shannon, Bell Telephone Laboratories, 31 agosto 1955

³⁰ELIZA era considerata una delle prime applicazioni dell'Intelligenza artificiale ed era in grado, attraverso l'analisi di parole chiave, di rispondere con frasi preimpostate ad alcuni tipi di domande

economica del mercato³¹, causata dalle alte aspettative che il periodo precedente si era inevitabilmente portato con sé. Nel 1980 nacque XCON, sistema contenente più di 2500 regole e capace di raggruppare in se più funzioni permettendo di risparmiare sia in termini di costi che di funzionalità necessarie per svolgere un determinato compito, e negli anni Novanta si riscontrò un rinnovato interesse per l'intelligenza artificiale, soprattutto grazie ad un approccio più settoriale con la consapevolezza di una percezione poco ottimistica da parte del settore economico e la definitiva rassegnazione sul superamento di limiti apparentemente invalicabili. Da qui in poi questa branca di ricerca scientifica iniziò a settorializzarsi sempre di più. Il campione mondiale di scacchi venne battuto da una CPU, nacque il primo veicolo a guida autonoma e nel 2002 si sviluppò addirittura il primo impiego di Intelligenze Artificiali in casa: Roomba progettò il primo robot pulitore automatico che grazie a dei sensori collegati ad un'IA riesce a mappare le stanze e individuare lo sporco da pulire.

Al giorno d'oggi, le tecnologie che usano IA per svolgere le proprie funzioni sono innumerevoli e alla portata di tutti. Da Alexa a Siri, dai Tools di Google sino ai sistemi di riconoscimento facciale, l'Intelligenza Artificiale è entrata a pieno nelle nostre vite. Anche nel campo della robotica i progressi sono sensazionali. Questo dicembre il Chief Technology Officer di Sony ha detto in un'intervista a Reuters che l'azienda ha gli strumenti per progettare robot umanoidi e che saranno fatti investimenti non appena verranno testate tutte le funzionalità potenzialmente utili³². Elon Musk questo settembre ha presentato il prototipo "Optimus", robot umanoide di ultima generazione. I settori in cui vengono impiegate le IA sono infiniti: e-commerce, ricerche, assistenti digitali, cybersicurezza, automobili, social, navigatori e perfino le notifiche delle app dei nostri cellulari utilizzano Intelligenze Artificiali per cogliere il momento migliore in base alle nostre abitudini. Le potenzialità sono molte e i rischi altrettanti. Urge la necessità di trovare un giusto equilibrio tra miglioramenti che le Intelligenze Artificiali possono portare in ambito lavorativo e nella vita di tutti i giorni e rischi ad esse correlati, come la privacy e la protezione dei dati, oltre che al possibile impatto negativo sull'occupazione. Persino l'Unione Europea ha iniziato a cogliere l'importanza strategica di queste tecnologie: il 12 febbraio del 2019 è stata approvata la "Risoluzione del Parlamento europeo su una politica industriale europea globale in materia di robotica e intelligenza artificiale" e il 21 aprile 2021 è stato proposto sempre al Parlamento Europeo un "Regolamento sull'approccio europeo all'Intelligenza Artificiale". L'ultimo Regolamento sta affrontando un iter abbastanza travagliato proprio per difficoltà nel bilanciare costi, rischi e opportunità. Molti paesi membri ne sottolineano le problematicità, anche se non possono ignorarne i benefici. Più avanti in questa sede analizzeremo le questioni più propriamente etiche e deontologiche, affrontando anche l'aspetto giuridico dell'uso di queste tecnologie. In ogni caso non si può ignorare la realtà dei fatti e accettare che queste innovazioni faranno sempre più parte della nostra realtà quotidiana e lavorativa.

³¹È un fenomeno originato dall'eccessivo entusiasmo nei confronti di un nuovo prodotto, tutti ne parlano, tutti ne vogliono fare parte e si perde di vista il valore reale dell'oggetto arrivando a sovrastimarne e ad innescare un cortocircuito economico (Alessandra Bianchi, Consulente Finanziaria di BancaWidiba)

³²Kiyoshi Takenaka, Mayu Sakoda, *Sony says it has technology for humanoid robots, needs to find usage* Intervista di Reuters, 2022

C'è un settore che, più di altri, pone in essere alcune questioni etiche sul bilanciamento tra logica del profitto e utilità per i cittadini e la società sull'utilizzo delle Intelligenze Artificiali: il giornalismo.

1.3 L'Intelligenza Artificiale nel giornalismo

Il giornalismo non è solo un settore economico delle moderne società. È uno degli aspetti fondamentali della democrazia: mettere a conoscenza i cittadini di ciò che accade nel mondo e nelle istituzioni, fornire dati accurati e dare voce a più opinioni possibile per promuovere il pluralismo sono solo alcuni degli obiettivi che una società democratica dovrebbe porsi. È ormai sotto gli occhi di tutti che al giorno d'oggi esso sia un settore in crisi e non riesca più ad avere quel ruolo centrale di mediatore che aveva un tempo. In una società in cui sempre di più si ha la percezione di avere a portata di mano qualunque tipo di informazione, in cui i politici ormai comunicano (o quantomeno danno l'impressione di comunicare) in maniera diretta con il proprio elettorato, in cui si ha a disposizione qualunque tipo di dato per qualunque tipo di ambito e in cui i social fanno ormai da padrone, il ruolo del giornalismo sembra aver perso la sua importanza.

È vero, la lettura di giornali cartacei è diminuita drasticamente negli ultimi anni e continua a scendere vertiginosamente³³, ma è in aumento la lettura di giornali online, siti web di informazione, radio online e telegiornali. Il modo di informarsi degli italiani è sempre più digitale e variegato ed è lampante che il problema non sia una mancata volontà dei cittadini ad informarsi, ma un adattamento del settore alle loro esigenze e una difficile sostenibilità economica delle redazioni. Il 12% degli italiani paga per avere notizie ma ciò non basta per mantenere economicamente in positivo un giornale, a maggior ragione se si tratta di piccole redazioni. Spesso ci si sostiene solo con inserti pubblicitari, con un inevitabile anche se minimo condizionamento da parte di chi paga indirettamente chi scrive. Nonostante lo scenario poco positivo, il numero di giornali nati negli ultimi anni è aumentato. Ciò vuol dire che non mancano all'attivo giornalisti e imprenditori che vogliono investire nel campo dell'informazione³⁴. Approfondiremo l'aspetto economico in un sottoparagrafo dedicato³⁵.

L'intelligenza artificiale nel giornalismo, per quanto recente e ancora poco sviluppata soprattutto nelle redazioni italiane, potrebbe essere utile proprio per soddisfare queste esigenze: potrebbe ridurre i costi nella produzione e distribuzione di notizie e potrebbe far sì che vengano soddisfatte a pieno le esigenze dei lettori. Sicuramente oltre alle opportunità, come detto, ci sono anche molti rischi. I pericoli di rimanere imbrigliati in una sezione di notizie eccessivamente personalizzata o di arrivare ad un mero report di dati senza argomentazioni o discussioni sul merito delle notizie sono dietro l'angolo, ma non per questo bisogna scoraggiarsi ed etichettare come negative tali tecnologie.

Facciamo però un passo alla volta. È necessaria un'ulteriore concettualizzazione e contestualizzazione sull'oggetto di questo lavoro.

³³Si è passati dal 67% nel 2007 ad un 29,1% nel 2021 (Sintesi per la stampa del Diciottesimo rapporto sulla comunicazione, 2021)

³⁴Censis, *I media dopo la Pandemia*, Sintesi per la stampa del Diciassettesimo rapporto sulla comunicazione, 6 ottobre 2021

³⁵Capitolo 4 paragrafo 4.2 sottoparagrafo 4.2.2

1.3.1 I termini appropriati

La tecnologia in generale è ormai parte integrante se non fondamentale del mondo dei media. Questo ha contribuito alla diffusione e alla prevalenza di forme quantitative di giornalismo. Nell'ultimo decennio il *Data Journalism*³⁶ è diventato la forma preponderante di giornalismo grazie alla grande mole di dati disponibili. Insieme ad esso, però, si sono diffuse altre forme di giornalismo digitale come l'*automated journalism*, il *metrics driven journalism* e l'*algorithmic journalism*. Ovviamente come per l'IA in generale, anche le forme di giornalismo non sono unanimemente condivise. Intese come categorie sono spesso convergenti tra loro. Il modo in cui le redazioni stesse definiscono le Intelligenze Artificiali, cioè come se ne parla tra colleghi o come gli addetti ai lavori ne parlano al loro pubblico di riferimento, è una buonissima indicazione di cosa si intende per IA nel giornalismo. Un terzo di chi lavora nelle redazioni sostiene che *Machine Learning* sia maggiormente utilizzato come termine e che la parola Intelligenza Artificiale sia poco usata perché troppo generale. Un quarto di loro dice di non saper rispondere a "che definizione darebbe di IA" perché vi è un ampio dibattito in materia. La BBC è l'eccezione che conferma la regola: sostiene che il termine sia utile per raggruppare, come fatto in questa sede, tutte le tecnologie avanzate³⁷. Il termine forse più utilizzato di quelli sopramenzionati è "*Automated Journalism*". Il *Metrics-Driven Journalism* si focalizza molto sui dati dei lettori al fine di influenzare il processo produttivo delle notizie sulla base di ciò, ma identifica solo un modo di operare, non una vera e propria forma di giornalismo. L'*Algorithmic Journalism* si riferisce a tutte quelle notizie nel quale il processo produttivo è avvenuto tramite l'aiuto di un algoritmo e potrebbe essere definita quasi come una categoria comprendente l'*Automated Journalism*, poiché esso si riferisce nello specifico a quei "processi algoritmici che convertono dati in testi di notizie narrative con limitato o nessun intervento umano oltre l'iniziale programmazione"³⁸. Per *Automated Journalism* si intende quindi una "specializzazione dell'informatica che mira a sostituire del tutto i giornalisti umani per sostituirli con degli algoritmi in grado di prelevare contenuti e riscriverli secondo dei criteri prestabiliti dall'uomo"³⁹. Sul fatto se sostituiscano o meno effettivamente il lavoro del giornalista discuteremo più avanti. Ci interessa ora sottolineare come vengano utilizzati i termini in questione. Un altro termine molto utilizzato è *NLG Journalism*. Il termine si riferisce agli algoritmi di generazione del linguaggio per aiutare gli esseri umani a scrivere articoli. La sigla NLG, come detto, significa "*Natural Language Generation*" e identifica però solo la parte dell'output del processo, cioè la generazione del linguaggio. Il termine più appropriato per identificare l'intero processo è *Natural Language Processing*, che comprende in sé sia il *Natural Language Generation* ma anche il *Natural Language Understanding*, cioè la parte dell'input.

³⁶È una categoria generale che raggruppa tutte quelle forme di giornalismo che traggono informazioni dai dati

³⁷Charlie Beckett, *New Power New Responsibilities, A global survey of journalism and artificial intelligence*, London School of Economics and Political Science, Google News Initiative, novembre 2019

³⁸Le definizioni di metrics-driven journalism, algorithmic journalism e automated journalism presentate fin qui sono tratte da "Artificial Intelligence and Automated Journalism: Contemporary Challenges and New Opportunities". Waleed ALI, Mohamed HASSOUN. *International Journal of Media, Journalism and Mass Communications (IJMJC)*, vol 5, no. 1, pp. 40-49, 2019

³⁹Arjen Van Dalen, *The algorithms behind the headlines, How machine-written news redefines the core skills of human journalists*, Journalism Practice, 2012



Fig.4 Natural Language Processing⁴⁰

Altro termine molto utilizzato è *robot journalism*⁴¹, ma in realtà nei processi descritti fino ad ora non è coinvolto nessun vero e proprio robot, ma codici e *script* di software. Linden mette in guardia contro una "concettualizzazione popolare ma banale dove si ritraggono robot che scrivono su tastiere del computer" che gioca sui timori dei giornalisti di essere licenziati e impedisce che le redazioni siano più innovative.

Un termine invece atipico che, a mio avviso, rispecchia molto il ruolo delle intelligenze artificiali nel giornalismo è l'"*exojournalism*", una proposta concettuale che connette le possibilità offerte dall'IA con la routine giornalistica, intendendo l'IA come una sorta di esoscheletro del giornalista. Questa proposta concettuale rispecchia a pieno, a differenza di chi, come sottolineava Linden, ne accentua le preoccupazioni piuttosto che sottolinearne i vantaggi⁴², il ruolo che hanno oggi le intelligenze artificiali nel giornalismo: uno strumento in mano ai redattori che hanno sempre l'ultima parola e possono sempre modificare e utilizzare a loro piacimento le tecnologie⁴³.

1.3.2 Tipologie di Intelligenze Artificiali utilizzate nel giornalismo

Dopo una breve rassegna generale analizzeremo ora come vengono usate le intelligenze artificiali nello specifico nelle redazioni. La BBC nello stesso studio precedentemente citato suddivide i tipi di IA utilizzate nelle redazioni in tre gruppi:

- Tecnologie basate su ML
- Tecnologie presentate come Agenti Intelligenti (Voice Agents e Chatbots ad esempio)
- Tecnologie che automatizzano compiti specifici

⁴⁰Mathias-Felipe de Lima-Santos and Wilson Ceron, *Artificial Intelligence in News Media: Current Perceptions and Future Outlook*, *Journal Media*, 13-26, dicembre 2021

⁴¹*A systematic review of automated journalism scholarship: guidelines and suggestion for future research*, Samuel Danzon-Chambaud, School of Communications, Dublin City University, 24 marzo 2021, open Research Europe 2021

⁴²*A systematic review of automated journalism scholarship: guidelines and suggestion for future research*, Samuel Danzon-Chambaud, School of Communications, Dublin City University, 24 marzo 2021, open Research Europe 2021

⁴³Tejedor, Santiago, and Pere Vila, *Exo Journalism: A Conceptual Approach to a Hybrid Formula between Journalism and Artificial Intelligence*, *Journalism and Media* 2: 830–840, 2021

Questa suddivisione si focalizza più sul metodo che sul merito. Lungi da me criticare tale suddivisione, sarà più utile alla nostra causa una suddivisione nel merito, cioè su che funzione svolgono effettivamente le IA nelle redazioni. Anche perché spesso i metodi, come già più volte ripetuto, sono ibridi o si incrociano tra loro.

Una suddivisione generale in macrocategorie, largamente condivisa e che può comunque esserci utile per capire meglio l'utilizzo di tali tecnologie, seguendo un criterio di suddivisione inerente agli step del processo di creazione di contenuti nelle redazioni è la seguente⁴⁴:

- *News Gathering*: le IA vengono utilizzate per cercare informazioni, eventi e storie con nuovi metodi di selezione; vengono usati spesso anche software ML che trovano *outliers*⁴⁵ tra i dati, che trovano automaticamente tendenze sui social e che estraggono informazioni da file
- *News Production*: i giornalisti possono ora avere l'opportunità di produrre contenuti automaticamente, utilizzando anche algoritmi che modificano diversi formati media in altri formati media (come testo dai dati, video dai testi, testi dai video, audio dai testi, testi da audio ecc.)
- *News Distribution*: i giornalisti hanno ora la possibilità di raggiungere nuovi consumatori su nuove piattaforme attraverso software di profilazione e ottimizzando le pubblicazioni, attraverso contenuti personalizzati e adattati all'utente

Rientrano nella prima fase di creazione di una notizia (*News Gathering*):

- Computer Vision
- Ricerca e verifica di informazioni (per combattere disinformazione e misinformazione)
- Previsione di tendenze

Rientrano nella seconda fase di una notizia (*News Production*):

- Software audio o video
- Scrittura automatica
- Deepfake

Rientrano nella terza fase di una notizia (*News Distribution*):

- Alert e personalizzazione di contenuti

⁴⁴Francesco Marconi, *Newsmakers. Artificial intelligence and the future of Journalism*, Columbia University Press, New York, 2020

⁴⁵Outlier è un termine utilizzato in statistica per definire, in un insieme di osservazioni, un valore anomalo e aberrante, ossia un valore chiaramente distante dalle altre osservazioni disponibili. (Wikipedia)

1.4 Storia delle Intelligenza Artificiale nel giornalismo

I quotidiani hanno attraversato quattro secoli e sono passati per diversi supporti, modificando le proprie caratteristiche e facendosi interpreti di cambiamenti sociali, politici ed economici, talvolta anche stimolandoli. Il primo quotidiano di cui si ha traccia nella storia fu pubblicato nel 1660 a Lipsia dal tipografo Timoteo Ritzsch, con una tiratura di 204 copie. Portava il titolo di “*Notizie fresche degli affari della guerra e del mondo*” e il testo era disposto su un'unica colonna. Sono passati più di trecentocinquant'anni e lo scenario è drasticamente cambiato. Si è passati dal cartaceo al digitale, dal testo all'ipertesto, dai salotti ai social network. Tralasciando cenni sulla storia del giornalismo, ci interessa principalmente capire come sono avvenute tali transizioni e quando sono state implementate le prime forme di IA nel mondo dell'informazione. La storia dei media mostra che i produttori di notizie, pur legati ai mezzi, si sono trasformati non in conseguenza di un'evoluzione a sé stante, ma per un complesso intreccio di relazioni tra dinamiche culturali, sociali ed economiche. È innegabile che internet abbia radicalmente cambiato il nostro modo di informarci ma anche la percezione di tali cambiamenti, l'utilizzo che è stato fatto delle nuove tecnologie e le loro dinamiche spontanee che hanno modificato in lungo e in largo le redazioni giornalistiche sia in quanto settori economici che in quanto strumenti fondamentali della democrazia, hanno influenzato l'evolversi delle tecnologie.

L'intelligenza artificiale si inserisce in questo contesto nella seconda metà del XX secolo, quando internet è entrato a far parte pian piano delle nostre vite e l'IA iniziava a svilupparsi come vero e proprio settore scientifico.

Sfogliando la letteratura scientifica e cercando sui motori di ricerca il primo esempio di utilizzo di intelligenze artificiali nel mondo dell'informazione è il prototipo di programma di scrittura automatica “Stats Monkey”, ideato da Kristian Hammond e Larry Birnbaum all'interno di un progetto dell'Intelligent Information Laboratory della Northwestern University.

Nonostante tale primato risulti corretto, non è il primo esempio di vera e propria scrittura automatica.

1.4.1 Articoli automatici di previsioni meteo

Il primo vero esempio di scrittura automatica di notizie sono previsioni prodotte in automatico da un computer contenente 80 frasi preimpostate che, partendo da input di dati rilevati ed elaborati dall'ESSA Weather Bureau's National Meteorological Center, produceva brevi articoli di previsioni meteo⁴⁶. Le previsioni erano prodotte da modelli di equazione primitiva⁴⁷ e avvezione sub sinottica⁴⁸ ed erano automatizzate nella scrittura di piccoli articoli meteorologici.

⁴⁶Harry R. Glahn, *Computer-produced Worded Forecast*, Techniques Development Laboratory, Systems Development Office, ESSA Weather Bureau, Silver Spring, Md, WEATHER SCIENCE, INC. 1225 WEST MAIN STREET NORMAN, OKLAHOMA 73069 1132 Vol 51, No. 12, December 1970

⁴⁷Serie di equazioni differenziali parziali non lineari usate per approssimare il flusso atmosferico globale e impiegate nella maggior parte dei modelli atmosferici.

⁴⁸Modello numerico per la previsione a breve termine (fino a 24 ore) utilizzato operativamente due volte al giorno

Il programma metteva in forma scritta le previsioni prodotte da queste equazioni, con tanto di valori numerici, virgole e punti. Erano utilizzati diversi titoli in maniera casuale per variare. L'unico errore che commetteva il programma era accorpare le probabilità di pioggia e di neve in unica "probabilità di precipitazioni".

Bulletin American Meteorological Society

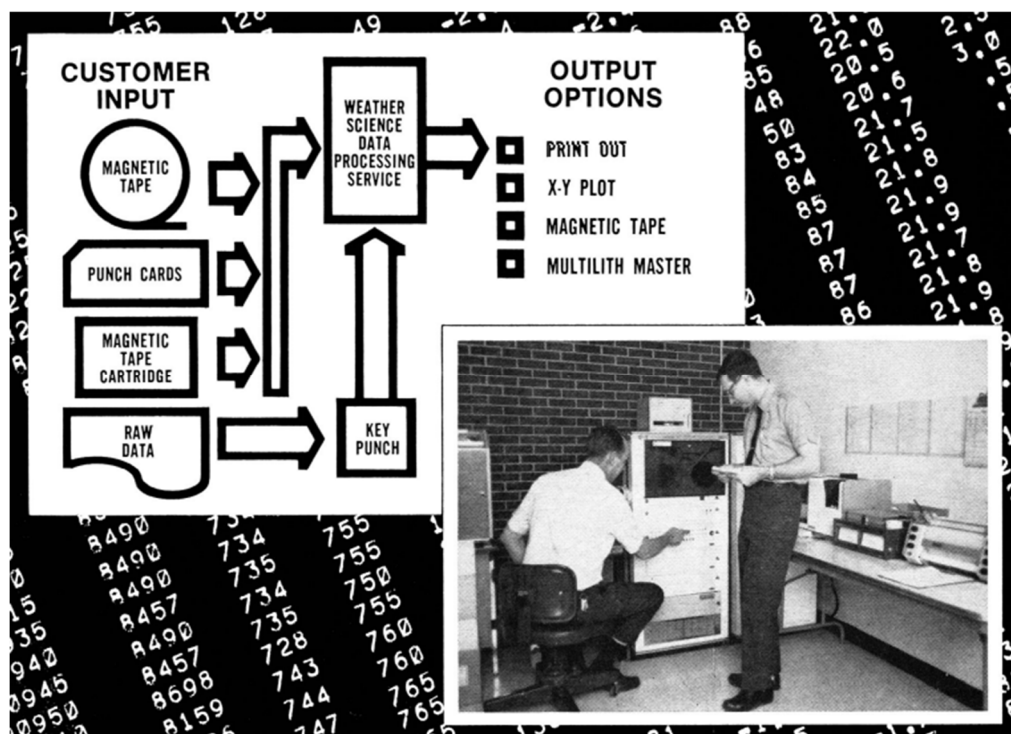
MAR 6, 1970

TDL EXPERIMENTAL FORECASTS

GOOD MORNING. THE SYSTEMS DEVELOPMENT OFFICE BRINGS YOU THE LATEST FORECAST FOR WASHINGTON, D. C. AND VICINITY. STRONG NORTHERLY WINDS THIS MORNING 20 MPH WITH GUSTS TO 30 MPH BECOMING NORTHERLY 5 MPH BY EVENING. COLDER TODAY, MAXIMUM TEMPERATURE 48 DEGREES. MOSTLY CLOUDY THIS MORNING WITH DECREASING CLOUDINESS THIS AFTERNOON. ONLY 2 PERCENT PROBABILITY OF PRECIPITATION TODAY.

GOOD MORNING. THE ESSA WEATHER BUREAU BRINGS YOU THE LATEST FORECAST FOR ATLANTA AND VICINITY. MOSTLY SUNNY THIS MORNING WITH A FEW MORE CLOUDS THIS AFTERNOON. COLDER TODAY, MAXIMUM TEMPERATURE 62 DEGREES. NORTHERLY WINDS 15 MPH THIS MORNING BECOMING NORTHEASTERLY 5 MPH BY LATE AFTERNOON. ONLY 2 PERCENT PROBABILITY OF PRECIPITATION TODAY.

*Fig. 5 Esempi di articoli metereologici
(Credits: Computer-produced Worded Forecast, Harry R. Glahn, ESSA Weather bureau)*



*Fig. 6 Struttura elementare del computer che programma gli articoli di previsione
(Credits: Techniques Development Laboratory, Systems Development Office, ESSA Weather Bureau, Silver Spring, Md, WEATHER SCIENCE, INC. 1225 WEST MAIN STREET • NORMAN, OKLAHOMA 73069 1132 Vol 51, No. 12, December 1970)*

Le previsioni potevano essere in ogni caso modificate ma lo stesso Glahn sosteneva che non sarebbe stato necessario in situazioni routinarie. Questa analisi sostiene inoltre una formidabile tesi sull'utilizzo di programmi simili che ancora oggi rispecchia a pieno come dovrebbero essere intese tali tecnologie: "i compiti di routine potranno essere gestiti dai computer, così da poter lasciare ai meteorologi le sfide più complesse come situazioni ad alto impatto meteorologico."

Gli articoli, come mostrato in figura 5, erano molto elementari e poco variegati. Si tratta comunque di un articolo del 1970 e fu una grande rivoluzione al tempo. La struttura del programma era molto elementare (figura 6) ma le prospettive future erano notevoli. Nell'articolo stesso si diceva che "in meno di tre anni questo tipo di messaggi avrebbero potuto essere disponibili anche in altre città degli Stati Uniti per previsioni del giorno prima, del giorno stesso e del giorno dopo". Ciò è realmente accaduto e a partire dal 1985 le previsioni trasmesse tramite articoli automatici hanno iniziato ad essere parte importante delle previsioni meteo. "Le parole possono essere molto espressive e possono comunicare il livello di fiducia delle previsioni. Esse sono spesso lette da persone che "fanno qualcos'altro". Questo suggerisce che i prodotti formulati continueranno ad avere un ruolo nella suite di prodotti del National Weather Service (Servizio Meteo Nazionale) anche se altre forme (tabella, grafica, griglia) sono diventate facilmente disponibili"⁴⁹.

Nei primi anni duemila è stata addirittura condotta uno studio sul confronto tra articoli di previsioni scritte da programmi o da meteorologi veri e propri: è stato dimostrato che gli utenti preferivano le previsioni scritte da un programma perché sceglieva meglio le parole. Fu la prima volta che uno studio mostrava che un testo NLG fosse meglio di un testo scritto da un umano⁵⁰.

Sicuramente le previsioni meteo non sono una vera e propria forma di giornalismo. Questo esempio ci serve per dimostrare come già negli anni Ottanta, in cui le abitudini dei consumatori in generale e dei consumatori di news nello specifico stavano cambiando⁵¹, fossero a portata di tutti brevi resoconti del meteo e che il lettore tipo di queste notizie automatiche fosse una "persona che faceva qualcos'altro" e che quindi fossero utili per comunicare dati in maniera veloce e senza approfondimenti.

Il primo esempio di giornalismo automatico in senso proprio avverrà, come detto, qualche anno più tardi.

1.4.2 Stats Monkey

Nel 2009 un gruppo di studenti di giornalismo e informatica seguiva un corso della scuola di giornalismo della Northwestern University, condotto da Kristian Hammond, informatico con dottorato a Yale, e Larry Birnbaum, suo collega. Stava nascendo un progetto all'interno dell'Intelligent Information Laboratory per un programma di scrittura automatica che, partendo da un set di dati statistici, produceva brevi storie base, giocata dopo giocata, di una partita di baseball. Il programma si basava su due sezioni di codice: la prima era composta

⁴⁹Matthew R. Peroutka, Ronla K. Meiggs and Michael B. Romberg, *The generation of products in interactive forecast preparation* Techniques Development Laboratory, Colorado, 1998

⁵⁰Ehud Reiter, Somayajulu Sripada, Jim Hunter, Jin Yu, Ian Davy, *Choosing words in computer-generated weather forecasts*, Department of Computing Science, University of Aberdeen, UK Aerospace and Marine International, Aberdeen, 15 agosto 2005

⁵¹Giovanni Ciofalo, *Infiniti anni Ottanta. Tv, cultura e società alle origini del nostro presente*, Mondadori, 2011

da algoritmi che sulla base di modelli statistici e analisi di precedenti match di baseball, insieme all'analisi dei cambiamenti della probabilità di vittoria e banalmente dei cambiamenti di punteggio, rilevava quali fossero le azioni rilevanti associandole anche ai giocatori partecipi di quella determinata azione; la seconda, partendo da una sorta di libreria di archi narrativi preimpostati, associava determinate parole all'azione rilevata dalla prima parte di codice, andando a produrre brevi frasi che descrivevano le azioni rilevanti della partita. Il risultato era una sorta di cronaca live con tanto di foto associate per alcuni giocatori.

Di seguito un esempio della partita di playoff dell'11 ottobre 2009 tra Los Angeles Angels e Boston Red Sox terminata con il risultato di 7-6 per la squadra californiana.

BOSTON — Things looked bleak for the Angels when they trailed by two runs in the ninth inning, but Los Angeles recovered thanks to a key single from Vladimir Guerrero to pull out a 7-6 victory over the Boston Red Sox at Fenway Park on Sunday.

Guerrero drove in two Angels runners. He went 2-4 at the plate.

“When it comes down to honoring Nick Adenhardt, and what happened in April in Anaheim, yes, it probably was the biggest hit (of my career),” Guerrero said. “Because I’m dedicating that to a former teammate, a guy that passed away.”

Guerrero has been good at the plate all season, especially in day games. During day games Guerrero has a .794 OPS. He has hit five home runs and driven in 13 runners in 26 games in day games.

After Chone Figgins walked, Bobby Abreu doubled and Torii Hunter was intentionally walked, the Angels were leading by one when Guerrero came to the plate against Jonathan Papelbon with two outs and the bases loaded in the ninth inning. He singled scoring Abreu from second and Figgins from third, which gave Angels the lead for good.

The Angels clinched the AL Division Series 3-0.

Angels starter Scott Kazmir struggled, allowing five runs in six innings, but the bullpen allowed only one runs and the offense banged out 11 hits to pick up the slack and secure the victory for the Angels.

J.D. Drew drove in two Red Sox runners. He went 1-4 at the plate.

Il sistema utilizzava un albero decisionale per selezionare l'arco narrativo appropriato. Questo individuava quindi i componenti principali della storia del gioco e consentiva al sistema di metterli insieme in modo coeso e convincente. Le storie potevano essere generate dal punto di vista di entrambe le squadre.

Il programma non poteva effettuare né previsioni né analizzare dettagli specifici della partita, oltre a non riuscire ad analizzare gli eventi inaspettati⁵². Il New York Times, pur riconoscendone l'ottimo lavoro, ne criticò l'impossibilità di valutare l'importanza della partita, che portò i Los Angeles Angels a giocarsi il titolo alla partita successiva⁵³.

In ogni caso lo strumento fu una grande innovazione e diede il via ad importanti sviluppi nel neonato settore dell'*automated journalism*. Il programma era stato progettato per coprire la Major League di baseball americana ma negli intenti iniziali del laboratorio c'era anche quello di fornire uno strumento a portata dei giornali locali, al fine di coprire anche le leghe minori con un minor impiego di risorse e permettere agli appassionati, agli amici e ai parenti dei giocatori o a chiunque fosse interessato di conoscere più del semplice risultato di una singola partita minore di baseball.

“Il nostro obiettivo è modellare realmente il pensiero, l'intelligenza e la ragione umana” “Devo ammettere che non lo stiamo facendo solo nello sport; stiamo valutando a quali altri ambiti potremmo applicare questa [tecnologia]”. Queste le parole di Hammond poco dopo la presentazione di StatsMonkey.

Stuart Frankel, presente alla presentazione del programma, ne intuì le potenzialità e parlando con Birnbaum e Hammond decisero di fondare Narrative Science, compagnia con sede a Chicago che sviluppò il brevetto di questa tecnologia e ne ottenne la proprietà intellettuale, iniziando le proprie attività commerciali nel 2010. Nonostante un iniziale cambio di direzione, dalla volontà di vedere dove le potenzialità di questa tecnologia potessero arrivare alla volontà di monetizzare tale invenzione⁵⁴, nel 2020 l'azienda ha lanciato “Data Storytelling for Good”, ramo no-profit che fornisce gratuitamente i suoi prodotti alle organizzazioni che fanno del bene nella loro comunità. Il competitor di Narrative Science in quegli anni fu Automated Insights, fondata da un ingegnere della Cisco nel 2007 con il nome di StatSheet e che cambiò nome nel 2011 proprio per la sua specializzazione in contenuti automatizzati. Nel 2015 Automated Insight lanciò il programma WordSmith, commercializzato su larga scala negli anni successivi e divenuto uno dei software più famosi di NLG, con clienti del calibro di Yahoo, Associated Press e Tableau.

⁵²Su molti articoli sono riportati due esempi: l'Home Run di Kirk Gibson nella World Series del 1988, divenuto famoso perché Gibson non era in condizioni ottimali e zoppicava, e l'Home Run di George Herman Ruth nelle World Series del 1932, prima del quale il giocatore puntò il dito nella direzione del battitore mostrando la sua convinzione. Questi due esempi sono riportati in quanto fuori dagli schemi rispetto ad una normale partita e considerati non rilevabili dal programma.

⁵³David Carr, giornalista del New York Times che scrisse l'articolo scrisse “Guess Algorithms can't do everything” quasi in tono provocatorio come a dire “Vedete che gli algoritmi non possono fare proprio tutto!?” probabilmente per esorcizzare i timori che tali strumenti portarono in quel periodo, come indica anche il titolo stesso dell'articolo “*The Robots Are Coming! Oh, They're Here*”. (*The Robots Are Coming! Oh, They're Here*. BY DAVID CARR, OCTOBER 19, 2009 9:51 AM.)

⁵⁴Ciò portò alla creazione di un'altra piattaforma, Quill, con l'obiettivo di venderla agli impiegati che non fossero pratici di analisi dei dati

Oggi le agenzie che lavorano e commercializzano software NLG sono molte⁵⁵ e sicuramente i programmi di Natural Language Generation sono tra i più importanti nel panorama delle IA nel giornalismo. Come abbiamo visto però i tipi di tecnologie utilizzate sono diverse e non si limitano solo alla scrittura automatica di articoli.

Analizzeremo nel prossimo capitolo diversi tipi di IA nel giornalismo attraverso esempi emblematici di esse. Le tecnologie analizzate comprenderanno tecnologie utilizzate in ogni step di creazione di una notizia, una per ogni tipo di tecnologia.

⁵⁵Secondo la "Market Guide for NLG" 2019 di Gartner, le principali aziende di NLG sono (in ordine alfabetico): Arria NLG, Automated Insights, AX Semantics, Narrative Science, vPhrase e Yseop. Altre aziende simili nel settore della generazione del linguaggio naturale sono Smartologic, Retresco, United Robots e Linguastat.

CAPITOLO 2 L' intelligenza artificiale nel giornalismo: utilizzi ed esempi virtuosi

Nel primo capitolo abbiamo introdotto concetti, definizioni e categorie. Ora illustreremo nello specifico che tipo di intelligenze artificiali si utilizzano nelle redazioni, anche attraverso alcuni esempi emblematici di importanti tg o giornali. La schematizzazione che abbiamo usato in questa sede, come detto, è quella utilizzata da Francesco Marconi⁵⁶ e da numerosi altri colleghi, che suddivide il *workflow giornalistico*⁵⁷ in *news gathering*, *news production* e *news distribution*. È doveroso sottolineare ancora una volta che queste categorie, per quanto utili e necessarie, sono soltanto indicative. Più è complessa una tecnologia e più è probabile che essa sia polifunzionale, cioè che assolva più funzioni contemporaneamente. Vedremo nello specifico più avanti come tecnologie utilizzate prima di produrre una notizia possano essere utilizzate anche durante o contemporaneamente alla sua creazione effettiva, o come alcune funzioni, anche se utilizzate comunemente in fasi differenti, siano in sostanza molto simili fra loro. Va inoltre precisato che il campo di ricerca nel quale stiamo operando si sviluppa ogni giorno sempre di più, quasi in maniera esponenziale e che quindi sperimentazioni di tecnologie, sia in termini di amplificazione di quelle già esistenti che in termini di creazione di nuovi software, stanno avvenendo proprio in questo periodo. Prima di analizzare nello specifico i tipi di tecnologie è necessaria una puntualizzazione. Nonostante la scrittura automatica sia la più menzionata e utilizzata nel mondo dell'informazione non è l'unica. Spesso in letteratura viene data maggiore enfasi a questo tipo di tecnologie⁵⁸ e la maggior parte della analisi a posteriori, che vedremo nei capitoli successivi, vengono effettuate su contenuti di questo tipo. Il numero di articoli scritti automaticamente per calcolare l'utilizzo percentuale delle IA nelle redazioni, il confronto tra un testo scritto da un IA o da un umano, il risparmio in termini di costi e chi è responsabile di quanto scritto sono tutti parametri che si focalizzano su questo ristretto tipo di tecnologia e che vengono largamente usati e menzionati per le analisi sull'oggetto del nostro lavoro. Non si sostiene in questa sede che esse non siano le più importanti e che non sia giusto utilizzare questi parametri. Si sostiene invece, e lo si corrobora attraverso le concettualizzazioni e categorizzazioni utilizzate e precedentemente menzionate, l'importanza di analizzare anche altri tipi di tecnologie. Un' impostazione dove sia la ricerca che le redazioni stesse analizzano e utilizzano principalmente solo le *machine-written news* risulterebbe limitante per diversi motivi. In passato questa impostazione era sicuramente quella maggioritaria e, più che impedire sviluppi ulteriori, ha rallentato sicuramente i processi di innovazione. Si può in ogni caso affermare con certezza che il sempre maggiore utilizzo di più tipi di contenuti informativi e la tendenza dell'ultimo decennio alla distribuzione di una stessa notizia in diversi format abbia permesso anche a questo ambito dispiegare le proprie ali e di poter esplorare nuove frontiere sia in termini di offerta per i lettori di news che in termini di risparmi e guadagni per le redazioni.

⁵⁶Francesco Marconi, *Newsmakers. Artificial intelligence and the future of Journalism*, Columbia University Press, New York, 2020, 6

⁵⁷Marconi nel suo libro intende per *workflow* la gestione dei processi lavorativi ridefiniti sulla base dell'utilizzo delle IA.

⁵⁸Noain-Sánchez, A, *Addressing the Impact of Artificial Intelligence on Journalism: the perception of experts, journalists and academics*, *Communication & Society*, 35(3), 105-121, 2022

2.1 Le intelligenze artificiali per il news gathering

Il termine *News Gathering* tradotto letteralmente vuol dire “raccolta di notizie”. Il termine *to gather* significa proprio radunare, raccogliere, riunire e nell’ambito del giornalismo indica il processo iniziale di creazione di una notizia. Pensiamo ad un giornalista che deve scrivere una notizia. Qual è la prima cosa che fa? Cerca la notizia, scova le tendenze del momento, si mette in condizione di avere più informazioni possibili su un determinato argomento. Tutta questa parte prima della creazione vera e propria del contenuto fa parte di quello che chiamiamo *News Gathering*.

Il merito di distinguere questa fase da quella di produzione effettiva fu di Abraham Z. Bass, il quale suddivise il *News Making* in generale in un processo a doppia azione: da una parte i *news gatherers* (reporters e line editor) e dall’altra i *news processors* (editori e traduttori). Si divideva così la “notizia grezza” dal prodotto completo. Enfatizzò l’importanza della prima rispetto alla seconda poiché senza di essa la seconda neanche sarebbe esistita⁵⁹. Ora entrambe queste due funzioni possono essere assolve da intelligenze artificiali.

Nel primo capitolo abbiamo suddiviso le tipologie di IA utilizzate per il *News Gathering* in:

- Computer Vision
- Previsione di tendenze
- Ricerca e verifica di informazioni (per combattere disinformazione e misinformazione)

La verifica di informazioni verrà analizzata nel terzo capitolo, nel quale parleremo di *fake news* e *fact checking* e quindi di verifica di informazioni. Si è deciso di dare più spazio a questo tipo di tecnologie e al problema in sé dal punto di vista etico e deontologico in quanto le *fake news*, dalla pandemia in poi, sono parte integrante se non maggioritaria delle notizie che circolano sul web. Inoltre, proprio negli ultimi mesi del 2022, la strategia di Elon Musk non appena acquistato Twitter, è stata quella di dare la possibilità di acquistare il verificato⁶⁰ per il proprio account per combattere i profili falsi: strategia che ha avuto invece il risultato opposto, cioè l’aumentare degli stessi⁶¹.

