

UNIVERSITÀ DEGLI STUDI DI PADOVA

Dipartimento di Tecnica e Gestione dei Sistemi Industriali

Corso di laurea magistrale in Ingegneria Gestionale

Tesi di Laurea

Comprendere i consumatori nell'era del digitale: netnografia e
etnografia mobile

Relatore
Ch. mo Prof. Stefano Biazzo

Laureando
Michele Menegatti

Anno Accademico 2023-2024

We are drowning in information, but starving for knowledge.

John Naisbitt

Indice

Introduzione.....	1
Capitolo 1 La netnografia	5
1.1. Definizione e origini della netnografia.....	5
1.2. Le comunità online	10
1.3. Metodologia della netnografia.....	15
1.4. Vantaggi e limiti	25
Capitolo 2 Evoluzione della netnografia	29
2.1. Sviluppo recente della netnografia	29
2.2. Intelligenza artificiale	33
2.3. Intelligenza artificiale a supporto della netnografia	46
2.4. Sfide e considerazioni etiche associate all'uso dell'IA	53
Capitolo 3 L'etnografia mobile.....	61
3.1. Definizione e principi chiave.....	61
3.2. Metodologia.....	67
3.3. Punti di forza e limiti della etnografia mobile.....	74
3.4. Scelta del provider per l'applicazione del metodo	81
Capitolo 4 IA e netnografia: sperimentazione per l'analisi dei testi.....	87
4.1. Strumenti disponibili per l'analisi dati	87
4.2. Caso Western Union	96
4.3. Vantaggi e limitazioni tra software e programmazione.....	112
Conclusioni.....	119
Bibliografia.....	127

Lista delle figure

Figura 1.3-1 - Classificazione delle tipologie di membri di una comunità online (Fonte: Kozinets R. V. (2010), Netnography: Doing Ethnographic Research Online).....	18
Figura 2.2-1 - Esempio di architettura della rete neurale profonda (Fonte: www.geeksforgeeks.org)	40
Figura 2.2-2 - Raffigurazione di una rete neurale convoluzionale (Fonte: www.intelligenzaartificialeitalia.net).....	41
Figura 2.2-3 - Esempio di rete CNN combinata con rete RNN (fonte: LeCun et al., (Vol. 521, 2015), Deep learning)	42
Figura 3.1-1 - Numero di vendite mondiali di smartphones dal 2007 al 2022 (Fonte: Statista.com)	63
Figura 3.4-1 - Rappresentazione dell'applicazione Indeemo su smartphone e su PC. Fonte(https://indeemo.com).....	85
Figura 4.2-1 - Screenshot della schermata del software Nvivo raffigurante l'analisi del sentimento.....	99
Figura 4.2-2 - Output analisi tematica con il software NVivo	100
Figura 4.2-3 - Rappresentazione aggiuntiva dell'analisi tematica fornita da NVivo...	101
Figura 4.2-4 - Conteggio e nuvola delle 10 parole più citate ottenute tramite NVivo .	102
Figura 4.2-5 - Istogramma raffigurante la distribuzione del sentimento calcolata tramite l'algoritmo VADER.....	105
Figura 4.2-6 - Grafico del codice Python delle dieci parole maggiormente citate.....	107
Figura 4.2-7 - Istogramma raffigurante il sentimento calcolato tramite la rete BERT delle recensioni raccolte	110

Lista delle tabelle

Tabella 4.2-1 - Temi individuati con LDA.....	106
Tabella 4.2-3 - Temi individuati con BERTopic	111
Tabella 4.3-1 - Confronto risultati delle due analisi con strumenti e metodi differenti	113

Introduzione

C'è un universo in piena evoluzione ed espansione, una realtà che attraversa costantemente il nostro quotidiano, diventata ormai parte integrante: quella virtuale. Tuttavia, soffermarsi ad analizzare nel dettaglio l'impatto di cambiamento, come le opportunità di sviluppo e le offerte esponenziali che si generano in maniera quasi automatica da quest'innegabile risorsa, consente di leggere in filigrana non solo quanto e come siano cambiate le regole nel mondo del business, ma anche di tratteggiare scenari futuri che illuminino e amplino le prospettive per le aziende.

Non si può omettere di considerare, insomma, quanto la Rete rappresenti una risorsa immensa, ma anche una trappola laddove non si sia in grado di "leggerne" correttamente orientamenti e mutazioni.

Questo strumento ha ottenuto una popolarità enorme, e viene considerato ormai un perno della nostra società, definibile "digitalizzata". Lo sviluppo esponenziale del mondo digitale ha permesso un accesso al mondo online di un immenso numero di persone, che conta su questo strumento per svolgere le proprie attività quotidiane, rivelando la sua natura estremamente duttile, che lo rende utilizzabile per un numero di funzioni pressoché infinito: ricerche professionali, personali, progettazioni, discussioni etc., insomma, ha sostituito gran parte delle attività che, in passato, implicavano un confronto umano oppure costosi spostamenti, con un semplice click.

Pertanto, la comprensione e l'approfondimento delle culture degli attori del web da parte delle aziende diventano imprescindibili e cruciali al fine di creare strategie di marketing che stiano al passo con la rivoluzione copernicana dell'internet. Il nuovo ambiente digitale ha di fatto rivoluzionato nel concreto la modalità stessa attraverso cui le persone interagiscono con i prodotti, i servizi e i marchi. La conseguenza innescata è che consumatori, esperti, cittadini, fornitori e ricercatori universitari si trovano così

virtualmente inseriti in un ecosistema digitale comune, una sorta di “agorà” in cui tutti possono interagire con tutti. In questo scenario, in cui dati e connessioni scorrono incessantemente online, si aprono, per le aziende, opportunità di crescita e sviluppo senza precedenti, ma al contempo emergono sfide significative che richiedono soluzioni innovative per mantenersi competitive sul mercato.

Comprendere a fondo i consumatori, i loro desideri, bisogni e persino paure, si pone ora come prerequisito necessario per il successo aziendale in un mondo che corre ad una velocità mai raggiunta.

È dunque possibile ravvisare quanto la competitività e la prosperità delle aziende dipendano in larga misura dalla loro capacità di interpretare e sfruttare al meglio le banche dati sul web, per riuscire ad innovare e aggiornare continuamente la propria offerta. La digitalizzazione ha determinato drastiche modifiche nei cicli di vita dei prodotti: questi si sono accorciati e di conseguenza le aziende hanno aumentato le spese per la ricerca e lo sviluppo. Per garantire che le innovazioni creino un valore effettivo, gli ideatori devono avere un'acuta consapevolezza delle preferenze dei clienti e questo implica uno sforzo, da parte delle aziende, maggiore rispetto a quello impiegato in precedenza. Gli antropologi sono perfettamente preparati ad esaminare questi mutamenti, in quanto i metodi etnografici da loro impiegati sono appositamente sviluppati per fornire dettagliate analisi su aspetti sconosciuti e insoliti.

La digitalizzazione ha altresì creato alcune facilitazioni aprendo nuove strade per l'acquisizione, l'analisi e l'utilizzo di dati sui comportamenti, le preferenze e le abitudini dei consumatori rendendo più semplice e precisa la reperibilità di queste informazioni. In passato, infatti, attraverso una classica ricerca di mercato etnografica, per comprendere le abitudini dei clienti era necessario ricorrere all'immersione del ricercatore all'interno della vita quotidiana del cliente, che impiegava molto tempo alle aziende e che coprivano un campione di acquirenti molto minore, portando così ad una analisi con un tasso di errore probabilmente maggiore.

Oggi sempre più clienti esprimono e scoprono le loro preferenze attraverso il passaparola sui prodotti, che è migrato nei mercati elettronici, creando interazioni attive sui social media che contengono una grande quantità di informazioni che ricercatori e marketer possono utilizzare (Ghose & Ipeirotis, 2011; Kozinets, 1999). Questi dati costituiscono

una miniera d'oro di informazioni che possono essere sfruttate per sviluppare e innovare in sintonia con le esigenze e le aspettative dei clienti trasformando in tal modo gli sforzi delle aziende in storie di successo sul mercato.

Lo scopo che mi sono prefisso con questo studio è dunque quello di identificare, approfondire e analizzare i principali sviluppi nell'uso dell'etnografia da parte degli antropologi, con l'intento di raggiungere una più profonda comprensione dei comportamenti dei consumatori nell'ambito digitale. Si tratta, in particolare, di un viaggio attraverso la netnografia e l'etnografia mobile, ovvero i due metodi fondamentali del campo della etnografia digitale¹.

Per quanto riguarda il tema della ricerca netnografica si farà in gran parte riferimento agli elaborati divulgati dal professore Rober V. Kozinets, l'ideatore di tale metodologia, quali "Netnography, Doing Ethnographic Research Online (2010)", "Netnography Redefined (2015)" e "Netnography Unlimited, Understanding Technoculture Using Qualitative Social Media Research (2020)". Invece, per l'approfondimento dell'etnografia mobile, verranno presi in considerazione articoli e spazi digitali specificatamente dedicati a questo argomento. Nello specifico, i quattro capitoli della tesi si articoleranno sull'esplorazione delle fasi cardine fondamentali per una corretta applicazione e i discussi punti di forza e debolezza di queste due metodologie emergenti. Successivamente l'attenzione verrà concentrata sul tema dell'evoluzione della netnografia, nello specifico all'impiego di strumenti di intelligenza artificiale utili al supporto e amplificazione delle capacità umane nella ricerca. In conclusione, verrà eseguita una sperimentazione pratica su un caso reale che renderà noto, tramite l'applicazione di un software con l'IA integrata e la apposita scrittura di un codice contenenti algoritmi di IA, l'impatto e l'utilità che tale tecnologia può avere nel campo della netnografia, oltre ad offrire una valida panoramica delle soluzioni disponibili a questa tecnologia utili ai ricercatori.

¹ Con etnografia digitale si identifica un'avanzata metodologia di ricerca che si avvale di tecniche etnografiche consolidate, opportunamente adattate per esplorare i comportamenti, le esperienze e le dinamiche culturali degli individui all'interno degli ambienti digitali. In contrasto con l'etnografia tradizionale, la quale richiede la presenza fisica e un'osservazione diretta in contesti specifici, l'etnografia digitale consente di effettuare ricerche in modalità remota. In un'era in cui le interazioni digitali costituiscono un pilastro centrale nella quotidianità, questo metodo di ricerca acquista un'importanza cruciale per la comprensione dei fenomeni sociali moderni, delle tendenze comportamentali dei consumatori e delle evoluzioni culturali in atto (Pink, 2016).

La tesi è stata dunque sviluppata per sostenere e rendere noto quanto l'integrazione della netnografia e dell'etnografia mobile nelle strategie aziendali di ricerca di mercato sia essenziale per le aziende nell'era digitale al fine di indirizzarle verso decisioni più informate e strategie di mercato più efficaci.

L'obiettivo finale, si concentra nella possibilità di evidenziare il percorso più soddisfacente e agile per le aziende interessate che vogliono attingere alla risorsa dell'universo di relazioni attivate online per delineare, incrementare e segmentare l'universo dei propri referenti-clienti, così da ridurre il rischio di un lancio di prodotti che resti poco seguito o che manchi il bersaglio prefissato e migliorare di conseguenza le strategie di marketing.

Capitolo 1

La netnografia

Il primo capitolo di questa tesi si dedica all'approfondimento della netnografia, una metodologia fondamentale per la comprensione dei consumatori nell'era digitale. Questo studio si concentrerà sulla classificazione delle comunità online, sull'illustrazione delle fasi necessarie per il suo corretto impiego, e sulla valutazione dei suoi punti di forza e debolezza. Si mostrerà dunque come la netnografia, radicata nei principi dell'etnografia ma specificamente adattata per l'ambiente virtuale, emerga come uno strumento potente per un'analisi incisiva e autentica del comportamento digitale dei consumatori.

1.1. Definizione e origini della netnografia.

Con il termine “netnografia” si indica il metodo di ricerca sviluppato dall’economista e professore di marketing della York University di Toronto, Robert V. Kozinets².

La parola, coniata nel 1998, è un neologismo che deriva dall’unione dei termini “net” ed “etnografia”. L’etimologia stessa anticipa il contenuto di questa nuova metodologia.

² Robert Kozinets ha conseguito una laurea in amministrazione aziendale presso la York University, Toronto, Canada nel 1987 e un master in amministrazione aziendale nel 1988. Nel 1997, ha conseguito il dottorato in marketing e comportamento organizzativo presso la School of Business della Queen's University, Kingston, Canada. Dal 1997 al 2004 è stato professore assistente presso la Northwestern University in Illinois; Nel 2004/2005 ha insegnato presso l'Università del Wisconsin-Madison. Dal 2005 insegna alla York University di Toronto. Kozinets si occupa, tra le altre cose, della gestione del marchio e la fedeltà al marchio, l'intrattenimento e il marketing tecnologico, lo sviluppo e l'introduzione di nuovi prodotti, nonché le culture del consumo, le sottoculture e l'attivismo dei consumatori. In particolare, insegna e ricerca il comportamento dei consumatori nel contesto delle nuove tecnologie, vari tipi di resistenza dei consumatori e il passaparola. Ha sviluppato il metodo di ricerca della netnografia, un metodo per studiare le comunità online, e, insieme a Russell Belk, il metodo della videografia.

Infatti, essa rappresenta uno sviluppo significativo e un adattamento della classica tecnica qualitativa definita “etnografia” tali da consentire ai ricercatori di esplorare il vasto mondo digitale del web, concentrandosi in particolare sullo studio dei comportamenti dei consumatori e delle dinamiche culturali all'interno delle comunità virtuali, come lascia intendere il prefisso “net”, che deriva da “internet” o “network”.

Sviluppata nel 1995, inizialmente come uno strumento per analizzare le discussioni online tra i fan di “Star Trek”, questa pratica si serve di un approccio antropologico, basato sull'analisi di informazioni pubbliche ottenute attraverso la spontanea condivisione degli utenti su piattaforme web. La definizione esatta della nuova disciplina è fornita direttamente dal suo creatore, Kozinets che nel suo primo scritto “ On Netnography: Initial Reflections on Consumer Research Investigations of Cyberculture” pubblicato nel 1998, afferma che la netnografia è “un nuovo metodo qualitativo ideato appositamente per investigare il comportamento dei consumatori, delle culture e comunità presenti in internet”.

L'avvento di Internet e delle nuove tecnologie, infatti, ha trasformato radicalmente la società, influenzando le modalità di comunicazione, di relazione e di informazione di coloro che ne fanno parte. In questo contesto, la netnografia si è affermata come uno strumento indispensabile per studiare questi fenomeni in evoluzione e ha trovato applicazione in diverse aree comprese nel campo delle scienze sociali.

Mediante l'impiego di metodologie non invasive, la netnografia permette al ricercatore di esplorare le dinamiche sociali online, attraverso una partecipazione diretta dello studioso nell'ambiente virtuale in cui ormai grande parte delle interazioni hanno luogo.

Tale tipologia di analisi si fonda, dunque, sull'utilizzo di “tecniche di osservazione naturalistiche” al fine di penetrare profondamente nei contesti in cui gli utenti generano e ricreano le loro esperienze di consumo. Il suo obiettivo principale consiste nel rivelare i preziosi insight sui consumatori, essenziali per guidare le strategie aziendali di innovazione e di marketing. Gli usi che possono essere fatti di questo strumento sono molteplici. La sua utilità non si limita solo al campo del marketing e delle strategie aziendali, ma si è ampiamente estesa, supportando lo sviluppo di ambiti quali l'educazione, le scienze librarie e dell'informazione, l'ospedalizzazione, il turismo,

le scienze informatiche, la psicologia, la sociologia, l'antropologia, la geografia, gli studi urbani, studi sul tempo libero e sul gioco, la sessualità umana e la ricerca sulle droghe.

Si tratta di un aspetto molto importante: attualmente la netnografia trova grande impiego in altri ambiti, alcuni dei quali sono stati precedentemente elencati. Infatti, è cruciale per capire la mente umana, l'atteggiamento nei confronti delle relazioni sociali e le modalità di interazione con altri soggetti quali le dinamiche dei gruppi virtuali, le comunità digitali e le comunicazioni digitali.

Si è accennato al fatto che, in una delle prime ricerche netnografiche condotte, se non la prima in assoluto, sono state analizzate le complicate trame del fandom legato alla celebre serie televisiva Star Trek, individuando in questo modo importanti fenomeni di consumo dei giorni nostri³.

L'impatto commerciale e culturale di questa serie ci insegna come le immagini, gli oggetti e i testi mediali abbiano una ricaduta fondamentale nelle nostre vite: sebbene siano connessi alla dimensione ludica, assumono il proprio rilievo cruciale nei processi identificativi, imponendosi come vere e proprie religioni civili. Questo, smentendo il famoso detto "non si giudica un libro dalla copertina". È emerso dalla ricerca quanto Star Trek sia presentato come un territorio espanso in cui tanto i fan quanto i semplici spettatori trovano rifugio, una via di fuga dal proprio contesto quotidiano, teletrasportandosi dalla realtà alla finzione.

³ Il fandom di "Star Trek" è noto per essere uno dei più attivi e coinvolti, con una vasta gamma di attività che vanno dalla creazione di fan fiction e fan art alla partecipazione a convention e alla discussione online. L'analisi di questo fandom attraverso la netnografia ha permesso di esplorare diversi aspetti, come:

- Modelli di Consumo: Come i fan interagiscono con il materiale della serie, inclusi gli episodi, i film, i libri e i merchandising.
- Creazione di Comunità: Le modalità con cui i fan si riuniscono per condividere la loro passione, sia online che offline.
- Identità e Appartenenza: Come l'essere fan di "Star Trek" influisce sull'identità personale e sul senso di appartenenza a una comunità più ampia.
- Influenza Culturale: L'impatto che "Star Trek" ha avuto sulla cultura popolare, sui media e persino sulla tecnologia.
- Fan Practice: Le diverse modalità con cui i fan esprimono il loro apprezzamento per la serie, come la scrittura di fan fiction, la creazione di fan art, la partecipazione a giochi di ruolo, ecc.

Dunque, il confine tra mondo virtuale e mondo “concreto” si presenta sempre più labile, con una difficoltà conseguente di attribuire l’aggettivo “reale” all’uno piuttosto che all’altro.

Questa acquisizione è stata possibile solo grazie all’utilizzo della ricerca netnografica e dei suoi strumenti, tra cui l’osservazione diretta delle interazioni online, l’analisi dei contenuti digitali come post, commenti, messaggi di chat e altre forme di comunicazione online, nonché l’uso di interviste e questionari online. Inoltre, i ricercatori spesso sfruttano la registrazione di dati digitali, l’analisi del sentiment, l’analisi di rete e altre tecniche di studio dei dati digitali per ottenere una comprensione approfondita delle dinamiche online.

Per decifrare al meglio l’utilità e gli straordinari benefici pratici che può portare l’uso corretto di uno strumento potente come la netnografia, è necessario partire dalle origini, dalla sua “genitrice” da cui prende anche una parte del nome, ovvero l’etnografia (dal greco: ethnos (ἔθνος) - "popolo", e grapho (γράφω) - "scrivo"; letteralmente "descrizione del popolo").

Come afferma Dick Hobbs⁴ l’etnografia può essere definita come "un cocktail di metodologie che condividono l’assunzione che, l’interazione personale con il soggetto è la chiave per comprendere una cultura o un contesto sociale specifico. L’osservazione partecipante è la componente più comune di questo cocktail, ma interviste, analisi conversazionali e del discorso, analisi documentaria, film e fotografia hanno tutti il loro posto nel repertorio dell’etnografo. La descrizione risiede al centro dell’etnografia, e in qualsiasi modo questa descrizione sia costruita, si cerca il significato intenso della vita sociale dal punto di vista quotidiano dei membri del gruppo" (Hobbs, 2006).

Il termine compare la prima volta nel 1767 in un libro dello storico tedesco Johann Friedrich Schöpperlin (1732-1772).

Tale metodo, è una tecnica di ricerca qualitativa comunemente usata nell’antropologia che si basa sull’osservazione diretta delle persone nei loro contesti di vita reale. Il metodo

⁴ Netnografo urbano specializzato in sociologia di Londra, criminalità organizzata e professionale, economia notturna, violenza, mercati della droga e metodologia di ricerca. Attualmente professore emerito all’Essex, professore di sociologia all’Università di Western Sydney e membro associato del Royal United Services Institute.

etnografico consiste principalmente in ricerche sul campo, gli etnologi infatti sono spesso immersi nelle comunità o nelle culture che studiano.

Questo approccio metodologico può essere fatto risalire all'epoca delle esplorazioni coloniali, durante la quale numerosi viaggiatori ed esploratori europei annotavano osservazioni meticolose sulle culture delle popolazioni con cui venivano in contatto nelle nuove terre appena scoperte, anche se qualcuno suppone una genesi ancora anteriore, risalente al 1733 quando il tedesco Gerhard Friedrich Müller⁵ ha elaborato la sua descrizione e categorizzazione di abbigliamento, religioni e rituali dei gruppi etnici siberiani.

Tali relazioni e registrazioni rappresentarono una fase primordiale nello sviluppo dell'etnografia, che, come ogni scienza, poi si è evoluta con i mutamenti culturali ed economici della società, aggiornando il proprio *modus operandi* per stare al passo con i tempi.

L'etnografia consente di raggiungere una comprensione profonda di una cultura o di un gruppo sociale basandosi su dati che emergono direttamente durante lo studio e non derivati a priori dalle ipotesi del ricercatore, quindi attraverso uno studio empirico.

La particolare precisione e utilità di questa tecnica deriva dal procedimento di immersione prolungata da parte del ricercatore in un contesto specifico: in questo modo, egli attinge a specifiche informazioni della vita sociale a cui risulterebbe difficile accedervi in altro modo. Infatti, il ricercatore, entrando a pieno nella vita del cliente risulta poco intrusivo nel contesto, ottenendo dunque informazioni personali veritiere e non forzate.

Tuttavia, è importante specificare che vi è un utilizzo piuttosto limitato di questa tecnica a causa della difficoltà di organizzazione e dell'elevato dispendio di tempo e risorse.

⁵ Geografo, ricercatore ed esploratore tedesco. È definito "il padre della storiografia siberiana".

1.2. Le comunità online

L'evoluzione esponenziale di internet e delle tecnologie digitali hanno portato a cambiamenti radicali nei comportamenti dei consumatori.

In particolare, lo sviluppo del WEB 2.0 ha permesso l'introduzione della comunicazione mediata da computer (CMC) trasformando completamente l'interazione online, consentendo agli utenti di partecipare attivamente alla creazione e alla condivisione di contenuti.

La comunicazione mediata da computer (CMC) rappresenta il processo mediante il quale le persone interagiscono e condividono informazioni utilizzando dispositivi informatici e tecnologie di comunicazione. Tale evoluzione ha consentito la nascita di forum, blog, podcast, chat e social network mettendo così in relazione gruppi di persone con interessi comuni. Questi incontri online, tra gruppi di consumatori, che si verificano sono stati designati come comunità online.

Oggetto principale per l'applicazione della netnografia, le comunità online, possono essere definite attraverso la precisa definizione di Howard Rheingold⁶, il quale le definisce come "aggregazioni sociali che emergono dalla rete quando abbastanza persone conducono discussioni pubbliche abbastanza a lungo, con sufficiente sentimento umano, per formare reti di relazioni personali nello spazio cibernetico" (Rheingold, 1993).

Analizzando in maniera più approfondita la definizione, è possibile sviscerare in maniera più completa il concetto di netnografia.

Nello specifico:

- Il termine "Aggregazioni sociali" fa riferimento al fatto che la netnografia è un approccio collettivo che esamina gruppi e raccolte di argomenti creati da persone e non che esamina pubblicazioni di singoli individui.

⁶ Howard Rheingold (Phoenix, 7 luglio 1947) è un critico letterario, sociologo e saggista statunitense, specializzato sulle implicazioni culturali, sociali e politiche dei nuovi media.

- "Emergere dalla rete": Come suggerisce il nome, la netnografia esamina le interazioni individuali che risultano dalle connessioni Internet o attraverso comunicazioni mediate dal computer come fonte focale di dati.

- "Discussioni" o comunicazioni costituiscono l'elemento fondamentale della netnografia. Attraverso le discussioni vi è uno scambio continuo di dati e informazioni tra gli utenti che costituiscono la base per l'uso della tecnica netnografica.

- "Un numero sufficiente di persone" implica che un gruppo, per essere definito comunità online, deve coinvolgere un numero minimo di utenti attivi che costantemente pubblicano informazioni. Si tratta, come afferma Robin Dunbar, di un numero di circa 20 persone come limite inferiore.

- L'espressione "Discussioni pubbliche" indica l'importanza e la necessità che le comunità online siano accessibili per la loro formazione e affinché i dati siano disponibili ai netnografi per le loro analisi.

- "A sufficienza a lungo" indica che la comunità per essere tale deve risultare attiva per un tempo sufficientemente lungo sia per creare un senso di comunità tra gli utenti sia per consentire uno studio netnografico che permetta contatti ripetuti e non occasionali.

- "Sentimento umano sufficiente" riguarda l'importanza degli utenti di esprimersi in modo autentico per generare fiducia e onestà tra i partecipanti della comunità.

- "Per formare reti di relazioni personali" indica l'esistenza di un legame sociale autentico tra i singoli componenti del gruppo, oltre a favorire la formazione di un forte senso d'appartenenza che va oltre l'idea di un'entità collettiva composta da tali relazioni. Tali connessioni, con frequenza, si estendono ben al di là dell'ambito virtuale, influenzando positivamente diversi aspetti della vita sociale degli individui.

In un contesto sempre più digitalizzato, è emerso così un ricco e diversificato panorama di comunità online.

Ciò che rende questi "habitat" online così speciali è la loro capacità di riunire persone provenienti da diversi luoghi geografici, che altrimenti non avrebbero modo di incontrarsi fisicamente. Le piattaforme web esistenti consentono agli utenti di accedere a una vasta gamma di comunità, ciascuna con un obiettivo specifico. Che si tratti di forum di

discussione, blog, social network o piattaforme di apprendimento online, esiste una comunità online per quasi qualsiasi interesse o passione.

È quindi importante definire l'esistenza di diverse tipologie di comunità che si possono trovare sul web. Si procede poi a suddividerle in base all'obiettivo e al tipo di transazione richiesta.

