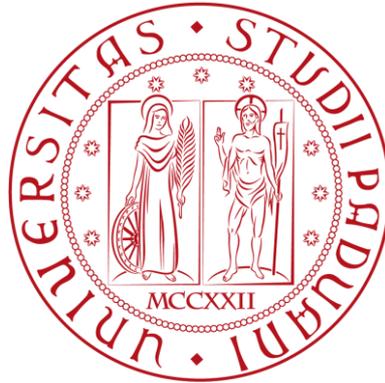


UNIVERSITÀ DEGLI STUDI DI PADOVA

DIPARTIMENTO DI DIRITTO PRIVATO E CRITICA DEL DIRITTO



CORSO DI LAUREA IN
CONSULENTE DEL LAVORO
A.A. 2021-2022

I MODERATORI DI CONTENUTI E IL MITO DELL'INFALLIBILITÀ
DELL'ALGORITMO

Relatrice: Prof.ssa Mingardo Letizia

Laureanda: Frison Sofia

Matricola: 1193664

Indice

Indice	2
Introduzione.....	4
1 I moderatori di contenuti	9
1.1 L'attività in generale	9
1.2 Assunzione e contratto	11
1.3 Formazione e aggiornamento.....	13
1.4 Turnazione e orario di lavoro.....	14
1.5 Retribuzione.....	18
1.6 Punteggi di precisione, licenziamenti e dimissioni.....	19
2 Il lavoro nell'era delle piattaforme digitali	21
2.1 Rivoluzione digitale	21
2.2 Lavoratori digitali	23
2.3 Invisibilità	25
2.4 Accordi di riservatezza	28
2.5 Addestrare l'algoritmo	30
3 Censurati e censura(tori)	32
3.1 Realtà datificate	32
3.2 Personalizzazione dei contenuti.....	34
3.3 Realtà differenti	35
3.4 La complicità dei moderatori	37
3.5 Le ricadute psicologiche della moderazione e i rimedi adottati	40
3.6 Missione.....	42
4 Conclusioni.....	43
4.1 Un caso attuale.....	43
4.2 Prospettive future	45

5	Bibliografia.....	48
---	-------------------	----

Introduzione

La diffusione di internet, e dei dispositivi che ne permettono l'accesso, ha comportato la necessità di una regolamentazione dei contenuti pubblicati online dagli utenti, onde evitare la condivisione di materiale inappropriato o illegale, a discapito delle persone più sensibili ovvero al fine di commettere crimini o azioni inopportune, i social network e le altre piattaforme digitali globali giocano un ruolo fondamentale nello scenario delineato, richiedendo una massiccia e penetrante regolamentazione. In particolare, piattaforme social, motori di ricerca ed applicazioni, i quali contano milioni (o miliardi) di persone collegate, richiedendo importanti interventi di regolamentazione dei contenuti pubblicati, necessità dovuta alla mancanza di barriere all'entrata, chiunque può condividere qualsiasi cosa, ma soprattutto, dalla capacità dell'algoritmo che governa la piattaforma di rendere virale qualsiasi contenuto, non riuscendo a distinguere fra un'immagine innocua ovvero una rappresentante violenza o altre barbarie. Una volta ricevute interazioni da parte di un numero sempre maggiore di utenti, interpretate dall'algoritmo come segnali di gradimento, un contenuto, a prescindere che esso sia appropriato o meno, sarà reso virale, raggiungendo milioni di persone nel paese da cui è partito, diffondendosi poi nel resto del mondo, per diminuire la portata e gravità di questi avvenimenti, le aziende statunitensi proprietarie delle piattaforme hanno fornito gli utenti della possibilità di 'segnalare' i contenuti. Possibilità da utilizzare nel caso vengano condivisi contenuti ritenuti inappropriati, sulla base di una serie di linee guida formulate dalla piattaforma stessa, utili ad orientare i comportamenti degli utenti e la pubblicazione di materiale online, la cosiddetta 'policy' contiene infatti indicazioni riguardanti contenuti condivisibili o proibiti, fra quest'ultimi figurano categorie quali fake news, violenza, pornografia, guerra, la segnalazione inviata è quindi utile per richiederne la rimozione.

Entrano qui in gioco i moderatori, addetti alla revisione ed eventuale rimozione dei contenuti segnalati, la cui sorte è decisa sulla base di quanto stabilito dalla policy, in particolare, i lavoratori compiono una delicata attività di interpretazione delle regole concordemente alle molteplici sfumature di significato e ai diversi contesti presentati dai contenuti, al fine di assumere la decisione più accurata possibile, onde evitare cancellazioni o limitazioni improprie.

La suddetta attività è però resa difficile dai continui cambiamenti delle regole contenute nella policy, sensibili agli avvenimenti esterni alla piattaforma, le quali possono addirittura variare da paese a paese in base a regole e leggi locali, senza contare le varie e delicate sfumature di significato caratterizzanti i contenuti, il tutto complicato dal fatto che molto spesso i moderatori hanno una manciata di secondi per prendere le decisioni corrette. Infatti, avendo milioni di iscritti in tutto il mondo, operativi in qualsiasi momento della giornata, le piattaforme necessitano di una sorveglianza attiva ventiquattro ore, sette giorni su sette, al fine di rispondere a quante più segnalazioni possibili, i moderatori si ritrovano così a dover visionare centinaia (se non migliaia) di contenuti ogni giorno, spendendo pochissimo tempo fra una decisione e l'altra, onde tentare di fermare l'inarrestabile diffusione del materiale segnalato. L'accuratezza nelle decisioni è pregiudicata altresì da una mancanza di preparazione specifica dei moderatori, nonostante un periodo iniziale di formazione ed il continuo aggiornamento riguardante i cambiamenti introdotti nella policy, molto spesso i lavoratori si ritrovano a dover revisionare contenuti a carattere scientifico, politico, medico senza avere alcuna competenza in merito, le decisioni assunte possono poi essere valutate da alcuni supervisori, esponendo i moderatori a possibili penalizzazioni o al licenziamento. L'algoritmo che governa la piattaforma organizza e suddivide i contenuti da revisionare fra i vari moderatori disponibili, programmando ogni momento della giornata lavorativa e sorvegliando sulla corretta esecuzione dell'attività, divenendo così un vero e proprio datore di lavoro, impersonale e distante, costringendo i lavoratori a sostenere ritmi frenetici, turni massacranti e controllando ogni loro azione, rilevando altresì ogni momento di inattività o di pausa non concordati. L'aspetto peggiore di questa professione riguarda, però, la natura dei contenuti revisionati, se buona parte di essi è innocua o noiosa, frutto di segnalazioni errate o in malafede, l'altra metà è composta da foto e video riprovevoli, violenze contro persone o animali, pedopornografia e attentati, sono solo alcune delle barbarie alle quali i moderatori sono sistematicamente esposti giorno dopo giorno, provocando sofferenze psicofisiche indelebili e difficilmente guaribili.

Il contesto lavorativo in cui sono inseriti i moderatori è evidentemente problematico, l'attività di revisione ed interpretazione dei contenuti è considerata troppo complicata per l'intelligenza artificiale, ma nonostante ciò, ai moderatori è richiesto di operare come macchine, mantenendo ritmi frenetici al fine di rimanere al passo con quell'algoritmo che in qualsiasi momento potrebbe rendere virale contenuti pregiudizievoli per la sensibilità degli utenti o infangare per sempre la reputazione della piattaforma.

L'attività di moderazione può essere svolta direttamente dai dipendenti delle aziende proprietarie delle piattaforme ma, in larga parte, è trasferita a società intermediarie di fornitura servizi, aventi sedi in tutti i paesi del mondo, le quali si occupano di assumere lavoratori e adibirli all'attività, moltissimi moderatori provengono da paesi poveri o in via di sviluppo, come le filippine, caratterizzati da un basso costo del lavoro, condizione sfruttata dalle suddette società. I lavoratori assunti da società terze sono definiti 'appaltatori' e costituiscono appunto la maggior parte della forza lavoro, in generale, i vari moderatori nel mondo, non percepiscono più del salario minimo del paese di riferimento, eccezione fatta per i dipendenti diretti delle aziende tecnologiche, i quali godono di stipendi più alti e vari benefit, mentre i contratti stipulati sono generalmente a termine; infatti, nemmeno i migliori moderatori resistono più di un paio d'anni. Nonostante la brevità dell'impiego, non mancano mai moderatori attivi e disponibili, le aziende riescono sempre a mobilitare grandi masse di lavoratori, infatti, non essendo richiesti specifici prerequisiti o particolari competenze, chiunque può accedere a questa professione purché sappia usare un computer, ciò garantisce alle aziende la possibilità di avere costantemente a disposizione lavoratori di 'ricambio' pronti a sostituire chi viene disconnesso, a seguito di licenziamento o altre vicissitudini.

La condizione in cui versano i moderatori è molto simile a quella patita dagli attuali lavoratori digitali, i cosiddetti gig workers, entrambi nati con l'avvento della rivoluzione digitale, al lavoro sulla piattaforma e per la piattaforma, considerati non più di un pezzo di ricambio nella macchina produttiva e vittime di un'organizzazione invisibile del lavoro, deliberatamente nascosti e sfruttati dalle aziende, le quali beneficiano

illegittimamente del loro lavoro. Pochissime persone sono a conoscenza delle dinamiche seguenti una segnalazione, questa confusione non è un caso, è frutto del sottile lavoro di occultamento effettuato dalle aziende proprietarie delle piattaforme, le quali per anni hanno marciato sulla promessa della totale automazione digitale in campo editoriale, la quale avrebbe comportato la sostituzione dei lavoratori con le macchine, promesse sfociate nella creazione di algoritmi imprecisi ed inclini alla commissione di errori, bisognosi dell'intervento umano.

Per assicurarsi la totale discrezione le aziende tecnologiche hanno quindi deciso di esternalizzare l'attività di moderazione, appaltandola a società terze, e costringere i lavoratori a firmare severi accordi di riservatezza al momento della assunzione, i quali comportano un obbligo di non divulgazione riguardante l'attività svolta, il committente e i materiali revisionati, la violazione dell'accordo implica l'immediato licenziamento e il pagamento di multe anche molto salate.

La moderazione umana sarà obbligata almeno fino a quando non sarà creata un'intelligenza artificiale in grado di assumere decisioni accurate, e forse questa prospettiva è più vicina di quanto si immagini: al momento della segnalazione l'utente è tenuto ad indicarne la motivazione (pornografia, spam, fake news, ecc), il moderatore durante la revisione, se deciderà di eliminare il contenuto, dovrà catalogarlo scegliendo una serie di alternative. Tutti i dati immessi sono archiviati dalle piattaforme ed elaborati dall'algoritmo, il quale li utilizzerà per 'imparare', infatti, grazie alle tecniche di machine learning è possibile addestrare la macchina al riconoscimento dei contenuti non ammessi, la quale analizza e confronta le scelte effettuate da utenti e moderatori, valutando in che percentuale determinati contenuti vengono rimossi e procedendo per analogia alla comparazione di questi con materiale simile.

I dati relativi alla pubblicazione e rimozione di contenuti non sono gli unici ad essere archiviati dalle piattaforme digitali, infatti, ogni azione compiuta in rete dagli utenti produce dati idonei a svelare informazioni, gusti e preferenze riconducibili ai suddetti soggetti, utili alla creazione di 'profili' digitali utilizzati dalle aziende al fine di predisporre pubblicità e offerte ad hoc, ed un catalogo di contenuti personalizzati pensati

su misura per ogni utente. La personalizzazione dell'esperienza nella rete comporta il fenomeno delle cosiddette bolle di filtraggio, creando tante realtà diverse quanti sono gli utenti collegati, in quanto, ogni individuo vedrà solo contenuti di suo gradimento, in linea con i propri gusti e visione del mondo, amplificando l'effetto conferma e la polarizzazione delle idee, impedendo l'incontro con il diverso, rendendo accessibili solo le informazioni che si preferiscono sapere e non quelle che si dovrebbero sapere.

Per anni, gli studiosi si sono focalizzati sull'analisi delle bolle automatiche, conseguenti alla personalizzazione di contenuti da parte degli algoritmi di profilazione, tralasciando completamente il lavoro effettuato dai moderatori, responsabili della creazione di bolle manuali per effetto della limitazione o eliminazione dei contenuti vietati dalle policy, i quali, seppur deplorabili, sono pur sempre un frammento della realtà, alla quale molti utenti non possono avere accesso. Ma la necessità di protezione degli utenti è forse utilizzata come escamotage per giustificare quel che potrebbe essere definito un servizio di censura globale, nel quale i moderatori hanno una parte di responsabilità, rendendosi partecipi della creazione di un'utopia, dove ciò che viene considerato sconveniente è fatto scomparire, le piattaforme divengono così realtà parallele, un mondo edulcorato, rappresentando una via di fuga dalle difficoltà e dall'asprezza della vita reale.

I moderatori proteggono gli utenti, ma nessuno protegge i moderatori dai contenuti con i quali entrano in contatto ogni giorno, minando la propria integrità psicofisica in via permanente e ricevendo scarsa assistenza, sviluppando ansia, stress e disturbi post traumatici, molti lavoratori sostengono di riuscire a gestire questo lavoro perché convinti di aver intrapreso una missione, quella di rendere internet, e forse, anche il mondo, un posto migliore.