2.1.1 La Computer Vision

“La visione artificiale è un campo dell’intelligenza artificiale che permette ai computer e ai sistemi di ricavare informazioni significative da immagini digitali, video e altri input visivi - e intraprendere azioni o formulare delle segnalazioni sulla base di tali informazioni. Se l’IA permette ai computer di pensare, la computer vision permette loro di vedere, osservare e capire”⁶². È un tipo di tecnologia basato sui *Convolutional Neural Network*.

⁵⁹Bass, A. Z., *Refining the “Gatekeeper” Concept: A UN Radio Case Study*, *Journalism Quarterly*, 46, 69–72, 1969

⁶⁰È un sistema destinato a comunicare l’autenticità di un account Twitter per distinguere i veri titolari di account di rilievo, come celebrità e organizzazioni, da impostori o parodie

⁶¹Tim Marcin, *People started spreading fake news on twitter the instant they could buy a blue check*, *mashable.com*, 9 nov. 2022

⁶²IBM, *Cos’è la Computer vision*, *ibm.com*

Rispetto all'essere umano, per quanto vi sia ampio dibattito sul fatto se sia qualitativamente migliore, è sicuramente quantitativamente più veloce. È impossibile per l'essere umano analizzare una grande mole di dati visuali in poco tempo. La Computer Vision riesce in pochissimo tempo ad analizzare innumerevoli immagini e video che richiederebbero giorni. Queste tecnologie necessitano di un ampio dataset di immagini da cui trarre le informazioni basilari per poter operare (distinzione colori, oggetti ecc.). Più è ampio il dataset, più sarà accurato il software. Si è deciso di inserire questo tipo di tecnologie nella fase antecedente la creazione della notizia poiché la maggior parte di esse vengono utilizzate per questo scopo. Quelle di analisi video e riconoscimento facciale sono invece ibride secondo questa categorizzazione, poiché possono anche essere utilizzati posteriormente ad una notizia, ad esempio se si commenta o analizza quanto visto, o essere utilizzate nel mentre di una notizia, come ad esempio se si affiancano a dirette streaming.

Il termine racchiude in sé tutte le tecnologie di analisi visuale, da quelle satellitari ai Qr Code. Rientrano nelle Tecnologie Computer Vision (CV) utilizzate per il giornalismo:

- Tecnologie per la classificazione di immagini: analisi di contenuti e attribuzione di etichette sulla base di determinate caratteristiche
- Tecnologie di riconoscimento di immagini per la ricerca di notizie
- Tecnologie di riconoscimento motorio: analisi video e riconoscimento di azioni
- Tecnologie di riconoscimento facciale: analisi video e riconoscimento dei volti attraverso determinate caratteristiche

Le tecnologie di classificazione di immagini possono essere utili per l'archiviazione di grandi quantità di immagini. Se ad esempio un giornalista deve catalogare un set di foto scattate da un reporter sul campo o da un fotografo ad un evento importante, attraverso questi software potrà farlo in maniera automatica settando solo pochi input iniziali.

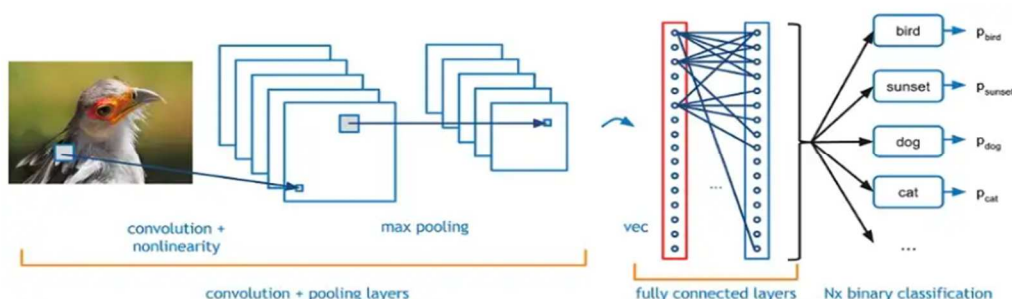


Fig.7 Funzionamento di un sistema di classificazione immagini⁶³

⁶³James Lee, *The 5 Computer Vision Techniques That Will Change How You See The World*, medium.com, Hearbeat, 12 aprile 2018

Per quanto riguarda le tecnologie di ricerca di immagini esse possono essere utilizzate soprattutto per la ricerca automatizzata di notizie sul web. Il funzionamento è simile a quello per la classificazione ma cambia lo step finale: invece di raggruppare le foto sotto una stessa etichetta, si ricercano attraverso quell'etichetta immagini simili sul web. Lo stesso Google Immagini, mentre prima basava le proprie ricerche su etichette apposte da chi caricava un'immagine sul web⁶⁴, utilizza ora un programma di Computer Vision per la ricerca di immagini. Questi software di ricerca sono utilizzati anche per immagini satellitari. Queste immagini si differenziano da quelle comuni perché meno accessibili, ma hanno potenzialmente maggiore valore informativo in quanto riguardano fenomeni di più ampia portata. Possono essere utilizzate per scopi giornalistici come il monitoraggio automatico del cambiamento climatico, della deforestazione, di zone di guerra o l'accesso ad immagini di luoghi inaccessibili all'uomo a causa di catastrofi naturali come incendi o terremoti (per questo scopo possono essere utili anche i droni⁶⁵). Spesso queste strumentazioni sono utilizzate parallelamente a dei sensori per il monitoraggio di determinati eventi. Un esempio di tali sensori è l'utilizzo di FeinstaubeRadar, un radar utilizzato nella città di Stoccarda a causa dell'inquinamento da polveri sottili. È un radar formato da 750 sensori apposti sulle case dei cittadini per rilevare costantemente e in tempo reale dati sulle polveri. Un software di scrittura automatica redige a fine giornata un report da destinare alle comunità locali. Le relazioni vengono aggiornate dagli umani con l'aggiunta di dati ufficiali di un'agenzia statale. Questo radar antinquinamento ha vinto il Konrad Adenauer Prize per il giornalismo locale.

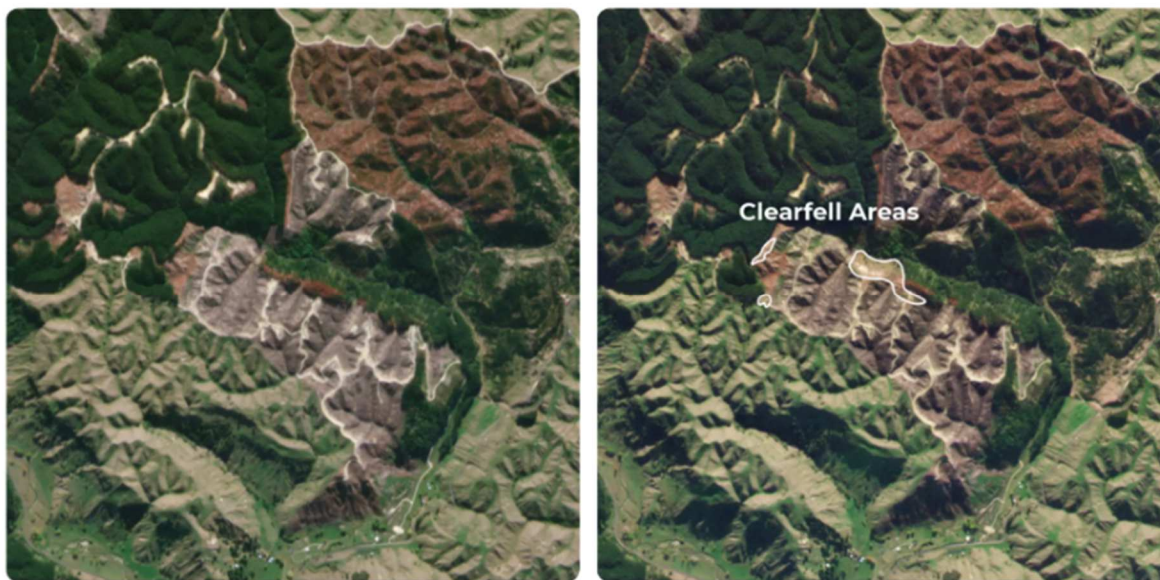


Fig.8 Monitoraggio deforestazione attraverso immagini satellitari⁶⁶

⁶⁴Gli Alt Tag sono degli attributi di codice che permettono di far sì che Google riconosca le immagini attraverso una breve descrizione di esse. Vengono ancora utilizzati per indicizzare in maniera ancora più precisa le immagini, parallelamente alla nuova tecnologia

⁶⁵In letteratura si parla anche di Aerial journalism e Drone journalism, un tipo di giornalismo volto ai report video di situazioni inaccessibili all'uomo fisicamente e raggiunte attraverso i droni.

⁶⁶*Powering Forestry 4.0: Planet Provides High-Frequency Insights Across Distributed Forest Assets*, Planet

È doveroso ricordare che gli sviluppi delle intelligenze artificiali in generale e della Computer Vision nello specifico, sono avvenuti negli ultimi anni. Nonostante i primi prototipi siano nati qualche decennio prima dei prototipi dei software di scrittura automatica, la Computer Vision è la sottobranca che ha avuto più difficoltà a svilupparsi. Negli anni Sessanta, come detto, l'IA divenne una disciplina accademica. È in questo periodo che Seymour Papert, professore del laboratorio di IA dell' MIT, decise di lanciare il Summer Vision Project e di risolvere in pochi mesi il problema della visione artificiale. Era convinto che un piccolo gruppo di studenti del MIT fosse in grado di sviluppare una parte significativa di un sistema visivo in un'estate. Gli studenti dovevano progettare una piattaforma in grado di eseguire automaticamente la segmentazione sfondo/primo piano e di estrarre oggetti non sovrapposti da immagini reali. Il progetto non ebbe successo. Tuttavia, quel progetto è stato, secondo molti, la nascita ufficiale della Computer Vision come campo scientifico⁶⁷.

Nonostante i recenti notevoli progressi, non siamo ancora vicini ad una Computer Vision senza errori. Tuttavia, sono già numerose le istituzioni sanitarie e le imprese che hanno trovato il modo di applicare i sistemi di CV, alimentati dalle CNN, ai problemi del mondo reale.

Le percentuali di errore dei software sono comunque notevolmente diminuite negli anni (Figura 9) ma non mancano esempi di errori grossolani, come quello commesso dal software JanetBot che non ha saputo distinguere il genere delle persone rappresentate in foto⁶⁸.

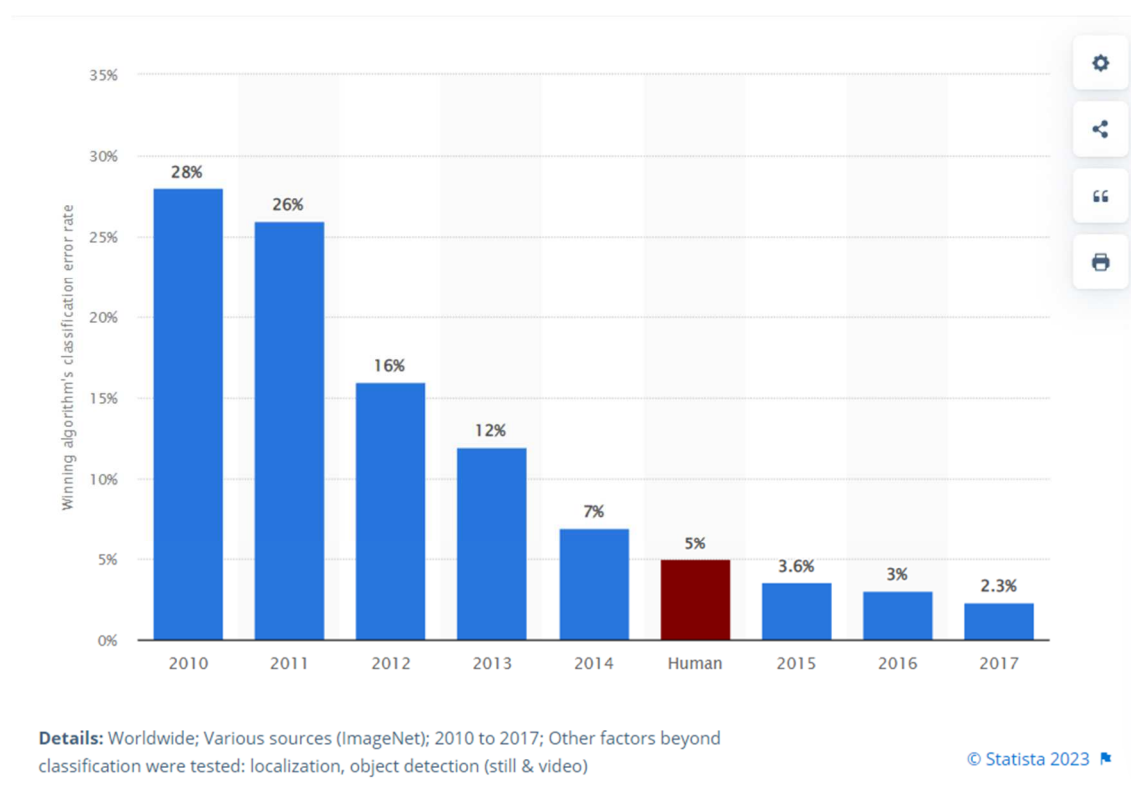


Fig.9 Miglioramento degli Error Rates dei software negli anni

⁶⁷Rostyslav Demush, *A brief history of Computer Vision (and Convolutional Neural Networks)*, Hackernoon, 26 febbraio 2019

⁶⁸JanetBot: *Analysing gender diversity on the FT homepage*, Financial Times Labs, 28 novembre 2022

Le tecnologie di riconoscimento motorio, insieme a quelle di riconoscimento facciale, sono tecnologie di analisi video e a differenza di quelle precedentemente menzionate sono più complesse, in quanto assimilabili ad un'analisi di innumerevoli immagini al secondo.

Le tecnologie di riconoscimento motorio analizzano i movimenti di un determinato soggetto andando a fornire dati che l'uomo da solo avrebbe difficoltà o impiegherebbe troppo tempo a sviluppare. Pensiamo alle analisi sportive. Durante una partita di calcio le app che segnalano live le statistiche della partita e dei singoli giocatori spesso ormai utilizzano questo tipo di software: riconoscono i giocatori e il tipo di azione di gioco che stanno svolgendo. Nella ginnastica è ancora più utilizzato poiché i punteggi dipendono proprio dal tipo di movimenti e da cambiamenti infinitesimali tra un movimento o un altro. Il New York Times qualche anno fa, in vista delle Olimpiadi del 2020, ha implementato Software di Computer Vision per monitorare le prestazioni delle atlete.

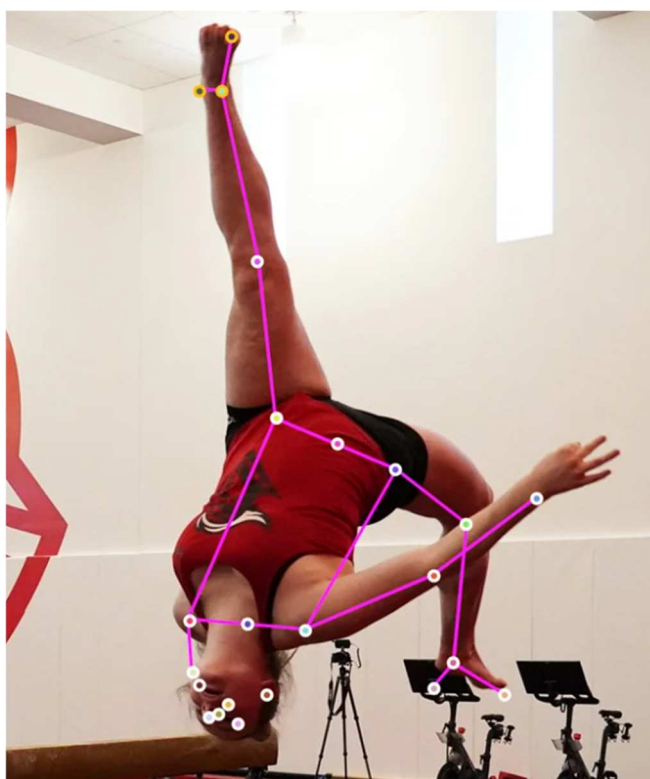


Fig. 9 Computer Vision utilizzata per monitorare ginnaste statunitensi⁶⁹

Le tecnologie di riconoscimento facciale sono spesso basate su software di analisi dei componenti principali⁷⁰ e riescono a identificare i volti delle persone, facendoli riscontrare con immagini simili e riuscendo quindi a identificarne l'identità. È intuitivo pensare che possano esserci implicazioni per la privacy delle persone, ma approfondiremo tale tema più avanti. Queste tecnologie possono essere utilizzate nell'ambito del giornalismo in diverse situazioni. Sono esempi emblematici di tale tecnologia due progetti avviati dal New York Times.

⁶⁹Dan Oved, Amelia Pisapia, Anna Gudnason, *Estimating 3D Poses of Athletes at Live Sporting Events*, R&D New York Times, 16 giugno 2020

⁷⁰Vedi Capitolo 1 Paragrafo 1.1.2

2.1.1.1 L'utilizzo della Computer Vision al New York Times

Il New York Times iniziò a sperimentare software di Computer Vision nel 2017. Il primo esperimento fu effettuato su una foto dell'inaugurazione del Presidente Trump. La foto poteva essere visualizzata in maniera interattiva con riconoscimento facciale dei presenti, potendoli anche filtrare per categorie. La seconda sperimentazione fu l'implementazione di un software capace di identificare i membri del Congresso americano tramite una singola foto. I membri del Congresso sono spesso poco conosciuti nella loro totalità perché vengono eletti ogni due anni a causa delle elezioni di metà mandato⁷¹. Il software aveva quindi l'obiettivo di mettere a disposizione l'identificazione immediata e fornire informazione sui singoli politici del Congresso americano. Era basato sulla tecnologia "Amazon Rekognition"⁷², specificatamente sulla sua sezione "RecognizeCelebrity" che utilizza foto di celebrità in generale. A causa di alcuni bias e di alcune foto oscurate in quanto scattate direttamente al Congresso, dove l'illuminazione non è altissima e dove si è impossibilitati ad avvicinarsi ai politici, inizialmente il software commise alcuni errori, confondendo alcuni politici con altre celebrità. Dopo alcune modifiche, tra cui l'utilizzo di foto ben definite ed illuminate, l'ampliamento del database e l'opportuna segnalazione di similitudini dell'aspetto di alcune celebrità, il software divenne più preciso. Questa tecnologia aiutò il New York Times a scovare notizie che altrimenti non sarebbe stato in grado di reperire. Il software si chiama "Who The Hill" e la sua formattazione è Open Source⁷³.



Fig. 10 Mitch McConnell Riconoscimento Facciale (R&D New York Times)

⁷¹Le elezioni di midterm si tengono a metà mandato presidenziale al fine di mettere alla prova dell'opinione pubblica l'operato dei primi anni del Presidente e del Congresso

⁷²Amazon Rekognition è un software Computer Vision lanciato da Amazon nel 2016

⁷³I software Open Source sono software decentralizzati basati sulla condivisione dei file e dei codici. Sono accessibili a tutti e ne sono concessi lo studio, l'utilizzo, la modifica e la redistribuzione.

L'utilizzo che è stato fatto di questo Software non si è fermato qui. Una volta assicurati i diritti per la diretta Live Stream del Royal Wedding tra Meghan e Harry nel 2018⁷⁴, Il New York Times ha deciso di tentare l'applicazione di questa tecnologia anche alla diretta video per infotainment. Dopo vari tentativi effettuati su video di altri matrimoni reali precedenti, il team di Research & Development⁷⁵ del New York Times ha deciso di aumentare l'esperienza degli utenti che avrebbero visto la diretta, con informazioni sugli invitati che andassero oltre la semplice identificazioni e fornissero dettagli sul rapporto con gli sposi. Dal momento che la lista ufficiale degli invitati non esisteva, collaborando con altre aziende, stilano una lista per eccesso di potenziali invitati andando a scovare le informazioni necessarie e immettendo nel database più foto possibile per poterli identificare con maggiore accuratezza. Il risultato fu un grande successo. Inoltre, se un utente aveva già visualizzato il nome e le informazioni di un ospite, alla comparsa successiva dello stesso ospite tali informazioni non si sarebbero riproposte. Il team ha dichiarato di aver rispettato il codice etico del giornale e di aver quindi aumentato l'esperienza di visione tramite questo software solo per personaggi famosi e maggiorenti.

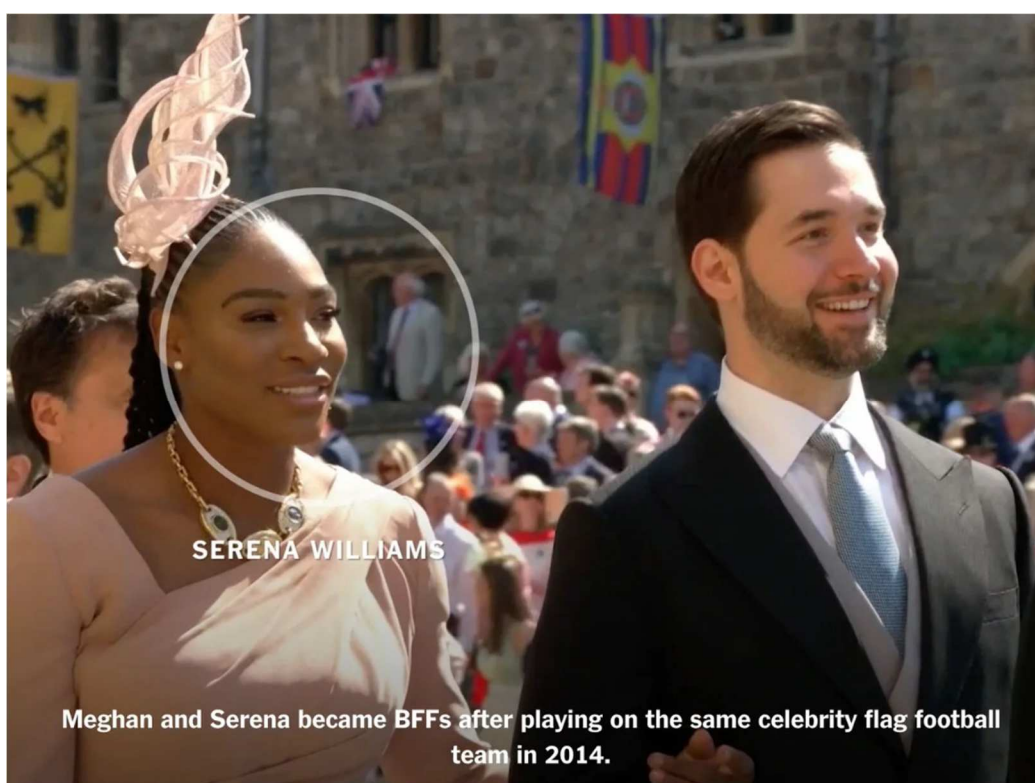


Fig.11 Riconoscimento Facciale di Serena Williams al Royal Wedding di Harry e Meghan (R&D New York Times)

La stessa tecnologia è stata proposta da Sky News con l'eccezione però che il testo visualizzato indicava unicamente il nome e il cognome dell'invitato, senza le motivazioni della sua presenza al matrimonio. Mostrava però le informazioni principali sulla vita dell'invitato e sul perché fosse famoso.

⁷⁴L'evento è stato visto in diretta da 31 mln di persone (fonte: startupitalia.eu)

⁷⁵Research and Development è il gruppo di ricerca e sviluppo del New York Times. A detta dello stesso team, il loro obiettivo è lo "sviluppo di tecnologie innovative al servizio della missione della nostra azienda di cercare la verità e aiutare le persone a capire il mondo"

2.1.2 Ricerca di informazioni e previsione di tendenze

La ricerca di immagini, come abbiamo visto, può essere utile anche per la ricerca di notizie sul web. Le immagini satellitari spesso aiutano a scovare eventi o cambiamenti di ampia portata. La ricerca di informazioni però non avviene solo attraverso l'analisi di pixel ma anche tramite elementi testuali.

Pensiamo ai motori di ricerca. Essi potrebbero essere tranquillamente etichettati come software di ricerca di notizie che, attraverso le cosiddette *keywords*⁷⁶, scandagliano al meglio il web per trovare qualcosa. Ovviamente svolgono tale ricerca nella totalità del web e non solo per le notizie. Se cercassi la parola "terremoti" probabilmente troverei tra i primi risultati della pagina di ricerca la definizione di terremoto, il link dell'istituto di geofisica o comunque informazioni rilevanti che gli algoritmi di Google ritengono utili alla nostra ricerca. Esistono software però che in maniera più complessa e accurata scovano notizie inerenti un determinato argomento e sono pensati per affiancare il giornalista nella ricerca di una notizia. Anzi, possono addirittura scovare quelle che saranno le notizie di domani. Si è ripetuto più volte in questa sede che le intelligenze artificiali non sono perfette. Commettono errori e si basano spesso su calcoli di probabilità. L'esempio emblematico di questo rischio è sicuramente la previsione del futuro. L'uomo può fare congetture, dire la sua sull'evolversi di una determinata situazione e provare a prevedere cosa accadrà sulla base di sensazioni. Le macchine possono fare lo stesso, ma sulla base di dati ed evoluzioni repentine di situazioni, inferendo su situazioni simili avvenute precedentemente. Prevedere il futuro è complicato per natura e solo di rado i giornalisti si avventurano in previsioni. Analizzeremo qui alcuni tipi di software che, sulla base di analisi web e social, prevedono quali saranno gli argomenti più caldi di un determinato periodo.

È intuibile pensare che tali strumenti possono essere utilissimi ad un giornalista al fine di rendere un proprio contenuto più desiderabile per i propri lettori. Capire quale potrà essere in termini probabilistici l'argomento più caldo nel breve termine può sicuramente rendere le *hard news* più accattivanti. Le *hard news* sono per definizione news che attecchiscono nel breve termine. Si differenziano dalle *soft news* perché l'interesse di questo tipo di notizie è forte ma di breve durata, mentre quello delle *soft news* è basso ma che rimane relativamente costante nel tempo. Le *breaking news* invece si discostano da tale dicotomia perché sono notizie particolarmente importanti per la loro urgenza, come lo scoppio di una guerra, un terremoto o altre catastrofi naturali.

La scelta di una notizia, prima dell'avvento di alcune tecnologie, avveniva tramite ricerche e sensazioni del giornalista. Spesso si aspettava la notizia da agenzie o da fonti dirette che il giornalista conosceva. Ora algoritmi e software possono consigliarci quale sia l'argomento migliore di cui parlare e quale sarà l'engagement del nostro articolo nel tempo.

⁷⁶Le *keywords* sono le parole chiave che permettono di individuare un determinato tipo di contenuto. Un esempio classico è quello delle parole che immettiamo nella barra di Google quando cerchiamo qualcosa

2.1.2.1 Asimov

Nicola Grandis, informatico abruzzese, insieme ad un gruppo di esperti e nuove leve, decise all'indomani della pandemia di lanciare Asc27, una startup con sede a Roma che si sarebbe occupata di intelligenze artificiali e cybersecurity. Nata nel marzo 2020 vincerà poi un premio già nel luglio 2021 al World Artificial Intelligence Congress a Shanghai come migliore startup europea e si classificherà nella top ten mondiale, sempre per la stessa categoria, insieme ad altre nove startup cinesi. Nell'ambito della ricerca e della sperimentazione di intelligenze artificiali, affiancati da un team di giornalisti, hanno creato Asimov⁷⁷, effettuando test tramite un software pilota per una rivista. Il programma è progettato per accompagnare e assistere il giornalista in tutte le fasi di creazione di una notizia ma ci concentreremo qui su quella iniziale, nello specifico nella previsione di tendenze. Il software è utilizzato anche da molti politici. Il software funziona in maniera simile ad un software di previsioni finanziarie, o ai software che identificano i truffatori di scommesse.

Il processo di analisi e scoperta di potenziali tendenze può essere così consequenzialmente riassunto:

1. Analisi web e social di testi, immagini e video
2. Analisi dell'audience del giornale
3. Suggerimenti su potenziali tendenze sulla base delle analisi pregresse
4. Assistenza alla scrittura
5. Suggerimenti sulle foto più adatte al tipo di articolo scritto
6. Previsione dell'engagement dell'articolo scritto

La descrizione del processo ci dice che Asimov assiste il copywriter anche in altre mansioni precedentemente approfondite. La funzionalità che ci interessa sottolineare qui è però il suggerimento dei potenziali trend. L'obiettivo è quello di scovare gli scoop dei prossimi giorni prima che arrivino su Google trend. Dopo vari studi e test si è potuto constatare che l'accuratezza delle previsioni effettuate in generale, quindi sia delle potenziali tendenze che dell'engagement del contenuto, varia dall' 85% all'88%⁷⁸.

Marco Pratellesi, giornalista della Repubblica, ha intervistato un anno e mezzo fa Nicola Grandis⁷⁹.

“I giornalisti spesso aspettano le notizie dall'agenzia o hanno una dritta da un amico o da una fonte. Quando il giornalismo entrerà completamente nel mondo dell'intelligenza artificiale avrà sullo schermo dei programmi di trending topic che potrà interrogare per venire a conoscenza di potenziali notizie e capire quali dovrà seguire. Certo, se improvvisamente Elon Musk decidesse di vendere la Tesla, questo non potremo prevederlo con la IA, ma a livello di trending topic e di scrittura ci sarà un potenziamento incredibile che potrà aiutare il giornalista anche a capire come il suo pubblico comprende meglio le notizie”

⁷⁷Il nome ha preso ispirazione dalla figura di Isaac Asimov, importante divulgatore scientifico del XX sec. nonché ideatore delle tre leggi della robotica

⁷⁸Alessandro Di Stefano, *Asc27, non è il robottino di Star Wars, ma una startup che porta l'IA in redazione*, StartupItalia, luglio 2021

⁷⁹Marco Pratellesi, *Asimov, intelligenza artificiale e voci neurali: la rivoluzione del giornalismo parte da Roma*, La Repubblica, 19 agosto 2021

2.2 Le intelligenze artificiali per la news production

La fase di creazione vera e propria del contenuto è la più importante del processo. Il modo in cui si scrive un articolo, si crea un podcast o si registra un servizio televisivo sarà fondamentale per la qualità della notizia. I software utilizzati in questa fase sono molteplici e abbiamo suddiviso in maniera generale e sintetica questa moltitudine in tre categorie. Quella con più sviluppi è sicuramente la scrittura automatica, che abbiamo approfondito concettualmente parlando di *automated journalism* e *automated news* nel primo capitolo. La fase di produzione vera e propria è anche quella eticamente e professionalmente più discussa. “Le IA sostituiranno i giornalisti?” “Chi ha scritto questo pezzo un giornalista o un’IA?” Sono solo alcune delle domande che si ripetono quando si parla di questo ambito. Approfondiremo gli aspetti etici e deontologici nel prossimo capitolo. Ora approfondiremo di più l’oggetto della questione.

Nel primo capitolo abbiamo suddiviso le tipologie di IA utilizzate per la *News Production* in:

- Software audio o video
- Scrittura automatica
- Deep Fake

2.2.1 Software audio e video

Questo tipo di software in realtà può essere utilizzato anche prima o dopo la creazione effettiva della news. Se devo trascrivere un’intervista, se voglio inserire la possibilità per un lettore online di sentire l’audio di un articolo piuttosto che leggerlo o se voglio creare un podcast partendo da un testo scritto, oggi si può fare con l’utilizzo di intelligenze artificiali. Anche negli strumenti che utilizziamo tutti i giorni ci sono strumentazioni simili: proprio mentre scrivo queste parole, invece di scriverle, potrei dettarle a Word che trascriverebbe, con tanto di virgole e punti, le parole che dico seguendo la sintassi indotta dalla mia intonazione. Esistono numerose app che svolgono funzioni simili, come Voice to Text, che supporta 150 lingue e può trascrivere anche file audio. Nella miriade di applicazioni, software e algoritmi che possono essere utilizzate per cambiare formato, quella che sta trovando negli ultimi tempi maggiore applicazione è sicuramente la produzione automatica di podcast. Partendo da un set di dati, tali strumenti permettono di tradurre in automatico prima i dati in testo e poi il testo in audio. È sostanzialmente l’unione di un software di scrittura automatica e un software text-to-speech.

I software video, invece, sono utilizzati per ottenere una migliore qualità visuale e per poter formattare i video senza che occupino troppo spazio. Le potenzialità della qualità video in generale si sono espresse al loro massimo negli ultimi anni, raggiungendo praticamente i livelli della visione umana. All’aumentare della qualità però, aumenta anche l’impellenza di processi che riescano a codificare i video in maniera opportuna, trovando un equilibrio tra qualità visiva e bit-rate⁸⁰.

⁸⁰Indicano la quantità di dati digitali che possono essere trasferiti, attraverso una connessione/trasmissione, su un canale di comunicazione in un dato intervallo di tempo

L'approccio tradizionale di codifica video non è più idoneo per video ad alta qualità e queste limitazioni hanno portato i ricercatori a rivolgere la loro attenzione alle intelligenze artificiali, nello specifico alle reti neurali. Le reti neurali hanno avuto eccellenti prestazioni in elementi importanti della codifica video. Molti di questi studi sono Open Source, ma spesso l'applicazione non è comunque di dominio pubblico perché i costi sono elevati e i dati necessari a comporre il dataset iniziale degli input del sistema sono molto più complessi e strutturati di quelli delle immagini. Inoltre, molti di questi software sono addestrati a partire da dati sintetici, creati cioè artificialmente, e che quindi riscontrano non poche complicazioni nell'applicazione reale. Nonostante tali difficoltà, è un campo in grande crescita e vengono effettuate sempre più sperimentazioni per raggiungere standard adeguati.

2.2.1.1 AI Anchor del Sole24Ore

AI Anchor è un podcast automatico del Sole24Ore. Creato all'interno di una collaborazione con Dataninja⁸¹ e con il supporto dell'Università di Pisa per le tecnologie del linguaggio naturale (NLG), posta ogni giorno notizie sulle chiusure dei mercati in Italia e nel mondo.

In tre minuti riporta le chiusure dei mercati finanziari e il cambio dei rapporti tra le valute principali. A detta degli stessi creatori del software, esso cerca di trovare il giusto equilibrio tra accuratezza del contenuto e naturalezza della forma, con particolare riguardo per la prima dal momento che si tratta sostanzialmente di un report di dati. Tali dati sono forniti direttamente dal Sole24Ore. I dati appena menzionati vengono arricchiti con ulteriori dati esterni per garantire un'opportuna contestualizzazione per il podcast. La produzione automatica di podcast è prodotta da due tecnologie contigue che trasformano i dati in testo e il testo in audio in maniera automatica. All'interno del software di scrittura vi è una sorta di modello di ricerca semantica che, per gestire tutte le informazioni e le indicazioni grammaticali, opera in maniera simile ad un motore di ricerca. Il sistema in generale è molto flessibile nella scelta dei termini. Sono impostati dei template che vengono modificati in base a quanto detto finora e al termine del processo vi è anche un servizio automatico di correzione grammaticale. Questo è un estratto del testo di una puntata.

Piazza affari ha chiuso la giornata di scambi con andamenti in calo. L'indice FTSE Mib ha registrato una variazione di -1,45% a 23.706,96 punti. L'indice FTSE All Share è in ribasso dell'1,4% a 25.719,37 punti. Le Mid Cap hanno chiuso in calo dell'1,07%. Il mercato delle Small Cap ha terminato la seduta in lieve calo dello 0,06%. Il segmento Star ha perso posizioni con un -0,82%. Tra le altre principali borse europee: Londra -0,81%, Parigi -1,52%, Madrid -1,13%, A Wall Street, il Dow Jones è in calo a 32.975,38 punti, -0,74%. Il Nasdaq è in ribasso a 10.841,49 punti, -1%. In rosso lo Standard and Poor's 500 a 3.817,87 punti, -0,82%. Sulle piazze asiatiche, chiusura poco mossa per il Giappone. L'indice Nikkei della borsa di Tokyo ha mantenuto le posizioni: 0% a 26.094,5 punti. In leggera contrazione dello 0,19% il Topix a 1.891,71 punti. In Cina, la piazza di Shanghai ha chiuso la giornata in territorio positivo, con una crescita dello 0,51%.

⁸¹È un'azienda nata in Italia nel 2012 con l'obiettivo di aiutare le persone a lavorare meglio usando i dati

2.2.2 La scrittura automatica

Le intelligenze artificiali nel giornalismo e la scrittura automatica nello specifico sono spesso identificate, nella letteratura scientifica e sul web, con altri termini. Il più utilizzato è *Natural Language Generation* che, come sottolineato nel primo capitolo, si focalizza sulla parte finale del processo. I software sono molto più complessi perché la parte iniziale, quella di *Natural Language Understanding*⁸², sarà fondamentale per la complessità e l'accuratezza del testo finale.

La scrittura automatica nacque e venne utilizzata inizialmente per lo più per report di dati. Da StatsMonkey sino all'esempio appena menzionato di AI Anchor, la capacità di un'intelligenza artificiale di analizzare una mole di dati e tradurli in testo è stata anche più accurata di quella umana. Le redazioni difficilmente scrivevano storie approfondite, creative o di opinione attraverso i neonati software. Le *machine-written news* si limitavano appunto al report di dati, anche per il timore che la complessità di un articolo che richiedesse una sintassi più articolata e accattivante non fosse alla portata di tali tecnologie. Il rischio era anche quello di scrivere articoli identici ad altre testate che usavano software simili, in quanto difficilmente le redazioni avevano un archivio ampio di propri dati raccolti differenti da quelli accessibili a tutti. Si prevedeva però, già allora, che all'aumentare degli investimenti e delle capacità dei software, essi sarebbero stati capaci di attingere da più fonti di dati con un miglioramento generale in accuratezza e complessità. Ovviamente questi miglioramenti sono avvenuti principalmente in chi ne ha avuto la possibilità, cioè nelle redazioni importanti. Per le testate locali spesso non c'è stata la possibilità economica di implementare software complessi o di sostenere il capitale umano necessario all'uso di essi. Con l'avanzare degli anni e una maggiore sostenibilità economica di tali software lo scenario è cambiato: l'utilizzo di questi software ha permesso di contenere i costi sia in termini di investimenti nelle tecnologie che di personale. Tralasciando per il momento analisi di natura deontologica e professionale sull'uso di tali tecnologie⁸³, possiamo affermare con certezza che oggi giorno queste tecnologie permettono di contenere i costi anche per le testate meno importanti. Sono quindi disponibili ora software che riescono a garantire degli standard minimi di "qualità giornalistica"⁸⁴ con costi accessibili anche alle *small newsroom*. Coloro i quali non riescono ad avere accesso a queste tecnologie non lo fanno per lo più per motivi culturali, cioè di percezione negativa delle intelligenze artificiali, o di mancanza di personale qualificato per implementare tali tipi di software, non per un costo eccessivo di essi. La figura professionale del futuro, infatti, sarà un giornalista competente anche dal punto di vista tecnico-informatico, capace di sfruttare al meglio le potenzialità delle tecnologie senza snaturare il lavoro umano di approfondimento e opinione di se stesso. Approfondiremo più avanti tali aspetti. Ora ci concentreremo sull'esempio forse più sorprendente di utilizzo delle intelligenze artificiali nel giornalismo.