Seguendo questi criteri si ottengono i seguenti tipi di comunità online⁷:

- Comunità di discussione
- Comunità orientate a compiti e obiettivi
- Mondi virtuali
- Comunità ibride

Le comunità di discussione si basano sullo scambio approfondito di informazioni concernente un argomento ben definito. L'obiettivo, dunque, di queste community è quello di creare uno spazio dedicato ad argomenti specifici dove i contenuti creati dai diversi utenti e le discussioni formatesi sono strutturati e facili da trovare. Inoltre, l'intento è quello di creare contenuti con un linguaggio semplice e generico e di conservare i dati più a lungo possibile in modo creare una sorta di manuale per gli utenti futuri.

In aggiunta è possibile distinguere delle sottocategorie alle comunità di discussione:

- Comunità di discussione con comunicazione diretta tra persone. Esempi di tali comunità sono le comunità di relazione così come le alleanze sociali e di aiuto.
- Comunità orientate all'argomento, che sono legate a determinati argomenti definiti. Esempi di tali comunità sono comunità sportive, comunità di prodotto, comunità ecologiche, ecc.
- Comunità di pratica, che emergono nelle organizzazioni intorno a determinati argomenti e know-how.

⁷ Classificazione elaborata da Katarina Stanoevska-Slabeva, professoressa di gestione dell'informazione e della comunicazione, Università di St. Gallen, Switzerland.

- Comunità di discussione indiretta con comunicazione indiretta tra i membri. Un esempio di tale comunità di discussione è la comunità delle recensioni di Amazon.

Tra queste sottocategorie di comunità di discussione un elemento differenziale caratterizzante è la volontà di rimanere anonimi durante l'interazione. Infatti, nelle comunità con comunicazione diretta tra le persone spesso si verifica il desiderio di utilizzare pseudonimi per la comunicazione mentre per le altre tipologie, la conoscenza dei partecipanti risulta necessaria contribuendo anche ad una maggiore fiducia.

Un ulteriore elemento differenziante, utile da specificare, è il canale di comunicazione che viene utilizzato.

Le discussioni tra persone richiedono essenzialmente un canale di comunicazione sincrono, mentre le comunità di pratica e le comunità orientate all'argomento richiedono di solito entrambi i tipi di canali di comunicazione: sincrono e asincrono.

Nelle comunità orientate a compiti e obiettivi gli utenti vi accedono per ottenere un obiettivo comune. Anche all'interno di queste comunità, vi sono suddivisioni in sottocategorie nella quale tutte condividono la stessa esigenza di accedere alle informazioni dei membri. Per ogni sottocategoria vi sono regole comportamentali, ruoli, processi e soluzioni differenti.

- **Comunità di transazione:** Luoghi di incontro in cui acquirenti e venditori si incontrano per determinare i prezzi di beni e servizi e scambiare informazioni e merci. Qui l'obiettivo principale è l'abbinamento dell'offerta e della domanda. La sensazione di comunità viene creata grazie alla comunicazione tra clienti nella fase di transazione. Un esempio di comunità di transazione sono i mercati online come eBay o Amazon.
- **Comunità di progettazione:** Qui i membri possiedono conoscenze e competenza complementari. Si tratta di comunità nelle quali l'obiettivo è la progettazione e lo sviluppo comuni di un prodotto. Le comunità open source fanno parte di questa sottocategoria.
- **Comunità di apprendimento online:** Piattaforme virtuali dove studenti e persone entrano e interagiscono per ottenere una maggiore comprensione su un argomento specifico. L'obiettivo principale è quello quindi, di creare uno spazio di

apprendimento comune. All'interno di questi spazi è necessaria però un'elevata passione personale per ottenere risultati significativi. Un esempio di comunità di apprendimento sono i forum dove più studenti si incontrano per discutere di argomenti con l'obiettivo di sostenersi reciprocamente durante la preparazione di un esame.

I mondi virtuali sono una tipologia di comunità online nella quale le informazioni degli utenti vengono comunicate mediante l'uso di avatar. Gli utenti interagiscono in una piattaforma tridimensionale sia essa un gioco online oppure uno spazio digitale creato appositamente.

Le comunità ibride sono formate da diverse comunità connesse tra loro, con membri che possono partecipare a più di una comunità. In questo ambiente vi è una combinazione di diverse comunità, come comunità di discussione, comunità di transazione e mondi virtuali, che ruotano attorno a un argomento comune.

Un esempio di comunità ibrida potrebbe essere un negozio online che offre un'esperienza tridimensionale simile a quella di un negozio fisico, con venditori virtuali che agiscono come assistenti di vendita. Gli utenti possono essere semplici visitatori, acquirenti o partecipanti a giochi. Possono ottenere informazioni sui prodotti attraverso un catalogo elettronico, chat, forum di prodotto e raccomandazioni. Possono ordinare prodotti e pagare online o tramite una bolletta mensile e persino creare liste della spesa per acquisti futuri.

Uno dei fattori chiave di successo delle comunità ibride è riuscire ad ottenere un design che permetta un'integrazione dei diversi tipi di comunità e delle tecnologie per il loro supporto. Nelle comunità ibride la conoscenza creata all'interno di una comunità deve essere condivisa tra le diverse comunità per massimizzare il valore delle informazioni interconnesse.

Questa classificazione fa riflettere su come i rapporti tra le persone si siano evoluti negli ultimi anni. È possibile, infatti, affermare che ora gran parte dei collegamenti/discussioni tra persone, anche vicine tra loro geograficamente, avviene tramite dispositivi digitali connessi al web e molto spesso partecipando a comunità online come parte regolare della loro attività sociale. In questo contesto, agli antropologi, volti ad analizzare il

comportamento delle persone, risulta imprescindibile osservare le modalità d'uso di internet e quindi di tecniche netnografiche.

Tuttavia, appare esponenzialmente più complicato definire una linea di demarcazione tra vita sociale reale e virtuale. Così, per esempio, una semplice ricerca etnografica, che abbia ad oggetto un gruppo di professionisti medici potrebbe rivelarsi limitata nei risultati se non si prendesse in considerazione anche l'analisi anche di e-mail, forum e chat varie.

Tutto, dunque, si sta fondendo in un unico mondo che prevede l'uso di strumenti tecnologici per comunicare, socializzare ed esprimersi.

1.3. Metodologia della netnografia

Per condurre una ricerca netnografica adeguata, il professor R. Kozinets (2010) suggerisce che, per i ricercatori, è essenziale seguire i sei step principali di una classica ricerca etnografica e adattarli al mondo digitale in analisi.

Si tratta di:

- Pianificazione della ricerca
- Entrée
- Raccolta dati
- Analisi dati e interpretazione
- Garantire gli standard etici
- Presentazione dei risultati

Di seguito una descrizione di ogni step.

Pianificazione della ricerca

La fase di pianificazione è considerata fondamentale al fine di ottenere un ingresso adeguato nella comunità e di assicurare una strutturazione ben definita della ricerca. A fronte di tale importanza è necessario, per il ricercatore o per il team in questione, andare a definire con precisione gli obiettivi della ricerca e gli argomenti chiave dello studio per evitare la navigazione a vuoto nella vastità del web. Nella pianificazione si passa poi all'uso dei motori di ricerca per individuare, attraverso l'inserimento di parole chiave dell'argomento di ricerca, le possibili comunità online.

Entrée

Una volta identificato l'obiettivo e le possibili comunità di analisi si passa all'identificazione del luogo di analisi della ricerca. Per individuare le community potenzialmente idonee alla specifica indagine, Kozinets ci suggerisce alcune linee guida. In particolare, consiglia di analizzare comunità online che sono:

- *Rilevanti*, cioè correlate al focus di ricerca
- *Attive*, cioè con comunicazioni recenti e regolari
- *Interattive*, cioè con un flusso di comunicazioni tra i partecipanti
- *Sostanziali*, cioè con una massa critica di comunicatori e un'atmosfera energica
- *Eterogenee*, cioè con diversi partecipanti
- *Ricche di dati*, offrendo dati dettagliati o ricchi di descrizioni

Le diverse comunità online attive nello spazio digitale, tra cui forum, blog, chat-room, social network e wiki, si rivelano sovente sospettose verso gli estranei, in particolare verso chi potrebbe essere percepito come un "intruso". Per il ricercatore, è dunque essenziale comprendere il contesto e le origini della comunità che intende studiare. Questo gli permette di comunicare in modo appropriato ed evitare di accedere a gruppi non adatti. È necessario inoltre evitare siti recentemente "sfruttati" da altri ricercatori o siti che sono stati "spenti" da ricerche scadenti in passato. Infatti, i membri potrebbero essere già esausti a causa degli sforzi di ricerca recenti. L'ingresso in una community inoltre non deve raffigurare come spam. È quindi necessario evitare intrusioni ripetute e non attente

alle norme. Inoltre, è molto importante per il ricercatore superare gli standard etici e ottenere una familiarizzazione⁸ con le comunità prese in considerazione per l'analisi.

Si rivela inoltre molto utile, per il ricercatore, riuscire a identificare sia piattaforme online dominate da clienti nelle quali sono i clienti a generare il contenuto sia ulteriori piattaforme nelle quali sono le aziende a generare il materiale di analisi con lo scopo di ottenere una visione più completa del problema e risultati più accurati.

Raccolta dati

Una volta identificata la community ideale da analizzare si procede alla fase successiva di raccolta dei dati. Nella netnografia la raccolta dati avviene grazie alla comunicazione con e tra i membri di una comunità e i dati generalmente si trovano all'interno delle conversazioni. Tale comunicazione è l'elemento fondamentale che differenzia i metodi della etnografia e netnografia dalle "semplici" tecniche di analisi del contenuto o analisi delle reti sociali.

Per la raccolta dati, il ricercatore può scegliere se adottare un approccio attivo, interagendo direttamente con la comunità e rendendo nota la sua attività di ricerca o passivo senza informare la comunità della ricerca in atto e leggendo solamente i messaggi scambiati tra gli utenti senza interazione. Tuttavia, anche se, l'approccio passivo, potrebbe sembrare limitante nella qualità dei dati raccolti, è il botta e risposta in una discussione che si crea tra le persone ad aggiungere valore e trasmettere informazioni rilevanti. Inoltre, molto spesso, come afferma anche Kozinets i ricercatori adottano un approccio passivo per la sua semplicità e convenienza e per ottenere una ricerca meno invasiva seguita da risultati meno influenzati e più autentici.

L'output del processo di raccolta dati prevede tre tipologie di informazioni:

- *I dati di archivio*: dati osservazionali abbondanti e facili da ottenere. Sono dati ottenuti dalle trascrizioni delle comunicazioni avvenute dai membri delle comunità online. Il ricercatore, dunque, non interagisce con gli utenti ma si occupa

⁸ Processo che include la comprensione dei membri della comunità, il loro linguaggio specifico e i loro interessi. Questo approccio garantisce non solo la validità e l'affidabilità della ricerca, ma contribuisce anche a costruire un rapporto di fiducia e rispetto reciproco tra il ricercatore e la comunità.

solamente del download dei testi. Si tratta di dati non influenzati dalle azioni del netnografo.

- *I dati indotti*: dati che il ricercatore ottiene grazie ad una partecipazione attiva nella comunità attraverso interazioni personali e comuni.
- *Le annotazioni dei ricercatori*: Note riflessive che il ricercatore si trascrive durante l'osservazione della comunità. Esse riguardano sensazioni, opinioni, significati e idee che emergono dall'indagine.

Durante il processo di raccolta dei dati, i dati riflessivi sono solitamente riservati per l'uso del ricercatore e non sono condivisi con la comunità.

Va sottolineato che gli utenti tendono spesso a pubblicare moltissimi messaggi e informazioni di poco rilievo o con presenza di errori. A fronte di tale problema è possibile categorizzare tali messaggi, in particolare categorizzare gli autori dei messaggi. Kozinets, prima nel suo articolo "E-tribalized marketing?: the strategic implications of virtual communities of consumption" (1999) e successivamente nel libro pubblicato nel 2010, identifica quattro tipologie di membri all'interno di una comunità. Essi sono stati classificati in base al grado di coinvolgimento sul tema trattato e al legame sociale nella comunità.

Di seguito l'immagine che mostra tale suddivisione.

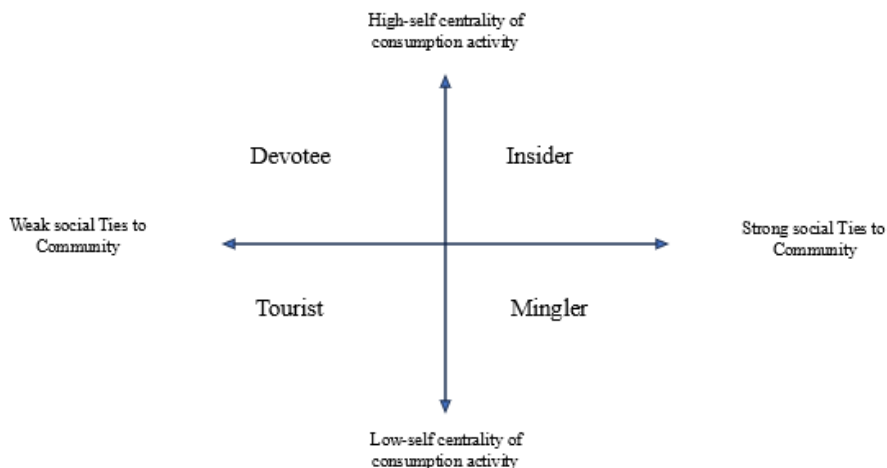


Figura 1.3-1 - Classificazione delle tipologie di membri di una comunità online (Fonte: Kozinets R. V. (2010), Netnography: Doing Ethnographic Research Online)

Vi sono dunque:

- I Tourist, che mancano di forti legami sociali con il gruppo e mantengono solo un interesse superficiale nell'attività di consumo in questione e inoltre possiedono abilità e competenze relativamente deboli.
- I Mingler, che mantengono forti legami personali con molti membri della comunità ma sono interessati o attratti solo superficialmente dall'attività di consumo centrale.
- I devotee, che invece hanno legami sociali relativamente superficiali con i membri della comunità, ma mantengono un interesse focale ed entusiasmo per l'attività di consumo nella comunità, così come competenze e conoscenze raffinate.
- Gli Insider, che sono coloro che hanno forti legami sociali con la comunità online e presentano un'attitudine e comprensione profonda dell'attività di consumo centrale.

Da questa semplice ma chiara classificazione è facile comprendere come i dati di analisi utili ai ricercatori vengano forniti principalmente dalle categorie di Devotee e Insider ovvero membri con interesse nel tema di analisi.

Per la raccolta dati il netnografo può optare tra due metodologie:

- 1) Metodo manuale: Con questa tecnica, la raccolta dati avviene manualmente tramite carta e penna oppure usando fogli di calcolo excel. Questa metodologia limita molto la ricerca netnografica in quanto solo limitate quantità di dati possono essere analizzate al fine di evitare tempistiche troppo elevate. Potrebbe anche alterare il focus della netnografia, che potrebbe essere più adatto ad analizzare solo messaggi o discussioni specifiche all'interno della comunità.
- 2) Metodo con software: In questo caso la raccolta dati può essere molto più prolifica estendendo così i confini della comunità online. Con il software si è in grado di raccogliere molte informazioni e di trarne un vantaggio per scovare sentimenti e opinioni nascoste. Il problema principale di tale metodo è la necessità effettuare un'azione di pulizia di dati estranei. Infatti, il software tende a raccogliere qualsiasi tipologia di dati, anche i non pertinenti alla ricerca netnografica. Dunque,

se si opta per questa metodologia il netnografo dovrà essere preparato a dedicare tempo per la pulizia dati.

Analisi dati e interpretazione

Per l'analisi e l'interpretazione dei dati i ricercatori si affidano a diverse tecniche. Lo studio effettuato da Kristina Heinonen e Gustav Medberg (2018) ha messo in luce che l'approccio di analisi più comune per gli studi netnografici è l'analisi tematica.

L'analisi tematica è un metodo qualitativo che prevede l'identificazione e l'analisi di temi emergenti che si trovano all'interno dei dati. L'obiettivo, dunque, è quello di evidenziare i significati più importanti presenti nei testi raccolti e analizzati.

Un altro strumento utilizzato spesso dai netnografi è l'analisi qualitativa del contenuto e l'analisi del discorso.

Con l'analisi qualitativa dei contenuti, i netnografi cercano di capire quali informazioni vengono trasmesse, mentre con l'analisi del discorso si cerca di individuare il significato culturale alla base della comunicazione tra i clienti.

Per l'analisi dei dati i ricercatori hanno la possibilità di ricorrere all'uso di un software di analisi qualitativa in grado di semplificare di molto il processo. Pur riconoscendo l'efficacia e il valore di questo strumento, va precisato che esso non possiede la stessa profondità interpretativa di un analista umano, il quale coglie le delicate sfumature di pensiero, sentimento ed emozione che possono sfuggire ad un processo di analisi privo di consapevolezza.

Come affermano Boellstorff et al. (2013) se un ricercatore fosse attivamente impegnato nella comunità attraverso l'osservazione partecipante, l'esame degli elementi tematici dovrebbe essere effettuato manualmente per ottenere maggiore accuratezza.

Invece, il software si rivela particolarmente efficace e vantaggioso quando si tratta di elaborare vasti volumi di dati e si adatta bene alla situazione in cui il ricercatore si dedica all'analisi testuale senza un coinvolgimento diretto all'interno della comunità studiata.

Se è vero che la netnografia è stata affrontata come metodo qualitativo basato su un approccio induttivo nel quale teorie e concetti emergono direttamente dai dati, non va tralasciato che con lo sviluppo crescente del mondo digitale e l'enorme quantità di dati

presenti nelle comunità online è possibile combinare la ricerca qualitativa con quella quantitativa per analizzare i dati. Si tratta dunque di fare riferimento a un metodo misto.

Mentre l'analisi qualitativa ha il potere di analizzare il contesto dei dati, le motivazioni, percezioni e opinioni dei consumatori, e non solo il contenuto in sé dei dati, l'analisi quantitativa può riguardare la misurazione di metriche come la frequenza di determinate parole o frasi, il numero di volte in cui un argomento viene discusso, le analisi sentiment per valutare toni positivi o negativi, ecc. La decisione di utilizzare solo l'analisi qualitativa, solo quella quantitativa o una combinazione di entrambe dipende dagli obiettivi della ricerca, dalla natura dei dati raccolti e dalle competenze del ricercatore.

Kozinets, inoltre, ha suggerito che per generalizzare i risultati netnografici e aumentare la qualità dei dati della ricerca è possibile combinare metodi tradizionali offline, per esempio interviste, focus group, etnografia tradizionale, con la netnografia e ottenere così una triangolazione. Questa tecnica applicata a un mondo sempre più digitalizzato in cui i confini tra esperienze offline e online sono sempre più sfumati potrebbe portare ad approfondimenti e risultati molto più profondi e precisi di quanto un singolo metodo produrrebbe. È infatti importante notare che i risultati di una ricerca qualitativa non sono generalizzabili in quanto le condizioni in cui la ricerca viene condotta non possono essere replicate.

Una volta analizzati i dati il ricercatore ha il compito di immergersi nel contesto e di interpretare al meglio tali dati per generare insight importanti sui consumatori. L'interpretazione rappresenta una comprensione sintetica, olistica e illuminante del significato, come nel decifrare un codice. In questo processo, è cruciale, per il ricercatore, mantenere flessibilità e sensibilità culturale, assicurandosi che le interpretazioni siano pertinenti e rispettose.

La comprensione approfondita del consumatore consente di generare e raffinare idee significative e soluzioni che portano allo sviluppo di prodotti e servizi innovativi, nonché a campagne di marketing mirate. Questo processo non solo porta a miglioramenti incrementali nelle prestazioni aziendali, ma contribuisce anche a un aumento del benessere nella società.

Garantire gli standard etici

Il mantenimento degli standard etici durante una ricerca netnografia appare determinante per la protezione della privacy dei soggetti in studio ed è tutt'ora l'area più incerta che i ricercatori si trovano ad affrontare prima di iniziare una ricerca.

L'etica non è solo una condizione che la netnografia deve rispettare, ma un atout che coinvolge tutto il procedimento della ricerca. Essa influisce sulla formulazione della domanda di ricerca, sulla scelta delle comunità da studiare, sui metodi impiegati, sul tipo di dati raccolti, sulle domande poste ai membri della comunità, sul tipo di analisi da condurre, e sulla natura del rapporto finale.

È possibile affermare che un principio fondamentale di una netnografia etica è la presenza dell'onestà tra il ricercatore e i membri della comunità online. Infatti, senza il rispetto dei principi etici si potrebbe facilmente danneggiare l'ambiente o la cultura di una comunità online.

Uno dei dilemmi alla base degli standard etici nella netnografia riguarda la divulgazione o meno ai membri di una comunità della presenza di un ricercatore. Diversamente dalla etnografia tradizionale nel quale il ricercatore deve palesarsi come tale in quanto svolge la sua ricerca in presenza fisica, nelle comunità online la presenza del ricercatore potrebbe anche essere facilmente mascherata.

A fronte di tale problematica Kozinets (2002) ha pubblicato alcune linee guida etiche che dovrebbero sempre essere rispettate nel momento in cui si intraprende una netnografia. Nello specifico esse sono:

- Rivelare pienamente la presenza e le intenzioni del ricercatore all'interno della comunità;
- Garantire la confidenzialità e l'anonimato degli informatori;
- Cercare e incorporare il feedback dei membri della comunità online;
- Chiedere il permesso ai membri della comunità prima di citare direttamente qualsiasi post specifico nella ricerca;

Seguendo queste linee guida il ricercatore sarà in grado di garantire gli standard etici evitando così problematiche sgradevoli con le comunità.

Molte, però, sono le idee contrastanti sul fenomeno di dichiarare o meno la presenza del ricercatore. Langer e Beckman (2005) sostengono che rendere nota la presenza del ricercatore potrebbe influenzare negativamente la partecipazione e alterare la natura organica della conversazione.

Altri, tra cui Goulding (2013), ritengono che il ricercatore dovrebbe valutare il livello di rischio per i partecipanti e decidere di conseguenza la necessità del consenso. Se il rischio è basso e gli argomenti non sono sensibili o personali, allora le informazioni disponibili pubblicamente potrebbero essere utilizzate liberamente. Vi è invece più chiarezza e unione, da parte dei ricercatori sopra elencati, sul fatto di proteggere l'anonimato dei partecipanti nascondendo identificatori evidenti come citazioni e nomi.

Un altro fattore chiave da rispettare è l'ottenimento del consenso informato da parte degli utenti di una comunità. Anche se le informazioni sono pubblicamente disponibili, gli standard etici nella ricerca su soggetti umani richiedono di ottenere il permesso dei partecipanti alla ricerca quando è appropriato. È quindi necessario determinare in quali contesti è appropriato o meno richiedere il permesso. Kozinets sostiene che, se un ricercatore partecipa attivamente in una comunità online interagendo come gli altri membri tramite commenti e partecipazione, mentre annota osservazioni su sensazioni, comportamenti e altri aspetti, potrebbe non essere indispensabile ottenere il consenso individuale da ogni partecipante. Questo perché l'interazione del ricercatore non si discosta da quella di un membro della comunità.

Invece quando il ricercatore indaga su comunicazioni sincrone e in tempo reale come chat e mondi virtuali non dovrebbe mai registrare le interazioni senza un permesso esplicito da parte del membro. Questo deriva proprio dalla natura effimera di queste comunicazioni nelle quali l'utente scambia informazioni e principi che non intende mantenere attive e disponibili nell'ambiente digitale in un tempo futuro.

È possibile affermare che navigare tra gli standard etici in netnografia comporta un equilibrio attento tra l'integrità della ricerca e il rispetto dei partecipanti. Sebbene alcune linee guida siano definite, altre sono suscettibili di interpretazioni diverse e necessitano di una ponderata riflessione etica da parte del ricercatore.

Confronto con i membri e presentazione dei risultati

Come descrive Kozinets nel suo elaborato del 2010, una volta che il ricercatore ha concluso il suo lavoro netnografico e identificato i principali insight sui consumatori, dovrebbe presentare i risultati della ricerca ai membri delle comunità analizzate con l'obiettivo di effettuare un controllo finale. Questo risulta essere un processo fondamentale in quanto i membri hanno la possibilità di validare le informazioni raccolte assicurandosi che le interpretazioni eseguite dai ricercatori siano adeguate. Durante il controllo, i membri, possono correggere eventuali errori e rivelare nuove prospettive sfuggite nell'analisi dei ricercatori, aumentando così l'affidabilità e la credibilità della ricerca. Vi è così la creazione di un flusso di informazioni bidirezionale tra ricercatore e la comunità in grado di arricchire la qualità della ricerca e di risolvere problemi legati alla privacy e alle norme etiche.

Con tale scambio di informazioni, infatti, il ricercatore mette in chiaro l'intento della ricerca, i metodi e l'entità delle informazioni utilizzati con gli utenti in modo tale da ottenere il consenso della privacy e dando l'opportunità ai membri di modificare/rimuovere la presenza di dati sensibili non graditi.

Una volta concluso il controllo con i membri i ricercatori possono procedere con la presentazione dei risultati. Spesso tale divulgazione viene eseguita tramite lo storytelling, una sorta di narrazione volta a descrivere il viaggio che il ricercatore ha effettuato nel mondo dei consumatori raccontando i principali insight ottenuti entrando in profonda empatia con essi.

Nello sviluppo della narrazione, Kozinets (2020) suggerisce l'inserimento di contenuti generati dagli utenti stessi come foto e video che mostrino le loro abitudini e routine in modo da trasmettere emozioni e problemi dei consumatori in modo autentico. È di fondamentale importanza far sentire i consumatori con la loro voce, la loro scelta di parole e con tutti i loro sentimenti che provano su un argomento specifico. Questo metodo permette facilmente ai clienti di sentirsi parte del mondo dei consumatori ed entrare velocemente in empatia con essi comprendendo al meglio i problemi presenti su un argomento specifico. Il tutto con l'obiettivo di sviluppare soluzioni innovative di prodotti e servizi.

1.4. Vantaggi e limiti

Adatta al nuovo mondo digitale, la netnografia, è in grado di offrire molti benefici sia economici che applicativi rispetto ad altri metodi di ricerca.

