La responsabilità dei fornitori di servizio

Giuseppe Vaciago Docente di Data Ethics and Data Protection, Politecnico di Torino

1. Metaverso centralizzato o decentralizzato?

Il progresso tecnologico, da sempre, ma mai come negli ultimi 30 anni, ha cambiato il comportamento, le abitudini e l'esperienza degli utenti. Al principio, i primi computer "desktop" avevano tenuto gli utenti inchiodati alla scrivania. Con il passare del tempo, i dispositivi sono divenuti sempre più piccoli, maneggevoli, infine tascabili. Si è quindi generata la necessità di avere una connessione a internet, colmata con il Wi-Fi, in grado di stare al passo con questa rivoluzione. L'utente ha iniziato ad apprezzare la possibilità di utilizzare un computer in movimento. Gli schermi touch screen, allo stesso modo, hanno rappresentato il punto di rottura rispetto alla precedente modalità di interazione con i dispositivi.
Il metaverso cambierà nuovamente la modalità di interazione con il dispositivo e con gli altri utenti-avatar. Anzitutto, cambia l'approccio dell'utente con lo spazio. L'esperienza immersiva garantita dai visori stimola un'esperienza sensoriale del tutto particolare e nuova, che per la prima volta coinvolge tutto il corpo.
Nel metaverso viene meno la barriera dello schermo, che per certi versi conferiva la rassicurante sensazione di essere invisibili, protetti. Nel metaverso, l'avatar dell'utente viene catapultato nel pieno dell'azione, in mezzo a piazze virtuali che non sono più piazze in senso metaforico, ma che conducono a strade, parchi, palazzi in cui è possibile entrare, esplorare le stanze e interagire con le persone in tempo reale.
Infine, cambia l'approccio relazionale con gli altri utenti. La comunicazione scritta diventerà sempre più marginale, lasciando spazio a un'interazione verbale, tramite *voice-chat*, più simile a quella offline. I visori in commercio sono infatti muniti di un microfono integrato, che permette di esprimersi a voce previa l'abilitazione dell'apposita funzione.
Le novità appena descritte hanno delle conseguenze immediatamente percepibili sulla natura delle violazioni e sulla potenziale responsabilità delle piattaforme. Prima, tuttavia, di entrare nel merito di tali distinzioni va chiarito uno dei tanti profili definitori del metaverso.
Complice l'ingente investimento di Meta in questo settore, percepiamo il metaverso come un luogo virtuale accessibile solo con particolari dispositivi che ci consente di muoverci nei confini e sulla base di regole definite da un unico gestore. Da ciò sorge la preoccupazione di creare un metaverso unico e interoperabile che giustifica la nascita di iniziative come il Metaverse Standards Forum[1].
Se è indubbio che ci sia e ci sarà sempre di più una mancanza di interoperabilità fra i vari metaversi, si potrebbe ipotizzare uno scenario simile a quello dei sistemi operativi attualmente presenti nei dispositivi mobili: poco cambia che si abbia uno smartphone Android o Apple. L'importante è che gli applicativi in esso installati siano in grado di funzionare tra diversi sistemi operativi. La sfida dell'interoperabilità si poggia, in modo maggiormente realistico ed economicamente sostenibile, più su questi presupposti che su quelli del sogno di un unico metaverso.

[1] Il Metaverse Standards Forum riunisce associazioni, società ed esperti del metaverso per favorire la creazione di standard di interoperabilità che consentano l'interazione fra i vari metaversi. Maggiori informazioni sono disponibili qui: https://metaverse-standards.org/.

Ciò che, invece, sembra non essere preso in considerazione è l'abissale differenza tra metaverso "centralizzato" e metaverso "decentralizzato". Sotto questo profilo abbiamo da un lato il mondo centralizzato di Meta, Roblox o di Fortnite e dall'altro quello di Decentraland e Sandbox e tanti altri. Le differenze sono davvero tante: al netto dell'utilizzo delle criptovalute o degli NFT che potrebbero un domani anche essere presenti in modo massiccio nel metaverso "centralizzato", ciò che cambia è che nell'universo decentralizzato, tutto è "costruito, governato e in potere degli utenti" come esplicita chiaramente Decentraland per presentare il suo metaverso.