⁸²Vedi Capitolo 1 paragrafo 1.3 sottoparagrafo 1.3.1

⁸³Vedi Capitolo 4

⁸⁴Vedi Capitolo 4 paragrafo 4.3

2.2.2.1 GPT-3

L'8 settembre 2020 il The Guardian pubblica un articolo interamente scritto da *Generative Pre-Trained Transformer 3*, meglio conosciuto come GPT-3. Al software è stato richiesto di scrivere un breve editoriale di 500 parole con linguaggio breve e conciso cercando di concentrarsi sul fatto che gli umani non abbiano nulla da temere dalle intelligenze artificiali. Il titolo, sempre scelto da GPT-3, è stato "A robot wrote this entire article. Are you scared yet, human?". Il risultato finale però non è stato interamente merito del software: partendo dagli otto output che il software aveva fornito, dei giornalisti in carne ed ossa hanno scelto le parti migliori di ognuno e "assemblato" il pezzo finale. Tutti gli articoli ricercati sull'argomento ci tengono a sottolineare come nonostante sia stato necessario l'intervento umano, c'è comunque stato un guadagno in termini di tempo. Intervenire su testi già scritti per modificarli o unirli impiega sicuramente meno tempo che scriverli di sana pianta. In molti poi sottolineano il tono inquietante dell'articolo, come esso ci tenga a rassicurare di non preoccuparsi delle intelligenze artificiali, quasi come se fosse un essere senziente. L'input concettuale è sì simile a quello umano a livello logico-razionale, ma essa non è consapevole di sé stessa e nessuna macchina ha ancora superato il Test di Turing⁸⁵, compresa GPT-3.

Rispetto a GPT-2 a livello sostanziale non è cambiato molto. Le migliorie apportate per l'avanzamento di livello sono per lo più inerenti alle capacità quantitative della macchina. GPT-3 è semplicemente più grande. Si basa su una struttura MLP 100 volte più grande, con 98 strati di *hidden layers*⁸⁶ e 175 miliardi di parametri. È una tecnologia che non necessita di supervisione ed è addestrata su molti più dati rispetto alle versioni precedenti: il 60% dell'addestramento si basa su un set di dati filtrati dell'intero Web dal 2011 al 2021, il 37% da altre librerie di dati e il 3% da Wikipedia. Uno dei punti di forza di questa versione è che è capace di svolgere compiti inusuali con grande successo. Quasi tutte le tecnologie NLP hanno bisogno di essere addestrate e testate. GPT-3, invece, non richiede questo tipo di messa a punto: sembra essere in grado di svolgere ragionevolmente bene un'intera gamma di compiti. Produzione di narrativa, poesie, comunicati stampa, codice funzionante, musica, barzellette, manuali tecnici, fino agli articoli di cronaca, GPT-3 sembra svolgere egregiamente qualsiasi compito. Il salto di qualità è avvenuto grazie ai tipi di modelli implementati, i cosiddetti *Transformer*⁸⁷, e al tipo di set di dati utilizzato (testi con dipendenze a lungo raggio⁸⁸).

⁸⁵Vedi Capitolo 1 sottoparagrafo 1.2.2

⁸⁶Fig. 3 Capitolo 1 Paragrafo 1.1.1

⁸⁷È un modello di apprendimento profondo che adotta il meccanismo dell'auto attenzione, ponderando in modo differenziato il significato di ciascuna parte dei dati in ingresso. I trasformatori sono progettati per elaborare dati di input sequenziali, come il linguaggio naturale, con applicazioni a compiti quali la traduzione e la sintesi di testi. Tuttavia, a differenza delle RNN, i trasformatori elaborano l'intero input in una sola volta. Il meccanismo di attenzione fornisce un contesto per qualsiasi posizione nella sequenza di input. Ad esempio, se i dati in ingresso sono una frase in linguaggio naturale, il trasformatore non deve elaborare una parola alla volta. Ciò consente una maggiore parallelizzazione rispetto alle RNN e quindi riduce i tempi di addestramento (Wikipedia)

⁸⁸La *long-range dependency*, cioè la dipendenza a lungo raggio, è una modalità di correlazione tra dati, in questo caso dati testuali quindi parole, che si riferisce a strutture di dipendenza tra parole che decadono lentamente con l'aumentare della distanza. Essendo a lungo raggio ha una capacità maggiore di correlare le parole in frasi anche più lunghe e implementare questo modello ha permesso una maggiore accuratezza e complessità

Uno dei problemi di questa tecnologia, o meglio, di OpenAI, la società no-profit fondata nel 2015 da Elon Musk e Sam Altman che ha progettato GPT-3, è che non è più no-profit⁸⁹. Alla sua creazione la società si finanziava attraverso un investimento di un miliardo di dollari da parte dei fondatori stessi, nonché per mezzo di ingenti finanziamenti da parte dei co-fondatori di PayPal, LinkedIn e da Amazon Web Services, lo stesso che ha fornito le tecnologie Computer Vision negli esempi precedentemente menzionati. Nel febbraio 2018 Elon Musk ha lasciato il CdA della società ma continua comunque a fare donazioni e fornire consulenza all'organizzazione. Nel 2019 è divenuta una società a scopo di lucro e ha raccolto un miliardo di dollari da Microsoft per finanziare i progetti e le sperimentazioni. GPT-3 è stato il primo prodotto commerciale della società. Una versione API (Application Programming Interface), cioè accessibile e usabile, è stata messa a disposizione degli utenti che ne facevano richiesta, con il nome di ChatGPT. La licenza è di esclusiva proprietà di Microsoft, ma la versione API è ormai divenuta accessibile a tutti sul sito ufficiale di OpenAI. Nelle prime settimane del 2023 ha avuto talmente tanto successo che i server non hanno retto la quantità di utenti che usavano tale tecnologia. Visitando oggi il sito si legge *“ChatGPT is at capacity right now”* e si visualizzano esempi diversi ad ogni visita senza però la possibilità di utilizzarla. A differenza delle sue versioni precedenti è meno soggetta a bias razzisti o misogini, dal momento che parte dei dati del dataset con cui è stata programmata sono venuti da test effettuati da una società kenyota. I test venivano effettuati su immagini e testi da valutare come razzisti e malevoli o da potersi considerare accettabili⁹⁰. Tralasciando per ora considerazioni sulla giustizia di tale procedimento, si può affermare con certezza che GPT-3 e la sua versione API ChatGPT siano i software di scrittura automatica migliori mai creati dall'uomo fino ad ora. In molti si pongono addirittura il dilemma se GPT-3 si avvicini ad essere AGI (Artificial General Intelligence), termine usato per indicare quei tipi di intelligenza artificiale che si avvicinano al concetto di IA forte⁹¹. Approfondiremo tutto questo nei prossimi capitoli utilizzando i parametri comunemente usati e provando ad usarne di altri per le altre tipologie di IA, per quanto possibile e per quanto disponibile in letteratura scientifica. Non potendo utilizzare un esempio di testo scritto da ChatGPT-3, riportiamo un esempio di chi ha provato questa versione appena resa disponibile da OpenAI⁹².

“Scene: A small dingy office, early morning, furniture of the Salvation Army store variety. Sordid atmosphere. Harry Potter, in ratty tweed suit, unpressed shirt, and unshined shoes, sits behind the desk looking haggard, ruffled, and embittered. In one corner of the office a torn shade rattles at the wind gusts outside. Harry has one long index finger on the butt of a .38 and he broods over a photograph. The photograph shows a gorgeous blonde wearing an evening gown. She is holding a cocktail in one hand and looking pensively at a cigarette in the other. A tear adorns her cheek. Harry looks at her for a long moment, sighs, then puts his gun down and stands. He straightens his clothes, scratches his head, then with more resolution turns around, picks up his pistol, and leaves the office. He passes through the front room and looks out the dirty window. An old woman is washing the sidewalk outside with a mop. A young man in a double breasted gray suit is leaning against the building. Harry sighs again and goes out the door. He walks up to the young man and without ceremony punches him in the jaw.”

⁸⁹Sam Sheard, *Why everyone is talking about the A.I. text generator released by an Elon Musk-backed lab*, CNBC, 23 luglio 2020

⁹⁰Billy Perrigo, *Exclusive: OpenAI Used Kenyan Workers on Less Than \$2 to Make ChatGPT Less Toxic*, TIME, 18 gennaio 2023

⁹¹Vedi Capitolo 1 Paragrafo 1.1.1

⁹²L' input fornito è stato *“Below is a screenplay for a film-noir hard boiled detective story by Raymond Chandler about the boy wizard Harry Potter: “Harry Potter” By Raymond Chandler*

2.2.3 Deep Fake

I *deepfakes* sono video falsi creati utilizzando software digitali simili alla Computer Vision basati sul *deep learning* e il *face swapping*⁹³. Le più moderne versioni sfruttano reti antagoniste generative⁹⁴ per creare video artificiali in cui le sembianze di una persona vengono sostituite con quelle di un'altra. La tecnologia *deepfake* è in grado di creare immagini convincenti ma completamente fittizie partendo da zero. I *deepfakes* vengono utilizzati per diversi scopi, dalla pornografia alla disinformazione politica, dal cyberbullismo ai crimini informatici, sino alle *fake news*. Le persone hanno iniziato a conoscere la tecnologia *deepfake* quando un utente di Reddit chiamato "Deepfakes" ha affermato di aver sviluppato un algoritmo di apprendimento automatico in grado di trasporre i volti delle celebrità senza soluzione di continuità nei video porno. Naturalmente sono stati forniti dei campioni e la sperimentazione è diventata presto molto popolare. Gli amministratori del sito hanno dovuto chiuderlo nel febbraio 2018, ma ormai la tecnologia era nota e disponibile. Nonostante il divieto, e azioni simili da parte di siti come Twitter e Pornhub, i *deepfakes* erano ancora in circolazione. I loro fan più accaniti li condividevano su altri siti di social media, oltre a una manciata di siti web e servizi di hosting dedicati. Ben presto la gente la utilizzò per creare video falsi, per lo più con protagonisti politici e attori.

Tuttavia, l'idea di manipolare i video non è nuova. Come detto, già negli anni Novanta alcune università conducevano importanti ricerche accademiche sulla Computer Vision. Nel 1997 Christoph Bregler⁹⁵, insieme a due colleghi, ha avviato un progetto per far dire a persone parole che non hanno mai detto e farlo sembrare vero. Le tecnologie utilizzate per tale progetto non erano sofisticate come quelle attuali. Veniva più o meno "ritagliata" graficamente la parte della bocca e venivano sovrapposti frame di video di addestramento aventi labiali per diverse parole. Quando non erano disponibili per determinate parole, il programma selezionava l'approssimazione più simile. A detta degli stessi progettatori⁹⁶ lo scopo della tecnologia madre del *deepfake* era in realtà fornire nuove forme di doppiaggio nel mondo cinematografico.

⁹³Il *face swapping* è una funzione che permette di sovrapporre le caratteristiche di una faccia su un'altra persona in tempo reale avendo solo alcuni dati visuali come foto o altri video della persona che si vuole visualizzare.

⁹⁴Una rete generativa avversaria (GAN) è una classe di framework per l'apprendimento automatico ideata da Ian Goodfellow e dai suoi colleghi nel giugno 2014. Si basa su un conflitto tra due reti neurali (nello specifico recurrent neural network) nella forma di un gioco a somma zero, in cui il guadagno di un agente è la perdita di un altro agente. Ha una struttura simile ai Dato un set di allenamento, questa tecnica impara a generare nuovi dati con le stesse statistiche del set di allenamento. Ad esempio, un GAN addestrato sulle fotografie può generare nuove fotografie che appaiono almeno superficialmente autentiche agli osservatori umani, avendo molte caratteristiche realistiche. L'idea centrale di una GAN si basa sull'addestramento "indiretto" attraverso il discriminatore, un'altra rete neurale in grado di dire quanto "realistico" sembra l'input, che a sua volta viene aggiornato dinamicamente. Ciò significa che il generatore non viene addestrato a minimizzare la distanza da un'immagine specifica, ma piuttosto a ingannare il discriminatore. Ciò consente al modello di apprendere in modo non supervisionato. Il modello generativo può essere considerato analogo a una squadra di falsari che cercano di produrre valuta falsa e di usarla senza essere scoperti, mentre il modello discriminativo è analogo a quello della polizia, che cerca di individuare la moneta falsa. La competizione in questo gioco spinge entrambe le squadre a migliorare i loro metodi fino a quando le contraffazioni non saranno indistinguibili dagli articoli autentici.

⁹⁵È Direttore e Principal Scientist di Google. È a capo dei team che si occupano di Misinfo Research, Media Integrity, DeepFakes, Cheapfakes, VFX Tech, AR/VR, Human Pose e Face Analysis & Synthesis presso Google AI con lanci in YouTube, DayDream, Photos, JigSaw e altre aree di prodotto. È stato professore alla New York University e alla Stanford University e ha lavorato per diverse aziende tra cui Hewlett Packard, Interval, Disney Feature Animation, LucasFilm's ILM, Facebook's Oculus e il New York Times

⁹⁶Christopher Bregler, Micheal Covell, Malcolm Slaney, *Video Rewrite: Driving Speech with Audio*, Interval Research Corporation, gennaio 1998

L'evoluzione di questo tipo di tecnologia è stata però ben diversa da quanto prospettato dagli ideatori.

Oggi questi software creano non pochi problemi se utilizzati in maniera inopportuna, tant'è che nella maggior parte dei casi vengono bannati o addirittura sanzionati. Pochi anni dopo la diffusione della tecnologia, circolarono sul web video di numerosi politici, tra cui Trump e Obama, sia per servizi volti a deridere la persona⁹⁷ ma anche per spot contro la tecnologia stessa⁹⁸. Circolarono, inoltre, video porno *deepfake* di numerose attrici hollywoodiane, soprattutto di Scarlett Johanson che definì la lotta contro i *deepfake* una "causa persa" e sottolineò come essi fossero una minaccia più per le donne comuni che per donne famose, delle quali era più facile capire che i video fossero falsi⁹⁹. Approfondiremo gli aspetti etici anche di ciò nel prossimo capitolo, nel paragrafo inerente alle *fake news*, alla lotta alla disinformazione e alla misinformazione. Possiamo affermare però che non tutto il mal vien per nuocere. Di questa tecnologia se n'è fatto un uso diverso negli ultimi anni, che si discosta notevolmente dal mondo della pornografia e dei crimini informatici, ma che solleva comunque dilemmi etici sul rapporto uomo macchina e sulla professione del giornalista, come d'altronde tutte le altre tecnologie: l'agenzia governativa cinese Xinhua nel 2018 ha presentato al quinto World Internet Congress mondiale il primo Anchor Man digitale.

2.2.3.1 L'Anchor Man digitale di Xinhua

È il 7 novembre 2018 e siamo a Wuzhen, nella provincia cinese dello Zhejiang. Si sta svolgendo la quinta edizione della World Internet Conference, un evento annuale, tenutosi per la prima volta nel 2014, organizzato dal governo cinese per discutere di questioni e politiche globali relative a Internet e per presentare nuove tecnologie. Tra di esse la quinta edizione ha visto presentarsi un Anchor Man digitale, una sorta di avatar creato tramite tecnologie *deepfake*.

"Hello Everyone. I'm an English Artificial Anchor. This is my very first day in Xinhua News Agency. My voice and appearance are modeled on Zhang Zhao, a real anchor at Xinhua. The development of media industry calls for continuous innovation and deep integration with the international advanced technology".

Si presenta così di fronte alla platea. Oltre alla versione inglese è disponibile anche quella cinese, che imita un altro presentatore in carne ed ossa dell'agenzia. Le movenze e le espressioni facciali sono un po' forzate e se si guarda bene si nota la differenza, anche per la cadenza monotona e artificiale della voce. "Lavorerò instancabilmente per tenervi informati man mano che i testi verranno digitati nel mio sistema senza interruzioni" spiega nel video di presentazione, togliendo ogni dubbio sul fatto che dietro di lui ci sia comunque un team umano. Il dubbio non è del tutto estinguibile in quanto il governo cinese non è nuovo alla mancanza di trasparenza nel proprio operato e in quello delle proprie istituzioni. Xinhua ci tiene a sottolineare che l'anchorman digitale possa lavorare 24 ore su 24 e che la giustificazione della sua creazione può essere

⁹⁷Kyle Swenson, *A Seattle TV station aired doctored footage of Trump's Oval Office Speech. The employee has been fired*, The Washington Post, 11 gennaio 2019

⁹⁸Aja Romano, *Jordan Peele's simulated Obama PSA is a double-edged warning against fake news*, Vox, 18 aprile 2018

⁹⁹Scarlett Johansson *on fake AI-generated sex videos: 'Nothing can stop someone from cutting and pasting my image'*, The Washington Post, 31 dicembre 2018

riassunta nel termine “efficienza”. Il governo cinese non si fa sicuramente scrupoli in termini di perdita del lavoro da parte dei presentatori in carne ed ossa ed è pressoché prevedibile che anche il gruppo di redattori che inserisce le notizie, alla luce delle tecnologie analizzate in questa sede, possa ben presto scomparire. Come si sia evoluta tale tecnologia non ci è dato saperlo. Di notizie dalla Cina, soprattutto per affari riguardanti il governo e le agenzie governative ne arrivano poche e quasi sempre poco attendibili. In questo caso quindi non ci soffermeremo neanche nei prossimi capitoli su aspetti professionali, di occupazione del settore giornalistico e di eticità, in quanto la Cina risulta essere un paese culturalmente e giuridicamente molto lontano dal nostro e dall’Europa.

I pareri degli esperti, infatti, sottolineano anche questa differenza nel descrivere le intelligenze artificiali. Aldo Fontanarosa, giornalista di La Repubblica nonché professore all’università Luiss Guido Carli di Roma, risponde così in un’intervista a ZetaLuiss: “Credo che se un giornale ammettesse di aver iniziato a far scrivere articoli alle macchine, potrebbero esserci scioperi e molte critiche. È giusto preservare i posti di lavoro, ma bisogna trovare un nuovo punto di equilibrio. Non siamo in Cina, dove non esistono limiti di questa natura e dove un’agenzia di informazione come Xinhua può dotarsi di un avatar poliglotta, indistinguibile da un conduttore in carne e ossa, per risparmiare sul personale. Qualcosa, però, può essere fatto”¹⁰⁰. Un ricercatore dell’Università di Oxford invece, ne sottolinea la non umanità e la possibilità che alla lunga gli spettatori possano stufarsi di un presentatore che non coinvolge, che non rende accattivante una notizia, che non è umano¹⁰¹. Questi aspetti, però, alla lunga, potranno sicuramente cambiare, sempre che non lo stiano già facendo.



Fig.12 Anchorman digitale ispirato ad un presentatore reale (Xinhua)

¹⁰⁰Francesco Stati, *Io, giornalista robot. La stampa tra progresso e tradizione*, ZetaLuiss, 23 aprile 2021

¹⁰¹Debutta in Cina il primo anchorman dotato di intelligenza Artificiale, Skytg24, 8 novembre 2018

2.3 Le intelligenze artificiali per la News Distribution

La *news distribution* è la fase finale del processo giornalistico. Capire che caratteristiche ha l'audience del nostro giornale, su che piattaforme caricarlo, come presentarlo e incentivarne la diffusione sono tutti aspetti fondamentali per far sì che i nostri lettori abbiano più probabilità di leggere, ascoltare o vedere la notizia in questione. Se per gli altri due paragrafi abbiamo rimandato tutti gli aspetti etici e deontologici o altre riflessioni che si discostassero dal merito della tecnologia, in questo paragrafo è necessaria una breve digressione su come sia cambiato questo aspetto del giornalismo negli anni a causa della digitalizzazione. Un tempo le notizie avevano un unico format, quello testuale, e i giornali erano solo cartacei. Il lettore comprava in edicola il proprio giornale. Con l'avvento della tv lo scenario è cambiato e sono sopraggiunti i telegiornali. La distribuzione avveniva quindi tramite etere e l'accessibilità era correlata al possesso del mezzo. Con la digitalizzazione lo scenario è drasticamente cambiato. Smartphone e computer hanno cambiato il modo di informarsi delle persone. I giornali col tempo hanno aperto i propri portali web e le persone, oltre che informarsi con i metodi tradizionali, reperiscono notizie anche su Google e sui social. È vero, l'innovazione digitale in Italia è stata più lenta che in Europa, a causa del dominio mediatico delle organizzazioni più importanti anche nei primi anni della trasposizione cartaceo-online, ma nel 2022, per la prima volta, un'agenzia nata in digitale ha ottenuto la più ampia visibilità online¹⁰², superando sia le emittenti più affermate che la prima agenzia di stampa italiana, l'Ansa. Stiamo parlando di FanPage, nata come pagina Facebook nel 2011 e divenuta ora una tra le testate italiane più seguite. FanPage è l'esempio emblematico del cambio di paradigma sulla *News Distribution*, sia per quanto riguarda i canali, sia per quanto riguarda il tipo di notizia scelto sulla base della propria audience e con l'obiettivo di maggiori visualizzazioni. La comunicazione di FanPage si basa molto sui social media, sui pettegolezzi delle celebrità, sulla cronaca nera e su grandi inchieste, tutti canali e argomenti che permettono una maggiore visibilità anche solo in quanto tali. Eterno dibattito in ambito giornalistico è se siano i giornalisti a doversi adattare alla propria audience o se sia l'audience a dover "imparare" dai giornalisti. Abbiamo già sottolineato in questa sede come i giornalisti abbiano in parte perso il ruolo di mediatori e di "filtratori" delle notizie e di come nonostante la percezione negativa dell'ambito giornalistico, la richiesta di news di qualità a pagamento non sia diminuita. Questo è inevitabilmente correlato alla diffusione di *fake news* esponenzialmente aumentata negli ultimi anni, ma anche ad una maggiore fiducia in chi si sostiene economicamente attraverso gli abbonamenti piuttosto che con pubblicità e sponsorizzazioni. La *News Distribution* è quindi strettamente correlata anche alle modalità di sostentamento delle redazioni. In ogni caso, sia che si decida di piegare la scelta delle news sui propri lettori, sia che si scelga di vestire i pochi panni rimasti dei *gatekeepers* e di provare a "educare" la propria audience, le tecnologie basate sull'intelligenza artificiale possono risultare fondamentali per questo ultimo step.

¹⁰²Nic Newman, Richard Fletcher, Craig T. Robertson, Kirsten Eddy, Rasmus Kleis Nielsen, *Reuters Institute Digital News Report 2022*, 2022

Tralasciando le modalità di distribuzione automatica, come quella di alcuni esempi riportati, in cui le notizie si pubblicano sostanzialmente da sole, analizzeremo come possono essere utili software di intelligenze artificiali per questa fase.

Nel primo capitolo abbiamo suddiviso le tecnologie utilizzate nella *News Distribution* in:

- Alert
- Personalizzazione di contenuti

2.3.1 Gli Alert e la personalizzazione dei contenuti

Gli alert e la personalizzazione dei contenuti sono l'essenza della *news distribution* del ventunesimo secolo. Il primo si differenzia dalla seconda perché attiene più ad una scelta dell'utente il quale, con opportuni comandi e impostazioni, può decidere di essere avvertito tramite apposita segnalazione di cambiamenti su determinate pagine web, tra cui quelle di news. La personalizzazione, invece, riguarda più ciò che la redazione fa al fine di rendere l'esperienza dell'utente la migliore possibile, sia in termini di usabilità e quindi di design della pagina o dell'app, sia in termini di suggerimenti di notizie. Il software alert più vecchio e anche il più conosciuto è sicuramente Google Alert, attivo dal 2003. Specificando un determinato tema e indicando il proprio indirizzo e-mail, Google ci segnalerà aggiornamenti in rete in qualsiasi lingua e provenienti da qualsiasi area geografica. Sembra però che esso trascuri gli aggiornamenti dei siti a basso traffico¹⁰³. Web Alert è un'app disponibile per Android che al cambiamento delle pagine non invia una notifica ma ci permette di monitorare i cambiamenti di più pagine visualizzandoli sulla homepage. Esistono, inoltre, estensioni di Browser come Google Chrome e Firefox che, opportunamente settate, ci tengono aggiornati sui cambiamenti. La personalizzazione vera e propria, invece, è maggiormente incisiva e diretta nel tenerci aggiornati.

La personalizzazione dei contenuti avviene in numerosi ambiti della nostra quotidianità: tutto ciò che consumiamo a livello mediatico, da Spotify a Netflix, dalla musica ai film, dagli acquisti online ai suggerimenti di amicizie sui social, è condizionato dalle nostre preferenze che vengono opportunamente rilevate ad ogni nostra azione da algoritmi che imparano sempre di più a conoscerci e ad affinare i suggerimenti e le pubblicità. Se per quanto riguarda gli altri ambiti professionali i problemi etici non sono poi tanto dibattuti, per il giornalismo è diverso. Vi sono opinioni discordanti sulla risposta alla domanda “Le news sono ormai divenute un prodotto come un altro per i consumatori?” La struttura che l'IA utilizza per creare notizie e suggerimenti personalizzati per i lettori è la stessa utilizzata per qualsiasi altro prodotto e per qualsiasi altra piattaforma di social media. In pratica, le IA sono in grado di rilevare la frequenza con cui un lettore legge una determinata sezione di un giornale e ricavare informazioni rilevanti per “vendere” meglio una notizia. I bot possono anche rilevare il tipo di articoli, la natura del giornale e la demografia del pubblico. Queste sezioni possono includere elementi come i tipi di pagine a cui un utente dedica più tempo, la natura delle pagine e tutti i dettagli legati alla presenza online dell'utente. In questo modo, l'intelligenza artificiale raccoglie tutte queste informazioni e

¹⁰³Aldo Fontanarosa, *Giornalisti Robot. L'Intelligenza Artificiale in redazione. Prove tecniche di news revolution*, UlisseAspettaPenelope, Capitolo 10 pg. 287, 14 ottobre 2020

genera design personalizzati che attraggono il lettore in base alle sue preferenze precedenti. Inoltre, la richiesta di personalizzazione non è più solo una tendenza, ma una vera e propria esigenza dei lettori. Una ricerca di McKinsey¹⁰⁴ mostra che il 71% dei consumatori si aspetta dalle aziende interazioni personalizzate. E il 76% si sente frustrato quando ciò non avviene.

La personalizzazione può quindi essere vista sia come un'opportunità che come una minaccia. Il dilemma è se le informazioni siano o no percepite come un qualsiasi prodotto o se, in quanto intrinsecamente diverse, siano percepite come un qualcosa che non debba essere dettato dalle preferenze e dal piacere. Può essere un'opportunità perché potrebbe permettere al sistema informativo di riconnettersi con la propria audience e trovare forme di sostentamento alternative, valorizzando gli abbonamenti e disconnettendosi da condizionamenti o presunte influenze. Può essere invece una minaccia perché si accentuerebbe ancora di più la perdita di ruolo democratico e quel ruolo culturalmente significativo che il giornalismo aveva e che, forse, già non ha più. Per di più un'eccessiva personalizzazione rischierebbe di tenere il lettore all'interno della propria bolla informativa, rafforzandone i confini.

Il problema delle bolle informative, meglio denominate *eco chambers*¹⁰⁵, è una questione che riguarda le informazioni nella loro totalità. Il pluralismo dovrebbe essere un valore cardine della democrazia ma la digitalizzazione, o meglio, i suoi aspetti negativi, hanno portato chiunque ad accentuare quel desiderio di rimanere nella propria zona di comfort, dove si ricerca più la conferma delle proprie opinioni piuttosto che la loro messa in discussione. Questa questione riguarda tanto i social quanto i giornali, con la differenza che i primi sono maggiormente legittimati a perseguire il profitto rispetto ai secondi in quanto visti, agli occhi dell'opinione pubblica, più come fornitori di beni ludico-ricreativi che come aventi un ruolo sociale, cosa che i secondi, pur nel loro declino, continuano in parte ad avere. Il dilemma allora si riduce a quale fine dovrebbe perseguire una redazione: utilità sociale e servizio democratico per i cittadini, promuovendo la messa in discussione delle opinioni e la messa a disposizioni di diversi punti di vista, o perseguimento del profitto e accentuazione del ruolo dell'informazione come ambito consumeristico come gli altri? Spesso la risposta a queste domande è condizionata dal fatto che in determinati casi non si tratta di logica del profitto ma per lo più di sopravvivenza. Abbiamo già descritto in questa sede la crisi del settore e le sue problematiche e approfondiremo ulteriormente tale tema in un capitolo dedicato. In ogni caso, le due finalità possono essere perseguite contemporaneamente. I lettori non hanno perso interesse in notizie di qualità, nonostante la sfiducia diffusa nel settore informativo. Una personalizzazione progettata sulla base delle esigenze del singolo che promuova comunque una pluralità di opinioni non è del tutto impossibile.

¹⁰⁴Nidhi Arora, Wei Wei Liu, Kelsey Robinson, Eli Stein, Daniel Ensslen, Lars Fiedler, Gustavo Schuler, *Next in Personalization Report*, McKinsey, 2021

¹⁰⁵Nella società contemporanea dei mezzi di comunicazione di massa, caratterizzata da forte interattività, le *eco chambers* sono situazioni in cui informazioni, idee o credenze più o meno veritiere vengono amplificate da una ripetitiva trasmissione e ritrasmissione all'interno di un ambito omogeneo e chiuso, in cui visioni e interpretazioni divergenti finiscono per non trovare più considerazione. La tendenza ad aggregarsi con persone con le stesse attitudini e interessi è determinante sia nel rinforzare le *echo-chambers* sia nel determinare la dimensione di un processo virale. Il numero di persone coinvolte in un processo virale è strettamente collegato alla dimensione dell'*echo-chamber* di cui fanno parte (Walter Quattrociocchi, Antonella Vicini, *Misinformation. Guida alla società della disinformazione e della credulità*, FrancoAngeli, p. 76, 2016)

2.3.1.1 Voitto: l'assistente per le notizie che propone alternative

Il progetto Voitto è stato avviato dalla compagnia finlandese Yle nel dicembre 2016. Oltre ad automatizzare alcune attività, come l'editing di brevi video o la scrittura di brevi articoli, aggiunge valore ai processi esistenti, come l'assistente personale nell'app Yle NewsWatch, che invia all'utente notifiche personalizzate sui contenuti consigliati. L'assistente, di nome Voitto¹⁰⁶, impara dalle usanze del lettore, da cosa ha letto in passato, dalle sue interazioni e ricerche sul web, ma anche da feedback diretti che il lettore stesso dà all'assistente. Fin qui sembrerebbe il classico *recommender system* che mira a mantenerti sulla piattaforma il più possibile, se non fosse che all'attivarsi della notifica di una notizia, e dopo l'apertura della stessa, Voitto consiglia altre due storie riguardanti lo stesso argomento ma che pongono la questione da diverse prospettive. Questo permette all'utente di rimanere aggiornato sui propri interessi e ad avere comunque la possibilità di evadere da quei confini rafforzati delle camere d'eco. Poter sentire diverse opinioni su una stessa storia è un valore che il giornalismo dovrebbe promuovere. Yle stessa ha dichiarato che tutte le linee guida del progetto rispettavano rigorosamente i valori del giornale, tra cui quello sì di far andare parallelamente sviluppo tecnologico e vicinanza ai lettori, ma anche quello di promuovere il pluralismo delle idee. L'assistente è stato un enorme successo: il 90% di chi lo ha provato ha continuato ad usarlo¹⁰⁷. Yle però non si è accontentata e ha fatto del miglioramento costante un mantra. Utilizzano dati in tempo reale per verificare la qualità delle raccomandazioni¹⁰⁸ ed effettuano A/B Test¹⁰⁹ su differenti modelli di raccomandazione. Oltre alla qualità delle raccomandazioni, ne viene valutata anche la giusta quantità e l'orario opportuno: ci sono lettori che preferiscono essere informati costantemente sulle novità di loro interesse, mentre altri preferiscono venire a conoscenza di poche notizie al giorno e in determinati momenti della giornata. L'obiettivo di Yle è stato anche quello di promuovere la trasparenza di queste tecnologie (Voitto, infatti, è Open Source) e di cambiare la percezione che l'opinione pubblica ha delle intelligenze artificiali: hanno dato un volto all'assistente (il volto sorridente di un robot simpatico e innocente) e un colore rassicurante (il blu). Personalizzare le notizie, in modo etico e promuovendo il pluralismo, è quindi possibile.



Fig.13 Yle's Voitto Smart Assistant¹¹⁰

¹⁰⁶Il nome deriva da un famoso commentatore sportivo finlandese, Voitto Liukkonen, morto nel 2007. Era noto per il suo stile minimalista di commento

¹⁰⁷Barbara Gruber, *Facts, Fakes and Figures: How AI is influencing journalism*, Goethe Institute

¹⁰⁸Jarno Kaponen, *First in the World: Yle's smart news assistant Voitto ensures that you don't miss the news you want to read*, News Lab Yle, 10 ottobre 2018

¹⁰⁹Il test A/B è un modo di confrontare due versioni di una singola variabile per testare la risposta del soggetto rispetto alla variabile A o B e determinare quale risulti la più efficace

¹¹⁰Jarno Kaponen, *First in the World: Yle's smart news assistant Voitto ensures that you don't miss the news you want to read*, News Lab Yle, 10 ottobre 2018

CAPITOLO 3 Legislazione e questione etiche

L'etica aiuta l'uomo a concettualizzare la differenza tra ciò che comunemente intendiamo per bene e ciò che comunemente intendiamo per male. Il punto focale dell'etica, intesa come insieme di studi e ricerche inerenti a ciò che è giusto e sbagliato, è la morale. La morale può essere individuale o collettiva e afferisce a quell'insieme di valori che si perseguono nella propria vita in quanto umani. Ciò che si vuole o si dovrebbe essere non riguarda solo l'ambito filosofico, ma anche quello giuridico-normativo e professionale. Questi due aspetti vanno spesso a braccetto e quando si affronta il problema della morale collettiva di una determinata professione, si finisce inevitabilmente a parlare anche degli aspetti giuridici. Abbiamo più volte ripetuto come le notizie di oggi vengano distribuite attraverso diversi canali e formati. L'informazione oggi usufruisce di più canali rispetto a prima e ci si è quindi posti il problema, nel momento storico di questo cambiamento, del regime giuridico che riguardava i nuovi mezzi di informazione: giornali online, blog, forum, social network come dovevano essere identificati a livello normativo? Ora tale questione si è posta per le intelligenze artificiali in generale e per il loro utilizzo negli ambiti professionali. Ci si pone il problema non solo di come esse debbano essere identificate giuridicamente, ma anche quali limitazioni apporre al loro operato e come disciplinare gli ambiti professionali per tutelare i lavoratori in carne ed ossa. Le questioni che il progresso tecnologico porta con sé, oltre a quelle giuridiche in senso stretto, sono molte e riguardano numerose dinamiche: trasparenza, discrezionalità ed ecologia sono solo alcune delle tematiche del dibattito. Le affronteremo cercando di rispondere o quantomeno porre le domande giuste per ognuna di esse nell'ambito del giornalismo. Nei precedenti capitoli abbiamo più volte accennato a dilemmi etici concernenti l'utilizzo delle intelligenze artificiali nel giornalismo, rimandandone l'approfondimento a questo capitolo. Affronteremo quindi tutte quelle discussioni e quei problemi lasciati in sospeso che riguardano l'oggetto di questo lavoro dal punto di vista filosofico e giuridico-normativo, andando ad analizzare diversi aspetti di come si dovrebbero utilizzare le tecnologie precedentemente analizzate e delle questioni che, non ancora risolte, sono oggetto dei dibattiti attuali.

Prima di entrare nello specifico è necessario sottolineare come l'etica nel giornalismo sia stata più volte concettualizzata e che esistono anche codici etici per l'utilizzo delle intelligenze artificiali in generale. Manca però un'unione delle due, cioè un codice etico per l'utilizzo delle intelligenze artificiali nel giornalismo. Nessuna delle redazioni che ha implementato o sperimentato intelligenze artificiali ha aggiunto qualcosa ai propri codici etici. A livello teorico concettuale, nella ricerca scientifica, esistono rappresentazioni scritte e ben categorizzate di quelli che dovrebbero essere i valori guida nell'utilizzo delle intelligenze artificiali nel giornalismo. Ci limiteremo quindi a riportare qui quelli che in letteratura scientifica vengono denominati *ethical framework*¹¹¹ e a fornire successivamente una nostra analisi di questo argomento.

¹¹¹Per *ethical framework* si intendono quegli schemi concettuali che fanno da guida nelle decisioni attraverso la distinzione di ciò che è giusto o ciò che è sbagliato in un determinato ambito e che riguardano i valori e le condotte che si perseguono

Il concetto fondamentale da tenere bene a mente quando si parla di tecnologie a livello etico è *human in the loop*¹¹². Nonostante la rappresentazione mediatica che ne viene fatta, gli umani hanno ancora un ruolo preponderante nell'utilizzo delle intelligenze artificiali. Joanna Bryson, professoressa di etica e tecnologia, sostiene che troppo spesso tendiamo ad esaltare il valore delle macchine in quanto tali sminuendo il nostro valore in quanto umani, ma sottolinea anche che le nuove tecnologie possono aiutarci a razionalizzare maggiormente le nostre decisioni¹¹³. Talvolta, sono stati utilizzati degli algoritmi per prendere delle decisioni ma il risultato a cui si è arrivati, sia concettualmente che giuridicamente, è che l'umano ha sempre l'ultima parola in merito.

Qualche anno fa, con la legge n. 107 del 2015, meglio conosciuta con il nome di “Buona Scuola”, si è dato vita oltre che ad una riforma scolastica senza precedenti, ad una serie di assunzioni di docenti e di copertura di posti vacanti. L'assegnazione, tenendo conto solo in parte delle preferenze dei singoli docenti, era stata affidata ad un algoritmo. L'algoritmo non ha tenuto conto in maniera appropriata neanche della collocazione geografica delle sedi scolastiche e ha assegnato scuole del Nord a docenti del Sud, commettendo anche altri errori simili. I docenti hanno fatto ricorso sul funzionamento dell'algoritmo, di cui tra l'altro non ne era stato conoscibile il processo. Il TAR del Lazio prima e il Consiglio di Stato poi hanno posto delle condizioni specifiche sull'utilizzo di questi algoritmi. I giudici di Palazzo Spada hanno quindi delimitato il perimetro entro il quale un algoritmo possa qualificarsi come regola giuridica amministrativa lecita, sottolineando come *trasparenza, conoscibilità e non esclusività*¹¹⁴, debbano essere dei principi cardine dell'utilizzo di queste tecnologie. È vero sì che si tratta di un ambito diverso dall'oggetto del nostro studio, ma questo esempio ci è utile per capire come l'utilizzo di queste tecnologie debba essere sempre vincolato all'assistenza dell'uomo e che, a maggior ragione in un periodo storico in cui gli algoritmi e i software commettono ancora errori, essi non possano essere lasciati operare e decidere in maniera autonoma. Nel giornalismo il discorso è ovviamente diverso, in quanto non si va ad incidere direttamente sulla vita delle persone come nelle decisioni della pubblica amministrazione. In maniera indiretta però, l'informazione è fondamentale per la democrazia e per la vita delle persone. Il giornalismo permette ai cittadini di decidere in maniera opportuna nella vita individuale e collettiva ed è quindi allo stesso modo fondamentale che i processi analizzati nei capitoli precedenti seguano un percorso eticamente giusto e formalmente corretto.

Una concettualizzazione che analizza il *workflow* giornalistico come presentato finora definisce così i valori etici dell'utilizzo delle IA nel giornalismo¹¹⁵:

- per il *news gathering*: valutare periodicamente l'input di dati e creare un processo *ex ante* che valuti tutto il procedimento e gli output

¹¹²Questo termine ha diverse accezioni. Nel Machine learning è usato per indicare che gli esseri umani aiutano i software anche dopo averli impostati, guidando l'apprendimento automatico e aiutandoli durante l'addestramento

¹¹³Bryson J., Kime, P., *Just another artifact: Ethics and the empirical experience of AI presented at Fifteenth International Congress on Cybernetics*, Edinburgh, January 1998

¹¹⁴Qui la “non esclusività” è intesa come l'impossibilità per la macchina di prendere decisioni in maniera autonoma senza l'ulteriore intervento dell'uomo

¹¹⁵ Kim, Haley, *AI in Journalism: Creating an Ethical Framework*, Syracuse University, Honors Program Capstone Projects, 2019

- per la *news production*: valutare costantemente l'ambito nel quale si producono contenuti in maniera totalmente automatica, al fine di evitare ad esempio che essi contravvengano a leggi appena entrate in vigore;
- per la *news distribution*: monitorare l'utilizzo di dati di terze parti e rendere chiaro ai lettori quale parti sono state create da tecnologie e quali invece sono frutto del lavoro umano.

