

**UNIVERSITÀ DEGLI STUDI DI MILANO-BICOCCA**

Scuola di Economia e Statistica

Corso di laurea in

**STATISTICA E GESTIONE DELLE INFORMAZIONI**



**ANALISI DEL METAVERSO: come è percepito dagli utenti del  
social media Twitter**

Relatore: Prof. Roberto Boselli

Tesi di Laurea di:

Aurora Musitelli

Matr. N. 856741

Anno Accademico 2021/2022



# Indice

<b>Introduzione.....</b>	<b>1</b>
 <b>Capitolo 1 - Evoluzione tecnologica fino a raggiungere il Metaverso.....</b>	<b>3</b>
1.1 Le quattro fasi della storia del Web (World Wide Web).....	3
1.2 L'innovazione e storia dei social media.....	11
1.2.1 Caratteristiche social network.....	16
1.3 Nuovo modo di vedere il Web: Realtà Virtuale.....	18
1.3.1 Definizione di Metaverso.....	24
1.3.2 Ambiti applicativi e ambienti di sviluppo del Metaverso.....	27
 <b>Capitolo 2 - Social Media Analytics: strumenti e analisi.....</b>	<b>31</b>
2.1 Richiami teorici: Text Mining per il trattamento dei dati non strutturati.....	33
2.2 Metodologia e domanda di ricerca.....	35
2.2.1 Raccolta dati e pre-processamento dei tweet.....	36
2.3 Sentiment Analysis: polarità dei tweet e risultati.....	42
2.4 Algoritmi Unsupervised Machine Learning: Topic Modelling.....	49
 <b>Capitolo 3 - Conclusioni e sviluppi futuri.....</b>	<b>55</b>
 <b>Bibliografia.....</b>	<b>61</b>
 <b>Ringraziamenti.....</b>	<b>62</b>

# Introduzione

Le innovazioni tecnologiche e informatiche rappresentano parti fondamentali nella vita di tutti i giorni poiché modificano e arricchiscono le comunicazioni, le abitudini lavorative e le interazioni sociali. Nel corso degli anni della storia di Internet sono state registrate tre grandi ondate di innovazione tecnologica in particolare, la prima fase riguarda la creazione della rete Internet il World Wide Web (comunemente chiamato Web) permettendo agli utenti di connettersi e svolgere le prime ricerche sul Web, la seconda fase relativa alla rivoluzione della comunicazione interpersonale con l'introduzione dei primi social media (Web 2.0.) ed infine il Web 3.0. ovvero la terza fase evolutiva riguardante il potenziamento del Web come un database (Web Database) capace di adattarsi a qualsiasi dispositivo elettronico presente sul mercato; attualmente si sta attraversando la quarta ondata di innovazione tecnologica (Web 4.0.) ancora in forte espansione, un nuovo modo di intendere la comunicazione attorno ai nuovi dispositivi come la Realtà Virtuale (Virtual Reality). In particolare, quest'ultima, fa riferimento alla multi-presenza spaziale, ambientale e sociale nei mondi virtuali, permettendo la creazione di un ambiente immersivo per connettere gli utenti in una realtà nuova e in via di sviluppo. Per tale motivo, il lavoro di tesi svolto si inserisce all'interno di una ricerca di carattere analitico volta a rispondere ad una specifica domanda di ricerca: ovvero comprendere se la Realtà Virtuale, in particolar modo il termine Metaverso (Metaverse), è stato compreso dal grande pubblico e quale tipo di percezione si è creata negli utenti che utilizzano quotidianamente i social media. Nello specifico l'idea è stata quella di dare valore alla conoscenza proveniente da dati non strutturati, presentando alcune metodologie di classificazione del testo (Text Mining) applicate alle opinioni che vengono espresse tramite dei tweet (in questo caso tweet tutti in lingua inglese), ovvero messaggi condivisi dagli utenti sulla piattaforma Twitter, per far emergere il pensiero del pubblico relativo alla tecnologia del Metaverso. In particolare, il social media Twitter può essere visto come una ricca fonte di informazioni poiché gli utenti comunicano ciò che pensano e lo fanno sapere "a tutti" fornendo dati utili da analizzare. Si ritiene quindi utile presentare in sintesi come è stato strutturato il presente lavoro di tesi.

Il primo capitolo “Evoluzione tecnologica fino a raggiungere il Metaverso” è stato suddiviso in tre paragrafi relativi alla parte teorica, in cui si ripercorre in maniera esaustiva lo sviluppo tecnologico dall’invenzione di Internet (World Wide Web) fino ad oggi, per mostrare l’importanza dell’era digitale e comprendere ciò che ha portato allo sviluppo della Realtà Virtuale e del Metaverso.

Il secondo capitolo “Social Media Analytics: strumenti e analisi” ha lo scopo di descrivere gli strumenti teorici e analitici, di analisi del testo (Text Mining), che sono stati utilizzati per rispondere alla domanda di ricerca iniziale.

Il terzo capitolo “Conclusioni e sviluppi futuri” si concentra sulle conclusioni per confrontare i risultati ottenuti delle analisi svolte e sulle considerazioni/prospettive future della tecnologia del Metaverso.

# Capitolo 1

## Evoluzione tecnologica fino a raggiungere il Metaverso

L'ambiente digitale si è diffuso nelle più disparate attività umane e per certi versi sta diventando quasi totalizzante; questa situazione è dovuta da numerose scoperte e invenzioni sviluppate circa un secolo fa, ma che negli ultimi anni, grazie ad una maggiore consapevolezza tecnologica, hanno portato a concretizzare la Realtà Virtuale e il Metaverso. Questi strumenti stanno prendendo sempre più spazio non solo nella comunicazione e nell'intrattenimento ma anche in contesti aziendali; quindi si ritiene fondamentale ripercorrere l'evoluzione tecnologica e le grandi fasi della storia di Internet per comprendere che il concetto "Metaverso", in realtà, è già stato trattato nell'immaginario dei primi romanzi del '900 e reso concreto e utilizzabile solo recentemente.

### 1.1 Le quattro fasi della storia del Web (World Wide Web)

Agli albori dei primi sviluppi dell'ambiente digitale e del Web, era difficile prevedere quello che sarebbe successo negli anni successivi; ma ben presto, considerando la progressiva diffusione di Internet nella società, si è osservato come in pochi decenni questa evoluzione abbia profondamente modificato il modo di comunicare degli individui cambiando le abitudini sia relazionali che lavorative. L'invenzione di Internet non rappresenta un'illuminazione improvvisa da parte di singoli individui, ma è frutto del risultato di un lungo e tormentato percorso, a cui hanno contribuito più ricercatori. Per tale ragione si possono identificare tre grandi fasi storiche caratterizzate da diverse invenzioni che hanno ampliato la gamma delle forme di comunicazione mediata dai dispositivi elettronici. La prima fase pionieristica che ha scatenato l'invenzione di Internet si può attribuire ai primi successi spaziali della Russia contro gli USA per la conquista dello spazio. Gli americani, per rispondere al primo successo Russo per il lancio dello Sputnik nello spazio risalente al 4 ottobre 1957, decisero di creare nel febbraio del 1958 il progetto ARPA (Advanced Research Projects Agency)<sup>1</sup>, progetto creato all'interno del Dipartimento di

---

<sup>1</sup> Ryan J., "Storia di Internet e il futuro digitale".

Difesa degli Stati Uniti, con la collaborazione di alcune università con lo specifico scopo di:  
*“Provide within the Department of Defense an agency for the direction and performance of certain advanced research and development projects.”*<sup>2</sup>

In questo modo si arrivò al primo collegamento in rete tramite un elaboratore del laboratorio dell'università di Los Angeles U.C.L.A. (University of California Los Angeles) e un calcolatore collocato all'istituto di ricerca di Stanford (Stanford Research Institute) nel 1969 e successivamente alla prima dimostrazione pubblica di ARPANET nel 1972, ovvero la prima rete operativa di comunicazione. La rete ARPANET, rinominata successivamente all'inizio del 1980 con il nome di Internet, continuò il suo processo di sviluppo all'interno dell'istituto di ricerca ARPA e, grazie ad una serie di innovazioni come la diffusione dei personal computer, portarono al primo e vero utilizzo della rete su larga scala: attraversando l'invenzione della posta elettronica e delle prime banche dati di messaggi che prevedevano l'interazione tra gli utenti, i BBS o Bulletin Board System (un sistema telematico in grado di consentire a computer di accedere ad un elaboratore centrale per condividere o prelevare risorse), fino a giungere all'invenzione nel 1989 da parte del fisico Tim Berners Lee del World Wide Web o comunemente chiamato Web. Fin da piccolo Tim Berners Lee aveva coltivato delle idee per rendere il computer più accessibile, difatti dopo la sua laurea in fisica, ha avuto l'opportunità di lavorare presso il CERN (Organizzazione europea per la ricerca nucleare) iniziando ad immaginare il Web: *“Suppose all the information stored on computers everywhere were linked. Suppose I could program my computer to create a space in which anything could be linked to anything.”*<sup>3</sup>

Nacque quindi il suo progetto marginale chiamato ENQUIRE, un sistema di sottorete di Internet che permetteva a qualsiasi utente di accedere alla consultazione delle informazioni circolanti in rete con una navigazione semplificata raggiungibile con “collegamenti ipertestuali” chiamati hyperlinks. La navigazione è stata resa semplificata per far risultare comprensibile a chiunque un nuovo approccio di ricerca delle informazioni che comunemente venivano trovate sui manuali. Trascorsero pochi anni quando nel 1989, Tim Berners Lee, propose la condivisione di documenti segnando la nascita di HTML che, negli anni successivi, avrebbe raffinato per creare il primo browser: Nexus al quale seguirono negli anni successivi Safari per Mac ed Explorer per

---

<sup>2</sup> William B. B., Windham P., Van Atta R., “The Darpa model for transformative technologies”;

<sup>3</sup> Tim Berners-Lee, “Weaving the Web - The original design and ultimate destiny of the World Wide Web”.

Windows. Negli anni Novanta iniziò la fase non pionieristica della storia di Internet, definita successivamente “Web 1.0” rinominato come “only read Web”, caratterizzato dalla presenza di siti realizzati in HTML in prevalenza “statici” dando la possibilità agli utenti di navigare in Internet tramite i motori di ricerca, permettendo la visualizzazione di documenti ipertestuali, ma senza poter interagire in modo attivo con essi; questo metodo venne definito approccio “top-down” ovvero processi di comunicazione verticali e gerarchici, in cui solo i webmaster avevano le competenze tecniche necessarie e gli strumenti per poter aggiornare le pagine di un sito Internet. Per interagire all’interno delle prime comunità virtuali, che erano basate sul sistema dei BBS, gli utenti dovevano accedervi attraverso un computer, in questo modo, iniziarono a diventare attivi sul Web. Nel corso del decennio, sono aumentati gli utenti che hanno fatto uso della posta elettronica, delle chat, dei forum di discussione e verso la fine degli anni ‘90, anche dei blog: queste rappresentavano tutte forme di comunicazione che offrivano la possibilità di esporre le proprie opinioni personali e soprattutto permettevano di tenersi in contatto con amici e familiari. Successivamente dagli anni 2000 si raggiunge una fase della storia di Internet definita come una vera e propria rivoluzione: “la rivoluzione del Web 2.0” definita anche “read-write Web”. La nota espressione “Web 2.0” è stata fondata durante una conferenza rivolta al mondo del business e delle ICT (Information and Communication Technology) nell’ottobre 2004 dall’editore irlandese Tim O’Reilly, per sottolineare il cambiamento da un Web “statico” (Web 1.0.) strutturato da siti fondati su una comunicazione unidirezionale, ad un Web prettamente “dinamico” (Web 2.0.) basato su applicazioni che permettevano agli utenti di interagire fra di loro. Quello che affermò Tim O’Reilly è la seguente frase: *“Web 2.0 is a set of economic, social, and technology trends that collectively form the basis for the next generation of the Internet - a more mature, distinctive medium characterized by user participation, openness, and network effects.”*<sup>4</sup>

Quello che emerge dalla celebre frase di Tim O’Reilly è la partecipazione da parte dell’utente nell’interagire con i contenuti che si trovano sul Web, per questo lui parlò di rivoluzione rispetto al paradigma precedente; questo passaggio da Web 1.0 a Web 2.0, difatti, evidenzia due fattori principali: Il primo aspetto è stato il successo dei social network, tra gli utenti della rete si è notato un elevato apprezzamento che ha portato a pensare ad un cambiamento irreversibile nella

---

<sup>4</sup> Tim O'Reilly risponde alla domanda:

"What's next for Web2.0?" <https://www.oreilly.com/pub/a/Web2/archive/what-is-Web-20.html>.



comunicazione interpersonale. Il secondo aspetto invece è rappresentato dalla diffusione su larga scala di dispositivi mobili, come smartphone, tablet e laptop, questo ha accelerato notevolmente il successo della nuova comunicazione via Web. Infatti, con l'introduzione dei linguaggi di programmazione dinamici gli sviluppatori hanno permesso all'utenza non tecnica di interagire con i contenuti dei siti Internet e quindi gli utenti della rete non solo hanno avuto l'opportunità di usufruire dei contenuti presenti sul Web (compresi anche quelli dei media tradizionali), ma anche di modificarli, commentarli e condividerli con gli altri utenti; è considerata la prima volta in cui si è dato importanza all'utilizzo e al modo di condividere i contenuti sul Web. Questa fase infatti è stata caratterizzata dall'aumento di "traffico" attivo online da parte degli utenti grazie all'espansione di forme di comunicazione preesistenti come l'invenzione e la diffusione sul Web di piattaforme di condivisione di contenuti, come YouTube, Flickr, Vimeo; di siti Wiki ovvero software di scrittura collettiva, come Wikipedia e soprattutto di nuove piattaforme sociali come reti di relazioni social come Facebook, Twitter e Instagram create con lo scopo di mettere le persone in comunicazione. Si sono così diffusi processi di comunicazione orizzontali, non gerarchici e "dal basso" chiamati "bottom-up", che hanno modificato i processi di comunicazione tradizionalmente intesi nel paradigma precedente, dando una visione completamente diversa del Web. I fattori di maggior interesse, dopo questa rivoluzione, sono stati i diversi approcci economici, tecnologici, comunicativi, culturali e lavorativi che hanno seguito una rete di scambio di informazioni non più rigida e verticale ma molto più flessibile e orizzontale; a partire dallo scambio delle informazioni tra le élite dominanti delle nazioni fino alla comunicazione quotidiana degli utenti della rete. Nel corso degli anni, lo sviluppo di Internet, ha subito un'accelerazione e non ha smesso il suo processo di evoluzione; nel 2010, il Web aveva già preso una nuova formulazione il cosiddetto "Web 3.0", l'epoca del "read-write-execute Web" anche denominato Web Semantico, termine coniato da Tim Berners-Lee, indicando la trasformazione del Web in un ambiente in cui si sottolinea l'importanza dei significati dei contenuti fruibili in rete. La differenza in questo caso non è l'introduzione di nuove tecnologie, ma l'introduzione di diversi fattori, quali<sup>5</sup>:

---

<sup>5</sup> Lombardi M., Giorgi D., Fiesoli I., "I territori connessi: L'impatto dei nuovi modelli open Innovation nelle dinamiche imprenditoriali e territoriali".

- **La rete come database (Web Database):** l'introduzione, prima degli RSS (Really Simple Syndication) uno dei più popolari formati per la distribuzione di contenuti Web e dei file XML, poi dei 'metadata' che offrono la possibilità di utilizzare Internet come un enorme database.
- **Web semantico:** l'introduzione dei contenuti correlati a determinate parole chiave, che permettono la ricerca di informazioni più appropriate dando importanza ai contenuti;
- **Web adattabile ai diversi dispositivi elettronici:** il Responsive Web Design e la grafica vettoriale hanno dato un'accelerazione per poter visualizzare, su diversi dispositivi, il contenuto presente sulla rete;
- **Web potenziato:** capace di "plasmare" la società grazie ai social network, la potenza dei mezzi di comunicazione digitale permettono uno scambio di informazioni tra azienda/consumatore e artista/fan;
- **Web a tre dimensioni:** le nuove tecnologie e l'elevato accesso alla rete Internet hanno permesso di replicare la realtà in formato digitale. Si può quindi accedere alla rete ed effettuare una buona parte delle interazioni che compongono la vita reale;
- **L'intelligenza artificiale:** rappresenta una disciplina appartenente all'informatica che si occupa di studiare le metodologie e le tecniche di calcolo per consentire la progettazione di sistemi hardware e programmi software in grado di simulare la capacità e il comportamento del pensiero umano<sup>6</sup>. Si hanno diversi esempi di programmi evoluti che vanno verso questa direzione dagli algoritmi di Google, che analizzano la rete per comprendere come posizionare i contenuti a seconda della qualità e della pertinenza in base a determinate parole chiave, all'algoritmo di Facebook che presenta notizie degli amici in linea con il pensiero dell'utente.

Dal 2020 fino ad oggi si è iniziato a discutere di una nuova forma di Internet ancora più immersiva; infatti, il fenomeno digitale sembra correre sempre a più elevate frequenze trasformando il paradigma nel "Web 4.0.", in particolare è in collegamento con il concetto di Industria 4.0 in cui la tecnologia sta diventando automatizzata grazie all'Intelligenza Artificiale capaci di aiutare e supportare l'avanzamento del processo produttivo di molte realtà aziendali. Tutto ciò che ruota intorno a questa nuova forma di Web ha avuto un punto di partenza, i

---

<sup>6</sup> Boldrini N., "Cos'è l'Intelligenza Artificiale, perché tutti ne parlano e quali sono gli ambiti applicativi".

“Big Data”<sup>7</sup>: oggi giorno vengono generati una varietà di dati in formati differenti ad una velocità che non mostra segni di rallentamento poiché ogni azione quotidiana lascia una traccia digitale. In particolare, i Big Data sono così chiamati per cinque caratteristiche che gli sono state attribuite le cosiddette cinque V: Varietà (dati con formati differenti); Velocità (dati generati e scambiati velocemente); Volume (principale caratteristica dei Big Data, vengono generati enormi quantità di dati); Veridicità (estrapolare le informazioni da dati puliti e realistici); Valore (osservare in maniera quantitativa i fenomeni della realtà per prendere delle decisioni attraverso i dati). La loro diffusione è esplosa poiché è diventato necessario estrapolare le informazioni non solo da dati di tipo numerici ma anche da dati testuali (ovvero dati non numerici). Quello che è successo è che le organizzazioni si sono aperte all’utilizzo di dati di terze parti, ovvero dati che provengono da altre sorgenti informative, in questo modo si sono create delle problematiche di utilizzo dei dati poiché per utilizzare questa sorgente dati risulta necessario avere il controllo del processo della costruzione del dato, si deve conoscere la struttura e avere inoltre specificità sul dato. In particolare, per poter comprendere i Big Data risulta necessario spiegare le tre tipologie di dati esistenti<sup>8</sup>: dati strutturati, dati semi-strutturati e dati non strutturati. I dati strutturati rappresentano tutti quei dati che si possono memorizzare in un database SQL (Structured Query Language), restituiscono informazioni organizzate che vengono caricate in database relazionali. Invece i dati non strutturati non sono inseriti in uno schema o tabella, per esempio un foglio di calcolo o a un database relazionale (per esempio sono dati in formato testuale). Infine, i dati semi-strutturati rappresentano una tipologia di dati strutturati che contengono tag o indicatori per separare elementi semantici (ad esempio dati in formato XML, JSON e file CSV). Una volta che si è compreso in che formati si possono presentare i dati digitali, è possibile delineare i diversi paradigmi che si sono sviluppati nel tempo per gestire e manipolare questa tipologia di dati, oltre ai database SQL si sono poi sviluppati i database NoSQL ovvero “Not Only SQL”, paradigma in cui non si utilizza uno schema relazionale fisso ma è lo schema che si adatta al dato e non viceversa quindi lo schema risulta “flessibile”; sono stati appositamente realizzati per modelli di dati specifici e per applicazioni che necessitano grandi volumi di informazioni. In questo modo, con lo sviluppo di ingenti quantità di dati, l’Industria 4.0 sta prendendo sempre più piede grazie alla disciplina dell’Intelligenza Artificiale (AI) che si spiega con la maturità tecnologica che è

---

<sup>7</sup> V. Cosenza, "Social media ROI";

<sup>8</sup> Atzei P., Ceri S., Fraternali P., Paraboschi S., Torlone R., “Basi di dati”.