Questo tipo di differenza ha un impatto non indifferente nella concreta possibilità di regolamentare eventuali attività illecite come vedremo nel prosieguo. Fatta questa doverosa distinzione, è importante sottolineare le sostanziali differenze tra il web 2.0 e il web3 in termini di attività illecita.

In primo luogo, le nuove dinamiche in molti casi faranno progressivamente venir meno, o perlomeno diminuire, una caratteristica propria del contenuto illecito per come lo abbiamo conosciuto finora sul web bidimensionale: la sua natura scritta, con portata lesiva protratta nel tempo, almeno finché non venga rimossa dall'utente o dal *provider* – come nei casi ormai familiari del commento diffamatorio, del post sul social network, del video caricato su YouTube.

Nel caso della comunicazione verbale o della condotta illecita commessa nel metaverso da un utente tramite il proprio avatar, dunque, l'illecito torna ad avere natura istantanea come nel mondo offline. Tuttavia, ciò non deve portare a sminuire le condotte così perpetrate, che possono avere un effetto disturbante e minaccioso non inferiore a quelle che si presentano in forma scritta – addirittura, a seconda dei casi, superiore, proprio per il fatto di essere così dirette ed estemporanee. Con la notabile differenza che di tali contenuti non rimane una traccia che si possa rileggere, fotografare o riascoltare, salvo in casi particolari – ed è in fondo proprio questo l'obiettivo che si pone il metaverso: ricreare nel modo più fedele possibili le dinamiche relazionali del mondo "reale". Per moderare tali contenuti sarebbe necessario registrarli, con un inevitabile *trade-off* in termini di privacy[2]. Per contro, il *content moderator* che aspiri a rendere il proprio metaverso un ambiente sicuro dovrebbe imporsi dei tempi di intervento particolarmente tempestivi, in quanto l'intervento tardivo offrirebbe al soggetto leso una tutela parziale: quella della sanzione magari, del blocco dell'account del trasgressore, ma con la consapevolezza di non aver impedito l'illecito in tempo. Naturalmente, un intervento tempestivo rischia di ledere, in una certa misura, l'accuratezza nella valutazione dell'illecito, correndo il rischio di censurare contenuti leciti.

In sintesi, in questi casi le piattaforme dovranno gestire due difficili bilanciamenti nel moderare contenuti: in primo luogo, il rapporto tra la sicurezza dell'utente e la sua privacy; in secondo luogo, la scelta tra rapidità e accuratezza nella rimozione dei contenuti. La scelta tra questi diversi bilanciamenti comporterà anche una diversa competizione tra le piattaforme del metaverso in termini di qualità di contenuti da offrire verso i propri utenti.

In secondo luogo, si apre la strada a tutta una serie di fatti illeciti completamente nuovi, che non erano concepibili nel web 2.0, ma che lo diventano in qualche modo nel metaverso: tutti quelli che coinvolgono

[2] È questa la scelta operata da Roblox, che nei propri termini e condizioni si riserva di registrare le comunicazioni a voce per esclusive ragioni di sicurezza e di moderazione contenuti, e non permette di disabilitare questa funzione, se non rinunciando del tutto alla comunicazione voice-chat. V. F.A.Q. di Roblox, disponibili su https://en.help.roblox.com/hc/en-us/articles/5704050147604-Chat-with-Voice-Recording-Frequently-Asked-Questions

la nuova dimensione meta-corporea, o, in altri termini, la possibilità di sfruttare il proprio avatar per compiere gesti fisicamente violenti nei confronti di altri utenti: attacchi fisici, con o senza l'uso di armi; risse, stupri e altre violenze. Attacchi di questo tipo sono, invero, già avvenuti e riportati nella cronaca[3].

È intuitivo che un attacco fisico nel metaverso non può riverberarsi sul corpo in carne e ossa dell'utente che si trovi al sicuro nella propria casa. Tuttavia, è anche dimostrato come l'esperienza immersiva, in particolare grazie all'uso dei visori, induca in uno stato di suggestione tale da rendere tali episodi autenticamente pericolosi per la psiche di chi li vive. Anche in questi casi si pongono i problemi già sopra descritti della moderazione dei contenuti: da una parte quello dell'accertamento del fatto illecito (forse ha più senso parlare di fatto, in questo caso, che di contenuto), accertamento che potrebbe assumere tratti più simili a quelli del mondo offline – magari con la raccolta di testimonianze – salvo che si adotti la scelta di registrare gli eventi, con conseguente sacrificio in termini di riservatezza; dall'altro quello della necessaria tempestività dell'intervento.