Nonostante verranno utilizzate anche in questa sede le nozioni teorizzate da Dörr, si è preferito suddividere l'analisi per argomenti, proprio perché si è ritenuto più opportuno, alla luce della trasversalità del monitoraggio e di altre funzioni utili all' "eticità" del processo giornalistico, analizzare per concetti e non per processi. Una categorizzazione simile a quella che utilizzeremo in questa sede è la seguente¹¹⁶:

- Leggi: l'utilizzo delle IA deve essere conforme al sistema giuridico all'interno del quale operano
- IA allineate ai valori del giornale: l'utilizzo delle IA deve essere conforme al codice etico del giornale
- Etica e sicurezza delle IA: l'utilizzo delle IA deve fornire output corretti e precisi nonché evitare bias come quelli di genere o razza
- Intelligibilità: le IA devono essere conoscibili e soprattutto comprensibili a tutti
- Trasparenza e responsabilità: il processo delle IA deve essere trasparente in ogni minimo dettaglio e deve essere possibile identificare se un determinato errore è stato commesso dall'IA per problemi tecnici, da chi ha fornito i dati o da chi ha impostato la tecnologia
- Governance dei dati e loro gestione: i dati che l'IA utilizza devono essere gestiti in maniera lecita e trasparente nonché rispettare i diritti di chi li ha raccolti

Questa suddivisione ci permette di capire quali siano i temi più dibattuti in materia e ci farà da stella polare nell'analisi specifica di questi concetti. La categorizzazione di questo capitolo, il cui riflesso sarà la suddivisione in paragrafi, prevederà analisi specifiche sulla legislazione italiana e come essa vada ad inficiare sull'utilizzo delle intelligenze artificiali nel giornalismo, sulla trasparenza dei processi che trascende la suddivisione del processo e deve contraddistinguere ogni step, sulla discrezionalità del loro utilizzo e quindi un approfondimento di cosa voglia dire nello specifico il già menzionato *human in the loop* e l'impatto ambientale delle intelligenze artificiali.

Tutti questi aspetti, più che essere delle limitazioni, sono delle opportunità per i giornalisti non solo per un discorso di coscienza professionale ma anche perché, in un "mercato informativo" pieno di attori diversi da soggetti istituzionali e redazioni giornalistiche vere e proprie, il rispetto di questi principi e di questi valori può porre gli organismi pubblici e i giornalisti un gradino sopra ai concorrenti. Va inoltre sottolineato che più si andrà avanti a livello tecnologico in questo campo, ancor più sarà necessaria una enfaticizzazione di tali aspetti.

¹¹⁶B. Jones, E. Luger, *AI and Journalism - Intelligible Cloud and Edge AI (ICE-AI)*, PETRAS National Centre of Excellence, 17 settembre 2021

3.1 La legislazione italiana sul giornalismo

3.1.1 Dal cartaceo al digitale

L'informazione oggi usufruisce di più canali rispetto a prima. Ci si è posti il problema, con il passaggio dal cartaceo al digitale, di quale dovesse essere dal punto di vista normativo il regime giuridico che riguarda giornali online, blog, forum e social network. Si è posta innanzitutto una questione. La Costituzione parla di "stampa" all' articolo 21, inerente alla libertà di espressione. Ognuno può esprimere la propria opinione e per la stampa valgono regole specifiche. Dopo il fascismo vi era la necessità di prevedere forme più ampie possibili di libertà del pensiero, soprattutto tramite la stampa che non poteva quindi essere sottoposta ad autorizzazioni e censure, salvo i reati previsti dalla legge sulla stampa. Fino al 2015, l'intenzione iniziale del legislatore quanto alla giurisprudenza era di negare l'assimilazione tra stampa cartacea e stampa online. Si poteva quindi parlare di stampa e valevano le garanzie previste dall'articolo 21, cioè divieto di autorizzazioni, censure o sequestri, solo per la stampa cartacea. Fino a quel momento quindi la stampa cartacea poteva essere sottoposta a censure e sequestri. La legge n. 62 del 2001 aveva in parte modificato il quadro senza minare la chiave differenzista¹¹⁷ dell'interpretazione della stampa online. Non c'era stata quindi ancora assimilazione tra stampa online e cartacea ma si poteva parlare di "prodotti editoriali" per i contenuti online. Il giornale online che pubblicava questi contenuti, alla ricorrenza di alcuni requisiti quali l'esistenza di un direttore responsabile e l'iscrizione al registro specifico per le pubblicazioni informative¹¹⁸, poteva accedere ad alcuni sovvenzionamenti pubblici. Con la pronuncia delle Sezioni unite n. 31022 del 2015 cambia l'interpretazione: si afferma che anche i giornali online, sempre al rispetto di alcuni requisiti, possano essere considerati allo stesso modo della stampa cartacea. L'accezione tecnica di "stampa", strettamente legata alla riproduzione tipografica, aveva corroborato fino a quel momento le tesi differenziste e quindi l'approccio della giurisprudenza. La legge n. 47 del 1948, infatti, disciplinava ciò che poteva essere chiamato stampa in senso giuridico secondo due criteri principali: la riproduzione tipografica (*prius*) e la pubblicazione e distribuzione (*posterus*). La Corte escludeva quindi inizialmente l'assimilabilità dei prodotti online alla stampa, facendo leva principalmente su innegabili rilievi tecnici. Nel 2015 però, constatata l'effettiva validità dei giornali online, per la prima volta si sottolinea, alla luce dell'evoluzione tecnologica e della conseguente evoluzione del settore giornalistico, la necessità di guardare ad un significato figurato di "riproduzione tipografica". La differenziazione era ormai divenuta obsoleta e i giornali online avevano ormai vita propria rispetto a quelli cartacei. La Cassazione sottolinea inoltre che social, blog e forum non sono mai assimilabili al concetto di stampa. Discorso diverso va fatto però per i social network. Nonostante essi non siano ritenuti "stampa" e possono quindi essere soggetti a censure e sequestri, nella pratica ciò non avviene. Rimane quindi il problema dell'inibizione di contenuti lesivi. Se per eventi diffamatori, e quindi concernenti una singola persona, si

¹¹⁷In giurisprudenza si parla in questo dibattito di tesi differenzista, quella sostenuta da chi non ritiene equiparabili stampa cartacea e stampa online, e tesi assimilista, quella sostenuta da chi ritiene che siano assimilabili.

¹¹⁸Il ROC (Registro unico degli Operatori di Comunicazione) ha la finalità di garantire la trasparenza e la pubblicità degli assetti proprietari, consentire l'applicazione delle norme sulla disciplina anticoncentrazione, la tutela del pluralismo informativo, il rispetto dei limiti previsti per le partecipazioni di società estere.

ricorre spesso al blocco dell'account, per i casi più generalizzati, ci si pone la domanda se sia il social stesso a favorire determinati tipi di condotte illecite. Quante volte vediamo fenomeni generalizzati di insulti sui social. Se un giudice ritenesse che i reati di diffamazione e discriminazione siano quantitativamente tanti a causa della natura stessa del social che incentiva tali condotte, potrebbe teoricamente sequestrarlo. Ovviamente ciò nella pratica non avviene ma tutto questo quadro giuridico appena descritto ha portato ad oggi ad un approccio collaborativo tra social network e istituzioni, principalmente europee, che prevede un monitoraggio costante delle condotte illecite sui social e la composizione di codici di condotta integrati all'interno delle piattaforme stesse.

Tutto questo ci è utile a capire il quadro giuridico che le redazioni oggi si trovano ad affrontare. Pur avendo maggiori tutele dei social, anch'essi fronteggiano la necessità di implementare forme di controllo sui commenti ai giornali online. Queste forme di controllo possono essere adottate attraverso l'utilizzo di intelligenze artificiali. Abbiamo visto nel precedente capitolo come esistano software di ricerca di testi e immagini che aiutano a scovare notizie, tendenze e potenziali contenuti. Questi software possono essere utilizzati per monitorare anche i casi di diffamazione e discriminazione online. Dal lato invece di chi le notizie le diffonde, il regime giuridico è diverso. Ci sono delle tutele a livello professionale che permettono al giornalista di "sbilanciarsi" e poter criticare una determinata persona. Queste esimenti sono il diritto di cronaca e di critica per i quali però devono valere dei presupposti: la struttura e la valenza informativa del giornale, cioè quindi che esso sia riconosciuto come tale, verità dei fatti, interesse pubblico e continenza verbale¹¹⁹. Tutti questi diritti e doveri valgono anche quando si utilizzano intelligenze artificiali.

3.1.2 Dal digitale alle intelligenze artificiali

In realtà sarebbe erroneo parlare già di un passaggio dal digitale alle intelligenze artificiali nel mondo del giornalismo. Non è equiparabile la portata del cambiamento che ha portato il web nel settore del giornalismo. È opportuno però fare il punto sugli sviluppi attuali del settore dal punto di vista legislativo. Dal momento che mancano codici etici e regolamenti sull'utilizzo delle intelligenze artificiali, riporteremo qua i punti fondamentali dei riferimenti legislativi in materia di intelligenze artificiali e di giornalismo come due ambiti distinti. Va detto che tentativi di codici etici per le intelligenze artificiali nel giornalismo sono stati quantomeno prospettati da alcuni ricercatori¹²⁰, ma nonostante i buoni presupposti essi non sono mai stati realizzati. Il punto di riferimento a livello di IA in generale è la proposta coordinata di Regolamento sulle intelligenze artificiali delle istituzioni europee. Sui punti di riferimento per la professione in sé rimandiamo la discussione al paragrafo dedicato alla deontologia. In ultima istanza va precisato che pur concentrandoci sul panorama mediatico italiano e conseguentemente su tutti gli aspetti paralleli che riguardano il nostro paese, l'Italia è molto indietro su tutti questi aspetti rispetto ad esempio agli Stati Uniti. Molti degli esempi riportati riguardano redazioni al di là dell'Atlantico e qui in Italia, nonostante sperimentazioni ed esempi virtuosi di alcune

¹¹⁹Intesa come proporzione e misura delle espressioni utilizzate

¹²⁰Mohamed Abdulzاهر, *Artificial Intelligence Journalism and Professional Code of Ethics*, Artificial Intelligence Journalism for Research and Forecasting, ottobre 2020

redazioni, le IA nel giornalismo non sono ancora una realtà consolidata. Risulta comunque utile analizzare il contesto legislativo nel quale si opera analizzando, come fatto finora, la legislazione italiana e quella europea. Il contesto normativo italiano non prevede norme sull'utilizzo di queste tecnologie, in attesa degli sviluppi dell'AI Act, la proposta di Regolamento appena menzionata.

In questa proposta di Regolamento, redatta dal Parlamento Europeo e del Consiglio, si stabiliscono regole armonizzate sull'intelligenza artificiale e la modifica di alcuni atti legislativi dell'Unione. All'interno dell'atto si legge: «Il termine intelligenza artificiale indica una famiglia di tecnologie in rapida evoluzione in grado di apportare una vasta gamma di benefici economici e sociali in tutto lo spettro delle attività industriali e sociali. [...] In considerazione della velocità dei cambiamenti tecnologici e delle possibili sfide, l'UE si impegna a perseguire un approccio equilibrato. [...] L'interesse dell'Unione è quello di preservare la leadership tecnologica dell'UE e assicurare che i cittadini europei possano beneficiare di nuove tecnologie sviluppate e operanti in conformità ai valori, ai diritti fondamentali e ai principi dell'Unione».

Questa proposta del 2021 prende spunto da un atto dell'anno precedente, il Libro bianco sull'intelligenza artificiale che prevedeva obiettivi specifici:

- Sistemi sicuri e che rispettino le normative
- Assicurare la certezza del diritto per facilitare gli investimenti
- Migliorare la governance e l'applicazione effettiva della normativa esistente in materia di diritti fondamentali e requisiti di sicurezza applicabili ai sistemi di intelligenza artificiale
- Facilitare lo sviluppo di un mercato unico per applicazioni di IA lecite, sicure e affidabili nonché prevenire la frammentazione del mercato.

L'obiettivo generale è quindi quello di bilanciare rischi in tutti gli ambiti di applicazione delle IA e favorire un approccio trasversale per l'applicazione dei requisiti minimi di sicurezza. La realtà dei fatti, quantomeno per ciò che concerne il settore giornalistico, vede i paesi del nord Europa essere ad un livello leggermente più avanzato rispetto a quelli mediterranei. In ogni caso nessuno dei paesi europei è al livello delle redazioni statunitensi. Va elogiato però l'approccio più garantista ed equilibratore dell'Unione Europea rispetto a quello statunitense. Alcuni degli approcci al progresso tecnologico che hanno accompagnato lo sviluppo di queste nuove tecnologie, hanno portato non pochi problemi nella loro maggiore libertà durante tutto il processo che ha accompagnato la creazione delle IA. Dai bias razzisti sino al percorso di addestramento di ChatGPT¹²¹, una maggiore libertà è stata talvolta sinonimo anche di maggiori complicazioni etiche. Nell'addestramento di ChatGPT è stato necessario inserire una sorta di filtro dei dati di addestramento. Bisognava quindi fornire ad un'intelligenza artificiale esempi etichettati di violenza, incitamento all'odio e abuso sessuale, così da "insegnare" all'IA a rilevare queste forme di tossicità.

¹²¹Billy Perrigo, *Exclusive: OpenAI Used Kenyan Workers on Less Than \$2 to Make ChatGPT Less Toxic*, TIME, 18 gennaio 2023

Il partner di OpenAI in Kenya era Sama, un'azienda americana che impiega lavoratori Africa e Asia per etichettare i dati per clienti come Google, Meta e Microsoft. Gli etichettatori di dati assunti da Sama per conto di OpenAI ricevevano un salario che oscillava tra 1,32 e 2 dollari l'ora, a seconda dell'anzianità e delle prestazioni. Tralasciando per ora discussioni su questi processi, va evidenziato come l'approccio legislativo americano sulle intelligenze artificiali sia più soft rispetto a quello europeo. Al fine di bilanciare i diversi diritti in gioco, la disciplina europea precedentemente menzionata prevede un'identificazione in base al rischio delle intelligenze artificiali: una piramide ascendente sino alle tecnologie ritenute inaccettabili¹²². Individuando quindi profili diversi in base alle problematiche che determinate tecnologie possono provocare si prevedono limitazioni e strumenti di prevenzione ad hoc. Tra le tecnologie considerate ad alto rischio che vengono utilizzate nelle redazioni giornalistiche vi è sicuramente il riconoscimento facciale, denominato nel regolamento come "identificazione biometrica e categorizzazione delle persone fisiche".

Le intelligenze ad alto rischio sono soggette a determinati obblighi:

- Individuazione e gestione dei rischi: capire quali possono essere i rischi in termini etici ed economici e predisporre opportune misure di prevenzione e programmi specifici di gestione dei rischi
- Qualità e appropriata gestione dei dati: i dataset utilizzati devono essere pertinenti, rappresentativi, corretti e completi al fine di limitare il più possibile la possibilità di bias
- Documentazione tecnica esaustiva: la creazione dei software e l'implementazione dei processi devono essere il più possibile dettagliati
- Tracciabilità: si deve far sì che determinati output possano essere tracciati e quindi si possa stabilire cosa ha originato cosa, in modo da poter identificare al meglio i processi; si differenzia dai concetti di individuazione e gestione dei rischi e dalla documentazione perché rispetto ad essi riguarda i nessi causa-effetto
- Trasparenza: riguarda non solo la conoscibilità delle strumentazioni e dei processi ma anche la comprensibilità di essi
- Supervisione umana: non sono consentiti per le Intelligenze Artificiali ad alto rischio sistemi non supervisionati. Il concetto di human in the loop come detto è fondamentale, e in questo caso diviene obbligatorio
- Accuratezza, robustezza e sicurezza informatica: riguarda la predisposizione dei software alla tutela dei dati nonché alla protezione da attacchi informatici e virus esterni potenzialmente possibili a causa dell'impossibilità di essi di operare offline

Questi obblighi devono fronteggiare non pochi ostacoli: dal bilanciamento di diritti confliggenti a limiti di tipo tecnico sino all'ottemperanza di tali obblighi, di cui la trasparenza è l'esempio emblematico.

¹²²La disciplina europea menziona tra le intelligenze artificiali nella categoria inaccettabili "pratiche online di manipolazione cognitiva che causano danni fisici o psicologici o sfruttano la vulnerabilità dovuta all'età o alla disabilità"

3.1.3 La responsabilità dell'operato dell'IA

Prima di affrontare la questione della trasparenza è necessario capire, per concludere l'analisi legislativa del regime giuridico che afferisce alle intelligenze artificiali, individuare chi è responsabile dell'operato di un'IA, con la conseguente individuazione anche di chi è giuridicamente opportuno incolpare per un output illecito. Gli output che un'intelligenza artificiale genera sono originati da processi complessi. Ogni step ha compreso più attori per la propria creazione, la loro modifica e il loro controllo. Abbiamo appena visto come sia necessario un monitoraggio il più costante possibile e una trasparenza che accompagni tutte le componenti e il percorso che un'IA intraprende per generare un output. Spesso questi output possono dare vita a conseguenze inaspettate che vanno a ledere diritti di terzi, come ad esempio un contenuto denigratorio o che travalica la tutela del diritto alla riservatezza, oppure un superamento dei limiti tracciati dagli esimenti di diritto di cronaca e di critica giornalistica. Nel momento in cui si verificano queste condotte illecite, chi ne è responsabile? La risposta, alla luce delle analisi sin qui compiute ci porterebbe a rispondere che la persona responsabile di tali condotte, nel caso ad esempio di articoli scritti da un'IA, sia il giornalista che ha "collaborato" con essa. Ci sono casi in cui però gli articoli sono completamente scritti da un'intelligenza artificiale e anche nel caso di collaborazione tra umano e software, non risulterebbe giusto attribuire colpe al giornalista quando magari sono state causate da un settaggio sbagliato. Tutto ci porta a constatare che l'attribuzione di responsabilità è alla fine un problema inevitabilmente collegato alla trasparenza. La conoscibilità dei processi ci permette di capire i nessi causa-effetto e di stabilire cosa abbia determinato una condotta illecita. Nello specifico però risulta difficile associare ogni piccolo step di un processo ad una persona fisica. La giurisprudenza cosa ci dice in merito?

Il testo di riferimento in questo senso è la Relazione recante raccomandazioni alla Commissione su un regime di responsabilità civile per l'intelligenza artificiale. Tale relazione prevede che debbano essere adeguati e coordinati i regimi di responsabilità al fine di garantire sempre il risarcimento alle persone lese e riconosce che l'utilizzo delle intelligenze artificiali possa far scaturire, in maniera diretta e indiretta, "danni e pregiudizi"¹²³.

La relazione "rileva a tale proposito che non è necessario conferire personalità giuridica ai sistemi di IA; è del parere che l'opacità, la connettività e l'autonomia dei sistemi di IA potrebbero rendere, nella pratica, molto difficile o addirittura impossibile ricondurre specifiche azioni dannose dei sistemi di IA a uno specifico input umano o a decisioni adottate in fase di progettazione; ricorda che, conformemente a concetti di responsabilità ampiamente accettati, è tuttavia possibile aggirare tale ostacolo considerando responsabili le varie persone nella catena del valore che creano il sistema di IA, ne eseguono la manutenzione o ne controllano i rischi associati". La disciplina Europea ritiene esaustive le normative nazionali per la tutela di tali dinamiche e che la "responsabilità per danno da prodotti difettosi" sia ancora la direttiva cardine per la tutela di danni e

¹²³Per danno o pregiudizio viene ad intendersi un qualunque effetto negativo che influisce sulla vita, la salute, l'integrità fisica di una persona fisica, il patrimonio di una persona fisica o giuridica o che causa un rilevante danno non patrimoniale comunque traducibile in una perdita economica verificabile (Altalex)

pregiudizi, ma sottolinea come debba essere adattata al progresso tecnologico e come “dovrebbe andare di pari passo con l'aggiornamento della direttiva 2001/95/CE del Parlamento europeo e del Consiglio, del 3 dicembre 2001, relativa alla sicurezza generale dei prodotti, al fine di garantire che i sistemi di IA integrino la sicurezza e la protezione fin dalla progettazione”. Unica lacuna che dovrebbe essere colmata quindi sono i casi in cui è irrintracciabile e insolubile determinare un terzo che ha interferito con il processo.

Viene individuata invece per tutti gli altri casi la figura dell'operatore. La relazione afferma che la responsabilità dell'operatore è giustificata dal fatto che tale persona sta controllando un rischio associato al sistema di IA, in modo analogo al proprietario di un'automobile e ritiene che, vista la complessità e la connettività di un sistema di IA, l'operatore sarà, in molti casi, il primo punto di contatto visibile. Suddivide i possibili operatori, contemplando l'ipotesi non rara in cui ci siano più operatori, in¹²⁴:

- Operatore di front-end: la persona fisica o giuridica che esercita un certo grado di controllo su un rischio connesso all'operatività e al funzionamento del sistema di IA e che beneficia del suo funzionamento
- Operatore di back-end: la persona fisica o giuridica che, su base continuativa, definisce le caratteristiche della tecnologia fornisce i dati e il servizio di supporto di back-end essenziale e pertanto esercita anche un elevato grado di controllo su un rischio connesso all'operatività e al funzionamento del sistema di IA

Si evidenzia inoltre come il successo di queste tecnologie sia collegata ad un'adeguata disciplina degli illeciti al fine di aumentare la fiducia delle persone per questi sistemi. Prevede inoltre che gli operatori si tutelino loro stessi prevedendo un'assicurazione su tale responsabilità, che non sia però finanziata da denaro pubblico.

A livello teorico Diakopoulos formalizza *l'accountability* dell'intelligenza artificiale considerando gli algoritmi come una creazione umana, con l'obiettivo di rendere conoscibile l'intento impostato nella loro creazione. Uno studio proposto evidenzia anche se il reverse-engineering conclude che i criteri incorporati in un algoritmo corrispondono all'intento pubblicizzato, è possibile che comportamenti avversari facciano deviare l'algoritmo dal suo funzionamento previsto. Si sottolinea quindi la necessità di dare rilievo al parametro della “robustezza” ovvero un'implementazione dei processi che preveda anche devianze e comportamenti inattesi al fine di limitarli¹²⁵.

¹²⁴Nel campo della progettazione software e sviluppo software il front end è la parte di un sistema software che gestisce l'interazione con l'utente o con sistemi esterni che producono dati di ingresso (es. interfaccia utente con un **form**), il back end è invece la parte che elabora i dati generati dal front end. Nei sistemi più complessi non è raro che i dati subiscano elaborazioni intermedie prima di passare al back end (ad es. su un middleware come un application server in una tipica architettura three-tier). La distinzione di una parte di ingresso e di una parte terminale nei sistemi software è un genere di astrazione che aiuta a mantenere le diverse parti di un sistema complesso logicamente separate e quindi più semplici. (Wikipedia)

¹²⁵Antonin Descampe, Clement Massart, Simon Poelman, Francois-Xavier Standaert, Olivier Standaert, *Automated news recommendation in front of adversarial examples and the technical limits of transparency in algorithmic accountability*, AI & Society, 13 marzo 2021

3.2 Questioni etiche

3.2.1 La trasparenza

“La trasparenza è generalmente considerata un mezzo per vedere la realtà, le motivazioni che hanno portato le persone a compiere una determinata azione e determinare le responsabilità di fronte al proprio pubblico, legittimando la fiducia di esso. A livello basilare, la trasparenza permette l’accesso a un maggior numero di informazioni che possono influenzare i rapporti di potere tra governi e cittadini, tra aziende e clienti, e nel nostro caso tra redazioni giornalistiche e audience. L’accesso ad un maggior numero di informazioni sui processi può essere visto positivamente per rafforzare la fiducia, rendendo visibile il lavoro nascosto di un governo, di un’azienda o di una redazione, ma può anche avere risvolti negativi come la scoperta di vantaggi competitivi da parte di attori che competono o il dover fronteggiare dei costi per rendere conoscibile il proprio operato senza però avere ritorni economici”¹²⁶. Così una ricerca del 2017 definiva la trasparenza degli algoritmi nel mondo di giornalismo. Come sappiamo la “disciplina” nacque nel 2009 e già negli anni successivi emersero i primi dilemmi etici e le questioni da affrontare con l’emergere di tali tecnologie. La trasparenza è forse quella più importante e allo stesso tempo quella di più complessa risoluzione. Come sottolineavano lo stesso Diakopoulos e Koliska, la trasparenza non è solo un discorso che riguarda l’*accountability* o la fiducia nelle istituzioni, nelle aziende o nelle redazioni, ma potrebbe ledere il diritto alla proprietà intellettuale e al segreto industriale. La conoscibilità dei dati è forse il più importante degli obiettivi in termini di bilanciamento di diversi diritti. Andiamo a vedere nello specifico come possano essere ben bilanciati.

3.2.1.1 La conoscibilità dei dati

I dataset che vengono utilizzati per far funzionare al meglio un’intelligenza artificiale sono spesso “prelevati” direttamente dal web. Negli esempi presentati nel precedente capitolo abbiamo più volte sottolineato come prelevare dati direttamente dal web sia una delle opzioni più gettonate poiché permette di attingere ad una sorta di dataset infinito e in continua espansione, senza costi di raccolta e selezione. Soprattutto per settori come l’oggetto del nostro studio è intuitivo comprendere come alcune tecnologie non possano prescindere dall’essere costantemente aggiornate. Come faccio a capire le tendenze future se non sono aggiornato costantemente su cosa accade sul web? Come posso aumentare la mia precisione e la mia somiglianza all’essere umano se non imparo dall’interazione con esso? L’attingere in maniera automatica dal web però, nonostante possa essere più semplice ed economico ed avere questi evidenti vantaggi, aumenta drasticamente i rischi. Oltre al rischio di attacchi informatici e alla vulnerabilità per i virus, che possono essere fronteggiati con opportuni sistemi di sicurezza, il collegarsi al web comporta problemi in termini di poca differenziazione tra tecnologie dello stesso tipo e nei rischi di incorrere in bias. Se tutti attingono dallo stesso dataset, daranno tutti output molto simili tra loro. La differenziazione degli input e una loro raccolta filtrata e selezionata avrebbe non pochi vantaggi in termini di originalità e eticità del software. Sarebbe difficile però bilanciare un processo di questo tipo con l’apprendimento e l’aggiornamento costante. La soluzione comporterebbe non

¹²⁶Nicholas Diakopoulos & Michael Koliska, *Algorithmic Transparency in the News Media*, Digital Journalism, 5:7, 809-828, 2017

pochi costi maggiori, nel senso che oltre ad essere necessaria la supervisione umana, dovrebbe aumentare anche la frequenza di questa supervisione, rendendola costante. Nel primo capitolo abbiamo visto che alcune tecnologie di *Artificial Neural Network* e *Deep Learning* hanno il vantaggio di auto modificare i propri input sulla base degli output¹²⁷, meccanismo che non rende necessaria la supervisione e che permette di risparmiare. La soluzione potrebbe quindi essere una via di mezzo, cioè una supervisione che non sia costante ma abbastanza frequente. Se nell'automigliorarsi i software predisponessero diverse opzioni già formattate da dover selezionare, questo potrebbe permettere sia di risparmiare in termini di tempo, sia di evitare errori che di differenziarsi sulla base di scelte umane. Tutti questi aspetti, che potremmo definire di "filtraggio" e "monitoraggio" degli input, permettono di fronteggiare i fattori di rischio all'origine, in modo da tutelarsi *ex ante* e non dover gestire successivamente le problematiche scaturite da processi inappropriati. La giurisprudenza europea ci dice però che questo non basta e che debbano essere comunque previsti programmi appositi di gestione dei rischi *ex post*.

Per quanto riguarda invece il processo in sé e per sé, la conoscibilità di esso è fondamentale. Gli ostacoli che si presentano in questo passaggio però sono a volte intrinseci nei software stessi. Ci sono delle tecnologie, strutturate con sistemi *black box*¹²⁸, che non permettono di conoscere alcune parti di codice, come se esso fosse parte intrinseca della tecnologia stessa. Questo crea non pochi problemi perché se neanche il programmatore è a conoscenza di anche una piccola parte del processo, potrebbe essere compromessa sia l'individuazione delle cause di determinati output, sia la loro comprensibilità al proprio pubblico di riferimento. Zuddas¹²⁹ riassume al meglio tre possibili cause di opacità degli algoritmi:

- La difficile comprensibilità a chi non ha la competenza tecnica per comprenderli, anche se resi il più possibile potabili in termini non tecnici
- La tutela di interessi giuridici confliggenti come quella dei dati trattati
- L'autoapprendimento rende spesso ignoti i processi agli stessi programmatori

Il primo ostacolo può essere fronteggiato con una maggiore competenza tecnica nel settore¹³⁰ e una maggiore digitalizzazione generale dei pubblici di riferimento, nonché della popolazione in toto. Non è necessario puntualizzare in questa sede su un dibattito trasversale e in corso da anni in merito alla digitalizzazione, sempre più necessaria parallelamente ad un adattamento dei percorsi didattici al progresso tecnologico. Dall'altra parte però è necessario che le nozioni tecniche vengano rese comprensibili ai più, rendendo semplici delle nozioni difficili, utilizzando termini di dominio pubblico.

¹²⁷Stiamo parlando dei Recurrent Neural Network vedi capitolo 1 paragrafo 1.1 sottoparagrafo 1.1.2

¹²⁸Nella teoria dei sistemi, un modello black box è un sistema che, similmente ad una scatola nera, è descrivibile essenzialmente nel suo comportamento esterno ovvero solo per come reagisce in uscita (*output*) a una determinata sollecitazione in ingresso (*input*), ma il cui funzionamento interno è non visibile o ignoto (Wikipedia)

¹²⁹Paolo Zuddas, *Brevi note sulla trasparenza algoritmica*, AmministrazioneInCammino, par. 5, 5 giugno 2020

¹³⁰Per quanto riguarda le competenze tecniche nel giornalismo vedi Capitolo 4

La seconda problematica può essere ovviata cercando di predisporre controlli e processi alla tutela dei dati e dei diritti dei singoli. Il GDPR¹³¹, testo di riferimento in tema di privacy, prevede che “al fine di garantire un trattamento corretto e trasparente nel rispetto dell'interessato, tenendo in considerazione le circostanze e il contesto specifici in cui i dati personali sono trattati, è opportuno che il titolare del trattamento utilizzi procedure matematiche o statistiche appropriate per la profilazione, metta in atto misure tecniche e organizzative adeguate al fine di garantire, in particolare, che siano rettificati i fattori che comportano inesattezze dei dati e sia minimizzato il rischio di errori e al fine di garantire la sicurezza dei dati personali secondo una modalità che tenga conto dei potenziali rischi esistenti per gli interessi e i diritti dell'interessato e che impedisca tra l'altro effetti discriminatori nei confronti di persone fisiche sulla base della razza o dell'origine etnica, delle opinioni politiche, della religione o delle convinzioni personali, dell'appartenenza sindacale, dello status genetico, dello stato di salute o dell'orientamento sessuale, ovvero che comportano misure aventi tali effetti. Il processo decisionale automatizzato e la profilazione basati su categorie particolari di dati personali dovrebbero essere consentiti solo a determinate condizioni”¹³².

L'opacità intrinseca delle tecnologie invece ha dato vita ad un nuovo vero e proprio campo di ricerca, conosciuto come *eXplainable Artificial Intelligence (XAI)*¹³³, che prescinde da sistemi Black Box e prevede una progettazione delle IA trasparente sin dal principio. Questo permetterebbe di raggiungere, o quantomeno di avvicinarsi, a quella *full disclosure* tanto ambita sia a livello tecnico che a livello legislativo, e di far sì che l'addestramento stesso delle IA possa essere seguito e controllato in tutti i suoi passaggi. Gli *Hidden Markov Models* non permettono questa possibilità in quanto sono contraddistinti da valutazioni di probabilità basate su variabili non osservabili. Sono molto utili in determinati ambiti ma non vengono utilizzati in ambiti in cui è prevista o necessaria per una maggiore legittimazione di fronte al proprio pubblico di riferimento (i lettori) la trasparenza algoritmica¹³⁴.

3.2.1.2 La trasparenza dell'addestramento

Le tecnologie ML e DL sono ormai quelle quantitativamente più utilizzate in numerosi settori produttivi. L'aumento dell'automazione con l'obiettivo di diminuire il più possibile l'intervento umano è stato per anni lo scopo primario di coloro che investivano nelle intelligenze artificiali. Lungi dal contestare che l'efficientamento dei costi sia un obiettivo da perseguire, si vuole presentare in questa sede un'integrazione di esso con la trasparenza dei processi. Spesso sono obiettivi in contrasto tra loro, poiché meccanismi *black box* prevedono una non piena conoscibilità ma un risparmio in termini di capitale umano investito.

¹³¹Il “General Data Protection Regulation” è il Regolamento europeo sulla privacy dei dati che è entrato in vigore a partire dal 25 maggio 2018. Questo regolamento è stato fissato dalla Commissione europea con la finalità di armonizzare e rafforzare la protezione dei dati personali di tutti i cittadini dell'Unione Europea, sia all'interno dell'Ue, sia all'esterno.

¹³²Condizioni a cui sono consentiti il trattamento dei dati e la profilazione secondo il GDPR

¹³³Amina Adadi, Mohammed Berrada, *Peeking Inside the Black-Box: A Survey on Explainable Artificial Intelligence (XAI)*, IEEE access, 6, 52138-52160, 12 ottobre 2018

¹³⁴Sono utilizzati in Bioinformatica dove la legittimazione non è necessaria. Non bisogna commettere l'errore di confondere la trasparenza con l'accuratezza poiché i sistemi in cui il processo non sia destrutturabile non vuol dire che non siano precisi

Come già detto in generale però, un bilanciamento delle due cose permette di raggiungere quell'equilibrio necessario tra risparmio e rispetto delle regole etiche e giuridiche. L'addestramento è il cuore di queste tecnologie più complesse. Da esso dipende l'accuratezza e la complessità di un'intelligenza artificiale. Nel momento in cui abbiamo presentato ChatGPT¹³⁵ abbiamo sottolineato come il dataset di partenza fosse quantitativamente enorme nonché qualitativamente differenziato, entrambi fattori che hanno permesso a questa tecnologia di essere considerata la migliore esistente tra quelle di scrittura automatica. Il suo addestramento però, nonostante sia stato fatto con l'obiettivo di evitare che essa incorresse in bias cognitivi, ha comunque incontrato problemi di natura etica. L'addestramento di ChatGPT è stato coadiuvato da una serie di dati raccolti direttamente da individui che giudicavano determinati frasi o immagini come razziste o denigratorie. I problemi etici sono scaturiti non dalla pratica in sé, che può essere un giusto filtro da applicare per bilanciare con il giudizio umano la raccolta di dati dal web, ma da come questi test sono stati condotti. I dati raccolti tramite questa procedura, oltre che permettere alla macchina di avere un dataset "filtrato" dal giudizio umano ed evitare così di incorrere in bias denigratori per determinate categorie, permettono di tutelarsi dall'interazione con l'utente.

Un altro fattore di rischio, per i software di scrittura automatica in versione AP, è l'interazione stessa con gli utenti. Essa espone a rischi quali la maliziosità degli utenti stessi nell'operare ai fini di un miglioramento in senso negativo del software: alcuni utenti, quale che fosse il loro intento, hanno appositamente inserito input all'interno di un software con frasi razziste, piene di frasi ingiuriose e che si allontanavano notevolmente dall'idea di pudore, in modo da rendere l'IA "addestrata ciò" e quindi predisposta a questo tipo di condotte¹³⁶. In ChatGPT questo non è stato possibile proprio per l'inserimento di dati ad hoc durante l'addestramento.

Un altro problema che scaturisce dall'utilizzo di queste tecnologie è il loro impatto ambientale. I server necessari al funzionamento hanno un impatto ambientale non irrilevante.

¹³⁵Motivo 71 del GDPR

¹³⁶Emanuele Capone, *Le intelligenze artificiali fra razzismo e questione etica*, La Repubblica, 6 aprile 2021

3.2.2 L'impatto ambientale delle intelligenze artificiali

Un aspetto che viene spesso trascurato quando si parla di intelligenze artificiali è l'impatto ambientale che essi hanno. Nel parlare di intelligenze artificiali e impatto ambientale, il termine più appropriato che identifica il rapporto tra questi due concetti è probabilmente "paradosso". Le intelligenze artificiali contribuiscono a fornirci più dati possibili sul cambiamento climatico e sull'inquinamento. Gli esempi riportati nel secondo capitolo sono solo alcune delle applicazioni di monitoraggio ambientale che vengono costantemente impiegate e senza le quali non potremmo fronteggiare al meglio una delle sfide più importanti del nostro secolo. Dall'altro lato però queste tecnologie utilizzano spesso data center di non indifferente grandezza che oltre a produrre calore, impiegano ingenti quantità di risorse. Secondo uno studio dell'Università del Massachusetts Ahmlet¹³⁷, creare una singola intelligenza artificiale provoca l'emissione di 284 tonnellate di anidride carbonica, cinque volte tanto l'emissione media di una macchina durante tutto il suo ciclo di vita.

Queste emissioni vengono prodotte per lo più durante la fase di addestramento. Come detto, attraverso lo sviluppo di tecniche di *reinforcement learning* è possibile scaricare sull'AI l'intero processo di ottimizzazione, ma questo avrebbe implicazioni sulla trasparenza del processo. Sistemi automatizzati che non vanno a incidere sull'addestramento in sé ma sugli effetti negativi di esso possono essere utilizzati riducendo al minimo gli effetti indesiderati. Google grazie agli algoritmi di DeepMind¹³⁸ ha ceduto totalmente il controllo del raffreddamento dei server nei propri datacenter all'intelligenza artificiale riducendo i costi e diminuendo l'impatto ambientale. Molte sono le strategie per ridurre i costi sul raffreddamento dei DataCenter: Facebook, ad esempio, li posiziona in zone fredde, come a Lulea in Svezia, o dove è possibile sfruttare al meglio energie rinnovabili, come quella eolica a Clonee, in Irlanda. Da questo punto di vista Microsoft ha avviato probabilmente il progetto più innovativo: immettere i datacenter nell'oceano in modo da poter quasi annullare i costi di raffreddamento e ridurre la probabilità di guasti causati da esso.



Fig. 14 Server sottomarini Microsoft¹³⁹

¹³⁷Mara Magistroni, *Le intelligenze artificiali inquinano più del previsto*, WiredItalia, 7 giugno 2019

¹³⁸DeepMind è un'azienda inglese di intelligenza artificiale controllata da Alphabet. La società ha sede a Londra, con centri di ricerca in Canada, Francia e Stati Uniti. Fondata nel 2010 come DeepMind Technologies, viene acquisita da Google nel 2014 (Wikipedia)

¹³⁹media.datacenterdynamics.com

Inoltre, più si sviluppa a livello tecnico il settore, più i processi di addestramento saranno efficienti e meno inquinanti. I dati sulle emissioni però sono negativi non solo per l'addestramento ma anche per l'impatto ambientale in generale: i data center inquinano per il 5% delle emissioni complessive e si stima che entro 5 anni si quadruplicheranno. Si constata poi che il 96% dei dati presenti sul web siano vecchi, definibili come spazzatura¹⁴⁰. Quante volte ci capita di tenere, con la scusa che gli spazi cloud siano infiniti¹⁴¹, foto vecchie o e-mail degli scorsi anni. Tutta questa mole di dati riempie quasi la totalità dello spazio web. Inoltre, spesso questi spazi sono di proprietà di big tech a danno di un business del cloud che promuova l'eticità. I costi di avviamento proibitivi per la creazione di risorse interne, costringe i gruppi poveri di risorse a fare affidamento su servizi di cloud come AWS, Google Cloud e Microsoft Azure. Sebbene questi servizi forniscano risorse di calcolo preziose, flessibili e spesso relativamente ecologiche, è più conveniente per i ricercatori accademici, che spesso lavorano per istituzioni educative senza scopo di lucro e la cui ricerca è finanziata da enti governativi, costruire centri di calcolo condivisi a livello di agenzie di finanziamento, come la National Science Foundation degli Stati Uniti. La maggior parte dei modelli di addestramento poi sono stati sviluppati al di fuori del mondo accademico. Limitare questo tipo di ricerca ai laboratori dell'industria danneggia la comunità di ricerca sulla NLP in molti modi. In primo luogo, soffoca la creatività. I ricercatori con buone idee ma senza accesso a calcoli su larga scala non saranno in grado di realizzare le loro idee, e saranno costretti a concentrarsi su problemi diversi. In secondo luogo, vieta alcuni tipi di ricerca sulla base dell'accesso alle risorse finanziarie. In questo modo si favorisce ancora di più il già problematico ciclo di finanziamento della ricerca "chi è ricco diventa più ricco", in cui i gruppi che hanno già successo e che quindi sono ben finanziati tendono a ricevere i finanziamenti¹⁴².