1 I moderatori di contenuti

1.1 L'attività in generale

I moderatori di contenuti sono addetti alla revisione e alla, eventuale, rimozione dei contenuti pubblicati online dagli utenti di social, motori di ricerca, app e qualsiasi altra piattaforma digitale globale di uso quotidiano governata da un algoritmo, in particolare, svolgono una attività di valutazione circa la conformità dei contenuti condivisi con la policy della piattaforma e procedono alla eliminazione dei suddetti nel caso in cui violino le regole di pubblicazione. Questo processo è innescato da una segnalazione anonima, effettuata da un utente, qualora ritenga che un determinato contenuto violi le politiche di pubblicazione, l'utente, durante questa azione, dovrà altresì indicare la "ragione" della segnalazione in modo da permettere all'algoritmo di suddividere automaticamente i contenuti, per i quali è stata richiesta revisione, all'interno di più categorie cosicché i più "delicati" ed urgenti vengano controllati ed eliminati nel minor tempo possibile

La policy può essere definita come un corpus di norme contenente le regole di pubblicazione della piattaforma volte ad orientare il comportamento degli utenti, ed è principalmente sulla base di questo documento che i moderatori indirizzano la loro attività e le loro scelte, andando ad interpretare quanto indicato per decidere, in relazione al caso concreto, quali contenuti eliminare e quali mantenere online, il tutto è, però, reso difficile dai costanti aggiornamenti delle policy. Come riporta Jacopo Franchi in "Gli obsoleti": "le indicazioni riguardanti i contenuti ammessi, ammessi con riserva e non ammessi cambiano continuamente a seconda dei principali avvenimenti esterni ed interni alla piattaforma stessa"¹, nonostante l'ininterrotto aggiornamento a cui sono sottoposti questi lavoratori, la poca chiarezza con cui sono definite le regole li rende vulnerabili alla commissione di errori che potrebbero comportare il licenziamento.

I contenuti pubblicati in rete e revisionati possono essere di varia natura, alcuni innocui altri, invece, a dir poco inquietanti, materiale pedopornografico, foto e video di suicidi, attentati, decapitazione e incitamento all'odio sono solo alcune delle barbarie con cui questi lavoratori convivono ogni giorno, le quali devono essere eliminate dalla

¹ Cfr. Franchi (2021), p. 68.

piattaforma nel minor tempo possibile onde evitare che un gran numero di utenti visualizzi quei contenuti.

L'algoritmo suddivide automaticamente il lavoro fra i vari moderatori disponibili assegnando ad ognuno una coda di revisione, scandendo i tempi e le modalità dell'attività lavorativa e contemporaneamente vigilando sulla stessa, i moderatori più esperti possono arrivare a visionare migliaia di contenuti al giorno anche grazie ad alcune tecnologie che permettono di scomporre i video in tante piccole immagini che svelano nell'immediato la natura del contenuto. Grazie a questa scomposizione il lavoratore non dovrà riprodurre l'intero video e in brevissimo tempo potrà deciderne la sorte, ovviamente i tempi di rimozione cambiano da caso a caso ed anche in base alle leggi dei vari paesi, le quali possono prevedere tempi diversi per la cancellazione dei contenuti, in Germania, ad esempio, vige una legge che impone la rimozione entro 24 ore dalla segnalazione di video e foto "illeghi"². Momenti di intenso lavoro si alternano a momenti più tranquilli, il ritmo è deciso dall'algoritmo e dalla sua capacità di rendere virale qualsiasi contenuto, dai più innocui ai più violenti, non essendo ancora stato creato un meccanismo di filtraggio automatico che impedisca la pubblicazione di determinati contenuti o riconosca quelli da rimuovere, tutto ciò rende, quindi, ancora indispensabili i lavoratori in carne ed ossa e la moderazione manuale. Il problema centrale ruota proprio intorno al fatto che qualsiasi utente potrebbe pubblicare qualsiasi cosa su qualsiasi piattaforma, non essendoci infatti barriere all'entrata, e che questi contenuti, una volta ricevute interazioni da parte degli altri utenti, potranno essere diffusi su scala mondiale dall'algoritmo, che non conosce la differenza fra contenuti consentiti e vietati, i quali, una volta diventati virali, potranno raggiungere un numero sempre maggiore di utenti.

Questo meccanismo potrà quindi interessare sia un'immagine innocua, sia un'immagine rappresentate violenza, ed è qui che, a seguito di una segnalazione, entra in gioco l'attività di moderazione, la quale ruota attorno a due finalità, la prima riguarda la necessità di evitare che contenuti "disturbanti" raggiungano gli utenti più sensibili, la seconda riguarda la cattiva pubblicità che potrebbe ricavarne l'azienda: se gli utenti

² La legge per la tutela dei diritti sui social network (NetzDG) obbliga i social network con più di due milioni di utenti registrati in Germania ad applicare una rimozione a livello locale dei contenuti "chiaramente illegali" (ad esempio un video o un commento) entro 24 ore dalla notifica.

accanto alle foto delle proprie vacanze vedessero immagini di violenze, sicuramente smetterebbero di collegarsi.

1.2 Assunzione e contratto

Per diventare moderatori di contenuti non sono richieste specifiche competenze o specializzazioni, fatta eccezione per alcune aziende che ricercano personale laureato, essendo una attività manuale e ripetitiva che necessita solamente una basilare conoscenza della rete e dei social, è quindi un impiego che può essere svolto da chiunque cerchi una occupazione, o perché alla prima esperienza lavorativa o perché disoccupato da tempo e in cerca di una attività per sostentarsi. Molti dei moderatori che negli ultimi anni hanno rilasciato dichiarazioni in merito alla loro professione, sostengono di aver trovato questo impiego in maniera quasi “casuale” e di non averlo cercato intenzionalmente, ad esempio perché iscritti a siti di domanda e offerta di lavoro, nei quali, dopo aver compilato il proprio curriculum, l’algoritmo trova automaticamente le offerte lavorative coerenti con le competenze dichiarate, per le quali è poi possibile candidarsi. Va precisato che generalmente le aziende che pubblicano annunci di lavoro di questo tipo non utilizzano il termine “moderatore” ma altre espressioni più “velate” e un po’ per curiosità, un po’ per l’allettante idea di lavorare nel settore tecnologico, o ancora per semplici ragioni di sostentamento, le persone decidono di inviare la propria candidatura a cui può seguire la chiamata da parte dell’azienda per una successiva assunzione.

Come riportato da Jacopo Franchi nel libro citato “nessuno vuole diventare moderatore di contenuti, eppure tanti lo sono già stati: basta avere un account su YouTube per aver rimosso almeno una volta un commento inappropriato di un altro utente...”³

ipoteticamente, tutti gli utenti della rete potrebbero essere idonei a svolgere questo lavoro ma solo alcuni trasformano questa attività in una professione remunerativa.

La presenza di moderatori professionisti non comporta il venir meno degli utenti che volontariamente e gratuitamente ripuliscono la rete da contenuti inappropriati, basti pensare agli amministratori dei gruppi Facebook, i quali hanno la facoltà di limitare o

³ Franchi (2021), p. 37.

eliminare un partecipante o ancora un commento, una foto, un video ritenuti inopportuni, andando ad “alleggerire” il carico di lavoro dei suddetti moderatori, il loro contributo è quindi molto prezioso.

Negli ultimi anni le grandi aziende tecnologiche, proprietarie delle suddette piattaforme, hanno optato per una tendenziale esternalizzazione dell’attività di moderazione, in particolare, collaborano con alcune tra le più grandi società statunitensi specializzate nella fornitura di servizi professionali, come TaskUs, Cognizant e Accenture, i moderatori assunti con queste modalità vengono definiti “appaltatori” e lavorano negli uffici dalle suddette società, in Nord America e in altre parti del mondo, realizzando una divisione globale del lavoro.

I moderatori possono essere altresì assunti a tempo pieno direttamente dalle grandi aziende tecnologiche, ciò che accomuna appaltatori e moderatori assunti in via diretta concerne la tipologia di contratto stipulato, generalmente a termine, vi sono però sostanziali differenze riguardo le condizioni di lavoro e le tutele accordate ai lavoratori assunti in via diretta, che godono di un miglior trattamento sotto tutti i punti di vista. Generalmente vengono preferiti gli appaltatori sia per ragioni di costo, in quanto lavorano per poco più del salario minimo del paese di riferimento, sia per ragioni di flessibilità organizzativa, periodi lavorativi più intensi sono fronteggiati con un numero maggiore di lavoratori che successivamente saranno eliminati; per Google i lavoratori temporanei, appaltatori e fornitori, i lavoratori che i Googler chiamano internamente TVC, costituiscono ora il 54% della forza lavoro dell'azienda⁴.

Facebook in controtendenza rispetto a qualche anno fa, quando molti degli uffici per la moderazione erano siti all’estero, ha deciso di riportare parte di questa attività negli Stati Uniti⁵, aprendo sedi in California, Florida, Arizona e Texas, partendo dalla considerazione per cui è più probabile che moderatori americani dispongano del contesto culturale necessario per valutare i contenuti pubblicati dai loro connazionali, grandi utilizzatori del social⁶, lo stesso è stato fatto da Google e YouTube.

⁴ Cfr. Newton (16 dicembre 2019).

⁵ Nel 2017 Facebook, dopo molti problemi riguardanti la sicurezza e la sorveglianza della piattaforma, ha deciso di reclutare migliaia di nuovi moderatori (perlopiù appaltatori) provenienti soprattutto dal Nord America, i quali, avendo una maggiore conoscenza dello slang caratterizzante i contenuti statunitensi, avrebbero potuto adottare decisioni più accurate.

⁶ Cfr. Newton (25 febbraio 2019).

Ogni contratto è accompagnato da un severo accordo di riservatezza, firmato al momento dell'assunzione, i moderatori non possono condividere con il mondo esterno i dettagli del proprio lavoro, non possono parlare del materiale con cui entrano in contatto ogni giorno, né delle reali modalità con cui i contenuti vengono eliminati dalla rete ed è soprattutto per questo motivo che le condizioni in cui versano sono rimaste così a lungo celate. L'accordo di riservatezza viene fatto firmare, oltretutto per la questione relativa alla segretezza che avvolge le modalità di rimozione dei contenuti, altresì per vietare ai moderatori di divulgare le informazioni private con cui entrano in contatto, ed infine per evitare ritorsioni personali di utenti scontenti delle decisioni prese circa la sorte dei contenuti da loro pubblicati e rimossi per qualsivoglia motivazione, la violazione di questo accordo comporta pesanti conseguenze.

1.3 Formazione e aggiornamento

Tutti i neoassunti, durante le prime settimane di lavoro, devono partecipare a corsi di formazione, tenuti dai loro supervisor⁷, che li instruiranno circa l'attività intrapresa, i corsi sono divisi in una parte teorica dove vengono illustrate la policy e le linee guida da seguire ed interpretare, ed una parte pratica dove viene testata la loro capacità di comprensione rispetto ad un caso concreto da moderare. Data la natura dei contenuti, molti lavoratori non riescono a superare questo periodo iniziale, solo chi rimane potrà iniziare a tempo pieno il lavoro, vi è da precisare che, date le continue modifiche apportate alla policy, la formazione, e l'aggiornamento, sono, in realtà, costanti in quanto i moderatori devono essere sempre preparati e flessibili per poter interpretare nel miglior modo possibile i casi concreti, ed evitare errori. La policy varia al variare del contesto sociale, economico e politico mondiale, le regole possono essere modificate, eliminate ed aggiunte anche giornalmente perciò è richiesto ai moderatori un enorme sforzo mnemonico ed anche una certa flessibilità mentale, senza considerare che, molto spesso, sono chiamati a dare un giudizio su argomenti che non conoscono, su cui non hanno alcuna conoscenza (contenuti scientifici, medici ecc), il margine di errore è quindi alto.

⁷ I supervisor sono generalmente moderatori con più esperienza.

In tema di modificazione della policy possiamo fare riferimento al caso della guerra in Siria e dei moderatori di YouTube: prima dello scoppio della guerra, tutti i contenuti che rientravano in questo tema erano in netta contraddizione con la policy della piattaforma ed andavano quindi rimossi, successivamente YouTube (come altre piattaforme) ha deciso di mantenere online foto e video ritraenti il conflitto, lasciando che l'algoritmo li rendesse virali. I contenuti più cruenti, per gli utenti maggiorenni, erano accompagnati da un messaggio di avvertimento circa le violenze rappresentate⁸, mentre erano vietati ai minori; tutto questo per dare più visibilità possibile a ciò che realmente stava accadendo e traghettare l'opinione pubblica verso quella parte del mondo, i moderatori si sono così ritrovati a dover mantenere online ciò che fino al giorno prima avrebbe dovuto essere immediatamente cancellato. Il punto più delicato riguardava l'applicazione delle nuove regole al caso concreto, in quanto per i moderatori non era facile distinguere fra un'immagine violenta collegata alla guerra ed un'altra fine a sé stessa, considerando la molteplicità di situazioni in corso, i vari contesti e le poche e contraddittorie informazioni a disposizione, il lavoro di interpretazione era divenuto quindi molto difficile ed insidioso. Le foto e i video documentati dai cellulari e caricati online erano considerati importanti prove riguardanti i crimini di guerra e la violazione dei diritti umani perpetrati nel paese, preziosi elementi per ricercare giustizia, nel 2017 però, dopo aver ricevuto forti pressioni per la cancellazione dei contenuti estremisti⁹, YouTube, attraverso un sistema di rimozione automatica, ha eliminato circa il 10% dell'archivio che documentava le violenze in Siria¹⁰

1.4 Turnazione e orario di lavoro

“La necessità di suddividere l'attività dei moderatori su più fasce orarie nel corso della giornata si spiega apparentemente con l'attività ininterrotta della macchina, la quale può rendere virale un contenuto problematico in qualsiasi ora del giorno e della notte[...]