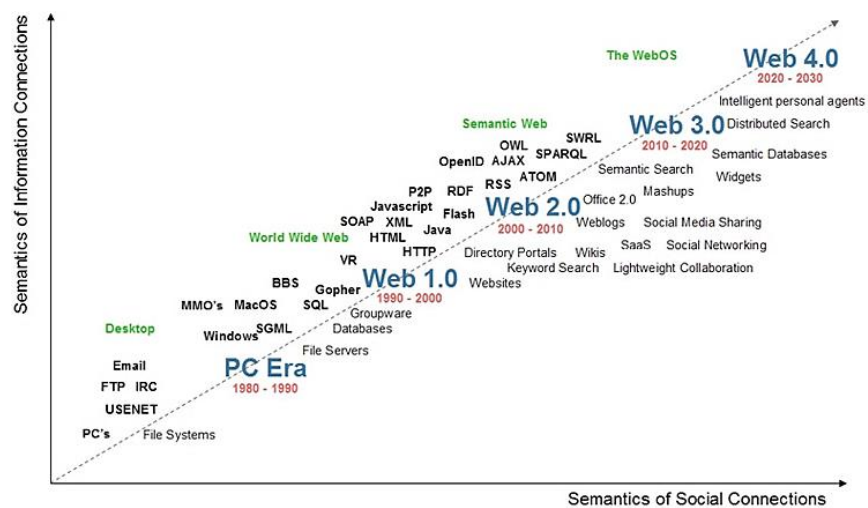
stata raggiunta sia nel calcolo computazionale, sia nella capacità di analisi in tempo reale ed in tempi brevi dei Big Data. In particolare, il primo e vero progetto di AI risale al 1943 quando la comunità scientifica inizia ad interessarsi all'Intelligenza Artificiale: due ricercatori proposero il primo neurone artificiale (Artificial Neural Network o ANN). A questa nuova ricerca seguì lo studio dello psicologo canadese Donald Olding Hebb, che diede il suo contributo per analizzare in dettaglio i collegamenti tra i neuroni artificiali ed i modelli complessi del cervello umano<sup>9</sup>: i primi veri prototipi funzionanti di reti neurali, arrivarono verso la fine del 1950, algoritmi matematici creati per riprodurre il meccanismo dei neuroni biologici. Questo generò l'interesse della comunità scientifica che aumentò gradualmente grazie al giovane londinese Alan Turing, un matematico considerato uno dei padri dell'informatica, che in quegli anni stava elaborando una spiegazione di come un computer potesse agire, a livello logico, come un essere umano. Qualche anno dopo John McCarthy, informatico statunitense, conia nel 1956 il termine "Artificial Intelligence" in occasione di un convegno di informatica al Dartmouth College negli USA al quale presero parte i maggiori esponenti dell'informatica. Nel convegno furono presentati alcuni programmi in grado di effettuare ragionamenti logici legati alla dimostrazione di teoremi matematici. Gli anni a venire furono caratterizzati da una grande spinta intellettuale e sperimentale in cui le maggiori università e le società informatiche (tra cui l'IBM) puntarono alla ricerca di nuovi programmi software che potessero essere in grado di pensare e agire come gli esseri umani. Dal 1970 in poi la sfida diventò quella di cercare di riprodurre software in grado di ragionare e prendere decisioni in base alle differenti situazioni, nella pratica si sarebbe dovuto tradurre in un linguaggio che permettesse di programmare le diverse possibilità di scelta previste dal ragionamento umano. In quegli anni, questo step si dimostrò particolarmente complesso, tanto che la ricerca subì un brusco rallentamento per poi riprendersi tra il 1980 e il 1990 con l'ingresso sul mercato dei processori grafici Gpu (Graphics Processing Unit) chip in grado di elaborare dati in maniera molto rapida, supportare processi complessi operando a frequenze basse e consumando meno energia rispetto alle Cpu. Il primo grande successo dell'AI risale al 1997 quando si osservò il confronto tra Deep Blue, una macchina realizzata dalla IBM, ed il campione di scacchi russo Garry Kasparov. I continui aggiornamenti aggiunti al sistema di ragionamento della macchina Deep Blue<sup>10</sup> permisero la sua vittoria e dimostrarono di aver

---

<sup>9</sup> Hebb, D. O., "Drives and the CNS (Conceptual Nervous System)" in Psychological review;

<sup>10</sup> Boldrini N., "Cos'è l'Intelligenza Artificiale, quali sono gli ambiti applicativi".

sorpassato il livello di creatività e di conoscenze del giocatore. Inoltre, negli anni con l'ingresso dei processori Gpu, si sono raggiunti enormi benefici in termini di efficienza e di potenza di calcolo fino allo sviluppo nell'ultimo decennio delle nanotecnologie, per emulare le funzioni sensoriali e cognitive del cervello umano. Dal precedente viaggio storico si può capire che il tema dell'AI e quindi del Web 4.0 è diventato sempre più condiviso e sta permettendo la costruzione di robot guidati dall'Intelligenza Artificiale rendendo noto il tema dell'automazione che porta un fattore di efficienza al sistema produttivo aziendale e non solo. I fattori che stanno conducendo verso questa evoluzione tecnologica sono, più in generale, l'espansione nel campo della robotica, nel campo biomedico, della sorveglianza e di tanti altri ambiti a cui possono essere applicate queste nuove tecnologie. Questa nuova fase del Web, quindi, è legata agli oggetti ovvero "Internet of things" (IoT): gli oggetti quotidiani che sono connessi alla rete diventano riconoscibili e "intelligenti" permettendo la condivisione di dati utili, relativi a sé stessi e, a proposito di questi oggetti, risulta interessante la teoria avanzata dall'economista statunitense Jeremy Rifkin di una società futura "a costo tendente allo zero": egli sostiene che l'espansione di Internet delle cose per cui miliardi di sensori collegati a ogni apparecchio e dispositivo elettronico unirà ogni persona e ogni cosa in un'unica rete neurale, proiettando l'essere umano verso una "Terza rivoluzione industriale": un nuovo sistema economico basato sulla produzione e la condivisione dei beni comuni (come l'energia e una serie di oggetti realizzati con stampanti 3D a costi marginali tendenti allo zero) da parte dei prosumer (i produttori-consumatori). In questo modo il "valore di scambio" nel mercato sarà sostituito dal "valore della condivisione".



**Figura 1.1** Sequenza temporale evoluzione del Web, fonte Radar Network & Nova Spivak

## 1.2 L'innovazione e storia dei social media

A pari passo con l'evoluzione del Web e dei dispositivi elettronici, si deve dare altrettanto importanza al panorama delle forme di comunicazione che ha subito un'evoluzione non indifferente: dispositivi che si basano su tecnologie digitali che hanno permesso di connettersi alla rete come personal computer, tablet, smartphone, fotocamere digitali, console per videogiochi e navigatori, costruendo un ecosistema connesso grazie anche al diffondersi dei social media. Un'affermazione di valore arriva dall'inventore del Web Tim Berners-Lee, riguardo al Web dichiara che sia un'innovazione sociale più che un'innovazione tecnica<sup>11</sup>, per tale motivo il termine “sociale” ha una doppia valenza: il primo valore è assegnare ad una tecnologia un aggettivo che la identifichi e il secondo valore è sottolineare il fatto che un'invenzione tecnologica è principalmente qualcos'altro, ovvero un nuovo modo di intendere le relazioni nel mondo. Questo per introdurre che i social media, oltre ad essere un'invenzione tecnologica sono soprattutto delle realtà di cambiamento nella società contemporanea: per comprendere il termine “social media” o “media sociale” è utile introdurre il concetto più generale di “reti sociali”<sup>12</sup>: per rete sociale si intende una struttura formale o informale che comprende un insieme di persone che collaborano fra di loro creando delle relazioni. Una rappresentazione grafica significativa delle reti sociali è data da una connessione di “nodi”, che corrispondono agli individui che sono in relazione tra di loro e operano in quella rete, ad esempio in relazioni esplicite, come legami di parentela, o in relazioni implicite, come per le amicizie che possono essere sia off-line (nel mondo reale) che on-line (nel mondo della rete virtuale). Inoltre, ad oggi, a causa del consistente utilizzo del social media, l'esperienza di comunicare online risulta in totale compenetrazione con quella off-line, portando le relazioni sociali su diversi piani comunicativi e il risultato che ne deriva da questa interazione è la nascita di un nuovo spazio sociale, “l'interrealtà”, molto più dinamico delle reti sociali precedenti. In questo senso le relazioni che si creano on-line sono direttamente collegate ai social media, che rappresentano, nello specifico, applicazioni basate sul Web che permettono di comunicare online includendo un insieme di utenti che consentono loro di creare, pubblicare, condividere ed entrare in contatto con altri utenti e/o organizzazioni e dare vita a interazioni/relazioni di vario tipo.

---

<sup>11</sup> Tim Berners-Lee, “Weaving the Web - The original design and ultimate destiny of the World Wide Web”;

<sup>12</sup> Ceron A., Curini L., Iacus S., “Social media e Sentiment Analysis - l'evoluzione dei fenomeni sociali attraverso la rete”.

Ecco, quindi, che si delinea la differenza tra social media e social network poiché vi è una netta differenza tra i due termini:

- **Social media:** rappresenta lo strumento, ovvero l'app/applicazione o il software fisico che permette lo scambio di contenuti da parte degli utenti via Web;
- **Social network:** rappresentano le persone (gli utenti) che utilizzano l'app per creare la propria rete di comunità online (le reti sociali).

Dopo aver chiarito quello che è la definizione di social media risulta utile creare una timeline che permette di comprendere l'evoluzione nella storia di queste piattaforme online; tutto partì nel 1991 quando l'High Performance Computing Act (HPCA)<sup>13</sup>, una legge statunitense, aprì la possibilità di ampliare la rete per finalità commerciali, traducendosi in grandi investimenti pubblici. Il primo servizio di social media della storia di Internet viene considerato Six-Degrees lanciato nel 1997 dall'avvocato statunitense Andrew Weinreich e basato sulla teoria elaborata nel 1967 dallo psicologo Stanley Milgram dei sei gradi di separazione chiamata anche "small world theory"<sup>14</sup>: teoria secondo la quale tutte le persone del pianeta terra sono collegate tra loro tramite una catena di conoscenze non superiore a più di 5 intermediari. I creatori di Six-Degrees hanno avuto l'idea di mappare un set di relazioni reali tra le persone però gli utenti non riuscirono a comprendere fino in fondo se Six-Degrees fosse una app per creare relazioni o un servizio legato al business. Difatti in quel periodo non vi era l'idea di "social network" (termine diventato di uso comune solo nel 2002). Quindi la prima vera applicazione ad essere utilizzata come social network, che riscosse molto successo, fu la piattaforma Friendster sviluppata dall'ingegnere canadese Jonathan Abrams. Per la prima volta un servizio Web mostrava le foto degli utenti e permetteva di cercare le persone, di vedere il loro profilo e decidere se connettersi alla loro rete. Il successo fu inaspettato tanto che l'azienda non è stata in grado di adeguare il proprio server scontentando gli utenti iniziali poiché in alcuni momenti il caricamento delle pagine impiegava più di venti secondi. Successivamente nell'agosto del 2003 approdò sul mercato MySpace ideata dallo studente universitario californiano Tom Anderson con l'aiuto dell'amico Chris DeWolfe, con lo scopo di offrire agli utenti la possibilità di creare blog personali, profili, accesso a gruppi e la visualizzazione di contenuti come foto, musica e video. Il successo della piattaforma arrivò

---

<sup>13</sup> HPCA rappresenta una legge statunitense per fornire un programma federale coordinato per garantire la continua leadership degli Stati Uniti;

<sup>14</sup> Arvidsson A., Delfanti A, "Introduzione ai media digitali".

quando gli utenti si accorsero che un bug del sistema dava loro la possibilità di personalizzare le pagine della piattaforma a loro piacimento, i fondatori anziché impedire questo comportamento decisero di incoraggiarlo dando libero sfogo alla creatività degli utenti. Proprio per questo MySpace diventò un servizio per artisti emergenti nel mondo della musica rivoluzionando il panorama musicale ma questo non ha garantito il successo nel tempo della piattaforma poiché nel 2005 fu venduto alla News Corporation di Rupert Murdoch per 600 milioni di dollari, da quel momento in poi MySpace iniziò a perdere una direzione strategica, e ad oggi risulta di proprietà di Specific Media un'agenzia che offre pubblicità online. Un grande salto nel mondo dei social network, incentrati sul settore del lavoro, fu LinkedIn ideato dall'imprenditore americano Reid Hoffman e da membri creatori di PayPal e Socialnet.com, in ordine di tempo non fu il primo social network orientato al mondo del lavoro ma divenne quello più utilizzato; infatti, LinkedIn ha permesso di far diventare il profilo personale dell'utente un vero e proprio curriculum vitae per poter creare relazioni utili per la propria carriera lavorativa. Nel 2005 è stato creato YouTube da Chad Hurley (amministratore delegato), Jawed Karim (consigliere) e Steve Chen (direttore tecnico), tre ragazzi che nel passato hanno lavorato per PayPal. La piattaforma è stata una tra le prime a permettere la condivisione e visualizzazione in rete di contenuti multimediali, ovvero videoclip, trailer, cortometraggi e live streaming. Successivamente però l'app fu acquistata nel 2006 dall'azienda Google indirizzando il servizio ad una strada più chiara e precisa. Ben presto, il panorama dei social network cambiò drasticamente quando arrivò Facebook di Mark Zuckerberg sul mercato mondiale, a partire dal 2008. Prima di essere la piattaforma social più conosciuta al mondo nacque con il nome di "The Facebook"<sup>15</sup> nel febbraio del 2004 quando il diciannovenne studente Mark Zuckerberg all'università di Harvard con l'aiuto di Andrew McCollum, Eduardo Saverin, Dustin Moskovitz e Chris Hughes, hanno ideato questa piattaforma che vedeva i primi iscritti gli studenti dell'università di Harvard poiché inizialmente si poteva accedere solo avendo un indirizzo di posta dell'università. Era basato su identità reali, un servizio di comunicazione per rimanere in contatto con gente conosciuta, dopo più di un anno la piattaforma si espanse in tutte le università statunitensi portando alla possibilità di iscriversi con un indirizzo di posta elettronica legato ad un dominio universitario. Nel 2006 invece ci fu l'apertura a tutti gli utenti che avessero un'età superiore a 13 anni, l'innovazione che in quel periodo ha reso Facebook unico e raro è stata la funzionalità foto, la prima ad avere

---

<sup>15</sup> V. Cosenza, "Social media ROI".



caratteristiche sociali come il tagging e il news feed. L'intuizione è stata poi trasformare il servizio in una vera e propria piattaforma, con un sistema operativo in grado di ospitare applicazioni di terze parti, avendo così un successo globale tanto da essere quotata in borsa al NASDAQ nel 2012 dopo che la piattaforma ha superato il miliardo di utenti attivi mensili. Da qui in poi il panorama dei social network aumentò a dismisura portando allo sviluppo di diverse app riscuotendo un successo a livello globale, si citano alcuni dei social network più noti in tutto il mondo:

→**Twitter** creato nel 2006 dall'imprenditore informatico statunitense Jack Dorsey durante una giornata di brainstorming tra i dirigenti dell'azienda di podcasting Odeo<sup>16</sup>. Questo social network permette agli utenti di interagire fra di loro "twittando" ovvero postando sul proprio profilo personale un messaggio di massimo 280 caratteri che viene chiamato "tweet" (tradotto dall'inglese significa cinguettio, che riprende il logo della piattaforma). La particolarità che caratterizza questo social network è il modo in cui vengono catalogati i contenuti testuali, poiché sono archiviati per argomento in base all'hashtag inserito dall'utente. Gli "hashtag" sono parole o frasi senza spaziatura precedute dal simbolo #;



**Figura 1.2** Logo di Twitter

→**WhatsApp** creata nel 2009 da due amici che lavoravano per Yahoo! Jan Koum (imprenditore informatico statunitense) e Brian Acton (programmatore statunitense). Si mettono al lavoro per la creazione di una piattaforma che consentisse agli utenti di chattare in tempo reale con gli altri utenti sostituendo gli sms tradizionali. Il nome, infatti, trova origine nell'unione dell'espressione inglese "what's up" che significa "che cosa succede/che c'è di nuovo?" e App che deriva da applicazione;



**Figura 1.3** Logo di WhatsApp

---

<sup>16</sup> D. Bennato, H. Benothman, A. Panconesi, "La diffusione delle informazioni online. Il caso Twitter".

→**Instagram** ideata nel 2010 da Kevin Systrom (imprenditore informatico statunitense) e Mike Krieger (imprenditore e ingegnere statunitense). I due fondatori inizialmente stavano lavorando alla realizzazione di Burbn (una app di check-in scritta in HTML5) quando però decisero di concentrarsi sulla funzionalità di condivisione di immagini hanno ricevuto dagli investitori 500.000 dollari lanciando così sul mercato l'applicazione. L'intuizione del nome Instagram nasce dall'unione di Instant Camera (fotocamera istantanea) e Telegram (telegramma) e, secondo la spiegazione di Systrom, trasferisce il concetto di immagini veloci che vengono pubblicate sul momento. Nel 2012 però l'app venne acquisita da Mark Zuckerberg, all'epoca già proprietario e fondatore di Facebook, per 1 miliardo di dollari. Da quel momento in poi all'interno dell'app gli utenti hanno potuto condividere e creare foto, video, commenti, chat private fino alla nascita di spot pubblicitari e Instagram Shop. Negli anni la piattaforma social è profondamente cambiata però quello che è certo è che ha portato all'utilizzo dell'app in tutto il mondo;



**Figura 1.4** Logo di Instagram

→**Telegram** ideata nel 2013 dai fratelli Nikolaj Durov e dall'imprenditore russo Pavel Durov, rappresenta un'applicazione di messaggistica istantanea in cui gli utenti possono chattare tra di loro senza possedere il contatto dell'altro poiché ci si iscrive attraverso un nickname. La particolarità di Telegram è quella di poter creare delle stanze virtuali che sono utilizzate per creare una community e tenere aggiornati periodicamente i contatti iscritti.



**Figura 1.5** Logo di Telegram

→**TikTok** fondata in Cina da Alex Zhu e Luyu Yang nel 2016. Rappresenta una app che permette agli utenti di creare brevi clip video di durata variabile (dai 15 ai 600 secondi). Inizialmente la piattaforma si chiamava Musical.ly, ma quando la società cinese Byte Dance acquistò Musical.ly, il nome venne cambiato in Tik Tok ("suono tremante") che fa riferimento al ticchettio delle lancette di un orologio.



**Figura 1.6** Logo di Tik Tok

### 1.2.1 Caratteristiche social network

Dopo aver chiarito la differenza tra social media e social network, si delineano le caratteristiche di quest'ultimi poiché l'evoluzione di queste piattaforme social sottolinea le principali differenze tra i social e i media tradizionali per aspetto e per tipologia di comunicazione. Nei media tradizionali come giornali, televisione e radio, solo chi ha le competenze ha la possibilità di creare contenuti; questa barriera invece è praticamente assente in tutti i tipi di social network creando in questo modo una parità di diffusione di testi, frasi, foto possedendo semplicemente un computer o cellulare interconnessi alla rete. Esistono vari tipi di social media che possono avere al loro interno la forma di una rete sociale definiti dai ricercatori Danah Boyd e Nicole Ellison anche Social Networking Sites (SNS)<sup>17</sup>; in particolar modo i ricercatori hanno delineato le caratteristiche dei social network che, per essere tali, devono rispettare tre condizioni principali:

1. Devono avere la possibilità di creare un profilo pubblico o semipubblico all'interno di un sistema vincolato;
2. Gli utenti devono essere collegati tra loro con una specifica relazione, ad esempio di "following", quindi un'articolazione di contatti;
3. Gli utenti devono avere la possibilità di interagire con una comunicazione interattiva (post, condivisione, messaggistica e interazione).