In primo luogo, con l'utilizzo di questa metodologia i ricercatori possono avere accesso e interagire ad un ampio panorama di comunità digitali presenti in tutto il mondo. La netnografia è in grado di sconfinare le barriere geografiche consentendo così di espandere l'area di ricerca. Inoltre, non avendo un contatto faccia a faccia con le comunità di analisi, il bisogno degli spostamenti fisici per i ricercatori svanisce rendendo questa tecnica economica e adatta a progetti con budget limitato. Questo garantisce anche una maggiore continuità della ricerca in quanto è possibile mantenere il contatto con i membri della comunità in qualsiasi luogo e quindi anche in presenza di spostamenti da parte dei ricercatori. Elevato è anche il risparmio di tempo che si verifica per la raccolta e l'analisi dei dati provenienti dalle fonti sul web. Inoltre, le interazioni e conversazioni possono essere monitorate in tempo reale.

Punto focale della netnografia è la capacità di ottenere dati ricchi e naturalistici. I dati sono raccolti direttamente dalle conversazioni naturali e casuali che le persone hanno nel loro ambiente digitale abituale. La ricerca è non invasiva e gli utenti, sentendosi non monitorati, sono più inclini a esprimersi liberamente offrendo uno sguardo spontaneo e autentico su ciò che pensano veramente e su come vivono la loro vita quotidiana. In questo modo i ricercatori sono in grado di captare meglio i comportamenti, preferenze e tendenze dei consumatori consentendo alle aziende di affinare le campagne di marketing e i progetti di sviluppo prodotto per avvicinarsi il più possibile alle volontà dei clienti target. In aggiunta i ricercatori con tutti questi dati raccolti possono sviluppare importanti conclusioni sulla concorrenza presente. Inoltre, grazie agli archivi digitali che permettono la permanenza nel tempo dei contenuti online è possibile effettuare una analisi in profondità osservando l'evoluzione delle percezioni e delle tendenze offrendo spunti preziosi per la comprensione della trasformazione culturale.

La netnografia è quindi in grado di fornire una comprensione profonda degli utenti online. Infatti, immergendosi nel contesto, non ci si limita solo a capire il "cosa" avviene, ovvero

i semplici dati pubblicati, ma vengono prese in considerazione tutte le 5W: What, When, Who, Why e Where. Si riesce così ad andare oltre al primo livello di conoscenza e ad entrare in empatia con il consumatore.

Nonostante i numerosi vantaggi, tale metodo presenta comunque delle criticità da non sottovalutare. In primo luogo, essendo la netnografia un metodo di ricerca basata sull'utilizzo di internet, si ha il grosso problema della quantità e dell'autenticità dei dati. Internet è uno spazio aperto a tutti coloro che dispongono di una connessione alla rete e molti sono gli utenti che ogni giorno pubblicano qualsiasi tipo di informazione. Il compito principale e complesso del ricercatore è quello di riuscire a identificare utenti con identità vere e non distorte. Molti, infatti, potrebbero essere bot ingannevoli che portano alla distorsione dei risultati finali. Il recente studio "Invasion of the PR-bots" (2018) pubblicato da Kozinets ha reso noto che circa il 52% di tutto il traffico web è composto da bot e il 56% di questi risulta essere dannoso per i siti web. Inoltre, è stato specificato che il circa il 15% degli account Twitter sono in realtà bot difficilmente individuabili in quanto il loro comportamento è molto simile alle attività degli esseri umani. È anche utile specificare che l'attività dei bot potrebbe rappresentare una grande porzione del traffico web in quanto risultano sempre attivi.

Un'altra sfida importante per il ricercatore è quella di navigare a fondo nel web e di identificare le comunità chiave che forniscano informazioni adeguate e utili alle esigenze della ricerca. Inoltre, a disturbare l'autenticità dei dati vi è stato l'ingresso della commercializzazione nelle comunità online. Molti post vengono pubblicati con il solo intento di guadagnare soldi. Si tratta di influencer che vengono pagati dalle aziende per pubblicare messaggi pubblicitari che "sporcano" la veridicità dei dati. Allo stesso modo è possibile imbattersi in dipendenti che nascondendo la loro vera identità partecipano ai blog e social media con l'intento di pubblicizzare i propri prodotti e magari screditare quelli dei concorrenti. È quindi necessario per il ricercatore porre maggiore attenzione nella selezione dei messaggi da analizzare.

I ricercatori, inoltre, per analizzare qualsiasi tipologia di comunità dovrebbero rispettare i principi etici. Un elemento chiave è dunque quello di ottenere il consenso informato da parte dei membri per procedere con la ricerca. Tale processo risulta però essere una sfida importante e molte volte impossibile da superare dal momento che molti utenti non

percepiscono la propria attività online come possibile oggetto di studio e si rivelano dunque contrari all'analisi.

Si entra poi nella problematica della privacy. I ricercatori hanno il dovere di mantenere l'anonimato sui membri; tuttavia, diversi sono gli utenti che preferiscono usare identità false per pubblicare informazioni personali impattando così sull'attendibilità dei risultati. A fronte di tale problematica è utile dunque cercare di studiare comunità con membri legati tra loro i quali difficilmente tendono a usare identità false e informazioni ingannevoli.

Inoltre, può evidenziarsi il problema di certi utenti che sapendo di essere osservati per motivi di ricerca tendono a mostrare una sorta di rifiuto evitando così di pubblicare post con pensieri personali al fine di proteggere la propria privacy (Kamboi et al., 2017).

Restando nelle tematiche riguardanti i soggetti aderenti alle comunità online, un altro motivo di distorsione della ricerca può cagionarsi dalla presenza di messaggi online da parte dei membri solo di opinioni estreme, o di soddisfazione o di reclami. Infatti, utenti con opinioni neutre tendono a rimanere passivi e solo osservatori, non interagendo con i propri pensieri.

Tale comportamento viene fatto risalire a diversi fattori, come il desiderio di evitare conflitti, la mancanza di una opinione forte su un argomento, o semplicemente il non sentirsi abbastanza informati per contribuire attivamente alla discussione. Inoltre, molte volte vi è il rischio che i messaggi pubblicati da alcuni utenti delle comunità tendono a esercitare un impatto significativo sulle opinioni e comportamenti degli altri membri.

Ulteriore elemento di debolezza può ravvisarsi nella limitazione della comunicazione tecnologica. Infatti, tramite il web è possibile analizzare solo informazioni di natura testuale o al massimo informazioni fotografiche. Si verifica così la perdita di tutte le informazioni comunicative provenienti dal linguaggio del corpo, tonalità vocali, pause, voci spezzate e espressività.

Capitolo 2

Evoluzione della netnografia

In questo capitolo, esploreremo gli aspetti fondamentali che stanno promuovendo e facilitando l'evoluzione della netnografia negli ultimi anni. Dopo un'introduzione sui nuovi contesti resi disponibili all'analisi attraverso questa tecnica verrà dedicata una particolare attenzione all'uso dell'intelligenza artificiale in questo campo. Offriremo dunque una panoramica generale sull'intelligenza artificiale, esaminando i principali concetti e tecnologie, per poi indagare come essa possa essere efficacemente integrata e utilizzata per potenziare la ricerca netnografica. Verranno infine discusse le sfide e le questioni etiche intrinseche all'uso dell'intelligenza artificiale, identificando anche alcune possibili strategie per mitigare eventuali violazioni degli standard etici.

2.1. Sviluppo recente della netnografia

Come osservato nel capitolo precedente, la netnografia fonda le sue radici sulla nascita ed evoluzione del WEB 2.0. Pertanto, è possibile affermare che lo sviluppo di tale metodo risulti guidata dall'espansione della comunicazione online e dalla crescente ed esponenziale diffusione dei social media. L'espansione del WEB 2.0, caratterizzata da una continua maggiore interattività e partecipazione degli utenti, ha creato nuovi spazi per le persone di interagire, condividere contenuti e formare comunità.

Con l'evoluzione della società e del web nuovi contesti si sono resi disponibili all'uso di questa tecnica netnografica.

Infatti, oltre alle già note applicazioni nell'ambito del marketing, turismo, salute e servizi finanziari è possibile introdurre tale metodologia nelle seguenti circostanze.

- Lo sviluppo di un'economia futura sostenibile vede una trasformazione dalla vendita di prodotti fisici alla vendita di servizi orientati al risultato. Si tratta di servizi che tendono a sostituire la proprietà e il funzionamento dei prodotti e che i clienti pagano solo per il risultato che esso offre. Molti di questi servizi vengono offerti su internet, piattaforma nella quale spesso i clienti si soffermano per discutere criticità, pensieri e novità. Dunque, attraverso la netnografia i ricercatori hanno la piena opportunità di studiare questo fenomeno e le relative implicazioni che si creeranno per la società.
- La crescente apertura mentale delle realtà aziendali verso nuovi approcci di ricerca permette agli studi netnografici di entrare nel dettaglio anche nel settore del B2B (Business-To-Business). Questo fornisce ai ricercatori la possibilità di ottenere una maggiore comprensione degli orientamenti di pensiero dei manager e ricavare informazioni sui processi decisionali e sulle strategie di leadership. Inoltre, analizzando le conversazioni online all'interno di chat B2B emergono anche gli sforzi che vengono eseguiti per le attività di marketing e come le tattiche vengono adattate alle tendenze emergenti. Con lo studio delle realtà aziendali, infine, vi è la possibilità di scoprire maggiori informazioni sui dipendenti, i quali molto spesso tendono a discutere e rivelare informazioni sul benessere del luogo di lavoro, sulle dinamiche di gruppo e sulla cultura aziendale. Un altro importante vantaggio può derivare dall'identificazione delle differenze tra i modelli mentali dei venditori e quelli degli acquirenti.

Affrontare dunque in una logica di netnografia il tema permette un importante arricchimento nella comprensione, da parte delle aziende, delle attese, esigenze, bisogni ma anche proiezioni immaginarie dei propri stakeholder, tanto in ambiente interno che esterno. Un know-how che, una volta acquisito, faciliterà decisioni di business e strategie implementative mirate e pertanto, con maggior grado di prevedibile successo, perché già plasmate su concrete informazioni apprese.

- L'ormai diffusione e comune presenza di siti, forum e social media internazionali crea enormi opportunità per analisi netnografiche basate su argomenti relativi alle

esperienze di servizi e prodotti tra clienti di cultura e nazionalità diversa. È possibile analizzare le diverse interpretazioni che vengono eseguite dagli utenti sui servizi al variare della cultura. Questa potrebbe essere una fonte di un valore inestimabile per le aziende che si occupano di creazione di campagne di marketing e prodotti mirati alle diverse nazioni.

Un altro aspetto notevole è l'impatto che i nuovi paradigmi di consumo hanno sulla netnografia. Si sta affermando una comprensione sempre più diffusa del fatto che le strategie di mercato delle aziende siano profondamente influenzate dalle evoluzioni nei modelli di consumo e acquisto, originati da una varietà di fenomeni. Un esempio è il consolidamento di atteggiamenti di consumo "razionale". «I consumatori sono più informati, più consapevoli, più attenti al “value for money” dei prodotti e dei punti di vendita. Utilizzano, per l'attività di scouting, molti touch point e di fronte agli scaffali si pongono più domande, leggono con più attenzione le etichette dei prodotti, fanno più confronti tra le marche» (Fornari e Ghezzi, 2023).

Parallelamente, assistiamo a una "socializzazione" dei comportamenti di consumo.

«Nei processi di acquisto i social sono diventati un touch point rilevante per raccogliere informazioni, suggerimenti e giudizi nei confronti dei prodotti e dei punti di vendita, arrivando a condizionare le performance commerciali dei produttori e dei distributori».

I consumatori tendono quindi a condividere sensazioni valutazioni sulle loro esperienze, tentando così di creare quel terreno di "dialogo aperto", che è diventata la nuova piazza concreta dove si gioca il successo o il fallimento di un'iniziativa commerciale.

In questo contesto, la netnografia emerge come strumento prezioso in grado di soddisfare i nuovi modelli di consumo. Infatti, attraverso l'analisi delle comunità online e dei comportamenti digitali, la netnografia offre alle aziende la possibilità di interpretare in tempo reale questi cambiamenti, anticipando le preferenze di acquisto. Questo approccio permette alle aziende di sviluppare strategie di marketing mirate, di migliorare la trasparenza e la responsabilità del brand e di interagire in modo più efficace con i consumatori.

Oltre alla nascita di nuovi contesti e ai cambiamenti dei modelli di consumo è fondamentale considerare la presenza di nuove tipologie di dati. Negli ultimi anni i dati

condivisi sul web sono diventati più completi. Oltre alla natura testuale i netnografi, per condurre analisi corrette, devono impegnarsi nell'affrontare dati che includono foto, video e audio. Risulta attualmente scontata la presenza di queste informazioni sul web, dal momento in cui sono nate comunità appositamente per la condivisione e scambio di contenuti visivi e audiovisivi come Instagram, TikTok e YouTube. È possibile affermare dunque che le informazioni testuali, da sole, possono non essere sufficienti per comprendere pienamente un contesto o un fenomeno culturale online.

L'introduzione di nuovi contesti di analisi e l'espansione sempre più ampia di spazi digitali frequentati da utenti moderni, conduce inevitabilmente a una presenza di un volume di dati disponibili e utili alla netnografia sempre maggiore.

Una netnografia "manuale" basata solo sul ruolo e l'attività dell'uomo per la raccolta e analisi dei dati risulta quindi sempre più complicata e svantaggiosa a causa dell'elevato impegno di personale necessario e per il limitato volume di contenuti analizzabili. Il metodo antropologico si trova, insomma, proiettato verso un necessario aggiornamento, che gli consenta di ottenere il massimo dalle informazioni disponibili generate volontariamente dagli utenti. Le modalità delle relazioni sociali, dello scambio di idee, di pensieri e di alcune attività lavorative appaiono oggi come incanalate dentro un tunnel che scaturisce verso un'enorme nuvola digitale pronta a raccogliere informazioni sotto forma di dati.

È necessario precisare che, poiché nativa dell'ambiente online, la netnografia ha sempre incluso l'uso di metodi assistiti da computer per la raccolta e analisi di dati, come supporto ai metodi tradizionali come l'osservazione partecipante del ricercatore e l'interpretazione qualitativa dei risultati.

Questa netnografia digitale, come la definisce Kozinets nel suo elaborato del 2015, rappresenta un'evoluzione in cui l'analisi qualitativa, focalizzata sulla comprensione del contesto e del significato delle interazioni online, è affiancata da tecniche quantitative. L'obiettivo è quello di ottenere informazioni più dettagliate e profonde attraverso la codifica e l'analisi statistica dei dati raccolti. L'approccio ibrido permette ai ricercatori di ottimizzare i pregi di ciascuna di queste metodologie, offrendo una comprensione più ricca e sfaccettata dei fenomeni online. Il grande plus della netnografia digitale è di combinare l'osservazione dettagliata e l'interpretazione tipiche della ricerca qualitativa

con la precisione e l'ampiezza delle analisi quantitative, consentendo una visione più profonda delle dinamiche culturali e sociali che si svolgono nell'ambiente online.

Tuttavia, per sincronizzarsi con la natura mutevole del contesto digitale e per raggiungere un livello superiore in termini di velocità, profondità e precisione di raccolta e analisi dati, la netnografia necessita di iniziare una collaborazione con programmi più strutturati di intelligenza artificiale.

2.2. Intelligenza artificiale

Il tema dell'Intelligenza Artificiale (IA) si rivela particolarmente delicato e stimolante, interessando sia l'attualità sia, ampiamente, il futuro. Questo è dovuto alle sue notevoli potenzialità e all'influenza significativa che esercita sulla società, nonché alle inevitabili implicazioni future che ne derivano.

Per una migliore comprensione del tema diventa fondamentale spiegare preliminarmente alcuni concetti chiave, quali le definizioni e le nozioni cardine che consentano di attraversare un territorio molto vasto e, per larga parte, ancora non delimitato.

«L'intelligenza artificiale è una disciplina appartenente all'informatica che studia i fondamenti teorici, le metodologie e le tecniche che consentono la progettazione di sistemi hardware e sistemi di programmi software capaci di fornire all'elaboratore elettronico prestazioni che, a un osservatore comune, sembrerebbero essere di pertinenza esclusiva dell'intelligenza umana.»⁹ Così la definisce Marco Somalvico, ingegnere italiano e specialista nel campo dell'intelligenza artificiale.

Le origini risalgono al 1943 quando Warren McCulloch e Walter Pitts condussero il primo lavoro riconosciuto come IA in cui mostravano come un semplice sistema di neuroni

⁹ Definizione di Marco Somalvico, ingegnere italiano, specialista nel campo dell'intelligenza artificiale. Somalvico è stato uno dei pionieri dell'IA in Italia. Nel 1973 ha fondato il Laboratorio di Intelligenza Artificiale e Robotica (AIRLab) al Politecnico di Milano, dove ha svolto attività di ricerca e didattica per oltre 30 anni.

artificiali connessi potesse essere in grado di eseguire delle funzioni logiche basilari. McCulloch e Pitts suggerirono anche che reti opportunamente definite potevano imparare.

Il lavoro più influente fu di Alan Turing¹⁰. Nel 1950 scrisse l'articolo "Computing Machinery and Intelligence" nel quale introdusse il test di Turing, l'apprendimento automatico, gli algoritmi genetici e l'apprendimento per rinforzo.

Il test di Turing è un esperimento per definire la capacità "mentale" di una macchina, ovvero la sua facoltà di pensare o meno.

La macchina supera il test se un intervistatore umano, dopo aver posto alcune domande scritte, non è in grado di definire se le risposte provengano da una persona o da un computer. In tal modo si vuole analizzare la capacità delle macchine di replicare il processo di pensiero umano, mascherando la loro vera identità di macchine nel farlo. (Turing, 1950)

Lo studio di Turing si focalizzò anche nel dimostrare la facilità di creazione di un'IA di livello umano sviluppando algoritmi di apprendimento e insegnando poi alla macchina piuttosto di programmare manualmente la sua intelligenza.

Molti furono i lavori esplorativi di sistemi con Intelligenza artificiale. Degno di nota il gioco della dama, nel quale Arthur Samuel utilizzò metodi di apprendimento per insegnare alla macchina a giocare a un livello amatoriale forte. Herbert Gelernter (1959) inoltre, costruì il Geometry Theorem Prover, in grado di dimostrare teoremi matematici piuttosto complicati.

Naturalmente lo sviluppo di sistemi di intelligenza artificiale si è imbattuto in sfide significative che hanno causato problemi, portando come conseguenza fallimenti e limitazioni di investimenti in ricerca sull'IA causando importanti ritardi nello sviluppo di tale disciplina. Le due ragioni, identificate come principali, per la *défaillance* originaria dell'intelligenza artificiale sono state l'uso di un'introspezione informata invece di

¹⁰ Alan Mathison Turing è stato un matematico, logico, crittografo e filosofo britannico, considerato uno dei padri dell'informatica e uno dei più grandi matematici del XX secolo. Il logico e pioniere informatico britannico Turing, a metà del XX secolo svolse il primo lavoro sostanziale nel campo dell'intelligenza artificiale.

un'analisi attenta e un atteggiamento di sottovalutazione dei problemi posti all'analisi dell'automazione intelligente.

Nello specifico, l'approccio basato sull'introspezione informata si focalizzava su come gli esseri umani svolgevano un compito, fornendo così ai ricercatori la capacità di emulare i processi cognitivi umani così come li comprendevano, senza necessariamente comprendere appieno la complessità o i dettagli specifici di quei processi. Questo poteva portare a soluzioni che sembravano logiche in teoria ma che non funzionavano bene nella pratica. Sarebbe stato necessario invece un'analisi oggettiva e dettagliata del compito stesso.

Quanto al secondo fattore di freno, ovvero un atteggiamento di sottovalutazione dei problemi posti all'analisi dell'automazione intelligente, si è evidenziato che la strategia inizialmente utilizzata dal sistema, consistente nella ricerca di diverse combinazioni di passaggi per tentativi fino a quando veniva trovata la soluzione, ha avuto esiti positivi solo nei primi tempi, in presenza di ambienti molto limitati e controllati (i cosiddetti "micromondi"), dunque dotati di pochi oggetti e sequenze di soluzioni brevissime. Tali teoremi di risoluzione sono stati presto abbandonati quando i ricercatori si sono trovati a fronteggiare scenari più complessi che coinvolgevano molti più dati come nel campo della dimostrazione di teoremi.

L'eccesso di ottimismo del passato nel campo dell'IA, derivante da una comprensione limitata sia della complessità dei compiti umani sia della natura articolata di molti problemi computazionali, ha rischiato di danneggiarne l'uso.

Negli ultimi anni, invece, grazie allo sviluppo del web e ai progressi della potenza di calcolo delle macchine, si è giunti a un notevole aumento dei dati disponibili come parole di testo, immagini, audio, video e dati di reti sociali, raggruppati nell'insieme sotto il termine di Big Data.

Un arricchimento che è stato reso possibile dall'aumento della connettività, da nuove tecnologie e processori più veloci, sensori, interconnessione e archiviazione economica. In questo modo abbiamo assistito all'intensificarsi di attività di sviluppo e adozione dell'intelligenza artificiale.

Da un punto di vista tecnico, l'intelligenza artificiale è composta da algoritmi, ovvero sequenze di istruzioni o insieme di regole che vengono eseguite al fine di completare un compito specifico. Tali algoritmi possono essere letti come modelli statistici complessi i quali utilizzano la logica e il calcolo delle probabilità per ottenere la migliore soluzione possibile a partire da una serie di dati input.

Affrontando il tema della IA, risulta utile specificare che essa non ha la capacità di generare una nuova conoscenza del mondo e dei suoi fenomeni. Più precisamente, la sua quota di creatività deriva soltanto dagli algoritmi e dai dati che utilizza. Tuttavia, grazie alla potenza delle macchine di individuare similarità e differenze, può scoprire le principali correlazioni che possono sfuggire all'occhio umano e ottimizzare la sistematizzazione delle nostre categorie di pensiero. Inoltre, la capacità computazionale delle macchine permette di elaborare e analizzare dati in modo molto più rapido ed efficiente rispetto all'uomo.

Se insomma, allo stadio di avanzamento attuale, le macchine mancano di creatività umana, atti irrazionali di volontà, emozioni e desideri e operano solamente in base a regole logiche e matematiche imposte, la loro "arte" è guidata da dati e algoritmi.

Questo non impedisce che l'intelligenza artificiale possa risolvere svariate attività e la conseguente capacità di svolgere determinate azioni dipende direttamente dalle tecnologie che utilizza (Burgess, 2017).

Più specificamente, per quanto attiene alle metodologie per comprendere i consumatori, sono tre i principali approcci di intelligenza artificiale utili a fornire supporto nella ricerca netnografica e che combinati tra loro offrono diversi sistemi di apprendimento aumentando così l'efficienza. Essi sono:

- **Machine Learning (apprendimento automatico)**

L'apprendimento automatico rappresenta un sottocampo dell'intelligenza artificiale. Si tratta di una serie di algoritmi con la capacità di apprendere automaticamente grazie alle informazioni raccolte, che possono provenire da qualsiasi tipologia di fonte, purché in formato digitale e che possono realizzare una vasta gamma di attività in diversi settori (medicina, finanza, marketing etc.) spaziando dall'analisi predittiva per prevedere dati

futuri, alla classificazione, al clustering, all'analisi delle serie temporali fino ad essere utilizzati in sempre più ambiti. Insomma, gli algoritmi di machine learning analizzano dati per riconoscere pattern e tendenze, che possono poi essere utilizzati per fare previsioni o raccomandazioni personalizzate in base alle esigenze dell'utilizzatore.

Nell'ambito della netnografia può ritornare molto utile l'utilizzo di questi algoritmi per eseguire diverse analisi, come ad esempio, l'analisi dei temi principali, l'analisi dei sentimenti o anche la classificazione e categorizzazione dei dati in base ai diversi criteri.

Si tratta di una tecnologia che dota i sistemi di una capacità quasi autonoma di apprendere direttamente dalle esperienze, bypassando la necessità di programmazioni dettagliate. È straordinariamente efficace nella risoluzione di problemi complessi, raggiungendo livelli di precisione che possono competere con, e in alcuni casi superare, quelli umani.

Vi sono due principali approcci del machine learning che si differenziano per le modalità in cui i dati vengono appresi.

- L'apprendimento supervisionato, il quale è caratterizzato dall'addestramento che avviene a partire da un set di dati già etichettato. Si tratta di dati in cui ogni esempio è associato a una specifica etichetta. L'obiettivo è che il modello riesca a prevedere l'etichetta a partire da nuovi esempi non etichettati. Gli algoritmi di apprendimento supervisionato possono essere utilizzati per esempio, per classificare lo spam in cartelle differenti dalla casella di posta in arrivo.
- L'apprendimento non supervisionato ha invece come input dati non categorizzati precedentemente dall'uomo e quindi privi di etichette a cui l'algoritmo deve cercare autonomamente di dare un senso. L'algoritmo si occupa di raggruppare i dati in base alle somiglianze, trovando strutture e schemi nascosti nei dati non etichettati. Gli algoritmi di apprendimento non supervisionato consentono anche compiti di elaborazione più complessi. Infatti, nessuna etichettatura significa che possono essere mappate relazioni complesse e cluster di dati e nessuna etichettatura dei dati significa nessun preconetto e nessun pregiudizio imposto dall'essere umano. La macchina impara senza sapere la risposta corretta.

Un esempio di apprendimento non supervisionato è l'applicazione del clustering; tecnica di data mining per raggruppare dati non etichettati in base alle loro

somiglianze o differenze. Un ulteriore esempio è il topic modeling con algoritmi come LDA (Latent Dirichlet Allocation), in grado di identificare automaticamente i temi dominanti all'interno di un insieme di documenti, senza la necessità di etichette predefinite.

L'apprendimento automatico risulta essere più veloce e semplice se basato su dati strutturati ovvero formattati in maniera ordinata come tabelle con righe e colonne ben definite. Tuttavia, nel caso della presenza di dati non strutturati come testo libero, immagini, audio, o video l'operazione di apprendimento risulta più complesso. In questo caso, dunque, risulta molto utile applicare l'elaborazione del linguaggio naturale (NLP) per la pulizia e la strutturazione dei dati e alimentare così in modo corretto i modelli di apprendimento automatico. Tale processo identificato come pre-elaborazione, consiste nella preparazione e nella pulizia dei dati testuali, rendendoli adatti all'analisi da parte della macchina.

- **Deep Learning (apprendimento profondo)**

Il deep learning o apprendimento profondo è una specializzazione del machine learning basato sull'utilizzo di architetture di reti neurali profonde.