⁸ Cfr. Adrian Chen (23 ottobre 2014).

⁹ Giornalisti ed altri esperti del settore hanno denunciato la massiccia presenza su YouTube di contenuti rappresentati violenze, soprusi ed incitamento all'odio, molti dei quali sono stati trovati sulla piattaforma YouTube kids, creata per offrire un catalogo di video sicuri per i bambini.

¹⁰ Cfr. Hadi Al Khatib (23 ottobre 2019).

non essendoci limite all'orario in cui un utente può connettersi per condividere un contenuto o inviare una segnalazione, la rapidità di risposta a queste ultime deve essere garantita in qualsiasi paese e in qualsiasi momento dell'anno"¹¹.

Partendo dalla considerazione che una piattaforma digitale globale, come Facebook, Twitter o Tik Tok, può arrivare a contare milioni di iscritti in tutto il mondo, è indubbia la necessità di un servizio di moderazione attivo ventiquattro ore su ventiquattro, sette giorni su sette, che possa, nel minor tempo possibile, rispondere alle innumerevoli segnalazioni inviate in qualsiasi momento della giornata e provenienti da ogni paese del mondo. La moderazione su di una singola piattaforma può contare su moltissimi lavoratori (appaltatori e non) sparsi in tutto il mondo, i quali potrebbero trovarsi a revisionare un contenuto proveniente dall'altra parte del globo, in quanto, per accelerare i tempi, il materiale è suddiviso fra i moderatori disponibili, a prescindere da dove questi sono localizzati e da quanti chilometri li dividono dall'utente oggetto di segnalazione e controllo. Alcuni moderatori Americani ed Europei, al lavoro per Twitter, YouTube e Facebook, intervistati dal 'Post' hanno riferito di moderare contenuti principalmente nella loro lingua, i moderatori filippini, al lavoro per le medesime aziende, che parlano solo Tagalog e inglese, hanno invece riferito di dover revisionare contenuti in dieci lingue diverse, e che, in mancanza di un traduttore esperto, sono costretti ad utilizzare strumenti quali Google translate, creando maggiore confusione¹². Negli ultimi anni, complice il grande aumento di utenti provenienti da paesi poveri e in via di sviluppo, come appunto le Filippine o l'India, e nell'ottica del risparmio economico, le grandi aziende hi-tec, tramite società fornitrici, hanno delocalizzato diversi siti di moderazione in queste aree dove, però, i lavoratori sono scarsamente tutelati e non possiedono gli strumenti adeguati a rivendicare un miglior trattamento economico e migliori condizioni lavorative.

Altresì per far fronte al problema delle barriere linguistiche, molti siti di moderazione, in Nord America e Europa, tendono ad assumere immigrati e stranieri, provenienti principalmente dal medio oriente, i quali sono poi assegnati alle code di revisione contenenti materiale di natura terroristica, estremista o di incitamento all'odio, molto

¹¹ Franchi (2021), p. 81.

¹² Cfr. Dwoskin, Whalen, Cabato, (25 luglio 2019).

frequenti nelle piattaforme, che necessitano di una comprensione della lingua e del contesto, utili alla identificazione della violazione e successiva rimozione.

Durante il turno di lavoro, che può quindi essere collocato in qualsiasi momento del giorno (e della notte), i lavoratori siedono alle proprie scrivanie per sei, sette, otto o nove ore, durante questo tempo hanno diritto ad una pausa pranzo e a vari momenti di relax, necessari data la natura stressante di questa professione, la quale richiede ai lavoratori un enorme sforzo mentale ed emotivo. I movimenti meccanici e ripetitivi, l'enorme attenzione ai dettagli ed alle regole e la natura dei contenuti visualizzati, provocano stress e tensione nei moderatori, per questo le aziende tecnologiche e i loro intermediari assicurano programmi di supporto e benessere, da realizzare internamente ed esternamente al luogo di lavoro, prevedendo sessioni di consulenza psicologica in azienda, attività distensive come lo yoga, benefici sanitari o ancora assistenza terapeutica dopo la fine della attività lavorativa. Tutto ciò non sempre è assicurato nella realtà, l'esempio più eclatante riguarda Facebook, dove i moderatori (specialmente gli appaltatori) vedono il loro tempo cronometrato al minuto, in particolare, hanno diritto ad una pausa pranzo di trenta minuti, due pause da quindici minuti ed una ulteriore pausa definita "tempo benessere" di nove minuti, la quale dovrebbe essere dedicata alla consulenza psicologica, se disponibile, o ad altre attività distensive¹³. Secondo alcune testimonianze, i moderatori sarebbero costretti a registrare tutti i momenti in cui si trovano lontani dalla scrivania, quando sono quindi inattivi, in particolare, Facebook starebbe utilizzando un "sistema di gestione del tempo" volto a monitorare ogni minuto delle loro giornate lavorative, comprese le pause bagno, le quali se ritenute troppo lunghe o frequenti, dovranno essere giustificate, cosa che lede la dignità dei lavoratori¹⁴. È quindi reso molto difficile l'effettivo godimento dei tempi di riposo e dei servizi di consulenza psicologica o delle attività distensive, in quanto, il capillare controllo dell'attività, non permette alcuna flessibilità, andando ad acuire lo stress e la tensione accumulati, senza contare che molto spesso i programmi benessere sono assenti o mal gestiti, non concedendo ai lavoratori nemmeno la possibilità di usufruire di quelle brevi sessioni di terapia.

¹³ Newton (25 febbraio 2019).

¹⁴ Cfr. David Gilbert (9 gennaio 2020).

Gli appaltatori al lavoro per Google e YouTube sono incaricati di revisionare almeno centoventi video in cinque ore (nonostante la società intermediaria neghi di imporre obiettivi di produttività e l'amministratrice delegata di YouTube avesse promesso di ridurre il carico di lavoro a quattro ore), ed hanno a disposizione un'ora per il pranzo (non retribuito) e due ore 'benessere' (retribuite) per rilassarsi o usufruire di una sessione di consulenza psicologica. Nei computer è installato un particolare software di monitoraggio volto a verificare quanto tempo i lavoratori hanno trascorso attivamente a moderare i video al fine di verificare il raggiungimento dell'obiettivo di cinque ore, vi è da precisare che vengono conteggiati solo i minuti di riproduzione effettiva dei video e non anche i tempi legati ad altre attività lavorative (le riunioni, il controllo delle e-mail ecc). Secondo alcune delle testimonianze di questi appaltatori, i tempi di riposo non sempre vengono garantiti, in particolare, durante periodi di intensa attività, viene chiesto ai lavoratori di continuare a moderare anche durante le pause, per riuscire a fronteggiare il numero di segnalazioni pervenute, senza contare che spesso non riescono a godere del tempo di riposo perché trascorso cercando di raggiungere l'obiettivo di moderazione attiva di cinque ore¹⁵.

Ai dipendenti diretti delle aziende hi-tec vengono garantite settimane di ferie pagate nonché la possibilità di richiedere congedi retribuiti per assentarsi dal lavoro, i quali generalmente sono di natura medica e vengono richiesti per poter processare i traumi derivanti dall'attività svolta, molto spesso durano mesi, tempo durante il quale i lavoratori potranno affrontare quanto visto, soprattutto grazie all'aiuto di terapisti e psicologici specializzati, e continuare ad essere pagati.

Tutte queste facoltà non sono riconosciute agli appaltatori che anzi, molto spesso, non riuscendo nemmeno ad accedere alla consulenza psicologica in azienda, sono costretti a richiedere congedi medici non retribuiti per processare i propri traumi e altresì vedono spesso negate le ferie in periodi di intensa attività, nonostante le aziende sostengano che il miglioramento continuo dei programmi benessere e la tutela psicofisica dei lavoratori sia il loro obiettivo primario.

¹⁵ Newton (16 dicembre 2019).

1.5 Retribuzione

La retribuzione di un moderatore è generalmente poco superiore allo stipendio minimo della nazione di riferimento, il cui ammontare può variare considerevolmente da paese a paese in base al costo della vita, per fare alcuni esempi concreti, gli appaltatori di Phoenix, Arizona, al lavoro su Facebook ed assunti tramite la società intermediaria Cognizant, guadagnano circa 15 \$ all'ora, mentre il compenso annuale si aggira intorno ai 28.000 \$¹⁶. Numeri completamente diversi da quelli previsti per i dipendenti a tempo pieno dell'azienda tecnologica, i quali arrivano a guadagnare fino a 240.000 \$ in un anno fra stipendio, bonus e sovvenzioni in azioni¹⁷, Facebook offre altresì innumerevoli benefit, tra cui quattro mesi di congedo retribuito per neo-madri e padri, cure odontoiatriche¹⁸ e moltissimi altri benefici, sconosciuti ai dipendenti delle società fornitrici del servizio. Diversa situazione quella degli appaltatori di Austin, Texas, assunti dalla società intermediaria Accenture e al lavoro su Google e YouTube, che arrivano a guadagnare 18,5 \$ l'ora e circa 37.000 in un anno, mentre un moderatore assunto in via diretta da Google guadagna quasi il triplo, dopo essere stato promosso, senza contare i bonus e altri svariati benefici¹⁹; un appaltatore irlandese percepisce circa 23.000 \$ annuali²⁰.

Ad ogni modo lo stipendio dei moderatori occidentali risulta molto più alto di quello dei loro collaboratori asiatici o africani a parità di mansioni, di contenuti revisionati ed errori commessi, in particolare, un appaltatore filippino, assunto da Accenture e al lavoro su Twitter, arriva a malapena a 450 \$ al mese (circa 2,60 \$ all'ora)²¹, mentre altri appaltatori asiatici possono arrivare a guadagnare un compenso maggiore, fino a 850 \$. Infine, vi è da precisare che, attualmente, oltre ai moderatori professionisti, assunti con regolari contratti direttamente dalle aziende tecnologiche o dalle società intermediarie di fornitura servizi, esiste un'ulteriore categoria di soggetti ai quali vengono trasferite alcune importanti componenti del processo di moderazione dei contenuti, questi ultimi

¹⁶ Newton (25 febbraio 2019).

¹⁷ Cfr. Rob Price (17 aprile 2018).

¹⁸ Cfr. Aine Caine (7 dicembre 2017).

¹⁹ Newton (16 dicembre 2019).

²⁰ Cfr. Olivia Solon (23 maggio 2017).

²¹ Dwoskin, Whalen, Cabato, (25 luglio 2019).

sono iscritti a piattaforme di gig working, sono attivabili al bisogno e vengono pagati a cottimo in base alle azioni effettuate.

Le suddette piattaforme possono essere definite “satelliti” rispetto alle principali sulle quali si svolge la vera e propria attività di revisione, un lavoratore a cottimo potrebbe trovarsi a controllare la corrispondenza fra una segnalazione e la categoria di appartenenza indicata dall’utente, tutto ciò per rendere il processo di moderazione più veloce ed immediato, evitando ai moderatori professionisti di dover sottrarre tempo alla rimozione di contenuti per controllare delle formalità.

Negli ultimi tempi Facebook ha annunciato che intende aumentare gli stipendi dei propri moderatori di contenuti statunitensi per adeguarli il più possibile al costo della vita nelle varie zone del Paese, nonché aumentare i compensi per i suoi appaltatori esteri, considerando però che, anche i migliori fra i moderatori non durano più di un paio d’anni, questo programma andrà a beneficio solo del personale assunto in futuro.

1.6 Punteggi di precisione, licenziamenti e dimissioni

Durante l’attività di revisione, un moderatore può ritrovarsi a dover controllare centinaia (se non migliaia) di contenuti al giorno, considerando l’enorme quantità di segnalazioni giornaliere e la necessità di far fronte al maggior numero possibile di queste, avrà a disposizione non più di qualche secondo per assumere la decisione corretta circa la foto o il video oggetto di controllo, decretandone l’eliminazione o il mantenimento sulla piattaforma. Quando un moderatore decide di rimuovere un contenuto deve altresì indicare il motivo della cancellazione, nel caso in cui avesse dubbi, circa l’eliminazione o il motivo di questa, potrà consultarsi con uno dei supervisori che lo guiderà nella decisione, quando le linee guida della policy non sono sufficienti, spesso si instaurano discussioni fra colleghi o fra questi e i loro superiori circa l’interpretazione di un contenuto.

Per testare l’accuratezza delle decisioni dei moderatori, ogni mese dei “policy manager”²² selezionano alcuni dei contenuti moderati da ogni lavoratore e ne verificano la correttezza, Facebook assegna ad ogni soggetto un “punteggio di qualità” che si aggira intorno al 98%: ogni mese ad un moderatore sarà permesso commettere circa un

²² Cfr. Elena Ciccarello (28 gennaio 2020).