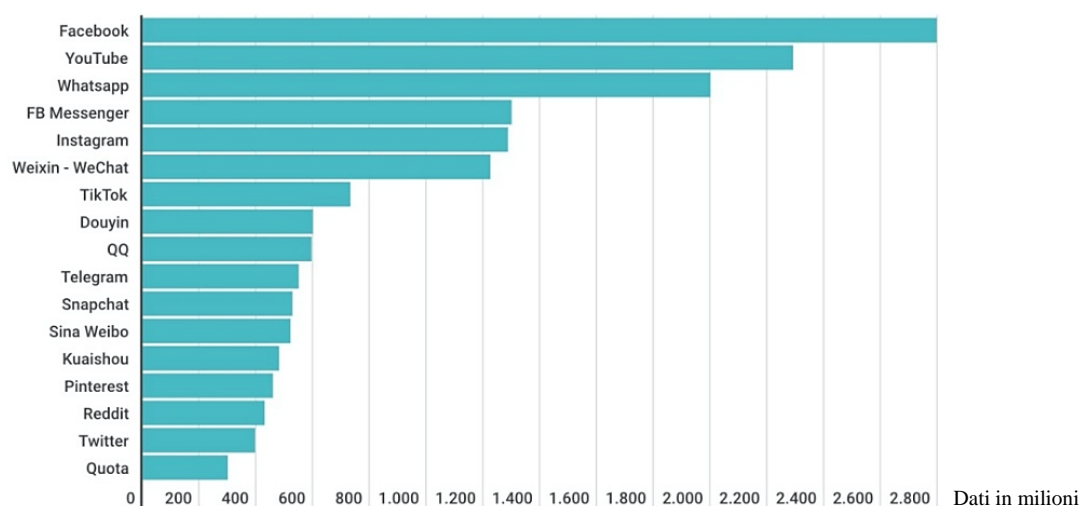
In particolar modo hanno degli elementi che li contraddistinguono<sup>18</sup>: sono digitali, cioè basati su codici numerici con unità discrete che, a differenza dei codici analogici, possono trasportare rapidamente immense quantità di informazioni, per questo sono anche replicabili e persistenti poiché le tracce sono bit che si possono facilmente riprodurre su diversi supporti. Ciò implica la possibilità, per qualsiasi contenuto, di essere decontestualizzato e remixato a piacimento da altri; sono ipertestuali, poiché permettono di fruire i diversi contenuti in maniera non lineare poiché il testo, ad esempio, non può essere stampato su una pagina cartacea ma ha una struttura più complessa; sono distribuiti, diversi dai media tradizionali centralizzati e unidirezionali - cioè distribuiscono informazioni "da uno a molti"-, poiché in essi i contenuti informativi sono prodotti e distribuiti "da molti a molti", ossia da milioni di individui che comunicano in una struttura orizzontale a rete; sono interattivi in quanto gli utenti nell'ambiente digitale possono

---

<sup>17</sup> Vincenzo Cosenza, "Social Media ROI";

<sup>18</sup> Arvidsson A., Delfanti A., "Introduzione ai media digitali".

interagire direttamente con i contenuti, selezionandoli e modificandoli, oltre a poterli produrre in prima persona e pubblicarli online; sono sociali, poiché i social media, permettono di creare un profilo personale attraverso il quale mantenere e gestire le relazioni con amici e familiari e di creare nuovi legami interagendo con una rete vasta di contatti; sono mobili in quanto smartphone, tablet, netbook, ecc., sono dispositivi che permettono agli individui di accedere alla rete da qualunque luogo e in qualsiasi momento, senza dovere utilizzare un personal computer collegato a una linea telefonica, questo aspetto permette agli individui di sganciarsi ancora di più dal “senso del luogo”, ossia dalle condizioni spazio-temporali e dagli altri elementi contestuali che caratterizzano la comunicazione faccia a faccia. I social network attraverso l'utilizzo del telefono, che si è trasformato nel più potente e versatile strumento di connettività, si ha a portata di mano una connessione “da persona a persona” in cui le persone sono in contatto come utenti piuttosto che come persone inserite in luoghi lavorativi o sociali. Quindi in qualche modo si eliminano quei confini spazio-temporali che nella vita reale sono presenti e limitanti; ecco che si creano dei confini sfumati tra pubblico e privato. I social network in questo senso hanno amplificato questa situazione, soprattutto nella fase del Web 2.0, poiché in prima persona i personaggi pubblici rendono visibile la propria vita creando così una community definita come una vera e propria rete sociale.



**Figura 1.7** Social Network utilizzati nel mondo report di luglio 2021, fonte Kepios

### 1.3 Nuovo modo di vedere il Web: Realtà Virtuale

Negli anni le forme di comunicazione si sono modificate ed evolute, prima con i media più tradizionali poi con i social media, però come insegna la storia di Internet, si è sempre alla ricerca di un cambiamento che possa rivoluzionare i paradigmi passati per crearne di nuovi, o meglio sfruttare l'evoluzione tecnologica per concretizzare le scoperte passate; portando a quella che oggi viene chiamata Realtà Virtuale (Virtual Reality, VR) associandola ad un concetto complesso e astratto, in realtà rappresenta un ramo dell'informatica caratterizzato da multidisciplinarietà che trova origine nella ricerca della grafica tridimensionale (3D). In particolar modo queste tecnologie si sono evolute con il passare degli anni poiché in passato sono stati diversi gli approcci utilizzati nell'elaborazione di diversi sistemi, dato che molti sono gli interessati nell'utilizzo della tecnologia. Tutti gli studiosi che hanno contribuito alla creazione di questo ambiente, riporta la propria personale percezione della tecnologia che varia dalla più semplice alla più sofisticata e astratta, proiettando e sviluppando varie aspettative sul tema. La definizione di Realtà Virtuale fa riferimento ad un sistema complesso composto da diverse tecnologie, quest'ultime si sono evolute negli anni e, al fine di comprenderne meglio la storia, è utile descrivere brevemente come funziona questo sistema. L'approccio alla base della tecnologia della Realtà Virtuale è dato da un computer che sintetizza un ambiente a tre dimensioni partendo dai dati numerici, l'ambiente è reso accessibile dall'operatore umano attraverso l'utilizzo di dispositivi progettati su misura, permettendo così di ricreare modelli che rappresentano in maniera fedele la realtà come per esempio una città esistente, o circostanze che ricreano l'ambiente che circonda un luogo in modo tale che possano essere comprese dall'uomo. Per la realizzazione di un ambiente virtuale è necessario<sup>19</sup>: la scelta dell'architettura di supporto e del sistema software in modo da assemblare gli elementi dell'ambiente 3D e permettere le interazioni al suo interno come luci, suoni e gravità. Una volta scelta la piattaforma e assemblati gli elementi sopra citati, è possibile assemblare l'ambiente VR che viene pubblicato e dev'essere accessibile e fruibile dagli utenti attraverso un device, ovvero la parte hardware che ricrea l'ambiente virtuale. Per poter usufruire dell'ambiente 3D, come strumento di interazione per l'utente, si deve utilizzare un avatar che può essere realizzato in ogni momento. Tutto questo vale per ambienti VR che sono realizzati per i singoli utenti (Single User Environnement),

---

<sup>19</sup> Milgram P., Takemura H., Utsumi A., Kishino F., "Augmented Reality: A class of displays on the reality-virtuality continuum".

mentre per riprodurre un ambiente con più utenti (Multi-User Environnement) è necessario creare un Network di avatar. La precedente descrizione del funzionamento della tecnologia VR rappresenta una breve premessa per la miglior comprensione dell'evoluzione storica della stessa. La storia della Realtà Virtuale risale tra gli anni '50 e '60 quando alcuni pionieri iniziano a sperimentarla: la prima importante invenzione è attribuibile a Morton Heilig, regista e cameramen statunitense, che nel 1956 parlò del cosiddetto "cinema esperienza"<sup>20</sup> che coinvolgeva in maniera realistica lo spettatore come mai fatto prima. Egli costruì un prototipo della sua visione, chiamato Sensorama brevettato nel 1962, un dispositivo meccanico a forma di cabina con schermi stereoscopici che ricreavano attraverso costruzioni 3D l'illusione della realtà e fu il primo esempio di tecnologia immersiva multisensoriale. L'idea del Sensorama<sup>21</sup> nasce dalla visione futuristica di Heilig e dal mondo del cinema, infatti egli prese ispirazione da un sistema a tre telecamere utilizzato per la proiezione di film su grande schermo. Però Heilig, a quel tempo, non trovò fondi sufficienti per fare in modo che la sua idea venisse applicata al cinema come voleva. Successivamente nel 1968 Ivan Sutherland, professore associato di Matematica applicata alla Harvard University, insieme all'aiuto di uno dei suoi studenti Bob Sproull realizzò il primo sistema di Realtà Virtuale con un visore appeso al soffitto poiché troppo pesante da indossare; quindi, risultava primitivo sia in termini di interfaccia utente sia di realismo (visore che qualche anno dopo venne chiamato "The Sword of Damocles"<sup>22</sup>). Nello stesso periodo, grazie alle ricerche svolte negli anni precedenti, le immagini sono diventate computerizzate e tridimensionali, gli schermi dei computer si sono trasformati in elaboratori complessi e la tecnologia informatica si avviò verso un universo virtuale che gradualmente ha aggiunto funzionalità con dispositivi sempre più complessi. Il dibattito si accese su questi temi dando pochi frutti all'inizio degli anni '80 poiché in quel periodo il Web era ancora ai suoi albori e sarebbe stato troppo precoce lanciare sul mercato un dispositivo che andava oltre all'immaginario comune e alla tecnologia presente di quel periodo. Nonostante questo, l'origine del termine "Virtual Reality"<sup>23</sup> fu utilizzato per la prima volta nel 1989 dall'informatico statunitense Jaron Lanier che venne considerato, in un secondo momento, il pioniere di questa espressione.

---

<sup>20</sup> Heilig M., "The Cinema of the Future";

<sup>21</sup> Sensorama è una macchina ideata per il cinema dell'esperienza creato da Morton Heilig;

<sup>22</sup> Sutherland I., "A head-mounted three-dimensional display";

<sup>23</sup> Enciclopedia "Breve storia dei supporti virtuali".

All'epoca Lanier fondò la VPL Research (Virtual Programming Languages) inventando dei guantoni e un visore per connettere il corpo ad uno scenario softwarizzato, in questo modo sintetizzò gli sforzi intellettuali e tecnici di tanti, come il visore di Ivan Sutherland e Bob Sproull. Altre definizioni del termine sono state formulate con il passare degli anni ma quella più significativa arrivò dallo scienziato informatico Jonathan Steuer, membro del reparto di ricerca sulla comunicazione dell'Università di Stanford, in cui notò che l'origine del termine provenisse dal settore IT (Information Technology) e non dalle istituzioni accademiche. Lo stesso Steuer diede una visione più tangibile e concreta dell'espressione sostenendo che, la definizione che è stata attribuita alla Realtà Virtuale (come un sistema hardware composto da computer, cuffie e guanti sensoriali) fosse riduttiva e focalizzata solo sulla tecnologia piuttosto che sulla parte importante che è l'esperienza dell'utente. Steuer sottolinea questo aspetto attraverso il concetto di immersione dell'utente nella Realtà Virtuale così esplicitato:

*“The key to defining virtual reality in terms of human experience rather than technological hardware is the concept of presence. It refers not to one’s surroundings as they exist in the physical world, but to the perception of those surroundings as mediated by both automatic and controlled mental processes. However, when perception is mediated by a communication technology, one is forced to perceive two separate environments simultaneously: the physical environment in which one is present, and the environment presented via the medium. The term Telepresence can be used to describe the precedence of the latter experience in favor of the former; that is, telepresence is the extent to which one feels present in the mediated environment, rather than in the immediate physical environment.”*<sup>24</sup>

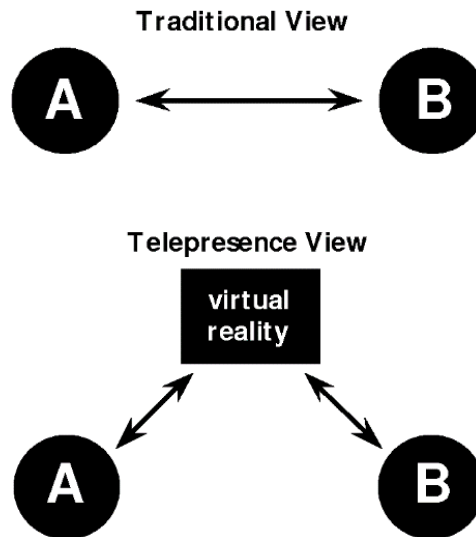
Da quello che emerge nella frase di Steuer, la combinazione del senso di immersione e il concetto di interattività sono indicate in italiano con il termine “telepresenza”, che in questo caso sottolinea come un'esperienza VR renda praticamente inconsapevoli dell'ambiente reale che ci circonda. Inoltre, vi è la necessaria esistenza di due componenti principali affinché si raggiunga l'immersione, ovvero:

- la profondità dell'informazione (quantità e qualità dei dati nei segnali che un utente riceve quando interagisce in un ambiente virtuale),
- l'ampiezza dell'informazione (vasta gamma di informazioni che stimola tutti i sensi).

---

<sup>24</sup> Steuer J., “Communication in the Age of Virtual Reality”.

Il concetto di Steuer venne reso noto da Krueger Myron, statunitense laureato in arti liberali e famoso per i suoi studi in ambienti interattivi e di interfacce uomo-computer, sintetizzandolo nella successiva rappresentazione grafica<sup>25</sup>:



**Figura 1.8** Due approcci/modelli di comunicazione a favore dell'idea di Steuer

Ci sono stati altri importanti contributi come quelli di Ken Pimentel e Kevin Teixeira appassionati di mondi virtuali, essi definiscono come: *"A new media for getting your hands on information, getting inside information, and representing ideas in ways not previously possible."*<sup>26</sup>

La loro visione, condivisa da molti ricercatori, si basa sulle potenzialità di questi strumenti che si focalizzano sull'illusione della realtà stessa. Molti altri ricercatori hanno dato la propria opinione sul tema, essendo così vasto e coinvolgente da trattare, ma in generale il termine immersivo può essere associato all'illusione dell'utente di far parte del mondo ricreato in digitale, circondato dalle immagini e i suoni che contribuiscono alla sensazione di presenza, difatti un sistema di Realtà Virtuale, ha lo scopo di aumentare il senso di immersione dell'utente. Se durante le prime scoperte della Realtà Virtuale il principale interesse si è concentrato nello sviluppare le parti hardware del sistema, dagli anni 2000 l'interesse si è spostato nell'ambiente software grazie alle

<sup>25</sup> Krueger M. W., "Artificial Reality II";

<sup>26</sup> Pimentel K., Teixeira K., "Virtual reality: through the new looking glass".



innovazioni tecnologiche che hanno reso possibili degli sviluppi significativi infatti, in quel periodo storico, stavano nascendo i primi engine 3D (un engine di grafica 3D è quella parte di software che permette di elaborare in tempo reale immagini, elementi e interazioni in base agli input che sono inviati dall'utente) portando l'interesse di diversi ricercatori su questo tema. Successivamente al 2009 l'entusiasmo nutrito negli anni precedenti iniziava a diminuire e molti si sono chiesti com'è stato possibile che il rapido successo dei sistemi VR e dei mondi virtuali abbia avuto un altrettanto rapido declino nel pensiero sociale. Una risposta a tale domanda può essere data, come suggerisce Galimberti filosofo e psicoanalista italiano, dal grande entusiasmo ed enormi aspettative che il pubblico ripone nell'arrivo di una nuova tecnologia spesso in modo affrettato ed esagerato. Le novità di solito sono accompagnate da molte informazioni che spesso sono parziali o non precise e la crescita dei mondi virtuali ne è la prova lampante. Ad esempio, nell'ambito aziendale, vi è stata un'ondata di articoli in cui venivano elogiate le grandi potenzialità offerte al mondo del business grazie all'ambiente VR, senza però specificare la strada da percorrere per raggiungere gli obiettivi prefissati. Inoltre, è utile ricordare che per lo sviluppo della tecnologia VR sono necessarie precise tecnologie avanzate che, storicamente, si sono evolute con modalità e tempi diversi per cui questo può essere uno dei diversi fattori che hanno contribuito a limitare l'utilizzo di tale tecnologia in quegli anni. L'entusiasmo generale continuò a diminuire fino al 2011<sup>27</sup>, anno in cui tocca il punto più basso di questo stato rimanendo invariato per i due anni successivi. Solo dal 2014, anno importante per la tecnologia VR, si è iniziata una graduale ascesa con l'invenzione di diversi dispositivi, nel marzo del 2014 durante l'evento "Driving the Future of Innovation at Sony Computer Entertainment", Shuhei Yoshida presidente dei Worldwide Studios di Sony, ha presentato l'invenzione di un caschetto di visualizzazione VR chiamato progetto "Morpheus" e, alla fine dello stesso mese, Mark Zuckerberg sul profilo personale di Facebook ha annunciato l'acquisizione dal valore di due miliardi di dollari di una giovane società del settore Oculus VR, il quale ha evidenziato i cambiamenti che avverranno nel futuro nel modo di giocare, comunicare e lavorare:

*"Oculus's mission is to enable you to experience the impossible. Their technology opens the possibility of completely new kinds of experiences. Immersive gaming will be the first, and Oculus already has big plans here that won't be changing, and we hope to accelerate. The Rift is highly anticipated by the gaming community, and there's a lot of interest from developers in*

---

<sup>27</sup> Analisi di Hype Cycle for Emerging Technologies 2011.

*building for this platform. Imagine enjoying a court side seat at a game, studying in a classroom of students and teachers all over the world or consulting with a doctor face-to-face -- just by putting on goggles in your home. This is really a new communication platform. Imagine sharing not just moments with your friends online, but entire experiences and adventures (...). The future is coming, and we have a chance to build it together.*"<sup>28</sup>

Dopo la precedente panoramica generale del mondo VR, si può comprendere che l'introduzione nel mercato di tecnologie nuove e innovative, prima di essere utilizzate dalle aziende, incontrano sul proprio percorso diversi ostacoli poiché devono essere realmente comprese, devono dimostrare il loro potenziale e la loro effettiva fattibilità di creazione per poterle utilizzare. Per esporre dei numeri concreti, delle analisi di mercato riguardanti il tasso di crescita della tecnologia VR, Bloomberg (nota multinazionale che opera nel settore dei mass media) ha stimato che il mercato globale del Metaverso potrebbe crescere fino a circa 800 miliardi di dollari entro il 2025<sup>29</sup>. E inoltre il mercato globale della tecnologia VR crescerà a un CAGR (Compound Annual Growth Rate) del 42,9% nel periodo 2020-2030<sup>30</sup>. Dalle analisi sopra esposte è possibile affermare che, la Realtà Virtuale ha continuato negli anni la sua evoluzione ma è proprio questo, e lo sarà negli anni a venire, il periodo storico in cui inizia la sua vera e propria espansione.

---

<sup>28</sup> Annuncio Facebook profilo di Mark Zuckerberg <https://www.facebook.com/zuck/posts/10101319050523971>;

<sup>29</sup> Analisi di Bloomberg <https://www.bloomberg.com/professional/blog/metaverse-may-be-800-billion-market-next-tech-platform/>;

<sup>30</sup> CAGR: rappresenta il tasso di crescita medio annuo di un investimento in un periodo di tempo specificato superiore ad un anno. È uno dei metodi più accurati per calcolare i rendimenti nel tempo.