Una rete neurale artificiale è una struttura computazionale che imita l'organizzazione e il funzionamento dei neuroni biologici. Tale rete artificiale è composta da diversi strati differenti nel quale più complesso è il problema che si sta risolvendo e più strati sono necessari. Nel complesso vi sono uno strato di input che riceve i dati e inizia a codificarli, uno o più strati intermedi nascosti dove tutto il lavoro di apprendimento ed elaborazione viene svolto e uno strato di output dove viene presentata la risposta. Tali strati sono formati da diversi nodi (chiamati "neuroni"). I nodi all'interno di questi strati sono interconnessi da collegamenti che hanno dei pesi associati, rappresentanti l'intensità dell'influenza di un nodo sull'altro. Un collegamento forte, caratterizzato da un peso elevato che aveva portato a una "risposta corretta" durante l'addestramento, significherà che le informazioni vengono propagate verso il layer successivo. Un collegamento debole con un peso basso, che aveva portato a una "risposta errata" nell'addestramento, non

passerà molte informazioni. Dunque, il peso di ciascuna connessione determina quanto l'input di un nodo influenzerà l'output. Nodi con pesi più forti contribuiscono quindi a guidare le decisioni o le previsioni finali della rete. L'output di un nodo diventa quindi l'input per altri nodi nello strato successivo della rete, e questo prosegue fino a quando i dati vengono trasformati in dati utili per l'output della rete. La rete neurale, dunque, impara sempre di più nel passare da uno strato all'altro.

Man mano che il modello viene addestrato con sempre più dati, i pesi vengono continuamente regolati attraverso un processo di ottimizzazione fino a raggiungere una soluzione ottimale. Quindi più dati si alimentano nel modello, più possibilità ha la rete di affinare i pesi e quindi più accurata sarà la soluzione (Burgess, 2018).

A differenza di algoritmi di apprendimento automatico, le reti neurali grazie ai diversi strati di neuroni collegati tra loro consentono di catturare relazioni più complesse tra i dati e offrire di conseguenza risultati più accurati. Inoltre, le reti neurali si dimostrano flessibili in quanto possono essere addestrate ad eseguire determinati compiti su alcuni dataset e poi essere adattate per svolgere altri compiti, comunque simili, usando una quantità minore di dati di addestramento. Infine, con l'aumento della quantità di dati e della potenza di calcolo, le reti neurali tendono a migliorare le loro prestazioni, mentre i modelli tradizionali possono raggiungere un limite di elaborazione. Tuttavia, va notato che le reti neurali richiedono in genere una maggiore quantità di dati per l'addestramento rispetto ai semplici modelli di apprendimento automatico e possono anche essere più complesse da interpretare e da implementare.

Di seguito viene mostrato un esempio di architettura di rete neurale profonda.

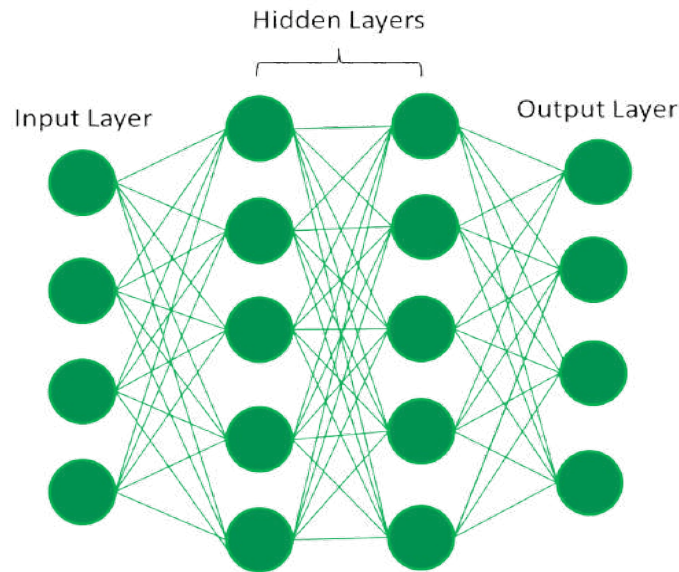


Figura 2.2-1 - Esempio di architettura della rete neurale profonda (Fonte: www.geeksforgeeks.org)

11

Il deep learning utilizza diverse reti neurali in base al compito che deve eseguire e molto spesso tende anche a combinare due reti per ottenere il massimo dell'efficienza.

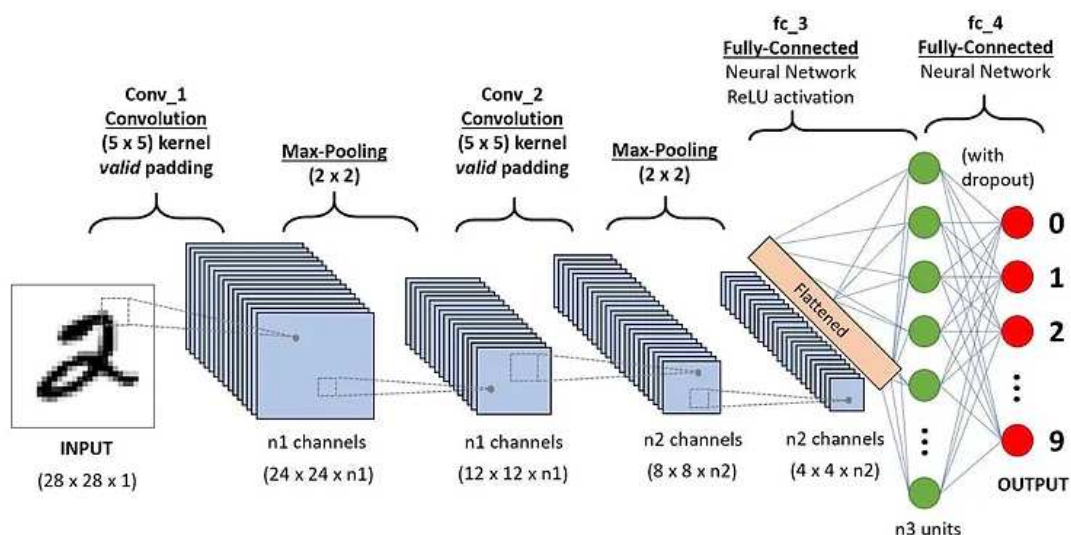
Se l'obiettivo è fornire un supporto durante la ricerca netnografica, e quindi comprendere i consumatori attraverso l'analisi di comunità online raccogliendo e analizzando i big data, le tipiche reti che vengono usate sono le:

Reti neurali feedforward (FNN): Sono la tipologia più semplice costituite da un flusso lineare di informazioni che scorre attraverso la rete. Sono largamente usate per compiti come classificazione delle immagini e per l'elaborazione del linguaggio naturale quando i dati di input non sono sequenziali.

Reti neurali convoluzionali (CNN): Sono progettate per elaborare dati che si presentano sotto forma di multiple matrici, come per esempio immagini o video a colori. Le reti CNN sono in grado di apprendere automaticamente caratteristiche dalle immagini, il che le rende particolarmente adatte per compiti come la classificazione delle immagini, il

¹¹ L'immagine illustra l'architettura di una rete neurale profonda, evidenziando i vari strati che la compongono. Mostra come i diversi livelli, dall'input ai livelli nascosti e fino al livello di output, siano interconnessi e distinti, ognuno con un ruolo specifico nella elaborazione dei dati. Le connessioni tra i nodi in ogni strato sono chiaramente visibili, dimostrando la complessità e l'efficacia del flusso di informazioni all'interno della rete.

rilevamento di oggetti e la segmentazione delle immagini. L'architettura è composta da una serie di stadi. Gli stadi fondamentali sono composti da due tipi di layer: layer convoluzionali e layer di pooling. Il convoluzionale si occupa di estrarre caratteristiche dell'immagine attraverso l'uso di filtri, creando una sorta di mappa di caratteristiche come bordi, angoli etc. dell'immagine. Il layer di pooling unisce le caratteristiche semanticamente simili in una sola. Riduce quindi la dimensionalità delle mappe di caratteristiche rendendo la rete più robusta a variazioni e distorsioni minori nell'immagine. I layer convoluzionali sono dunque responsabili dell'estrazione di caratteristiche e della costruzione di una rappresentazione gerarchica delle immagini, mentre i layer di pooling consolidano queste informazioni migliorando la generalizzazione della rete. Possono essere applicati più strati convoluzionali e di pooling al fine di ridurre la dimensionalità e ridurre distorsioni. Le mappe vengono infine appiattite in un vettore unidimensionale e poi passate in strati completamente connessi per un'ulteriore elaborazione. Con strati completamente connessi si fa riferimento agli strati utilizzati nelle reti neurali, dove ogni nodo (o neurone) è connesso a ogni nodo dello strato precedente. Questo significa che ogni valore nel vettore appiattito diventa un input per questi strati. Di seguito una raffigurazione di tale rete.



2.2 - Raffigurazione di una rete neurale convoluzionale (Fonte: www.intelligenzaartificialeitalia.net)

Reti neurali ricorrenti (RNN): Reti in grado di elaborare input sequenziali come il linguaggio naturale. Le RNN hanno la capacità di elaborare una sequenza di input un elemento alla volta, mantenendo nelle loro unità nascoste un "vettore di stato" che contiene implicitamente le informazioni sulla storia di tutti gli input precedenti. Tale rete è in grado di trasformare una sequenza di elementi come delle parole, in un'altra sequenza di elementi. Ha dunque diverse capacità come la previsione del carattere successivo nel testo, la traduzione linguistica e il riconoscimento vocale. Tuttavia, queste reti sono sempre più difficili da gestire all'aumentare della lunghezza delle sequenze.

Inoltre, se combinata con la rete CNN è in grado di tradurre testualmente qualsiasi immagine. L'encoder in questo caso è una rete convoluzionale profonda che converte i pixel dell'immagine in un vettore di caratteristiche nel suo ultimo strato nascosto. Tali caratteristiche vengono poi trasmesse al decoder, ovvero la rete RNN simile a quella per la traduzione automatica e l'analisi del linguaggio naturale. La RNN interpreta dunque il vettore di caratteristiche fornito dalla CNN e genera una descrizione testuale della scena.

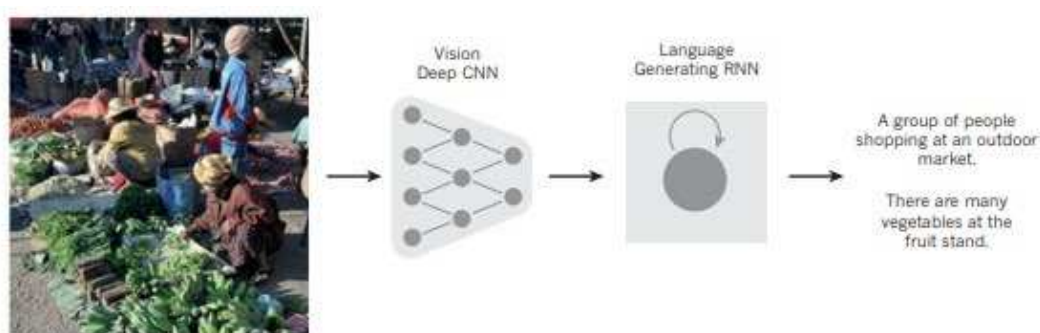


Figura 2.2-2 - Esempio di rete CNN combinata con rete RNN (fonte: LeCun et al., (Vol. 521, 2015), Deep learning)

12

Reti neurali basate su Transformer: Il Transformer è un modello di rete neurale che utilizza la meccanica dell'attenzione, in cui il modello assegna un peso alle diverse parti

¹² L'illustrazione rappresenta graficamente l'interazione tra una rete neurale convoluzionale (CNN) e una rete neurale ricorrente (RNN) nel processo di conversione di un'immagine in testo. Inizia con un'immagine di un mercato di verdure, che funge da input per la CNN. Dopo aver elaborato l'immagine, la CNN passa le informazioni pertinenti estratte alla RNN. La RNN elabora queste informazioni sequenzialmente, tenendo conto del contesto e della struttura delle informazioni ricevute, e procede a generare una descrizione testuale dell'immagine in inglese, sintetizzando le caratteristiche visive in parole.

dell'input in base alla loro rilevanza per la comprensione del testo. Questo meccanismo permette alla rete di concentrarsi su elementi specifici della sequenza e stabilire riferimenti tra di essi, simili a come il cervello umano focalizza l'attenzione su parole chiave durante l'ascolto o la lettura. I Transformer elaborano gli input in modo parallelo e possono gestire sequenze di lunghezza arbitraria, a differenza delle RNN che procedono sequenzialmente e hanno limitazioni sulla lunghezza delle sequenze. Questa caratteristica li rende particolarmente vantaggiosi nel campo dell'elaborazione del linguaggio naturale (NLP) in quanto consente al modello di acquisire una migliore comprensione del contesto del testo analizzato. Questi modelli utilizzano una serie di blocchi di trasformazione che permettono l'elaborazione parallela dell'input. L'architettura di un Transformer è generalmente divisa in due parti principali: un codificatore, che elabora l'input, e un decodificatore, che genera l'output. Questa struttura consente ai Transformer di gestire compiti complessi di elaborazione del linguaggio e di altre sequenze dati in modo molto efficace.

Per la netnografia queste reti basate su transformer risultano quindi molto utili. Per esempio, i modelli di analisi del sentimento basati sui transformers sono in grado di catturare le sfumature del linguaggio naturale, come le ironie e le ambiguità, che possono influenzare la comprensione del sentimento di un testo.

Come affermano dunque Zhang et al., (2017) nell'articolo "Deep Learning for Sentiment Analysis: A Survey", il deep learning risulta essere estremamente adatto per analizzare ed estrarre conoscenze da grandi quantità di dati, raccolti anche da fonti diverse.

- **Natural Language Processing (NLP)**

Con il termine NLP si identifica la tecnologia basata sull'intelligenza artificiale in grado di consentire alle macchine di leggere, comprendere e infine generare il linguaggio naturale nelle stesse modalità con cui viene fatto dall'essere umano.

Secondo quanto affermato da Stuart Russell e Peter Norvig nel loro elaborato "Artificial Intelligence: a modern approach", sono tre le ragioni principali per cui una macchina dovrebbe eseguire l'elaborazione del linguaggio naturale:

- 1) Comunicare con gli esseri umani¹³
- 2) Apprendere¹⁴
- 3) Avanzare nella comprensione scientifica delle lingue e dell'uso del linguaggio, utilizzando gli strumenti dell'IA in collaborazione con la linguistica, la psicologia cognitiva e le neuroscienze.

Tuttavia, il linguaggio naturale dell'essere umano risulta essere ambiguo e vago. I significati linguistici variano spesso da persona a persona e da momento a momento e spesso due occorrenze della stessa parola possono riferirsi a due aspetti differenti.

Il processo che consente di comprendere e interpretare il linguaggio naturale dell'uomo avviene attraverso due fasi fondamentali: la preelaborazione dei dati e lo sviluppo di algoritmi specifici. La fase di preelaborazione si occupa di preparare e "pulire" i dati testuali in modo che possano essere facilmente analizzati dalla macchina. Per tale operazione le tecniche più diffuse sono:

- Tokenizzazione: Suddivisione del testo in unità più piccole (come parole o frasi).
- Rimozione delle stop-words: Eliminazione di parole comuni che non aggiungono valore significativo all'analisi.
- Lemmatizzazione o stemming: Riduzione delle parole alla loro forma base.
- Part-of-Speech Tagging: Assegnazione di etichette grammaticali alle parole.

Dopo la preelaborazione, si sviluppano algoritmi in grado di elaborare e processare le informazioni. Questi algoritmi si possono basare sia su regole e lessici o modelli linguistici sia su tecniche di apprendimento automatico per comprendere, interpretare e generare risposte al linguaggio naturale.

Tali modelli basati sull'apprendimento automatico necessitano però di addestramento tramite set di dati in modo da sviluppare le abilità. Questi set di dati includono una vasta

¹³ In molte situazioni, è conveniente, per gli esseri umani, interagire con i computer tramite il linguaggio parlato utilizzando il linguaggio naturale invece del linguaggio simbolico.

¹⁴ Molte sono le informazioni e conoscenze generate e disponibili sul web sotto forma di linguaggio naturale. La macchina, per apprendere al meglio tali informazioni e per riusarle in compiti specifici, deve riuscire a comprendere il linguaggio utilizzato.

gamma di testi che aiutano il modello a comprendere come vengono comunemente usate le parole, come vengono formate le frasi, e come si strutturano i testi in diverse lingue.

Tuttavia, come specificato prima, le lingue naturali sono molto complesse e qualsiasi modello verrà utilizzato sarà sempre comunque una approssimazione.

Gli algoritmi sviluppati per elaborare e interpretare il linguaggio naturale possono essere di diversi tipi, ognuno con capacità e applicazioni specifiche. Di seguito qualche esempio di ciò che tali algoritmi sono in grado di svolgere.

- Analisi del testo e individuazione di elementi chiave.
- Interpretazione di un testo per classificarlo in una categoria predefinita (es. spam).
- Analisi del sentimento, ovvero il rilevamento dell'umore all'interno di un testo.
- Generazione di un testo.
- Traduzione di testi
- Riconoscimento del parlato

Oggi, le tecnologie di elaborazione del linguaggio naturale hanno la capacità di esaminare quantità illimitate di dati testuali senza stancarsi, mantenendo un approccio coerente e neutrale. Considerando l'immensa quantità di dati non strutturati generati quotidianamente, come i post sui social media, l'automazione fornita da queste tecnologie risulta fondamentale per un'analisi efficiente di dati basati su testo.

Inoltre, con l'avvento dei big data, la combinazione dell'algoritmo NLP con modelli di machine learning, deep learning risulta molto utile e quasi obbligatoria per comprendere profondamente il linguaggio umano, interpretando non solo le parole, ma anche il contesto e le sfumature del linguaggio. L'applicazione delle reti neurali, per esempio, alla elaborazione del linguaggio naturale comporta sempre più un potenziamento della capacità di comprendere e generare testi naturali.

2.3. Intelligenza artificiale a supporto della netnografia

L'esplosione del volume di informazioni disponibili sul web e sui social media avvenuta negli ultimi anni ha reso la tecnica della netnografia fondamentale e sfidante per il mondo attuale del business. Oggi tutti questi flussi di dati digitali disponibili e tutti i dati provenienti dall'infinità di fonti in grado di generare contenuto digitale rientrano nella categoria di big data. Come afferma anche IBM, i big data possono essere considerati tali se caratterizzati dalle tre V: elevato volume, elevata varietà ed elevata velocità, dove con volume si intende un aumento della quantità di dati generati, con varietà ai diversi tipi di dati che vengono usati per descrivere una situazione e con velocità la rapidità di generazione dati.

Tutte le informazioni utili a un ricercatore netnografico rientrano in questa categoria. Volume, dovuto all'incremento di dati disponibili nel web; Varietà, dovuta alla presenza nei blog, forum e social media di dati strutturati¹⁵ e non strutturati¹⁶; Velocità, dovuta all'aumento della frequenza di informazioni pubblicate dagli utenti nelle comunità.

In questo contesto, la netnografia è diventata un ambito di applicazione per l'intelligenza artificiale, la quale è in grado, grazie ad appositi algoritmi e reti neurali, di fornire un supporto al lavoro umano indispensabile. Infatti, combinando la capacità dell'IA di fornire supporto quantitativo e qualitativo, con la raccolta di dati qualitativi e l'empatia del ricercatore, è possibile ottenere una comprensione più olistica del comportamento umano.

Inoltre, è possibile affermare che l'avvento dei big data e l'avanzamento delle tecniche di intelligenza artificiale, assottigliano sempre più il confine tra netnografia (osservazione e analisi qualitativa dei comportamenti sociali in ambienti online) e analisi quantitativa.

¹⁵ I dati strutturati sono dati che sono stati predefiniti e formattati in una struttura impostata prima di essere pubblicati online. Sono dunque, dati che aderiscono a un modello o schema predefinito, rendendoli facilmente leggibili e manipolabili da sistemi informatici. Un esempio sono le statistiche del weblog e i dati dei punti vendita, come i codici a barre e la quantità o i fogli di calcolo.

¹⁶ I dati non strutturati non seguono un modello o schema specifico e, pertanto, non sono organizzate in un formato facilmente leggibile da sistemi informatici. Questo tipo di dati rappresenta la maggior parte dei dati disponibili nel mondo digitale di oggi. Sono i dati che vengono pubblicati dalla maggior parte degli utenti delle comunità online e includono testi, immagini video e audio.

I ricercatori stessi, come afferma Kozinets in “Netnography Unlimited”, con l’aumento del volume di dati da usare nelle proprie ricerche e con l’aumento del tempo necessario per svolgere tale operazione, hanno espresso la loro volontà nell’utilizzare strumenti di IA. Tale tecnologia, infatti, non solo riduce notevolmente il tempo necessario per elaborare grandi volumi di dati, ma ottimizza anche la selezione delle fonti. Ciò permette una maggiore focalizzazione su argomenti di rilevante interesse, affinando l’analisi su temi emergenti e tendenze significative. In aggiunta, le analisi effettuate con strumenti basati sull’IA si rivelano spesso più complete e specifiche. Gli algoritmi avanzati utilizzati nell’IA possono infatti identificare pattern e connessioni che potrebbero sfuggire all’analisi umana, portando a scoperte più approfondite. L’IA, inoltre, ha la capacità di adattarsi e apprendere dai dati, migliorando costantemente in qualità ed efficienza nell’analisi. Tali aspetti rendono nel 2023 l’IA uno strumento indispensabile nella netnografia, facilitando un’analisi più rapida, precisa e focalizzata.

Le condizioni che hanno permesso l’introduzione di queste tecnologie basate sull’IA sono state:

- L’aumento esponenziale del volume di dati generati dagli utenti
- L’avanzamento nella qualità degli algoritmi i quali hanno consentito ottimizzazioni in termini di accuratezza, apprendimento automatico e profondo ed elaborazione del linguaggio naturale. Questi progressi hanno reso gli algoritmi in grado di analizzare dati complessi, interpretare il linguaggio umano con maggiore precisione e funzionare in modo più efficiente.
- La sempre più diffusa disponibilità di computer con potenza elevata.

L’introduzione dell’uso dell’intelligenza artificiale come strumento per la netnografia è tuttavia limitato solamente ad alcune fasi del processo di ricerca. Questo perché, ad oggi, le macchine non hanno ancora acquisito la capacità di empatizzare con gli esseri umani e con i dati raccolti. Questa mancanza si traduce in una limitazione significativa: l’incapacità delle macchine di generare intuizioni profondamente rivoluzionarie riguardo ai comportamenti e alle preferenze dei consumatori. Tale osservazione pone quindi in evidenza il valore insostituibile della comprensione umana nell’ambito dell’analisi dei dati. Ne consegue che compiti come la generazione di insight, assieme alla comprensione approfondita del contesto e alla identificazione di parole pertinenti per la raccolta dei dati

necessari all'analisi, rimangono prerogativa essenziale del ricercatore. In questo contesto la componente umana risulta dunque ancora fondamentale. Come afferma Camilla Masala¹⁷ nel dibattito "Reimagine Netnography with AI", umano e macchina non si sostituiscono ma si completano e ognuno trae vantaggio da quello che l'altro sa fare meglio.

L'intelligenza artificiale risulta quindi molto utile e ormai fondamentale nelle fasi di raccolta e analisi dati.

I ricercatori, con l'attuazione del metodo misto (IA e umano), hanno la possibilità quindi di investire molto più tempo nella fase "creativa" della ricerca ovvero nel tradurre i risultati offerti dall'intelligenza artificiale in insight distintivi e migliorare così l'evoluzione di prodotti, servizi e campagne di marketing.

Entrando più nello specifico andiamo ora ad analizzare le diverse fasi delle netnografia identificando dove e con quali strumenti l'IA è in grado di supportare il ricercatore.

Nella fase di pianificazione, il ricercatore deve focalizzarsi sull'identificazione degli obiettivi della ricerca e dell'argomento da trattare in base alle proprie necessità e sulla definizione delle parole chiave per la ricerca delle comunità e dei contenuti. Essendo input definiti a partire da un'elaborazione neurale dell'essere umano a fronte di un bisogno specifico, l'intelligenza artificiale può fornire un supporto minimo solamente nella ricerca di parole chiave maggiormente rilevanti.

Si tratta però di un aiuto minimo in quanto si rende necessario l'inserimento manuale di almeno una parola chiave per ottenere il risultato dalla macchina. Inoltre, il ricercatore deve assicurarsi che le parole generate coprano l'intero argomento di ricerca. Questo perché la validità di questa fase è direttamente collegata a tutte le successive e una definizione errata comprometterebbe tutta la ricerca, specialmente se è previsto l'uso di intelligenza artificiale nelle fasi di raccolta e analisi dati. Infatti, il processo del metodo basato sulle macchine risulta essere molto più rigido rispetto all'approccio incentrato sull'uomo nel quale sarebbe possibile includere aspetti dell'argomento in un secondo momento (Kozinets, 2020).

¹⁷ Camilla Masala ha il ruolo di Service Design Lead nell'azienda Avenade.

La fase di entrée, essendo un'attività delicata, è riservata ancora per la maggior parte alla azione umana del ricercatore. Tuttavia, l'IA, grazie alle tecniche di NLP, riesce a supportare il ricercatore nell'identificazione rapida dei luoghi dove i contenuti correlati alle parole chiave vengono discussi.

L'intelligenza artificiale offre invece un supporto fondamentale nei due step successivi della netnografia: la raccolta e l'analisi dati.

Per la raccolta dati, una volta individuata/e la o le fonti online da analizzare è possibile utilizzare metodi come il web scraping per scaricare e raccogliere automaticamente il maggior numero di informazioni in forma "grezza".

Il web scraping è una procedura che utilizza strumenti tecnologici per estrarre automaticamente e organizzare dati del web con l'obiettivo di analizzare in profondità tali dati (Krotov e Tennyson 2018). Tale procedura è composta dalle seguenti fasi interconnesse tra loro: analisi del sito web, crawling del sito web e organizzazione dei dati.

Con analisi del sito web si identificano le procedure di analisi della struttura HTML della pagina web per comprendere le modalità e la posizione di memorizzazione dei dati di interesse. Ottenuta la posizione nel quale sono archiviati i dati si passa alla fase del crawling. Il crawling del sito web implica lo sviluppo e l'esecuzione di uno script in grado di navigare automaticamente nel sito web e di recuperare i dati necessari. Infine, dopo l'estrazione, si passa alla pulizia e all'organizzazione dei dati. In questa fase, come anche nel crawling, vengono utilizzati linguaggi di programmazione spesso integrando anche l'elaborazione del linguaggio naturale per effettuare la pre-elaborazione dei testi. I dati vengono così preparati per la successiva analisi. Strumenti di intelligenza artificiale vengono quindi direttamente applicati ai dati scaricati per aumentare la qualità, l'efficienza e l'efficacia di raccolta.