2% di errori rispetto al materiale revisionato, ciò significa che potrebbe essere licenziato per aver assunto anche pochissime decisioni sbagliate. Anche una decisione corretta motivata in modo sbagliato è conteggiata come errore, i moderatori hanno la possibilità di appellarsi alle valutazioni dei suddetti soggetti entro quarantotto ore dall'assegnazione dell'errore, per cercare di recuperare il proprio punteggio, l'accuratezza è difficile da raggiungere soprattutto per le continue modifiche alla policy, per le regole scarse e contraddittorie e per il brevissimo tempo a disposizione per la valutazione del contenuto. Questo meccanismo comporta frequenti licenziamenti e un forte turnover, senza contare che la pressione ingenerata nei lavoratori, per il raggiungimento di decisioni accurate e precise, costringe alcuni a dimettersi prima della scadenza del contratto, ed anche se il periodo da moderatori dura relativamente poco, il forte stress e la natura del materiale revisionato hanno conseguenze psicofisiche durature, che potrebbero accompagnarli anche per anni dopo la fine del lavoro.

2 Il lavoro nell'era delle piattaforme digitali

2.1 Rivoluzione digitale

La moderazione di contenuti è una professione nata e sviluppatasi con l'avvento della rivoluzione digitale, internet è divenuto il luogo di incontro per milioni di persone, le quali hanno la possibilità di interagire e contribuire dinamicamente al suo sviluppo, ciò ha comportato la necessità di una regolamentazione specifica di ciò che viene condiviso, al fine di garantire spazi digitali sicuri ed in linea con le regole delineate dalle varie comunità della rete.

L'avvento della rivoluzione digitale ha comportato profondi e rilevanti cambiamenti nella società, la rete internet, e i dispositivi che ne permettono l'accesso, hanno stravolto quelle che erano la vita quotidiana, la realtà lavorativa e le dinamiche relazionali dalle persone nell'era predigitale, a partire dagli anni Novanta la rete è andata consolidandosi e diffondendosi esponenzialmente, abbracciando, oggi, (quasi) ogni singolo aspetto della vita di milioni di persone in tutto il mondo. Molte delle azioni che compiamo quotidianamente sono mediate dalla rete, definita da Stefano Quintarelli "la sede della dimensione immateriale del mondo che, nel ventunesimo secolo, è la maggior base delle relazioni sociali ed economiche delle persone e delle aziende"²³, comunicare, informarsi, vendere o acquistare prodotti e servizi, attività divenute immediate, accessibili a milioni di utenti, trovano il proprio punto di convergenza nella rete. Con internet cadono le barriere geografiche e politiche che impediscono alle persone di dialogare fra loro²⁴, semplificando la circolazione di notizie, informazioni e conoscenza, contribuendo alla creazione di un modo interconnesso, globalizzato, dove le differenze sembrano scomparire e tutti hanno diritto di parola ed espressione, cambia il modo di fare impresa, di produrre valore, di lavorare e nascono nuove figure professionali legate alla rete e alle sue implicazioni. In questo nuovo scenario rivestono un ruolo fondamentale le piattaforme digitali, rese disponibili da determinate aziende, oggi diffusissime, basti pensare ai social network, alle app, ai motori di ricerca, spazi digitali di mediazione in

²³ Cfr. Quintarelli (2019).

²⁴ Cfr. Formenti (2011) p. 7.

grado di connettere persone, lavoratori, produttori, consumatori e chiunque voglia usufruire di determinati servizi, programmi e strumenti tecnologici, gratis o a pagamento, le piattaforme estraggono valore dalla presenza in esse degli utenti e dalle loro interazioni. Basti pensare al caso dei social network o dei motori di ricerca – oggi gli strumenti più utilizzati per comunicare ed informarsi – dove il servizio offerto è apparentemente libero, chiunque ha la possibilità di creare un profilo social o effettuare una semplice ricerca su Google, in realtà, nel momento in cui le persone accedono ad una piattaforma, l’algoritmo che la governa rileva ed archivia ogni informazione relativa alla loro presenza in rete, dati inseriti, contenuti visualizzati, ricerche effettuate. Vengono poi estratte dagli archivi specifiche informazioni relative a singoli utenti grazie al meccanismo di data mining, un metodo automatico di estrapolazione dati che sfrutta precisi algoritmi adatti allo scopo, utili alla profilazione dell’utente, al quale saranno poi mostrati contenuti personalizzati, pubblicità mirate ed offerte ad hoc, nell’ottica della vendita di prodotti e servizi, perché più si conosce una persona e più sarà semplice predire e soddisfare i suoi bisogni.

Le finalità di questa operazione possono essere diverse (di marketing, economiche, scientifiche..) il tratto comune afferisce all’estrazione di valore dalle informazioni, che gli utenti, molto spesso inconsciamente, forniscono attraverso la loro presenza online e le interazioni effettuate²⁵, molto sentita, è la questione relativa alla privacy, la possibilità di essere costantemente connessi viene pagata con l’archiviazione di molti dati personali, estratti e condivisi con modalità non sempre chiare ed utilizzati al fine di produrre profitti. Come suddetto, la maggior parte delle operazioni compiute quotidianamente passano dalla rete (e dalle piattaforme), comunicare, informarsi, comprare prodotti, richiedere servizi e consulenze sono tutte attività trasferitesi nella dimensione immateriale del mondo, ciò ha comportato la crisi di molte attività tradizionali e professioni intellettuali, le quali offrivano servizi oggi mediati e custoditi dalle grandi aziende proprietarie delle piattaforme, le quali operano molto spesso in regime di monopolio o di oligopolio²⁶.

²⁵ Cfr. Armano, Murgia, Teli (2017).

²⁶ La possibilità di attirare e fidelizzare un largo bacino di utenti, la personalizzazione di prezzi ed offerte e l’autoregolamentazione dell’attività da parte di queste aziende, rendono molto difficile l’entrata nel mercato di potenziali concorrenti. Con la progressiva migrazione delle relazioni commerciali dalla dimensione materiale a quella immateriale, le suddette aziende conquistano il rispettivo mercato, imponendo i prezzi finali anche agli intermediari materiali. Imporranno anche le condizioni di acquisto ai fornitori, che non potranno prescindere da questo canale commerciale, onde evitare la scomparsa del mercato. Quintarelli (2017) p. 49.

Lo sviluppo e la diffusione delle piattaforme digitali, oltre al controllo del mercato e all'estrazione di valore dal lavoro inconsapevole svolto dagli utenti in rete²⁷, hanno comportato profondi mutamenti in ambito lavorativo, riguardati non solo l'automazione delle attività, ma altresì l'organizzazione e la dimensione relazionale del rapporto di lavoro, comportando, come suddetto, la creazione di nuove figure professionali, i cosiddetti "lavoratori digitali", e la modificazione di quelle già esistenti.

2.2 Lavoratori digitali

La creazione e lo sviluppo delle piattaforme digitale, intese come luogo di incontro della quasi totalità delle interazioni sociali e lavorative attuali, ha ribaltato quella che era la visione del lavoro nei decenni precedenti, i contratti a tempo indeterminato, la lealtà verso un'unica azienda e i percorsi di carriera lunghi anni, hanno ceduto il passo a lavori saltuari, precari e poco tutelati, nell'ottica della iper-flessibilità organizzativa e del risparmio sui costi di lavoro. Il nuovo modello organizzativo è quello della cosiddetta 'gig economy' (letteralmente 'economia del lavoretto') che vede mobilitate masse di individui addetti alle più svariate attività, dalla consegna di pacchi o generi alimentari, al trasporto di persone o ancora alla consulenza tecnica e artistica, lavoratori più o meno qualificati accomunati dallo svolgimento di una attività mediata da una piattaforma digitale, accessibile in qualsiasi momento tramite una semplice registrazione.

Con contratti a progetto e rapporti di freelancing²⁸, molto spesso i gig workers sono iscritti a diverse piattaforme e svolgono contemporaneamente più attività, 'saltando' da un datore di lavoro all'altro, agli individui è lasciata la facoltà di decidere quando lavorare (quando connettersi alla piattaforma), che compiti effettuare e quando disconnettersi, rappresentando un'ottimo compromesso per il bilanciamento della vita lavorativa e personale, nell'aspettativa di impersonare il proprio 'micro-imprenditore'²⁹. Nella realtà, le piattaforme di gig working si sono trasformate in strumenti per lo sfruttamento dei lavoratori, nonostante le promesse di indipendenza e auto-gestione della prestazione lavorativa, sono le aziende che forniscono la piattaforma a regolare in maniera capillare

²⁷ Armano, Murgia, Teli (2017) p. 10.

²⁸ Armano, Murgia, Teli (2017) p. 43.

²⁹ Armano, Murgia, Teli (2017) p 46.

ogni aspetto della attività ed accaparrarsi la maggior parte dei profitti, nell'ottica della iper-flessibilità produttiva e della riduzione dei costi di lavoro, offrendo ai gig workers salari bassi e poche tutele.

In particolare, è l'algoritmo che governa la piattaforma a stabilire i tempi e le modalità della singola prestazione lavorativa, coordinando e sorvegliando ogni azione, verificando la corretta esecuzione dell'attività e tracciando ogni pausa o momento di inattività, penetranti sono, quindi, i poteri di direzione e controllo esercitati dalla piattaforma, la quale è oggi paragonabile ad un vero e proprio datore di lavoro, impersonale, lontano ed inaccessibile. Lontani da colleghi e vertici aziendali, i gig workers, molto spesso, vivono in una condizione di persistente precarietà, caratterizzata dal susseguirsi di lavoretti mal pagati, ricevendo poco più del salario minimo o una paga a cottimo, in base al lavoro correttamente svolto – correttezza valutata dall' algoritmo, il quale, in caso di errori, disconnette (licenzia) il lavoratore, senza la necessità di ulteriori comunicazioni³⁰ - intrappolati in un circolo di sfruttamento, dal quale risulta difficile uscire. Questi lavoratori possiedono ancora scarse capacità contrattuali, difficilmente riescono a rivendicare trattamenti migliori, considerando la brevità dei contratti e la possibilità di essere sostituiti in ogni momento, le piattaforme, infatti, mobilitano migliaia di individui, potendo facilmente rimpiazzare chi viene disconnesso o conclude il lavoro, i gig workers sono quindi intercambiabili fra loro, considerati non più di un ingranaggio nella macchina produttiva, vittime di un'organizzazione 'invisibile' del lavoro³¹.

Lilly Irani definisce il lavoro digitale invisibile "Black box labour"³² sottolineando come le aziende che (formalmente) occupano pochi lavoratori attraggono numerosi investimenti, in quanto gli investitori le percepiscono come aziende tecnologiche, non basate sulla manodopera³³, le quali hanno quindi un grande interesse nel nascondere i lavoratori e trarre vantaggio dalla loro invisibilità; anche per questo motivo, molto spesso, non si conosce l'esatto numero dei gig workers attivi sulle piattaforme.

³⁰ Armano, Murgia, Teli (2017) p. 14.

³¹ Franchi (2021) p. 23.

³² Letteralmente 'lavoro in scatola nera'. Si usa l'espressione scatola nera per riferirsi ad un sistema il cui funzionamento interno è ignoto.

³³ Irani (2016).

Volendo fare un paragone ‘audace’, la vita e l’attività condotte oggi dai lavoratori digitali potrebbero essere accostate all’ambiente in cui erano inseriti gli operai durante la rivoluzione industriale, i quali erano considerati non più di semplici ingranaggi di una catena di montaggio, in un’industria governata dalla ideologia capitalista del consumo di massa, com’è noto, si è assistito ad un processo di alienazione e disumanizzazione di questi lavoratori. Intercambiabili fra loro, sottopagati e soggetti a turni massacranti, gli operai venivano sfruttati nell’ottica della produzione del più alto profitto possibile a favore degli imprenditori, i quali utilizzavano a loro vantaggio tutti i frutti prodotti dal lavoro, senza considerare minimamente chi li aveva creati, per certi versi, la storia sembra oggi riproporsi, nonostante la rivoluzione digitale portasse con sé la promessa di una società più equa.

Nel contesto delineato sembrano inserirsi perfettamente i moderatori di contenuti, al lavoro sulla piattaforma e per la piattaforma, migliaia di lavoratori sottoposti a turni massacranti, costantemente sorvegliati (anche durante le pause bagno), pagati non più del minimo sindacale, poco tutelati e licenzianti per la commissione anche i pochissimi errori, lavoratori sfruttati ed invisibili impegnanti nella moderazione di milioni di contenuti al fine di correggere gli errori di quell’algoritmo descritto troppo precocemente come “intelligente”³⁴

2.3 Invisibilità

Gli utenti delle piattaforme sono consapevoli del fatto che, quando un contenuto viene segnalato, per qualsivoglia motivo, questo probabilmente finirà con l’essere eliminato, ma quasi nessuno conosce le esatte modalità con cui ciò avviene, è però diffusa l’idea secondo cui, dopo l’invio della segnalazione, i contenuti vengano analizzati ed eliminati in modo automatico dall’algoritmo, nessuno pensa mai che dietro a questo meccanismo ci possano essere delle persone in carne ed ossa. Gli utenti non avrebbero nemmeno motivo di pensare che questa attività è svolta da moderatori umani, considerando i ritmi di lavoro che questi sono tenuti a sopportare per mantenere il passo con l’algoritmo, il quale lavora senza sosta e richiede moderatori disponibili ventiquattro ore al giorno in

³⁴ Franchi (2021) p. 23.

tutti i paesi del mondo, risulta facile presumere che, data la velocità e la precisione nelle risposte, la moderazione avvenga in modo automatico.