Per altro verso, una volta accertati i fatti, si pone il problema dell'eventuale applicabilità delle fattispecie penali a quanto accade nel metaverso. È facile immaginare l'insorgenza di conflitti di giurisdizione, superati i quali, comunque, rimane il problema della concreta sussunzione dei fatti in norme giuridiche concepite per crimini aventi un forte elemento di materialità. Si ponga il caso della violenza sessuale, di cui già si registrano episodi[4]. L'ordinamento italiano, all'articolo 609 *bis* c.p., descrive la condotta nei seguenti termini: costringere taluno a compiere o subire atti sessuali, con violenza, minaccia o abuso di autorità. Tale formulazione potrebbe far pensare alla necessarietà di un contatto fisico tra l'autore e la vittima. Al tempo della redazione della norma, nel 1996, non si era prevista la possibilità di una commissione tramite mezzo elettronico. Per colmare il vuoto di tutela è intervenuta la Corte di Cassazione, che già a partire dal 2013 aveva ritenuto commesso il reato in oggetto da parte di un soggetto che, dietro minaccia, costringeva le vittime a inviare materiale pornografico[5]. È ragionevole pensare che il solco tracciato da tali pronunce possa spianare la strada al riconoscimento della punibilità di talune condotte a contenuto sessuale commesse nel metaverso; tuttavia, condotte diverse costringeranno inevitabilmente i giudici a nuovi sforzi interpretativi (nello specifico, condotte emulative di atti sessuali, ma commesse da avatar) onde evitare di cadere in quello che correttamente il compianto Prof. Sgubbi definiva il "Diritto Penale Totale"[6].

In terzo luogo, si pensi al tema della sostituzione di persona, che porrà dei problemi assolutamente nuovi e innovativi nel mondo del metaverso. Come già accennato, nel metaverso gli utenti interagiscono attraverso degli avatar, ossia personaggi virtuali personalizzabili. Tali avatar possono essere creati in modo che

---

[3] V. SumOfUs, *Metaverse: another cesspool of toxic content* (2022), contenente un report degli attacchi registrati nel metaverso Horizon Worlds di Meta, in https://www.sumofus.org/images/Metaverse_report_May_2022.pdf

[4] V. PATEL, *Reality or fiction?* su Medium, 2021, in cui l'autrice racconta una violenza sessuale personalmente subita nel metaverso di Meta: «entro 60 secondi dall'iscrizione, sono stata molestata verbalmente e sessualmente – 3 o 4 avatar maschili, con voci maschili, hanno violentato virtualmente il mio avatar e hanno scattato delle foto. Mentre cercavo di fuggire mi hanno urlato "non fingere che non ti sia piaciuto"». Patel riporta di essersi sentita spiazzata, congelata, e di non aver pensato immediatamente a effettuare il *logout* dalla piattaforma. «La realtà virtuale è stata progettata in modo che la mente e il corpo non possano distinguere le esperienze virtuali da quelle reali. In qualche modo, la mia risposta fisiologica e psicologica è stata come se fosse accaduto nella realtà». La tendenza a essere influenzati, anche nella vita reale, dalle esperienze vissute nella realtà virtuale prende il nome di effetto Proteus. Disponibile su https://medium.com/kabuni/fiction-vs-non-fiction-98aa0098f3b0.

[5] V. Cassazione penale, sez. III, sentenza n. 19033/2013, e Cassazione penale, sez. III, sentenza n. 17509/2018.

[6] F. Sgubbi, *Il diritto penale totale. Punire senza legge, senza verità, senza colpa*, Il Mulino, 2019.