Tuttavia, nel momento di applicazione del web scraping, il ricercatore deve tenere in considerazione le numerose limitazioni della privacy presenti nei diversi siti web. Infatti, il processo di estrapolazione dei dati generati da altri utenti per usi personali può essere visto come un atto di violazione. Verranno, nel paragrafo successivo, affrontate anche queste problematiche.

Nel processo di analisi dei dati, che si attua successivamente alla loro raccolta, l'intelligenza artificiale gioca un ruolo cruciale nel semplificare e velocizzare l'esame delle informazioni. Questa tecnologia è in grado di rilevare elementi essenziali sia nei dati qualitativi (come opinioni o descrizioni) sia nei dati quantitativi (come cifre e statistiche). Inoltre, attraverso l'impiego combinato delle varie metodologie di IA precedentemente menzionate, si può massimizzare l'efficacia di questo strumento nell'ambito della netnografia, permettendo di ottenere un'analisi dettagliata e profonda del comportamento dei consumatori.

Un primo esempio di algoritmo a supporto della fase iniziale di analisi, come suggerisce Kozinets (2020), è il Support Vector Machine (SVM), basato sull'apprendimento supervisionato e in grado di classificare con ottimi risultati i testi (Armancio et. al., 2014). Nel caso di una ricerca netnografica può sostituire l'azione umana nell'etichettare, come validi per l'analisi, i contenuti generati dai consumatori e, come non validi, quelli non generati dai consumatori. Tuttavia, per una corretta applicazione, l'algoritmo SVM necessita di un addestramento su un set di dati etichettati manualmente.

Identificati i contenuti utili all'analisi è possibile applicare alcuni strumenti per sviscerare le informazioni che si nascondono tra i dati. Uno di questi è il text mining.

Il Text Mining, una tecnica avanzata nell'ambito dell'Intelligenza Artificiale, impiega l'elaborazione del linguaggio naturale (NLP) e, in certi casi, anche il machine learning per analizzare e trasformare l'ampia mole di testi non strutturati prodotti dagli utenti. Questi testi includono dunque i contenuti di pagine web, forum, blog, e-mail, oltre a post e commenti sui social media. Attraverso il text mining, questi dati vengono in primo luogo convertiti in un formato strutturato e normalizzato, facilitando così l'identificazione successiva di pattern, tendenze, e sentimenti espressi. Tale tecnica consentendo un'analisi e un'organizzazione dei grandi volumi di informazioni testuali raccolti fornisce di conseguenza un enorme supporto nell'identificare insight preziosi sul comportamento e le preferenze dei consumatori.

Nello specifico, nel text mining, la fase di pre-elaborazione dei dati raccolti prevede, grazie all'uso dell'elaborazione del linguaggio naturale, azioni come la rimozione delle

stop-words¹⁸, la rimozione di punteggiatura, la tokenizzazione¹⁹ e la lemmatizzazione²⁰. Successivamente, con a disposizione un testo più strutturato, è possibile procedere a svariate analisi come per esempio all'analisi della semantica, all'analisi del sentiment e all'identificazione di temi principali che emergono in una discussione

L'analisi semantica è il processo che consente di stabilire il significato delle parole contenute in un testo scritto. Le reti neurali hanno la capacità di decifrare il significato contestuale nascosto nelle parole, superando i limiti dell'analisi sintattica di base. Questa capacità risulta estremamente vantaggiosa nell'interpretare testi che presentano complessità o ambiguità, consentendo una comprensione più profonda e accurata del loro contenuto.

L'analisi del sentiment, che si avvale di metodi statistici, di apprendimento automatico e, in alcuni casi, di tecniche di deep learning (reti neurali), permette di valutare il tono emotivo dei testi raccolti. Questo processo consente di determinare se il sentimento espresso dai clienti nei confronti di un prodotto o marchio sia positivo, negativo o neutro. Tale analisi offre alle aziende la possibilità di comprendere meglio le percezioni e le opinioni dei consumatori, fornendo insight preziosi per la strategia di marketing e lo sviluppo del prodotto.

Per l'identificazione dei temi principali, il topic modeling rappresenta un approccio efficace. Questa tecnica, appartenente al campo dell'elaborazione del linguaggio naturale (NLP), ha l'obiettivo di scoprire e categorizzare le tematiche principali presenti in un insieme di documenti. I temi vengono identificati dagli algoritmi di topic modeling basandosi sulla frequenza e sulla co-occorrenza delle parole nei testi. Uno degli algoritmi più noti e utilizzati per questo scopo è il Latent Dirichlet Allocation (LDA)²¹, un modello

¹⁸ Si tratta di congiunzioni, articoli, preposizioni, ausiliari, pronomi, aggettivi possessivi, etc. Sono parole frequenti utilizzate per la mera strutturazione sintattica della frase con scarso apporto di semantica propria

¹⁹ Si tratta di un processo che suddivide il testo in unità più piccole chiamate "token". Questi token sono tipicamente parole, ma possono anche includere punteggiatura e altri caratteri.

²⁰ Considera il contesto della frase e riduce una parola alla sua forma base o lemma. Viene così mantenuto il significato contestuale delle parole.

²¹ Algoritmo ideato da David Blei, Andrew Ng e Michael Jordan e presentato sul "Journal of Machine Learning Research" nell'articolo del 2003 intitolato "Latent Dirichlet Allocation". Il modello LDA presuppone che ogni documento sia composto da un insieme di argomenti, e che ogni parola in un documento sia associata a uno dei temi. I temi sono latenti, ovvero non sono direttamente osservabili, ma

statistico basato sull'apprendimento automatico. Tuttavia, è bene affermare che è possibile utilizzare anche modelli di reti neurali per il topic modeling consentendo così, nella maggior parte dei casi, una migliore contestualizzazione dei temi.

In aggiunta, l'uso delle reti neurali può ritornare molto utile nell'analisi visiva dei contenuti generati dagli utenti delle comunità online. Come abbiamo visto in precedenza con la combinazione delle reti convoluzionali e delle reti ricorrenti è possibile ottenere automaticamente il significato mostrato nelle immagini. Per un ricercatore netnografico risulta quindi molto più facile e veloce analizzare anche i post contenuti immagini.

Nel quarto capitolo, verranno approfondite e applicate, a un caso di studio specifico, alcune delle analisi sopra menzionate, utilizzando gli strumenti di IA rilevanti nel campo della netnografia.

Tuttavia, è utile precisare che tutte queste informazioni generate dagli algoritmi non offrono la soluzione finale su quello che il consumatore desidera o sulle problematiche esistenti. Le analisi che l'intelligenza artificiale è in grado di effettuare con estrema velocità e completezza costituiscono “solo” una base avanzata, piena di suggerimenti ma l'elaborazione finale tocca ancora e sempre all'uomo.

Come afferma Zanzotto²², una quantità considerevole di dati può essere analizzata con strumenti di apprendimento automatico, ma la complessità dell'interpretazione può essere produttiva solo con un orientamento umano (Zanzotto, 2019). Infatti, gli insight dei consumatori sono, come già ripetuto in precedenza, identificabili solo attraverso una interpretazione umana dei dati raccolti. Il talento e la capacità dei netnografi umani di “leggere tra le righe”, identificare pezzi unici di contenuti generati dai consumatori, dedurre esigenze tacite, applicare ingegnose capacità di generazione di insight e condividere gli insight sotto forma di narrazioni avvincenti che stimolano il processo di innovazione sono ancora insostituibili (Kozinets, 2020).

In conclusione, è possibile affermare che l'IA supera e può essere combinata all'azione dell'uomo nelle attività di raccolta, raggruppamento e analisi dati mentre il ricercatore ha

possono essere inferiti dal testo. Attraverso l'analisi statistica il modello cerca di scoprire i temi più rilevanti in un insieme di testi.

²² Fabio Massimo Zanzotto è professore associato presso il Dipartimento di Ingegneria dell'Impresa dell'Università di Roma Tor Vergata.

ancora in mano le attività di identificazione contenuti unici, l'analisi empatica dei dati, la generazione di insight e lo storytelling.

È infine utile che ricordare che la ricerca netnografica non consiste solo nell'analizzare le tracce lasciate dai consumatori sul web, ma prevede anche l'osservazione partecipante del ricercatore nella comunità di analisi, per ottimizzare i risultati offrendo informazioni anche sul contesto e dinamiche sociali ottenendo così una visione più completa del fenomeno studiato.

2.4. Sfide e considerazioni etiche associate all'uso dell'IA

Naturalmente un sistema così pervasivo come quello dell'intelligenza artificiale solleva acuti interrogativi e pone problematiche complesse. È impossibile, infatti, non domandarsi come un simile strumento di analisi dei dati possa essere impiegato senza violare sfere umane, quali l'etica e la privacy.

Insomma, l'evoluzione di tecnologie atte a identificare e tracciare abitudini, desideri, comportamenti, intenzioni non può stravolgere sacrosanti diritti che sono stati conquistati nella vita di collettività.

La discussione sull' Artificial Intelligence act²³ è, per esempio, da molto tempo al centro dell'attenzione delle principali istituzioni dell'Unione Europea. Consiglio europeo, Commissione e Parlamento, infatti, si stanno confrontando sulle sorti del ruolo dell'Europa nel digitale, con particolare attenzione alla protezione dei cittadini dai rischi visibili e invisibili che questa potentissima tecnologia conduce con sé.

²³ L'Artificial Intelligence Act è un regolamento dell'Unione Europea che mira a promuovere lo sviluppo e l'uso sicuro e responsabile dell'intelligenza artificiale (IA) nell'UE. Esso si applica a tutti i sistemi di IA che vengono sviluppati, immessi sul mercato o utilizzati nell'UE. Il regolamento stabilisce un quadro normativo per garantire che i sistemi di IA siano sicuri, responsabili ed etici.

Uno degli ostacoli principali a ottenere un accordo sulle modalità viene dalla corrente che ritiene sia necessario ridurre al minimo l'azione legislativa, lasciando spazio all'autoregolamentazione delle aziende al fine di prevenire utilizzi scorretti.

Due sono le grandi questioni sul tavolo: la prima è di evitare che l'intelligenza artificiale venga usata per riconoscimento facciale nei luoghi pubblici in modo troppo invasivo per la privacy dei cittadini, mentre l'altra, più aderente al tema della netnografia, si concentra sui rischi sistemici connessi ai grandi modelli generativi del tipo di chatGPT²⁴, rischi che sono riconosciuti anche dai produttori stessi.

Se la ricerca di un codice normativo condiviso non solo è auspicabile ma necessaria perché i giochi siano chiari, va comunque riconosciuto che non è affatto facile per un'impresa definire una governance capace di controllare, o comunque rallentare lo sviluppo innovativo in nome della protezione dei consumatori.

Qui, certamente, l'etica di ogni singola realtà che operi diventa fondamentale per un'efficace protezione, ma non ci si può affidare soltanto a questo. Tanto più che l'universo delle aziende vede la presenza di gruppi potentissimi rispetto a realtà più piccole, tra queste anche molte start-up, che da sole non saprebbero difendersi dagli assalti, o abusi da parte di soggetti più potenti.

Come ha chiarito il professor Dino Pedreschi dell'Università di Pisa²⁵, «L'autoregolazione non basta. Le intelligenze artificiali generative hanno prestazioni sorprendenti, ma sono immature e hanno dimostrato i loro difetti: allucinazioni, pregiudizi, discorsi antisociali. Possono avere effetti straordinari ma anche generare rischi potenti. Una regolamentazione è necessaria per garantire che i prodotti escano sul mercato a un livello di sviluppo relativamente sicuro per gli utenti».

Oggi sappiamo che anche l'assunzione di informazioni di fatto costituisce un prodotto, ovvero un gruppo di dati da cui partire per elaborare strategie e nuovi approcci commerciali. Risulta pertanto che la questione dell'uso di queste tecnologie impatta in

²⁴ ChatGPT è una chatbot basata su intelligenza artificiale e apprendimento automatico sviluppato da OpenAI. È un modello di linguaggio generativo pre-addestrato, ovvero è stato addestrato su un enorme dataset di testo e codice. Questo gli permette di generare testo, tradurre lingue, scrivere diversi tipi di contenuti creativi e rispondere alle domande di un cliente in modo informativo.

²⁵ Dichiarazione avvenuta nell'articolo "L'Intelligenza artificiale e i nodi dell'approvazione di un regolamento europeo" pubblica dal ilsole24ore.

maniera intensa sulla modalità in cui un'azienda deciderà di orientare, sviluppare e potenziare il proprio business.

Non è cosa da poco, invece la voce capace di muovere investimenti importanti e, proprio per questo, bisogna che la modalità in cui i dati vengono raccolti così come processati e comunicati avvenga su base non solo legittima ma anche legale.

Nel campo della ricerca netnografica l'IA può accelerare e semplificare molte fasi del processo, ma è importante e fondamentale considerare gli aspetti etici che ne derivano. In aggiunta ai pilastri etici forniti da Kozinets, elencati in precedenza, per una classica ricerca netnografica è necessario tenere in considerazione le nuove sfide che emergono con l'applicazione di strumenti di intelligenza artificiale a supporto di essa.

Come afferma il professor Herbert, «l'uso etico dell'IA si riferisce alla sfida di costruire la fiducia tra l'azienda e il cliente riguardo alle informazioni utilizzate dall'IA.» (Herbert, 2022)

Nello specifico di una ricerca netnografica le principali problematiche etiche che si sollevano a causa dell'intelligenza artificiale sono dovute all'applicazione del web scraping e all'automatizzazione di analisi che avviene di conseguenza.

Il web scraping, processo di estrazione di grandi quantità di dati dai siti web, potrebbe involontariamente provocare conseguenze dannose per gli utenti a causa del non rispetto dei principi etici e di privacy. Il dilemma etico sorge quando i dati pubblicamente disponibili, che possono contenere informazioni personali, vengono utilizzati senza il consenso individuale. Tuttavia, il legame presente tra i dati pubblici e il consenso da parte degli utenti alla loro raccolta e analisi su larga scala resta tutt'ora un'area grigia in ambito legale, che interpella direttamente le normative sulla privacy, come il GDPR²⁶.

²⁶ Il Regolamento Generale sulla Protezione dei Dati (GDPR), è una normativa dell'Unione Europea che ha un ruolo fondamentale nella regolamentazione dell'uso dei dati personali, inclusi quelli raccolti e analizzati nell'ambito della netnografica. Questa normativa impone rigorosi requisiti da rispettare sul consenso informato, sulla trasparenza e sulla responsabilità nell'uso ed elaborazione dei dati introducendo anche il concetto di diritto all'oblio. Queste normative creano dunque di conseguenza importanti sfide nell'applicazione dell'intelligenza artificiale nella netnografia. Per esempio il GDPR afferma che “le persone hanno il diritto di non essere soggette a decisioni basate esclusivamente su processi automatizzati” e ancora: “I processi decisionali completamente automatizzati dovrebbero essere soggetti a salvaguardie, che dovrebbero includere informazioni specifiche da fornire alla persona interessata e al suo diritto di ottenere un intervento umano, di esprimere il suo punto di vista e di ottenere una spiegazione della decisione presa – in base a quali valutazioni – e di poter quindi contestare tale decisione”.

L'applicazione del web scraping può compromettere involontariamente la privacy delle persone che partecipano alle attività offerte dal sito web (Mason 1986), svelando, per esempio, l'identità dell'utente che ha creato il contenuto. Inoltre, potrebbe accadere che i visitatori di un sito web non abbiano fornito il loro consenso per l'utilizzo dei loro dati da parte di entità terze, e dunque l'utilizzo di tali informazioni senza il consenso costituirebbe una violazione dei diritti dei soggetti di ricerca (Buchanan, 2017), causando successivamente gravi conseguenze per il proprietario del sito web e per il sito web stesso. Inoltre, le tecniche automatiche di raccolta dati possono inavvertitamente causare danni a organizzazioni terze, per esempio rivelando segreti commerciali. L'applicazione di tecniche come il web scraping può quindi entrare in conflitto con alcune normative come il GDPR il quale richiede che il trattamento dei dati personali sia legale, trasparente e giusto.²⁷ La questione diventa quindi ancora più complicata e di difficile gestione. A causa di tutte queste implicazioni, molti siti web limitano o vietano l'estrazione automatica di dati attraverso i limiti di utilizzo. Tuttavia, come ha affermato il garante della privacy nelle linee guida del 2022²⁸, “in alcuni casi, è comunque possibile ottenere un'autorizzazione speciale per la raccolta automatica dei dati dal proprietario del sito web. Questo può essere realizzato attraverso negoziazioni o accordi che rispettano sia i termini di utilizzo del sito sia le leggi sulla protezione dei dati come il GDPR.”

Risulta quindi fondamentale applicare una sofisticata procedura di entrée nella comunità garantendo una chiara comunicazione sui metodi di raccolta dati usati e sull'utilizzo degli algoritmi di IA per generare fiducia e trasparenza con i membri e riuscire così a ottenere il permesso per la raccolta dati. Questo approccio etico non solo facilita la conformità alle normative sulla privacy dei dati, ma aiuta anche a mantenere l'integrità e l'accuratezza della ricerca.

A supporto di ciò, i professori Krotov e Silva nel loro articolo “Legality and Ethics of Web Scraping”, hanno presentato una lista di domande utili che ogni ricercatore dovrebbe porsi per poter garantire gli standard etici nell'applicazione del web scraping. Esse sono:

²⁷ Affermazione del garante privacy italiano sul web scraping nell'esposizione delle linee guida pubblicate nel 2022.

²⁸ Le Linee Guida del Garante Privacy italiano sul web scraping sono un documento ufficiale che forniscono indicazioni su come il web scraping può essere utilizzato in modo conforme al GDPR.

- Il Web Scraping è esplicitamente vietato dalla politica "termini di utilizzo" del sito web?
- I dati del sito web sono esplicitamente coperti da copyright?
- Il progetto di ricerca coinvolge un uso illegale o fraudolento dei dati?
- Lo scraping può potenzialmente causare danni materiali al sito web o al server web che ospita il sito?
- I dati ottenuti dal sito web possono compromettere la privacy individuale?
- I dati ottenuti dal sito web possono rivelare informazioni confidenziali sulle operazioni delle organizzazioni che forniscono i dati o sull'azienda proprietaria del sito web?
- Il progetto che richiede i dati web può potenzialmente diminuire il valore del servizio fornito dal sito web?

È intuibile come una risposta positiva a una di queste domande potrebbe compromettere il rispetto etico dell'azione di web scraping.

Altrettanto fondamentale, nella procedura di analisi dati, vi è il mantenimento dell'anonimato e della confidenzialità dei membri. La netnografia, assistita dall'IA, deve assicurare che le informazioni identificative siano adeguatamente mascherate o rimosse.

Un'ulteriore questione etica che si verifica con l'utilizzo di strumenti di intelligenza artificiale nella fase di analisi dei dati è la problematica dei bias. Tale questione nasce dai dati utilizzati per l'apprendimento degli algoritmi. Tali dati possono contenere i bias (pregiudizi) dei creatori e di conseguenza gli algoritmi tendono ad ereditarli nel processo di addestramento. Un algoritmo di IA che perpetua o amplifica i bias esistenti, può portare, nelle fasi di analisi dei dati raccolti, a decisioni e conclusioni potenzialmente inaccurate o discriminatorie verso determinati gruppi di persone. Si tratta di una sfida molto importante e complessa da risolvere in quanto è necessario agire sui dati di input.

È importante affermare che l'IA non raggiungerà mai il suo pieno potenziale a meno che i ricercatori non eliminino le fonti di pregiudizio, e il pubblico non venga a sapere che molti pregiudizi derivano dagli errori umani nell'operare gli algoritmi e nell'interpretare i risultati dell'IA.

Infatti, un fattore che tende a bloccare e a non accettare l'uso di strumenti con IA in tutte le sue applicazioni da parte degli utenti è la paura umana.

Al di là delle storie di fantascienza e rappresentazioni mediatiche, questo timore ha radici arcaiche, persino ancestrali e affonda da sempre nella tematica di quanto è nuovo, sconosciuto. Un sentimento che viene alimentato non solo dall'immaginazione, ma in gran parte dai pregiudizi che stanno accompagnando lo sviluppo di questa tecnologia a causa dei molteplici e frequenti errori umani in fase di addestramento.

Per mitigare questi pregiudizi e paure, possono essere attivate diverse strategie illustrate nell'articolo "Pseudo AI Bias"²⁹:

- Osborne, ricercatore dell'Università di Oxford, raccomanda certificazioni professionali per utenti delle applicazioni dell'IA per migliorare la loro comprensione e ridurre i pregiudizi.
- Fornire una guida personalizzata agli utenti per aiutarli a comprendere meglio le limitazioni e i meccanismi dell'IA e per garantire un uso adeguato dell'intelligenza artificiale.
- Sviluppare approcci sistematici per monitorare e ridurre i pregiudizi, come l'uso di strumenti appropriati per ispezionare i pregiudizi negli algoritmi di apprendimento automatico.

Queste strategie sono essenziali per creare un ambiente in cui l'IA possa essere utilizzata in modo efficace e responsabile, aumentando la fiducia del pubblico in questa tecnologia.

A fronte di queste informazioni, è possibile affermare che l'intelligenza artificiale offre opportunità significative per l'avanzamento delle ricerche netnografiche, consentendo la raccolta e l'analisi di grandi quantità di dati in modo più approfondito ed efficiente rispetto ai metodi tradizionali. Tuttavia, può diventare uno strumento per l'infrangimento dei diritti umani, come la violazione della privacy o la discriminazione. Dunque, il delicato equilibrio tra lo sfruttamento dei vantaggi dell'IA e la protezione dei diritti degli

²⁹ L'articolo "Pseudo AI Bias" di Xiaoming Zhai e Joseph Krajeck si concentra sull'analisi del bias artificiale nella letteratura scientifica, definito come Pseudo Artificial Intelligence Bias (PAIB). Lo studio esamina sistematicamente la letteratura per identificare tre tipi di PAIB causati da: a) incomprensioni, b) bias pseudo-meccanico, e c) eccessive aspettative. Vengono inoltre elencate tre possibili soluzioni per eliminare/ridurre i pregiudizi nei confronti dell'intelligenza artificiale.

individui richiede un codice normativo chiaro, che regoli l'uso dell'IA nel rispetto della privacy e dell'etica.

Capitolo 3

L'etnografia mobile

Nel presente capitolo, si procede all'esame e all'esposizione di un approccio analitico per la comprensione dei comportamenti dei consumatori nell'ambito digitale: l'etnografia mobile. Questa sezione mira a offrire una disamina esaustiva di tale metodologia emergente, ponendo le fondamenta per il suo efficace impiego in contesti aziendali. Saranno esplorati in dettaglio i principi fondamentali, gli obiettivi specifici, la metodologia impiegata, nonché i relativi vantaggi e svantaggi. Inoltre, verranno presentate alcune piattaforme rilevanti che facilitano l'attuazione di questa forma di analisi.

3.1. Definizione e principi chiave

Negli anni recenti, l'espansione dell'uso di dispositivi mobili e l'adozione di approcci digitali da parte di un ampio pubblico, unitamente all'aumento dello scambio di contenuti visivi e audiovisivi su piattaforme dedicate, hanno richiesto, grazie anche all'evoluzione di reti internet sempre più affidabili e rapide, la necessità di sviluppare un nuovo approccio di ricerca etnografica.

La nascita della etnografia mobile è considerata dunque un'evoluzione della metodologia etnografica tradizionale, sviluppata appositamente per inseguire le mutevoli dinamiche della società attuale. Come afferma il nome stesso, questa metodologia qualitativa si avvale di tecnologie mobili per la raccolta di dati etnografici al fine di ottenere una

comprensione più profonda del contesto, dei bisogni, delle volontà e delle problematiche sui prodotti e servizi osservate dai consumatori.

Insomma, in linea con la tradizione dell'osservazione partecipante, i dispositivi mobili, come i telefoni cellulari, sono diventati un "estensione del sé nell'osservatore partecipante" negli studi etnografici mobili (Hein, 2011).

L'idea e la nascita di questa metodologia si è basata infatti nel riuscire a scavare nel profondo e ottenere più informazioni possibili sui consumatori, osservando e analizzando ciò che le persone fanno effettivamente nel loro contesto quotidiano, sfruttando l'immensa potenza, sempre più elevata, degli smartphone. È facilmente osservabile, nell'immagine allegata, come negli ultimi anni gli smartphone abbiano avuto un notevole incremento diventando così parte integrante della vita di ciascuno di noi, riuscendo anche a supportare molte delle azioni/compiti da svolgere. Un report pubblicato da Statista afferma inoltre che il tempo medio giornaliero di un utente passato sui social network è di circa due ore. Questi dati riescono facilmente a rendere l'idea del fatto che, al giorno d'oggi, quando un essere umano ha il desiderio di comunicare e condividere alcuni fatti della propria vita, il dispositivo mobile diventa la piattaforma di riferimento. Si tratta di un potente strumento che ora, anche grazie a questa nuova metodologia etnografica, le aziende o qualsiasi altro ente voglia effettuare indagini di mercato, possono sfruttare per ottenere maggiori profitti e maggiore soddisfazione da parte degli utenti finali.