Questa generale confusione sulle modalità di eliminazione dei contenuti non è un caso, è stata voluta dalle grandi aziende tecnologiche proprietarie delle piattaforme, la motivazione è da ricercare nella enorme propaganda effettuata dalle suddette riguardante la promessa della completa automazione digitale nel campo dell'editoria e più in generale in qualsiasi altro ambito della vita quotidiana, prospettiva che comporterebbe la creazione di algoritmi in grado di svolgere molte funzioni autonomamente. In particolare, l'obiettivo delle aziende hi-tec, è quello di quadruplicare i profitti a fronte di un limitato utilizzo di manodopera umana, attraverso la creazione di sistemi di intelligenza artificiale in grado di autogovernarsi e svolgere automaticamente ed efficientemente i compiti affidati, compiti in precedenza svolti da lavoratori in carne ed ossa, questo processo comporta quindi l'obsolescenza dei suddetti, i quali sembrano essere destinati a vedersi sostituiti dalle macchine. I lavoratori divenuti obsoleti non sono solo i meno qualificati o aggiornati, basti pensare alla automazione in campo editoriale: quando le persone utilizzano i social come fonte primaria di informazioni e notizie, rimane poco spazio per i professionisti come giornalisti, editori e scrittori, il processo di sostituzione interessa quindi, indistintamente, diverse frangie di lavoratori, i quali possono appunto vedersi rimpiazzati da una macchina intelligente in costante evoluzione.

L'intelligenza artificiale creata per agevolare le attività rendendole più veloci ed efficienti, una volta penetrata nel mondo del lavoro, ha iniziato a diffondersi in maniera esponenziale, anche grazie al costante aggiornamento e miglioramento a cui è sottoposta, andando a rimpiazzare molte professioni tradizionali, nel campo della moderazione di contenuti però, si è assistito ad un fenomeno particolare: le macchine intelligenti hanno comportato la creazione di migliaia di posti di lavoro. In particolare, nonostante l'obiettivo delle aziende tecnologiche di aumentare esponenzialmente i profitti assumendo pochissimi lavoratori qualificati, la necessità di revisionare i milioni di contenuti pubblicati ogni giorno sulle piattaforme, ha costretto le aziende a dover fare affidamento su di un esercito di moderatori umani, i soli in grado di valutare esattamente la natura di una foto o di un video sospetti ed eliminarli dalla rete. L'algoritmo che governa le piattaforme, infatti, non è (ancora) in grado di effettuare questa valutazione, non riuscendo a classificare un contenuto come violento o innocuo, l'esponenziale

crescita degli utenti dei social, del materiale condiviso e dei servizi offerti, ha costretto le aziende ad assumere migliaia di moderatori in tutti i paesi del mondo, non riuscendo altrimenti a rispettare i propri standard di sicurezza e protezione della piattaforma.

La massiccia assunzione di moderatori, soprattutto tramite società terze, volta a rimediare agli errori dell'algorithm, è stata tendenzialmente oscurata dalle aziende hi-tec, le quali non sono mai state particolarmente trasparenti in tema di moderazione, è soprattutto per questo che, ad oggi, non conosciamo con esattezza il numero di moderatori nel mondo, secondo alcune stime questo numero si aggira intorno alle 100.000-150.000³⁵ unità. Numeri in continua evoluzione (senza contare le migliaia di lavoratori a cottimo operanti tramite piattaforme digitali che forniscono prestazioni occasionali) per far fronte all'enorme carico di lavoro e agli errori di quell'algorithm descritto dalle aziende come intelligente ed automatico, che non necessita di lavoratori umani per poter prendere la miglior decisione possibile, il quale si è invece rivelato impreciso e bisognoso dell'intervento di moderatori in carne ed ossa. Oltre a proteggere la promessa della completa automazione digitale, le aziende beneficiano della condizione di invisibilità dei moderatori per lo stesso motivo spiegato in precedenza relativamente ai gig workers: le aziende percepite come tecnologiche attirano ingenti investimenti, le quali hanno quindi interesse a non far trapelare alcuna informazione relativa alla manodopera utilizzata, né agli utenti e nemmeno ai mercati, potendo produrre enormi profitti a spese dei lavoratori sfruttati per questo fine.

L'invisibilità, però, non sempre è un fattore negativo, in alcune occasioni gioca a favore dei moderatori, basti pensare al momento in cui vengono eliminati dei contenuti dalla piattaforma, l'utente oggetto di revisione non sa (e mai saprà) chi, o cosa, ne ha deciso la rimozione, anche se la decisione era sbagliata o ingiusta, finché i moderatori applicano correttamente la policy sono inattaccabili, evitando qualsiasi ritorsione o responsabilità.

L'invisibilità giova al moderatore anche quando si tratta di scavare nelle informazioni private degli utenti, come suddetto, le piattaforme estraggono enormi quantità di dati, sfruttando il lavoro 'gratuito' delle persone collegate, durante l'attività è quindi scontato che i lavoratori entrino in contatto con alcune di queste informazioni, gli utenti sarebbero

³⁵ Chen (2014).

molto meno inclini ad utilizzare la piattaforma sapendo che vengono controllati da altre persone in carne ed ossa e non da un algoritmo automatico³⁶.

Un moderatore esercita quindi un enorme potere su utenti segnalati e relative informazioni personali, per contro, l'utente che segnala esercita un enorme potere sul moderatore, come suddetto, i ritmi di lavoro sono decisi dalla capacità dell'algoritmo di rendere virale qualsiasi contenuto, bello o brutto che sia, ma, altresì, dalla facoltà delle persone di inviare segnalazioni illimitatamente, ad ognuna delle quali dovrà essere data risposta. Vi è da dire che molte delle segnalazioni inviate sono sbagliate o arbitrarie, i moderatori non hanno alcun modo per comunicare con gli autori della segnalazione (dopotutto gli utenti non sono a conoscenza della loro esistenza) e chiedere motivazioni o spiegazioni, la piattaforma munisce le persone di poteri illimitati e si adatta alle loro richieste e bisogni, a discapito dei lavoratori, costretti ad operare come fossero macchine.

2.4 Accordi di riservatezza

La condizione di invisibilità in cui i moderatori sono stati costretti per molto tempo deriva altresì dal severo accordo di riservatezza che accompagna il contratto di lavoro, accordo che questi lavoratori sono costretti a firmare al momento dell'assunzione, il quale li obbliga a non divulgare alcuna informazione relativa all'attività svolta ed ai materiali revisionati, né ad amici e familiari e nemmeno ad altre persone, è infatti richiesta una assoluta discrezione.

Paradossalmente, i moderatori, sottoposti sistematicamente a contenuti tremendi, i quali hanno effetti irreversibili sulla psiche e sulla salute mentale e fisica degli individui, non possono, in alcun modo, confidarsi o sfogarsi con i propri affetti più stretti, le uniche persone con le quali possono parlare di tutto ciò sono i consulenti aziendali messi a disposizione dal datore di lavoro, che, come suddetto, spesso sono mancati o inefficienti. L'accordo di riservatezza comporta altresì l'impossibilità per i moderatori di qualificarsi come tali, dovranno definirsi in alti modi, utilizzando espressioni "opache" (le stesse utilizzate dalle aziende nelle offerte di lavoro) come contractor, social media analyst,

³⁶ Franchi (2021).

legal removals associat ed altre ancora³⁷, terminologie volutamente oscure e non direttamente riconducibili all'effettiva attività svolta, in modo tale che nessuno, né fra gli utenti e nemmeno i giornalisti, possa rintracciarli. I moderatori, molto spesso, non possono nemmeno nominare il nome dell'azienda per la quale lavorano, soprattutto se questa è un gigante tecnologico che non intende far trapelare troppe informazioni riguardanti la propria attività di moderazione, un problema sorge, in particolare, alla fine del periodo da moderatori, quando i suddetti dovranno tentare di spiegare, al nuovo potenziale datore di lavoro, il precedente impiego svolto, senza violare l'accordo di riservatezza. Secondo alcune testimonianze, alcune delle aziende tecnologiche che appaltano l'attività di moderazione a società intermediarie, per assicurarsi riservatezza ed anonimato, utilizzano parole ed espressioni in codice per riferirsi alle operazioni di moderazione, il "the honey badger project" è il nominativo coniato da Facebook, il quale deve essere utilizzato da tutti gli appaltatori dell'azienda ogni qual volta vogliano parlare del lavoro per il quale sono assunti, nell'ottica della totale segretezza³⁸.

Alcune società intermediarie, al fine di evitare il rischio che i moderatori condividano o trascrivano qualche informazione sensibile, riguardante la società o gli utenti controllati, hanno costretto i lavoratori a depositare in appositi spazi il proprio cellulare all'inizio del turno ed hanno altresì vietato, nella postazione di lavoro, fogli di carta e penne, mentre qualsiasi altro oggetto personale dovrà essere controllato e posto sempre in vista³⁹.

Altre società che impiegano moderatori in outsourcing hanno fatto addirittura ricorso alla sorveglianza privata dei lavoratori, per dissuaderli dal parlare della loro attività con i giornalisti⁴⁰, in molte altre aziende i moderatori sono letteralmente separati dal resto dei dipendenti della piattaforma (per evitare che questi ultimi vengano a contatto con il materiale segnalato), i quali molto spesso non sanno che questo lavoro è svolto da persone in carne ed ossa⁴¹. L'obiettivo è quindi la totale discrezione, l'invisibilità assoluta di questi lavoratori, i quali, nel caso in cui decidessero di violare l'accordo di riservatezza, andrebbero incontro a licenziamenti e pesanti sanzioni economiche⁴², infatti, sono stati

³⁷ Franchi (2021) pp. 36-37.

³⁸ Cfr. Wilson (11 agosto 2018).

³⁹ Newton (2019).

⁴⁰ Cfr. Wray (26 giugno 2018).

⁴¹ Franchi (2021) p 100.

⁴² Le multe previste da alcuni accordi di riservatezza in caso di violazione della regola di non divulgazione arrivano fino a 10.000 dollari. Wray (26 giugno 2018).

molto pochi i moderatori che intenzionalmente hanno violato l'accordo per parlare apertamente con giornalisti e ricercatori dell'attività svolta e delle condizioni di lavoro, spesso degradanti, alle quali sono sottoposti.

Lontani da familiari e amici, lontani dagli altri lavoratori, dagli utenti, da qualsiasi altra persona (e molto spesso anche dai contenuti revisionati), i moderatori sono stati a lungo tenuti nascosti dai vertici aziendali, gli unici a conoscenza della loro reale posizione, i quali hanno cercato in qualsiasi modo di proteggere il mito della completa automazione digitale, trasferendo l'attività all'esterno ed assicurandosi il silenzio di società intermedie e lavoratori collegati.

2.5 Addestrare l'algoritmo⁴³

Come suddetto, l'attività di revisione e rimozione dei contenuti pubblicati nelle piattaforme è, necessariamente, stata affidata ad un esercito di moderatori umani, in quanto, non sono ancora stati creati (nonostante le promesse) algoritmi in grado di assumere decisioni accurate, decifrando in modo corretto le varie sfumature e i molteplici contesti caratterizzanti i contenuti segnalati, l'interpretazione è, quindi, ancora oggi, una attività mentale riservata ai lavoratori. La moderazione di contenuti, differentemente rispetto a quello che si potrebbe pensare, richiede attenzione e capacità di analisi, operazioni mentali ritenute troppo complesse per gli algoritmi, nonostante ciò, ai lavoratori è chiesto di agire come macchine, considerando il poco tempo a disposizione, i controlli e le pressioni subite, un compito intellettualmente impegnativo diviene un'azione automatizzata, la cui ripetizione provoca un senso di alienazione e frustrazione⁴⁴. Circostanze che si ripeteranno almeno fino a quando non sarà creata un'intelligenza artificiale guidata da algoritmi in grado di assumere decisioni accurate, e forse, questa prospettiva è più vicina di quello che si pensa: come suddetto, quando un utente invia una segnalazione, deve indicarne la motivazione (pornografia, violenza, spam...), durante la revisione, il moderatore, se deciderà di procedere alla eliminazione, dovrà catalogare il contenuto selezionando una serie di alternative. Tutti i dati immessi

⁴³ Franchi (2021) p 223.

⁴⁴ Punsmann (2018).

sono archiviati nel sistema ed elaborati dall'algoritmo, il quale li utilizza per 'imparare', infatti, grazie alle tecniche di machine learning è possibile addestrare la macchina al riconoscimento dei contenuti non ammessi, la quale analizza e confronta le scelte effettuate da utenti e moderatori, valutando in che percentuale determinati contenuti vengono rimossi e procedendo per analogia alla comparazione di questi con materiale simile. Il suddetto approccio all'intelligenza artificiale è per ora utilizzato al sol fine di bloccare la pubblicazione di contenuti già in precedenza eliminati, ma potrebbe rappresentare un'enorme svolta in questo campo, infatti, i moderatori stanno letteralmente allenando un software che un giorno potrebbe rimpiazzarli⁴⁵, perché più la macchina svilupperà la capacità di apprendere, o meglio, imitare⁴⁶ le capacità intellettive umane e meno spazio ci sarà per questi lavoratori.

⁴⁵ Gilbert (2020).