### 1.3.1 Definizione di Metaverso

Da qualche anno a questa parte, parlando di Realtà Virtuale, è stato diffuso il termine “Metaverso”, a partire dall’ottobre 2021, secondo i rapporti sulle tendenze di Google, c’è stato un forte aumento delle ricerche su Metaverse nel motore Google Search. Quello che ha attirato l’attenzione su questa parola è stato il comunicato stampa di Mark Zuckerberg, facendo conoscere al grande pubblico questo termine. Zuckerberg ha annunciato il 28 ottobre 2021 che avrebbe cambiato il nome dell’azienda proprietaria di Facebook, Messenger, Instagram, WhatsApp, Quest e Portal in Meta e inoltre ha dichiarato che farà investimenti significativi in Metaverse. La scelta di Zuckerberg è stata vista inizialmente come un tentativo di rebranding per ripulirsi dagli scandali che hanno coinvolto la compagnia negli anni precedenti, quello che è certo è che il cambio nome di Facebook ha contribuito a portare l’attenzione del grande pubblico sul tema del Metaverso. Per fare chiarezza però l’origine del termine, come si è potuto notare dalla storia del Web e dai primi studi pionieristici della Realtà Virtuale, risale a quasi 20 anni prima poiché è entrato nel vocabolario popolare quando è stato definito per la prima volta dell’autore statunitense Neal Stephenson nel suo romanzo di fantascienza "Snow Crash", pubblicato nel 1992. In questo romanzo Stephenson immagina il Metaverso come un mondo virtuale popolato che opera in un sistema di cultura digitale in parallelo al mondo fisico. Il termine è utilizzato per indicare uno spazio tridimensionale (3D) e mondi virtuali (Virtual World) all’interno del quale persone fisiche possono muoversi senza limiti fisici del mondo reale, interagire fra di loro tramite avatar personalizzati. Per darne una definizione più chiara la parola Metaverso, dall’inglese Metaverse, formata dal prefisso “meta” che indica un qualcosa che va oltre/dopo e il sostantivo “(uni)verse” che indica universo/realtà, ovvero una realtà parallela anche cosiddetta virtuale. Tanti ricercatori hanno espresso le loro definizioni su questo concetto e, tra di esse, di particolare interesse risulta la definizione esposta dall’imprenditore statunitense John Smart, CEO di Foresight Company e presidente dell’Acceleration Studies Foundation:

*“The Metaverse is the convergence of 1) virtually enhanced physical reality; 2) physically persistent virtual space. It is a fusion of both, while allowing users to experience it as either.”*<sup>31</sup>

---

<sup>31</sup> Smart J., Cascio J., “A Cross-Industry Foresight Project, Metaverse Roadmap”.

Secondo l'affermazione di Smart il Metaverso rappresenta la connessione che fonde la realtà fisica con la realtà digitale. Per parlare adeguatamente di Metaverso bisogna tenere in considerazione almeno quattro elementi secondo una ricerca focalizzata sulla realizzazione del Metaverso condotta da Dionisio, Burn e Gilbert, queste quattro caratteristiche sono<sup>32</sup>:

- 1) **Realismo:** associato all'esperienza che viene vissuta direttamente dagli utenti, permettendo il coinvolgimento psicologico ed emotivo grazie all'utilizzo della vista, udito e tatto. Un maggiore realismo viene associato ad una maggiore naturalezza nella gestualità dell'avatar di riferimento;
- 2) **Ubiquità:** capacità di accesso degli utenti dai diversi dispositivi digitali agli spazi virtuali che formano il Metaverso, affinché sia uno strumento efficace deve assicurare agli utenti di rimanere interconnessi nel sistema;
- 3) **Scalabilità:** questa caratteristica rappresenta la sfida più grande poiché il Metaverso, in termini di livelli e dimensioni, è potenzialmente infinito. Quindi deve permettere a tutti gli utenti presenti di essere utilizzato in maniera efficiente nonostante il numero elevato di flusso;
- 4) **Interoperabilità:** capacità di rendere continuo e trasparente l'interazione e lo scambio di informazioni tra sistemi distinti e piattaforme differenti per fare in modo di creare e spostare oggetti e utenti senza interruzioni all'interno del sistema.

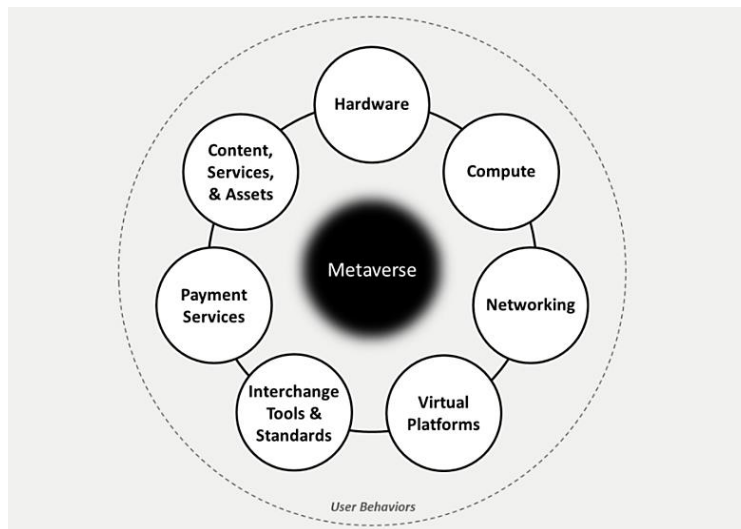
Questi sono i quattro fattori fondamentali per la realizzazione fisica di un Metaverso efficiente però è utile anche ricordare che, a livello tecnologico, per l'implementazione del Metaverso sono necessari otto elementi, come specificato da Matthew Ball analista e amministratore delegato di Epyllion una holding che fornisce servizi di consulenza nel settore tecnologico; gli otto elementi esposti da Ball sono:

- *hardware* (per coinvolgere l'utente a 360° si deve utilizzare visori per la Realtà Virtuale e dispositivi in grado di fornire sensazioni tattili e più vicini al reale);
- *infrastrutture di rete* (ovvero la messa in pratica del sistema infrastrutturale di rete con una banda molto ampia per gestire le connessioni nel Metaverso in tutto il mondo);
- *compute* (la potenza computazionale per gestire una tale mole di dati);

---

<sup>32</sup> Dionisio J., Burn W., Gilbert R., "3D Virtual Worlds and the Metaverse".

- *piattaforme virtuali* (le aziende che costruiranno questi mondi devono strutturare il proprio modello di business su queste piattaforme virtuali);
- *tool e standard comuni* (al fine di garantire l'operabilità tra le varie piattaforme e costruire un Metaverso che sia il più possibile accessibile e omogeneo);
- *pagamenti* (che potranno variare significativamente nel caso di monete digitali "ufficiali", criptovalute o altri strumenti appositi);
- *creazione di contenuti e servizi ad hoc* (come asset digitali e NFT per garantire l'unicità di ciò che si fa e si possiede nel Metaverso);
- *comportamento degli utenti* (se saranno disposti ad utilizzare questa tecnologia e a che vantaggio).



**Figura 1.9** Tecnologie e sfide del Metaverso, fonte MatthewBall.vc

Si tratta quindi di qualcosa che va ben oltre lo specifico utilizzo della Realtà Virtuale utilizzata per semplificare alcuni compiti della vita quotidiana, bensì di una vera e propria collisione tra digitale e analogico. Se negli anni precedenti gli studiosi avevano già teorizzato questa fusione, il Metaverso in realtà la porterebbe a livelli successivi, creando una realtà dinamica in cui analogico e digitale si co-costruiscono costantemente. È possibile quindi immaginare un network digitale dato da un insieme di mondi virtuali che possono essere interconnessi tra di loro dando vita al Web 3D utilizzando la tecnologia USD (Universal Scene Description)<sup>33</sup>, che rappresenta l'evoluzione del Web che viene utilizzato al giorno d'oggi.

<sup>33</sup> "Mondi virtuali 3D: USD, l'HTML del Metaverso".

### 1.3.2 Ambiti applicativi e ambienti di sviluppo del Metaverso

#### Ambiti applicativi

La tecnologia del Metaverso, come già trattato precedentemente, non deve ambire a diventare capillare in tutti i campi; infatti molti imprenditori e ricercatori della tecnologia hanno già effettuato importanti investimenti nei settori che ritengono di maggior sviluppo. Risulta chiaro ormai che sia da un punto di vista tecnologico che da un punto di vista dell'esperienza vissuta da parte degli utenti, si conferma il pensiero che questa nuova tecnologia abbia tutto il potenziale per diventare nel futuro prossimo una grande piattaforma che cambierà alcuni aspetti della vita quotidiana delle persone. Negli anni sono state condotte diverse ricerche per ottenere un'elaborazione delle previsioni sui principali ambiti applicativi che questa tecnologia ingloba con sé e sono state individuate delle macroaree comuni. Il settore del gaming già nel 2016 era valutato come il campo d'esordio poiché permette una reale immersione nel gioco stimolando "l'engagement" degli utenti. Una ricerca pubblicata da Deloitte, "Technology, Media and Communication Prediction"<sup>34</sup>, individua il gaming come settore principale: *"Deloitte Global predicts that Metaverse will have its first billion-dollar year in 2016, with about \$700 million in hardware sales. In 2016 we expect the vast majority of commercial activity on video games."*<sup>35</sup>

Un'altra industria che risulta tra le prime a cogliere la potenzialità di questa tecnologia è quella dell'entertainment, i settori del cinema, dei video, degli eventi, della musica, dello sport e del turismo. L'industria del cinema e, in generale, la produzione di video è tra gli esempi più significativi poiché non tutti utilizzano i videogames ma molti usufruiscono del cinema e dei contenuti video sul Web. Un film o un video fruito attraverso l'utilizzo di un HMD<sup>36</sup> stravolge il modo di guardare i contenuti poiché la sensazione è quella di immergersi all'interno della scena. L'utente, quindi, non vede semplicemente immagini tridimensionali, bensì vive un altro tipo di sensazione, quella di essere parte del film e delle sue ambientazioni. Un altro settore d'interesse è quello del lavoro che dall'inizio della pandemia di Covid-19 ha rivoluzionato le abitudini lavorative di miliardi di persone in tutto il mondo, rendendo quotidiano il lavoro da remoto e le riunioni sulle varie piattaforme Teams, Zoom, Skype e WebEx; in questo senso gran parte del

---

<sup>34</sup> Deloitte è un'azienda prima al mondo, in termini di ricavi, di servizi di consulenza;

<sup>35</sup> Lee P., Steward D., "A billion-dollar niche, TMT prediction, Deloitte";

<sup>36</sup> HMD (Head-Mounted Display) è uno schermo montato sulla testa dello spettatore attraverso un casco ad hoc, può essere monoculare o binoculare.

lavoro si è già spostato in parte online ed il Metaverso potrebbe renderlo ancora più interattivo, portando ad un'ulteriore collisione tra reale e digitale: utilizzando un ambiente di lavoro con avatar 3D, questo approccio potrebbe rivelarsi utile per chi ha intenzione di sviluppare abilità pratiche e manuali utili nell'ambiente di lavoro, che si potrebbero simulare nel Metaverso. Di ampio utilizzo è anche il settore dell'educazione, EdTech (Education Technology) in cui molte imprese di tutto il mondo stanno concentrando le proprie energie investendo il 47% del proprio budget in formazione e soluzioni di digital learning, con un fatturato che vale circa 19,5 miliardi di euro<sup>37</sup>. Come osservato nella ricerca questa nuova tecnologia potrebbe essere un valore aggiunto per migliorare la parte educativa degli studenti: *“Oggi l'EdTech può essere definito come l'insieme delle soluzioni a supporto del processo educativo, volto a massimizzare la qualità dell'esperienza di apprendimento e abilitato dalle tecnologie hardware e software.”*

Infine, il settore tra i più importanti che abbraccia tutti quelli sopra citati, rimane quello commerciale. La maggiore attenzione sul tema del Metaverso è arrivata dal settore dei mercati finanziari, che hanno visto un'opportunità di business straordinaria. Secondo il report di Citigroup “Metaverse and Money, Describing the Future”<sup>38</sup>, prendendo in considerazione una definizione molto più ampia di Metaverso, ovvero come una vera e propria estensione del Web, si tratta di un mercato che già nel 2030 potrebbe valere tra gli 8 e i 13 trilioni di dollari. Da una parte vi è la possibilità di creare mercati del tutto nuovi, a livello di hardware e di software per il funzionamento del Metaverso e dall'altra vi è un rapido aumento nell'utilizzo di NFT (Non Fungible Token)<sup>39</sup>. In particolar modo tutte le transizioni che un utente compie sono “datificabili” ovvero risultano tracciate e l'utente viene profilato e quindi si potrebbe arrivare alla monetizzazione di qualsiasi cosa. Per concludere è interessante osservare più da lontano, chi sono gli investitori principali per la creazione del Metaverso, in particolar modo possono essere distinti in tre tipologie: le grandi aziende tecnologiche come Meta, Microsoft, Nvidia, Apple, Alphabet, i Venture Capitalist che stanno finanziando iniziative e imprese nel campo adiacente degli NFT ed infine i Brand che sperimentano con progetti innovativi. Una ricerca condotta da

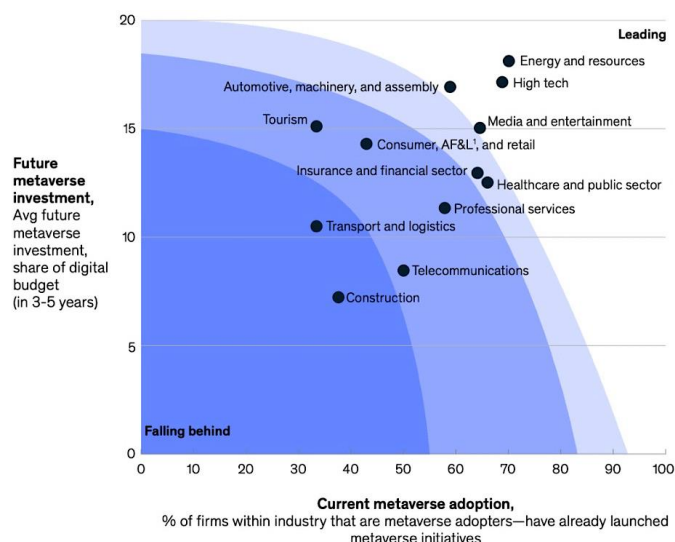
---

<sup>37</sup> Ricerca condotta da Osservatori.net Digital Innovation “Cresce il mercato dell'Educational Technology”;

<sup>38</sup> Ricerca condotta da City Group “Metaverse and Money, Describing the Future” <https://www.citivelocity.com/citigps/metaverse-and-money/>;

<sup>39</sup> NFT (Non Fungible Token) “certificati digitali” basati sulla tecnologia blockchain volti a identificare in modo univoco, non replicabile la proprietà di un prodotto digitale.

McKinsey<sup>40</sup> afferma che nei primi mesi del 2022, 120 miliardi di dollari sono stati investiti nel “campo largo” del Metaverso (risulta più del doppio di quanto speso nel 2021).



**Figura 1.10** Settori che guidano l’adozione del Metaverso, fonte McKinsey & Company

## Ambienti di sviluppo

Sono diversi gli ambienti di sviluppo Open Source che si sono creati per i Metaversi; si pone però attenzione su uno in particolare High Fidelity (HiFi) una piattaforma sviluppata da Philip Rosedale, imprenditore e fisico statunitense fondatore di Second Life<sup>41</sup>, con l’aiuto dei suoi colleghi Ryan Karpf e Freddy Heiberger nel 2013 a San Francisco. Fin dall’inizio la piattaforma ha ricevuto l’interesse di investitori che, ad oggi, hanno finanziato il progetto per più di 18 milioni di dollari<sup>42</sup> principalmente da True Ventures (società con sede nella Silicon Valley focalizzata in startup tecnologiche), Linden Lab (società americana che sviluppa giochi online), Vulcan Capital (attività di investimento in capitale di rischio di aziende in fase di start up, caratterizzate da un elevato potenziale di sviluppo) e Google Venture (azienda che si occupa di finanziare progetti nuovi e fornire agli imprenditori risorse tecnologiche). Il concetto che sta alla base di High Fidelity è quello di creare una piattaforma in cui qualsiasi utente ha la possibilità di

<sup>40</sup> McKinsey & Company è una multinazionale di consulenza strategica leader nel mondo per quota di mercato, fondata da James O. McKinsey e Marvin Bower;

<sup>41</sup> Second Life è un mondo virtuale lanciato il 23 giugno 2003, si tratta di una piattaforma informatica nel settore dei nuovi media che trova applicazione in molteplici campi;

<sup>42</sup> <https://www.businessintelligencegroup.it/metaverso-il-ritorno-di-second-life/>.



accedere a mondi virtuali tra loro interconnessi, secondo questo pensiero quindi, High Fidelity rappresenta l'infrastruttura tecnologica che renderà possibile la creazione del Metaverso, grazie all'uso di tecniche e linguaggi comuni si potranno generare gli spazi virtuali accessibili e in comunicazione tra loro. Lo scopo di HiFi è quello di poter realizzare ogni spazio virtuale all'interno di uno specifico mondo virtuale che, a sua volta, farà parte nel suo complesso del Metaverso. Come l'attuale Web, che è composto da diversi domini, anche i diversi spazi virtuali verranno identificati con il proprio dominio creando così un sistema di domini. I principali fattori che caratterizzano e rendono unica la tecnologia della piattaforma:

- la natura Open Source per permettere di arrivare ad un numero elevato di utenti;
- la scalabilità infinita fornita dalla potenza di calcolo dei dispositivi connessi;
- l'audio 3D spaziale che garantisce agli utenti un'esperienza audio immersiva grazie alla riproduzione di suoni provenienti da diverse fonti;
- il riconoscimento delle espressioni facciali e del movimento del corpo grazie all'uso di videocamere 3D e dei dispositivi sensoriali che si trovano sui visori di ultima generazione come Valve Index, HTC Vive Pro 2, Oculus Quest 2, PlayStation VR e molti altri.

Ciò che rende la piattaforma High Fidelity unica rispetto alle altre, è la tecnologia d'interfaccia utente che permette un'interazione più vicina all'esperienza reale con l'ambiente e con gli altri utenti. Difatti il sistema permette all'avatar di riprodurre i movimenti del corpo dell'utente ma anche le espressioni facciali e il movimento di occhi e bocca: per poter fare ciò vi è una telecamera frontale collegata allo schermo del PC che permette così al sistema di riconoscere le espressioni del volto dell'utente e ricrearle in tempo reale nel volto dell'avatar presente all'interno dell'ambiente virtuale. La tipologia di telecamere, a tecnologia avanzata, è già da qualche anno in dotazione nei dispositivi di ultima generazione come PC, tablet e laptop, che hanno sostituito le meno recenti webcam<sup>43</sup>. Questa evoluzione è frutto della tecnologia sviluppata nel 2014 da Intel Corporation (multinazionale statunitense fondata nel 1968), come spiega un articolo pubblicato sulla MIT Technology Review<sup>44</sup>. Questo sistema di telecamere offre molteplici opportunità agli utenti, relative ad un'esperienza ancor più immersiva e realistica se percepita attraverso l'utilizzo di un HMD poiché è come se l'utente vivesse l'ambiente virtuale con gli occhi dell'avatar.

---

<sup>43</sup> Una webcam (parola composta che unisce "web" abbreviazione di World Wide Web e "cam" come camera) è una piccola telecamera utilizzabile solo (o principalmente) come dispositivo di input per un elaboratore;

<sup>44</sup> <https://www.technologyreview.com/2014/01/06/82566/ces-2014-intels-3-d-camera-heads-to-laptops-and-tablets/>.