¹⁴⁰Emily Chasan, The environmental cost of keeping mail and files online keeps rising, The Japan Times, 26 gennaio 2020

¹⁴¹In realtà anche lo spazio web è una risorsa scarsa e dipende dagli investimenti stanziati per le infrastrutture fisiche

¹⁴²Emma Strubel, Ananya Ganesh, Andrew McCallum, *Energy and Policy Considerations for Deep Learning in NLP*, College of Information and Computer Sciences, 5 giugno 2019

CAPITOLO 4 Il cambiamento della professione giornalistica e la sua percezione

In questo capitolo tenteremo di analizzare come tutti i cambiamenti, il progresso tecnologico e le questioni dibattute in materia abbiano influito sulla professione giornalistica in sé e su come il cambiamento di essa abbia impattato sul pubblico di riferimento, cioè i lettori in generale. Analizzeremo prima, sulla scia delle questioni etiche affrontate nel precedente capitolo, i fondamenti valoriali e legislativi della professione, andando poi ad analizzare le novità e l'impatto che le intelligenze artificiali hanno avuto sul settore. Discuteremo anche della percezione dell'opinione pubblica su questi cambiamenti, riportando diverse ricerche sul confronto tra prodotti editoriali di un giornalista e di un'intelligenza artificiale.

4.1 Deontologia

Nonostante la deontologia sia una questione propriamente etica, riguardante i principi e le regole, si è preferito discuterne in questo capitolo al fine di capire come essi si debbano necessariamente adattare ad un lavoro giornalistico che sempre di più si trasformerà. La legge che incide sulla professione giornalistica è la n. 69 del 1963. È intuibile comprendere che in quell'anno la professione giornalistica era completamente diversa da quella attuale. Abbiamo visto nel capitolo tre come la rivoluzione digitale prima e le intelligenze artificiali poi abbiano cambiato oltre che la professione in sé anche il regime giurisprudenziale dell'ambito giornalistico. I diritti e i doveri di un giornalista però, nonostante si debbano necessariamente integrare con i cambiamenti, rimangono sempre quelli dettati dalle norme del XX secolo. La libertà d'informazione e il diritto di critica sono stati già discussi precedentemente. Sottolineiamo ora importanti principi che, alla luce dei cambiamenti, sono forse quelli che si sono dovuti maggiormente adattare al progresso. L'articolo 2 della legge professionale del 1963, oltre a ricordarci che la libertà di informazione e di critica sono diritti insopprimibili, formalizza l'obbligo inderogabile alla verità sostanziale dei fatti osservati, l'obbligo di rettifica delle notizie che risultino inesatte, la riparazione di eventuali errori, il rispetto del segreto professionale sulla fonte delle notizie, quando ciò sia richiesto dal carattere fiduciario di esse, e la promozione della fiducia tra stampa e lettori. La verità sostanziale dei fatti, nonostante sia inderogabile e democraticamente presunta, ha subito una distorsione enorme negli ultimi decenni. Dalla fine degli anni Novanta ha iniziato a diventare sempre più di uso comune il termine *post-verità*¹⁴³, indicando la propensione dell'opinione pubblica ad accettare notizie come vere solo sulla base di emozioni¹⁴⁴ o della leva sui propri pregiudizi¹⁴⁵. Questa propensione ha portato negli anni alla diffusione esponenziale di *fake news*, problema divenuto ormai di importanza centrale per il mondo dell'informazione.

¹⁴³La post-verità è un termine ombrello volto ad indicare la degenerazione del mondo dell'informazione

¹⁴⁴Martel C, Pennycook G, Rand DG. Reliance on emotion promotes belief in fake news. *Cogn Res Princ Implic*. 2020;5(47):1–20. pmid:33026546

¹⁴⁵*Anatomia di una fake news: l'importanza della disinformazione sulla nostra società, Confindustria*, maggio 2022

4.1.1 Le fake news

Napoleone non era basso. Maria Antonietta non ha mai detto di dare al popolo le brioche. Einstein non era un cattivo studente in matematica. Questi sono solo alcuni degli esempi storici che nel tempo sono divenuti parte dei racconti popolari su alcune figure illustri della storia mondiale ma che in realtà non hanno nessun fondamento storico. Perché allora si sono diffuse come vere? Questa domanda non ha una risposta semplice. Prima di analizzare il problema nello specifico è necessaria una puntualizzazione concettuale. La disinformazione è la creazione o la diffusione intenzionale di notizie false. La misinformazione è la creazione o diffusione accidentale di notizie false. In entrambi i casi è intuibile il danno che si reca a livello informativo alla società. Risulta quindi opportuno, per i cittadini in generale, oltre che per i giornalisti stessi, avere gli strumenti adeguati a fronteggiare questo tipo di dinamiche che possono ledere la valenza informativa delle notizie e il ruolo dell'informazione all'interno di una società democratica.

4.1.1.1 La psicologia delle fake news

Spesso tendiamo a voler sembrare attendibili in quello che diciamo mostrando di sapere aneddoti o eventi solo per il gusto di dimostrare qualcosa. O semplicemente abbiamo creduto ad una notizia accattivante riportata da un amico o un parente, diffusa sul web con l'intento di avere risonanza. Le notizie false sono sempre esistite ma mai come oggi vengono diffuse in lungo e in largo, spesso con dei secondi fini. L'intento della risonanza, e quindi di un ritorno in termini economici o di influenza dell'opinione pubblica, ha ora più che mai strada libera vista la dimensione e la pervasività della rete. Spesso sono progettate proprio per essere virali cercando di fare leva su sentimenti quali la rabbia, la solidarietà, l'indignazione o la sorpresa. In generale la nostra mente agisce secondo ciò che viene chiamata "economia cognitiva". A differenza dei database, la mente umana ha dei criteri, talvolta inconsci, attraverso cui filtra le informazioni che riceve, riuscendo poi a generalizzare in maniera quasi automatica su determinati aspetti della nostra vita. Se ad esempio devo decidere dove fare la spesa, non penserò in quale luogo fisico andare, ma a quale mercato o supermercato, dando per scontato che il luogo migliore per fare la spesa sia un mercato o un supermercato e non un ferramenta. Lo diamo per scontato proprio grazie ai meccanismi inconsci di economia cognitiva. Questi meccanismi funzionano sia per decisioni banali che per decisioni più complesse. Ovvio che se devo decidere che tipo di università frequentare uscito dal liceo, la riflessione sarà nella maggior parte dei casi profonda e prolungata. Se ad esempio però devo rispondere alla domanda "ci sono più politici uomini o donne?" risponderò tramite un'euristica¹⁴⁶ di disponibilità, partendo da quello che la mia esperienza e la mia percezione mi suggerisce quasi in maniera automatica, rispondendo probabilmente uomini solo perché in passato non vi erano le quote rosa in politica e sono convinto che il cambiamento sia ancora in atto. Tralasciando analisi psicologiche sulle euristiche e quale fosse la risposta alla domanda posta come esempio, ci interessa capire come spesso senza rendercene quasi conto agiamo in una determinata maniera senza riflettere troppo su cosa stiamo facendo. Questi meccanismi appena elencati portano alcune volte a degli errori, denominati *bias*, cioè errori che commettiamo per non

¹⁴⁶In psicologia le euristiche sono semplici ed efficienti regole che sono state proposte per spiegare come le persone risolvono, danno giudizi, prendono decisioni di fronte a problemi complessi o informazioni incomplete (Wikipedia)

sovraccaricare il nostro cervello e per quell'economia cognitiva di cui in breve abbiamo descritto le caratteristiche principali. Tutto questo accade anche quando la nostra mente riceve informazioni o notizie. Se la notizia ci arriva da un nostro parente o dal giornale che leggiamo tutti i giorni, tenderemo a non verificare l'attendibilità di essa e a darla per scontata. Uno dei meccanismi che in questi casi agisce da filtro è il *confirmation bias*, per il quale tendiamo a credere a quello di cui già siamo convinti. Se ho votato per un partito rispetto ad un altro, è istintivo reputare un politico di quel partito più affidabile rispetto ad un altro. La propensione di cui prima parlavamo, quella di accettare notizie sulla base di criteri diversi da quelli comunemente usati per definire attendibile una fonte, è suscitata anche da una tendenza umana a rimanere nella propria bolla informativa¹⁴⁷. Il nostro cervello ritiene più "comodo" fare così piuttosto che esplorare altre alternative e mettere in discussione le proprie credenze e i propri pregiudizi. Le *fake news* sfruttano questo meccanismo. Tutti siamo soggetti a crederci. Per di più, in psicologia vi è un ulteriore "effetto" che paradossalmente ci rende più vulnerabili: l'effetto terza persona¹⁴⁸. Tendiamo a pensare che gli altri siano più vulnerabili di noi, più soggetti ad influenze e condizionamenti esterni, dando fiducia alla nostra percezione ed essendo convinti di essere poco influenzati ed influenzabili dall'ambiente circostante. Questo effetto è confermato anche da un recente studio¹⁴⁹ che dimostra come il 73% degli intervistati ritenga di saper distinguere notizie reali da notizie false, mentre il 35% è convinto che l'italiano medio sappia fare tale distinzione. La convinzione di essere poco vulnerabili ci rende quindi più vulnerabili e ci porta a non tutelarci da questi rischi con gli opportuni strumenti, credendo di averli già.

Ci sono diversi studi che constatano come, oltre a queste caratteristiche che ci portano ad essere predisposti e vulnerabili in quanto umani all'accettazione di *fake news*, ci siano alcune caratteristiche individuali che predispongono di più o di meno all'accettazione di esse. Alcuni ricercatori parlano di propensione ad accettare acriticamente affermazioni con basso valore epistemico, etichettando tale caratteristica come *reflexive open-mindedness*, da contrapporre agli individui più analitici, dotati di *reflective open-mindedness*. Bronstein e altri ricercatori¹⁵⁰ hanno aggiunto tendenze cognitive come l'ideazione illusoria, il dogmatismo e il fondamentalismo religioso all'elenco dei tratti a livello individuale debolmente associati a una maggiore credenza nelle *fake news*, mentre il pensiero analitico e aperto riduceva leggermente questa credenza. Anche il vedere più volte una notizia ci porta ad assumerla come vera¹⁵¹. La diffusività però non è sempre sinonimo di verità in quanto, come detto, le *fake news* sono spesso progettate per essere virali. Un fattore importante per

¹⁴⁷Capitolo 3 Paragrafo 2.3 Sottoparagrafo 2.3.1

¹⁴⁸Pennycook G, Rand DG. *Lazy, not biased: Susceptibility to partisan fake news is better explained by lack of reasoning than by motivated reasoning*, *Cognition*, 188:39–50, 2019

¹⁴⁹Sondaggio realizzato da Ipsos per IDMO (Italian Digital Media Observatory) l'hub nazionale contro la disinformazione, coordinato dal centro di ricerca Data Lab dell'Università Luiss Guido Carli e formato da Tim, Rai, Università di Tor Vergata, Gruppo Gedi la Repubblica, i fact-checkers NewsGuard e Pagella Politica e il think tank T6 Ecosystems presso un campione casuale nazionale rappresentativo della popolazione italiana tra i 18 e i 65 anni secondo genere, età, livello di scolarità, condizione occupazionale, area geografica di residenza e dimensione del comune di residenza. Sono state realizzate 1.000 interviste, condotte con tecnica CAWI (Computer Assisted Web Interview) tra l'1 e il 4 febbraio 2022.

¹⁵⁰Bronstein MV, Pennycook G, Bear A, Rand DG, Cannon TD. *Belief in Fake News is Associated with Delusionality, Dogmatism, Religious Fundamentalism, and Reduced Analytic Thinking*. *J Appl Res Mem Cogn.* ,8(1):108–17, 2019

¹⁵¹Pennycook G, Cannon TD, Rand DG., *Prior Exposure Increases Perceived Accuracy of Fake News*, *J Exp Psychol Gen.* 2018,147(12):1865–80, 2018

discernere notizie vere da notizie false è la conoscenza procedurale delle notizie¹⁵², cioè il conoscere come funziona il processo informativo. Tutti questi fattori individuali che contraddistinguono una maggiore o una minore propensione al valutare *fake news* come vere, pur avendo premesso che siamo tutti potenziali vittime di questo fenomeno, sono attribuibili ad un maggiore o ad un minore livello di istruzione. Tra l'altro, chi è meno scolarizzato attribuisce un'immagine complessivamente più negativa e/o erronea alla scienza di chi è più istruito.

4.1.1.2 Scienza e fake news

La scienza produce continuamente materiale volto ad informare la popolazione su progressi e attività scientifiche e di ricerca spesso correlati a scoperte sensazionali o disastrose. I media spesso diffondono nel corso della giornata informazioni scientifiche che possono enfatizzare o meno le dimensioni del rischio e della paura. Tutti siamo alla ricerca di rassicurazioni da parte della scienza. Secondo il giornalista Francesco Giorgino¹⁵³, si era sviluppato negli anni antecedenti la pandemia un divario tra l'importanza assunta dalla scienza nella vita di ognuno e il persistere di indicatori di scarsa informazione e di scarso interesse per la ricerca, per responsabilità soprattutto dei media. Quello che egli auspicava era un processo di mediatizzazione della scienza, in modo da adattare e trasformare le notizie sulla base delle esigenze cognitive di tutti, in alcuni casi anche quelle emozionali. La pandemia ha cambiato questa dinamica? Apparentemente sembrerebbe di sì. Più che aumentare la fiducia nella scienza o nei media, il dibattito sulle *fake news* si è spostato dalla politica alla scienza e alla medicina¹⁵⁴, spostando il focus dell'individuazione delle *fake news* su questo tipo di notizie e contribuendo a quella mediatizzazione tanto auspicata. Il Covid-19 ha portato con sé un enorme dibattito sulla validità scientifica dei vaccini e sull'attendibilità dell'informazione scientifica in sé. È necessario sottolineare che nonostante ci fossero nicchie di popolazione che non credevano nell'efficacia dei vaccini e del contributo della scienza per fronteggiare la pandemia, la maggior parte della popolazione italiana si è comunque vaccinata, ridando lustro non solo al contributo che la scienza può dare alla società, ma anche all'informazione scientifica di per sé. Il dibattito, nonostante fosse già presente prima della pandemia¹⁵⁵, ha avuto molta più risonanza dopo il periodo delle campagne vaccinali. Si è posto il tema di quanto e come diffondere il sapere scientifico da parte delle istituzioni, ristabilendo quel rapporto trasversale con la comunità di cittadini. In molte auspicano la creazione di nuove arene di confronto partecipativo per la comunicazione della scienza come ambito di (ri)costruzione del rapporto fra cittadini e istituzioni del sapere scientifico. In questo contesto può l'intelligenza artificiale aiutare il cittadino a non essere fuorviato dalle *fake news*?

¹⁵²Amazeen MA, Bucy EP. Conferring Resistance to Digital Disinformation: The Inoculating Influence of Procedural News Knowledge. *J Broadcast Electron Media*, 63(3):415–32, 2019

¹⁵³Francesco Giorgino, *Giornalismo e Società. Informazione, politica, economia, cultura*, febbraio 2017

¹⁵⁴Fake news, *ItaliaInDati*, italiaindati.com, 2020

¹⁵⁵Ceravolo F., Rostan M., *Raccontare la scienza: opportunità e rischi nella società dell'informazione*, in «Cambio. Rivista sulle trasformazioni sociali» Vol. 9, n. 18: 5-8., 2019

Essa può analizzare titoli, contenuti, commenti e fonti ma non può rilevare l'intenzionalità di una notizia falsa¹⁵⁶. Vedremo però che sia a livello giuridico sia a livello etico, fa poca differenza se una notizia è falsa volutamente o accidentalmente: il rischio per la democrazia rimane sempre lo stesso.

4.1.1.3 Il regime giuridico delle fake news

L'articolo 656 del nostro Codice penale prevede l'arresto fino a tre mesi o l'ammenda fino a 309 € per chi diffonde notizie false, esagerate o tendenziose, "atte a turbare l'ordine pubblico", con pene più gravi in caso di disfattismo politico, aggrigotaggio, attività antinazionale di un cittadino all'estero, procurato allarme presso le autorità, abuso della credulità popolare o diffamazione. Inoltre, il codice non distingue tra persone che creano notizie false o che diffondono tali notizie. Le pene sono le stesse sia per chi diffonde accidentalmente notizie false e per chi le crea intenzionalmente. Il suddetto articolo definisce la notizia come "un annuncio o un'informazione dal contenuto in ogni caso preciso e riconoscibile, stante la non configurabilità della norma in esame in relazione a semplici dicerie". A livello europeo si è recentemente rafforzato il codice di condotta sulla disinformazione e contiene 44 impegni e 128 misure specifiche, nelle seguenti aree:

- Demonetizzazione e trasparenza della pubblicità politica
- Garantire l'integrità dei servizi e responsabilizzare gli utenti e i ricercatori
- Potenziare la comunità del fact-checking
- Creare un centro per la trasparenza e rafforzare il monitoraggio

Il Codice di condotta sulla disinformazione del 2018 ha riunito per la prima volta gli operatori del settore, secondo il già decantato approccio collaborativo, a livello mondiale, al fine di impegnarsi a contrastare la disinformazione. Al centro della strategia dell'UE contro la disinformazione, il Codice si è dimostrato uno strumento efficace per limitare la diffusione della disinformazione online, anche durante i periodi elettorali e per rispondere rapidamente alle crisi, come la pandemia di coronavirus e la guerra in Ucraina. In Italia ci sono stati diversi tentativi di formalizzare meglio le fattispecie della disinformazione e della misinformazione. Nel 2017 il d.d.l. Gambaro voleva introdurre una art. 656-bis con l'aggiunta del mezzo "piattaforme informatiche, andando a configurare un reato più specifico. Il d.d.l. Zanda-Filippin proponeva di introdurre forme di responsabilizzazione per i fornitori di servizi di social network sull'impronta del modello sanzionatorio amministrativo tedesco e il d.d.l. De Girolamo che voleva introdurre il divieto di anonimato online. Durante il secondo Governo Conte si è inoltre proposta una "Commissione parlamentare di inchiesta sulla diffusione seriale e massiva di contenuti illeciti e di informazioni false attraverso la rete di internet, le reti sociali telematiche e le altre piattaforme digitali. Alla luce della recente esplosione dei partiti populistici, secondo alcuni giuristi l'attenzione alla diffusione di notizie false o decontestualizzate deve essere ancora maggiore e riguardare anche e soprattutto l'ambito politico e le campagne elettorali¹⁵⁷.

¹⁵⁶Luigi Ceccarini, Taina Pihlajarine, Anette Alén-Savikko, *Digital Citizenship and artificial intelligence: Information and disinformation*, European Journal of Communication, Vol. 37(5) 563-568, 2022

¹⁵⁷Tommaso Guerini, *Fake news e diritto penale. La manipolazione digitale del consenso nelle democrazie liberali*, Itinerari di Diritto Penale, 2020

4.1.2 Le intelligenze artificiali per le fake news

Le *fake news* possono provenire, purtroppo, anche da organi di informazione ufficiali: giornali, TG e testate online. Il giornalista, quindi, non è estraneo a tutte queste dinamiche. Anzi, il ruolo che dovrebbe avere, alla luce di quanto detto sui principi deontologici, è anche quello di *fact-checker*. Il ruolo del giornalista, quindi, nonostante abbia perso in parte l'accezione di *gatekeeper*, può ancora essere quello di intermediario tra chi legge gli articoli e la realtà, tramite un filtraggio delle notizie, derubricando quelle false e promuovendo la diffusione di quelle vere.

Il rilevamento delle *fake news* in generale, a prescindere dall'utilizzo di software automatizzati, risulta essere un compito intrinsecamente difficile in quanto le notizie false sono spesso progettate in modo da sembrare vere e in modo da diventare virali. Sui social questo è ancora più difficile in quanto spesso si presuppone l'aiuto degli utenti che in realtà non sempre posseggono gli strumenti cognitivi adatti per poterle rilevare.

Nonostante si sia detto che la richiesta di informazione a pagamento sia aumentata, la percentuale di persone che si informano sui social rimane comunque la maggioranza in quanto mezzo più tempestivo, più accessibile e meno costoso. In più le dinamiche stesse dei social, come sottolineato nel capitolo tre, incentivano la discussione accesa ed è quindi più facile, oltre che diffondere e ricondividere notizie, anche commentarle e discuterne come se fossero vere. Il filone di ricerca che ha affrontato tale problema a livello sia teorico-concettuale che pratico, parte dal presupposto che le *fake news* abbiano caratteristiche diverse dalle notizie vere e che queste caratteristiche siano rilevabili, anche tramite le intelligenze artificiali.

Gli approcci generali, in merito alle modalità che si utilizzano a priori per scovare le notizie, riguardano la decisione su chi dovrà svolgere tale compito. Esse si dividono in:

- *Expert-oriented*: gli esperti dei sistemi informativi e medialti nonché specializzati in *fake news* negli ultimi anni possiedono gli strumenti opportuni per rilevare questo tipo di notizie e saranno loro, come ad esempio i siti di *debunking*¹⁵⁸, a svolgere questo compito
- *Crowd-sourcing*: gli utenti, nel momento in cui dubitano della veridicità della notizia, sfruttano la caratteristica trasversale del web in generale e dei social nello specifico al fine di incentivare il dibattito non sull'argomento in sé, ma sulla veridicità di esso
- *Computational*: sistemi automatizzati, attraverso strumenti concettualmente simili a quelli degli esperti ma notevolmente più veloci e con maggiori capacità di analisi, riescono ad individuare le notizie false

Tra i tre approcci non ce n'è uno oggettivamente migliore e maggiormente utilizzato rispetto agli altri. Vanno sottolineati però i vantaggi e gli svantaggi di ognuno di essi. L'*expert-oriented* ha il vantaggio di avere a disposizione capacità umane quali l'intuizione e l'analisi delle emozioni su cui si fa leva con le *fake news*.

¹⁵⁸Opera di demistificazione e confutazione di notizie o affermazioni false o antiscientifiche, spesso frutto di credenze, ipotesi, convinzioni, teorie ricevute e trasmesse in modo acritico (Treccani). Il termine ha assunto negli anni un'accezione sempre più digitale, con la nascita di numerosi siti di debunking assimilabili oggi a siti di fact-checking. I due termini potrebbero essere anche sinonimi.

Per di più la collaborazione tra comunità di esperti e siti di *debunking* può portare a maggiori certezze sulla veridicità e sulla falsità delle notizie. Il *crowd-sourcing* ha il vantaggio di comprendere nell'analisi la partecipazione del cittadino. L'approccio *computational* ha il vantaggio della rapidità e della quantità nelle analisi che svolge. Un filone di ricerca in questo ambito sta tentando di avvicinare le capacità computazionali delle macchine anche a quelle emozionali dell'umano, tramite la *sentiment analysis*¹⁵⁹. Probabilmente l'approccio migliore è un uso coordinato di tutte e tre le tecniche, il quale però prevederebbe un ingente utilizzo di risorse. L'utilizzo più appropriato quindi in termini di costi/benefici, ancora una volta, è quello di lasciar operare le macchine con la supervisione dell'uomo e far sì che la decisione finale spetti ad esso. I software che analizzano le notizie sono molto simili a quelli di ricerca di notizie tramite immagini o parole analizzati nel secondo capitolo. La differenza sta nel fatto che, invece di ricercare contenuti per keywords o immagini simili, analizzano le notizie sulla base di caratteristiche specifiche.

Queste caratteristiche possono essere suddivise in tre macroaree di analisi:

- Chi ha pubblicato il contenuto
- Il contenuto
- A chi è destinato il contenuto

Il cuore dell'analisi è ovviamente il contenuto in sé ma non vanno sottovalutate anche le altre due parti.

L'analisi di chi ha pubblicato il contenuto concerne:

- L'analisi del nome: si cerca di individuare nomi simili, se esso possa essere vero, se non sia un furto di identità o un bot¹⁶⁰
- L'analisi del dominio: alcune volte le notizie sono diffuse da account falsi con domini inventati o che rimandano a siti intenzionalmente creati per essere falsi, i quali hanno domini simili ad altri di siti a scopo informativo, ma sono leggermente diversi.
- L'età e la locazione geografica: sono funzionali alla identificazione di profili simili o a furti d'identità finalizzati alla diffusione di *fake news*

L'analisi del contenuto prevede:

- Analisi del titolo: spesso le notizie in generale, ma quelle false ancora di più, utilizzano titoli accattivanti per il *click-bait*¹⁶¹, con l'obiettivo di fare più visualizzazioni e rendere virale la notizia. Spesso capita che non sia il contenuto stesso ad essere falso ma che il titolo sia fuorviante e non rispecchi il contenuto della notizia

¹⁵⁹Alonso M.A., Vilares D., Gómez-Rodríguez C., Vilares J., *Sentiment Analysis for Fake News Detection*, Electronics, 2021

¹⁶⁰Il bot (abbreviazione di robot) in terminologia informatica in generale è un programma che accede alla rete attraverso lo stesso tipo di canali utilizzati dagli utenti (per esempio che accede alle pagine Web, invia messaggi in una chat, si muove nei videogiochi, e così via). Programmi di questo tipo sono diffusi in relazione a molti diversi servizi in rete, con scopi vari, ma in genere legati all'automazione di compiti che sarebbero troppo gravosi o complessi per gli utenti (Wikipedia)

¹⁶¹Adescamento a visitare le pagine di un sito Web, finalizzato all'aumento delle rendite pubblicitarie (Treccani)

- Analisi del testo: per quando riguarda l'analisi testuale spesso si ricerca per scovare le *fake news* un linguaggio di opinione e "infiammatorio" volto a scatenare un dibattito acceso e a creare fazioni. La ricerca è spesso orientata anche a ricercare linguaggi adesivo/oppositivi sempre per la stessa finalità. Questo tipo di ricerca si divide in analisi lessicale, ricercando termini di un determinato tipo, e sintattica, ricercando frasi strutturate in una certa maniera
- Analisi delle immagini: le immagini anche possono essere false oppure possono essere utilizzate per contenuti errati o non affini alle immagini. Oltre alla correlazione tra immagine e contenuto, viene quindi analizzata anche la veridicità dell'immagine stessa attraverso parametri come la similarità, la chiarezza, l'associazione plausibile di pixel, istogrammi di distribuzione media di immagini¹⁶².
- Analisi delle fonti a cui si rimanda: spesso nelle notizie sono presenti link ad altre notizie o a determinati account o siti, o viene semplicemente citata la fonte dal quale si è attinto. L'analisi non deve quindi fermarsi al contenuto in sé ma anche ai suoi collegamenti, con lo stesso tipo di approccio.

L'analisi del pubblico riguarda la stessa procedura dei primi due ma per chi visualizza o commenta il contenuto. Ovviamente il fatto che profili falsi o bot visualizzino o commentino un contenuto non basta per etichettarlo come falso ma la quantità e le modalità di interazione di essi possono essere un campanello d'allarme per la verifica della veridicità del contenuto.

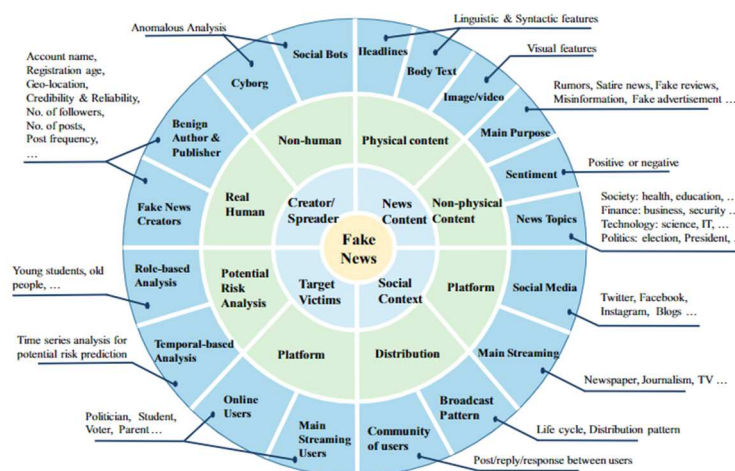


Fig. 2. Fake news and everything related to it.

Fig. 15 Funzionamento IA per la fake news detection
(Fake News Detection on Social Media: A Data Mining Perspective nota 163)

Uno dei casi esemplari di *fact-checking*, e anche uno dei primi, fu la campagna elettorale americana del 2016 durante cui alcuni indicatori che fecero saltare all'occhio determinati contenuti, successivamente rivelatisi falsi, furono proprio l'engagement eccessivamente elevato e la quantità surreale di utenti che commentò alcuni post¹⁶³.

¹⁶²È una sorta di inferenza statistica sulla struttura media delle immagini sul web considerando tutti gli elementi in gioco (colori, pixel ecc.). Se l'immagine si discosta da tale struttura media di superando un valore soglia viene considerata falsa

¹⁶³Kai Shu, Amy Sliva, Suhang Wang, Jiliang Tang, and Huan Liu, *Fake News Detection on Social Media: A Data Mining Perspective* Computer Science & Engineering, Arizona State University, Tempe, AZ, USA Charles River Analytics, Cambridge, MA, USA, 2017

Gli indicatori quantitativi e quelli temporali, quindi, sono anch'essi parametri utili e spesso utilizzati per la *fake news detection*. La quantità, la frequenza e il momento di effettiva pubblicazione o diffusione di contenuti possono essere importanti indicatori per una *fake news*. È stato dimostrato che i bot operano sempre in orari specifici essendo settati a priori e la frequenza con cui lo fanno è sempre la stessa. Bisogna stare attenti a non confondere ciò con i software di pubblicazione automatica: qui ritornano in gioco l'importanza della trasparenza e l'opportuna segnalazione dell'utilizzo di queste tecnologie. Nell'identificazione di profili falsi e bot si utilizzano anche altri indicatori come la percentuale di followers, la percentuale di followers fittizi e la percentuale di amici nel caso di celebrità.

È ovviamente importante, in tutti questi processi, il dataset che si utilizza. Se parliamo di Machine Learning e *Deep Learning* utilizzando un dataset accurato e ampio, con abbondanti esempi di notizie vere e di notizie false, si riuscirà a generalizzare in maniera precisa le caratteristiche di esse su base prettamente statistica, senza incorrere in *bias cognitivi* come quelli di conferma delle proprie convinzioni o di adesione/opposizione ad un determinato pensiero. L'intelligenza artificiale ha quindi il vantaggio di valutare sulla base non della notizia in sé, ma sulla base delle caratteristiche che contraddistinguono tale notizia come falsa e come vera. Ovviamente esistono diversi tipi di software di analisi per il fact-checking. Ci sono quelli che utilizzano solo dataset di esempi, altri che vengono impostati sulla base di parametri specifici, altri ancora che danno un peso maggiore ad una macroarea piuttosto che ad un'altra. Anche qui la soluzione migliore sarebbe combinare tutti gli approcci¹⁶⁴. Va fatta però una valutazione in termini di costi e vantaggi. Va sottolineato poi come questi software siano il riflesso dei parametri umani e come sia curioso che le intelligenze artificiali vengano utilizzate in questo caso per rilevare altre IA utilizzate però per scopi malevoli come la diffusione di notizie false. Una delle lacune di questi software, come accennato, è l'analisi dei sentimenti. Essa può risultare utile sia per la rilevazione delle notizie stesse che per i contenuti in sé. Rilevare attraverso le parole utilizzate e la sintassi adottata il "tono" e l'accezione che una frase o un commento hanno, è possibile tramite le macchine, ma è sicuramente il compito che esse svolgono meno bene¹⁶⁵. È di per sé difficile comprendere tali caratteristiche da un testo scritto, per di più la contestualizzazione risulta più complessa per questi software. Da ciò scaturisce che la rilevazione della disinformazione, e quindi anche delle notizie vere ma decontestualizzate e "falsamente distribuite", sia più complicata. Risultando ad oggi giuridicamente limitati i software di analisi emotiva, la nuova sfida del fact-checking è quella di migliorarsi attraverso la sentiment analysis senza incorrere nella violazione di diritti e rispettando le normative europee.

¹⁶⁴Kai Shu, Suhang Wang, Huan Liu, *Beyond News Contents: The Role of Social Context for Fake News Detection*, Twelfth ACM International Conference on Web Search and Data mining, 312-320, gennaio 2019

¹⁶⁵Alonso, M.A.; Vilares, D.; Gómez-Rodríguez, C.; Vilares, J. Sentiment Analysis for Fake News Detection. Electronics 2021

Per valutare le prestazioni degli algoritmi per il rilevamento delle *fake news*, sono state utilizzate diverse metriche di valutazione. Le metriche più utilizzate per il rilevamento delle *fake news* sono¹⁶⁶:

- Vero positivo (TP): quando le *fake news* predette sono effettivamente annotate come *fake news*;
- Vero Negativo (TN): quando le notizie vere previste sono effettivamente annotate come notizie vere;
- Falso Negativo (FN): quando le notizie vere previste sono in realtà annotate come *fake news*;
- Falso positivo (FP): quando le notizie false predette sono in realtà annotate come notizie vere.

$$\text{Precisione } P = \frac{|T P|}{|T P| + |F P|}$$

$$\text{Richiamo} = \frac{|T P|}{|T P| + |F N|}$$

$$\text{Accuratezza} = \frac{|T P| + |T N|}{|T P| + |T N| + |F P| + |F N|}$$

Tutti questi parametri sono direttamente proporzionali alla qualità del software utilizzato. L'accuratezza di diversi software è stata più volte testata e i risultati generali danno prova della loro indiscutibile validità. Una ricerca del 2022 ha analizzato e confrontato due algoritmi diversi di fact-checking per giudicarne l'accuratezza. Il primo, ANN, si basa in linea generale sulla frequenza di determinate parole, il secondo, BERT, sull'analisi semantica. Utilizzando lo stesso dataset di news mainstream, la ricerca ha dimostrato notevole similarità tra ANN e BERT. Il secondo presenta un più alto parametro nel riconoscimento della credibilità della fonte, ma i livelli di accuratezza sono in entrambi i casi molto alti (91,2% BERT, 82,75% ANN)¹⁶⁷. Un altro Software, IREX di Lore.Ai, viene utilizzato per rintracciare gli articoli di opinione che riportano le analisi soggettive dei giornalisti, con un'accuratezza del 95% dopo sedici sessioni di addestramento¹⁶⁸. Ovviamente non sempre i contenuti opinionistici sono *fake news*. I contenuti di opinioni sono forse la più importante funzione giornalistica, alla luce del progresso descritto in questo lavoro. Risulta comunque importante poter discernere tra un contenuto di opinione e un fatto riportato in maniera imparziale e fattuale. In generale i modelli hanno raggiunto precisioni paragonabili alla capacità umana di individuare i contenuti falsi¹⁶⁹.

Strumenti di questo genere vengono utilizzati anche per scovare i *deepfakes*. Sono molto simili ai software video analizzati nel secondo capitolo e rintracciano i video falsi. In un'analisi sulle modalità con cui grandi redazioni e piattaforme online affrontano questo fenomeno si constata come l'identificazione dei *deepfakes* è una pratica comune per entrambi i tipi di organizzazione. Le differenze sono nel fatto che mentre i media sono concentrati nell'insegnare ai giornalisti la loro identificazione, le piattaforme online tendono a finanziare

¹⁶⁶Kai Shu, Amy Sliva, Suhang Wang, Jiliang Tang, and Huan Liu *Fake News Detection on Social Media: A Data Mining Perspective* Computer Science & Engineering, Arizona State University, Tempe, AZ, USA Charles River Analytics, Cambridge, MA, USA, 2017

¹⁶⁷Chiang T.H.C., Liao C.S., Wang W.C., *Investigating the Difference of Fake News Source Credibility Recognition between ANN and BERT Algorithms in Artificial Intelligence*, Appl. Sci., 12, 7725. 2022

¹⁶⁸Samhir Vasdev, *Can machine learning help us measure the trustworthiness of news?*, IREX, settembre 2018

¹⁶⁹Perez-Rosas, Bennett Kleinberg, Alexandra Lefevre, *Automatic Detection of Fake news*, Computer Science and Engineering, University of Michigan, University of Amsterdam, 2017

progetti di ricerca con l'obiettivo di sviluppare e migliorare gli strumenti di indagine¹⁷⁰. Responsabilizzazione delle piattaforme e formazione digitale dei giornalisti sono quindi le chiavi per combattere la diffusione di video, immagini e articoli falsi. Un'intelligenza artificiale è stata utilizzata per analizzare le conversazioni su Twitter inerenti al fenomeno dei *deepfakes*. I risultati mostrano come più di metà degli attori che fanno da ponte nelle interazioni sulla rete per questo argomento sono media e giornalisti¹⁷¹.

Tutto questo ci porta alla necessità di analizzare come il ruolo del giornalista sia cambiato non solo nel passaggio dal ruolo di *gatekeeper* a quello di *fact-checker*, ma anche nell'incombenza di imparare nuove skills digitali che non riguardano più solamente la capacità di curare un sito, caricare un articolo sul web o saper distribuire un contenuto in diverse forme e su più piattaforme. Gli iscritti all'ordine, in Italia, sono tenuti a seguire corsi di aggiornamento ogni anno. È stato inoltre messa in campo, dall'ordine stesso, un progetto di intelligenza artificiale per "certificare" l'autore delle notizie: il Protocollo Informazione Certificata. Questo Protocollo si traduce a livello pratico in un Qr Code che l'utente-lettore potrà scannerizzare e ricevere tutte le informazioni rilevanti inerenti all'autore dell'articolo. Un ex consigliere nazionale dell'ordine sostiene che ci sarà bisogno anche di una certificazione geografica e temporale dei contenuti, alla luce delle difficoltà recentemente riscontrate nell'analisi della veridicità di immagini di guerra nel conflitto tra Ucraina e Russia.¹⁷²

¹⁷⁰Angel Vizoso, Martin Vaz-Alvarez, Xosè Lopez-Garcia, *Fighting Deepfakes: Media and Internet Giants' Converging and Diverging Strategies Against Hi-Tech Misinformation*, Media and Communication, Volume 9, Issue 1 Pages 291-300, 3 marzo 2021

¹⁷¹Jesus Perez Dasilva, Koldobika Meso Ayerdi, Terese Mendiguren Galdospin, *Deepfakes on Twitter: Which Actors Control Their Spread?*, Media and Communication, Volume 9, Issue 1, Pages 301-312, 3 marzo 2021

¹⁷²Aldo Fontanarosa, *Giornalisti Robot. L'intelligenza Artificiale in redazione. Prove tecniche di news revolution*, Capitolo 18, pagina 273, 2020

4.2 L'impatto delle intelligenze artificiali sul settore giornalistico

Alla luce di quanto appena visto, il giornalista ha dovuto ridimensionare le proprie skills in funzione di un panorama mediatico che ha visto crescere esponenzialmente notizie false. Il giornalista è passato nell'era digitale dal decidere lui quali notizie dovessero essere al centro dei dibattiti, a subire tali decisioni e ritrovarsi quindi solo a "filtrare" le notizie più discusse, fornendo la propria opinione o etichettando una notizia come vera o falsa. Un tempo, come constata la teoria dell'agenda setting¹⁷³, ciò che sapevamo lo sapevamo tramite i mass media. Solo attraverso la stampa venivamo a conoscenza di eventi ed argomenti altrimenti estranei¹⁷⁴. I mass media offrivano riferimenti contestuali all'interno dei quali collocare e dar senso agli eventi stessi. Non che i media persuadessero i lettori, ma il "monopolio" dell'informazione da parte dei media tradizionali permetteva loro di determinare la presenza o meno di specifici temi e costituire quindi ciò che per l'individuo sarebbe diventato il tema su cui dibattere. Già al tempo alcuni ricercatori avevano notato come nel momento in cui il sistema mediatico avrebbe avuto bassa credibilità, questa influenza unidirezionale si sarebbe interrotta¹⁷⁵. Con l'avvento dei social media e degli *user generated content*, quel "monopolio" si è arrestato e sono entrati nel "mercato informativo" milioni di utenti pronti a condividere per primi un video di un evento e ricondividere contenuti scovati da altri. Tutto ciò porta i giornalisti ad essere degli attori all'interno di un sistema informativo ampio e variegato.