⁴⁶ Kaplan (2018).

3 Censurati e censura(tori)

3.1 Realtà datificate

Come esposto in precedenza, l'avvento della rivoluzione digitale, o digitalizzazione⁴⁷, ha comportato un cambio di passo significativo nella società, la diffusione su larga scala di strumenti tecnologici unita alla possibilità di accedere alla rete in qualsiasi momento, hanno dato vita ad un mondo globalizzato ed interconnesso, in cui le distanze e le differenze sembrano azzerarsi, le possibilità moltiplicarsi e dove tutti hanno diritto ad esprimersi, informarsi, comunicare ed associarsi liberamente.

Oggi, buona parte della vita privata e professionale di milioni di persone è mediata dai dispositivi tecnologici, quali computer e cellulari, e dalla rete, i quali accompagnano e scandiscono (quasi) ogni momento della giornata, indipendentemente dalla attività svolta, nell'ottica della connettività *always on*⁴⁸, rendendo così sempre più labile e sfocato il confine fra online e offline⁴⁹, fra immateriale e materiale, dove il mondo immateriale (o cybersphera) diviene così teatro della gran parte degli eventi di vita delle persone.

I suddetti dispositivi tecnologici, oramai diffusissimi ed indispensabili allo svolgimento di molteplici attività, traducono ed inviano nella cybersphera informazioni datificate⁵⁰ riguardanti eventi riconducibili agli utenti, la quale diventa così un immenso spazio pubblico, accessibile a chiunque ed in continua espansione, nel quale ogni persona riversa inarrestabilmente dati che la riguardano, i quali sono poi resi disponibili, per chiunque possa accedervi, e rielaborati per svariati fini⁵¹. Dai cellulari in grado di registrare la posizione e gli spostamenti, alle ricerche effettuate su internet, fino alla creazione di identità professionali o social, ogni dato offerto è idoneo a veicolare informazioni sulle persone, ma la questione rilevante non riguarda tanto il fatto che i dati esprimono

⁴⁷ Primo livello dell'informatizzazione, necessario per l'inserimento progressivo del computer nella vita di ciascuna persona. Cfr. Negroponte (1995).

⁴⁸ Collegamento permanente alla rete. Quintarelli (2018).

⁴⁹ Cfr. Floridi (2007).

⁵⁰ Datificazione (dataficazione) s. f. Processo tecnologico che trasforma vari aspetti della vita sociale o della vita individuale in dati che vengono successivamente trasformati in informazioni dotate di nuove forme di valore anche economico. Cfr. Treccani.

⁵¹ Cfr. Sarra (2019).

informazioni, quanto i modi con cui essi vengono utilizzati, i quali sono estratti e rielaborati al fine di scoprire collegamenti e notizie prima ignoti⁵². In particolare, i dati prodotti dalle manifestazioni degli utenti, sono archiviati nei database delle aziende⁵³ che erogano servizi e gestiscono siti web o piattaforme digitali, le quali estraggono informazioni e conoscenza dalla enorme mole di dati a disposizione, sfruttando il lavoro inconsapevole compiuto dagli utenti, nell'ottica della produzione di profitti e valore, infatti, nonostante l'utilizzo della rete sembri facile e gratuito, la connessione viene pagata con l'archiviazione di molti dati personali.

Attraverso il meccanismo del data mining, dai dati presenti nei database vengono estratte, con sofisticate tecniche matematico-statistiche, informazioni potenzialmente proficue, implicite e finora ignorate, al fine di creare modellizzazioni, utili alla pre comprensione specifica dei soggetti e del loro comportamento futuro⁵⁴, operazioni affidate ad algoritmi adatti allo scopo ed effettuate direttamente dalle aziende che intendono servirsene oppure da altri intermediari, come società addette alla gestione dati⁵⁵.

Come precedentemente esposto, in un contesto del genere, molto sentita, è la questione relativa alla privacy, l'Unione europea nel 2016 ha emanato il Regolamento Generale per la Protezione dei Dati personali⁵⁶, sostituendo la precedente direttiva, il quale ha delineato un nuovo quadro normativo in materia di trattamento e protezione dei dati personali delle persone fisiche, limitando le suddette operazioni nell'ottica della tutela dei diritti e delle libertà fondamentali degli individui. Il regolamento disciplina altresì la libera circolazione dei suddetti dati affermando, però, che “la libera circolazione dei dati personali nell'Unione europea non può essere limitata né vietata per motivi attinenti alla protezione delle persone fisiche con riguardo al trattamento dei dati personali”⁵⁷ segnando quella che sembrerebbe una vera e propria contraddizione, in quanto, nonostante la circolazione configuri una modalità di trattamento dei dati, essa non sembra andare incontro a restrizioni. Claudio Sarra ne ‘il mondo dato’ propone una soluzione

⁵² Sarra (2019).

⁵³ I dati sono acquisiti e archiviati in quantità altrimenti inimmaginabili grazie al fatto che archiviare ha un costo nullo. Quintarelli (2019).

⁵⁴ Sarra (2019).

⁵⁵ Nonostante alcune aziende abbiano promesso di non divulgare dati personali, altri siti ed applicazioni non garantiscono nulla del genere, dietro molte pagine e siti web si annida un enorme mercato di informazioni su ciò che gli utenti fanno online, controllato da società per la raccolta di dati, poco conosciute ma molto redditizie. Cfr. Pariser (2012).

⁵⁶ Cfr. Reg. 2016/679, GDPR.

⁵⁷ GDPR, Art. 1, n 3.

interpretativa della disciplina affermando che, fermi i principi di rilevanza costituzionale riguardanti i diritti e le libertà fondamentali degli individui, il legislatore europeo ha voluto riconfigurare il concetto di protezione della persona per inserirlo nel contesto attuale, caratterizzato dalla continua ed inarrestabile circolazione di dati personali, dove sono le stesse regole che governano questi processi a porsi come limite esterno⁵⁸.

3.2 Personalizzazione dei contenuti

Il data mining, è quindi utilizzato al fine di estrarre informazioni utili o precedentemente ignote riconducibili ad una o più persone, le quali concorrono alla profilazione⁵⁹ dei soggetti, in particolare, viene creato un profilo digitale dell'utente composto da tutti i dati di partenza e dalle informazioni ricavate da questi, evidenziando età, nazionalità, professione, stato di salute, gusti e interazioni effettuate, senza contare le nuove notizie emerse grazie alle tecniche di manipolazione dei dati. Il profilo creato è analizzato e valutato al fine di inquadrare comportamenti e bisogni presenti del soggetto, e, tentare di predire quelli futuri, orientando le decisioni in base al giudizio su di esso scaturito dall'esame della sua identità digitale, la profilazione è generalmente utilizzata dalle aziende al fine di personalizzare i contenuti con i quali l'utente entra in contatto, creando pubblicità ed offerte ad hoc.

Secondo Eli Pariser, la strategia di mercato dei colossi di internet è semplice: «più informazioni personali sono in grado di offrire, più spazi pubblicitari possono vendere e più probabilità ci sono che gli utenti comprino i prodotti mostrati»⁶⁰, offerte mirate proposte al momento perfetto, aziende come Google o Facebook attirano ingenti investimenti, ma la pubblicità rimane sicuramente al centro del modello commerciale, generando enormi profitti. Grazie alla profonda conoscenza di bisogni, gusti e preferenze dell'utente, è possibile predire quelle che saranno le sue esigenze, creando offerte mirate

⁵⁸ Sarra (2019).

⁵⁹ «Profilazione»: qualsiasi forma di trattamento automatizzato di dati personali consistente nell'utilizzo di tali dati personali per valutare determinati aspetti personali relativi a una persona fisica, in particolare per analizzare o prevedere aspetti riguardanti il rendimento professionale, la situazione economica, la salute, le preferenze personali, gli interessi, l'affidabilità, il comportamento, l'ubicazione o gli spostamenti di detta persona fisica; GDPR Art 4, n 4.

⁶⁰ Pariser (2012) p 13.

per ogni consumatore: i prezzi sono fissati dinamicamente grazie a sistemi di dynamic pricing che, grazie all'analisi dei dati effettuata da parte di intelligenze artificiali, determinano in tempo reale il prezzo corrispondente al massimo valore che il soggetto è disposto a pagare⁶¹.

Le aziende non si servono della profilazione al sol fine di condizionare gli acquisti, infatti, volendo creare spazi digitali il più possibile aderenti alle esigenze e caratteristiche dei singoli utenti, utilizzano le informazioni a disposizione e le previsioni degli algoritmi al fine di personalizzare ogni singolo contenuto con il quale l'utente entra in contatto, offrendogli un servizio su misura, in grado di soddisfare e predire ogni suo bisogno.

Come precedentemente esposto, considerando che oggi molte persone utilizzano social network o altre piattaforme digitali come fonte primaria di notizie ed informazioni, la personalizzazione dei contenuti diviene un problema, in quanto, se ad ogni utente vengono offerte solo informazioni personalizzate, create ed organizzate in base al profilo personale ed linea con interessi ed aspirazioni, allora, sarà oscurata o limitata la visione di molti contenuti, seppur importanti. L'utente è così immerso in una realtà digitale confezionata su misura, composta da notizie ed informazioni organizzate automaticamente dall' algoritmo, il quale seleziona esclusivamente contenuti coerenti con l'identità e le convenzioni del soggetto, escludendo e limitando ciò che è considerato estraneo e distante dalla sua visione del mondo, esponendolo solo a ciò che vuole sapere e non a quello che dovrebbe sapere, condizionando così il suo accesso alla conoscenza.

3.3 Realtà differenti

L'utente esposto solo a contenuti personalizzati, in linea con la sua visione del mondo, si ritrova rinchiuso in quella che Eli Pariser ha definito 'bolla dei filtri', la quale altera il modo con cui egli entra in contatto con idee ed informazioni⁶², condizionando la sua realtà, le piattaforme veicolano ad ogni persona messaggi diversi, ciò comporta la creazione di tante realtà differenti quanti sono gli utenti della rete⁶³. La questione riveste

⁶¹ Quintarelli (2019) p. 48.

⁶² Pariser (2012) p. 15.

⁶³ Per tentare di capire la portata di questo fenomeno è sufficiente ricordare che gli utenti di Facebook hanno da poco sfiorato i due miliardi.

un'enorme importanza, in quanto, le informazioni condizionando la nostra visione del mondo, formandoci come persone, aiutandoci a comprendere l'ambiente in cui siamo inseriti e la portata dei problemi da affrontare, nonostante l'accesso alla conoscenza sia oggi illimitato ed immediato, grazie ad internet, esso è ostacolato dalla personalizzazione dell'informazione, dove le preferenze e le interazioni espresse in passato determinano i contenuti che saranno visualizzati in futuro.

La nascita della bolla dei filtri non influisce solo sulle modalità con cui vengono trattate le notizie, alterando il modo in cui le persone entrano in contatto con idee ed informazioni, può anche influire sul loro modo di pensare, limitandole in diversi modi, innanzitutto amplificando il cosiddetto 'effetto conferma'⁶⁴, infatti, consumare solo informazioni conformi alla propria idea di modo non fa altro che rafforzare le opinioni sedimentarie di ciascuno. In generale, ogni persona tende a ricercare conferme alle proprie convinzioni, ma una volta dentro la bolla dei filtri, questa tendenza si amplifica considerevolmente, i contenuti presenti in essa sono selezionati appositamente per essere in linea con l'identità del soggetto ma egli non conoscendo i criteri utilizzati per filtrare le informazioni è portato a pensare che esse siano obiettive e neutrali, andando così a rafforzare ulteriormente le proprie credenze.

In secondo luogo, la bolla dei filtri limita qualsiasi interferenza esterna, riducendo le possibilità di apprendimento e rimuovendo gli stimoli che favoriscono la curiosità, indispensabili al fine di entrare in contatto con nuove idee, diverse prospettive e modi di pensare, il web personalizzato è utile, ma incoraggia le persone a passare meno tempo in atteggiamento di scoperta, rendendole un po' più passivi⁶⁵, limitando conoscenza e creatività.

Nonostante la diffusione di internet promettesse la diffusione su larga scala di informazioni e conoscenza, accessibili direttamente da chiunque in qualsiasi momento, e l'eliminazione della mediazione dei mass media, ha comportato la creazione di nuovi intermediari, le piattaforme digitali, le quali rilevano ogni traccia digitale lasciata dagli utenti, poi memorizzate ed elaborate con sistemi di intelligenza artificiale al fine di immergere le persone in una info sfera personalizza su base individuale⁶⁶. I nuovi

⁶⁴ Gli psicologi chiamano la tendenza a credere a ciò che rafforza le proprie opinioni, errore di conferma (confirmation bias).

⁶⁵ Pariser (2012).

⁶⁶ Quintarelli (2019) p. 89.

intermediari custodiscono le identità digitali di milioni (o miliardi) di persone e sono al corrente di un numero di informazioni forse inimmaginabile, se la conoscenza è potere, le asimmetrie di conoscenza sono asimmetrie di potere e se la persuasione personalizzata funziona bene con prodotti e servizi, può funzionare anche con le idee⁶⁷, meccanismo sottile in quanto le persone non sanno che i dati raccolti vengono utilizzati per condizionare il loro comportamento.