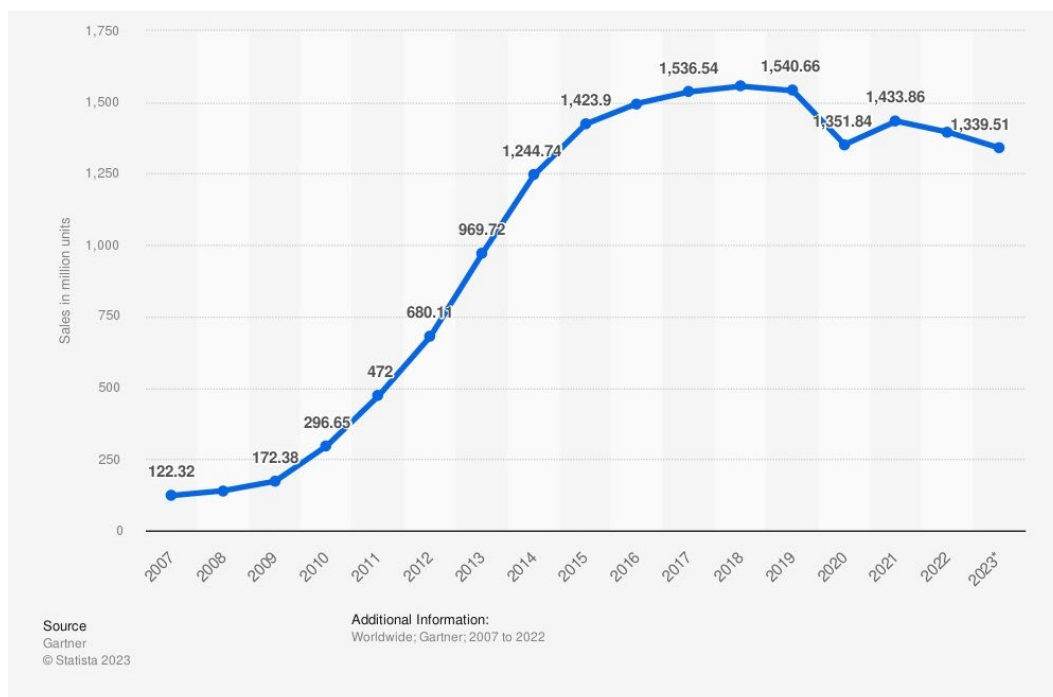


Figura 3.1-1 - Numero di vendite mondiali di smartphones dal 2007 al 2022 (Fonte: Statista.com)

Dunque, l'etnografia mobile, attraverso l'uso di applicazioni per smartphone e altri dispositivi affini (come tablet, laptop e PC), consente ai partecipanti di auto-documentare e condividere, in contesti specifici, le proprie esperienze, comportamenti e sentimenti in tempo reale riguardo a un determinato argomento di ricerca. Tutto questo tramite la condivisione di immagini, video, audio e testi.

Il progetto prevede il reclutamento di intervistati target, accuratamente selezionati secondo criteri specifici della ricerca. Questi individui saranno incaricati di documentare e condividere sottoforma di post le proprie azioni quotidiane su una piattaforma appositamente designata. Questo impegno, che durerà per un periodo di tempo prestabilito, consentirà loro di creare una sorta di diario multimediale (Bartlett e Milligan 2015). I ricercatori avendo poi accesso costante all'applicazione hanno la possibilità di visualizzare e analizzare tali post in tempo reale direttamente dal proprio ufficio o in qualsiasi posto si trovano, eliminando così la necessità di viaggiare fino alla località del partecipante. Ciò consente, oltre a un notevole risparmio di tempo e denaro, di seguire i soggetti di ricerca nel loro ambiente naturale, catturando le dinamiche della loro vita in modo più spontaneo, completo e approfondito.

Allo stesso modo della netnografia, un simile approccio si rivela non invasivo per gli intervistati. Infatti, agendo in autonomia, nel proprio ambiente di riferimento, il partecipante registra comportamenti e sentimenti con un alto grado di veridicità. Questo metodo minimizza perciò le potenziali distorsioni che potrebbero emergere a causa della presenza o dell'influenza del ricercatore, comunemente note come "effetto ricercatore". Distorsione comunemente presente invece nella tradizionale ricerca etnografica.

In alcuni casi, vi sono piattaforme che offrono la possibilità ai ricercatori di interagire con i partecipanti durante la ricerca sul campo tramite notifiche push, con l'obiettivo di fornire orientamenti continui, assegnare compiti o chiedere ulteriori dettagli sulle esperienze riportate. In questo caso si tratta di un costante "botta e risposta" tra ricercatore e partecipante, utile a fornire il giusto output più approfondito possibile.

I dati raccolti tramite l'etnografia mobile offrono dunque un quadro fedele e non alterato delle reali esperienze e percezioni del partecipante arricchendo così di molto il valore che verrà ottenuto dalla ricerca. A confermare questa teoria, Ideemo³⁰, società specializzata in etnografia mobile e in altri servizi di ricerca, ha pubblicato il feedback di alcuni ricercatori, i quali affermano che i comportamenti auto-registrati, catturati nel contesto familiare della vita quotidiana del partecipante, sono più intimi, autentici e rivelatori di qualsiasi altra tecnica di raccolta dati che i ricercatori abbiano utilizzato prima.

Insomma, il metodo dell'etnografia mobile risulta essere un valido successore della classica ricerca etnografica, poiché, essendo basata sulle recenti tecnologie e sui principi cardine del suo predecessore riesce a offrire un grado di qualità nei risultati senza paragoni. Tuttavia, vi è il rischio, da parte degli utilizzatori, di dimenticare le vecchie metodologie di raccolta dati, che invece in questo caso risultano comunque molto utili e, addirittura, rafforzano l'impatto delle recenti ricerche se affiancate ad esse. Mentre vengono acquisite informazioni in tempo reale, gli studiosi hanno infatti notato che i metodi tradizionali di raccolta dati etnografici, come l'intervista, sono utili come supporto. I dati aggiuntivi ricavati in modo tradizionale possono aiutare i ricercatori a comprendere e interpretare le conoscenze raccolte in tempo reale (quindi con la nuova metodologia).

³⁰ Azienda Irlandese specializzata nel fornire soluzioni software di ricerca video in-the-moment sfruttando tecnologie mobili. Il loro sito web offre una sezione di blog e risorse nella quale pubblicano metodologie e casi studio.

Ciò è stato provato in particolare nell'ambito del turismo: Ad esempio, Birenboim et al. (2015) hanno condotto un sondaggio prima di monitorare le attività in tempo reale. Shoval et al. (2018) hanno condotto sia un sondaggio tramite questionario che un'intervista semi strutturata all'inizio della raccolta dati. Un approccio simile è stato utilizzato anche per comprendere gli spostamenti e i viaggi dei residenti in una città.

L'etnografia mobile rappresenta così un metodo flessibile e adatto a diversi contesti. Segue un'esposizione dettagliata degli scenari applicativi potenziali.

Se l'obiettivo è conoscere meglio i propri clienti, tale metodologia attraverso gli studi diario³¹ di un periodo prefissato, è in grado di comprendere i modelli di consumo in evoluzione identificando anche le esigenze emergenti. Un aspetto fondamentale è lasciare che gli utenti si esprimano liberamente e naturalmente. Questo significa non imporre schemi predefiniti di risposta, ma piuttosto consentire l'interazione con i prodotti o i servizi secondo i loro tempi e modi. Così facendo questo approccio può portare a scoperte più autentiche e significative.

Le esigenze emergenti identificate durante l'applicazione dell'etnografia mobile, possono essere sfruttate da team specifici di progettazione e sviluppo prodotto per tradurre i risultati in caratteristiche concrete di prodotto. In questo caso il cliente si trova al centro del processo di sviluppo prodotto al fine di rendere il bene più utile, intuitivo e piacevole possibile per l'utente finale. Il tutto, sempre, senza essere fisicamente presenti con l'utente target.

L'etnografia mobile inoltre può essere sfruttata per eseguire dei test prodotto o concetto prodotto. Attraverso la richiesta ai partecipanti di documentare le loro impressioni sui nuovi concetti di prodotto o sull'uso di determinati prototipi, si è in grado di determinare in anticipo il possibile futuro successo di un progetto evitando così di intercorrere in inutili e dannosi fallimenti per il business. Così facendo si riesce a raggiungere maggiore agilità nei cicli di progettazione creativa evitando anche costosi test da parte degli utenti in laboratorio.

³¹ Durante lo studio diario i partecipanti sono invitati a documentare autonomamente le proprie esperienze quotidiane. Questo può includere la scrittura di note, la registrazione di voci, la cattura di immagini o video, o l'utilizzo di app specifiche per diario.

Un vantaggio ulteriore è costituito dalla possibilità di tracciare e comprendere al meglio il percorso di acquisto del cliente offrendo una consapevolezza più profonda del comportamento e delle interazioni dell'intervistato con il marchio attraverso diverse piattaforme. Negli ultimi anni quest'attività è diventata sempre più complessa e non lineare. I clienti, grazie all'avvento di internet, risultano avere maggiore consapevolezza di ciò che desiderano comprare e tendono sempre più a interagire con il marchio di interesse in modi diversi, ognuno dei quali tende, di conseguenza, a contribuire a creare l'esperienza complessiva con il marchio.

Con l'etnografia mobile i ricercatori hanno la possibilità di monitorare e registrare le esperienze dei clienti in tempo reale durante il loro processo di acquisto.

I risultati ottenuti forniscono di conseguenza importanti nozioni su tutte le fasi del processo: consapevolezza³², considerazione³³, decisione³⁴, conservazione³⁵ e advocacy.³⁶ Le aziende che attuano questa metodologia per scovare tali informazioni si trovano a detenere l'enorme potere e vantaggio di poter influenzare il consumatore nel suo processo d'acquisto. Per esempio, le aziende potrebbero concentrarsi su strategie come il marketing di influenza, la creazione di comunità intorno al marchio, e l'uso efficace di politiche di cashback e voucher per incoraggiare l'acquisto (Canvas8, 2020)³⁷. Inoltre, un'azienda dovrebbe pianificare e gestire attentamente ogni interazione tra il marchio e il cliente; Interazioni che possono avvenire in diverse fasi del percorso del cliente. L'obiettivo di questo approccio strategico è assicurarsi che ogni punto di contatto contribuisca positivamente all'esperienza del cliente con il marchio, rafforzando così la

³² Fase in cui un utente riconosce un bisogno o desiderio da soddisfare. Esso inizia dunque a cercare informazioni di base per una maggiore comprensione del bisogno e le prime possibili soluzioni. Vi è il primo contatto con i brand e per le aziende è importante posizionarsi come scelta ideale.

³³ Fase nella quale i clienti iniziano a valutare e confrontare le opzioni disponibili atte a soddisfare i loro bisogni. Posizionarsi efficacemente e influenzare positivamente le decisioni di acquisto è fondamentale per le aziende.

³⁴ Momento cruciale in cui il consumatore sceglie se procedere o meno con l'acquisto di un prodotto o servizio. In questa fase il cliente tende a pesare tutte le informazioni e le influenze per fare una scelta. Fase fondamentale per le aziende che devono ottimizzare le proprie strategie di marketing per convertire l'interesse dell'utente in acquisto.

³⁵ La fase di conservazione, nota anche come fase di fidelizzazione, si riferisce al mantenimento di un rapporto a lungo termine tra il cliente e il brand dopo che il cliente ha effettuato un acquisto.

³⁶ La fase di advocacy si verifica dopo che un cliente ha completato un acquisto. Se il cliente è soddisfatto, può diventare un "advocate" del brand, condividendo la sua esperienza positiva con amici, familiari o attraverso i social media e influenzando di conseguenza le decisioni di acquisto di altri potenziali clienti.

³⁷ Canvas8 è una società con sede a Londra che studia gli insight comportamentali aiutando le organizzazioni a comprendere le persone

loro fiducia e la loro relazione con l'azienda. Questo può includere aspetti come la qualità del servizio clienti, l'efficacia della pubblicità e la facilità di utilizzo del sito web.

3.2. Metodologia

Un ricercatore che voglia mettere in atto una corretta applicazione di tale etnografia tramite dispositivi mobili si trova ad affrontare, prima della raccolta e della successiva analisi dei dati, diverse importanti pianificazioni e decisioni, fondamentali per l'ottenimento di risultati significativi.

Di primaria importanza è l'identificazione e la definizione dell'obiettivo della ricerca, elemento che costituisce la base per la costruzione dell'intero processo. Comprendere al meglio lo scopo della ricerca è cruciale per capire il contesto specifico dove operare e immergersi così nel corretto ambiente di riferimento. Tenendo a mente l'obiettivo, il ricercatore deve poi focalizzarsi nell'identificazione dei clienti target da selezionare.

La selezione degli utenti, considerata la parte più critica di pianificazione di una ricerca etnografica mobile, avviene tramite tecniche di campionamento³⁸; Diverse sono le tecniche, sia probabilistiche che non, che i ricercatori possono mettere in atto per la selezione del campione. Alcune di queste elencate di seguito:

- **Campionamento Casuale:** Adottando questa tecnica ogni individuo della popolazione ha uguali probabilità di essere scelto per il campione. Assegnando un numero a ciascuna persona l'estrazione avviene in modo casuale senza fare riferimento a determinate caratteristiche. Questo metodo è utile per evitare bias nella selezione e per ottenere un campione rappresentativo dell'intera popolazione.

³⁸ Il campionamento è un approccio usato per estrarre un sottoinsieme di individui (campione) da un gruppo più grande (popolazione) costituenti l'oggetto in studio, per partecipare a uno studio. Questo sottoinsieme, selezionato attraverso differenti tecniche, ha l'obiettivo di rappresentare in modo accurato il gruppo di riferimento.

- Campionamento Stratificato: In questo caso l'intera popolazione viene divisa in sottogruppi o strati in base a determinate caratteristiche (p.e. genere, età, ecc.). I partecipanti alla ricerca vengono selezionati estraendo un campione casuale in modo proporzionale da ciascuno strato e il campione finale sarà caratterizzato dall'unione dei singoli campioni. Metodo che risulta utile quando i ricercatori vogliono assicurarsi che determinati segmenti della popolazione siano adeguatamente rappresentati nel campione.
- Campionamento a Scelta ragionata: In questo campionamento non probabilistico i partecipanti vengono selezionati sulla base di alcune loro caratteristiche ritenute importanti per la ricerca.
- Campionamento per Convenienza: La selezione ricade sui partecipanti che sono più facilmente accessibili per i ricercatori. Anche se questo metodo è il meno costoso e il più semplice in termini di logistica, può portare a un campione meno rappresentativo e, quindi, a risultati meno generalizzabili.

Anche se scontato, è strettamente fondamentale reclutare utenti tecnologicamente preparati in grado di utilizzare programmi software e strumenti dei propri dispositivi mobili e soprattutto disposti a fornire dati di elevata qualità (Cribbett, 2015).

A supportare il processo di selezione può tornare utile anche l'adozione di agenzie esterne specializzate in grado, grazie alle loro esperienze e risorse, di fornire individui che soddisfano i criteri specifici richiesti per lo studio.

Con il metodo della etnografia mobile i partecipanti soggetti allo studio ottengono di conseguenza il ruolo di co-ricercatori nella raccolta dati e quindi il processo di selezione diventa automaticamente estremamente importante; Tutti i dati che si ottengono dipendono dalle loro abilità di documentare, eseguire istruzioni e fornire una visione il più reale possibile del contesto richiesto.

In aggiunta, per incoraggiare gli utenti a partecipare con entusiasmo e ad impegnarsi attivamente nel progetto, i ricercatori devono prevedere di offrire incentivi, necessari anche per il riconoscimento del tempo e sforzo dedicato. Necessario è anche il bisogno di stabilire un rapporto di fiducia con i partecipanti dimostrando il valore e l'apprezzamento per i loro contenuti al fine di incrementare il loro coinvolgimento nella ricerca. Importante sottolineare che "la qualità della relazione tra ricercatore e partecipanti è uno dei fattori

più importanti che determinano il successo di una ricerca sociale” (Babbie, 2010). Dunque, un approccio rispettoso e inclusivo può fare una grande differenza nel coinvolgimento e nella qualità dei risultati ottenuti.

Una volta completato il processo di selezione, prima di procedere con gli ulteriori step di pianificazione, è strettamente necessario per i ricercatori considerare e verificare gli aspetti legali dei partecipanti. Infatti, è necessario assicurarsi la possibilità di procedere alla successiva fase di raccolta di dati personali. Come prima cosa è fondamentale quindi verificare le leggi locali o dell’istituzione in analisi sulla privacy e le possibili restrizioni legali nello scattare foto o video. Alcuni paesi, infatti, proteggono le persone da riprese non autorizzate in luoghi pubblici o privati. Si procede successivamente nel verificare l’ottenimento di un consenso e/o di un accordo di riservatezza per i partecipanti. Richiedere il consenso nell’ utilizzare le immagini e video forniti includendo, nella richiesta, anche i dettagli di come i dati verranno utilizzati e conservati è estremamente necessario. In alcuni casi, potrebbe essere appropriato stabilire un accordo di riservatezza, specialmente se le informazioni raccolte sono sensibili o confidenziali. In aggiunta, per proteggere la privacy, è possibile procedere con alcune limitazioni mirate al caso specifico, come per esempio la disabilitazione del sistema GPS o l’anonimato attraverso pseudonimi. Quest’ultima è l’operazione che è stata eseguita nel caso dello “Studio delle pratiche quotidiane di lettura cartacea e digitale delle ragazze adolescenti” (Loh, Sun and Lim, 2023), nel quale venivano sottoposte alla ricerca ragazze minorenni. Attraverso, quindi, un accordo stipulato con i genitori è stato possibile procedere con la ricerca.

È dunque possibile affermare che ogni scelta di protezione dati e privacy dipende strettamente dal contesto geografico e sociale in cui si opera e dalle condizioni richieste dai partecipanti.

Al termine della fase di selezione partecipanti, si procederà alla preparazione del progetto. Ciò comporta, da una parte l’esigenza di procedere alla selezione accurata di un software da personalizzare, che verrà poi implementato sui dispositivi dei partecipanti dall’altra, l’elaborazione di domande o compiti specifici destinati a essere presentati agli utenti.

La selezione appropriata di un software si basa sulle specifiche esigenze della ricerca in questione e sulle caratteristiche degli utenti coinvolti. Numerose soluzioni software sono disponibili sul mercato, offrendo una varietà di funzionalità. Queste possono includere

funzioni comuni quali la capacità di caricare foto e video, la possibilità di interagire e funzioni più avanzate come il tracciamento GPS. Un criterio fondamentale nella scelta del software è la sua usabilità. È imprescindibile che il software sia intuitivo e di facile impiego sia per i ricercatori che per i partecipanti. Una interfaccia utente ben progettata e un'esperienza di navigazione fluida e comprensibile sono cruciali per assicurare il coinvolgimento attivo e continuativo dei partecipanti durante il processo di raccolta dati.

I partecipanti vengono così invitati e tramite un piccolo briefing viene spiegato loro l'obiettivo del progetto, i compiti da svolgere e l'incentivo che riceveranno. Prima dell'inizio della raccolta dati può essere utile per i ricercatori intervistare i partecipanti per conoscere le loro esperienze e aspettative riguardo al progetto. Conoscere in anticipo queste informazioni consente di ottenere una formulazione adeguata della domanda da effettuare. I compiti da assegnare devono essere brevi e mirati al fine di aumentare le possibilità di partecipazione. Ogni richiesta ai partecipanti deve essere pensata attentamente, inserendo più dettagli possibili nella domanda così da evitare la perdita di sfumature o intuizioni nelle risposte. Carey Palmer³⁹ di Dscout, consiglia inoltre di inserire anche domande a risposta multipla così da avere già sottomano il materiale per l'analisi. Inoltre, l'esecuzione di un piccolo progetto pilota è molto utile per verificare se le istruzioni sono chiare e se i dati raccolti sono effettivamente utili per l'obiettivo di ricerca.

Conclusa la progettazione della ricerca e invitato i partecipanti, la raccolta dei dati, tramite la pubblicazione di foto, video, risposte a domande specifiche e tramite la geolocalizzazione dei movimenti degli utenti, può iniziare. Il processo di raccolta dati si basa spesso sull'azione del "diario". In questo contesto, i partecipanti sono invitati a condividere le loro giornate attraverso l'applicazione dedicata alla ricerca. In aggiunta alla documentazione giornaliera, i partecipanti possono anche essere invitati a svolgere determinati compiti specifici richiesti dai ricercatori. Per esempio, le applicazioni mobili possono essere utilizzate per somministrare sondaggi e questionari in modo efficiente aggiungendo così anche una dimensione strutturata alla raccolta dei dati. Questo consente di raccogliere rapidamente feedback e opinioni dai partecipanti. Uno degli aspetti chiave

³⁹ Carey Palmer è la responsabile della ricerca di Dscout. Dscout è una piattaforma di etnografia mobile in grado di consentire la raccolta di dati in tempo reale.

della etnografia mobile è la capacità di interagire con gli utenti in tempo reale grazie all'invio di notifiche push. Tali interazioni aggiuntive possono tornare utili come promemoria e/o solleciti o per la richiesta di approfondimenti riguardo ai dati condivisi. Infatti come afferma anche l'agenzia Angelfish Fieldwork⁴⁰ “è difficile sapere esattamente cosa potrebbe fare un partecipante se lasciato a se stesso. Alcuni potrebbero non seguire esattamente le loro istruzioni, concentrarsi sulle cose sbagliate o commettere errori nella loro documentazione.”

Questa fusione di approcci, combinando sia tecniche di diario che di interazione supplementare consente di ottenere una maggiore completezza e profondità nell'ambito della ricerca condotta.

Ponendo l'accento sul principio “mostra, non raccontare”, si richiede ai partecipanti di documentare il loro comportamento, le azioni intraprese e le emozioni provate in situazioni reali durante lo svolgimento di compiti o attività specifiche. Quest'ultime meticolosamente progettate per garantire che i partecipanti agiscano il più autenticamente possibile.

La durata dello studio è tendenzialmente di breve durata se comparata a una classica analisi etnografica o ad uno studio netnografico e questa può variare da qualche giorno fino ad una settimana. È importante, infatti, limitare lo stress degli intervistati. Inoltre, è sufficiente e importante avere una panoramica quotidiana reale di quello che l'utente svolge con lo strumento/oggetto/servizio in analisi. All'utente viene infatti richiesto di non alterare i suoi comportamenti naturali al fine di evitare una visione distorta della realtà. In aggiunta come afferma la piattaforma Further⁴¹ nell'articolo “An Essential Guide to Mobile Ethnography”, viene anche richiesto di condividere i momenti di forte emozione come la frustrazione o delusione in quanto questi tendono ad offrire migliori opportunità di innovazione.

⁴⁰ Angelfish Fieldwork è un'agenzia Inglese specializzata nel reclutamento di partecipanti per ricerche di mercato e studi qualitativi. La loro attività principale è quella di identificare e selezionare individui che si adattino ai criteri specifici di un determinato studio o progetto di ricerca.

⁴¹ Further è una piattaforma digitale specializzata nella ricerca di mercato qualitativa che offre strumenti avanzati per la gestione di comunità online e studi etnografici. Offre servizi professionali che promettono rapidità ed efficienza, sostenuti da un team di esperti del settore e accademici, per fornire analisi approfondite e favorire decisioni strategiche basate sui dati.

Le richieste specifiche vengono poi presentate in modo graduale e non massivamente per prevenire stress non necessario e per evitare di scoraggiare gli utenti nella condivisione di contenuti.

In conclusione, è utile stabilire una scadenza definitiva per i partecipanti. Questo approccio garantisce una comprensione del lasso di tempo previsto e comunica chiaramente quando terminerà la raccolta dei dati. Impostare una scadenza specifica comporta anche un senso di urgenza e impegno tra i partecipanti, portando a processi di raccolta dati più efficienti e organizzati.

A supporto delle informazioni raccolte esclusivamente dagli intervistati, i ricercatori spesso utilizzano le cosiddette interviste di follow-up per generare discussioni con gli utenti sui dati da loro condivisi al fine di ottenere una comprensione più profonda sull'argomento e sui comportamenti dei partecipanti. Si tratta di vere e proprie interviste basate sullo stesso principio delle interviste di foto-elicitazione utilizzate anche dalla ricercatrice Analía Inés Meo⁴². La foto-elicitazione si riferisce all'uso di una singola o insieme di fotografie come stimolo durante un'intervista di ricerca. Ha l'obiettivo di suscitare risposte e ricordi e svelare le opinioni, le credenze e i significati dei partecipanti o di indagare sulle dinamiche di gruppo (Hurworth, 2003).

L'etnografia mobile per sua natura offre meno controllo sull'ambiente di ricerca rispetto ai metodi etnografici tradizionali, in quanto basa i suoi risultati sui contesti naturali dei partecipanti. Questo a volte può portare a una mancanza completa di contesto nei comportamenti osservati e nelle azioni e decisioni prese dai partecipanti. Le interviste di follow-up vengono quindi effettuate per colmare queste lacune.

Ad esempio, MDRG⁴³ nella sua ricerca etnografia mobile ha svolto uno studio sul consumo di caffè. Come supporto a tale ricerca ha condotto interviste telefoniche di follow-up sia prima che dopo l'applicazione etnografica. Le interviste iniziali sono state

⁴² Professore ordinario incaricato della materia Sociologia dell'educazione presso la Facoltà di scienze Sociali dell'UBA. Meo nel 2010 ha scritto l'articolo "Picturing Students' Habitus: The Advantages and Limitations of Photo-Elicitation Interviewing in a Qualitative Study in the City of Buenos Aires", nella quale esamina i vantaggi e svantaggi delle interviste di foto-elicitazione nell'ambito delle scuole secondarie di Buenos Aires.

⁴³ MDRG è una delle migliori società, con sede a New Orleans, di ricerche di mercato specializzata nella scoperta di approfondimenti per clienti di vari settori e dimensioni.

utilizzate per trarre le prime scoperte e coinvolgere i partecipanti. Dopo che i partecipanti avevano condiviso le loro esperienze di acquisto e consumo di caffè attraverso video, foto e testi tramite un'applicazione mobile, è stato condotto un secondo set di interviste. Queste interviste di follow-up hanno permesso ai ricercatori di approfondire i comportamenti e le abitudini osservate durante il periodo di etnografia mobile, fornendo una comprensione più ricca e sfumata dell'argomento.

In sintesi, le interviste di follow-up nell'etnografia mobile svolgono un ruolo vitale nel completare i dati raccolti tramite dispositivi mobili, affrontando così anche i limiti dell'etnografia mobile.

Dopo la raccolta, i dati raccolti attraverso l'etnografia mobile richiedono un'analisi approfondita e un'interpretazione attenta per la trasformazione in intuizioni significative.

In questa fase, i ricercatori effettuano una triangolazione tra le informazioni raccolte mediante l'azione di diario con i dispositivi mobili e le più classiche interviste di "follow-up" o questionari. Nella raccolta dati tutto viene registrato nell'applicazione o tramite le note sul campo durante le interviste e i ricercatori grazie alla tecnologia hanno la possibilità di iniziare ad analizzare i dati caricati in tempo reale. Infatti, alcune applicazioni offrono la funzione di analisi dei dati integrata, offrendo così ai ricercatori la possibilità di eseguire l'analisi preliminare direttamente dal dispositivo. Strumenti software e di intelligenza artificiale forniscono dunque anche in questa metodologia un supporto non indifferente.