## Capitolo 2

### Social Media Analytics: strumenti e analisi

La parte dedicata al progetto analitico è illustrata in questo capitolo, come già accennato nell'introduzione, i social media rappresentano una ricca fonte di opinioni espresse dagli utenti, poiché essi espongono ciò che pensano, e questo risulta fondamentale per raccogliere informazioni e utilizzarle per creare delle analisi utili: in questo caso il social media preso in considerazione è Twitter riguardo i temi Metaverso e Realtà virtuale. A primo impatto questi due temi sono al centro del dibattito dei nostri giorni poiché il grande pubblico è interessato alle nuove tecnologie e affascinato dal potenziale che possono trarne in prima persona, per questo motivo si è reso utile analizzare ciò che pensano. La necessità di analizzare l'opinione degli utenti/consumatori c'è sempre stata, ma si è evoluta con il Web poiché porta con sé diversi aspetti: dalla Web Analytics nel Web 1.0. in cui la modalità di identificazione del consumo di contenuti in rete era la misurazione del numero di pagine viste dagli utenti, fino ad arrivare alla Social Media Analytics<sup>45</sup> del Web 2.0, in cui si è passati ad una misurazione del tempo di fruizione dei contenuti da parte degli utenti, per esempio, il tempo di permanenza all'interno di un sito, il numero di interazioni e il tempo di fruizione di un video online. Questo passaggio è avvenuto poiché si è avuto un arricchimento dei contenuti non strutturati, ovvero non solo testuali ma anche multimediali e application-based<sup>46</sup>. Con questa nuova misurazione sono sorte nuove opportunità e sfide, in particolare la Social Media Analytics rappresenta una disciplina legata ad analizzare le opinioni espresse sui social media poiché questa è in grado di: *“Aiutare le aziende a misurare, valutare e spiegare le performance delle iniziative sui social media nel contesto di specifici obiettivi di business”*<sup>47</sup>. In questo senso, attraverso i nuovi metodi di comunicazione, le aziende hanno la possibilità di catturare l'opinione del consumatore in maniera rapida per poter offrire il miglior potenziale dei prodotti e soddisfarne le esigenze. Però per estrapolare informazioni da dati testuali è necessario utilizzare diverse tecniche di analisi del

---

<sup>45</sup> Arvidsson A., Delfanti A., “Introduzione ai media digitali”;

<sup>46</sup> Application-based sono software accessibili attraverso la rete Internet che non richiedono installazioni;

<sup>47</sup> Lovett John, “Social Media metrics secrets”.

testo come il Text Mining, che ha l'obiettivo di estrarre informazioni utili dal testo contenuto in un insieme di documenti; in genere queste informazioni testuali vengono trasformate in dati strutturati attraverso l'impiego di metodi di vettorizzazione, che consentono al computer di "leggere" ed analizzare i dati testuali, riconoscendo dei concetti simili, anche se espressi con modi differenti. Per tale motivo le analisi che si sono svolte si concentrano nell'analizzare e indagare se la tecnologia del Metaverso e della Realtà Virtuale sono state comprese dal grande pubblico e soprattutto come vengono percepite. In particolare, si è posta attenzione al comunicato stampa di Mark Zuckerberg del 28 ottobre 2021 (che annunciava il cambio di nome dell'azienda in Meta) in modo da analizzare la reazione degli utenti di Twitter, confrontando le differenze tra i tweet pubblicati un mese dopo l'annuncio e i tweet pubblicati a distanza di un anno dall'annuncio. Per l'acquisizione dei dati, come già riportato precedentemente, è stato preso di riferimento il social media Twitter poiché secondo quanto riportato da Hootsuite<sup>48</sup> ci sono più di 4,70 miliardi di utenti che utilizzano i social media nel mondo, numero in continua crescita, di cui solo Twitter conta più di 300 milioni di utenti attivi al mese. Con questi numeri si riesce solo a dare un'idea parziale di quanto traffico di informazione esista sui social e che, se utilizzate in maniera utile, possono essere sfruttate in vari ambiti lavorativi, poiché molte notizie al giorno d'oggi vengono rese disponibili sui social media prima di qualsiasi altro canale tradizionale. Per riassumere, l'attività di analisi ha previsto:

- Acquisizione dei dati (tweet 2022), riguardanti la tecnologia del Metaverso e Realtà Virtuale, in particolare scaricando i testi con il software open source RapidMiner. Per lo scopo finale, inoltre, è stato necessario ottenere un dataset da Kaggle (tweet 2021);
- Preparazione e pre-processamento dei dati per poter poi applicare le analisi di Text Mining (estrarre informazioni da dati testuali non strutturati) come Sentiment Analysis e Tweet Profiler;
- Rappresentazione del testo dei tweet attraverso le tecniche di Information Retrieval; il metodo TF-IDF (Term Frequency), per poter svolgere la Topic Modelling con il metodo Latent Dirichlet Allocation (LDA).

---

<sup>48</sup> Hootsuite è un software di gestione di profili sociali che crea statistiche relative all'utilizzo dei social media, creato da Ryan Holmes nel 2008.

## 2.1 Richiami teorici: Text Mining per il trattamento dei dati non strutturati

Al giorno d'oggi, vengono generate enormi quantità di dati e in vari formati, per questo motivo non solo è utile analizzare dati di tipo strutturato (dati numerici) ma risulta anche utile estrapolare informazioni da dati testuali, come in questo caso quelli presenti sui social media. L'importanza che nel tempo è stata riconosciuta ai dati ha portato alla nascita di nuove tecniche:

- **Text Mining** o KDT (Knowledge Discovery in Texts) rappresenta un campo di ricerca di analisi del testo che aiuta a estrarre le informazioni utili da dati di testo che, altrimenti, rimarrebbero “non raccolte” e sepolte sotto una massa di dati testuali non strutturati. Per produrre conoscenza da dati non strutturati ed estrapolare le informazioni in modo efficiente, vengono utilizzate tecniche statistiche e linguistiche. Uno degli scopi principali del Text Mining è quello di classificazione dei dati testuali e valutazione dell'appartenenza di un testo ad una determinata categoria per produrre conoscenza, in particolare il TM rappresenta un'attività di ricerca di informazioni utili all'interno di una “miniera” di dati da esaminare. In genere, il Text Mining viene utilizzato per due scopi principali: individuare i dati per estrarre automaticamente le informazioni dai testi e classificare i documenti per archiviare le informazioni estratte. Quattro funzionalità comunemente utilizzate sono<sup>49</sup>:

1. Pre-processamento: processo con finalità di organizzazione e di pulizia del materiale per l'analisi. Questa fase può richiedere molto tempo, poiché i dati di testo possono presentare irregolarità e presentarsi in formati insoliti;
2. Extraction: processo per estrarre keyword per la scelta delle parole e delle frasi che identificano meglio la natura del testo;
3. Categorization: per la classificazione del testo in una o più categorie. Se queste categorie sono già note è possibile utilizzare algoritmi di apprendimento supervisionato, mentre se le categorie sono sconosciute è necessario utilizzare algoritmi di apprendimento non supervisionato per classificare il testo in gruppi;

---

<sup>49</sup> Vincenzo Cosenza, “Social Media ROI”.

4. Modelling: processo che ottimizza la previsione o semplicemente fornisce una descrizione più dettagliata tra l'unione dei componenti essenziali del testo con attributi di dati più strutturati che vengono tipicamente usati nel Data Mining.

Storicamente il Text Mining nasce nei primi anni '60, quando lo studio del linguaggio naturale era eseguito da psicologi, sociologi e linguisti. Ad oggi rappresenta un oggetto di studio della statistica testuale e dell'analisi automatica dei testi poiché lo sviluppo di diverse tecnologie ha permesso la disponibilità di enormi quantità di documenti in formato elettronico e per tale motivo è cresciuta la richiesta di avere dei metodi più efficienti che facessero risparmiare anche del tempo materiale. Negli anni, si attraversano diversi metodi di analisi, agli albori della tecnica si utilizzava uno studio quantitativo con un approccio linguistico dove venivano rilevate le frequenze delle parole, ma queste analisi presentavano dei problemi legati all'ambiguità, poiché una stessa parola può assumere significati diversi (per esempio la parola squadra fa riferimento allo strumento tecnico ma anche ad un gruppo di persone). Per sorpassare questi limiti, negli anni si sono sviluppate nuove tecniche, arrivando ad uno studio lessico-testuale molto più approfondito e complesso, che richiede procedure più specifiche e raffinate per l'analisi testuale come strumenti di carattere linguistico: la normalizzazione, la lemmatizzazione, la tokenizzazione e altri strumenti che verranno trattati nel paragrafo relativo alle analisi che sono state applicate ai dati.

- **Data Mining** (“to mine” in inglese “scavare per estrarre”), rappresenta l'applicazione di algoritmi volti ad estrarre informazioni e pattern contenuti nei dati caricandoli poi in repository, i Data Warehouse<sup>50</sup>. A differenza dei sistemi di Text Mining, quelli di Data Mining sono volti all'analisi di dati quantitativi omogenei di grandi dimensioni, per esplorare automaticamente modelli e relazioni non noti a priori nei dati o, in alternativa, per obiettivi predittivi. I passaggi tipici del Data Mining possono essere due categorie: le tecniche di apprendimento supervisionato (Classificazione, Regressione ecc.) utilizzano specifiche variabili per predire il valore nascosto di altre variabili obiettivo, e le tecniche di apprendimento non supervisionato (il Clustering, la ricerca di regole associative e di ricerca di pattern ecc.) che ricercano dei pattern interpretabili per descrivere i dati.

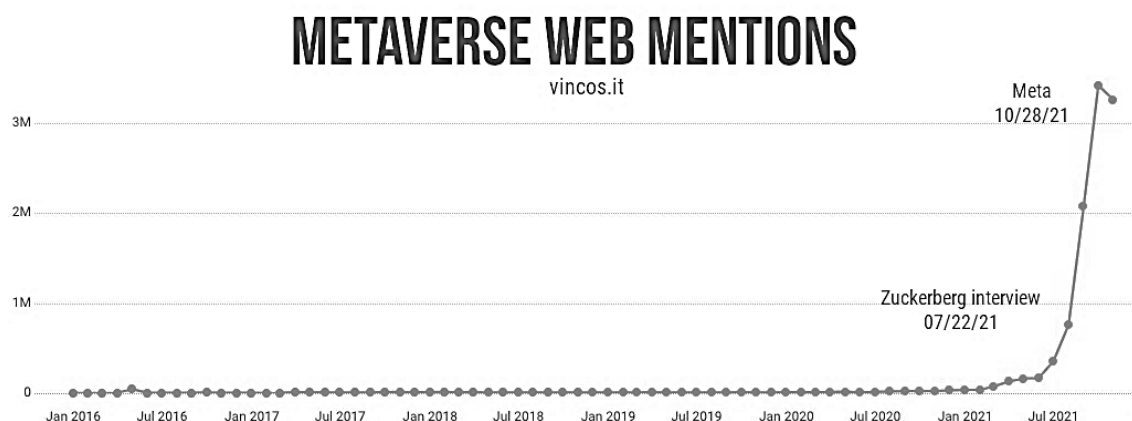
---

<sup>50</sup> Data Warehouse (“magazzino di dati”): una collezione di dati provenienti da fonti interne operazionali (DBMS) ed esterne al sistema informativo aziendale, utili ad analisi e rapporti informativi.

L'obiettivo comune di entrambe le tecniche è quello di sfruttare i dati per estrarre informazioni per produrre conoscenza utile al fine di prendere decisioni, nonostante questo però le due tecniche differiscono in maniera significativa per i vincoli di complessità dei dati e per i tempi di implementazione pratica di essi. In altre parole, il Data Mining rappresenta una tecnica già collaudata e robusta svolta su dati strutturati mentre il Text Mining è considerata una tecnica particolare di Data Mining, pensata per svolgere analisi specifiche, semantiche e sintattiche, su un insieme di documenti testuali liberi.

## 2.2 Metodologia e domanda di ricerca

Come è già stato trattato nel primo capitolo, il Web ha cambiato il modo di comunicare delle persone attraverso l'utilizzo dei social media, per questo gli utenti si scambiano opinioni, emozioni e critiche poiché si sentono liberi di esprimere le proprie opinioni. La domanda di ricerca, che ha condotto il filo delle analisi che sono state svolte, è nata dall'esigenza di voler comprendere e chiarire cosa sia il Metaverso e perché il comunicato stampa di Mark Zuckerberg ha attirato l'attenzione della maggior parte di utenti che utilizzano almeno un social media, anche di coloro che non sono a stretto contatto tutti i giorni con il settore tech. Infatti, da come si può osservare nella figura successiva (figura 2.1), le menzioni del termine Metaverso hanno iniziato ad aumentare nel periodo di ottobre 2021.



**Figura 2.1** Menzioni nel Web riguardo il Metaverso osservato negli anni, fonte Vincos Blog

Nel Web negli ultimi anni, a causa di video ironici o di incomprensioni, si è diffusa un'idea non del tutto chiara su ciò che effettivamente si sta creando con il Metaverso. Ad esempio, una domanda che sorge spontanea da parte degli utenti può essere: “Ma quindi, che cos'è il Metaverso?”, in una puntata di un podcast italiano di nome “Muschio selvaggio”, si è parlato proprio di questo argomento. Gli ospiti del podcast hanno risposto: *“Far sì che il nostro corpo sia obsoleto e che iniziamo a vivere in un mondo virtuale”*. La visione e percezione che è stata esposta in questo podcast, in realtà, sembra essere un pensiero dilagante tra gli utenti, ma non è del tutto corretta poiché il Metaverso non vuole sostituirsi con il mondo reale: una spiegazione a questa incomprensione collettiva arriva da Bradford Lee Smith, presidente di Microsoft, in un'intervista rilasciata a Reuters<sup>51</sup> (durante il Web Summit<sup>52</sup> del 2021) ha voluto chiarire alcuni punti con le seguenti parole: *“I think the Metaverse will be very big and quite important (...). We're all talking about the Metaverse as if we're entering some new dimension. This is not like dying and going to heaven. We're all going to be living in the real world with people.”* Questo per sottolineare il fatto che non si vuole “abbandonare” il mondo reale, anche perché sarebbe quasi impossibile, si tratta di rendere più “immersivo” qualcosa che già esiste. Lo stesso Zuckerberg, nella conferenza stampa di ottobre 2021, ha definito il Metaverso come la prossima rivoluzione di Internet, facendo intendere che può essere un supporto per affrontare alcune azioni quotidiane anche quando ci si trova a distanza.

## **2.2.1 Raccolta dati e pre-processamento dei tweet**

### **Raccolta dati**

Partendo dalla domanda di ricerca iniziale, la scelta di voler utilizzare due dataset si è resa necessaria per poter confrontare e osservare le differenze dell'opinione degli utenti, svolgendo le analisi sui tweet condivisi a novembre del 2021 e un dataset riguardante i tweet condivisi durante il 2022: entrambi i dataset analizzati presentano i tweet in lingua inglese.

---

<sup>51</sup> Reuters è un'agenzia di stampa britannica, fa parte del gruppo Reuters Group plc con sede a Londra;

<sup>52</sup> Web Summit è la più grande conferenza tecnologica Europea che si svolge ogni anno a Lisbona.

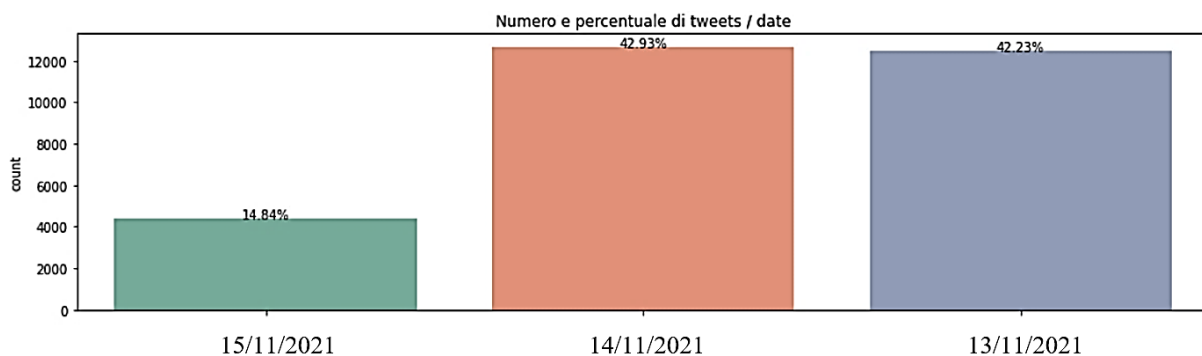
- **I tweet relativi all'arco temporale 2021** sono stati presi da Kaggle<sup>53</sup> e sono stati raccolti utilizzando l'API (Application Programming Interface) di Twitter. Un dataset in formato csv che è stato poi trasformato nel formato excel. L'approccio che è stato utilizzato per recuperare i tweet, da chi ha scaricato il seguente dataset, è stato quello di estrarli utilizzando una keyword, ovvero è stato utilizzato l'hashtag #Metaverse. Il dataset<sup>54</sup> è composto da 13 colonne:

```

1  text: testo del tweet condiviso;
2  user_name: identifica l'utente autore del tweet condiviso;
3  user_location: luogo geografico da dove è stato condiviso il tweet;
4  user_description: bio dell'utente, ovvero una stringa di testo che si trova nel profilo;
5  user_created: data in cui l'utente ha creato il suo account Twitter;
6  user_followers: rappresenta il numero di utenti che seguono il profilo;
7  user_friends: rappresenta il numero di utenti che il profilo segue;
8  user_favourites: numero di volte che il tweet è stato cliccato come preferito;
9  user_verified: indica se il profilo dell'utente è verificato;
10 date: data di creazione del tweet;
11 hashtags: hashtag utilizzati nel tweet;
12 source: riporta il dispositivo da cui è stato generato il tweet;
13 is_retweet: riporta "True" o "False" in base al tweet "retweettato" da altri utenti.

```

Composto, prima di essere pre-processato, da 50.003 tweet totali che coprono un arco temporale di tre giorni, come si può osservare dalla figura 2.2: 13/11/2021 in cui sono stati scaricati il 42,23% dei tweet, 14/11/2021 il 42,93% dei tweet, 15/11/2021 il 14,84% dei tweet.



**Figura 2.2** Percentuale di tweet scaricati per ogni data del 2021

<sup>53</sup> Kaggle, fondata nel 2010, è una consociata di Google LLC rappresenta una comunità online di data scientist e professionisti del machine learning. Consente agli utenti di trovare e pubblicare dataset;

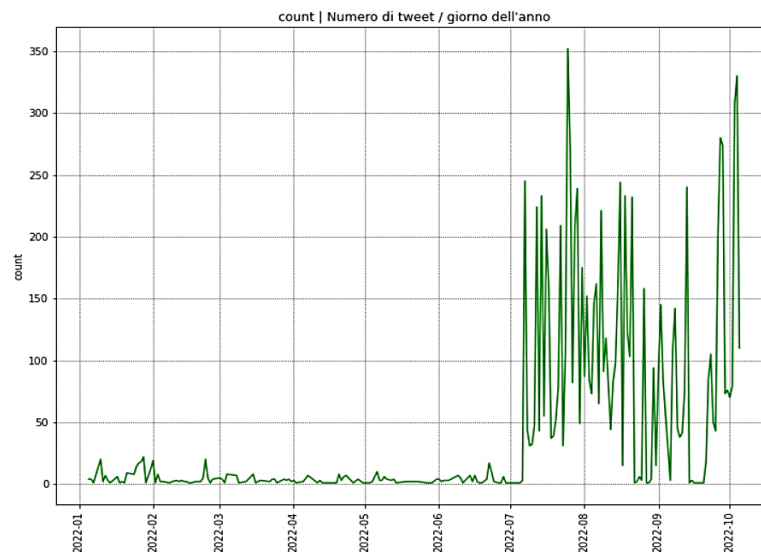
<sup>54</sup> Dataset di Kaggle: <https://www.kaggle.com/datasets/mathurinache/metaverse-tweets>.



- Mentre i tweet relativi all'arco temporale 2022 rappresentano un dataset composto, prima di essere pre-processato, da 31.258 righe totali e con 12 attributi. Essi sono i tweet scaricati una volta a settimana a partire da inizio luglio 2022 fino a inizio ottobre 2022. I seguenti dati sono stati scaricati con il software RapidMiner<sup>55</sup> via API di Twitter utilizzando i seguenti hashtag: #meta, #metaverse, #virtualreality. Il software permette di scaricare l'elenco dei tweet che contengono gli hashtag che vengono specificati, in ordine cronologico inverso rispetto alla data di inizio scaricamento:

- 1 Text: testo del tweet condiviso;
- 2 Id: identifica un numero univoco dell'utente autore;
- 3 Created-At: data di creazione del tweet;
- 4 From-User: identifica il nome dell'utente autore del tweet condiviso;
- 5 From-User-Id: identifica il numero dell'utente autore del tweet condiviso;
- 6 To-User: identifica il nome a cui è stato indirizzato un tweet condiviso da un altro utente;
- 7 To-User-Id: identifica il numero a cui è stato indirizzato un tweet condiviso da un utente;
- 8 Language: indica la lingua del tweet;
- 9 Source: riporta il dispositivo da cui è stato generato il tweet;
- 10 Geo-Location-Latitude: latitudine geografica dove è stato condiviso il tweet;
- 11 Geo-Location-Longitude: longitudine geografica dove è stato condiviso il tweet;
- 12 Retweet-Count: riporta il numero di volte che il messaggio è stato "retweettato".