Lo snodo fondamentale per comprendere il ruolo che ancora possono avere i giornalisti è la comprensione del fatto che non tutti abbiamo gli strumenti necessari per capire la realtà e cercare di trasmetterla il più oggettivamente possibile agli altri. Abbiamo più volte ripetuto in questa sede quali sono state le degenerazioni del cambiamento digitale: dalle *eco chambers* alla faziosità dei dibattiti sul web, se da un lato tutto ciò ha reso possibile una maggiore accessibilità ai contenuti e una diffusione trasversale di essi, dall'altro ha portato gli individui a chiudersi nelle proprie bolle informative e ad aderire acriticamente a determinate opinioni. Il ruolo del giornalista, quindi, potrebbe e dovrebbe essere ancora quello di incentivare un dibattito polarizzato, comprendente al suo interno opinioni diverse e vari spunti di riflessione. Altro nodo cruciale è sicuramente la sostenibilità economica di un tipo di informazione certificata e professionale. Se la decisione degli utenti di informarsi tramite il web è spesso dettata dal portafoglio, bisogna necessariamente trovare un modo per rendere più accessibile questo tipo di informazione. Ultima ma non meno importante sfida è riguadagnare la fiducia dei lettori per questo tipo di informazione. Affronteremo ognuna di queste sfide nello specifico capendo lo stato dell'arte e discutendo di come le intelligenze artificiali possano aiutare a raggiungere gli obiettivi sperati.

¹⁷³L'agenda-setting è la teoria delle comunicazioni che ipotizza la possibile influenza dei mass-media (mass-news) sull'audience in base alla scelta delle notizie considerate "notiziabili" e allo spazio e preminenza loro concessa. Il postulato principale dell'Agenda-setting è il salience transfer, cioè il rendere la notizia saliente rispetto alle altre, quindi indica l'abilità dei mass media a trasferire un argomento da una agenda privata a quella pubblica d'interesse generale più elevato. (Wikipedia)

¹⁷⁴Walter Lippmann, *Public Opinion*, Macmillan, New York, 1922

¹⁷⁵Dearing J.W., Rogers E.M., *Communication Concepts: Agenda-setting*, Thousand Oaks, 1996

4.2.1 L'impatto occupazionale delle intelligenze artificiali nelle redazioni

“Solo negli Usa BuzzFeed ha licenziato il 15% della forza lavoro, 220 persone, e Verizon (proprietaria di Huffington Post, Yahoo News e molte altre testate) ha annunciato il taglio di circa 800 impieghi: un migliaio di posti andati in fumo”¹⁷⁶. “Di chi è la colpa di tutto ciò? In primis, dell’incapacità di trovare un nuovo modello di business adatto all’epoca digitale, ma anche del duopolio Google e Facebook, che si mangia il 75% della torta pubblicitaria. Non solo: un altro fattore, almeno secondo una bella analisi di Michele Mezza, è l’automazione. “A essere sostituiti con bot e intelligenze artificiali, in larga parte, sono proprio le funzioni desk, l’ultimo ridotto in cui si erano rifugiati i giornalisti dopo le decimazioni di inviati e corrispondenti”¹⁷⁷. “Una caratteristica non da poco quando ai giorni nostri bisogna fare il confronto con i nuovi arrivati. È di fine maggio 2020 la notizia del Seattle Times, secondo la quale MSN, il portale di news di Microsoft, avrebbe licenziato 50 dipendente addetti sostituiti dall’intelligenza artificiale”¹⁷⁸

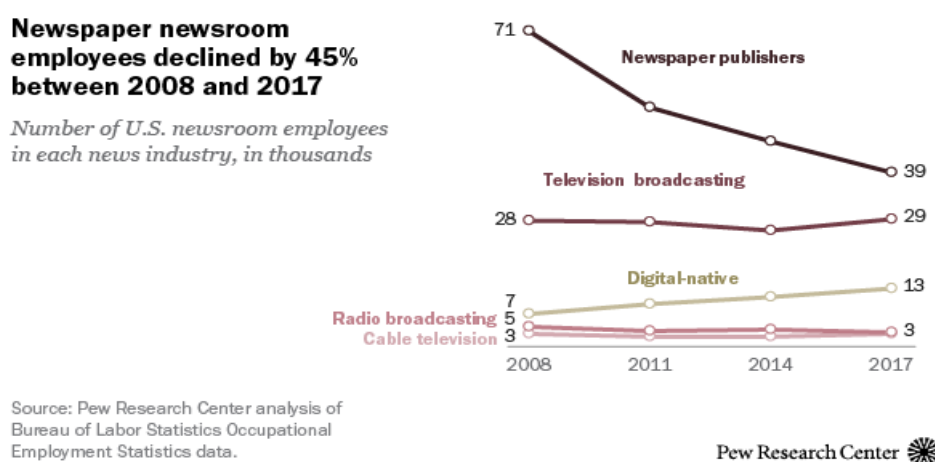


Fig. 16 Occupazione nelle redazioni dal 2008 al 2017 in America

La domanda che si legge sempre più spesso sul web: è “le intelligenze artificiali sostituiranno gli esseri umani”? Anche per i più ottimisti risulta difficile nascondersi dietro un dito. La risposta è sì, le intelligenze artificiali sostituiranno alcune mansioni svolte oggi dagli umani. Le sostituiranno tutte? È ancora presto per dirlo. L’analisi svolta in questa sede ci dice che non tutti gli ambiti possono essere coperti dalle intelligenze artificiali. Quantomeno non ancora. La *sentiment analysis* come detto è un ambito ancora in via di sviluppo e risulta difficile sostituire la capacità dell’uomo a livello emotivo-relazionale. Per di più, tutti gli aspetti etici e deontologici descritti sin qui non possono prescindere dall’intervento umano. Se la visione del settore giornalistico rimane quella di un mercato come un altro, orientato al profitto e al “consumo” di notizie, le intelligenze artificiali troveranno sicuramente sempre più spazio.

¹⁷⁶Stefano Tenedini, *Giornalismo e AI: un rapporto complicato ma inevitabile. E utile.*, impactscool magazine, robotica e ai, 20 febbraio 2019

¹⁷⁷Andrea Daniele Signorelli, *Questo articolo non è stato scritto da un robot*, Wired Italia, Scienza, 13 febbraio 2019

¹⁷⁸Simone Catania, *Giornalismo automatizzato: l’intelligenza artificiale genera notizie*, news.srl, 23 luglio 2020

Ma se gli obiettivi, come detto, sono altri e allora la capacità umana di portare avanti valori etici e democratici, di capire quale sia la cosa più giusta da fare in relazione ai lettori, di riuscire a conciliare i diversi interessi in gioco rimarrà sempre un elemento importante all'interno delle diverse redazioni. C'è chi opina questa tesi sostenendo che le macchine prima o poi arriveranno anche a sostituire tutto questo e che con lo scorrere del tempo ci avvicineremo sempre di più a quell'idea di IA forte presentata all'inizio di questa analisi. Vedremo a breve qual è lo stato dell'arte sulle potenzialità delle intelligenze artificiali e quanto si avvicinano a queste mansioni etichettabili come "umane". Prima però è doveroso, alla luce della situazione occupazionale appena descritta, capire quale sia oggi il ruolo del giornalista in redazione. Va detto anche che questa rivoluzione riguarda tutti gli ambiti lavorativi. Il profilo lavorativo che sta emergendo sempre con più trasversalità occupazionale proprio grazie a questa rivoluzione è sicuramente quello riguardante gli operatori menzionati nel terzo capitolo¹⁷⁹. Programmatori, informatici, tecnici in generale sono tutti mestieri che sempre di più sono richiesti nelle aziende e nei pubblici uffici. Secondo LinkedIn, le professioni in ascesa in Italia nell'ultimo anno sono¹⁸⁰:

1. Ingegnere robotico
2. Ingegnere del machine learning
3. Cloud architect
4. Data engineer
5. Sustainability manager

Tutte queste mansioni sono direttamente o indirettamente connesse con quanto discusso sinora. Il giornalismo necessiterà sempre di più figure di questo tipo? Se nei progetti di ricerca sulle intelligenze artificiali nel giornalismo è sempre più comune trovare tecnici e giornalisti fianco a fianco per indagare i processi di implementazione e le potenzialità delle nuove tecnologie, è difficile pensare a profili di questo tipo all'interno delle redazioni, soprattutto per la capacità retributiva di esse: persone specializzate in questi ambiti, a parità di mansioni e carico di lavoro, sceglieranno inevitabilmente posti di lavoro dove possono percepire un reddito più alto. All'inizio di questo potenziale cambiamento questo non sarà possibile. In realtà in cui le intelligenze artificiali nel giornalismo sono già una realtà affermata, come l'agenzia JX Press, lavorano anche ingegneri informatici¹⁸¹. La già descritta crisi del settore non permette dunque di integrare questi profili all'interno delle redazioni italiane. Le intelligenze artificiali hanno solo reso più ripida una discesa occupazionale che era già iniziata negli scorsi anni e che già si preventivava al nascere di queste tecnologie. Un report dell'I.N.P.G.I.¹⁸² evidenziava già nel 2018 come in 5 anni si fosse diminuita del 15% l'occupazione giornalistica, incolpando, in un'Italia dove le intelligenze artificiali non sono ancora una realtà propriamente affermata, i nuovi sistemi di informazione tecnologica.

¹⁷⁹Capitolo 3 paragrafo 3.1 sottoparagrafo 3.1.3

¹⁸⁰LinkedIn, *Lavori in crescita nel 2022: le 25 professioni in ascesa*, LinkedIn Notizie, 18 gennaio 2022

¹⁸¹Aldo Fontanarosa, *Giornalisti robot. L'intelligenza Artificiale in redazione. Prove Tecniche di news revolution*, UlisseAspettaPenelope, Capitolo 11 pg.149, 14 ottobre 2020

¹⁸²*In 5 anni perso il 15% di occupazione in Italia*, I.N.P.G.I. Notizie, 18 maggio 2018

Riconosce come si siano poi “consolidate nuove figure professionali legate al mondo dell’informazione e della comunicazione che sfuggono ai tradizionali sistemi di classificazione nell’ambito delle ordinarie categorie dell’attività giornalistica”.

Le figure professionali del mondo dell’informazione e della comunicazione, però, non sono la stessa cosa rispetto al profilo del giornalista. Il giornalista ha delle regole da rispettare, ha una missione democratica da portare avanti, non è orientato solo al profitto. Le redazioni poi sono da anni funzionali al perseguimento di questi obiettivi ed hanno gli strumenti e il *modus operandi* per farlo. Il problema è quindi quello di riconquistare il ruolo di *gatekeepers* e ridare lustro all’importanza del giornalismo professionale, raggiungere la sostenibilità economica per le redazioni e riguadagnare la fiducia dei lettori. Le intelligenze artificiali possono essere funzionali al raggiungimento di tutti e tre gli obiettivi. Anzi, si può affermare che senza di esse se ne lascerebbe sicuramente qualcuno per strada o non si compirebbe a pieno questa missione. Se si facessero campagne di sensibilizzazione sull’importanza del giornalismo professionale per la democrazia, sia come *gatekeepers* che come *fact-checkers*, si potrebbe sì riguadagnare la fiducia dei lettori, ma si finirebbe inevitabilmente sul lastrico. Se si riguadagnasse la fiducia dei lettori con prodotti accattivanti e altamente *tailored*, proposti su misura per l’utente-lettore, si riuscirebbe sì a ritornare in pista dal punto di vista economico ma si rischierebbe di perdere quel rispetto dei valori democratici e quell’importanza etica che ha il giornalismo all’interno della società. Se si cercasse di ritrovare il ruolo di mediatori che si aveva una volta riuscendo a trovare forme di sostentamento, ciò non potrebbe prescindere dal ritrovare la fiducia dei lettori.

Partendo quindi dal presupposto che fiducia dei lettori e riconquista dell’importanza del ruolo di giornalista professionale siano due facce della stessa medaglia, riuscire a perseguire queste tre strade contemporaneamente sarà possibile solo grazie alle intelligenze artificiali. La domanda che sorge spontanea ora è: come si fa a ridare lustro ad una professione in crisi se l’occupazione è in drastico calo? La risposta è semplice. Adattando quella professione alle esigenze del momento. Il giornalista non dovrà più essere solo una persona socializzata ai valori giornalistici, capace di coltivare una fonte, in grado di comprendere al meglio la realtà e riportarla ai lettori. Il giornalista dovrà essere una persona capace di saper introdurre degli input all’interno di un programma, in grado di saper cogliere i trend del momento. Dovrà saper collaborare fianco a fianco ad una intelligenza artificiale facendone non un nemico, bensì uno strumento funzionale al raggiungimento degli obiettivi. L’intelligenza artificiale permetterà di risparmiare in termini di costi e quindi rendere più sostenibile e autosufficiente una redazione. Permetterà al giornalista di adattarsi al cambiamento digitale del sistema informativo e permetterà di riguadagnare la fiducia dei lettori ritrovando prossimità con i loro interessi, riuscendo allo stesso tempo a consigliarli o indirizzarli quantomeno nella determinazione degli argomenti di cui discutere. Il giornalismo professionale può quindi avere ancora il suo ruolo importante, e non solo di *fact-checker*.

4.2.1.1 La necessità di un nuovo ruolo per il giornalista

Il giornalista, per riuscire ad adattarsi a tutto ciò, deve essere predisposto a questi cambiamenti. Deve socializzarsi digitalmente ai nuovi strumenti e deve essere disposto a mettere in discussione la concezione tradizionale della professione. Il processo di apprendimento non è differente da quello delle altre tecnologie¹⁸³. Nei progetti avviati in Inghilterra dalla BBC e The Times, si riscontra una propensione da parte dei giornalisti ad imparare ad utilizzare queste tecnologie, consapevoli che possano apportare benefici sia in termini di tempo che di qualità del lavoro, ma i tecnologi che hanno partecipato a questi progetti riscontrano difficoltà nell'integrare queste tecnologie all'interno del lavoro quotidiano delle redazioni¹⁸⁴. In altre parti del mondo c'è chi già parla di quarta rivoluzione industriale e sottolinea come il progresso tecnologico abbia aggiornato e digitalizzato le nostre hard-skills attraverso le intelligenze artificiali che consentono all'industria dei media di cercare, creare, distribuire e misurare la comunicazione in modo efficace. Per quanto riguarda le soft skills, bisognerebbe migliorare la capacità di scrittura creativa, non più di esclusivo dominio dell'uomo ma in ogni caso compatibile con il nuovo pubblico, che dispone di molte più opzioni¹⁸⁵. Per fare tutto questo c'è bisogno di un aggiornamento costante. A priori però, in paesi come l'Italia in cui questa rivoluzione fa fatica ad iniziare, bisognerebbe partire dai piani di studi. L'IA è assente in quasi tutti i piani di studi delle università. In Spagna, invece, paese simile al nostro dal punto di vista delle difficoltà nell'affermazione delle IA nell'ambito giornalistico, si sono avviati progetti che approcciano il tema delle IA nel giornalismo e ce la introducono negli insegnamenti. I docenti spagnoli affrontano l'IA con due prospettive: una critica che riguarda le conseguenze sociali e una propositiva con l'obiettivo di formare gli studenti nel trattamento dei dati, nella creazione di contenuti automatizzati e nella verifica dei contenuti¹⁸⁶.

Alcuni sostengono che il giornalista sopravviverà solo in quanto inviato. I più pessimisti direbbero che anche quel ruolo sarà prima o poi intrapreso dalle IA, alla luce del recente sviluppo del drone journalism e del *sensor journalism*¹⁸⁷. I più ottimisti però, con la realtà dei fatti che corrobora la loro tesi, direbbero che l'intervistatore non potrà mai essere sostituito. E se i *deepfakes* di Xinhua iniziassero a condurre anche interviste? È possibile, ma l'unione dei due ruoli appena descritti non sarà mai possibile a meno del raggiungimento, ancora molto lontano, di un robot umanoide identico a noi. L'inviato-intervistatore è un ruolo che spetta, forse ancora per molto tempo, all'umano.

¹⁸³Joo-Wha Hong, *I Was Born to Love AI: The Influence of Social Status on AI Self-Efficacy and Intentions to Use AI*, International Journal of Communication 16(2022), 172-191, 2022

¹⁸⁴Marisela Gutierrez Lopez, Colin Porlezza, Glenda Cooper, Stephann Makri, Andrew MacFarlane, Sondess Missaoui, *A question of Design: Strategies for Embedding AI-Driven Tools into Journalistic Work Routines*, Digital Journalism, 2022

¹⁸⁵Mohamed Abdulzaher, *Artificial Intelligence Journalism and Media Digital Soft-Skills*, gennaio 2020

¹⁸⁶Gloria Gomez-Diago, *Perspectives to address artificial intelligence in journalism teaching. A review of research and teaching experiences*, Revista Latina de Comunicacion Social, 80, 29-46, 2022

¹⁸⁷Capitolo 2 paragrafo 2.1 sottoparagrafo 2.1.1

4.2.2 La sostenibilità economica delle redazioni

Negli anni del covid sono più che raddoppiati i finanziamenti pubblici ai giornali. Si è passati da 175,6 mln a 386,6 mln. Nel 2021 poi sono stati stanziati 232,9 mln sotto forma di crediti di imposta¹⁸⁸. Le redazioni riescono inoltre ad avere ulteriori introiti principalmente tramite due modelli. Il primo è quello degli abbonamenti. Nonostante la sfiducia degli italiani nei confronti della stampa ci sono ancora alcuni giornali che hanno un pubblico che potremmo definire fidelizzato. L'altro modello è quello di lasciare spazi pubblicitari cartacei o digitali a soggetti terzi che pagano quegli spazi pubblicitari. Questi due modelli possono anche essere complementari e in alcuni casi ci sono grandi gruppi editoriali che riescono a coprire ulteriori spese. Ma allora qual è il problema delle redazioni? Perché abbiamo più volte parlato di crisi del settore? Le spese che sono state compiute in questi ultimi periodi hanno dovuto per lo più risanare situazioni già molto gravi. La capacità di spesa dei giornali quindi, nonostante i numeri appena elencati, non è stata comunque all'altezza per poter investire sull'innovazione digitale. Per di più, la vera crisi riguarda i lettori: "dal 2013 al 2019 la gran parte dei quotidiani italiani aveva perso fra il 30 e il 50 per cento delle copie diffuse (la "diffusione" comprende sia le copie vendute in edicola o in abbonamento che quelle omaggio). Con una crisi di tale portata in atto da prima dell'inizio della pandemia, le perdite dei giornali nel 2020 sono state consistenti: l'andamento negativo si è riconfermato nel 2021 e – salvo qualche caso – non c'è stato un "rimbalzo" dopo la fine dei lockdown né le copie digitali hanno compensato la minor diffusione cartacea"¹⁸⁹. Questo vuol dire quindi meno abbonamenti e anche meno aziende che vogliono investire sugli spazi pubblicitari dei giornali, poiché la loro propensione a ciò è direttamente legata alle statistiche della diffusione dei giornali. Il sistema non può quindi sostenersi in eterno tramite sussidi statali o sovvenzionamenti esterni. La sopravvivenza di questo settore passa quindi inevitabilmente dalla capacità di saper riguadagnare la fiducia dei lettori. I dati ci suggeriscono che adattando le news al cambiamento digitale si percorre quantomeno una strada meno ripida¹⁹⁰. Video, immagini, news interattive sono la strada giusta da percorrere per conformarsi ad un cambiamento che non bisogna osteggiare, ma di cui bisogna prendere atto. Il numero di persone che ascolta notizie tramite podcast, come di quelli che preferiscono lo streaming video piuttosto che la lettura della notizia, sono sempre di più, all'interno però sempre di quella cerchia ristretta di persone intenzionate ad informarsi.

La sfiducia nelle élite professionali e nelle istituzioni è un problema generale. Non è in questa sede che affronteremo tale problematica e come riacquistare la fiducia dell'opinione pubblica nella politica, nella scienza o nella medicina. In parte questo processo, come abbiamo visto, è stato incentivato dalla necessità durante la pandemia quantomeno di discutere di argomenti inerenti questi ambiti. La strada però rimane sempre in salita. Quello che possiamo fare è suggerire delle soluzioni per il mondo dell'editoria.

¹⁸⁸ *Approvazione dell'elenco delle imprese editrici di quotidiani e periodici cui è riconosciuto, per l'anno 2021, il credito d'imposta per i servizi digitali, a norma dell'articolo 190 del decreto-legge 19 maggio 2020, n. 34, convertito, con modificazioni, dalla legge 17 luglio 2020, n. 77, prorogato, per gli anni 2021 e 2022, dall'articolo 1, comma 610, della legge 30 dicembre 2020, n. 178*

¹⁸⁹ Alice Lanciano, *La pandemia ha aggravato la crisi dei giornali*, *IlPost*, 17 luglio 2022

¹⁹⁰ Censis, *I media dopo la Pandemia*, Sintesi per la stampa del Diciassettesimo rapporto sulla comunicazione, 6 ottobre 2021

Il fascino delle tecnologie e la propensione dei giovani ad utilizzarle è un fatto innegabile¹⁹¹. La rinascita dei giornali potrebbe quindi passare da contenuti interattivi, innovativi, presentati in diverse forme e su diversi mezzi. Tutto questo può essere perseguito tramite l'utilizzo delle intelligenze artificiali. Ovviamente se fosse tutto così semplice i giornali avrebbero già adottato questa soluzione. I principali problemi riguardano l'accessibilità economica alle strumentazioni ma anche le modalità con cui integrarle nelle redazioni. Se i sovvenzionamenti statali più che tenere a galla una nave che sta già affondando fornissero ad essa gli strumenti per tenersi a galla da sola, si potrebbero fare dei passi avanti. Gli esempi dall'altra parte dell'Atlantico sembrano di buon auspicio per un approccio simile, ovviamente con tutte le cautele e le differenze del caso. Non dimentichiamoci che, come abbiamo visto, il sistema legislativo e quello informativo in generale sono molto diversi di paese in paese e ogni nazione ha la sua diversa realtà mediatica. Ciò non vuol dire però che non si possa trarre spunto da esempi virtuosi, come abbiamo fatto nel secondo capitolo.

In una survey mondiale¹⁹² i consigli strategici per implementare un'intelligenza artificiale in redazione sono:

1. Valutare il vostro stadio e il vostro stato di preparazione all'IA
2. Comprendere e classificare il tipo di tecnologie di IA che si stanno considerando
3. Decidere come l'IA potrebbe essere collegata al vostro marchio e alla vostra strategia generale, ai problemi che potrebbe risolvere o ai bisogni che potrebbe soddisfare
4. Valutare quali aree della vostra organizzazione potrebbero utilizzare l'IA e perché
5. Identificare gli ostacoli principali: risorse, competenze, cultura, gestione, ecc. e come affrontarli in modo sistematico
6. Assegnare ruoli e responsabilità e creare una struttura di comunicazione all'interno dell'organizzazione per includere tutti gli stakeholder
7. Stabilire sistemi di monitoraggio e revisione delle prestazioni e delle priorità
8. Creare un ruolo per le relazioni esterne con i partner, i clienti e le risorse più ampie dell'AI, con la missione di indagare e analizzare i problemi e i problemi di salute

Francesco Marconi propone all'interno del suo libro un modello denominato "iterative journalism" per rendere il processo di creazione di una news il più responsivo¹⁹³ possibile alle esigenze dei lettori. L'approccio richiede una sistematica integrazione delle nuove tecnologie nei processi tradizionali, piuttosto che implementare una sola tecnologia in uno step specifico. L'idea è anche quella di adattare la copertura delle notizie ai rapidi cambiamenti nelle richieste degli utenti. Ciò è possibile solo integrando le statistiche editoriali con i *feedback* dei lettori.

¹⁹¹Eurobarometer of European Commission, *European citizens' knowledge and attitudes towards science and technology*, 2021

¹⁹²Charlie Beckett, *New Power New Responsibilities, A global survey of journalism and artificial intelligence*, London School of Economics and Political Science, Google News Initiative, novembre 2019

¹⁹³Capacità di sapersi adattare dinamicamente al cambio di una situazione esterna. Viene utilizzato per lo più in informatica per indicare un sito web che si adatti al tipo di dispositivo che si utilizza ridimensionato il tutto sulla base delle dimensioni di esso. In questa sede è utilizzato per indicare la capacità di un modello di sapersi adattare ai cambiamenti e agli inconvenienti

Il processo di implementazione di questo modello inizia con il definire le opportunità di sperimentazione e identificare le risorse e le competenze tecniche necessarie. Si faranno sperimentazioni valutando anche in scala gli investimenti necessari.

Il processo pratico di sperimentazione della notizia può essere invece così riassunto:

- 1) Definire l'area nel quale si sta operando (*gathering, production o distribution*)
- 2) Identificare il tipo di tecnologia da utilizzare e quali sono gli strumenti e le competenze necessarie
- 3) Lanciare una sorta di “notizia pilota” e tracciare l'engagement
- 4) Analizzare la risposta del pubblico di riferimento e analizzare le statistiche
- 5) Decidere dove investire risorse sulla base di questo processo

Questo permette alle redazioni di valutare la sostenibilità economica dell'implementazione delle intelligenze artificiali e di investire in maniera proporzionata e adeguata.

Una delle innovazioni proposte per modelli sostenibili per le redazioni negli ultimi decenni è il concetto di *membership*. Un articolo sul Nieman Lab ponendosi la domanda della sostenibilità delle imprese giornalistica propose questo relativamente nuova idea. Relativamente perché in realtà si tratta di un'idea di 40 anni fa che andrebbe solo riadattata al contesto odierno. Nel 2013 The Guardian ha deciso di intraprendere questa strada creando un ruolo apposito per l'implementazione di questa strategia, per far sì che i guadagni non dipendessero più da paywall o sottoscrizioni chieste ai lettori. Il concetto trascende il rapporto tra giornali e lettori andando a superare quell'egocentrismo che lo vede al centro della comunità. Questo modello mette tutti allo stesso livello, cercando di rendere i lettori-utenti parte attiva del processo giornalistico, fornendo strumenti ad essi per condividere al meglio la loro conoscenza¹⁹⁴. Il giornale potrebbe così essere percepito come un valore aggiunto per il lettore e far sì che esso si senta ancora più fidelizzato. Tutto ciò potrebbe aiutarci a fronteggiare la carenza di fiducia da parte dei lettori. L'approccio potrebbe essere quindi quello di utilizzare strumenti innovativi ma di farli usare anche al lettore nei forum o nei blog di discussione dei giornali.

Approfondiamo ora la questione della fiducia, nello specifico la percezione che le persone hanno dei contenuti prodotti dalle intelligenze artificiali. Abbiamo detto sì che è innegabile la “propensione tecnologica” ma ciò non vuol dire che possano piacere in automatico notizie presentate come “più tecnologiche”. Vedremo ora qual è la percezione degli utenti-lettori sui contenuti delle intelligenze artificiali.

¹⁹⁴Lelio Simi, *Membership e giornali* (due idee per discuterne), DataMediaHub, 22 maggio 2014

4.3 La percezione delle IA nel giornalismo

ChatGPT è stato un successo enorme. Abbiamo evidenziato come addirittura non si potesse usare la sua versione AP nelle prime settimane del 2023 dal momento che troppi utenti la utilizzano e i server non sono riusciti a reggere tale mole di richieste¹⁹⁵. Questo entusiasmo per le tecnologie però non ci deve far trarre conclusioni affrettate. Un conto è la percezione dell'utilizzo di alcune tecnologie, un conto è la percezione di notizie scritte da una tecnologia. Le intelligenze artificiali in generale sono considerate positive da più di metà degli italiani¹⁹⁶. Per quanto concerne invece la propensione all'uso? L'eurobarometro ci suggerisce che il 58% degli italiani non si sentirebbe a disagio nel farsi assistere nello svolgimento del proprio lavoro da una intelligenza artificiale.

In generale, la propensione all'uso delle IA dipende in parte dall'istruzione e dal reddito, mentre l'età non sembra avere nessun tipo di incidenza¹⁹⁷. Fattore ancora più importante è sicuramente il contesto socioculturale¹⁹⁸ e l'accezione del dibattito sulle intelligenze artificiali in un paese. Se sui media e sui social network si parla delle IA in senso negativo, molto probabilmente esse saranno percepite come tali. In molti articoli italiani sul web consultati per condurre questa analisi, a differenza della letteratura scientifica, il tono con cui si parla delle intelligenze artificiali è spesso pessimistico. Molti media si concentrano più sulle difficoltà che sulle opportunità che le intelligenze artificiali possono portare al giornalismo. Si sottolinea come essendo un settore già in crisi ed essendo stati fatti dei tagli al personale nelle redazioni quando sono state adottate intelligenze artificiali, esse potranno essere più un male che un bene. Si tende poi a scrivere titoli funzionali alle visualizzazioni con termini eccessivamente spettacolari o tragici, senza analizzare nello specifico i pro e i contro del loro utilizzo. Molti cercano di fare leva sull'immaginario cinematografico fantascientifico di robot umanoidi e intelligenze artificiali che raggiungono le capacità umane. Hanno ragione?

Recentemente Google ha "intervistato" Chat-GPT per un colloquio di lavoro come ingegnere. L'intelligenza artificiale più discussa del momento ha raggiunto, sulla base delle risposte, il punteggio necessario all'assunzione per una posizione di ingegneria di livello tre. La capacità della macchina di fornire una risposta concisa e altamente affidabile potrebbe essere utilizzata da Google per migliorare i propri motori di ricerca¹⁹⁹.

Il tutto ovviamente pone problemi come l'utilità che ingegneri dello stesso livello possono avere alla luce di quanto detto. Tutto ciò è valido anche per il mondo del giornalismo?

¹⁹⁵Capitolo 2 Paragrafo 2.1 sottoparagrafo 2.2.1 sotto-sottoparagrafo 2.2.2.1

¹⁹⁶Eurobarometer of European Commission, *European citizens' knowledge and attitudes towards science and technology*, 2021

¹⁹⁷Joo-Wha Hong, *I Was Born to Love AI: The Influence of Social Status on AI Self-Efficacy and Intentions to Use AI*, *International Journal of Communication* 16(2022), 172-191, 2022

¹⁹⁸María T. Soto-Sanfiel, Adriana Ibiti, Mabel Machado, Beatriz Elena Marín Ochoa, María Mendoza Michilot, Claudio Guillermo Rosell Arce & Ariadna Angulo Brunet, *In Search of the Global South: Assessing Attitudes of Latin American Journalists to Artificial Intelligence in Journalism*, *Journalism Studies*, 23:10, 1197-1224, 2022

¹⁹⁹Jennifer Elias, *Google is asking employees to test potential ChatGPT competitors, including chatbot called "Apprentice Bard"*, CNBC, gennaio 2023

Per rispondere a questa domanda dovremmo capire le potenzialità dei software nei compiti che svolge normalmente un giornalista. Il parametro più utilizzato sono i testi scritti, nello specifico la loro credibilità, la loro qualità e la loro leggibilità. In una ricerca basata su 12 studi sperimentali con un totale di più di 4000 partecipanti, si nota come i testi siano percepiti allo stesso modo per quanto riguarda la credibilità. Sulla percezione della qualità gli umani hanno un leggerissimo vantaggio mentre per quanto riguarda la leggibilità il divario risulta più evidente²⁰⁰. La comparazione dei 12 studi ci suggerisce inoltre che la percezione dei parametri cambia quando ai partecipanti viene detto che la notizia è stata scritta da un umano, a vantaggio di quest'ultimo. Già nei primi anni di sviluppo dei software di scrittura automatica, nello specifico con una ricerca del 2014, si constatava come la percezione della differenza tra un contenuto umano e uno scritto da una intelligenza artificiale non fosse poi così diverso²⁰¹.

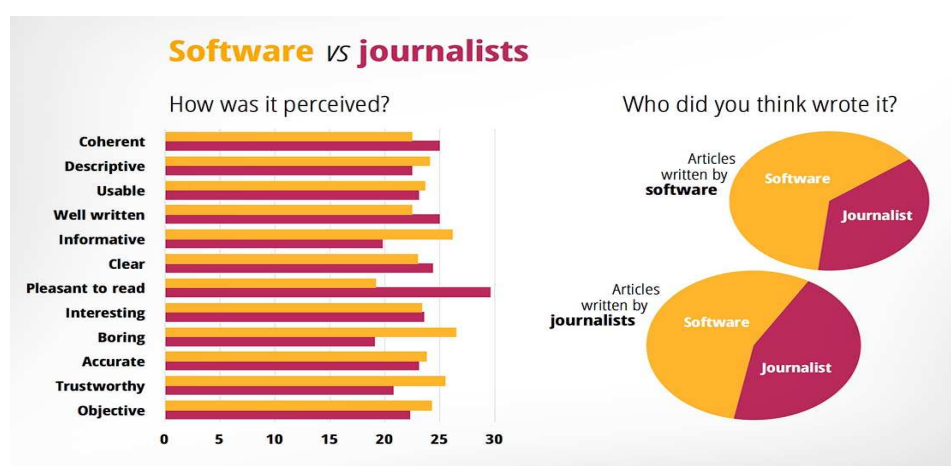


Fig.17 La percezione dei testi scritti da un software e da un umano

(Credit: Thomas Baekdal, *Robots vs Journalists? It's happening*, baekdal.com, marzo 2014)

La differenza più ampia si poteva rilevare su “*pleasant to read*”, cioè piacevole alla lettura. Le IA sono quindi allo stesso livello di scrittura degli umani? La risposta è indicativamente sì. Consultando la letteratura scientifica attuale²⁰² i risultati sono più o meno gli stessi. Nuove ricerche indicano nuove prove a favore dell'avversione per l'IA, ricordando però che la ricerca passata ha mostrato casi sia di apprezzamento dell'IA sia di avversione²⁰³. Ulteriori ricerche constatano questa avversione. Si specifica anche che nel momento in cui è presentato un contenuto scritto da un software, nel quale processo però è stato presente l'uomo secondo il tanto decantato approccio *human in the loop*, l'avversione scompare quasi del tutto. Inoltre, questa avversione riguarda più i contenuti definibili come creativi²⁰⁴, constatando quanto detto in questa sede sulla raggiungibilità di alcune capacità umane da parte delle macchine²⁰⁵.

²⁰⁰Andreas Graefe and Nina Bohlken, *Automated journalism: A Meta-Analysis of Readers' Perceptions of Human-Written in Comparison to Automated News*, Media and Communication, Volume 8, Issue 3, pages 50-59, ISSN: 2183-2439, 2020

²⁰¹Thomas Baekdal, *Robots vs Journalists? It's happening*, baekdal.com, marzo 2014

²⁰²Nils Kobis, Luca D. Mossink, *Artificial intelligence versus Maya Angelou: Experimental evidence that people cannot differentiate AI-generated from human-written poetry*, Computer in Human Behaviour, Elsevier, settembre 2020

²⁰³Chiara Langoni, Andrey Fradkin, Gordon Pennycook, *News from Generative Artificial Intelligence Is Believed Less*, ACM Conference on Fairness, Accountability and Transparency, giugno 2022

²⁰⁴Si tratta di analisi di poesie

²⁰⁵Nils Kobis, Luca D. Mossink, *Artificial intelligence versus Maya Angelou: Experimental evidence that people cannot differentiate AI-generated from human-written poetry*, Computer in Human Behaviour, Elsevier, settembre 2020

4.3.1 La percezione della personalizzazione delle notizie

Quali sono gli atteggiamenti, le percezioni e i comportamenti del lettore di oggi in merito a servizi di notizie personalizzati tramite IA? Gli effetti riscontrati sono quasi tutti positivi. Gli utenti di portali personalizzati sono più attenti alle fonti e sono più propensi al consumo di notizie offline²⁰⁶. I contenuti personalizzati hanno un effetto positivo sugli utenti in merito alla rilevanza che danno alle notizie, al coinvolgimento e alla propensione all'interazione²⁰⁷. Gli utenti che usufruiscono di contenuti personalizzati hanno una soddisfazione maggiore²⁰⁸. I contenuti personalizzati tramite intelligenza artificiale giovano anche alle redazioni in termini di fidelizzazione e abbonamenti²⁰⁹.

I mass media pubblici sono inoltre tenuti a progettare algoritmi che rispettino i valori fondamentali del giornalismo, garantendo così la promozione della diversità dei contenuti, anche se gli utenti decidono di personalizzare la selezione dei contenuti. La personalizzazione dei contenuti rappresenta una minaccia per la diversità e la pluralità dell'informazione ma, come abbiamo visto, ci sono alternative alla personalizzazione univoca²¹⁰.

Il New York Times ha vinto 125 premi Pulitzer di cui 3 solo nel 2018. I giornali americani possono quindi essere un modello per gli editori online di tutto il mondo, soprattutto per quanto riguarda le loro strategie economiche. Il rapporto 2020 del New York Times, ad esempio, che fa riferimento ai cambiamenti delle strategie e ai nuovi obiettivi dell'azienda, è fonte di ispirazione per ogni editore, come la *user experience* dei siti web del The New York Times, The Washington Post e The Wall Street Journal e la conversione dei lettori in abbonati. Vedremo come questi giornali riescono a far capire ai lettori l'importanza di pagare per accedere a contenuti di qualità, come convincono il loro pubblico utilizzando il paywall e come hanno sviluppato le loro strategie di paywall per garantire all'utente un'esperienza ottimale in ogni contesto. I pilastri delle loro strategie sono²¹¹:

- la raccolta di dati come alternativa all'abbonamento (questa strategia comporta probabilmente un form di registrazione rispetto al paywall).
- l'adozione di strategie diverse per il sito web e per l'applicazione mobile, poiché questi media sono utilizzati dai lettori in modi diversi.
- proposta di offerte adatte al profilo dell'utente

²⁰⁶M.A. Beam, G.M. Kosicki, Personalized news portals: filtering systems and increased news exposure, *Journal. Mass Commun.* Q. 91, 59–77, 2019

²⁰⁷S. Kalyanaraman, S.S. Sundar, The psychological appeal of personalized content in web portals: does customization affect attitudes and behavior? *J. Commun.* 56, 110–132, 2019

²⁰⁸Q. Ye, Y. Luo, G. Chen, X. Guo, Q. Wei, S. Tan, Users intention for continuous usage of mobile news apps: the roles of quality, switching costs, and personalization, *J. Syst. Sci. Syst. Eng.* 28, 91–109, 2019

²⁰⁹B. Bodo, ' Selling news to audiences – a qualitative inquiry into the emerging logics of algorithmic news personalization in European quality news media, *Digit. Journal* 7, 1054–1075, 2019

²¹⁰Capitolo 2 paragrafo 2.3 sottoparagrafo 2.3.3 sotto-sottoparagrafo 2.3.3.1

²¹¹*White paper: The New York Times, The Washington Post and The Wall Street Journal Paywall Analysis*, Pool, 30 maggio 2018

In che modo le organizzazioni giornalistiche progettano e implementano servizi giornalistici personalizzati da algoritmi? Vediamo qual è il ruolo della personalizzazione nel contribuire al successo finanziario delle redazioni, nel raggiungere e fidelizzare il pubblico e nel rispettare i valori deontologici. Un paper del ANNOX, con un approccio di tipo economico, analizza i modelli delle redazioni europee. Questo approccio, dal quale in parte ci siamo discostati in questa sede, vede le redazioni come attori economici di un mercato informativo. L'analisi rileva che invece che concentrarsi sull'aumento del coinvolgimento degli utenti a breve termine, i media europei di qualità cercano di utilizzare la personalizzazione delle notizie per aumentare la fedeltà del pubblico a lungo termine. Guidati da valori diversi, alla fine del 2017 sembrano essere emersi almeno due filoni di personalizzazione nettamente differenti.