È infatti possibile affermare che, anche se i due metodi descritti, la netnografia e la etnografia mobile, si differenziano molto nelle procedure operative di raccolta dati e della natura di essi, offrendo vantaggi e limiti differenti, risultano simili nell'operazione di analisi e interpretazione dei dati. Entrambi, infatti, necessitano di un'analisi qualitativa come la codifica, l'analisi tematica e l'analisi semantica⁴⁴ e di un'analisi quantitativa per i dati derivanti dai questionari o dalle informazioni di geolocalizzazione.

Insomma, strumenti manuali o di IA utilizzati in precedenza nella netnografia per l'analisi sono spesso i medesimi che vengono presi in considerazione per dedurre importanti

⁴⁴ L'analisi semantica è fondamentale nell'etnografia mobile per interpretare accuratamente i testi raccolti come interviste e note sul campo fornendo una comprensione più profonda del contesto, delle interazioni, delle percezioni e delle esperienze dei partecipanti.

insight e per supportare il ricercatore nella etnografia mobile. Tuttavia, come già descritto in precedenza quest'ultima metodologia fonda i suoi principi nella condivisione di dati visivi, audio, GPS e audiovisivi oltre che testuali, raccolti in ambienti fisici. È quindi necessario l'utilizzo di strumenti/algoritmi dedicati maggiormente ad un'analisi adatta a questa tipologia.

Quest'esigenza è efficacemente supportata da software avanzati, spesso integrati direttamente nelle applicazioni mobili, che facilitano processi come la codifica, l'ordinamento⁴⁵, il filtraggio⁴⁶, la trascrizione di audio e l'analisi delle immagini⁴⁷. È necessario precisare però che, nonostante l'avanzamento tecnologico di questi sistemi, è l'analisi umana a rappresentare il fulcro per una comprensione completa e autentica del significato in quanto gli aspetti culturali, sociali e contestuali possono essere pienamente interpretati solo attraverso la lente dell'esperienza umana. Questa capacità unica del cervello umano è in grado di fornire insight preziosi per lo sviluppo di strategie volte al miglioramento e all'innovazione aziendale.

3.3. Punti di forza e limiti della etnografia mobile

Terminata l'analisi sulla metodologia dell'etnografia mobile, è bene soffermarsi adesso sui possibili punti di forza e di debolezza di questo nuovo strumento.

Partendo dai punti di forza, il primo da ricordare è la possibilità, grazie alla sempre più diffusione dei dispositivi mobili con funzionalità avanzate, di considerare ed analizzare contemporaneamente un elevato campione di individui posizionati in qualsiasi luogo del

⁴⁵ L'ordinamento permette di organizzare i dati in un ordine specifico (ad esempio, alfabeticamente, per data, per frequenza, ecc.)

⁴⁶ Attraverso il filtraggio, i ricercatori possono isolare dati che sono particolarmente pertinenti a specifiche domande di ricerca o temi. Questo è utile per esaminare in dettaglio determinate aree di interesse senza essere sopraffatti dall'intera massa di dati raccolti.

⁴⁷ Attraverso l'uso di algoritmi di IA, nello specifico usando le reti neurali convoluzionali combinate con le reti neurali ricorrenti è possibile ottenere il riconoscimento di oggetti, il tracciamento del movimento e l'analisi delle espressioni facciali nelle immagini e nei video.

mondo, il tutto ad un costo estremamente ridotto (Bosio, 2017). Infatti, l'osservazione virtuale permette ai ricercatori di svolgere il proprio lavoro direttamente dal proprio ufficio evitando spese inutili e perdite di tempo a non a valore nel viaggio. È dunque possibile affermare che tutto il tempo che il ricercatore impiega nella ricerca, dalla progettazione, alla raccolta fino all'analisi è considerato a valore.

I dati vengono registrati sull'applicazione e condivisi in tempo reale dagli intervistati e i ricercatori possono così osservare, analizzare e condividere tali informazioni direttamente con gli stakeholders senza aspettare la fine del progetto. Un altro vantaggio che deriva dalla condivisione in tempo reale dei contenuti è il fatto di ricordarsi tutti i dettagli dell'esperienza appena vissuta. Infatti, uno dei vantaggi di usare il proprio dispositivo mobile (p.e. lo smartphone) come strumento di registrazione dati è che gli intervistati riescono a fornire informazioni in qualsiasi momento della propria giornata. Hanno la possibilità di scattare foto o registrare audio non appena provano un'emozione su un evento specifico. Segnalando e condividendo in tempo reale hanno modo di accedere completamente a tutti i dettagli possibili che altrimenti potrebbero risultare sfocati se documentati su un periodo di riferimento più ampio. Il bias del ricordo viene così minimizzato (Schwarz, 2007) in quanto i partecipanti perdono la necessità di recuperare le informazioni dopo un tempo relativamente lungo dalla loro memoria.

I dispositivi mobili, oltre a ricoprire un ruolo centrale nella raccolta dati, consentono una sorta di barriera protettiva contro la presenza del ricercatore. Si tratta infatti di una metodologia di natura non intrusiva nel quale le azioni degli utenti non vengono influenzate dal comportamento dei ricercatori. Inoltre, nello studio condotto da Hein e altri nel 2011, viene condiviso che l'uso di un dispositivo ormai integrato nell'ambiente naturale del ricercatore come il telefono cellulare, risulta essere meno invasivo rispetto ad altri dispositivi estranei ed innaturali da utilizzare. Si riesce così ad ottenere risultati più rappresentativi del comportamento, garantendo la scoperta di insight più genuini e accurati. La validità della ricerca viene inoltre garantita dalle possibili triangolazioni che possono essere effettuate durante l'esecuzione di questa ricerca etnografica mobile:

- Triangolazione dell'investigatore: Diversi ricercatori si impegnano ad analizzare lo stesso set di dati disponibili. In questo modo si confrontano i differenti risultati per ottenere una maggiore qualità e affidabilità.

- Triangolazione del metodo: Si combina l'etnografia mobile con altri metodi di analisi come sondaggi e interviste.
- Triangolazione dei dati attraverso diversi media catturati con il telefono cellulare come testo, immagini, video e valutazioni quantitative.

L'uso di dispositivi mobili per la raccolta dati richiede inoltre minore sforzo e minori competenze necessarie per i partecipanti (Zhang et al., 2021). Gli sforzi richiesti per la raccolta dati si limitano infatti al solo fornire il consenso e a condividere i dati. Inoltre, la presenza di applicazioni sempre più complesse e piene di funzionalità permette ai partecipanti di ridurre le proprie competenze linguistiche necessarie, e rispetto alla stesura di un diario di viaggio testuale, scattare foto e video risulta essere un compito meno impegnativo e meno dispendioso in termini di tempo.

Un ulteriore vantaggio è che i diversi tipi di dati (foto, audio posizione, etc.) raccolti contemporaneamente dallo stesso partecipante consentono ai ricercatori di sviluppare maggiori insight rispetto a quanto si potrebbe fare se gli stessi dati venissero raccolti in studi differenti. Un esempio concreto di questo vantaggio è stato fornito dallo studio di Birenboin et al. (2015) i quali combinando la posizione di alcuni turisti con la loro condivisione istantanea dell'esperienza, sono riusciti a creare una mappa di uno zoo nella quale sono state identificate le aree più attraenti e quelle meno.

Aggiuntivo punto di forza della etnografia mobile, di matrice prettamente psicologica, è il coinvolgimento degli utenti che si crea nella raccolta dei dati. Il partecipante, infatti, utilizzando i propri dispositivi digitali per contribuire alla ricerca, assume il ruolo di "co-ricercatore". Questo senza dubbio riveste di una maggiore autorità e libertà il soggetto, incrementando notevolmente il livello di coinvolgimento e la qualità dell'esperienza vissuta, oltre che l'attendibilità dello studio. Infatti, sentendosi parte attiva del progetto di ricerca, la persona sarà stimolata a offrire indicazioni molto più precise e corrette. Inoltre, in alcune casistiche, come nel campo del turismo, alcune persone sono disposte a partecipare al progetto di ricerca, anche senza incentivi. Sono infatti realmente entusiasti di condividere le proprie esperienze con l'azienda, con l'obiettivo di contribuire al miglioramento del servizio. Spesso manifestano anche un vivo interesse nel testare il metodo, considerando la partecipazione un'esperienza piacevole e stimolante (Stickdorn e Frischhut, 2012).

Infine, il valore dell'etnografia mobile deriva anche dalla capacità di produrre risultati sempre più vicini alla naturalezza e autenticità del comportamento umano. Negli ultimi cinque anni circa, infatti, grazie all'aumento della messaggistica video i partecipanti si sono trovati a esprimere le loro opinioni ed emozioni tramite i video-selfie. Questa modalità di comunicazione si traduce in un enorme vantaggio in quanto lo strumento audiovisivo limita la capacità di autocurare il contenuto, al contrario dei testi scritti che la alimentano, incrementando così la spontaneità nei dati caricati sull'applicazione. Tuttavia, non tutti i video risultano utili a questo fine, infatti sono preferibili clip brevi e non invece più estese, riassuntive della giornata. In questo modo, infatti, aumenta la possibilità di ottenere risposte "sul momento" dell'esperienza, mantenendo così elevata l'emozione appena vissuta. Inoltre, durante l'analisi dei video gli esperti hanno la possibilità di identificare elementi apparentemente nascosti, non altrimenti identificabili, come lo sguardo o il tono di voce sviluppando così importanti insight.

Accanto a questo, per rigore scientifico, è necessario considerare anche alcuni limiti che ne derivano.

Il primo che verrebbe in mente a qualsiasi ricercatore è il ruolo chiave che assume l'intervistato nel processo di ricerca. Infatti, la raccolta dati e la qualità dei risultati che si ottengono dipendono in larga scala dall'impegno e dalle capacità che i partecipanti presentano, dall'inizio alla fine. Poiché sono solo i partecipanti ad avere il potere di raccogliere i dati, i ricercatori si devono impegnare nel fornire una guida chiara e concisa seguita da una corretta formazione di come procedere e utilizzare il software nel modo più adeguato. Il successo della formazione e la ben riuscita del progetto può essere facilitata notevolmente dalla facilità d'uso dello strumento di raccolta dati. Tuttavia, anche se la formazione e le istruzioni risultano essere perfettamente eseguiti, i partecipanti potrebbero non sempre rispettare le regole e le tempistiche fornendo così disagi sia nella raccolta ma anche nella successiva analisi. Per esempio, nel rispondere a un sondaggio con trigger di posizione, i partecipanti, a causa di eventi inaspettati che potrebbero verificarsi nelle singole casistiche, potrebbero accumulare esperienze/emozioni e registrare successivamente più domande nello stesso momento e nello stesso luogo falsando così la qualità dei risultati. In tutto questo il ricercatore non ha possibilità di intervenire se non grazie a istruzioni e incentivi adeguati.

Collegato a questo fattore vi è il rischio che, nonostante le possibili triangolazioni con le interviste di follow-up, i ricercatori possano mancare di chiarimenti o indagini approfondite a causa dell'assenza fisica durante la raccolta dei dati, come avviene invece nella ricerca etnografica tradizionale.

Un'ulteriore sfida che può emergere nella raccolta di dati in tempo reale è la potenziale alterazione dei comportamenti degli utenti durante l'esecuzione delle loro attività quotidiane. La buona riuscita della raccolta non può essere pienamente assicurata anche se durante il briefing viene più volte ripetuto di evitare l'alterazione delle proprie azioni. Risulta infatti difficile trovare un giusto equilibrio tra l'intrusione nell'esperienza dei consumatori e l'efficacia della raccolta dati (Zhang et. al., 2021).

Per quanto riguarda invece l'analisi dei dati un limite che si verifica è il fatto che i partecipanti hanno la libertà di condividere dati multimediali nel formato che essi preferiscono (audio, foto, video). Questa eterogeneità tra dati raccolti da individui differenti può dunque causare una sfida per il processo di analisi. A fronte di ciò, Bosio et al. (2017) nell'elaborato "Customer experience research with mobile ethnography: a case study of the alpine destination serfaus-fiss-ladis" suggeriscono di includere una domanda di valutazione su scala Likert per misurare, al di là della tipologia di contenuto condiviso, la positività o meno dell'esperienza.

Infine, come già descritto in precedenza vi è la necessità di proteggere la privacy degli utenti sottoposti all'analisi. Questo può essere risolto tramite il consenso informato dei partecipanti e in alcuni casi tramite accordi di riservatezza. Il processo implica informare i partecipanti sulla natura dello studio, su cosa comporterà la loro partecipazione, sui potenziali rischi e benefici, e su come saranno gestiti e protetti i loro dati. Successivamente a queste informazioni i partecipanti decidono se dare o meno il loro consenso esplicito per partecipare allo studio.

Si procede ora all'introduzione di una problematica che coinvolge non tanto la metodologia della etnografia mobile, bensì i partecipanti del progetto di ricerca. Nello specifico si tratta del tema dell'intelligenza artificiale combinato alla proattiva partecipazione del consumatore nel fornire dati che nutriranno la miniera in cui la IA pescherà.

Grazie all'etnografia mobile, il consumatore si mette in gioco, esprimendosi e facendosi presente, condividendo dati. Tuttavia, spesso non è del tutto consapevole della modalità in cui quelle informazioni potranno venire usate anche nei suoi stessi confronti. Ottenuto il consenso informato dai partecipanti, questo dovrebbe specificare come saranno utilizzati e conservati i dati. In mancanza di una clausola che richiede l'eliminazione dei dati dopo un certo periodo o dopo il completamento dello studio, i dati possono essere impiegati dalle aziende per altri scopi.

Non può essere, infatti, sottovalutato il tema della persuasione, che può sfociare, quando gli strumenti sono molto sofisticati e forti, in quello della coartazione della volontà.

Il "nodo nevralgico", ovvero il punto che resta scoperto e che può dare adito a valutazioni difensive da parte dei consumatori viene ben esplicitato in un testo scritto a quattro mani da Kai-Fu Lee, a lungo presidente di Google China, dirigente di Microsoft e in Apple, attualmente co-presidente del consiglio per l'intelligenza artificiale del World Economic forum, e Chen Quifan⁴⁸, tra i più noti narratori cinesi di fantascienza, con il titolo "AI 2041, scenari dal futuro dell'intelligenza artificiale".

Nel primo racconto-saggio, "L'elefante d'oro", vengono esplicitate le conseguenze che possono determinarsi per il consumatore quando le basi della IA si associano a quello del Deep Learning.

Come spiegano gli autori: «se addestrato su un grande ammontare di dati, il Deep Learning può essere personalizzato sul singolo utente, in base ai pattern di quell'utente e di altri pattern simili individuati in altri utenti. Per esempio, quando visiti Amazon, la IA del sito evidenzia specifici prodotti allo scopo di attirarti e massimizzare la tua spesa. Quando apri una pagina Facebook, ti vengono mostrati i contenuti progettati per massimizzare il numero dei minuti che trascorrerai su Facebook. Le IA di Amazon e Facebook sono in grado di targhettizzare, vale a dire che mostrano a ogni singola persona dei contenuti personalizzati. Quindi, il contenuto che mi viene mostrato ha un grande effetto su di me, ma potrebbe non averne su di te. Questa accuratezza è molto più efficace

⁴⁸ Kai-Fu Lee, Chen Quifan, AI 2041, scenari dal futuro dell'intelligenza artificiale, Luiss UniversityPress

a produrre clic e acquisti di un unico approccio valido per tutti, come quelli utilizzati dai siti web statici tradizionali»⁴⁹.

Nel caso della etnografia mobile, dunque, l'offrire la propria compartecipazione, condividere l'esperienza, aprirsi ad una sorta di "dialogo" con la macchina che funge da intermediario su desideri e ricerche, consentono di rendere noti, se consentito dall'utente, i propri dati all'interno di un database per molto tempo. Tali dati possono però, in certi casi, venire utilizzati come fonte di addestramento per determinati algoritmi di apprendimento automatico o profondo. Il tutto per poter utilizzare quegli stessi dati nei confronti dell'utente.

A seconda della filosofia che sta dietro l'azienda che li impiega e quindi che programma la IA, questo know-how potrà essere sfruttato non necessariamente solo a favore, bensì anche generando un insieme di conseguenze negative, in quanto quei dati sono suscettibili di essere sfruttati in qualche modo per condizionare il cliente. Il tema del condizionamento è estremamente delicato, perché i confini tra persuasione e forzatura a volte sono labili, non definiti.

In concreto, le informazioni che il cliente ha volontariamente ceduto alla azienda si trasformano di fatto in "armi" di pressione, al fine di condurre il cliente a compiere ciò che è nell'intenzione, negli obiettivi della azienda, ma non nei propri. Il risultato, dunque, è che il cliente si trova così vessato dal sistema a cui ha volontariamente collaborato.

Naturalmente si tratta di un racconto che esaspera i toni di quanto ci troviamo ad affrontare nella realtà commerciale oggi, ma mette giustamente in allerta su quali e quante storture un sistema di raccolta di dati come quello dell'etnografia mobile potrebbe cagionare.

Dunque, a fronte di un'incentivazione del sistema dell'etnografia mobile che per le aziende è importante, funzionale, rapido, così come di soddisfazione per il consumatore che si sente parte del processo, resta il tema degli obiettivi complessi, che hanno poi a che fare con gli scopi ultimi che le aziende si pongono quando utilizzano queste strade.

⁴⁹ Ivi, p.44

3.4. Scelta del provider per l'applicazione del metodo

Aziende private che intendono condurre un progetto di ricerca di etnografia mobile per adattarsi all'era digitale, devono fare riferimento a provider esterni, i quali si occupano di gestire al meglio il lavoro fornendo piattaforme adeguate alla raccolta di dati. Gli intervistati utilizzeranno i loro propri dispositivi digitali come smartphone e laptop per raccogliere dati. Tuttavia, il dispositivo risulta solo uno strumento; la vera raccolta avviene grazie all'applicazione installata o al software che si intende utilizzare. Queste piattaforme facilitano l'interazione tra il ricercatore e l'intervistato, consentendo di condurre interviste, sondaggi e altre forme di raccolta dati in modo efficiente e in tempo reale.

Risulta però spontanea la domanda: Chi svolge il ruolo di ricercatore? Il provider esterno oppure l'azienda che richiede la ricerca?

Non esiste una regola specifica da adottare, ma tutto dipende dal contesto specifico che varia da caso a caso. Tale ruolo può infatti essere assunto da entrambe le parti.

Nello specifico sarà il provider a fornire, oltre alla tecnologia, un team di ricercatori partecipando attivamente alla ricerca, se direttamente coinvolto nello sviluppo della metodologia, nell'analisi dati e nella fornitura di insight. Per assolvere a questo ruolo in modo efficace, è fondamentale che il fornitore esterno di servizi mantenga una comunicazione stretta con l'azienda cliente, acquisendo una comprensione approfondita degli obiettivi del progetto e delle caratteristiche dei potenziali intervistati. La delegazione del gruppo di ricerca è spesso dovuta alla necessità di colmare una mancanza di risorse o competenze specifiche necessarie allo svolgimento di tale ruolo all'interno dell'azienda.

In altre casistiche invece il ricercatore sarà interno alla stessa azienda che ha richiesto la ricerca. In questo contesto, il provider del software svolgerà il solo ruolo di supporto fornendo gli strumenti necessari per la raccolta dati. Questo accade se l'azienda ha già, al suo interno, un team specializzato in ricerca di mercato o se preferisce appoggiarsi a ricercatori esterni.

È quindi possibile affermare che la scelta volta a identificare il soggetto cui assegnare il ruolo di ricercatore dipende dalle competenze, dalle risorse disponibili e dagli obiettivi specifici della ricerca.

Essendo presenti diverse aziende fornitrici di propri strumenti in grado di svolgere la raccolta dati in tempo reale, risulta fondamentale per il richiedente della ricerca focalizzarsi su quale dei software esistenti condurre il progetto. Come anticipato precedentemente la scelta del software è una fase chiave della progettazione di ricerca. Così nella valutazione e scelta tra i diversi provider disponibili e quindi tra le diverse applicazioni/software è necessario considerare con attenzione i seguenti fattori:

Facilità d'uso: Dato che i partecipanti baseranno la loro ricerca su un'applicazione, è necessario che sia intuitiva, semplice da usare e che agevoli il più possibile il completamento delle proprie attività. Se il partecipante è confuso o ha difficoltà nell'utilizzare l'applicazione, è più probabile che abbandoni la ricerca. La facilità d'uso è intesa anche dal punto di vista dei ricercatori. È infatti consigliato selezionare una piattaforma che agevoli il più possibile anche il lavoro amministrativo come, per esempio, la possibilità di impostazione di notifiche di promemoria automatiche.

Accessibilità: Vi è la possibilità di scegliere tra aziende che offrono strumenti basati sul web e ottimizzati per dispositivi mobili oppure aziende che offrono direttamente l'applicazione scaricabile per smartphone. La differenza sostanziale risiede nel fatto che i primi funzionano su ogni dispositivo digitale (telefono, tablet, PC) ma necessitano sempre di una connessione al web per la condivisione di informazioni. In questo contesto se i partecipanti si trovano in un luogo senza accesso ad Internet, ad esempio un ospedale o un trasporto pubblico sotterraneo, non saranno in grado di utilizzare lo strumento. Di conseguenza i partecipanti dovranno aspettare a condividere con il rischio di dimenticare quello che stavano provando in quel momento. Le applicazioni per smartphone invece, poiché memorizzate sul telefono hanno la capacità di essere utilizzate offline, consentendo ai partecipanti di continuare a utilizzare l'app per registrare e acquisire dati anche quando fuori dalla copertura di rete. Tuttavia, è utile considerare il fatto che esistono meno utenti di smartphone rispetto alla totalità dei dispositivi digitali connessi. Selezionando dunque l'azienda fornitrice di un'applicazione solo per telefoni cellulari si limiterebbe il campione da poter considerare per la ricerca. La selezione di un fornitore

di software o di applicazione dipende quindi dall'obiettivo della ricerca e da quale tipologia di campione si intende analizzare.

Funzionalità: Nel processo di selezione della piattaforma è utile considerare quali funzionalità è in grado di offrire. Un requisito fondamentale oltre alla possibilità di caricare foto, registrazioni e testi è la presenza della configurabilità per i video. Ciò significa riuscire a specificare la durata, la risoluzione e a effettuare la compressione per adattare i video ai requisiti del progetto. La funzione compressione (ossia la riduzione delle dimensioni del file prima del caricamento) permette il risparmio di tempo durante il caricamento soprattutto se si sta facendo uso della rete 3G/4G. Molto utili sono le piattaforme che offrono la possibilità di interagire con gli utenti tramite l'invio di notifiche push. Tale funzionalità è però generalmente presente solo nelle app native. Un'altra funzionalità ormai necessaria è la possibilità di impostare diverse lingue per poter ottenere informazioni da utenti di tutto il mondo.

Facilità di Analisi: Ulteriore aspetto importante da considerare è la presenza di strumenti di analisi integrati nella piattaforma che consentono di procedere istantaneamente alla fase che segue la raccolta dati. Indeemo consiglia la verifica della presenza di potenti capacità di filtraggio e codifica, della possibilità di taggare facilmente i dati e di esportare con velocità e facilità foto, audio, video e testi. Un'opzione aggiuntiva potrebbe anche essere la possibilità di trascrizione video e analisi delle parole chiave con l'obiettivo di ridurre il tempo di visione video e trascrizione manuale.

È facilmente intuibile come la scelta tra le diverse piattaforme disponibili con le loro capacità dipendono dall'obiettivo della ricerca, dalle proprie competenze e dal campione che si andrà ad analizzare.

A tale scopo è interessante analizzare, nello specifico, a titolo di esempio, due delle piattaforme comunemente usate per studi etnografici con l'obiettivo di rendere più chiaro quali strumenti esistono a supporto di tale metodo di ricerca.

Indeemo:

Indeemo si afferma come una piattaforma di etnografia mobile all'avanguardia, progettata specificamente per l'utilizzo su dispositivi digitali mobili. La sua applicazione può essere facilmente scaricata e installata su dispositivi palmari, offrendo un'esperienza utente

ottimizzata e intuitiva. La piattaforma si distingue per la sua notevole flessibilità, essendo in grado di adattarsi a vari obiettivi di ricerca, inclusi studi di diario e analisi di test prodotto.

I ricercatori hanno la possibilità di creare compiti personalizzati per i partecipanti, sfruttando un'ampia gamma di funzionalità. I partecipanti, a loro volta, possono condividere facilmente video, immagini, testi e registrazioni di schermate. Un elemento distintivo di Indeemo è la sua capacità di facilitare l'interazione in tempo reale tra ricercatori e partecipanti attraverso l'uso di commenti e notifiche push, gestibili attraverso l'applicazione del moderatore.

In termini di capacità multilingue e sicurezza, Indeemo supporta dieci lingue diverse e offre standard di sicurezza di livello aziendale, garantendo così una copertura globale e la protezione dei dati. In aggiunta, la piattaforma consente ai ricercatori di gestire più studi simultaneamente, ottimizzando così l'efficienza e la produttività del lavoro di ricerca.

Nella fase di analisi dei dati, Indeemo eccelle nel ridurre i tempi di elaborazione grazie alla sua dashboard unificata, che centralizza tutte le foto e i video raccolti. La funzione di filtraggio avanzato, la trascrizione automatica dei video, l'analisi delle parole chiave e gli strumenti di intelligenza artificiale integrati semplificano ulteriormente il processo, permettendo di ottenere approfondimenti ricchi di contesto. Inoltre, la piattaforma è capace di importare e analizzare rapidamente interviste aggiuntive, come quelle condotte tramite chiamate Zoom o focus group, grazie alle sue avanzate capacità di elaborazione basate sull'intelligenza artificiale.

In sintesi, Indeemo si presenta come una soluzione completa e versatile per la ricerca etnografica mobile, offrendo strumenti potenti e flessibili per una vasta gamma di applicazioni di ricerca.