I tweet scaricati coprono un arco temporale più ampio, rispetto ai tweet del 2021, poiché si è voluto osservare il pensiero degli utenti durante tutto l'arco del 2022.



**Figura 2.3** Numero di tweet che sono stati scaricati, mesi dell'anno 2022

<sup>55</sup> RapidMiner Studio rappresenta una piattaforma open source di data science molto completa, progettata per caricare e analizzare dati strutturati e non strutturati in grosse quantità. Tramite un'interfaccia facilmente usabile si può progettare il processo di analisi dei dati.

## Pre-processamento dei tweet

In una prima fase di pre-processamento del testo è stato utilizzato il software RapidMiner per la rimozione di tweet duplicati, unendo in un unico file excel i tweet che sono stati scaricati in giorni e con hashtag differenti. Inoltre, in questa fase, è stato necessario rimuovere le mentions di altri utenti (ad esempio @WorldMetaverse), e/o hashtag (#...) e/o links (http://...) che non risultano utili all'estrazione e all'analisi del testo (tweet 2021 analizzati dopo la rimozione dei duplicati 29.541; tweet 2022 analizzati dopo la rimozione dei duplicati 10.015). Successivamente, dopo la prima fase di pulizia dei dati, il pre-processamento è stato svolto con il software Orange<sup>56</sup> ed ha riguardato:

- **trasformazione** del testo per uniformarlo in caratteri minuscoli (lower case);
- **tokenizzazione** per suddividere il testo in parole, in base all'espressione regolare del testo, usando spazi bianchi senza punteggiatura;
- rimozione delle **stopwords** in inglese per eliminare tutte quelle parole tipiche della lingua che risultano ridondanti e non utili all'estrazione del significato del testo, inserendo una lista di stopwords personalizzata;
- inserimento degli **N-grams** ovvero tecnica di categorizzazione che crea n-grammi dai token realizzati precedentemente, tiene in considerazione non solo la singola parola ma anche quella che la precede, in questo caso bi-grammi;
- la **normalizzazione** con WordNetLemmatizer che applica una rete di sinonimi dei token, basati su un ampio database lessicale in inglese, applica la lemmatizzazione, cioè il processo di riduzione di una forma flessa di una parola alla sua forma canonica, detta lemma. Per esempio "I will live", "lives" vengono trasformati in "to live". È stata utilizzata tale normalizzazione perché è importante per testi come i commenti sui social media e messaggi di testo in cui prevalgono abbreviazioni ed errori di ortografia;
- è stata applicato un filtro Regexp: `\b\w{1,2}\b` per eliminare i token composti da 1 a 2 caratteri presenti nel testo.

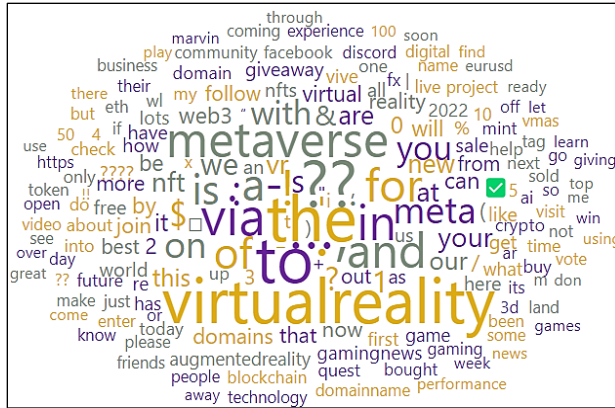
---

<sup>56</sup> Orange rappresenta un software open source di programmazione visiva basato su componenti per la visualizzazione dei dati, l'apprendimento automatico e l'analisi dei dati. I componenti di Orange sono chiamati "widget" e vengono utilizzati collegandoli tra di loro per creare un flusso di lavoro.

**Figura 2.5** Tweet 2021 dopo il pre-process

<sup>57</sup> Word Cloud: visualizzazione grafica del testo analizzato a forma di nuvola, mostra le parole più frequenti all'interno dei tweet, più un termine appare grande e maggiore è la sua frequenza all'interno del tweet;

40



**Figura 2.6** Tweet 2022 prima del pre-process



**Figura 2.7** Tweet 2022 dopo il pre-process

Mentre per i tweet del 2022, come si osserva nella figura 2.7, i temi che appaiono più frequentemente sono: l'aspetto di una tecnologia del futuro con i termini new, web3 domain, world best, best metaverse, future, project, explore, immersive e metaverse performance. Mentre l'altro tema che appare più frequentemente è quello del gaming, come nei tweet del 2021, con i seguenti termini: gamingnews web3 domain, virtualreality gamingnews, game e winner.

In questa prima fase, quindi, si possono notare i termini più frequenti presenti nei testi dei tweet e si osserva che in entrambi i dataset, com'era prevedibile in base al tema trattato, il termine metaverse è il più frequente sia nei tweet del 2021 sia nei tweet del 2022 ed in particolare per i tweet del 2022 anche il termine virtualreality. Questo lo si è capito osservando la grandezza delle parole nelle Word Cloud che risultano più grandi rispetto alle altre, ma lo si può anche osservare nella figura 2.8 e nella figura 2.9, che rappresentano le Word Count dei termini più frequenti all'interno dei tweet.

	Word	Word Count
1	metaverse	31335
2	project	14524
3	airdrop	8003
4	gamefi	6823
5	flux	4228
6	gamefi metaverse	4201

**Figura 2.8** Word Count tweet 2021

	Word	Word Count
1	virtualreality	2880
2	metaverse	2136
3	meta	1413
4	nft	823
5	new	731
6	domain	649

**Figura 2.9** Word Count tweet 2022

## 2.3 Sentiment Analysis: polarità dei tweet e risultati

Dopo aver svolto il pre-processamento dei tweet è possibile svolgere le diverse analisi di Text Mining, in particolare, è necessario scoprire cosa dicono gli utenti in termini d'idea, polarità ed intensità dei messaggi che condividono tramite un'analisi linguistica del testo. In generale è utile fare una premessa sull'analisi del testo di tweet condivisi sui social media: in particolare, i testi che vengono estratti dai social network risultano avere delle problematiche poiché, oltre a presentare frasi brevi e molto spesso incomplete, creano alcune difficoltà di analisi perché presentano una loro sintassi particolare data dall'utilizzo del parlato da parte degli utenti, ad esempio la presenza di slang, dialetti, abbreviazioni, un mix di lingue differenti, ironia, sarcasmo, giochi di parole, assenza di punteggiatura ed in generale errori ortografici. Premesso questo, ora si presentano i metodi di analisi semantica del testo che sono stati utilizzati, rispondendo ad una prima domanda iniziale: che cosa si intende per Opinion Mining e più in generale per Sentiment Analysis<sup>59</sup>. La prima fa riferimento alle motivazioni che stanno dietro ad un sentimento positivo/negativo, mentre la seconda si riferisce alla tipologia e all'intensità positiva/negativa di tale sentimento. Più nello specifico per poter svolgere la Sentiment Analysis ci sono due possibilità, l'analisi del testo supervisionata (i metodi supervised) e l'analisi non supervisionata (i metodi unsupervised). Nella seguente tesi ci si è focalizzati sulle tecniche non supervisionate poiché le categorie semantiche dei tweet non sono note a priori, ciò significa che sono emerse a posteriori applicando la Sentiment Analysis con il metodo VADER che funziona molto bene con la lingua inglese. In particolare, il metodo VADER (Valence Aware Dictionary and Sentiment Reasoner) rappresenta un lessico e uno strumento di analisi del sentimento dei testi basato su regole, progettato appositamente per i messaggi espressi sui social network. Nello specifico VADER utilizza un vocabolario lessicale relativo al sentimento, in formato testo, all'interno del quale sono presenti una serie di combinazioni di parole che vengono riconosciute ed etichettate secondo il loro orientamento semantico come positivo, negativo e neutro (un valore di sentimento medio varia nel range [-4; +4] e un valore della deviazione standard basato su quanto il sentimento stimato di quella parola varia nei testi raccolti). Inoltre, VADER restituisce quattro diversi elementi di sentimento in cui espone quanto sia positivo o negativo un sentimento rilevato all'interno del testo: la polarità positiva, la polarità negativa, la neutralità e un valore che

---

<sup>59</sup> Ceron A., Curini L., Iacus S., "Social media e Sentiment Analysis: l'evoluzione dei fenomeni sociali attraverso la rete".

viene chiamato compound che riflette il punteggio complessivo. Il valore di compound calcola la somma di tutte le valutazioni del lessico normalizzate tra -1 (estremo negativo) e +1 (estremo positivo), il valore compound è calcolato nel seguente modo:

$$compound = \frac{score}{\sqrt{score^2 + \alpha}}$$

Dove score rappresenta il valore normalizzato del sentimento di una frase, mentre  $\alpha$  rappresenta un valore di livellamento selezionato per impostazione predefinita a 15. In altre parole, il valore compound, è una normalizzazione del punteggio di una frase che si calcola dalla somma dei sentimenti presenti nel vocabolario tenendo presente la loro deviazione standard. Mentre le soglie tipiche standardizzate per classificare le frasi come positive, negative, neutre sono:

1. Sentimento positivo: punteggio compound  $\geq +0,5$
2. Sentimento neutro: punteggio compound  $> -0,5$  e punteggio compound  $< +0,5$
3. Sentimento negativo: punteggio compound  $\leq -0,5$

Per sintetizzare il metodo VADER esso calcola i sentimenti sulla base di questi punti chiave<sup>60</sup>:

- Punteggiatura: l'uso di un punto esclamativo (!) aumenta l'intensità senza modificare l'orientamento della semantica;
- Capitalizzazione: l'utilizzo di lettere maiuscole enfatizza il valore e l'intensità del sentimento aumentandone la grandezza;
- Modificatori di grado (chiamati anche intensificatori): hanno un impatto sull'intensità del sentimento aumentando o diminuendo l'intensità (ad esempio "estremamente", "marginalmente");
- Congiunzioni: l'utilizzo di congiunzioni come "ma" diventa dominante in una frase spostando in avanti la polarità del sentimento, ad esempio: "Il cibo qui è ottimo, ma il servizio è orribile", questa frase ha due sentimenti contrastanti e la congiunzione nella seconda parte determina la valutazione complessiva del sentimento.

---

<sup>60</sup> Hutto C., Gilbert E., "VADER: A Parsimonious Rule-based Model for Sentiment Analysis of Social Media Text".

## Distribuzioni di frequenza dei tweet

Dopo aver introdotto come funziona il metodo VADER per la Sentiment Analysis, si mostrano ed elaborano le parti che hanno caratterizzato il lavoro di analisi dei testi relativi al Metaverso (le analisi sono state svolte con il software open source Orange). Il primo step in Orange è stato quello di importare i dataset e, dopo aver svolto il pre-processamento spiegato precedentemente, si è aggiunto il widget “Sentiment Analysis” con il metodo VADER.

TWEET 2021

Sentiment	Text
Neutrale	13365
Positivo	15623
Negativo	553

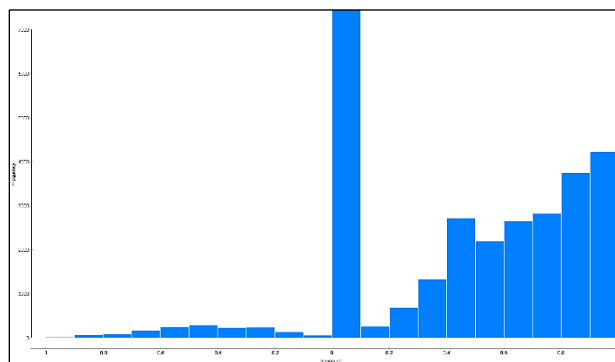
**Figura 2.10** Risultati Sentiment

TWEET 2022

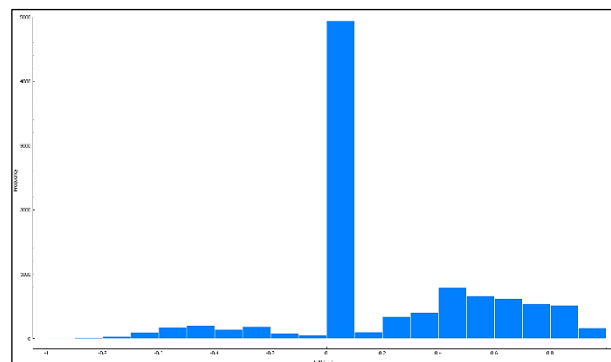
Sentiment	Text
Neutrale	7213
Positivo	2497
Negativo	305

**Figura 2.11** Risultati Sentiment

Nelle figure 2.10 e 2.11 si osservano i tweet negativi, positivi e neutri che sono stati classificati con il metodo VADER, si vede una predominanza di tweet neutrali per entrambi i dataset, questo significa che per questi tweet non prevale un sentimento specifico da poterlo collocare in negativo o positivo. A questo punto si è collegato al widget “Sentiment Analysis” il widget “Distribuzioni” per osservare il grafico di distribuzione del valore di compound relativo al punteggio complessivo della Sentiment Analysis. In questo modo i grafici (simmetrici -1; +1) che vengono riportati nelle figure 2.12 e 2.13 mostrano la distribuzione di frequenza dei valori di compound all’interno dei tweet, siccome in questo caso non sono presenti domini con classi, le barre sono visualizzate in blu da Orange.



**Figura 2.12** Distribuzione di frequenza in base al valore compound, tweet 2021



**Figura 2.13** Distribuzione di frequenza in base al valore compound, tweet 2022

Nella figura 2.12 e nella figura 2.13 si possono osservare che entrambi le colonne vicino allo zero sono molto alte rispetto alle altre, questo significa che ci sono molti tweet con un punteggio complessivo neutro (come è stato osservato precedentemente). Sia nei tweet del 2021 che in quelli del 2022 prevalgono quelli positivi rispetto a quelli negativi. A questo punto, dopo aver osservato la distribuzione di frequenza del valore di compound, risulta utile riportare alcuni esempi di tweet che sono stati condivisi dagli utenti con punteggio di compound molto alto (positivo) e un punteggio di compound molto basso (negativo), dove si può osservare che la tecnica di Sentiment Analysis con il metodo VADER ha funzionato.

## TWEET 2021

Punteggio compound = +0.9548



Đỗ Quỳnh Dương  
@hp359khaipro

In risposta a @polkawarft

Very good project with specific goals, a good product and a great team that will certainly achieve its goals. All the successes of the project related to their team. Go to the moon?

#GameFi #Metaverse

[Traduci il Tweet](#)

3:27 PM · 15 nov 2021

Punteggio compound = -0.9349



Hollywood Style 🎬 New York Speed 🚗  
@NikkoQuinto

The whole #Metaverse be like:

Let's do the whole world in post.

Ask any cinematographer, but that just doesn't work... there is no substitute for real life.

If you think planet Earth is fucked up beyond repair, I hate to break it to you: You can't fix it in post either.

[Traduci il Tweet](#)

4:19 AM · 14 nov 2021

Il tweet con un punteggio di compound positivo si riferisce al tema commerciale/gaming del Metaverso poiché utilizzando l'hashtag #GameFi fa riferimento a minigiochi che offrono all'utente di guadagnare criptovalute giocando. In questo tweet l'utente ha utilizzato parole positive facendo intendere che il gioco di cui sta parlando ha attirato la sua attenzione grazie a degli specifici obiettivi che possiede. Mentre il secondo tweet, con un punteggio di compound negativo, espone una riflessione relativa alla perplessità che questo utente ha nello sviluppo del Metaverso. Spicca il suo dissenso nella creazione dei Mondi Virtuali, cercando di esprimere che in qualche modo la vita reale è complicata ma questo non vuol dire che si troverà una soluzione per aggiustarla nel Metaverso: infatti scrive “non c'è sostituto per la vita reale. Se pensi che il pianeta terra sia incasinato irreparabilmente, odio dirtelo: non puoi nemmeno aggiustarlo via post.”



## TWEET 2022

Punteggio compound = +0.9066



Avril Lavigne  
@AvrilLavigne

Super excited to finally be entering the Web3 space!  
I've teamed up with [@metacityTabs](#) and you can mint  
premium land right now.

Check it out: [metacity.tokau.io/#/Mint?landcod...](https://metacity.tokau.io/#/Mint?landcod...)  
[#Tlabs](#) [#MetaCityTlabs](#)

3:48 AM · 16 lug 2022

Punteggio compound = -0.9027



Dragon  
@Dragonsoul9Fire

A coincidence???

Can You see the obvious?

Do You understand what it means?

REJECT IT no matter what because it means DEATH of  
the body

[#Metaverse t.co/XfTxZ5k2GC](https://metaverse.t.co/XfTxZ5k2GC)

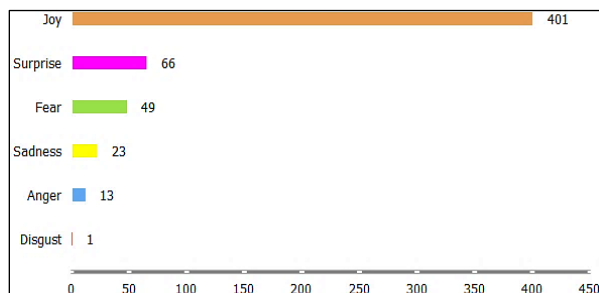
6:20 PM · 15 ago 2022

Un punteggio di compound positivo si è osservato nel tweet condiviso da Avril Lavigne, nota cantautrice canadese, in cui espone il suo entusiasmo riguardo la sua collaborazione con MetaCity. TOKAU, una società di blockchain con sede in Giappone, sta lanciando il proprio progetto nel Metaverso chiamato MetaCity la prima piattaforma incentrata sulle celebrità. In questo modo le celebrità si impegnano a fornire contenuti e partecipare alla progettazione di spazi ed eventi in MetaCity 3D. Mentre nel tweet con un punteggio di compound negativo l'utente esprime il suo dissenso per il Metaverso, scrivendo una provocazione relativa alla "morte del corpo" facendo intendere che, immergendosi nel Metaverso e nei mondi virtuali, molti utenti non riescono a vedere che si sta andando verso una negazione del proprio corpo.

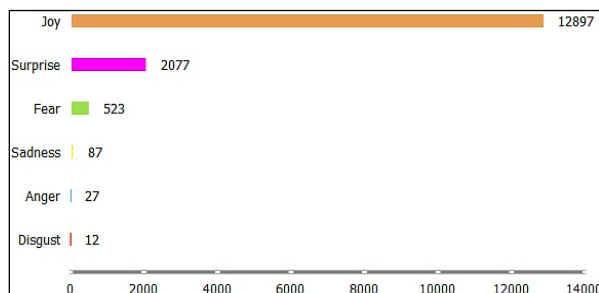
## Tweet Profiler

Successivamente per analizzare i sentimenti più rilevanti all'interno dei tweet è stato utilizzato il widget "Tweet Profiler". Come primo passaggio è stato collegato al widget "Distribuzioni" il widget "Seleziona righe" dove sono stati suddivisi i tweet negativi (valore di compound  $\leq -0,5$ ) e i tweet positivi (valore di compound  $\geq +0,5$ ). In questo modo si sono formati due percorsi, quello dei tweet negativi e quello dei tweet positivi, a cui ad entrambi è stato collegato il widget "Tweet Profiler". Il seguente widget rappresenta un operatore che è collegato alla Sentiment Analysis e, per recuperare le informazioni sul sentimento dei tweet, viene estrapolato dal server dove è presente un modello di classificazione del testo per le emozioni che calcola le probabilità e/o i punteggi delle emozioni, in questo caso è stato utilizzato Ekman Multi-class dove ha prodotto classi di emozioni categoriche. Le emozioni rilevate vengono rappresentate collegando il widget Box Plot.