rispecchino il proprio aspetto reale (la piattaforma di Meta, ad esempio, contiene un *tool* che permette di caricare una propria fotografia dalla quale generare un avatar somigliante) oppure possono essere di pura fantasia, dando sfogo alla propria creatività, in modo da prioritizzare la propria percezione di sé più che il proprio vero aspetto. A ciò si aggiunga che gli avatar possono essere modificati in ogni momento, oppure se ne possono creare varie versioni – ad esempio, una per le relazioni personali e una per quelle commerciali. Tutto questo pone il problema della conoscenza della reale identità di coloro con cui si interagisce. La questione, nei suoi tratti più generali, non è nuova rispetto all'internet tradizionale, dove già era possibile celarsi dietro nickname di fantasia, ma assume tratti decisamente più marcati nel metaverso, in parte perché esiste la possibilità di creare degli avatar graficamente rassomiglianti alle persone fisiche, creando una suggestione ben maggiore rispetto a un semplice nickname o a una firma, e in parte perché è legittimo aspettarsi dagli utenti catapultati in questa nuova realtà una maggiore ingenuità. Questo fenomeno potrebbe dunque essere sfruttato per indurre in errori di valutazione o, nel peggiore dei casi, per architettare raggiri e condotte illecite. L'interrogativo che ci si pone, e a cui non si sa ancora offrire una risposta precisa, è se si possa configurare il reato di sostituzione di persona anche in caso di avatar "sostituiti". L'ordinamento italiano prevede a tal fine la fattispecie di cui all'articolo 494 c.p., ma potrebbero venire in rilievo anche le fattispecie di truffa (articolo 640 c.p.) o il trattamento illecito di dati (articolo 167 del Codice Privacy).

2. Il DSA e il DMA nel contesto del metaverso

Tali regolamenti, lungamente attesi, si affacciano nel panorama legislativo in un momento particolarmente delicato, quello in cui il mondo digitale minaccia di essere rivoluzionato dal metaverso. Il quesito da porsi è se le neonate norme possano reggere la prova di resistenza a cui il metaverso le sottoporrà, o se viceversa siano destinate a rivelarsi obsolete già dal momento della loro entrata in vigore.

I profili di maggior impatto sul metaverso di queste due normative saranno presumibilmente le seguenti cinque.

In primo luogo, il cambio di paradigma da *notice and take down* a *notice and action* avrà un impatto importante in una realtà in cui le necessità di rimozione di contenuti possono avere caratteristiche diverse rispetto al mondo tradizionale del Web. Tale apertura permetterà di modulare la reazione del prestatore del servizio in modi diversi e ritagliati su misura sul caso specifico. La disabilitazione temporanea dei contenuti, ad esempio, potrebbe essere di grande utilità per la gestione di tutte quelle situazioni in cui è necessario un approfondimento prima di arrivare ad una conclusione definitiva circa l'illiceità o meno dei contenuti. In taluni casi, invece, la reazione più appropriata potrebbe essere la segnalazione alle forze dell'ordine, oppure l'acquisizione di prove forensi utili a identificare l'autore dell'illecito.

In secondo luogo, gli obblighi supplementari in materia di trasparenza della pubblicità online giocheranno un ruolo fondamentale in un settore delicatissimo e sempre più in fermento. Se è vero che stiamo andando verso una monetizzazione del dato personale[7], e conseguentemente verso la fine del modello per cui i dati personali sono conferiti alle piattaforme unicamente come corrispettivo di un servizio reso da queste ultime,

[7] Ossia verso una cessione volontaria da parte dell'utente dei propri dati personali (utilissimi alle piattaforme per finalità commerciali) dietro remunerazione economica, spesso corrisposta in forma di criptovaluta.

sarà ancora più importante rendere edotti gli utenti di come verranno gestite le informazioni pubblicitarie all'interno delle piattaforme.

In terzo luogo, le norme in materia di *accountability* e di valutazione dei rischi sistemici significativi derivanti dal funzionamento e dall'uso dei servizi delle piattaforme di metaverso, di cui agli articoli 26 e seguenti, potrebbero divenire una grande opportunità per monitorare attività che in passato hanno portato a situazione estremamente complesse, come, ad esempio, il caso "Cambridge Analytica". La vera sfida sarà comprendere con quale metodologia svolgere tali analisi in un contesto come quello del metaverso. Se, tuttavia, dovessimo mutuare dalla normativa in materia di protezione dei dati personali, si ritiene che un ruolo importante potrebbe essere rappresentato dall'individuazione delle misure a protezione dell'utente virtuale che oggi il metaverso sta iniziando a creare, ma che dovrà rafforzare sia da un punto di vista legale che etico.