Recenti studi hanno avanzato alcune critiche relativamente al fenomeno delle bolle di filtraggio, con particolare riferimento al rafforzamento dell'effetto conferma che esse indurrebbero, in particolare, i Dr. Dubois e Black⁶⁸ sostengono che internet ha creato un ambiente mediatico molto ampio, il quale non ricomprende solo i social network ma altresì media online o cartacei, televisioni e radio, e all'interno di questo ambiente le persone non sono rinchiusi in bolle. Infatti, in luogo così complesso gli utenti consumano molti media, interagendo con altre persone che hanno punti di vista diversi, cambiando idea, incontrando informazioni contraddittorie e controllando le notizie trovate nei social network⁶⁹, nonostante si riescano comunque a rintracciare forme di polarizzazione e rafforzamento di idee o pregiudizi su alcune piattaforme social, è più probabile incontrare diverse opinioni e visioni del mondo, ampliando l'orizzonte della conoscenza.

3.4 La complicità dei moderatori

Il fenomeno delle bolle dei filtri (o echo chamber) è stato per anni considerato conseguenza del lavoro di selezione automatica effettuato dagli algoritmi di profilazione sulla base delle identità digitali e delle iterazioni degli utenti, ma mai definitivamente confermato, ed effettivamente più che parlare di bolle automatiche sarebbe maggiormente corretto parlare di bolle manuali⁷⁰, frutto del lavoro di revisione ed eliminazione dei moderatori di contenuti.

⁶⁷ Pariser (2012).

⁶⁸ Dubois, Black (2018).

⁶⁹ La contribuzione di notizie da parte dei singoli nelle piattaforme social alimenta il fenomeno delle fake news. Quintarelli (2019).

⁷⁰ Franchi (2021).

Probabilmente, la condizione di invisibilità in cui versano i lavoratori, ha determinato il passaggio in sordina del ruolo che questi rivestono nella fase di selezione dei contenuti offerti dalle piattaforme, sono i moderatori i veri responsabili delle limitazioni alle quali sono sottoposti gli utenti, decidendo cosa è possibile vedere e cosa no, avvolgendo persone, o intere comunità⁷¹, in bolle e rappresentando l'unico vero filtro fra individui e realtà. Ovviamente, le scelte dei moderatori sono frutto di un lavoro di interpretazione della policy, ma, come suddetto, essa è molto spesso complessa, dettagliata o esageratamente generica, senza contare il fatto che subisce continue modifiche e rivisitazioni conseguenti ad avvenimenti esterni alla piattaforma stessa, perciò, in ultima istanza spetta ai lavoratori decidere quali contenuti salvare e quali eliminare, precludendo la possibilità degli utenti ad avervi accesso. La policy di una piattaforma cambia altresì da paese a paese, onde rispettare gli standard e le leggi locali, i moderatori hanno il potere (dovere) di oscurare determinati contenuti in alcune zone del mondo, Facebook è stata accusata di collaborare con regimi oppressivi, come Russia o Cina, andando incontro alle loro esigenze di censura politica⁷², nonostante la libertà di parola sia uno dei principi cardine della piattaforma⁷³ (e di internet).

Fra la mancanza di preparazione specifica su determinati argomenti e settori⁷⁴, le regole eccessivamente rigide (o generiche) della policy e le infinite sfumature di significato che possono caratterizzare un contenuto, i moderatori sono spesso soggetti ad errori di valutazione, nonostante la difficile ed attenta attività di interpretazione, svariate decisioni non vengono assunte correttamente e ciò comporta la censura (ingiusta) di contenuti innocui o comunque ammissibili. La visione della sessualità o di altre forme d'espressione artistica e umana vengono da sempre catalogate come contenuti vietati e rimossi onde incontrare il favore della maggioranza degli utenti e degli inserzionisti, nascite, allattamenti, foto di statue o dipinti senza veli e molti altri contenuti simbolo di una rivoluzione culturale e dei costumi vengono censurati a seguito di segnalazioni anonime, gratuite e molto spesso in malafede⁷⁵. Nonostante il fine ultimo dei moderatori

⁷¹ I moderatori hanno il potere di oscurare determinati contenuti in base alla zona geografica o alla fascia d'età, a seconda delle regole contenute nella policy.

⁷² Solon (23 maggio 2017).

⁷³ Gli interessi commerciali sembrano, molto spesso, venire prima di alcuni diritti fondamentali come la libertà di parola ed espressione.

⁷⁴ Molto spesso i moderatori sono chiamati a valutare contenuti a carattere scientifico, economico, medico, argomenti su cui non hanno alcuna preparazione.

⁷⁵ Franchi (2021).

sia di salvaguardare la sensibilità di milioni di utenti dalla visione di determinate circostanze, vi è da considerare che anche i contenuti orribili sono pur sempre un frammento della realtà, la quale è così preclusa a milioni di utenti, impossibilitati ad avere accesso a determinate informazioni e rinchiusi in bolle, create giorno dopo giorno dalla minuziosa opera di revisione ed eliminazione dei contenuti considerati ‘inappropriati’.

Le piattaforme divengono così un luogo edulcorato, una realtà parallela, una via di fuga rispetto alle nefandezze che la vita esterna potrebbe riservare, ciò è sicuramente un incentivo per gli utenti a connettersi più e più volte, infatti, senza i moderatori, social, motori di ricerca e siti diverrebbero in poco tempo luoghi inospitali grazie all’algoritmo che, se lasciato libero di agire, renderebbe virale qualsiasi contenuto, colmando le piattaforme di materiale riprovevole.

I moderatori arrivano dove gli algoritmi sbagliano⁷⁶, anche se da un punto di vista esterno, non sarebbe possibile sapere con certezza chi ha agito, se i lavoratori o quell’intelligenza artificiale tanto cara alle aziende proprietarie delle piattaforme, ma l’invisibilità non è solo un vincolo, permette ai moderatori di insinuarsi nelle identità digitali di milioni di utenti, visualizzare foto, messaggi, decidendone le sorti e selezionando le informazioni alle quali possono avere accesso. Sono i moderatori a decidere quali contenuti possono diffondersi nelle bolle e quali meritano di essere eliminati, attraverso un processo editoriale poco trasparente, abituando gli utenti all’illusione di vivere in un modo in cui tutto ciò che è ambiguo o spaventoso non esiste o esiste in misura minore rispetto ai contenuti più piacevoli, condizionando la loro vita, senza che essi siano del tutto consapevoli di essere manipolati (e manipolabili)⁷⁷.

I moderatori, nonostante siano non più di pedine utilizzate al fine di preservare l’illusione dell’automazione editoriale, partecipano alla realizzazione delle suddette dinamiche, diventano gli esponenti di un servizio di censura globale⁷⁸, nonostante internet promettesse di rendere il mondo più aperto e consapevole, esso sembra invece andare nella direzione opposta, dove ogni evento considerato sconveniente è eliminato, onde preservare la reputazione della piattaforme e il favore di utenti e mercato.

⁷⁶ Franchi (2021).

⁷⁷ Franchi (2021).

⁷⁸ Franchi (2021) p 97.

I moderatori prendono parte alla realizzazione di un'utopia globale, escludendo dall'orizzonte degli eventi tutto ciò che ostacola la creazione di una nuova società, beata ed ignara, propensa al consumo di servizi e contenuti organizzati da un algoritmo secondo regole valide per tutti⁷⁹, secondo Tarleton Gillespie, le piattaforme divengono così, «non un mezzo per diffondere ciò che pubblichiamo, quanto uno strumento che definisce la realtà che vediamo»⁸⁰.

3.5 Le ricadute psicologiche della moderazione e i rimedi adottati

Come precedentemente esposto, molto del materiale revisionato è innocuo e noioso, fake news, profili o annunci fasulli, oppure frutto di segnalazioni errate o in mala fede, d'altro canto, molti altri contenuti sono a dir poco deprecabili e se, i moderatori impediscono agli utenti la visione di ciò, nessuno impedisce ai moderatori la visione dei suddetti, non esiste alcun filtro per loro, nessuna limitazione o avviso. I lavoratori sono tenuti a revisionare ore di video e migliaia di foto con la consapevolezza che da un momento all'altro potrebbe apparire sullo schermo qualcosa che non dimenticheranno mai e che li segnerà per sempre, evidentemente, le poche ore di assistenza offerte in azienda, non sono sufficienti e l'impossibilità di confidarsi con amici e familiari proietta i moderatori in uno stato di ansia e stress perenni. Questa condizione non può che avere ricadute gravi e permanenti sull'integrità fisica e mentale dei lavoratori, molti dei quali sviluppano un disturbo da stress post traumatico che li accompagnerà per il resto della loro vita, a prescindere che essi siano stati moderatori per un mese o un anno, quanto visto, anche se per un breve periodo, non può essere dimenticato facilmente, nemmeno con l'aiuto mirato di professionisti. Se la macchina è programmata per dimenticarsi dei contenuti revisionati dopo averli utilizzati per addestrare l'intelligenza artificiale, la memoria dei moderatori sembra andare incontro ad un processo inverso⁸¹, dove le immagini viste e i suoni uditi continueranno ad infestare le loro giornate e i loro incubi, sopravvivendo per anni a

⁷⁹ Franchi (2021) p 171.

⁸⁰ Gillespie (2018) p 21.

⁸¹ Franchi (2021) p. 238.

qualsiasi tentativo di dimenticare o guarire, pregiudicando la possibilità di tornare a condurre un'esistenza degna di questo nome.

Molti ex moderatori hanno deciso di citare in giudizio le aziende tecnologiche e le società di intermediazione per vedersi riconosciuto un risarcimento a seguito dell'accertamento di un disturbo da stress post traumatico causato dal lavoro svolto, accusando i datori di lavoro di non prevedere adeguati programmi di formazione, non implementare le misure di sicurezza e non fornire adeguato sostegno psicologico ai lavoratori⁸², addetti ad una delle professioni più usuranti al mondo.

In seguito a svariate denunce, provenienti da diversi paesi nel mondo e in continuo aumento, trasformantesi molto spesso in vere e proprie class action, alcune società di moderazione di outsourcing hanno deciso di obbligare i lavoratori in fase di assunzione a firmare un documento che li avvisa dei pericoli del mestiere, mettendoli in guardia sulle alte probabilità di sviluppare un disturbo post traumatico da stress⁸³.

L'elevato numero di denunce e la sempre maggiore attenzione mediatica attirata dalle condizioni dei moderatori, ha portato le aziende a prendere coscienza della situazione, le quali hanno promesso maggiore attenzione ai bisogni psicologici dei lavoratori e una implementazione dei sistemi di sicurezza e benessere, sviluppando altresì alcuni programmi volti a ridurre l'impatto che determinati contenuti hanno sui moderatori, aumentando allo stesso tempo la loro resistenza fisica e mentale. Alcune aziende hanno condotto esperimenti su diversi gruppi di lavoratori al fine di testare le loro reazioni a determinati espedienti tecnici volti a diminuire l'impatto dei contenuti visualizzati, ad esempio, le immagini e i video erano convertiti in bianco e nero o in grigio, oppure veniva data loro la possibilità di eliminare l'audio o introdurre sfocature selettive, mentre il sangue era colorato di verde⁸⁴.

Gli esperimenti condotti stanno restituendo buoni risultati, comprovando il miglioramento dell'umore di molti moderatori coinvolti nei test, ma, una soluzione che le aziende non hanno ancora messo in gioco è il tentativo di diminuire il numero di contenuti inquietanti con cui i lavoratori entrano in contatto ogni giorno e che li porteranno, chi presto chi tardi, a sviluppare disturbi e malattie mentali, dalle quali difficilmente guariranno.

⁸² Koebler (24 settembre 2018).

⁸³ Murgia (26 gennaio 2020).

⁸⁴ Newton (2019).

3.6 Missione

Il mestiere del moderatore sembra un incubo sotto tutti i punti di vista, ma se tante persone decidono di intraprendere questa professione, rimanendo mesi o addirittura anni alla propria scrivania, allora forse qualche lato positivo deve esserci, oltre alla paga assicurata a fine mese e alla stabilità lavorativa, molte persone resistono nel tempo per via dello scopo, dell'obiettivo di vita che questa attività sembra regalargli. Infatti, diversi moderatori attuali (ed ex) hanno dichiarato di comparare il proprio lavoro ad una missione, la missione di rendere internet un posto più sicuro per tutti⁸⁵, paragonando sé stessi a dei supereroi⁸⁶, agendo nell'ombra e salvando milioni di utenti dai contenuti aberranti pubblicati dai loro coetanei, compito difficile, se non impossibile, ma comunque in grado di spronare migliaia di persone a proseguire questa attività. Seppur, come suddetto, la finalità di protezione degli utenti è molto spesso utilizzata al fine di giustificare azioni di censura a livello globale, molti moderatori credono davvero in questa missione, e grazie al loro potere di eliminazione dei contenuti illeciti, di penalizzazione degli utenti che violano la policy e di denuncia alle autorità (per particolari crimini come la pedopornografia) forse riescono davvero a rendere il mondo un posto migliore.

⁸⁵ Bruni, Chemaly (2006).

⁸⁶ Franchi (2021).

4 Conclusioni

4.1 Un caso attuale

In modo dove dati ed informazioni sono centrali, traducendo e rappresentando ogni esperienza umana nella cybersfera, anche l'esperienza della guerra cambia volto, venendo, in parte, combattuta attraverso una serie di strategie comunicative propagandistiche volte a manipolare la diffusione di informazioni e notizie grazie allo sfruttamento di strumenti tecnologici, ne è un esempio l'attualissimo conflitto in corso fra Russia ed Ucraina, dove la cosiddetta 'guerra dell'informazione' sta giocando un ruolo fondamentale.