## TWEET 2021



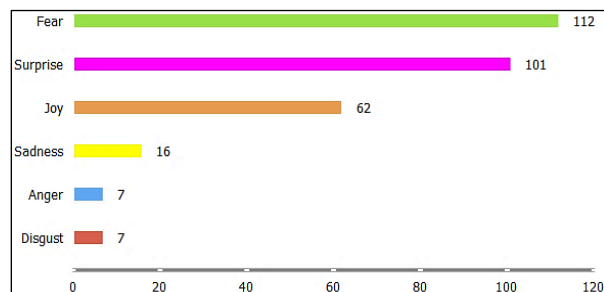
**Figura 2.14** Tweet profiler negativi



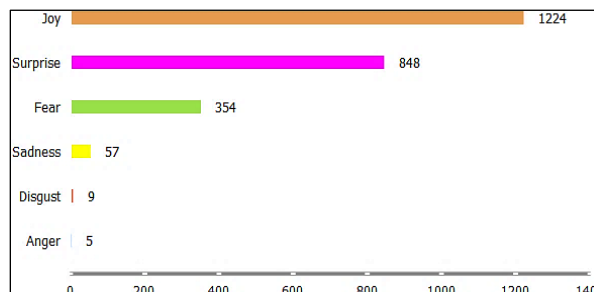
**Figura 2.15** Tweet profiler positivi

Tweet 2021 negativi: su 553 tweet classificati come negativi dal metodo VADER predomina un sentimento di gioia in 401 tweet, questo potrebbe significare che gli utenti, nonostante abbiano espresso un parere negativo riguardo al Metaverso, siano inclini ad esprimere emozioni positive essendo tweet condivisi subito dopo il comunicato stampa di Mark Zuckerberg. Gli altri sentimenti rilevanti sono sorpresa, paura e tristezza che possono essere associati a sentimenti per cui gli utenti non sanno cosa aspettarsi dal Metaverso; Tweet 2021 positivi: in 12897 tweet prevale un sentimento di gioia mentre le altre emozioni rilevanti sono sorpresa e paura. Poco rilevanti sono le emozioni tristezza, rabbia e disgusto.

## TWEET 2022



**Figura 2.16** Tweet profiler negativi



**Figura 2.17** Tweet profiler positivi

Tweet 2022 negativi: in 112 tweet si osserva che l'emozione predominante è la paura e in 101 la sorpresa, da questo si osserva immediatamente la differenza con i tweet del 2021 classificati come negativi. Si può dedurre che dopo quasi un anno dal comunicato stampa di Mark Zuckerberg, in alcuni utenti predomina un sentimento di "paura" nei confronti di ciò che può essere creato con il Metaverso ma allo stesso tempo anche di sorpresa per alcuni; Tweet 2022 positivi: nei tweet classificati come positivi dal metodo VADER prevale la gioia e la sorpresa e quest'ultima sembra essere dominante sia nei tweet classificati come positivi che negativi.

A questo punto, per osservare i termini più frequenti in entrambi le polarità dei tweet, si riportano le Word Cloud per mostrare quali termini vengono più utilizzati dagli utenti nell'esporre il proprio pensiero per entrambi le polarità.

In questo modo si osserva che nella figura 2.18 i termini più utilizzati sono per lo più dispregiativi come: hell, facebook death, death reality, kill e shit meta. La maggior parte delle parole riportate esprimono odio nei confronti della realizzazione del Metaverso ma in particolare l'espressione "facebook death", come se gli utenti volessero una chiusura di Facebook, fa pensare che il Metaverso appartenga solo a Meta ma in realtà non è così. Mentre i termini positivi più utilizzati sono: project future, flux airdrop, airdrop gamefi e good future. La maggior parte di questi fa riferimento al tema gaming e commerciale, per questo gli utenti hanno una prospettiva di speranza nei confronti dello sviluppo di questi due settori.

[illegible][illegible]

48

Per quanto riguarda i tweet del 2022 la differenza si nota rispetto ai tweet del 2021. Questo fatto potrebbe far riflettere, poiché solamente dopo un anno dal lancio di Meta, le idee degli utenti evolvono con una rapidità non indifferente. I termini più utilizzati nei tweet con un Sentiment negativo sono: platform failing current, failing user, virtual reality e meta; questi termini si concentrano più sull'aspetto di non fiducia della tecnologia tanto da pensare che fallirà. Mentre i termini più utilizzati in positivo sono: best, free, best metaverse, follow like, nft e join. Il significato di questi termini risulta molto simile ai tweet del 2021, questo fa pensare che gli utenti che hanno un approccio positivo alla tecnologia del Metaverso non abbiano abbandonato tale pensiero con il tempo.

## **2.4 Algoritmi Unsupervised Machine Learning: Topic Modelling**

Nel seguente paragrafo si illustra una delle analisi di Unsupervised Machine Learning che è stata applicata al testo dei tweet per estrarre argomenti, detti topic, siccome si tratta di un approccio non supervisionato non necessita di un dataset già etichettato. Il metodo in questione fa riferimento alla branca dei Topic Models detta Topic Modelling con la tecnica Latent Dirichlet Allocation (LDA)<sup>61</sup>: rappresenta una tecnica popolare di analisi testuale per estrarre le strutture semantiche latenti/nascoste all'interno del testo di documenti, raggruppandoli tra di loro in base al loro argomento principale (il termine latente indica qualcosa che esiste ma non è visibile direttamente). Con questa tecnica si suppone che esista un modello statistico basato su una distribuzione chiamata Dirichlet, nella teoria della probabilità. Questi metodi rappresentano una famiglia di processi stocastici le cui realizzazioni sono distribuzioni di probabilità ovvero basate sull'occorrenza delle parole in un testo. Ogni testo è composto da diversi argomenti latenti, chiamati topic a cui viene associata una sequenza di parole. Ogni topic, che è presente nel testo, rappresenta una distribuzione multinomiale sulle parole e queste vengono raggruppate all'interno di un vocabolario che è definito in precedenza sulla base dei testi analizzati. In altre parole, si assume la presenza di una distribuzione di topic nel corpus dei tweet e ad ognuno di questi è associata una sequenza di parole. Ora che si è illustrata la tecnica che è stata utilizzata si espongono le analisi: una volta che è stato pre-processato il testo (come è stato illustrato in

---

<sup>61</sup> SYSDATA, "Topic Modeling con Latent Dirichlet Allocation".

precedenza), si è creato un corpus formato da tutti i tweet che devono essere trasformati in vettori e, per continuare le analisi, si è utilizzato il corpus con l'aiuto di Gensim<sup>62</sup> di Python (elaborato in Google Colab). Per prima cosa si è dovuto costruire un vocabolario dei tweet trasformati, ovvero rappresentarli in vettori e per fare ciò è stato utilizzato il metodo TF-IDF (Term Frequency - Inverse Document Frequency), tale metodo misura l'importanza di un termine rispetto ad un documento e aumenta proporzionalmente al numero di volte che una determinata parola è contenuta nel documento ma cresce in maniera inversamente proporzionale con la frequenza della parola nella collezione di documenti. Questo accade perché si vuole dare più importanza ai termini che compaiono nel documento, ma che in generale sono poco frequenti. La formulazione matematica della funzione di peso di TF-IDF può essere vista come una combinazione di due fattori<sup>63</sup>:

1. Numero di parole presenti nel documento (nei tweet):

$$TF_{i,j} = \frac{n_{i,j}}{|d_j|}$$

$n_{i,j}$  rappresenta il numero di occorrenze della parola  $i$  nel tweet  $j$ ;

$|d_j|$  rappresenta la dimensione, espressa in numero di parole, del tweet  $j$ .

2. Importanza generale della parola  $i$  nel corpus:

$$IDF_i = \log \frac{|D|}{|\{d : i \in d\}|}$$

$|D|$  rappresenta il numero di tweet  $j$  nel corpus;

il denominatore rappresenta il numero di tweet che contengono il termine  $i$ .

Quindi la formula finale risulta essere:

$$(TF - IDF)_{i,j} = TF_{i,j} \times IDF_i$$

Per sintetizzare la tecnica di vettorizzazione dei tweet e l'applicazione del modello LDA si riportano i codici utilizzati.

```
corpus = [dictionary.doc2bow(text) for text in texts] #corpus dei tweet
tfidf = models.TfidfModel(corpus) #inizializzazione modello tfidf
corpus_tfidf = tfidf[corpus] #per trasformare il testo in vettori
```

<sup>62</sup> Gensim è un progetto open source di Python che lavora con dati testuali, quindi non strutturati, e permette di svolgerci sopra molti task di Natural Language Processing;

<sup>63</sup> Scielo Analytics "Latent Dirichlet Allocation complement in the vector space model".

A questo punto, quando i tweet sono stati vettorizzati con il metodo TF-IDF, si è passati all'addestramento del Topic Model con la tecnica Latent Dirichlet Allocation (LDA). Per il suo calcolo, il modello LDA prende come input i valori TF-IDF dai documenti e il numero di topic che sono stati forniti (per iniziare in entrambi i dataset sono stati scelti cinque topic), successivamente il modello generativo determina la probabilità di appartenenza di un documento per ogni topic:

```
topics = 5  #numero di topic da estrarre
lda_model = models.LdaModel(corpus, id2word=dictionary, num_topics=topics,
passes=30, iterations=100, chunksize=10000)
```

All'interno della creazione del modello si devono specificare determinati elementi quali:

- `corpus` rappresentano i tweet trasformati in formato vettoriale;
- `id2word` dizionario che definisce il mapping degli id con le parole;
- `passes` controlla la frequenza con cui si addestra il modello sull'intero corpus;
- `iterations` controlla la frequenza con cui si ripete un particolare ciclo su ciascun documento;
- `chunksize` numero di documenti da elaborare in ciascun blocco durante la fase di train.

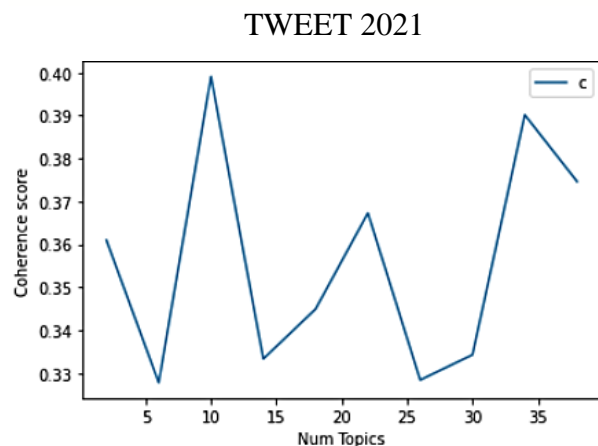
Una volta terminato l'addestramento del modello, si vuole capire se il numero di topic estratto risulta essere corretto misurando la coerenza dei topic. La coerenza (`coherence`) valuta un singolo topic misurando il grado di somiglianza semantica tra le parole che hanno un elevato punteggio nel topic. Questa misurazione aiuta a distinguere tra argomenti semanticamente interpretabili e argomenti che sono artefatti di inferenza statistica. La misura `c_v` si basa su una "finestra scorrevole", la segmentazione delle parole principali e una misura di conferma indiretta che utilizza l'informazione reciproca puntuale normalizzata (NPMI) e la somiglianza del coseno.

```
coherence_model_lda = CoherenceModel(model=lda, texts=cleaned_summary_list,
dictionary=dictionary, coherence='c_v')
coherence_lda = coherence_model_lda.get_coherence()
```

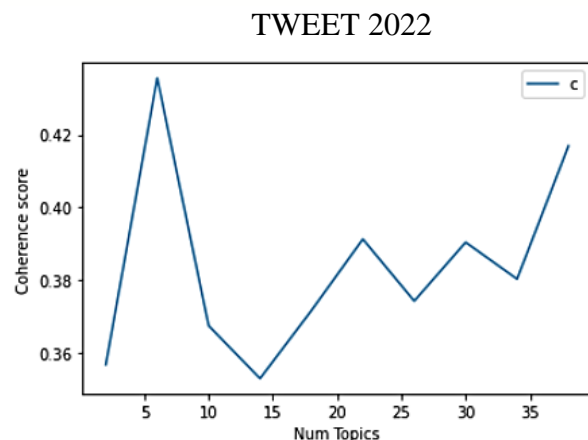
Ora che si è ottenuto il punteggio di coerenza per il modello LDA, si osserva se tale punteggio può essere migliorato aumentando o diminuendo il numero di topic da estrarre applicando la funzione seguente, osservando l'andamento della coerenza in base al numero di topic

```
def compute_coherence_values(dictionary, corpus, texts, limit, start=2,
step=3)
```

Il punteggio più elevato di coerenza dei topic genererà un buon modello.



**Figura 2.22** I 10 topic danno il punteggio di coerenza più alto di 0.3991



**Figura 2.23** I 6 topic danno il punteggio di coerenza più alto di 0.4356

Dopo aver migliorato il numero di topic da estrarre, si cerca di capire se possono essere ricondotti ad una determinata categoria.

#### TWEET 2021

*Topic 1:* “Progetto GameFi”. È rappresentato da parole riguardanti il tema del gaming in particolare GameFi riguarda i giochi blockchain che offrono degli incentivi economici (NFT o crypto) ai giocatori.

*Topic 2:* “Airdrop di Crypto”. Nel tema crypto, gli airdrop rappresentano un metodo di marketing che prevede la distribuzione gratuita di nuove criptovalute a dei wallet al fine di avviare dei progetti in cambio della condivisione di un post sui social media.

*Topic 3:* “NFT community”. È rappresentato da parole relative al mondo commerciale, in particolare gli NFT (Non Fungible Token) che in italiano significa gettone non copiabile ovvero qualcosa di unico che non può essere sostituito.

*Topic 4:* “Criptovalute e futuro”. Termini che si riferiscono alle criptovalute che sono token (gettoni virtuali) e sono fungibili ovvero si possono moltiplicare.

*Topic 5:* “Potenziale del Metaverso”. Differenti approcci utilizzati per esporne il potenziale.

*Topic 6:* “Approccio innovativo e commerciale”.

*Topic 7:* “Bitcoin e token”. Topic relativo ai Bitcoin (criptovalute con un sistema di pagamento valutario internazionale), mentre i token sono gettoni virtuali usati come moneta virtuale.

*Topic 8:* “Mercato del trading”. Termini inerenti al settore degli investimenti, il trading online.



*Topic 9: “Member e community”.* Termini che si riferiscono ai giochi del settore gaming interagendo con la community che si crea.

*Topic 10: “Futuro e mercato del Metaverso”.* In generale fa riferimento a termini relativi al mercato nel Metaverso e se in questo mondo lo renderà migliore.

#### Termini per Topic

Topic1	project, metaverse, airdrop, gamefi, great, opportunity, future, success, thanks, amazing, hopefully, believe, support, strong, wonderful, better, happy, always, metaverseinu, development
Topic2	metaverse, crypto, bitcoin, cryptocurrency, memecoin, interested, until, look, space, altcoin, bullish, shiba, ethereum, dogelon, bscgems, early, coin, binancesmartchain, according, after
Topic3	metaverse, nftcommunity, nftart, nftcollector, nftartists, nftcollectors, crypto, ethereum, opensea, blockchain, nftcollection, solana, collection, bitcoin, nftdrop, check, nftcollectibles, nftartist, openseanft, beginning
Topic4	metaverse, gaming, crypto, about, blockchain, playtoearn, there, people, future, world, nftgaming, gamefi, token, first, their, money, facebook, virtual, before, community
Topic5	metaverse, listed, potential, twitter, holder, telegram, marketcap, list, 🗨️ talking, yooshi, profit, check, token, memecoin, world, going, already, still, hello, dogelon
Topic6	metaverse, airdrop, project, their, multichain, crosschain, going, definitely, innovative, metaverseinu, gamefi, towards, inspiration, approach, achieving, goal, giveaway, vision, congratulation, crypto
Topic7	metaverse, market, price, holder, address, bitcoin, contract, current, pugcoin, token, domain, casino, mean, supply, top1nft, discount, building, million, value, 0.0005
Topic8	metaverse, gaming, terrabots, asset, holder, farming, pitswap, pitsafe, pitfarm, pitsafe, pitcharts, pitstop, pitlottery, charting, pittracker, security, lunatic, manager, tokenarmy, about, coming
Topic9	metaverse, playtoearn, gamefi, presale, nftdrop, binancesmartchain, growing, project, amazing, maticverse, quite, supported, extraordinary, community, which, socialfi, member, show, rapidly, kmnos
Topic10	metaverse, project, future, really, playtoearn, airdrop, gamefi, think, nftgame, strong, bright, giveaway, because, bscgems, planned, super, transparent, perfect, excellent, happy

Topic	Doc_Count	Total_Docs_Perc
1	3989	13.50
2	1427	4.83
3	1825	6.18
4	1695	5.74
5	1803	6.10
6	1633	5.53
7	2415	8.18
8	3889	13.16
9	1147	3.88
10	9718	32.90

Si sono osservati i conteggi dei documenti per ciascun topic e la loro percentuale nel corpus. Il topic maggiormente rappresentato è il numero 10 “Futuro e mercato del Metaverso”, mentre il meno rappresentato è il numero 9 “Member e community”.

## TWEET 2022

*Topic 1: “Metaverso e business creators”.* È rappresentato da parole relative alla condivisione e alla creazione di account di creators nel Metaverso.

*Topic 2: “Esperienza nel Metaverso”.* All’interno di questo topic prevalgono termini relativi all’esperienza della tecnologia del Metaverso, quindi, inerenti all’utilizzo di un visore e di conseguenza come è percepita la Realtà Virtuale da parte dell’utente.

*Topic 3: “Meta account”.* Il topic fa riferimento a parole relative all’account di Meta, alla sicurezza e alle informazioni di privacy.

*Topic 4: “Sicurezza nella Realtà Virtuale”.* Il topic fa riferimento alla sicurezza di privacy e di comportamento all’interno della realtà virtuale.

*Topic 5: “Account personale nel Metaverso”.* Alcuni termini risultano simili al *Topic 3* inerenti agli account personali degli utenti.



*Topic 6: “Futuro del Metaverso”.* Sono racchiuse parole relative all’idea iniziale per chi si avvicina per la prima volta alla tecnologia Metaverso, proiettando la tecnologia nel futuro.

#### Termini per Topic

<b>Topic1</b>	black, business, community, creator, history, gaming, update, world, month, every, making, their, story, program, facebook, voice, leader, since, investing, celebrating
<b>Topic2</b>	virtualreality, could, happy, would, issue, assist, about, reality, virtual, metaverse, headset, world, first, augmentedreality, health, technology, training, experience, model, quest
<b>Topic3</b>	check, their, facebook, using, community, after, center, please, forward, advice, still, shared, shutting, restaurant, grateful, paying, coogan, iconic, door, domain
<b>Topic4</b>	about, share, still, center, detail, listed, issue, please, option, filling, problem, galaxy, quest, disney, tale, spring, headphonethoughts, boris, brejcha, island
<b>Topic5</b>	account, please, about, think, facebook, thank, disabled, secure, profile, sharing, taken, experiencing, problem, there, visit, everyone, report, review, using, creating
<b>Topic6</b>	metaverse, virtualreality, follow, today, first, crypto, future, blockchain, token, augmentedreality, discord, project, enter, video, friend, world, coming, event, marvin, learn

Topic	Doc_Count	Total_Docs_Perc
1	595	5.94
2	2586	25.82
3	1626	16.24
4	518	5.17
5	252	2.52
6	4438	44.31

Il topic maggiormente rappresentato è il numero 6 “Futuro del Metaverso”, mentre il meno rappresentato è il numero 5 “Account personale nel Metaverso”.

## Risultati LDA

L’analisi Topic Modelling con il metodo LDA ha lo scopo di individuare i topic presenti nei dati scaricati da Twitter. È necessario fare una considerazione: nei topic individuati dal modello viene catturata la correlazione tra le parole ma non tra i topic, questo perché vengono estratti i topic da una singola distribuzione Dirichlet e il modello non riesce a gestire quei casi in cui i topic si sovrappongono. Si ottengono quindi informazioni importanti, ma poco variabili. Detto ciò, si sono ottenuti dei risultati che rilevano una differenza nella percezione, da parte degli utenti, del Metaverso relativi ai tweet pubblicati nel 2021 e nel 2022. I tweet pubblicati nel 2021 fanno riferimento ad argomenti vasti quali: il gaming, crypto e NFT, tecnologia del Metaverso e del futuro del mercato finanziario relativo al Metaverso. Questo fa comprendere che, dal momento in cui Mark Zuckerberg ha cambiato il nome della società in Meta, questo abbia scaturito la curiosità degli utenti tanto che hanno condiviso nei loro tweet diversi argomenti che toccano la tecnologia del Metaverso e della Realtà Virtuale. Mentre nei tweet pubblicati nel 2022 sembrerebbe che gli utenti si stanno chiedendo come funziona realmente la tecnologia del Metaverso, domandandosi se è possibile creare il proprio account personale e le possibili interazioni che si possono avere con gli altri utenti.