- La "logica di personalizzazione delle piattaforme" è caratterizzata da un'abbondanza di dati degli utenti, un'immensa base di utenti e contenuti, un modello di business aggressivo e di successo basato sugli annunci pubblicitari, risorse finanziarie e tecnologiche quasi illimitate e una forte resistenza a qualsiasi controllo editoriale o supervisione delle raccomandazioni algoritmiche.
- La "logica di personalizzazione delle notizie" è caratterizzata da un insieme limitato di dati sugli utenti (limitato da risorse finanziarie insufficienti), una base di utenti e di contenuti limitata, un modello di business basato sugli annunci in difficoltà, con l'emergere di notizie a pagamento come alternativa, risorse finanziarie e tecnologiche limitate, un forte controllo editoriale e un rispetto dei valori editoriali.

L'autore invita a distinguere bene tra i due tipi di personalizzazione, poiché differiscono sostanzialmente nelle minacce e nelle opportunità che creano. Giornalisti e redattori hanno iniziato a costruire le loro competenze e a sviluppare modi per utilizzare e controllare le tecnologie di personalizzazione. Lo sviluppo di servizi di informazione personalizzati può essere letto anche come un processo attraverso il quale i professionisti dell'informazione rivendicano alcuni strumenti algoritmici per contrastare il dominio delle tecnologie di personalizzazione messe in campo dalle piattaforme e dalle aziende di annunci, e per mitigare il loro impatto negativo sul business dell'informazione. Ciò che è cambiato è che le organizzazioni giornalistiche si sono rese conto che per loro questa manipolazione della domanda riguarda la vendita di informazioni di qualità, la coltivazione di un gusto per le *hard news*²¹², la giustificazione degli investimenti pubblici o il mantenimento dell'autorità e dell'affidabilità giornalistica nell'era delle *fake news*, piuttosto che la competizione con le piattaforme per il coinvolgimento dell'audience e la vendita di pubblicità.

Per raggiungere questi obiettivi, le organizzazioni giornalistiche, i loro editori, i giornalisti e i professionisti della distribuzione stanno lottando per riguadagnare la fiducia del pubblico.

²¹²Hindman, Matthew, *Journalism Ethics and Digital Audience Data*. In *Remaking the News: Essays on the Future of Journalism Scholarship in the Digital Age*, edited by Pablo J. Boczkowski and C. W. Anderson, 177–195. Cambridge, MA: MIT Press, 2017

Conclusione

Si è tentato di capire come sfruttare al meglio le intelligenze artificiali per il sostentamento economico delle redazioni e come offrire all'utente contenuti innovativi e accattivanti, il tutto rispettando le regole giuridiche ed etiche del settore giornalistico.

Per fare ciò il primo passo, corrispondente al primo capitolo, è stato quello di concettualizzare al meglio il termine intelligenza artificiale e di ripercorrere la strada che ci ha portato sino alle moderne intelligenze artificiali, sia in generale che nel mondo del giornalismo. Abbiamo ritenuto opportuno proporre categorizzazioni e terminologie al fine di evitare fraintendimenti ed equivoci nell'uso di determinati termini e puntualizzare al meglio sulle tipologie di software. Le intelligenze artificiali sono un campo vastissimo e senza una lente di ingrandimento sui concetti basilari non saremmo riusciti a comprendere le tecnologie più complesse, pur analizzandole principalmente nelle funzioni che svolgono. Dopo questa rassegna generale, ma necessaria, ci siamo addentrati nel mondo del giornalismo provando a partire dal primo esempio comunemente indicato come utilizzo di intelligenza artificiale. Se ne è trovato un altro risalente a quarant'anni prima, raramente menzionato in letteratura scientifica, utilizzato per le previsioni meteorologiche. Analizzando poi nel secondo capitolo le intelligenze artificiali utilizzate nel mondo del giornalismo, si sono compresi i rischi e le potenzialità di esse. Gli esempi virtuosi presentati sono stati funzionali a constatare che software utili al perseguimento dei nostri obiettivi iniziali esistono e se utilizzati nel rispetto di quanto presentato nel terzo capitolo, cioè della legislazione e dei principi etici, si può perseguire il profitto economico senza ledere i diritti di nessuno, né di terze persone né del pianeta. La parte finale del lavoro ha fatto luce su come inevitabilmente debba cambiare la professione giornalistica per andare in questa direzione. Non è sufficiente solo una propensione positiva all'uso delle tecnologie da parte dei giornalisti, ma è necessario un inserimento di queste materie nelle scuole di giornalismo e un aggiornamento costante delle skills utilizzate in redazione, che saranno sempre più tecniche e digitali. Le intelligenze artificiali potranno sì sostituire il giornalista nello svolgimento di alcune mansioni, ma non arriveranno mai a sostituirlo del tutto, quantomeno non ancora. La sua vena creativa e le sue intuizioni professionali non sono ancora appannaggio delle intelligenze artificiali. Per di più, un testo scritto da un giornalista, da solo o in collaborazione con queste tecnologie, è percepito più positivamente di contenuti interamente prodotti dalle intelligenze artificiali. Si riscontra la carenza di studi sugli altri tipi di tecnologie, con un'eccessiva focalizzazione della ricerca scientifica sulla scrittura automatica di contenuti. Le prospettive future sono quindi di cavalcare l'onda del cambiamento, prendendone atto e sfruttando al meglio le sue potenzialità, transcendendo un dibattito pubblico che troppo spesso si concentra sui rischi oscurandone le opportunità. Un'implementazione graduale delle intelligenze artificiali delle redazioni, consapevoli delle possibilità economiche di esse, può essere la strada giusta in questa direzione. Grazie ad esse si potranno offrire agli utenti contenuti innovativi e di qualità incentrati sulle loro esigenze, proponendo al contempo alternative ai loro interessi e invitandoli ad uscire da una bolla informativa sempre più ristretta che rischia di soffocare il diritto all'informazione.

Bibliografia

- W. Zhao, T. He and L. Xu, *Enhancing Local Dependencies for Transformer-Based Text-to-Speech via Hybrid Lightweight Convolution*, *IEEE Access*, vol. 9, pp. 42762-42770, 2021
- Harry R. Glahn, *Computer-produced Worded Forecast*, Techniques Development Laboratory, Systems Development Office, ESSA Weather Bureau, Silver Spring, Md, WEATHER SCIENCE, INC. 1225 WEST MAIN STREET NORMAN, OKLAHOMA 73069 1132 Vol 51, No. 12, December 1970
- Daniel Fagella, *What is Machine Learning*, Emerj The Ai Research and Advisory Company, 26 febbraio 2020
- Oliver Jeffery, *Explain Artificial Intelligence an history of Artificial intelligence*, University of Alberta, novembre 2022
- Thomas Hobbes, *Leviathan or The Matter, Forme and Power of a Common Wealth Ecclesiastical and Civil*, 1651
- Vincenzo Tagliasco, Annamaria D'Ursi, Riccardo, *Macchine che pensano: dalla cibernetica alla coscienza artificiale*, Media Duemila, 2002
- Martijn Kuipers, Ramjee Prasad, *Journey of Artificial Intelligence*, Published online 29 October 2021, under exclusive licence to Springer Science+Business Media, LLc, Part of Springer Nature, 2021
- Turing, A.M., *Computing machinery and intelligence*. *Mind*, 433-460, 1950
- J. McCarthy, Dartmouth College M. L. Minsky, *A proposal for the Dartmouth Summer Research Project on Artificial Intelligence*, Harvard University N. Rochester, I.B.M. Corporation C.E. Shannon, Bell Telephone Laboratories, 31 agosto 1955
- Kiyoshi Takenaka, Mayu Sakoda, *Sony says it has technology for humanoid robots, needs to find usage* Intervista di Reuters, 2022
- Charlie Beckett, *New Power New Responsibilities, A global survey of journalism and artificial intelligence*, London School of Economics and Political Science, Google News Initiative, novembre 2019
- Waleed Ali, Mohamed Hassoun, *Artificial Intelligence and Automated Journalism: Contemporary Challenges and New Opportunities*, *International Journal of Media, Journalism and Mass Communications (IJMJMC)*, vol 5, no. 1, pp. 40-49, 2019
- Arjen Van Dalen, *The algorithms behind the headlines, How machine-written news redefines the core skills of human journalists*, *Journalism Practice*, 2012
- Mathias-Felipe de Lima-Santos and Wilson Ceron, *Artificial Intelligence in News Media: Current Perceptions and Future Outlook*, *Journal Media*, 13-26, dicembre 2021
- Samuel Danzon-Chambaud *A systematic review of automated journalism scholarship: guidelines and suggestion for future research*, School of Communications, Dublin City University, 24 marzo 2021, open Research Europe 2021

- Samuel Danzon-Chambaud ,*A systematic review of automated journalism scholarship: guidelines and suggestion for future research*, , School of Communications, Dublin City University, open Research Europe 2021, 24 marzo 2021
- Tejedor, Santiago, and Pere Vila, *Exo Journalism: A Conceptual Approach to a Hybrid Formula between Journalism and Artificial Intelligence*, *Journalism and Media* 2: 830–840, 2021
- Francesco Marconi, *Newsmakers. Artificial intelligence and the future of Journalism*, Columbia University Press, New York, 2020
- Matthew R. Peroutka, Ronla K. Meiggs and Michael B. Romberg, *The generation of products in interactive forecast preparation* Techniques Development Laboratory, Colorado, 1998
- Ehud Reiter, Somayajulu Sripada, Jim Hunter, Jin Yu, Ian Davy, *Choosing words in computer-generated weather forecasts*, Department of Computing Science, University of Aberdeen, UK Aerospace and Marine International, Aberdeen, 15 agosto 2005
- Giovanni Ciofalo, *Infiniti anni Ottanta. Tv, cultura e società alle origini del nostro presente*, Mondadori, 2011
- Noain-Sánchez, A, *Addressing the Impact of Artificial Intelligence on Journalism: the perception of experts, journalists and academics*, *Communication & Society*, 35(3), 105-121, 2022
- Bass, A. Z., *Refining the “Gatekeeper” Concept: A UN Radio Case Study*, *Journalism Quarterly*, 46, 69–72, 1969
- Tim Marcin, *People started spreading fake news on twitter the instant they could buy a blue check*, mashable.com, 9 novembre 2022
- James Lee, *The 5 Computer Vision Techniques That Will Change How You See The World*, medium.com, Hearbeat, 12 aprile 2018
- *Powering Forestry 4.0: Planet Provides High-Frequency Insights Across Distributed Forest Assets*, Planet
- Rostyslav Demush, *A brief history of Computer Vision (and Convolutional Neural Networks)*, Hackernoon, 26 febbraio 2019
- *JanetBot: Analsing gender diversity on the FT homepage*, Financial Times Labs, 28 novembre 2022
- Dan Oved, Amelia Pisapia, Anna Gudnason, *Estimating 3D Poses of Athletes at Live Sporting Events*, *R&D New York Times*, 16 giugno 2020
- Alessandro Di Stefano, *Asc27, non è il robottino di Star Wars, ma una startup che porta l’IA in redazione*, StartupItalia, luglio 2021
- Marco Pratellesi, *Asimov, intelligenza artificiale e voci neurali: la rivoluzione del giornalismo parte da Roma*, La Repubblica, 19 agosto 2021
- T. De Mauro, *Grande dizionario italiano dell’uso*, Torino, 2000
- Sam Shead, *Why everyone is talking about the A.I. text generator released by an Elon Musk-backed lab*, CNBC, 23 luglio 2020

- Billy Perrigo, *Exclusive: OpenAI Used Kenyan Workers on Less Than \$2 to Make ChatGPT Less Toxic*, TIME, 18 gennaio 2023
- Christopher Bregler, Micheal Covell, Malcolm Slaney, *Video Rewrite: Driving Speech with Audio*, Interval Research Corporation, gennaio 1998
- Kyle Swenson, *A Seattle TV station aired doctored footage of Trump's Oval Office Speech. The employee has been fired*, The Washington Post, 11 gennaio 2019
- Aja Romano, *Jordan Peele's simulated Obama PSA is a double-edged warning against fake news*, Vox, 18 aprile 2018
- Francesco Stati, *Io, giornalista robot. La stampa tra progresso e tradizione*, ZetaLuiss, 23 aprile 2021
- Nic Newman, Richard Fletcher, Craig T. Robertson, Kirsten Eddy, Rasmus Kleis Nielsen, *Reuters Institute Digital News Report 2022*, 2022
- Aldo Fontanarosa, *Giornalisti Robot. L'Intelligenza Artificiale in redazione. Prove tecniche di news revolution*, UlisseAspettaPenelope, 14 ottobre 2020
- Nidhi Arora, Wei Wei Liu, Kelsey Robinson, Eli Stein, Daniel Ensslen, Lars Fiedler, Gustavo Schuler, *Next in Personalization Report*, McKinsey, 2021
- Walter Quattrociochi, Antonella Vicini, *Misinformation. Guida alla società della disinformazione e della credulità*, FrancoAngeli, 2016
- Barbara Gruber, *Facts, Fakes and Figures: How AI is influencing journalism*, Goethe Institute
- Jarno Kaponen, *First in the World: Yle's smart news assistant Voitto ensures that you don't miss the news you want to read*, News Lab Yle, 10 ottobre 2018
- Bryson J., Kime, P., *Just another artifact: Ethics and the empirical experience of AI presented at Fifteenth International Congress on Cybernetics*, Edinburgh, January 1998
- Kim, Haley, *AI in Journalism: Creating an Ethical Framework*, Syracuse University, Honors Program Capstone Projects, 2019
- B. Jones, E. Luger, *AI and Journalism - Intelligible Cloud and Edge AI (ICE-AI)*, PETRAS National Centre of Excellence, 17 settembre 2021
- Mohamed Abdulzاهر, *Artificial Intelligence Journalism and Professional Code of Ethics*, Artificial Intelligence Journalism for Research and Forecasting, ottobre 2020
- Billy Perrigo, *Exclusive: OpenAI Used Kenyan Workers on Less Than \$2 to Make ChatGPT Less Toxic*, TIME, 18 gennaio 2023
- Antonin Descampe, Clement Massart, Simon Poelman, Francois-Xavier Standaert, Olivier Standaert, *Automated news recommendation in front of adversarial examples and the technical limits of transparency in algorithmic accountability*, AI & Society, 13 marzo 2021
- Nicholas Diakopoulos & Michael Koliska, *Algorithmic Transparency in the News Media*, Digital Journalism, 5:7, 809-828, 2017
- Paolo Zuddas, *Brevi note sulla trasparenza algoritmica*, AmministrazioneInCammino, 5 giugno 2020

- Amina Adagi, Mohammed Berrada, *Peeking Inside the Black-Box: A Survey on Explainable Artificial Intelligence (XAI)*, IEEE access, 6, 52138-52160, 12 ottobre 2018
- Emanuele Capone, *Le intelligenze artificiali fra razzismo e questione etica*, La Repubblica, 6 aprile 2021
- Mara Magistroni, *Le intelligenze artificiali inquinano più del previsto*, WiredItalia, 7 giugno 2019
- Emily Chasan, *The environmental cost of keeping mail and files online keeps rising*, The Japan Times, 26 gennaio 2020
- Emma Strubel, Ananya Ganesh, Andrew McCallum, *Energy and Policy Considerations for Deep Learning in NLP*, College of Information and Computer Sciences, 5 giugno 2019
- Martel C, Pennycook G, Rand DG., *Reliance on emotion promotes belief in fake news*, Cogn Res Princ Implic,5(47):1–20, 2020
- Pennycook G, Rand DG., *Lazy, not biased: Susceptibility to partisan fake news is better explained by lack of reasoning than by motivated reasoning*, Cognition, 2019
- Bronstein MV, Pennycook G, Bear A, Rand DG, Cannon TD., *Belief in Fake News is Associated with Delusional, Dogmatic, Religious Fundamentalism, and Reduced Analytic Thinking.*, J Appl Res Mem Cogn., 2019
- Pennycook G, Cannon TD, Rand DG., *Prior Exposure Increases Perceived Accuracy of Fake News.*, J Ex Psychol Gen.;147(12):1865–80, 2018
- Amarene MA, Bucy EP., *Conferring Resistance to Digital Disinformation: The Inoculating Influence of Procedural News Knowledge.*, J Broadcast Electron Media, 2019
- Francesco Giorgino, *Giornalismi e Società. Informazione, politica, economia, cultura*, febbraio 2017
- Ceravolo F., Rosta M., *Raccontare la scienza: opportunità e rischi nella società dell'informazione*, in «Cambio. Rivista sulle trasformazioni sociali», 2019
- Luigi Ceccarini, Taina Pihlajarine, Anette Alén-Savikko, *Digital Citizenship and artificial intelligence: Information and disinformationI*, European Journal of Communication, Vol. 37(5) 563-568, 2022
- Tommaso Guerini, *Fake news e diritto penale. La manipolazione digitale del consenso nelle democrazie liberali*, Itinerari di Diritto Penale, 2020
- Alonso M.A., Vilares D., Gómez-Rodríguez C., Vilares J., *Sentiment Analysis for Fake News Detection*, Electronics, 2021
- Kai Shu, Amy Sliva, Suhang Wang, Jiliang Tang, and Huan Liu, *Fake News Detection on Social Media: A Data Mining Perspective* Computer Science & Engineering, Arizona State University, Tempe, AZ, USA Charles River Analytics, Cambridge, MA, USA, 2017
- Kai Shu, Suhang Wang, Huan Liu, *Beyond News Contents: The Role of Social Context for Fake News Detection*, Twelfth ACM International Conference on Web Search and Data mining, 312-320, gennaio 2019
- Alonso, M.A.; Vilares, D.; Gómez-Rodríguez, C.; Vilares, J. *Sentiment Analysis for Fake News Detection*. Electronics 2021

- Chiang T.H.C., Liao C.S., Wang W.C., *Investigating the Difference of Fake News Source Credibility Recognition between ANN and BERT Algorithms in Artificial Intelligence*, Appl. Sci. 2022
- Samhir Vasdev, *Can machine learning help us measure the trustworthiness of news?*, IREX, settembre 2018
- Perez-Rosas, Bennett Kleinberg, Alexandra Lefevre, *Automatic Detection of Fake news*, Computer Science and Engineering, University of Michigan, University of Amsterdam, 2017
- Angel Vizoso, Martin Vaz-Alvarez, Xosè Lopez-Garcia, *Fighting Deepfakes: Media and Internet Giants' Converging and Diverging Strategies Against Hi-Tech Misinformation*, Media and Communication, Volume 9, Issue 1 Pages 291-300, 3 marzo 2021
- Jesus Perez Dasilva, Koldobika Meso Ayerdi, Terese Mendiguren Galdospin, *Deepfakes on Twitter: Which Actors Control Their Spread?*, Media and Communication, Volume 9, Issue 1, Pages 301-312, 3 marzo 2021
- Stefano Tenedini, *Giornalismo e AI: un rapporto complicato ma inevitabile. E utile.*, impactscool magazine, robotica e ai, 20 febbraio 2019
- Andrea Daniele Signorelli, *Questo articolo non è stato scritto da un robot*, Wired Italia, Scienza, 13 febbraio 2019
- Simone Catania, *Giornalismo automatizzato: l'intelligenza artificiale genera notizie*, news.srl, 23 luglio 2020
- LinkedIn, *Lavori in crescita nel 2022: le 25 professioni in ascesa*, LinkedIn Notizie, 18 gennaio 2022
- Joo-Wha Hong, *I Was Born to Love AI: The Influence of Social Status on AI Self-Efficacy and Intentions to Use AI*, International Journal of Communication 16, 172-191, 2022
- Marisela Gutierrez Lopez, Colin Porlezza, Glenda Cooper, Stephann Makri, Andrew MacFarlane, Sondess Missaoui, *A question of Design: Strategies for Embedding AI-Driven Tools into Journalistic Work Routines*, Digital Journalism, 2022
- Mohamed Abdulzاهر, *Artificial Intelligence Journalism and Media Digital Soft-Skills*, gennaio 2020
- Gloria Gomez-Diago, *Perspectives to address artificial intelligence in journalism teaching. A review of research and teaching experiences*, Revista Latina de Comunicacion Social, 80, 29-46, 2022
- Alice Lanciano, *La pandemia ha aggravato la crisi dei giornali*, IlPost, 17 luglio 2022
- Censis, *I media dopo la Pandemia*, Sintesi per la stampa del Diciassettesimo rapporto sulla comunicazione, 6 ottobre 2021
- Eurobarometer of European Commission, *European citizens' knowledge and attitudes towards science and technology*, 2021
- Charlie Beckett, *New Power New Responsibilities, A global survey of journalism and artificial intelligence*, London School of Economics and Political Science, Google News Initiative, novembre 2019
- Lelio Simi, *Membership e giornali* (due idee per discuterne), DataMediaHub, 22 maggio 2014

- Eurobarometer of European Commission, *European citizens' knowledge and attitudes towards science and technology*, 2021
- Joo-Wha Hong, *I Was Born to Love AI: The Influence of Social Status on AI Self-Efficacy and Intentions to Use AI*, *International Journal of Communication* 16(2022), 172-191, 2022
- María T. Soto-Sanfiel, Adriana Ibiti, Mabel Machado, Beatriz Elena Marín Ochoa, María Mendoza Michilot, Claudio Guillermo Rosell Arce & Ariadna Angulo Brunet, *In Search of the Global South: Assessing Attitudes of Latin American Journalists to Artificial Intelligence in Journalism*, *Journalism Studies*, 23:10, 1197-1224, 2022
- Jennifer Elias, *Google is asking employees to test potential ChatGPT competitors, including chatbot called "Apprentice Bard*, CNBC, gennaio 2023
- Andreas Graefe and Nina Bohlken, *Automated journalism: A Meta-Analysis of Readers' Perceptions of Human-Written in Comparison to Automated News*, *Media and Communication*, Volume 8, Issue 3, pages 50-59, ISSN: 2183-2439, 2020
- Thomas Baekdal, *Robots vs Journalists? It's happening*, baekdal.com, marzo 2014
- Nils Kobis, Luca D. Mossink, *Artificial intelligence versus Maya Angelou: Experimental evidence that people cannot differentiate AI-generated from human-written poetry*, *Computer in Human Behaviour*, Elsevier, settembre 2020
- Chiara Langoni, Andrey Fradkin, Gordon Pennycook, *News from Generative Artificial Intelligence Is Believed Less*, *ACM Conference on Fairness, Accountability and Transparency*, giugno 2022
- Nils Kobis, Luca D. Mossink, *Artificial intelligence versus Maya Angelou: Experimental evidence that people cannot differentiate AI-generated from human-written poetry*, *Computer in Human Behaviour*, Elsevier, settembre 2020
- M.A. Beam, G.M. Kosicki, *Personalized news portals: filtering systems and increased news exposure*, *Journal. Mass Commun.*, 2014
- S. Kalyanaraman, S.S. Sundar, *The psychological appeal of personalized content in web portals: does customization affect attitudes and behavior?* *J. Commun.* 56, 2006
- Q. Ye, Y. Luo, G. Chen, X. Guo, Q. Wei, S. Tan, *Users intention for continuous usage of mobile news apps: the roles of quality, switching costs, and personalization*, *J. Syst. Sci. Syst. Eng.*, 2019
- B. Bodo, *Selling news to audiences – a qualitative inquiry into the emerging logics of algorithmic news personalization in European quality news media*, *Digit. Journal* 7 (2019) 1054–1075
- Hindman, Matthew, *Journalism Ethics and Digital Audience Data. In Remaking the News: Essays on the Future of Journalism Scholarship in the Digital Age*, edited by Pablo J. Boczkowski and C. W. Anderson, 177–195. Cambridge, MA: MIT Press, 2017

Sitografia

- https://blog.osservatori.net/it_it/intelligenza-artificiale-funzionamento-applicazioni
- <https://www.cwi.it/tecnologie-emergenti/intelligenza-artificiale/da-talos-turing-la-storia-dellintelligenza-artificiale-136258>
- https://it.frwiki.wiki/wiki/Histoire_de_l%27intelligence_artificielle
- <https://www.marinellionoranze.it/due-millenni-di-robot/>
- <https://www.vanillamagazine.it/le-geniali-invenzioni-di-erone-di-alessandria-una-tecnologia-che-non-aveva-scopi-pratici/>
- <https://www.euautomation.com/it/automated/magazine/article/breve-storia-dellintelligenza-artificiale>
- <https://www.cyberlaws.it/2018/la-storia-dellintelligenza-artificiale-da-turing-ad-oggi/>
- https://it.wikipedia.org/wiki/Conferenza_di_Dartmouth
- <https://www.civiltadellemacchine.it/it/news-and-stories-detail/-/detail/ritorno-della-metafisica-passa-per-l-intelligenza-artificiale>
- <https://www.euautomation.com/it/automated/article/le-tre-tappe-fondamentali-dellintelligenza-artificiale>
- <https://www.studenti.it/discorso-sul-metodo-di-cartesio-descrizione-e-riassunto.html>
- <https://formazione.indire.it/paths/testo-1-cartesio-discorso-sul-metodo>
- <https://gtfondazione.org/industria-4-0/intelligenza-artificiale-dallantichita-al-xix-secolo/>
- <https://www.logogramma.com/antichi-e-recenti-creatori-di-intelligenze-artificiali/>
- <https://www.varesenews.it/2017/02/intelligenza-artificiale-un-sogno-delluomo-sin-dallantichita/595906/>
- <https://www.laltragenesi.it/2016/06/i-robot-degli-antichi-dei.html>
- <https://www.romatoday.it/eventi/cultura/silvestro-ii-tesoro-campo-marzio.html#:~:text=Solitamente%20identificato%20come%20una%20testa,proveniente%20dal%20tesoro%20di%20Augusto.>
- <https://mailchi.mp/lse.ac.uk/closing-the-ai-gapresources-for-journalists?e=77ac9a9f6b>
- <https://blogs.lse.ac.uk/polis/2022/08/02/introducing-the-participants-of-the-journalism-academy-for-small-newsrooms/>
- <https://aibstudi.aib.it/article/view/13365>
- <https://www.jstor.org/stable/23004044>
- https://journals.sagepub.com/doi/full/10.1177/1464884918757072?utm_source=summon&utm_medium=discovery-provider&
- <https://airtable.com/shrvumtqhlJMnUx7g/tblDDdPAvmtJLpKq1>
- <https://www.hoepli.it/libro/robot-journalism/9789813237339.html>
- <https://www.worldscientific.com/worldscibooks/10.1142/10913#t=aboutBook>
- https://www.researchgate.net/publication/265340945_Enter_the_Robot_Journalist

- <https://portal.findresearcher.sdu.dk/en/publications/the-algorithms-behind-the-headlines-how-machine-written-news-rede>
- <https://www.semanticscholar.org/paper/Mapping-the-field-of-Algorithmic-Journalism-D%C3%B6rr/2b7728fd35983ef227de8adbf10b76a2976e0c97>
- <https://journalismai.com/2015/11/03/mapping-the-field-of-algorithmic-journalism/>
- <https://www.semanticscholar.org/paper/computer-produced-worded-forecasts-Glahn/8b32b141d99b3d6e87c86ed6e403a441b5427d47>
- https://journals.ametsoc.org/view/journals/bams/51/12/1520-0477_1970_051_1126_cpwf_2_0_co_2.xml
- https://en.wikipedia.org/wiki/Outline_of_artificial_intelligence
- https://en.wikipedia.org/wiki/Dartmouth_workshop
- https://en.wikipedia.org/wiki/Computing_Machinery_and_Intelligence
- https://en.wikipedia.org/wiki/Intelligent_agent
- https://en.wikipedia.org/wiki/Turing_test
- https://en.wikipedia.org/wiki/Alan_Turing
- https://en.wikipedia.org/wiki/Artificial_intelligence
- <https://www.ancientfaces.com/person/harry-h-glenn-birth-1913-united-states/154966542>
- <https://www.metoffice.gov.uk/weather/learn-about/how-forecasts-are-made/computer-models/history-of-numerical-weather-prediction>
- https://en.wikipedia.org/wiki/Pamela_McCorduck
- <https://www.tandfonline.com/doi/abs/10.1080/21670811.2015.1096748?journalCode=rdij20>
- <https://www.ibm.com/cloud/learn/deep-learning>
- <https://www.ai4business.it/intelligenza-artificiale/deep-learning/deep-learning-cose/>
- <https://250.dartmouth.edu/highlights/artificial-intelligence-ai-coined-dartmouth>
- <https://web.archive.org/web/20080702224846/http://loebner.net/Prize/TuringArticle.html>
- <https://www.agendadigitale.eu/cultura-digitale/intelligenza-artificiale-i-concetti-chiave-per-comprenderla-meglio/>
- https://www.jstor.org/stable/j.ctv65swg4.9?seq=1#metadata_info_tab_contents
- https://www.thestar.com/sports/baseball/2009/10/23/can_a_computer_cover_a_ball_game.html#article
- <https://singularityhub.com/2009/11/09/is-software-set-to-replace-sports-journalists/>
- <https://www.wbur.org/npr/122424166/story.php>
- https://www.boorp.com/notizie_articoli_news_post/articolo_Un_algoritmo_puo_scrivere_meglio_di_un_giornalista.php
- <https://archive.nytimes.com/mediadecoder.blogs.nytimes.com/2009/10/19/the-robots-are-coming-oh-theyre-here/>

- <https://www.ibm.com/it-it/topics/computer-vision>
- https://blog.osservatori.net/it_it/computer-vision-definizione-applicazioni
- <https://www.internet4things.it/industry-4-0/computer-vision-cose-come-funziona-e-applicazioni/>
- <https://www.mobiuslabs.com/resources/4-ways-superhuman-computer-vision-is-transforming-journalism-and-broadcasting>
- <https://www.ibm.com/it-it/topics/computer-vision>
- <https://www.sas.com/content/dam/SAS/documents/infographics/2019/en-computer-vision-110208.pdf>
- <https://vitolavecchia.altervista.org/differenza-tra-computer-vision-e-machine-vision-in-informatica/www.netapp.com>
- <https://www.dmif.uniud.it/magistrale/artificial-intelligence-cybersecurity/piano-di-studio/computer-vision/>
- https://blog.osservatori.net/it_it/computer-vision-definizione-applicazioni
- <https://open.nytimes.com/how-the-new-york-times-uses-software-to-recognize-members-of-congress-29b46dd426c7>
- <https://www.dataninja.it/portfolio/ai-anchor-sole24ore-it/>
- <https://magazine.dataninja.it/2021/04/22/intelligenza-artificiale-podcast-automatici/>
- <https://stream24.ilsole24ore.com/video/italia/royalty-e-anchor-ecco-progetti-innovazione-digitale-gruppo-24ore-finanziati-google/AEsEmjRF>
- https://cdn.openai.com/research-covers/language-unsupervised/language_understanding_paper.pdf
- <https://proceedings.neurips.cc/paper/2014/file/5ca3e9b122f61f8f06494c97b1afccf3-Paper.pdf>
- <https://www.primaonline.it/2018/11/09/280453/ecco-il-giornalista-tv-artificiale-lagenzia-di-stampa-cinese-xinhua-sperimenta-il-primo-anchor-man-virtuale/>
- <https://tg24.sky.it/tecnologia/2018/11/08/debutta-cina-primo-anchorman-intelligenza-artificiale>
- <https://www.ai4business.it/intelligenza-artificiale/governo-il-futuro-dellintelligenza-artificiale-e-nella-trasparenza-degli-algoritmi/>
- <https://lavalibera.it/it-schede-749-algoritmi-governano-il-mondo-intelligenza-artificiale>
- <https://www.amministrazioneincammino.luiss.it/wp-content/uploads/2020/06/ZUDDAS.pdf>
- <https://www.wfw.com/articles/the-transparency-of-algorithms-between-the-artificial-intelligence-act-and-the-italian-courts/?l=it>
- <https://www.studiolegaledelliponti.eu/rendere-sempllice-qualcosa-di-complesso-la-trasparenza-degli-algoritmi-come-prerequisito-del-consenso/>
- https://fpsshare.it/intelligenza-artificiale-emissioni/?cli_action=1675351203.375
- <https://www.thewisemagazine.it/2022/01/01/intelligenza-artificiale-e-cambiamento-climatico-il-dilemma-del-futuro/>
- <https://www.japantimes.co.jp/news/2020/01/26/business/tech/email-global-warming/>

- <https://www.agendadigitale.eu/cultura-digitale/impatto-ambientale-la-grande-sfida-dellintelligenza-artificiale/>
- <https://www.agendadigitale.eu/infrastrutture/intelligenza-artificiale-cosi-diventera-una-tecnologia-verde/>
- <https://fattinofake.federchimica.it/articolo/le-fake-news-non-sono-lunico-problema/>
- <https://www.ekr.admin.ch/i850.html>
- <https://www.agendadigitale.eu/cultura-digitale/fake-news-possibile-soluzione-algoritmi-piu-trasparenti/>
- <https://www.agendadigitale.eu/cultura-digitale/perche-ci-credo-la-credulita-online-come-strumento-di-influenza-sociale/>
- <https://www.agcom.it/documents/10179/12791486/Pubblicazione+23-11-2018/93869b4f-0a8d-4380-aad2-c10a0e426d83?version=1.0>
- <https://www.sciencedirect.com/science/article/pii/S1364661321000516>
- <https://emplus.egeaonline.it/it/61/archivio-rivista/rivista/3446560/articolo/3446580>
- <https://www.europarl.europa.eu/news/it/headlines/society/20200827STO85804/che-cos-e-l-intelligenza-artificiale-e-come-viene-usata>

Riassunto

Questo riassunto è stato scritto con l'ausilio della funzione “*summurize*” di QuillBot AI, strumento di parafrasi con intelligenza artificiale che aiuta studenti e professionisti a riscrivere, modificare e cambiare il tono del testo per migliorarne la chiarezza. L'approccio utilizzato è quello sostenuto all'interno della tesi, ovvero di utilizzare le intelligenze artificiali come “esoscheletri” e rivedere in ultima analisi il loro operato per rifinirlo al meglio, secondo il concetto di *human in the loop*. La versione disponibile su Chrome, per riassumere, seleziona le frasi principali di un testo, commettendo talvolta qualche errore. Anche per questo si è deciso di utilizzare l'approccio appena menzionato.

Capitolo 1 Concetti, definizioni e storia dell'Intelligenza Artificiale nel giornalismo

La moderna concezione di Intelligenza Artificiale è più complessa di quanto si possa pensare. Si può parlare di intelligenza artificiale forte, cioè che riproduce in tutto e per tutto azioni tipiche dell'uomo, e intelligenza artificiale debole, cioè un'intelligenza che riproduce solo in parte compiti umani e che tipicamente viene affiancata da un vero e proprio essere umano nello svolgimento della propria funzione. Talvolta questo termine viene accostato all'idea di robot ma è in realtà diversa da essa. Il termine robot ha un'accezione più inerente allo spazio fisico e alle azioni pratiche all'interno di esso, mentre l'intelligenza artificiale riguarda per lo più sistemi e programmi capaci di assolvere azioni accostabili a processi mentali. Spesso anche altri termini, come *Machine Learning*, *Deep Learning* e *Natural Language Processing* vengono scambiati, inglobati, sovrapposti tra loro, talvolta in maniera opportuna, quasi sempre in maniera indebita. L'ostacolo da tenere bene a mente è che anche se esistono suddivisioni e concettualizzazioni condivise, non lo sono mai in tutto e per tutto e risulta impossibile incasellare, categorizzare e gerarchizzare in maniera precisa tali concetti. Nonostante ciò, è stato comunque utile cercare di esemplificare il più possibile i termini in questione, oltre che a darne una gerarchia indicativa e fornire più elementi possibili al lettore al fine di una comprensione a 360° degli ambiti dell'Intelligenza Artificiale.

Il termine “Intelligenza Artificiale” è un termine ombrello, omnicomprensivo e comunemente usato per racchiudere tutti gli altri.

Il *Machine Learning* (ML) è una branca dell'Intelligenza Artificiale. Il termine è utilizzato per identificare quelle tecnologie che auto-apprendono: in base a come sono state impostate sono in grado, di volta in volta, di generalizzare in maniera sempre più precisa, senza bisogno di modificare gli input.

Artificial Neural Network (ANN) sta ad indicare le reti neurali, sorta di reti neuronali artificiali che hanno una complessità maggiore rispetto a software o algoritmi elementari di auto-apprendimento.

Il *Deep Learning* è anch'esso una sotto branca del *Machine Learning*, forse la più innovativa e quella con più prospettive future, che sfrutta reti neurali capaci di apprendere senza supervisione, senza cambiamenti di input e senza la necessità di dati strutturati.

La differenza tra software e algoritmo, invece, è molto banale: un software (o un programma) può essere definito in maniera esemplificativa come un insieme di algoritmi. La maggior parte delle tecnologie che menzioneremo o analizzeremo saranno software, ma non mancano anche esempi di singoli algoritmi che vengono utilizzati per determinati scopi.

Le categorie presentate all'interno del lavoro sono state molte e, come detto, alcune di esse comprendono le altre, si intersecano tra di loro o variano per piccolissime differenze tecniche. Sono comunque utili alla comprensione dell'ambito scientifico delle IA, anche se presentate in linea generale e suddivise secondo criteri di specificità e complessità.

Per andare ancora più a fondo e capire a pieno la natura delle Intelligenze Artificiali è stato necessario tornare indietro di millenni, andando a capire cosa ha motivato la creazione di strumenti di questo genere e come siamo arrivati all'odierno *status quo*.

Il concetto di intelligenza artificiale in senso lato ha avuto origine nell'antichità. L'idea che qualcosa di altro, di diverso dall'uomo, potesse assumere tratti simili ad esso ha sempre caratterizzato la mitologia e la storia umana. A livello teorico, essa pone le sue radici nelle teorie filosofiche del meccanicismo e del determinismo del XVII secolo. Queste dottrine definiscono il mondo e l'essere umano come una serie di meccanismi causa-effetto individuabili, conoscibili e riproducibili. A tali scuole di pensiero si rifacevano Cartesio ed Hobbes, che abbiamo definito in questa sede come i padri della concezione meccanica del pensiero umano. Nel XVIII secolo, invece, sono stati effettuati diversi tentativi di ricreare figure simili all'umano. Nell'Ottocento lo studio su queste tematiche ha iniziato a riguardare di più il mondo "scientifico" in senso stretto, in particolare matematici e fisici e gli studi nello specifico hanno iniziato a riguardare principalmente le teorie del pensiero deduttivo e le connessioni neurali, oltre che la logica e la probabilità. Oggi diremmo che si stava passando concettualmente da un interesse per i robot ad un interesse per le intelligenze artificiali.