L'Ucraina nasce come stato indipendente all'inizio degli anni Novanta a seguito della dissoluzione dell'Unione sovietica, ma fin dagli albori l'integrità del paese è minacciata dai continui scontri fra fautori dell'avvicinamento dell'Ucraina all'Unione europea e sostenitori dello storico legame con la Russia, dissapori culminati nel 2014 a seguito del rifiuto del presidente ucraino di firmare l'accordo di associazione e libero scambio con l'Ue. A seguito di questo rifiuto si sono verificati una serie di scontri di piazza, i quali hanno visto coinvolti alcuni nazionalisti ucraini (fra cui alcuni neonazisti), terminati con l'uccisione di diverse persone e la fuga del presidente, successivamente diverse regioni ucraine di confine (a maggioranza russofona) sono state occupate da militari russi e attraverso referendum illegittimi dei cittadini, perché non previsti dalla costituzione Ucraina, sono state dichiarate annesse alla Russia. Le suddette vicende hanno dato il via a conflitti armati nei territori in questione e nonostante i tentativi di accordo fra paesi, la guerra civile è continuata quasi ininterrottamente fino ad oggi, culminando nell'invasione dell'Ucraina da parte della Russia, la quale rivendica l'annessione dell'intero territorio ucraino giustificando l'attacco con la necessità di 'liberare' i cittadini⁸⁷ e fermare l'avanzata della NATO, e dell'Occidente, verso est.

In Russia, ormai da anni, ogni notizia pubblicata dalla stampa è filtrata e richiede la preventiva approvazione del Cremlino, il cui apparato informativo non lascia particolare

⁸⁷ Putin ha usato l'espressione 'de-nazificazione'.

spazio all'interpretazione essendo generalmente univoco, una narrazione diversa di ciò che realmente accade nel paese è disponibile grazie ai media stranieri e alla diffusione di informazioni per il tramite delle piattaforme digitali, in grado di aggirare le politiche di censura russe.

La propaganda russa accompagna ormai da anni la crisi ucraina e nelle settimane antecedenti l'inizio della guerra, il presidente Vladimir Putin attraverso una serie di conferenze e discorsi pubblici ha preparato le basi per l'invasione e ricercato il sostegno fra la popolazione interna russa, mentre Ucraina e Stati Uniti tentavano di scongiurare il conflitto attraverso vie diplomatiche, tentativo fallito a seguito dell'annuncio televisivo di un'operazione militare speciale da parte del presidente russo. La guerra in atto è oggi combattuta su due fronti, da un lato il conflitto armato e dall'altro la propaganda, in Russia, come suddetto, la stampa è sotto il controllo dello stato, il quale filtra ogni notizia ed applica politiche di censura molto rigide, i media russi riportano solo la versione ufficiale del Cremlino sulla guerra⁸⁸ e questo ha spinto i cittadini del paese a ricercare informazioni sui social network. Le piattaforme social sono quindi state oscurate dallo stato centrale, nonostante questo i cittadini più giovani continuano ad accedere attraverso server privati, aggirando la censura e rimanendo informati grazie alle notizie provenienti da fonti internazionali rese disponibili attraverso il deep web, mentre chi non ha queste possibilità vede la propria unica fonte di informazione nei media russi, rimanendo all'oscuro su ciò che realmente sta accadendo.

Le strategie utilizzate dalla Russia in questa guerra delle informazioni sono varie e hanno diversi fini, perlopiù vengono pubblicate e rese virali negli altri paesi notizie false o fuorvianti accompagnate da foto e video 'dimostrativi', lo scopo è di disorientare il 'nemico', il quale sarà costretto a mettere in discussione tutte le informazioni ricevute, creando il caos, nel paese la comunicazione è utilizzata come una vera e propria arma. Né i politici e nemmeno i giornalisti stranieri riescono a confermare con certezza la veridicità o meno di certe notizie, in arrivo sia dalla Russia sia dall'Ucraina, non essendoci molto tempo per approfondire o smentire le informazioni ricevute, mentre i media e i social

⁸⁸ Censurando il numero reale di caduti, la resistenza Ucraina, i bombardamenti ed ogni altra informazione sulla crisi umanitaria in corso.

network internazionali agiscono sostanzialmente da megafono per la propaganda di entrambi i paesi, diffondendo⁸⁹, e poi eventualmente smentendo, un'enorme mole di notizie. La massa di fake news rilasciate dalla Russia, durante un preciso arco di tempo su specifici paesi, è paragonabile ad un vero e proprio attacco, il quale richiede misure controffensive; infatti, anche l'Ucraina fa massiccio uso dei mezzi di comunicazione a fini propagandistici, con dichiarazioni, videoconferenze, ma soprattutto attraverso post sui social network utilizzati al fine di diffondere foto e video della guerra e della crisi umanitaria in corso.

I social sono divenuti centrali in questa guerra, in quanto primo strumento di informazione e comunicazione utilizzato da persone comuni, politici e giornalisti, il presidente ucraino Zelenskij li utilizza per richiamare il proprio popolo all'unità e alla resistenza, i cittadini al fine di condividere in diretta le atrocità in atto nel loro paese, lo scopo comune è di mobilitare l'opinione pubblica, ricercare alleati e tentare di fermare la guerra.

Se i social sono divenuti centrali, anche i proprietari delle piattaforme lo sono, racchiudendo nelle loro mani importanti poteri, ad esempio, Facebook e Instagram, entrambi parte del gruppo Meta, hanno visto cambiare le proprie policy, permettendo in alcuni paesi la pubblicazione di post anti-russia, colmi di odio o minacce, decisione poi ritirata dai vertici del gruppo⁹⁰, ma le dinamiche dietro queste scelte sono ancora opache.

4.2 Prospettive future

Per quanto riguarda i moderatori, le condizioni in cui versano sono ancora, molto spesso, sfavorevoli, ma il gran numero di inchieste e pubblicazioni sull'argomento, le quali hanno raccolto svariate testimonianze di (ex) moderatori, hanno contribuito allo spostamento dell'attenzione di opinione pubblica, giornalisti e governi su questo delicato argomento, costringendo le aziende tecnologiche e le società intermediarie ad utilizzare un differente approccio in tema di moderazione di contenuti.

Le aziende puntano ancora molto solo sul versante retribuzione, ad esempio, come precedentemente esposto, Facebook si è mostrata particolarmente orgogliosa di aver

⁸⁹ Natali (2022).

⁹⁰ Natali (2022).

innalzato gli stipendi di tutti i suoi moderatori statunitensi, rendendoli più aderenti al costo della vita nelle varie zone dello stato, risultato al quale dovrebbe seguire anche l'obiettivo di un miglioramento nelle retribuzioni degli appaltatori stranieri, avvicinandole a determinati standard internazionali, offrendogli così condizioni lavorative più favorevoli. Anche se un aumento in busta paga può sicuramente far comodo, sono altri i versanti su cui le aziende dovrebbe agire al fine di salvaguardare la salute psicofisica dei propri lavoratori, particolarmente provati dall'attività svolta, cominciando ad esempio a ridurre i ritmi frenetici di lavoro, lasciando più flessibilità ai lavoratori, offrendo un serio e realmente utile supporto psicologico, ma soprattutto, riducendo il numero dei contenuti da visionare. Nonostante gli espedienti tecnici sperimentati dalle aziende, quali riduzione dell'audio o sfocature selettive, l'unica soluzione per migliorare realmente l'umore e la salute mentale dei lavoratori sarebbe regolamentare in maniera puntuale la coda di revisione di ogni moderatore, diminuendo il numero di contenuti ai quali sono costretti ad assistere ed allentando i rigidi sistemi di monitoraggio e controllo inseriti nei computer, dando loro la possibilità di fermarsi qualora si sentano eccessivamente scossi. Un'organizzazione del lavoro flessibile e realmente attenta ai bisogni dei lavoratori è forse l'unica soluzione possibile in questo contesto, nonostante l'incessante lavoro di quell'algorithm sempre bisognoso di sorveglianza, perché il benessere delle persone dovrebbe sempre essere messo al primo posto e non sacrificato al fine di correggere gli errori di una macchina impersonale, la quale ha causato fin troppi danni all'incolumità dei moderatori (e degli altri lavoratori digitali). Forse allora dovrebbe essere la correzione della macchina il punto nevralgico su cui dovrebbero concentrarsi le aziende, ad esempio attraverso l'inserimento di barriere all'entrata, impedendo la pubblicazione di contenuti vietati grazie a sistemi di intelligenza artificiale in grado di riconoscerli e bloccarli ancor prima che vengano condivisi, liberando di qualche contenuto l'infinita coda di revisione dei moderatori, sicuramente i punti di azione potrebbero essere tanti. Nonostante le tecniche di machine e deep learning utilizzate, reali risultati con riguardo all'accuratezza delle decisioni automatiche sono, per ora, lontani e il loro utilizzo è ancora limitato, vi è infatti da considerare che, anche

introducendo o apportando importanti miglioramenti all'algoritmo, i contenuti continueranno a presentare infinite sfumature di significato e svariati contesti, decifrabili (almeno per ora) solo dalle capacità intellettive e mentali umane, rendendo ancora indispensabili i moderatori.

5 Bibliografia

- Armano, E., Murgia, A., & Teli, M. (2017). *Platform Capitalism e confini del lavoro negli spazi digitali*. Milano: "Mimesis".
- Bruni, C., & Chemaly, S. (2016, aprile 13). *The Secret Rules of The Internet*. Tratto da "The Verge".
- Caine, A. (2017, dicembre 7). *The 11 Coolest Perks at Facebook, The Best Place to Work in America*. Tratto da "Insider".
- Chen, A. (23 ottobre 2014). *The Laborers Who Keep Dick Pics and Beheadings Out of Your Facebook Feed*. "Wired".
- Ciccarello, E. (2020, gennaio 28). *Io, moderatrice di Facebook all' lavoro contro l' odio. Salvini, protetto speciale*. "La via libera".
- Dubois, E., & Blank, G. (2018). *The myth of the echo chamber*. "Oxford Internet Institute".
- Dwoskin, E., Jeanne, W., & Regine, C. (2019, luglio 25). *Content Moderators at Youtube, Facebook and Twitter See the Worst of the Web - And Suffer Silently*. "The Washington Post" .
- Floridi, L. (2007). *A look into the future impact of ICT on our lives*.
- Formenti, C. (2011). *Felici e sfruttati. Capitalismo digitale ed eclissi del lavoro* . Milano: "Egea".
- Franchi, J. (2021). *Gli obsoleti. Il lavoro impossibile dei moderatori di contenuti*. Milano: "Agenzia X".
- Gilbert, D. (2020 , gennaio 9). *Facebook Is Forcing Its Moderators to log Every Second of Their Days - Even in the Bathroom*. "Vice".
- Gillespie, T. (2018). *Custodians of the Internet: Platforms, Content Moderation, And The Hidden Decisions That Shape Social Media*. "Yale University Press".
- Irani, L. (2015, gennaio 1). *Difference and Dependence Among Digital Workers: The Case of Amazon Mechanical Turk*. "South Atlantic Quarterly", 114(1).
- Kaplan, J. (2018). *Intelligenza artificiale: guida al futuro prossimo*. Roma : "LUISS".
- Khatib, H. A. (2019, ottobre 23). *Youtube Is Erasing History*. "The New York Times".
- Koebler, J., & Cox, J. (2018, settembre 24). *Content Moderator Sues Facebook, Says Job Gave Her PTSD*. "Vice".

- Murgia, M. (2020, gennaio 26). *Facebook Content Moderators Required to Sign PTSD Forms*. "Financial Times".
- Natali, G. (2022, marzo 24). *La disfatta dell'informazione sulla guerra in Ucraina*. "Atlante".
- Negroponte, N. (1995). *Being Digital*. Londra : " Hodder & Stoughton".
- Newton, C. (2019, dicembre 16). *The terror queue*. Tratto da "The Verge".
- Newton, C. (2019, febbraio 25). *The trauma floor. The Secret Lives of Facebook Moderators in America*. Tratto da "The Verge".
- Pariser, E. (2012). *Il Filtro: Quello che Internet ci Nasconde*. Milano: "Il saggiatore".
- Price, R. (2018, aprile 17). *The Median Pay at Facebook Is More Than \$240.000 a Year*. Tratto da "Insider".
- Punsmann, B. (2018, gennaio 6). *Three Month in Hell* . "Suddeutsche Zeitung".
- Quintarelli, S. (2019). *Capitalismo immateriale. Le tecnologie digitali e il nuovo conflitto sociale* . Torino: "Bollati Boringhieri".
- Sarra, C. (2019). *Il mondo-dato. Saggi su datificazione e diritto* . Padova : "Cleup".
- Solon, O. (2017, maggio 23). *To Censor Or Sanction Extreme Content? Either Way, Facebook Can't Win* . "The Guardian".
- Wilson, M. (2018, agosto 11). *The Hardest Job in Silicon Valley Is a Living Nightmare*. "Fast Company".
- Wray, D. D. (2018, giugno 26). *The Companies Cleaning the Deepest, Darkest Part of Social Media*. "Vice".