# Capitolo 3

## Conclusioni e sviluppi futuri

Per concludere il seguente lavoro di tesi, in questo capitolo si vuole: esporre i principali vantaggi e gli ostacoli incontrati nello sviluppo delle analisi, riassumere i risultati delle analisi che sono state svolte con i rispettivi confronti dei dataset utilizzati e per concludere esporre le considerazioni e le prospettive future della tecnologia del Metaverso. Innanzitutto, lo scopo principale dell'elaborato è stato quello di utilizzare un approccio quantitativo nello sviluppo dell'analisi testuale, riportando la teoria degli strumenti analitici che sono stati utilizzati ed esaminando i risultati rispetto alla domanda di ricerca iniziale.

### Vantaggi e svantaggi nella raccolta dati provenienti da Twitter

La scelta di utilizzare come metodo di raccolta dati il software open source RapidMiner tramite API di Twitter porta con sé numerosi vantaggi e diversi svantaggi computazionali, ecco quindi che se ne elencano alcuni. Per quanto riguarda i vantaggi, il Social Network Twitter rappresenta una fonte di dati più utilizzata in assoluto, questo per una serie di fattori che caratterizzano l'applicazione. Come fattore principale vi è una percentuale di profili pubblici decisamente più elevata rispetto agli altri Social Network presenti sul mercato, questo significa che, chi ha la necessità di scaricare i dati, può avere a disposizione il contenuto dei tweet attraverso API che viene fornito direttamente da Twitter, senza dover chiedere nessuna licenza per ottenere i dati. Un'altra particolarità ed un grosso vantaggio nello scaricare i dati da Twitter è rappresentato dall'utilizzo degli hashtag, con essi i diversi tweet pubblicati da utenti differenti possono essere rintracciati e possono essere letti da un pubblico molto più vasto. Questo fattore rafforza le tematiche che vengono affrontate e condivise dagli utenti, rendendo Twitter uno spazio pubblico dove le opinioni degli utenti si possono modificare attraverso lo scambio di pensieri. In altre parole, al social Network Twitter gli si attribuisce la capacità<sup>64</sup> di identificare quali saranno i temi oggetto di discussione sugli altri media e l'abilità di regolare il traffico di informazioni tra

---

<sup>64</sup> Ceron A., Curini L., Iacus S., "Social media e Sentiment Analysis: l'evoluzione dei fenomeni sociali".

tutte le piattaforme presenti online. Invece, per quanto riguarda gli svantaggi che si possono riscontrare scaricando i dati da Twitter, sono molteplici: avere a disposizione una grande quantità di dati, a volte, non significa ottenere informazioni migliori bisogna cercare di ponderarne l'utilizzo e capire se effettivamente la provenienza dei dati è verificata. Un'altra questione importante è la rappresentazione dei dati, nello specifico caso di Twitter, i pensieri dei diversi utenti non rappresentano in modo omogeneo l'opinione dell'intera popolazione, ma appartengono a specifiche fasce demografiche da cui provengono i tweet. Per tale motivo è bene specificare la provenienza geografica dei tweet, se presente, poiché si possono creare problemi nell'estendere i risultati all'intera società. I pensieri che vengono espressi online non sono rappresentativi né dell'intera popolazione, perché non tutte le persone hanno un profilo Social, né dell'insieme degli utenti del Social Network preso in considerazione, poiché solo gli utenti che decidono di esporre il proprio pensiero online verrà rappresentato. Uno studio condotto da Yannick Bouissiere, fondatore di ProInfluent.com agenzia specializzata nella generazione di clienti professionali, dichiara che non tutti gli utenti che hanno un account Twitter pubblicano messaggi e nemmeno tutti gli utenti che espongono la propria opinione trattano ogni argomento; infatti, nel secondo semestre del 2022 Twitter ha dichiarato che ci sono più di 1,3 miliardi di account sulla piattaforma e tra di essi sono più di 300 milioni gli utenti attivi al mese. A questo, però, si deve considerare il fatto che vengono pubblicati online post relativi a pensieri non veritieri oppure post pubblicati da account fasulli. Ci sono casi in cui un utente possiede più account, oppure che un solo account viene utilizzato da più utenti o in cui alcuni account sono bot ovvero che producono contenuti in maniera automatica senza coinvolgere delle persone reali. Per concludere, un altro degli svantaggi presenti nella raccolta dati da Twitter è rappresentato dall'accessibilità dei dati. Molto spesso quando si scaricano dati dalle API di Twitter non sempre si ottengono dati completi, ma presentano alcuni dati mancanti relativi, ad esempio, alla geolocalizzazione del tweet. Per ottenere dati completi per realizzare analisi più specifiche si deve richiedere l'aiuto di aziende specializzate, in modo da ricevere autorizzazioni adeguate. In particolare, si elencano le tipologie di problematiche che si sono incontrate nell'utilizzo dei dati scaricati dal Social Network Twitter:

- Avendo selezionato inizialmente gli hashtag per l'estrazione dei tweet può capitare che alcuni, nonostante siano in linea con il tema d'interesse, non vengono inclusi. Può inoltre accadere che tra i tweet estratti molti siano irrilevanti al fine della ricerca, questo può

accadere a causa della generalizzazione dell'identificazione dei tweet; infatti, secondo una stima condotta da Ceron, Curini e Iacus<sup>65</sup> circa l'11% dei tweet che vengono estratti da Twitter non sono inerenti al tema d'interesse;

- Siccome Twitter rappresenta una piattaforma di microblog, l'architettura dell'applicazione non permette di generare conversazioni articolate con più dettagli di opinione personale;
- Infine, pochi tweet sono resi geolocalizzati, questo limita alcune analisi che si possono condurre basandosi sulla localizzazione degli autori dei messaggi pubblicati su Twitter, questo accade poiché l'utente ha la possibilità di negare l'accesso a questa informazione.

Elencati i vantaggi e gli svantaggi, nell'ottenere i dati pubblicati su Twitter, è utile considerare che anche se i dati ottenuti sono accurati, è giusto dare importanza alla loro interpretazione che non risulta sempre così immediata poiché come dice Andreas Weigend, esperto di Big Data e autore del libro "Data for the people"<sup>66</sup>: *"I dati sono il nuovo olio, e come l'olio devono essere raffinati prima di essere utilizzati."*

## Confronto dei risultati

Per quanto riguarda l'interpretazione dei risultati, si evidenziano gli aspetti principali e le differenze riscontrate nei tweet pubblicati nel 2021 e nel 2022. Dal punto di vista quantitativo si nota subito la differenza dei due dataset: nel dataset scaricato da Kaggle, sono stati registrati dei picchi del numero di tweet pubblicati nel 2021 (29.541 tweet totali in lingua inglese, pubblicati in soli 3 giorni) relativi al tema del Metaverso, in particolare a metà novembre 2021 in corrispondenza del comunicato stampa di Mark Zuckerberg avvenuto il 28 ottobre 2021 per il cambio del nome della sua azienda in Meta. Mentre i tweet del 2022, sono stati pubblicati dal mese di gennaio al mese di ottobre (10.015 tweet totali in lingua inglese, relativi a 9 mesi), periodo storico in cui per la prima volta l'azienda Meta, subisce un calo dei ricavi. Infatti, nel secondo trimestre del 2022, i ricavi sono stati di 28,8 miliardi di dollari<sup>67</sup>, inferiori dell'1% rispetto allo stesso periodo del 2021. I passaggi successivi sono stati relativi all'approfondimento della comprensione del contenuto dei tweet, applicando strumenti di analisi del testo. Per prima

---

<sup>65</sup> Ceron A., Curini L., Iacus S., "Social media e Sentiment Analysis: l'evoluzione dei fenomeni sociali";

<sup>66</sup> Weigend A., "Data for the people: how to make our post-privacy economy work for you";

<sup>67</sup> "Meta registra una riduzione dei propri ricavi" <https://www.ilpost.it/2022/07/28/meta-facebook-calo-ricavi/>.

cosa si sono create delle Word Cloud per evidenziare i termini che, con la frequenza maggiore, compaiono all'interno del corpo dei tweet. Attraverso questa visualizzazione dei dati si è osservato che i vocaboli più frequenti fanno riferimento al tema del gaming e quello commerciale del Metaverso. Successivamente, attraverso il software Orange si sono create le Word Count, ovvero sono state considerate le frequenze più alte dei termini osservati nelle Word Cloud al fine di capire i temi più rilevanti in base ai termini utilizzati. A questo punto delle analisi è già emerso che i temi rilevanti su cui gli utenti esprimono il proprio pensiero sono: come poter guadagnare nel Metaverso attraverso le monete virtuali e come espandere la propria rete utilizzando le attività del Metaverso, accessibili a tutti, come il settore del Gaming. Successivamente è stata applicata l'analisi semantica qualitativa condotta con il metodo VADER per evidenziare i punteggi positivi, negativi e neutrali dei tweet per poter applicare il widget Tweet Profiler e osservare i sentimenti predominanti all'interno dei messaggi degli utenti. Nei tweet relativi al 2021 prevalgono dei sentimenti positivi come la gioia e la sorpresa, essendo pensieri condivisi dagli utenti subito dopo il comunicato stampa di Mark Zuckerberg, anche se questi potrebbero essere influenzati proprio dal comunicato poiché si sposta l'attenzione solamente sul concetto Meta=Metaverso, ma bisogna specificare che il Metaverso non è di dominio dell'azienda Meta. Mentre per quanto riguarda i tweet del 2022 prevalgono sentimenti come gioia, paura e sorpresa e in questi messaggi sembrerebbe che l'attenzione si è spostata dal comunicato stampa di Meta concentrandosi invece più sull'aspetto del Metaverso stesso e sul futuro della tecnologia. La relativa Sentiment Analysis condotta con il metodo VADER, ha permesso di sottolineare la varietà dei tweet che sono stati pubblicati, ampliandone la visione finale del tema. Infine, con la tecnica di Topic Modelling (Unsupervised Machine Learning) si è fatto emergere a posteriori i topic latenti tramite la tecnica Latent Dirichlet Allocation. I risultati che si sono ottenuti con questa tecnica hanno confermato quello che si era osservato all'inizio, ovvero che nei tweet pubblicati nel 2021 gli argomenti maggiormente trattati sono relativi alla curiosità degli utenti generata dal cambio di nome della società in Meta, e al contrario molti utenti ironizzano e sono in disaccordo con la tecnologia verso cui si stanno dirigendo le grandi società. Mentre, per i tweet pubblicati nel 2022, gli utenti stanno nutrendo un reale interesse nel conoscere cosa accadrà in futuro, alimentando diverse questioni relative al proprio account personale e le possibili interazioni che si possono avere con gli altri utenti.

Per concludere, i risultati ottenuti sono stati utili per confermare le numerose potenzialità che si possono ottenere utilizzando le tecniche di Text Mining su dati testuali.

## **Considerazioni finali**

Si ritiene necessario esprimere le considerazioni e le prospettive future della tecnologia del Metaverso e della Realtà Virtuale. Si può dire che il Metaverso, da pochi anni a questa parte, viene spesso descritto come il prossimo futuro. Però, ad ostacolare questo pensiero, vi sono dei problemi tecnici<sup>68</sup>: il primo risulta essere la modalità di accesso nel Metaverso e il secondo l'infrastruttura di Internet. Per quanto riguarda la modalità di accesso al Metaverso si fa riferimento all'utilizzo dei visori che risultano necessari per la creazione della Realtà Virtuale e per un'immersione a 360° nel Metaverso. Però ad oggi, nonostante i continui progressi nella tecnologia dei visori, risultano esserci ancora alcuni problemi per fare in modo che più persone possibili acquistino tali dispositivi, ecco alcuni problemi che vengono riscontrati dagli utenti: il senso di estraneità che comporta l'idea di indossare un visore per immergersi nel Metaverso e il limite di doverlo utilizzare in ambienti sicuri e protetti come la propria abitazione e non, come accade con lo smartphone, in qualsiasi contesto in cui si possono trovare gli utenti. Il secondo aspetto invece risulta essere quello relativo all'infrastruttura, come la velocità di Internet. L'attuale infrastruttura di Internet non risulta essere sufficiente per permettere di raggiungere una collisione tra il digitale e l'analogico, ad esempio una videochiamata su qualsiasi social media può arrivare ad avere una latenza<sup>69</sup> di 75-150 millisecondi, ossia il tempo impiegato dal segnale audio e video per arrivare da un punto all'altro di Internet e tornare indietro. Mentre alcuni studi espongono che un'esperienza immersiva nel Metaverso dovrebbe essere al di sotto di 12 millisecondi, per questo le recenti infrastrutture di Internet non risultano essere ancora all'altezza di questo traguardo, soprattutto nelle aree meno sviluppate del pianeta Terra. Per questo dovrebbe verificarsi un salto tecnologico non indifferente che però porterebbe ad avere delle conseguenze sull'ambiente, sia per le maggiori emissioni di CO2 scaturite dal settore digitale, sia per l'estrazione di materie prime e metalli rari che risultano essere indispensabili per la creazione

---

<sup>68</sup> Nasi F. "Across the metaverse: un'introduzione al metaverso";

<sup>69</sup> La latenza indica il tempo impiegato dal proprio segnale video e audio a raggiungere il server, e di conseguenza quanto tempo ha impiegato quello dell'interlocutore ad arrivare all'altro interlocutore.

delle nuove tecnologie<sup>70</sup>. Oltre a questi due elementi tecnici che necessitano di essere ulteriormente sviluppati, vi è una un'altra questione meno tecnica da affrontare per lo sviluppo del Metaverso. Riferita alla domanda di ricerca iniziale della seguente tesi, ovvero se gli utenti abbiano capito realmente di che cosa si tratta e se nutrono l'interesse ad utilizzare il Metaverso negli ambienti in cui verrà sviluppato. Uno dei termini chiave per far riferimento al Metaverso è "interattività"<sup>71</sup>, però è anche legittimo che da parte degli utenti ci sia la curiosità di capire se sia davvero questo l'obiettivo perseguito o se, si ricerchi una semplice e comoda passività. Ci si chiede se realmente un lavoratore preferirebbe indossare un visore rispetto che premere i link e connettersi alla rete dal proprio dispositivo. Questo concetto si può anche estendere alla messaggistica, alle riunioni di team o alla partecipazione di qualsiasi altro evento fisico. Per fare un paragone, il successo che hanno riscontrato i social media, che vengono tutt'ora utilizzati, è stato grazie alla possibilità data a miliardi di persone di condividere in modo attivo le proprie vite e anche di vedere in modo passivo le vite condivise dagli altri utenti. In questo caso nel Metaverso si avrebbe una maggiore interattività e un elevato livello di engagement<sup>72</sup>, questo potrebbe portare all'allontanamento della maggior parte degli utenti piuttosto che avvicinarli alla tecnologia. Molto spesso la comodità e l'interattività non vanno sempre nella stessa direzione: nel momento in cui uno dei due aspetti trionferà, si potrà affermare quale variabile decisiva porterà all'espansione del Metaverso.

---

<sup>70</sup> Guillaume Pitron, "La guerra dei metalli rari. Il lato oscuro della transizione energetica e digitale";

<sup>71</sup> Interattività: La capacità di un determinato sistema di scambiare informazioni con un operatore umano;

<sup>72</sup> Engagement: misura il grado di coinvolgimento nel compiere una certa azione, come l'attenzione e la partecipazione individuale (like e share).

# Bibliografia

- Arvidsson A., Delfanti A., *“Introduzione ai media digitali”*, Il Mulino, 2013;
- Atzei P., Ceri S., Fraternali P., Paraboschi S., Torlone R., *“Basi di dati”*, McGraw-Hill, 2018;
- Bennato D., Benothman H., Panconesi A., *“La diffusione delle informazioni online - Il caso Twitter”*, AIS (Associazione Italiana di Sociologia), 2010;
- Berners-Lee T., *“Weaving the Web - The original design and ultimate destiny of the World Wide Web”*, Harper Collins, 2000;
- Boldrini N., *“Cos'è l'Intelligenza Artificiale, perché tutti ne parlano e quali sono gli ambiti applicativi”*, Consoft Sistemi, 2019;
- Bonvillian W. B., Van Atta R., Windham P., *“The Darpa model for transformative technologies”*, The journal of technology transfer, 2011;
- Ceron A., Curini L., Iacus S., *“Social media e Sentiment Analysis: l'evoluzione dei fenomeni sociali attraverso la rete”*, Springer Editore, 2014;
- Cosenza V., *“Social media ROI”*, Apogeo, 2012;
- Dionisio J., Burn W., Gilbert R., *“3D Virtual Worlds and the Metaverse”*, ACM Computing Surveys, 2013;
- Guillaume P., *“La guerra dei metalli rari. Il lato oscuro della transizione energetica e digitale”*, Luiss University Press, 2019;
- Hebb, D. O., *“Drives and the CNS in Psychological review”*, McGill University, 1955;
- Heiling M., *“The Cinema of the Future”*, Communication Design, 1992;
- Hutto C., Gilbert E., *“VADER: A Parsimonious Rule-based Model for Sentiment Analysis of Social Media Text”*, AAAI Conference, 2014
- Krueger M. W., *“Artificial Reality II”*, Addison-Wesley Publishing Company, 1991;
- Lee P., Steward D., *“Virtual Reality: a billion-dollar niche, TMT predictions”*, Deloitte, 2016;
- Lombardi M., Giorgi D., Fiesoli I., *“I territori connessi. L'impatto dei nuovi modelli open innovation nelle dinamiche imprenditoriali e territoriali”*, Contributo di ricerca, 2018;
- Lovett J., *“Social Media metrics secrets”*, Wiley, 2011;
- Milgram P., Haruo Takemura, Akira Utsumi, Fumio Kishino, *“Augmented Reality: A class of displays on the reality-virtuality continuum”*, ATR Communication Systems Research Laboratories, 1994;
- Pimentel K., Teixeira K., *“Virtual reality: through the new looking glass”*, McGraw-Hill, 1994;
- Ryan J., *“Storia di Internet e il futuro digitale”*, Piccola biblioteca Einaudi, 2011;
- Smart J., Cascio J., *“A Cross-Industry Public Foresight Project, Metaverse Roadmap”*, Acceleration Studies Foundation, 2007;
- Steuer J., *“Communication in the Age of Virtual Reality”*, Lawrence Erlbaum Associates, 1995;
- Sutherland I., *“A head-mounted three-dimensional display”*, The University of Utah, 1968.



# Ringraziamenti

A conclusione di questo elaborato desidero dedicare questo spazio a tutti coloro che, con dedizione e pazienza, hanno partecipato alla realizzazione del lavoro svolto.

Un sentito ringraziamento al professor Boselli Roberto, relatore della tesi, che in questi mesi di lavoro ha saputo guidarmi con suggerimenti pratici, nelle ricerche e nella stesura dell'elaborato.

Ringrazio di cuore i miei genitori, Michele e Giusy, per il supporto e per avermi permesso di portare a termine gli studi universitari.

Un particolare ringraziamento va a mia sorella Sandra, compagna di avventure, grazie per supportarmi e per rimanere sempre al mio fianco, nonostante i chilometri che ci separano.

Un ringraziamento ai miei zii, Edvige e Davide, che mi hanno aiutata e spronata durante questa esperienza universitaria.

Ringrazio tutti coloro che mi stanno vicino e tutti i colleghi che ho incontrato durante il percorso di studi, grazie per i momenti passati insieme e per quelli che verranno.

Infine, vorrei dedicare questo piccolo traguardo a me stessa, che possa essere l'inizio di una lunga e brillante carriera professionale.