Il primo vero e proprio tentativo di spaccettare la logica di pensiero venne effettuato nel 1854. Per l'inizio del Novecento era ormai comunemente accettato che la maggior parte dei principi matematici potesse essere espresso da un insieme di nozioni primitive accostabili a logicismi. Alan Turing, considerato il padre dell'Intelligenza Artificiale, insieme ad altri studiosi, consapevoli dell'incompletezza dei modelli precedenti, tentò a metà Novecento di introdurre un metodo per testare una macchina, provando a capire se essa fosse capace di svolgere qualsiasi calcolo computazionale. I risultati constatarono che ci fossero dei limiti, al tempo considerati invalicabili, che la logica matematica non avrebbe potuto superare. Il focus in questo periodo si spostò dalla volontà di superare i limiti alla possibilità di, consapevoli di questa limitatezza, dare risvolto pratico a tutto questo progresso teorico creando effettivamente delle macchine che potessero eseguire calcoli. Da lì in poi, fino agli anni Settanta, si attraversò un periodo d'oro per l'intelligenza artificiale: nuovi programmi vennero avviati, tra cui computer che risolvevano problemi di algebra e imparavano a parlare inglese, accompagnati dall'entusiasmo dell'opinione pubblica.

In prossimità degli anni Ottanta, però, il governo britannico e quello statunitense smisero di finanziare le ricerche riguardanti questo ambito. Parallelamente, le potenzialità di tali progetti si scontrarono con una realtà dei fatti che vide le macchine non avere la capacità di raggiungere gli obiettivi sperati.

Negli anni Novanta si riscontrò un rinnovato interesse per l'intelligenza artificiale con la consapevolezza di una percezione poco ottimistica da parte del settore economico e la definitiva rassegnazione sul superamento di quei limiti apparentemente invalicabili. Da qui in poi questa branca di ricerca scientifica iniziò a settorializzarsi sempre di più.

Al giorno d'oggi, le tecnologie che usano IA per svolgere le proprie funzioni sono innumerevoli e alla portata di tutti. Da Alexa a Siri, dai Tools di Google sino ai sistemi di riconoscimento facciale, l'Intelligenza Artificiale è entrata a pieno nelle nostre vite. Anche nel campo della robotica i progressi sono sensazionali. Le potenzialità sono molte e i rischi altrettanti. Urge la necessità di trovare un giusto equilibrio tra i miglioramenti che le Intelligenze Artificiali possono portare e i rischi ad esse correlati, come la privacy e la protezione dei dati, oltre che al possibile impatto negativo sull'occupazione. In ogni caso non si può ignorare la realtà dei fatti e accettare che queste innovazioni faranno sempre più parte della nostra realtà quotidiana e lavorativa.

C'è un settore che, più di altri, pone in essere alcune questioni etiche sul bilanciamento tra logica del profitto e utilità per i cittadini sull'utilizzo delle Intelligenze Artificiali: il giornalismo. Il giornalismo non è solo un settore economico delle moderne società. È uno degli aspetti fondamentali della democrazia: mettere a conoscenza i cittadini di ciò che accade nel mondo e nelle istituzioni, fornire dati accurati e dare voce a più opinioni possibile per promuovere il pluralismo sono solo alcuni degli obiettivi che una società democratica dovrebbe porsi. È ormai sotto gli occhi di tutti che al giorno d'oggi esso sia un settore in crisi e non riesca più ad avere quel ruolo centrale di mediatore che aveva un tempo. La lettura di giornali cartacei è diminuita drasticamente negli ultimi anni e continua a scendere vertiginosamente, ma è in aumento la lettura di giornali online, siti web di informazione, radio online e telegiornali. Nonostante lo scenario poco positivo, il numero di giornali nati negli ultimi anni è aumentato. Ciò vuol dire che non mancano all'attivo giornalisti e imprenditori che vogliono investire nel campo dell'informazione. L'intelligenza artificiale nel giornalismo, per quanto recente e ancora poco sviluppata soprattutto nelle redazioni italiane, potrebbe essere utile proprio per soddisfare queste esigenze: potrebbe ridurre i costi nella produzione e distribuzione di notizie e potrebbe far sì che vengano soddisfatte a pieno le esigenze dei lettori. Prima di analizzare le intelligenze artificiali utilizzate nel giornalismo, è stata necessaria un'ulteriore concettualizzazione e contestualizzazione sull'oggetto di questo lavoro. Nell'ultimo decennio il *Data Journalism* è diventato la forma preponderante di giornalismo grazie alla grande mole di dati disponibili. Insieme ad esso, però, si sono diffuse altre forme di giornalismo digitale come l'*automated journalism*, il *metrics driven journalism* e l'*algorithmic journalism*. Il termine forse più utilizzato di quelli sopramenzionati è *Automated Journalism*. Per *Automated Journalism* si intende una "specializzazione dell'informatica che mira a sostituire del tutto i giornalisti umani per sostituirli con degli algoritmi in grado di prelevare contenuti e riscriverli secondo dei criteri prestabiliti dall'uomo".

Sul fatto se sostituiscano o meno effettivamente il lavoro del giornalista abbiamo discusso nell'ultimo capitolo. Ci interessava sottolineare inizialmente come venissero utilizzati i termini in questione.

Un altro termine molto utilizzato è *NLG Journalism*. Il termine si riferisce agli algoritmi di generazione del linguaggio per aiutare gli esseri umani a scrivere articoli. La sigla NLG, come detto, significa “*Natural Language Generation*” e identifica però solo la parte dell'output del processo, cioè la generazione del linguaggio. Il termine più appropriato per identificare l'intero processo è *Natural Language Processing*, che comprende in sé sia il *Natural Language Generation* ma anche il *Natural Language Understanding*, cioè la parte dell'input. Un termine invece atipico che, a mio avviso, rispecchia molto il ruolo delle intelligenze artificiali nel giornalismo è l'“*exojournalism*”, una proposta concettuale che connette le possibilità offerte dall'IA con la routine giornalistica, intendendo l'IA come una sorta di esoscheletro del giornalista.

Partendo da questi presupposti teorici, abbiamo analizzato come si è arrivati al moderno utilizzo delle intelligenze artificiali nel giornalismo.

I quotidiani hanno attraversato quattro secoli e sono passati per diversi supporti, modificando le proprie caratteristiche e facendosi interpreti di cambiamenti sociali, politici ed economici, talvolta anche stimolandoli. Si è passati dal cartaceo al digitale, dal testo all'ipertesto, dai salotti ai social network. La storia dei media mostra che i produttori di notizie, pur legati ai mezzi, si sono trasformati non in conseguenza di un'evoluzione a sé stante, ma per un complesso intreccio di relazioni tra dinamiche culturali, sociali ed economiche. L'intelligenza artificiale si inserisce in questo contesto nella seconda metà del XX secolo, quando internet è entrato a far parte pian piano delle nostre vite e l'IA iniziava a svilupparsi come vero e proprio settore scientifico.

Sfogliando la letteratura scientifica e cercando sui motori di ricerca, il primo esempio di utilizzo di intelligenze artificiali nel mondo dell'informazione è sembrato essere il prototipo di programma di scrittura automatica “*Stats Monkey*”, ideato da Kristian Hammond e Larry Birnbaum all'interno di un progetto dell'Intelligent Information Laboratory della Northwestern University.

Nonostante tale primato risulti corretto, non è stato il primo esempio di vera e propria scrittura automatica.

Il primo vero esempio di scrittura automatica è stato un computer contenente 80 frasi preimpostate che, partendo da input di dati rilevati ed elaborati dall'ESSA Weather Bureau's National Meteorological Center, produceva brevi articoli di previsioni meteo. Le previsioni erano automatizzate nella scrittura di piccoli articoli meteorologici.

Nel 2009, invece, un gruppo di studenti di giornalismo e informatica seguiva un corso della scuola di giornalismo della Northwestern University, condotto da Kristian Hammond, informatico con dottorato a Yale, e Larry Birnbaum, suo collega. Stava nascendo un progetto all'interno dell'Intelligent Information Laboratory per un programma di scrittura automatica che, partendo da un set di dati statistici, produceva brevi storie base, giocata dopo giocata, di una partita di baseball. Stava nascendo StatsMonkey.

Capitolo 2 L'intelligenza artificiale nel giornalismo: utilizzi ed esempi virtuosi

Una suddivisione generale in macrocategorie, largamente condivisa e che ci è stata utile per capire meglio l'utilizzo delle tecnologie, seguendo un criterio di suddivisione riguardante gli step del processo di creazione di contenuti nelle redazioni è la seguente: *News Gathering*, *News Production* e *News Distribution*.

È doveroso sottolineare ancora una volta che queste categorie, per quanto utili e necessarie, sono soltanto indicative. Più è complessa una tecnologia e più è probabile che essa sia polifunzionale, cioè che assolva più funzioni contemporaneamente. Prima di analizzare nello specifico i tipi di tecnologie è necessaria una puntualizzazione. Nonostante la scrittura automatica sia la più menzionata e utilizzata nel mondo dell'informazione, non è l'unica. Spesso in letteratura viene data maggiore enfasi a questo tipo di tecnologie e la maggior parte della analisi a posteriori, che abbiamo visto nei capitoli successivi, vengono effettuate su contenuti di questo tipo. Non si sostiene in questa sede che esse non siano le più importanti e che non sia giusto utilizzare questi parametri. Si sostiene invece, e lo si corrobora attraverso le concettualizzazioni e categorizzazioni utilizzate e menzionate in questo lavoro, l'importanza di analizzare anche altri tipi di tecnologie.

Il termine *News Gathering* tradotto letteralmente vuol dire "raccolta di notizie". Il termine *to gather* significa proprio radunare, raccogliere, riunire e nell'ambito del giornalismo indica il processo iniziale di creazione di una notizia. Abbiamo suddiviso le tipologie di IA utilizzate per il *News Gathering* in: *Computer Vision* (o visione artificiale), previsione di tendenze e ricerca e verifica di informazioni (per combattere disinformazione e misinformazione). La verifica di informazioni è stata analizzata nel terzo capitolo riguardante l'etica, nel quale abbiamo parlato di *fake news* e *fact checking* e quindi di verifica di informazioni.

La visione artificiale è un campo dell'intelligenza artificiale che permette ai computer e ai sistemi di ricavare informazioni significative da immagini digitali, video e altri input visivi - e intraprendere azioni o formulare delle segnalazioni sulla base di tali informazioni. La *Computer Vision* riesce in pochissimo tempo ad analizzare innumerevoli immagini e video che richiederebbero giorni. Queste tecnologie necessitano di un ampio dataset di immagini da cui trarre le informazioni basilari per poter operare (distinzione colori, oggetti ecc.). Il termine racchiude in sé tutte le tecnologie di analisi visuale, da quelle satellitari ai Qr Code.

Rientrano nelle tecnologie *Computer Vision* (CV) utilizzate per il giornalismo: tecnologie per la classificazione di immagini, tecnologie di riconoscimento di immagini per la ricerca di notizie, tecnologie di riconoscimento motorio e tecnologie di riconoscimento facciale.

Le tecnologie di classificazione di immagini possono essere utili per l'archiviazione di grandi quantità di immagini. Per quanto riguarda le tecnologie di ricerca di immagini esse possono essere utilizzate soprattutto per la ricerca automatizzata di notizie sul web. Questi software di ricerca sono utilizzati anche per immagini satellitari. Spesso queste strumentazioni sono utilizzate parallelamente a dei sensori per il monitoraggio di determinati eventi, come il monitoraggio ambientale o le catastrofi naturali.

La *Computer Vision* è la sottobranchia che ha avuto più difficoltà a svilupparsi. Nonostante i recenti e notevoli progressi, non siamo ancora vicini ad una *Computer Vision* senza errori. Tuttavia, sono già numerose le istituzioni sanitarie e le imprese che hanno trovato il modo di applicare i sistemi CV ai problemi del mondo reale.

Le tecnologie di riconoscimento motorio, insieme a quelle di riconoscimento facciale, sono tecnologie di analisi video e a differenza di quelle precedentemente menzionate sono più complesse, in quanto assimilabili ad un'analisi di innumerevoli immagini al secondo. Le tecnologie di riconoscimento motorio analizzano i movimenti di un determinato soggetto andando a fornire dati che l'uomo da solo avrebbe difficoltà o impiegherebbe troppo tempo a sviluppare. Le tecnologie di riconoscimento facciale sono spesso basate su software di analisi dei componenti principali e riescono a identificare i volti delle persone, facendoli riscontrare con immagini simili e riuscendo quindi a identificarne l'identità.

Sono stati riportati esempi di utilizzi di queste tecnologie da parte del New York Times.

La ricerca di immagini, invece, può essere utile anche per la ricerca di notizie sul web. Le immagini satellitari spesso aiutano a scovare eventi o cambiamenti di ampia portata. La ricerca di informazioni non avviene solo attraverso l'analisi di pixel ma anche tramite elementi testuali. Esistono software, però, che in maniera più complessa e accurata scovano notizie inerenti un determinato argomento e sono pensati per affiancare il giornalista nella ricerca di una notizia. Anzi, possono addirittura scovare quelle che saranno le notizie di domani. La scelta di una notizia, prima dell'avvento di alcune tecnologie, avveniva tramite ricerche e sensazioni del giornalista. Spesso si aspettava la notizia da agenzie o da fonti dirette che il giornalista conosceva. Ora algoritmi e software possono consigliarci quale sia l'argomento migliore di cui parlare e quale sarà l'engagement del nostro articolo nel tempo.

È stato riportato un esempio di software di questo tipo progettati da una start-up italiana.

La fase di creazione vera e propria del contenuto è la più importante del processo. Il modo in cui si scrive un articolo, si crea un podcast o si registra un servizio televisivo sarà fondamentale per la qualità della notizia. I software utilizzati in questa fase sono molteplici e abbiamo suddiviso in maniera generale e sintetica questa moltitudine in tre categorie. Abbiamo suddiviso le tipologie di IA utilizzate per la *News Production* in: software audio o video, scrittura automatica e *deepfake*. Questi tipi di software, in realtà, possono essere utilizzati anche prima o dopo la creazione effettiva della news. Se devo trascrivere un'intervista, se voglio inserire la possibilità per un lettore online di sentire l'audio di un articolo piuttosto che leggerlo o se voglio creare un podcast partendo da del testo scritto, oggi si può fare con l'utilizzo di intelligenze artificiali.

È stato riportato un esempio di podcast automatico utilizzato dal Sole24Ore.

I software video, invece, sono utilizzati per ottenere una migliore qualità visuale e per poter formattare i video senza che occupino troppo spazio. Le potenzialità della qualità video in generale si sono espresse al loro massimo negli ultimi anni, raggiungendo praticamente i livelli della visione umana.

I software di scrittura automatica sono utilizzati per scrivere notizie in maniera automatica o semiautomatica. È stato riportato l'esempio di Chat-GPT, la più famosa e utilizzata non solo nell'ambito giornalistico ma anche sul web.

I *deepfakes* sono video falsi creati utilizzando software digitali simili alla Computer Vision basati sul *deep learning* e il *face swapping*. La tecnologia *deepfake* è in grado di creare immagini convincenti ma completamente fittizie partendo da zero. I *deepfakes* vengono utilizzati per diversi scopi, dalla disinformazione politica al cyberbullismo ai crimini informatici, sino alle *fake news*. Di questa tecnologia si è fatto un uso diverso negli ultimi anni, che si discosta notevolmente dal mondo dei crimini informatici, ma che solleva comunque dilemmi etici sul rapporto uomo-macchina e sulla professione del giornalista, come d'altronde tutte le altre tecnologie: l'agenzia governativa cinese Xinhua, nel 2018, ha presentato al quinto World Internet Congress mondiale il primo Anchor Man digitale.

La *News Distribution* è la fase finale del processo giornalistico. Capire che caratteristiche ha l'audience del nostro giornale, su che piattaforme caricarlo, come presentarlo e incentivarne la diffusione sono tutti aspetti fondamentali per far sì che i nostri lettori abbiano più probabilità di leggere, ascoltare o vedere la notizia in questione. La *News Distribution* è strettamente correlata anche alle modalità di sostentamento delle redazioni. In ogni caso, sia se che si decida di piegare la scelta delle news sui propri lettori, sia che si scelga di vestire i pochi panni rimasti dei *gatekeepers* e di provare a "educare" la propria audience, le tecnologie basate sull'intelligenza artificiale possono risultare fondamentali per questo ultimo step. Abbiamo suddiviso le tecnologie utilizzate nella *News Distribution* in: alert e personalizzazione di contenuti.

Gli alert e la personalizzazione dei contenuti sono l'essenza della *News Distribution* del ventunesimo secolo. Il primo si differenzia dalla seconda perché attiene più ad una scelta dell'utente il quale, con opportuni comandi e impostazioni, può decidere di essere avvertito tramite apposita segnalazione di cambiamenti su determinate pagine web, tra cui quelle di news.

La personalizzazione, invece, riguarda più ciò che la redazione fa al fine di rendere l'esperienza dell'utente la migliore possibile, sia in termini di usabilità e quindi di design della pagina o dell'app, sia in termini di suggerimenti di notizie. Un'eccessiva personalizzazione rischierebbe di tenere il lettore all'interno della propria bolla informativa, rafforzandone i confini. Il dilemma allora si riduce a quale fine dovrebbe perseguire una redazione: utilità sociale e servizio democratico per i cittadini, promuovendo la messa in discussione delle opinioni e la messa a disposizione di diversi punti di vista, o perseguimento del profitto e accentuazione del ruolo dell'informazione come ambito consumeristico come gli altri.

Abbiamo approfondito questo aspetto in un paragrafo dedicato nell'ultimo capitolo.

In ogni caso, una personalizzazione progettata sulla base delle esigenze del singolo che promuova comunque una pluralità di opinioni è possibile, come l'esempio riportato di Voitto, l'assistente alle notizie che propone alternative.

Capitolo 3 Legislazione e questioni etiche

L'etica aiuta l'uomo a concettualizzare la differenza tra ciò che comunemente intendiamo per bene e ciò che comunemente intendiamo per male. Ciò che si vuole o si dovrebbe essere non riguarda solo l'ambito filosofico, ma anche quello giuridico-normativo e professionale. L'informazione oggi usufruisce di più canali rispetto a prima e ci si è quindi posti il problema, nel momento storico di questo cambiamento, del regime giuridico che riguardava i nuovi mezzi di informazione: giornali online, blog, forum, social network come dovevano essere identificati a livello normativo?

Ora tale questione si è posta per le intelligenze artificiali in generale e per il loro utilizzo negli ambiti professionali. Ci si pone il problema non solo di come esse debbano essere identificate giuridicamente, ma anche quali limitazioni apporre al loro operato e come disciplinare gli ambiti professionali per tutelare i lavoratori in carne ed ossa. Le questioni che il progresso tecnologico porta con sé, oltre a quelle giuridiche in senso stretto, sono molte e riguardano numerose dinamiche: trasparenza, discrezionalità ed ecologia sono solo alcune delle tematiche del dibattito. L'etica nel giornalismo è stata più volte concettualizzata ed esistono anche codici etici per l'utilizzo delle intelligenze artificiali in generale. Manca però un'unione delle due, cioè un codice etico per l'utilizzo delle intelligenze artificiali nel giornalismo. Nessuna delle redazioni che ha implementato o sperimentato intelligenze artificiali ha aggiunto qualcosa ai propri codici etici. Ci siamo quindi limitati a riportare qui quelli che in letteratura scientifica vengono denominati *ethical framework* e a fornire successivamente una nostra analisi di questo argomento.

Il concetto fondamentale da tenere bene a mente quando si parla di tecnologie a livello etico è *human in the loop*. Nonostante la rappresentazione mediatica che ne viene fatta, gli umani hanno ancora un ruolo preponderante nell'utilizzo delle intelligenze artificiali. Talvolta, sono stati utilizzati degli algoritmi per prendere delle decisioni ma il risultato a cui si è arrivati, sia concettualmente che giuridicamente, è che l'umano deve avere sempre l'ultima parola in merito.

Una concettualizzazione che analizza il *workflow* giornalistico come presentato finora definisce così i valori etici dell'utilizzo delle IA nel giornalismo. Per il *News Gathering*: valutare periodicamente l'input di dati e creare un processo *ex ante* che valuti tutto il procedimento e gli output. Per la *News Production*: valutare costantemente l'ambito nel quale si producono contenuti in maniera totalmente automatica, al fine di evitare ad esempio che essi contravvengano a leggi appena entrate in vigore. Per la *News Distribution*: monitorare l'utilizzo di dati di terze parti e rendere chiaro ai lettori quale parti sono state create da tecnologie e quali invece sono frutto del lavoro umano.

La categorizzazione di questo capitolo, il cui riflesso è stata la suddivisione in paragrafi, ha previsto analisi specifiche sulla legislazione italiana e come essa vada ad incidere sull'utilizzo delle intelligenze artificiali nel giornalismo, sulla trasparenza dei processi, sulla discrezionalità del loro utilizzo e un approfondimento sull'impatto ambientale delle intelligenze artificiali.

Dal momento che mancano codici etici e regolamenti sull'utilizzo delle intelligenze artificiali, abbiamo riportato i punti fondamentali dei riferimenti legislativi in materia di intelligenze artificiali e di giornalismo come due ambiti distinti. Il punto di riferimento a livello di IA in generale è la proposta coordinata di Regolamento sulle intelligenze artificiali delle istituzioni europee.

In questa proposta di Regolamento, redatta dal Parlamento Europeo e dal Consiglio, si stabiliscono regole armonizzate sull'intelligenza artificiale e la modifica di alcuni atti legislativi dell'Unione. Questa proposta del 2021 prende spunto da un atto dell'anno precedente, il Libro bianco sull'intelligenza artificiale che prevedeva obiettivi specifici: sistemi sicuri e che rispettino le normative; assicurare la certezza del diritto per facilitare gli investimenti; migliorare la governance e l'applicazione effettiva della normativa esistente in materia di diritti fondamentali e requisiti di sicurezza applicabili ai sistemi di intelligenza artificiale; facilitare lo sviluppo di un mercato unico per applicazioni di IA lecite, sicure e affidabili nonché prevenire la frammentazione del mercato. L'obiettivo generale è quindi quello di bilanciare i rischi e le opportunità in tutti gli ambiti di applicazione delle IA e favorire un approccio trasversale per l'applicazione dei requisiti minimi di sicurezza. La disciplina europea appena menzionata prevede un'identificazione in base al rischio delle intelligenze artificiali: una piramide ascendente sino alle tecnologie ritenute inaccettabili. Individuando quindi profili diversi in base alle problematiche che determinate tecnologie possono provocare si prevedono limitazioni e strumenti di prevenzione ad hoc. Tra le tecnologie considerate ad alto rischio che vengono utilizzate nelle redazioni giornalistiche vi è sicuramente il riconoscimento facciale.

Prima di affrontare la questione della trasparenza è stato necessario capire, per concludere l'analisi legislativa del regime giuridico che afferisce alle intelligenze artificiali, individuare chi è responsabile dell'operato di un'IA, con la conseguente individuazione anche di chi è giuridicamente opportuno incolpare per un output illecito. Nel momento in cui si verificano condotte illecite, chi ne è responsabile? La risposta, alla luce delle analisi sin qui compiute ci porterebbe a rispondere che la persona responsabile di tali condotte, nel caso ad esempio di articoli scritti da un'IA, sia il giornalista che ha "collaborato" con essa. Ci sono casi in cui però gli articoli sono completamente scritti da un'intelligenza artificiale e anche nel caso di collaborazione tra umano e software, non risulterebbe giusto attribuire colpe al giornalista quando magari sono state causate da un settaggio sbagliato. Nello specifico, però, risulta difficile associare ogni piccolo step di un processo ad una persona fisica. La giurisprudenza cosa ci dice in merito? Il testo di riferimento in questo senso è la Relazione recante raccomandazioni alla Commissione su un regime di responsabilità civile per l'intelligenza artificiale. Tale relazione prevede che debbano essere adeguati e coordinati i regimi di responsabilità al fine di garantire sempre il risarcimento alle persone lese e riconosce che l'utilizzo delle intelligenze artificiali possa far scaturire, in maniera diretta o indiretta, "danni e pregiudizi". Nella pratica, è molto difficile ricondurre specifiche azioni dei sistemi di IA a uno specifico input umano. È tuttavia possibile considerare responsabili le varie persone nella "catena del valore" che "creano il sistema di IA, ne eseguono la manutenzione o ne controllano i rischi associati".

Unica lacuna che dovrebbe essere colmata, quindi, sono i casi in cui è irrintracciabile e insolubile determinare un terzo che ha interferito con il processo. Viene invece individuata per tutti gli altri casi la figura dell'operatore. Tutto ciò è ovviamente collegato al concetto di trasparenza, senza il quale non si potrebbero ritracciare le persone nella "catena del valore".

La trasparenza è generalmente considerata un mezzo per vedere la realtà, le motivazioni che hanno portato le persone a compiere una determinata azione e determinare le loro responsabilità di fronte al proprio pubblico, legittimandone la fiducia. La trasparenza non è solo un discorso che riguarda l'*accountability* o la fiducia nelle istituzioni, nelle aziende o nelle redazioni, ma potrebbe ledere il diritto alla proprietà intellettuale e il segreto industriale. La conoscibilità dei dati è forse il più importante degli obiettivi in termini di bilanciamento di diversi diritti. I dataset che vengono utilizzati per far funzionare al meglio un'intelligenza artificiale sono spesso "prelevati" direttamente dal web. Negli esempi presentati nel precedente capitolo abbiamo più volte sottolineato come prelevare dati direttamente dal web sia una delle opzioni più gettonate poiché permette di attingere ad una sorta di dataset infinito e in continua espansione, senza costi di raccolta e selezione. Soprattutto per settori come l'oggetto del nostro studio è intuitivo comprendere come alcune tecnologie non possano prescindere dall'essere costantemente aggiornate. Va considerato comunque il rischio di attacchi informatici e la vulnerabilità per i virus. Se tutti attingono dallo stesso dataset, poi, daranno tutti output molto simili tra loro. La differenziazione degli input e una loro raccolta filtrata e selezionata avrebbe non pochi vantaggi in termini di originalità e eticità del software. Sarebbe difficile, però, bilanciare un processo di questo tipo con l'apprendimento e l'aggiornamento costante. La soluzione comporterebbe costi maggiori, nel senso che oltre ad essere necessaria la supervisione umana, dovrebbe aumentare anche la frequenza di questa supervisione, rendendola costante. La soluzione potrebbe quindi essere una via di mezzo, cioè una supervisione che non sia costante ma abbastanza frequente. La giurisprudenza europea ci dice però che questo non basta e che debbano essere comunque previsti programmi appositi di gestione dei rischi *ex post*.

Per quanto riguarda invece il processo in sé e per sé, la conoscibilità di esso è fondamentale. Ci sono delle tecnologie, strutturate con sistemi *black box*, che non permettono di conoscere alcune parti di codice, come se esso fosse parte intrinseca della tecnologia stessa. L'opacità intrinseca delle tecnologie ha dato vita ad un nuovo vero e proprio campo di ricerca, conosciuto come *explainable Artificial Intelligence (XAI)*. La trasparenza dell'addestramento comporta problematiche simili e deve essere il più possibile conoscibile per lo stesso ordine di motivi. Gli si è dedicato in questa sede un sotto-sottoparagrafo a parte poiché negli ultimi anni l'addestramento è risultato lo step fondamentale sia in termini di costi che in termini di ottimizzazione dei software. Un aspetto che viene spesso trascurato quando si parla di intelligenze artificiali è l'impatto ambientale che esse hanno. Nel parlare di intelligenze artificiali e impatto ambientale, il termine più appropriato che identifica il rapporto tra questi due concetti è probabilmente "paradosso", poiché oltre ad inquinare le intelligenze artificiali contribuiscono a fornirci più dati possibili sul cambiamento climatico e sull'inquinamento. Sistemi automatizzati che non vanno a incidere sull'addestramento in sé ma sugli effetti negativi di esso possono essere utilizzati riducendo al minimo gli effetti indesiderati.

Capitolo 4 Il cambiamento della professione giornalistica e la sua percezione

In questo capitolo abbiamo analizzato, sulla scia delle questioni etiche affrontate in quello precedente, i fondamenti valoriali e legislativi della professione, andando poi ad analizzare le novità e l'impatto che le intelligenze artificiali hanno avuto sul settore, sia dal punto di vista degli input, e quindi del ruolo del giornalista, che dal punto di vista degli output, e quindi della qualità dei contenuti. Nonostante la deontologia fosse una questione propriamente etica, riguardante i principi e le regole, si è preferito discuterne in questo capitolo al fine di capire come essi si debbano necessariamente adattare ad un lavoro giornalistico che sempre di più si trasformerà. I diritti e i doveri di un giornalista però, nonostante si debbano necessariamente integrare con i cambiamenti, rimangono sempre quelli dettati dalle norme del XX secolo. Ci riferiamo alla libertà di informazione e di critica, che sono diritti insopprimibili, all'obbligo inderogabile alla verità sostanziale dei fatti osservati, all'obbligo di rettifica delle notizie che risultino inesatte, alla riparazione di eventuali errori, al rispetto del segreto professionale sulla fonte delle notizie (quando ciò sia richiesto dal carattere fiduciario di esse) e alla promozione della fiducia del pubblico. La verità sostanziale dei fatti, nonostante sia inderogabile e democraticamente presunta, ha subito una distorsione enorme negli ultimi decenni. Dalla fine degli anni Novanta ha iniziato a diventare sempre più di uso comune il termine *post-verità*. Questa distorsione riguarda la diffusione esponenziale di *fake news*, problema divenuto ormai di importanza centrale per il mondo dell'informazione.

Risulta opportuno, per i cittadini in generale, oltre che per i giornalisti stessi, avere gli strumenti adeguati a fronteggiare questo tipo di dinamiche che possono ledere la valenza informativa delle notizie e il ruolo dell'informazione all'interno di una società democratica. Le notizie false sono sempre esistite ma mai come oggi vengono diffuse in lungo e in largo, spesso con secondi fini. L'intento della risonanza, e quindi di un ritorno in termini economici o di influenza dell'opinione pubblica, ha ora più che mai strada libera vista la dimensione e la pervasività della rete. Nell'analizzare l'approccio psicologico alle *fake news*, abbiamo capito come spesso, senza rendercene quasi conto, agiamo in una determinata maniera senza riflettere troppo su cosa stiamo facendo. Questo meccanismo appena elencati porta alcune volte a degli errori, denominati *bias*, cioè errori che commettiamo per non sovraccaricare il nostro cervello. Uno dei meccanismi che in questi casi agisce da filtro è il *confirmation bias*, per il quale tendiamo a credere a quello di cui già siamo convinti. Il nostro cervello ritiene più "comodo" fare così piuttosto che esplorare altre alternative e mettere in discussione le proprie credenze e i propri pregiudizi.

Tendiamo inoltre a pensare che gli altri siano più vulnerabili di noi, più soggetti ad influenze e condizionamenti esterni, dando fiducia alla nostra percezione ed essendo convinti di essere poco influenzati ed influenzabili dall'ambiente circostante. Ci sono diversi studi che constatano come, oltre a queste caratteristiche che ci portano ad essere predisposti e vulnerabili in quanto umani all'accettazione di *fake news*, ci siano alcune caratteristiche individuali che predispongono di più o di meno all'accettazione di esse. Negli anni della pandemia tutto ciò è accaduto soprattutto per informazioni di tipo scientifico.

La scienza produce continuamente materiale volto ad informare la popolazione su progressi e attività scientifiche e di ricerca spesso correlati a scoperte sensazionali o disastrose. I media spesso diffondono nel corso della giornata informazioni scientifiche che possono enfatizzare o meno le dimensioni del rischio e della paura. Tutti siamo alla ricerca di rassicurazioni da parte della scienza. È necessario sottolineare che nonostante ci fossero nicchie di popolazione che non credevano nell'efficacia dei vaccini e nel contributo della scienza per fronteggiare la pandemia, la maggior parte della popolazione italiana si è comunque vaccinata, ridando lustro non solo al contributo che la scienza può dare alla società, ma anche all'informazione scientifica di per sé. Si è posto il tema di quanto e come diffondere il sapere scientifico da parte delle istituzioni, ristabilendo quel rapporto trasversale con la comunità di cittadini.

In questo contesto può l'intelligenza artificiale aiutare il cittadino a non essere fuorviato dalle *fake news* e riguadagnare la sua fiducia?

Essa può analizzare titoli, contenuti, commenti e fonti ma non può rilevare l'intenzionalità di una notizia falsa. Abbiamo visto però che sia a livello giuridico sia a livello etico, fa poca differenza se una notizia è falsa volutamente o accidentalmente: il rischio per la democrazia rimane sempre lo stesso. Le pene, infatti, sono le stesse sia per chi diffonde accidentalmente notizie false che per chi le crea intenzionalmente. L'articolo 656 del nostro Codice penale prevede l'arresto fino a tre mesi o l'ammenda fino a 309 € per chi diffonde notizie false, esagerate o tendenziose, "atte a turbare l'ordine pubblico", con pene più gravi in caso di disfattismo politico, aggio, attività antinazionale di un cittadino all'estero, procurato allarme presso le autorità, abuso della credulità popolare o diffamazione. Inoltre, il codice non distingue tra persone che creano notizie false o che diffondono tali notizie. Il Codice di condotta sulla disinformazione del 2018 ha riunito per la prima volta gli operatori del settore, secondo un approccio collaborativo, a livello mondiale, al fine di impegnarsi a contrastare la disinformazione. Al centro della strategia dell'UE contro la disinformazione, il Codice si è dimostrato uno strumento efficace per limitare la diffusione della disinformazione online, anche durante i periodi elettorali e per rispondere rapidamente alle crisi, come la pandemia di coronavirus e la guerra in Ucraina.

Abbiamo visto come il ruolo che il giornalista dovrebbe avere, alla luce di quanto detto sui principi deontologici, è anche quello di *fact-checker*. Il rilevamento delle *fake news* in generale però, a prescindere dall'utilizzo di software automatizzati, risulta essere un compito intrinsecamente difficile in quanto le notizie false sono spesso progettate in modo da sembrare vere e in modo da diventare virali.

Il filone di ricerca che ha affrontato tale problema a livello sia teorico-concettuale che pratico, parte dal presupposto che le *fake news* abbiano caratteristiche diverse dalle notizie vere e che queste caratteristiche siano rilevabili, anche tramite le intelligenze artificiali. Gli approcci generali, in merito alle modalità che si utilizzano a priori per scovare le notizie, riguardano la decisione su chi dovrà svolgere tale compito. Esse si dividono in: *Expert-oriented*, *Crowd-sourcing* e *Computational*.

I software che analizzano le notizie sono molto simili a quelli di ricerca di notizie tramite immagini o parole analizzati nel secondo capitolo. La differenza sta nel fatto che, invece di ricercare contenuti per keywords o immagini simili, analizzano le notizie sulla base di caratteristiche specifiche.

Queste caratteristiche possono essere suddivise in tre macroaree di analisi: chi ha pubblicato il contenuto, il contenuto e a chi è destinato il contenuto.

Gli indicatori quantitativi e quelli temporali sono anch'essi parametri utili e spesso utilizzati per la *fake news detection*. È ovviamente importante, in tutti questi processi, il dataset che si utilizza. L'intelligenza artificiale ha quindi il vantaggio di valutare sulla base non della notizia in sé, ma sulla base delle caratteristiche che contraddistinguono tale notizia come falsa o come vera. Ovviamente esistono diversi tipi di software di analisi per il *fact-checking*. Ci sono quelli che utilizzano solo dataset di esempi, altri che vengono impostati sulla base di parametri specifici, altri ancora che danno un peso maggiore ad una macroarea piuttosto che ad un'altra. Per valutare le prestazioni degli algoritmi per il rilevamento delle *fake news*, sono state utilizzate diverse metriche di valutazione. L'accuratezza di diversi software è stata più volte testata e i risultati generali danno prova della loro indiscutibile validità.

Strumenti di questo genere vengono utilizzati anche per scovare i *deepfakes*. Sono molto simili ai software video analizzati nel secondo capitolo e rintracciano i video falsi.

Tutto questo ci porta alla necessità di analizzare come il ruolo del giornalista sia cambiato non solo nel passaggio dal ruolo di *gatekeepers* a quello di *fact-checker*, ma anche nell'incombenza di imparare nuove skills digitali che non riguardano più solamente la capacità di curare un sito, caricare un articolo sul web o saper distribuire un contenuto in diverse forme e su più piattaforme. Il giornalista ha dovuto ridimensionare le proprie skills in funzione di un panorama mediatico che ha visto crescere esponenzialmente notizie false. Il giornalista è passato nell'era digitale dal decidere lui quali notizie dovessero essere al centro dei dibattiti a subire tali decisioni e ritrovarsi quindi solo a "filtrare" le notizie più discusse, fornendo la propria opinione o etichettando una notizia come vera o falsa.

Lo snodo fondamentale per comprendere il ruolo che ancora possono avere i giornalisti è la comprensione del fatto che, alla luce di quanto detto sulla psicologia delle *fake news*, non tutti abbiamo gli strumenti necessari per capire la realtà e cercare di trasmetterla il più oggettivamente possibile agli altri. Il ruolo del giornalista, quindi, potrebbe e dovrebbe essere ancora quello di incentivare un dibattito polarizzato, comprendente al suo interno opinioni diverse e vari spunti di riflessione.

Altro nodo cruciale è sicuramente la sostenibilità economica di un tipo di informazione certificata e professionale. Se la decisione degli utenti di informarsi tramite il web è spesso dettata dal portafoglio, bisogna necessariamente trovare un modo per rendere più accessibile questo tipo di informazione. Ultima ma non meno importante sfida è riguadagnare la fiducia dei lettori per questo tipo di informazione.

Recentemente Google ha “intervistato” Chat-GPT per un colloquio di lavoro come ingegnere. L’intelligenza artificiale più discussa del momento ha raggiunto, sulla base delle risposte, il punteggio necessario all’assunzione per una posizione di ingegneria di livello tre. La capacità della macchina di fornire una risposta concisa e altamente affidabile potrebbe essere utilizzata da Google per migliorare i propri motori di ricerca.

Il tutto ovviamente pone problemi come l’utilità che ingegneri dello stesso livello potranno avere alla luce di quanto appena detto. Tutto ciò è valido anche per il mondo del giornalismo?

In una ricerca basata su dodici studi sperimentali con un totale di più di 4000 partecipanti, si nota come i testi siano percepiti allo stesso modo per quanto riguarda la credibilità. Sulla percezione della qualità gli umani hanno un leggerissimo vantaggio mentre per quanto riguarda la leggibilità il divario risulta più evidente. La comparazione dei dodici studi ci suggerisce, inoltre, che la percezione dei parametri cambia quando ai partecipanti viene detto che la notizia è stata scritta da un umano, a vantaggio di quest’ultimo.

Le IA sono quindi allo stesso livello di scrittura degli umani? La risposta è indicativamente sì. Consultando la letteratura scientifica attuale i risultati indicano nuove prove a favore dell’avversione per l’IA, ricordando però che la ricerca passata ha mostrato casi sia di apprezzamento dell’IA sia di avversione. Inoltre, questa avversione riguarda più i contenuti definibili come creativi, constatando quanto detto in questa sede sull’irraggiungibilità di alcune capacità umane da parte delle macchine, e si attenua quando il contenuto è stato editato da un giornalista in “collaborazione” con un’intelligenza artificiale.

Quali sono, invece, gli atteggiamenti, le percezioni e i comportamenti del lettore di oggi in merito a servizi di notizie personalizzati tramite IA?

Gli effetti riscontrati sono quasi tutti positivi. Gli utenti di portali personalizzati sono più attenti alle fonti e sono più propensi al consumo di notizie offline. I contenuti personalizzati hanno un effetto positivo sugli utenti anche in merito alla rilevanza che danno alle notizie, al coinvolgimento e alla propensione all’interazione. Gli utenti che usufruiscono di contenuti personalizzati hanno una soddisfazione maggiore. I contenuti personalizzati tramite intelligenza artificiale giovano anche alle redazioni in termini di fidelizzazione e abbonamenti.

I mass media pubblici sono inoltre tenuti a progettare algoritmi che rispettino i valori fondamentali del giornalismo, garantendo così la promozione della diversità dei contenuti, anche se gli utenti decidono di personalizzare la selezione dei contenuti. La personalizzazione dei contenuti rappresenta una minaccia per la diversità e la pluralità dell’informazione ma, come abbiamo visto, ci sono alternative alla personalizzazione univoca.