In quarto luogo, il tema del "segnalatore attendibile" o "*trusted flagger*" rappresenta una sfida rilevante per il metaverso. A tal riguardo, è interessante domandarsi quale sarà il ruolo degli Stati nelle piattaforme, sia con riguardo alla possibilità di somministrare servizi, sia in relazione alla presenza digitale delle forze dell'ordine, in ottica special-preventiva. Non si può escludere, infatti, che si istituzionalizzi una presenza statuale – di veri e propri agenti di polizia, in borghese e/o in divisa, eventualmente munita di permessi speciali rispetto ai comuni utenti – che possa essere chiamata a intervenire nei casi di violazioni più eclatanti. Alcuni hanno già ipotizzato l'idea di un poliziotto virtuale che vaga nei vari luoghi del metaverso alla ricerca di potenziali illeciti. L'idea è certamente affascinante, e ci avvicina sempre di più all'immaginario più fantascientifico e "asimoviano" del metaverso, ma al tempo stesso non appare più peregrina. La complessità del metaverso, a maggior ragione quando l'utenza attiva raggiungerà numeri ragguardevoli, potrebbe imporre la presenza di "sentinelle" in grado di intercettare comportamenti illeciti. Il vero problema sarà comprendere se tali sentinelle dovranno essere persone fisiche o software di intelligenza artificiale.

In quinto e ultimo luogo, rimane da chiarire uno dei dilemmi cruciali del metaverso e citato in premessa, ossia se ci si dirige verso una privatizzazione di tali ecosistemi, mediante la creazione di protocolli chiusi e *hardware* a uso esclusivo delle (varie) singole piattaforme, o se, viceversa, si riuscirà a indirizzare il mercato verso la creazione di un unico protocollo all'interno del quale i vari *player* possano inserire la loro soluzione. La capacità di trasportare dati e asset virtuali da una piattaforma all'altra, è molto importante in un panorama frazionato in molteplici metapiattaforme, dove gli utenti sono portati a spendere somme di denaro anche cospicue per acquistare prodotti e servizi – in particolare da quando gli NFT sono divenuti un aspetto integrante delle transazioni del mondo digitale[8]. Poter trasportare i propri contenuti tra le varie piattaforme permetterebbe agli utenti di intraprendere progetti – anche a carattere socio-culturale – a cavallo tra diverse piattaforme, creando un ecosistema fluido ove l'esperienza dell'utente non sia ostacolato da fastidiose soluzioni di continuità. Ridurrebbe i c.d. *switching cost* degli utenti, favorendo la concorrenza tra i fornitori. Il DMA ambisce a evitare questa deriva. Tra le previsioni a favore dell'interoperabilità ricordiamo l'articolo 6, paragrafo 1, lettera e) che impone ai *gatekeeper* di astenersi dal rendere tecnicamente difficile per gli utenti passare a servizi o operatori concorrenti, e la lettera h), che impone di

[8] MANGADA REAL DE ASÙA, OTTER, TSUKUDA, VIVENOT, *The Metaverse Challenges and Regulatory Issues*, student paper published on SciencesPo, 2022

garantire l'effettiva portabilità dei dati generati dall'utente commerciale o finale, e impone ai *gatekeeper* la condivisione delle API, ossia delle interfacce di programmazione, i protocolli che permettono ai diversi operatori di trovare un linguaggio informatico comune.

Prevedere il futuro non è mai semplice, ma un dato è certo: queste due normative, con particolare riferimento al DSA, sono state pensate in un contesto "centralizzato". Il vero quesito è quindi capire se potranno avere valore sotto il profilo dell'*enforcement* normativo anche nel contesto "decentralizzato" dove attualmente una realtà come il gruppo hacker russo "Wizard Spider", noto per aver guadagnato 150 milioni di dollari con il ransomware Ryuk, ha un suo legittimo spazio in Digitible[9] promuovendo petizioni improbabili sulla necessità di scarcerare pericolosi cybercriminali[10].

[9] Digitible è un "ufficio virtuale" presente in decentraland dove chi vuole può presentare le petizioni che desidera per promuovere le più varie iniziative culturali, sociali e politiche. Maggiori informazioni qui: https://digitible.com/.

[10] TARA ANNISON, *The Future of Financial Crime in the Metaverse*, in Elliptic Metaverse Report 2022, disponibile a questo indirizzo: https://www.elliptic.co/hubfs/Crime%20in%20the%20Metaverse%202022%20final.pdf