La piattaforma si presenta nel seguente modo:

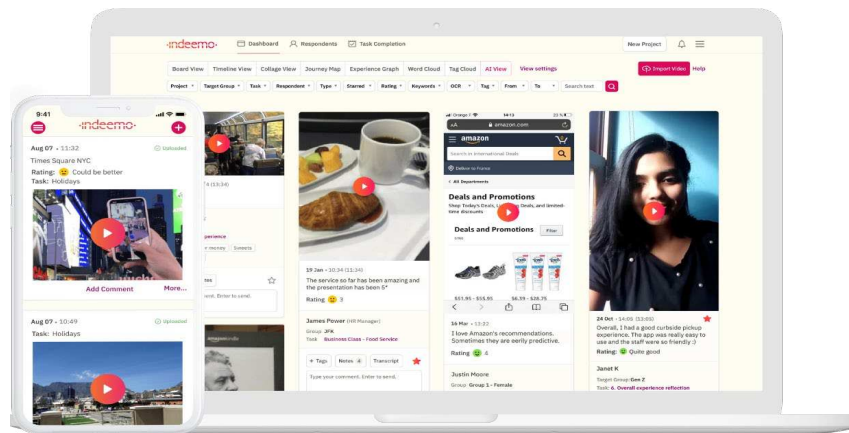


Figura 3.4-1 - Rappresentazione dell'applicazione Indeemo su smartphone e su PC.
Fonte(<https://indeemo.com>)

Dscout:

Dscout è una società di Chicago che offre una piattaforma di diario digitale per svolgere ricerche di etnografia mobile su dispositivi Android e iOS. Tale applicazione risulta essere molto chiara e facile da usare sia per i ricercatori ma anche per gli intervistati. In primo luogo, offre un metodo altamente efficace e semplice per personalizzare gli screener utilizzati nella selezione dei partecipanti, consentendo l'inserimento di parametri specifici quali il range di età e il paese di provenienza. Essa privilegia l'inclusione di partecipanti attivi e dotati di esperta conoscenza nel settore specifico, assicurando così elevata qualità nei partecipanti selezionati. Questa applicazione consente inoltre di definire in anticipo le ricompense per i partecipanti, fornendo loro l'opportunità di valutare la propria partecipazione al progetto. In aggiunta, offre ai ricercatori la flessibilità di optare per metodi di raccolta dati, sia con che senza la supervisione di un moderatore.

Gli utenti, mediante il download dell'applicazione sul proprio dispositivo mobile, l'autenticazione tramite login e il rispetto dei criteri di accesso definiti dal ricercatore, acquisiscono la possibilità di partecipare a "missioni", note anche come sondaggi retribuiti. Queste missioni richiedono agli utenti di documentare le proprie esperienze quotidiane relative a un prodotto o servizio specifico, utilizzando video, immagini e rispondendo a quesiti mirati. Inoltre, tale strumento offre una funzionalità di interazione

in tempo reale, attraverso la quale i partecipanti possono essere intervistati o interagire direttamente con i ricercatori, fornendo così ulteriori approfondimenti e feedback.

Per la sezione di analisi dati la piattaforma Dscout integra l'apprendimento automatico con funzionalità avanzate e all'avanguardia con l'obiettivo di facilitare e aumentare il successo di questo processo. Questa combinazione permette agli analisti di estrarre informazioni preziose e tendenze emergenti con una rapidità e precisione. La capacità di convertire automaticamente il contenuto di video e audio in testo trascritto, la creazione di nuvole di parole e l'utilizzo di filtri di espressività⁵⁰, consentono di cogliere sfumature e insights che altrimenti potrebbero sfuggire. Un esempio concreto dell'applicazione di questa piattaforma si trova nello studio condotto da Loh et al. (2023) intitolato "Because I'm always moving!: a mobile ethnography study of adolescent girls' everyday print and digital reading practices". Gli autori hanno utilizzato Dscout per la sua interfaccia user-friendly e la sua intuitività, caratteristiche che hanno reso la piattaforma particolarmente adatta per i partecipanti coinvolti nello studio.

La scelta di questa piattaforma ha permesso una raccolta dati fluida e un'analisi approfondita dei comportamenti e delle preferenze di lettura delle adolescenti. In definitiva, anche Dscout risulta uno strumento fondamentale per gli studiosi che cercano di comprendere meglio i comportamenti e le esperienze degli utenti in un'ampia varietà di contesti.

Questi due esempi illustrano solamente una piccola parte dell'ampio spettro di applicazioni disponibili nel mercato attuale. Inoltre, il crescente interesse verso questo metodo di ricerca sta portando a un significativo aumento del numero di piattaforme specializzate disponibili. Tale tendenza evidenzia un'evoluzione del seguente campo di ricerca, sottolineando l'importanza e la popolarità delle metodologie di raccolta dati basate su strumenti digitali.

⁵⁰ I filtri di espressività si riferiscono ai metodi utilizzati per analizzare l'espressione e le reazioni emotive dei partecipanti in studi di etnografia mobile. Questi possono includere l'analisi del linguaggio, l'espressione facciale, il tono della voce e altri indicatori comportamentali. Vengono utilizzati per filtrare le espressioni desiderate e accelerare così il processo di analisi video.

Capitolo 4

IA e netnografia: sperimentazione per l'analisi dei testi

L'analisi finora condotta sulla netnografia è stata orientata a preparare la base per affrontare ora, con particolare attenzione, quello che risulta essere il centro più sperimentale e concreto dello studio, ovvero l'esplorazione di strumenti avanzati che incorporano algoritmi di intelligenza artificiale, essenziali per l'analisi dei dati ottenuti attraverso la ricerca netnografica. Verranno esaminati in dettaglio due diversi strumenti di text mining applicati a un caso reale: il primo è una piattaforma software esistente, accessibile alle aziende tramite un abbonamento a pagamento; il secondo è basato sull'utilizzo di un linguaggio di programmazione nel quale si adotterà un approccio sperimentale utilizzando il linguaggio Python per sviluppare un codice specifico. L'intento è di comprendere, in termini pratici, le possibilità disponibili, a pagamento e non, per analizzare automaticamente i testi condivisi all'interno delle comunità online, sfruttando queste tecnologie.

4.1. Strumenti disponibili per l'analisi dati

Nell'ambito della netnografia, l'impiego di avanzate tecnologie basate sull'utilizzo di algoritmi di intelligenza artificiale è diventato un pilastro fondamentale per il supporto di un'analisi efficace e approfondita dei grandi volumi di dati generati dalle comunità digitali. I diversi algoritmi esistenti forniscono la possibilità, infatti, di ridurre il tempo di analisi e di aumentare il livello di dettaglio nei risultati. Durante il processo text mining,

utile per l'analisi della ricerca netnografica, due dei principali algoritmi di IA che vengono spesso utilizzati sono gli algoritmi che consentono l'analisi del sentimento e gli algoritmi che identificano i temi principali.

Nello specifico, l'analisi del sentimento si riferisce all'uso di algoritmi progettati per valutare e classificare le emozioni e i sentimenti espressi in un testo. Questo processo impiega tecniche di elaborazione del linguaggio naturale (NLP), che vanno dalle reti neurali e altri metodi avanzati di machine learning fino a metodi statistici e basati su regole più tradizionali. Queste tecniche sono utilizzate per leggere e comprendere il linguaggio umano, interpretando le espressioni e identificando la positività o la negatività dei toni nei testi. Tale analisi viene spesso utilizzata quindi per passare al vaglio recensioni, feedback dei clienti e commenti sui social media con lo scopo di ottenere importanti intuizioni sulle percezioni del pubblico. Questo processo è prezioso per gli analisti di dati nelle grandi imprese, in quanto consente di valutare l'opinione pubblica, condurre ricerche di mercato dettagliate, monitorare la reputazione di marchi e prodotti e comprendere le esperienze dei clienti (Chadha et. al.,2023).

Per l'identificazione dei temi sono invece utilizzati algoritmi di topic modeling, utili per identificare e categorizzare i principali argomenti o temi che si presentano nei dati testuali raccolti. I ricercatori, dunque, tramite questa analisi sono in grado di identificare con estrema velocità gli argomenti chiave che i consumatori hanno esposto nelle loro discussioni, consentendo la focalizzazione sui principali punti di forza e debolezza del prodotto/servizio.

Alcuni di questi algoritmi si fondano sull'apprendimento automatico senza supervisione per identificare cluster o gruppi di parole simili all'interno di un corpo di testo. Un noto algoritmo che utilizza tale tecnologia, fruibile per svolgere tale compito, è il Latent Dirichlet Allocation (LDA). Fondandosi su un principio probabilistico, tratta ogni documento come un insieme di diversi argomenti in proporzioni variabili. In pratica, considera che ogni argomento sia caratterizzato da un gruppo specifico di parole chiave. L'algoritmo assegna quindi le parole ai vari argomenti in base alla loro probabilità di appartenenza, raggruppando le parole simili sotto lo stesso argomento.

È inoltre necessario specificare che, oltre a LDA e altri algoritmi tradizionali, il topic modeling può essere eseguito anche tramite le moderni reti neurali profonde,

particolarmente efficaci nel catturare relazioni complesse e sottili nel testo, offrendo una categorizzazione dei topic più accurata e contestualizzata. Questo processo automatizzato facilita dunque l'analisi dei testi, permettendo di scoprire i temi prevalenti nei documenti senza la necessità di una categorizzazione manuale.

Per l'applicazione concreta di questi algoritmi, il ricercatore può scegliere tra diversi strumenti disponibili. Esistono infatti svariati software di analisi dati che si concentrano nell'identificare le preziose informazioni che si nascondono dietro i dati testuali. Alcuni, nati come supporto al ricercatore per la codifica manuale delle informazioni raccolte si stanno sempre più evolvendo integrando automatismi grazie all'apporto della cosiddetta intelligenza artificiale. Attraverso algoritmi di questa tipologia il ricercatore ha facoltà di scavare nelle informazioni con elevata rapidità e con altrettanta precisione ottenendo risultati magari difficilmente reperibili con la sola manodopera dell'operatore. È utile quindi precisare che questi software disponibili all'acquisto nascono come supporto alle attività di analisi e di ricerca di insight e rimangono tali anche con l'introduzione dei nuovi algoritmi basati su IA. Si tratta infatti di un significativo progresso nelle capacità del software dove gli algoritmi di intelligenza artificiale sono progettati per processare, analizzare e identificare pattern nei dati. Tuttavia, è fondamentale riconoscere che la comprensione accurata del contesto, l'interpretazione approfondita dei dati e il giudizio finale restano domini dove l'intervento umano è insostituibile.

Di seguito vengono ora presentati alcuni software disponibili che integrano nelle loro attività di analisi manuale, algoritmi basati su intelligenza artificiale.

NVivo

NVivo, sviluppato dalla società Lumivero⁵¹, è un rinomato pacchetto software di analisi qualitativa e quantitativa dei dati. Nato nel 1999, grazie alle sue capacità, è diventato uno strumento essenziale per i ricercatori in diversi campi, tra cui la scienza sociale, l'antropologia, il turismo e la sanità. Questo strumento è infatti specializzato nell'analisi di dati non strutturati e aiuta di conseguenza i ricercatori ad organizzare, analizzare e

⁵¹ Nata nel 1995, Lumivero è una compagnia che sviluppa software per la ricerca e l'analisi di dati. È nota per essere la società madre di alcuni strumenti di ricerca qualitativa e quantitativa, tra cui NVivo. Lumivero riunisce strumenti di ricerca, decisionali e organizzativi affidabili che consentono ai professionisti di svolgere il proprio lavoro in modo efficiente e produttivo e di prendere decisioni con maggiore sicurezza.

approfondire tali informazioni provenienti da testi di interviste, recensioni, sondaggi, riviste, ma anche da contenuti multimediali come immagini, video e registrazioni audio.

Il pacchetto software offre un piano di pagamento annuale o mensile e prevede diverse licenze a seconda che si tratti di un utente individuale, un team di ricerca oppure un'istituzione accademica. A disposizione vi sono inoltre 14 giorni di prova gratuita.

Tra le funzionalità principali di NVivo, si evidenzia la capacità di organizzare e codificare dati complessi, rendendo più semplice per i ricercatori categorizzare e analizzare grandi quantità di informazioni. All'interno della piattaforma sono disponibili, inoltre, potenti strumenti di visualizzazione, quali grafici e mappe mentali, che aiutano a presentare i risultati secondo logiche più chiare e intuitive. Il pacchetto offre inoltre la funzionalità di ricerca di testo e di calcolo della frequenza di parole, quest'ultime visualizzabili anche tramite il grafico "nuvola di parole". In aggiunta vi è la possibilità di riprodurre audio o video direttamente nel software in modo da produrre la trascrizione. Tale operazione può essere svolta manualmente o, invece, supportata dalla trascrizione automatica tramite l'uso di IA. Infatti, un aspetto distintivo di NVivo è la presenza di algoritmi di intelligenza artificiale. In particolare, di algoritmi per l'analisi del sentimento e di topic modeling. Nel primo caso vi è come output il numero di frasi identificate come positive e negative. Si offre a chi lo utilizza anche la facoltà di accedere alla lettura delle frasi catalogate con un tono positivo o negativo, dando così la possibilità al ricercatore di capire il contesto nel quale queste frasi sono state espresse. L'identificazione rapida dei temi nei dati testuali invece consente di capire quali argomenti sono più discussi o rilevanti all'interno del loro set di dati. Questo consente di orientare più efficacemente l'analisi successiva o di identificare aree che richiedono ulteriori approfondimenti. Tale piattaforma è quindi costituita da una automazione di una parte del processo di analisi che consente di migliorare la gestione di grandi set di dati. Si rivela dunque un ottimo software per l'analisi dei testi in un contesto di ricerca netnografia.

Brandwatch

Brandwatch, fondata nel 2007 da Giles Palmer a Brighton, Regno Unito, è una piattaforma di ascolto e analisi dei social media, nota per le sue capacità avanzate di monitoraggio e analisi dei dati online. Tale software prevede tre piani di pagamento differenti a seconda dello scopo della ricerca (Consumer Intelligence, Social Media

Management, Influencer Marketing). In primis, il software è stato progettato per raccogliere, monitorare e analizzare le conversazioni sui social media fornendo insight sulle opinioni e i comportamenti dei consumatori. Vi è però la possibilità anche di caricare direttamente i dati privati raccolti tramite interviste, sondaggi, recensioni etc. e di analizzarli separatamente o tramite la combinazione con le informazioni raccolte sui social media migliorando l'accuratezza dei risultati. L'analisi avviene prevalentemente utilizzando algoritmi di intelligenza artificiale. Si effettua infatti l'analisi del sentimento espresso sui social media o sui dati privati tramite tecniche di NLP. Tra le funzioni disponibili si distingue in particolare il riconoscimento di immagini e video identificando loghi, prodotti e scene principali attraverso l'uso di algoritmi basati su reti neurali convoluzionali (CNN). Tale piattaforma offre anche la possibilità di analisi predittiva consentendo di identificare le tendenze emergenti e prevedere sviluppi futuri nel comportamento dei consumatori. Si tratta dunque di ottenere informazioni estremamente importanti per decisioni strategie di sviluppo prodotto e di marchio.

Il software declina diverse soluzioni in base al caso d'uso, al settore e al ruolo di chi utilizza il pacchetto e i diversi step di raccolta e analisi possono essere supportati da un assistente digitale basata sull'IA, chiama Iris.

Tuttavia, non tutte le operazioni prevedono l'uso di algoritmi di intelligenza artificiale. Infatti, escluse le funzionalità avanzate come analisi sentimento e riconoscimento immagini, alcuni strumenti di raccolta e analisi sono basati su metodi tradizionali. Per esempio, Brandwatch consente agli utenti di segmentare i dati in base a vari criteri, come demografia, ubicazione o fonte, utilizzando metodi di classificazione e filtraggio semplici. Ancora, la piattaforma fornisce dashboard personalizzabili che consentono agli utenti di visualizzare e organizzare i dati raccolti in modi significativi. Queste visualizzazioni possono includere grafici e nuvole di parole. Questo rende la piattaforma versatile e utile anche per utenti che non necessitano di analisi complesse basate su IA.

Ampiamente utilizzato per la ricerca di mercato e per ottenere insight approfonditi sui comportamenti dei consumatori, è dunque possibile affermare che Brandwatch si rivela un ottimo software per l'analisi netnografica che consente oltretutto l'integrazione dell'IA.

MAXQDA

MAXQDA è un rinomato software di analisi qualitativa e quantitativa, sviluppato dalla VERBI Software. Questo strumento si è affermato nel tempo come uno dei principali software nell'ambito della ricerca sociale, fornendo soluzioni affidabili e versatili per l'analisi dei testi, sondaggi, immagini, audio e video. Le licenze per tale software variano a seconda dell'utente. Si passa infatti da piani di pagamento per studenti e istituzioni scolastiche a piani per aziende e il conseguente acquisto può avvenire sia tramite l'acquisto di una licenza permanente sia tramite abbonamenti annuali.

Al centro delle capacità di MAXQDA spicca la sua robusta funzione di codifica, che permette agli utenti di organizzare e categorizzare i dati qualitativi non strutturati facilitando l'analisi dettagliata. Il software supporta l'analisi di una varietà di formati di dati, inclusi testi, immagini, file audio e video. Questa flessibilità è particolarmente utile in contesti di ricerca che richiedono una combinazione di dati qualitativi e quantitativi. Inoltre, MAXQDA offre una varietà di strumenti visivi come grafici, mappe e diagrammi che consentono di ottenere una rapida panoramica dei dati e aiutano a identificare i modelli. Un altro strumento visivo disponibile, utile per l'analisi dei contenuti, è la nuvola di parole che visualizza le parole più frequenti e permette di esplorare i termini chiave e i temi centrali di uno o più articoli. Negli ultimi anni, MAXQDA ha integrato funzionalità basate sull'intelligenza artificiale. Si tratta più nello specifico di un assistente in grado di supportare l'azione dell'uomo nella fase di riepilogo testi e nella funzione di codifica. Infatti, nella funzione di codifica l'assistente con IA è in grado di analizzare il testo e suggerire i sottocodici. Inoltre, l'algoritmo basato su IA risulta in grado di fornire la trascrizione automatica di audio o video. Infine, anche MAXQDA integra la funzionalità automatica di analisi del sentimento utilizzando algoritmi per analizzare i dati testuali e identificare automaticamente i sentimenti espressi, classificandoli come positivi, negativi o neutri.

Il software trova applicazione in una vasta gamma di contesti, dalla ricerca accademica, dove è utilizzato per esplorare complesse questioni sociali, psicologiche e politiche, alla ricerca di mercato, dove aiuta le aziende a comprendere meglio i comportamenti e le percezioni dei consumatori.

In conclusione, MAXQDA rappresenta una soluzione completa per chiunque necessiti di analizzare dati qualitativi e misti, combinando un approccio tradizionale all'analisi qualitativa con le più recenti innovazioni nell'ambito dell'intelligenza artificiale.

Nella scelta tra i software disponibili supportati dall'IA per l'analisi qualitativa in una ricerca netnografica, esiste per il ricercatore anche un'altra soluzione, ricorrendo a un differente strumento, ovvero la programmazione informatica di un codice specifico. Questo permette al ricercatore di personalizzare l'analisi in base alle specifiche esigenze del progetto, permettendo una maggiore flessibilità e potenzialmente una maggiore profondità nell'analisi dei dati raccolti online.

Il linguaggio di programmazione Python, considerato il principale linguaggio per l'intelligenza artificiale, offre, grazie alle sue potenti librerie⁵², diverse funzionalità che lo rendono particolarmente utile per la raccolta e analisi di una ricerca netnografica supportata da IA. Essendo uno dei linguaggi di programmazione più popolari e con un'ampia comunità di sviluppatori, permette, attraverso determinate librerie, di implementare algoritmi di IA per l'analisi testuale, l'analisi del sentimento, il topic modeling e molto altro. Oltre alle sue applicazioni nell'intelligenza artificiale, Python offre la capacità di automatizzare efficacemente la raccolta di grandi quantità di dati da una varietà di fonti online, inclusi social media, blog, forum e siti di recensioni. Questo processo, comunemente noto come "web scraping", si avvale di specifiche librerie disponibili in Python, tra cui "BeautifulSoup". Questa permette di analizzare ed estrarre informazioni utili dalle pagine web in modo efficiente e affidabile, facilitando significativamente il processo di raccolta dati su larga scala.

Python, grazie alla sua vasta gamma di librerie, supporta anche la visualizzazione avanzata dei dati, consentendo agli utenti di rappresentare i risultati raccolti sotto forma di grafici intuitivi. In particolare, si possono creare grafici lineari e a colonna, tra le altre forme di visualizzazione.

Di seguito alcune delle librerie più note utili all'analisi dei dati supportata da algoritmi basati su IA:

⁵² Una libreria è una raccolta di moduli e funzioni prescritte che gli sviluppatori possono utilizzare per eseguire una varietà di compiti senza dover scrivere codice da zero. Le librerie Python possono essere scritte da diversi gruppi o individui, tra cui comunità open source e sviluppatori core di Python.

NLTK (Natural Language Toolkit): si tratta di una delle librerie più note e utilizzate per lavorare con il linguaggio umano, data la sua ricchezza di moduli e funzionalità per l'interpretazione e analisi di testi. Tale libreria consente infatti di eseguire diverse operazioni su testi non strutturati, come la preparazione del testo tramite la tokenizzazione, la rimozione delle stopwords, la lemmatizzazione etc. Queste operazioni contribuiscono di conseguenza a rendere il testo in un formato più strutturato e analizzabile. Inoltre, tale libreria consente anche l'analisi del sentimento tramite l'uso di specifici algoritmi. Uno di questi è VADER sentiment analyzer⁵³, progettato per identificare e classificare il sentimento espresso nei testi.

Pandas: pandas è una libreria di Python open source, potente e popolare, usata per la manipolazione e l'analisi dei dati. La libreria è ampiamente utilizzata per la pulizia dei dati, offrendo funzionalità per gestire valori mancanti, rimuovere duplicati, filtrare dati e trasformare i tipi di dati. È inoltre in grado di fornire una varietà di funzioni e metodi integrati per l'analisi e la visualizzazione dei dati, come il calcolo delle statistiche di riepilogo, ed essendo combinabile con altre librerie come "matplotlib" è in grado di fornire anche la creazione di grafici. È particolarmente utile per gestire grandi dataset, che sono comuni in ricerche netnografiche.

TensorFlow e Pytorch: si tratta di due delle più popolari librerie open source per applicazioni delle reti neurali profonde. Entrambe offrono diversi modelli pre-addestrati per i vari utilizzi. In particolare, modelli come BERT⁵⁴ (Bidirectional Encoder Representations from Transformers), che sono stati addestrati su vasti insiemi di dati, possono essere facilmente utilizzati con queste librerie per vari compiti di NLP, come la classificazione del sentimento. Entrambe inoltre offrono la flessibilità e una vasta gamma di strumenti per costruire e addestrare modelli di deep learning.

Gensim: il nome identifica una libreria open source specificamente progettata per il topic modeling e altre applicazioni di elaborazione del linguaggio naturale come l'indicizzazione di documenti e il recupero di similarità. Supporta inoltre algoritmi

⁵³Vader è un modulo all'interno della libreria NLTK. Lo strumento è specificamente progettato per i sentimenti espressi nei social media e utilizza una combinazione di un lessico di sentimento e un elenco di caratteristiche lessicali che sono generalmente etichettate in base al loro orientamento semantico come positivo o negativo.

⁵⁴ Rete neurale pre-addestrata su 2.5 miliardi di parole in Wikipedia e 800 milioni sui libri .

popolari come Latent Dirichlet Allocation (LDA) e Latent Semantic Analysis (LSA), utilizzati per scoprire la struttura tematica nascosta in grandi collezioni di testo (Řehůřek et. al, 2010).

Usando queste librerie, un ricercatore può personalizzare l'analisi in base alle esigenze specifiche del suo progetto. Ad esempio, possono essere sviluppati algoritmi personalizzati sia per l'analisi del sentimento, che per il topic modeling.

È importante notare che, a seconda delle operazioni da eseguire, il ricercatore può optare per l'uso di librerie che contengono algoritmi con o senza tecniche di apprendimento automatico. Ad esempio, VADER, che si basa su regole e lessico per l'analisi del sentimento, o LDA, un modello di apprendimento automatico che utilizza approcci probabilistici, per l'analisi dei topic, oppure optare per librerie che utilizzano reti neurali profonde, offrendo così funzionalità più avanzate e potenzialmente più accurate. Esistono infatti molti compiti nel campo del Natural Language Processing (NLP) come l'analisi del sentimento e il topic modeling che possono essere eseguiti efficacemente sia con metodi tradizionali che con reti neurali. In poche parole, è corretto identificare gli algoritmi tradizionali (p.e. Vader e LDA) come una forma basilare di IA, mentre algoritmi basati su reti neurali profonde come intelligenza artificiale avanzata.

La scelta del ricercatore nell'utilizzare reti neurali più complesse o algoritmi più basilari dipende da vari fattori. Tra questi spicca sicuramente la tipologia di informazioni da processare e dall'obiettivo dell'analisi. Infatti, se è necessario analizzare in rapidità testi brevi e diretti sarà sufficiente implementare un algoritmo basato su regole, lessico e probabilità e che non necessita di un addestramento profondo. Qualora, invece, l'obiettivo principale si concentri verso la comprensione di testi lunghi e complessi dove è necessario identificare precisamente il contesto al fine di ottenere un'analisi sofisticata, allora la scelta ricadrà nell'uso di reti neurali. La scelta dipende inoltre anche dalle risorse disponibili. L'uso di reti neurali e il conseguente addestramento richiedono risorse computazionali maggiori, specialmente per il trattamento di grandi set di dati necessari per l'addestramento. Non va sottovalutato che, per poter garantire risultati soddisfacenti, le reti neurali profonde necessitano di un propedeutico addestramento con dati di esempio. Questo processo, però, impone l'impiego di molto tempo e molta energia da parte del ricercatore. Nonostante queste sfide, la comunità Python offre una soluzione pratica

attraverso la disponibilità di reti neurali open source già addestrate. Queste reti, pre-addestrate su specifici insiemi di dati, permettono agli utenti di accelerare notevolmente il processo di analisi. Tuttavia, non va sottovalutato che l'utilizzo di reti pre-addestrate può talvolta comportare un compromesso in termini di precisione⁵⁵.

Ciò nonostante, queste risorse rappresentano un valido strumento per coloro che desiderano applicare rapidamente le tecnologie di rete neurale profonda senza l'onere di un lungo e complesso processo di addestramento.

È infine possibile affermare che in molti casi, gli approcci tradizionali possono essere sufficienti e più pratici, mentre in altri scenari le reti neurali possono offrire vantaggi significativi in termini di accuratezza e profondità di analisi.

In conclusione è quindi possibile affermare che la scrittura di un corretto codice con linguaggio di programmazione Python permette al ricercatore netnografico un supporto non indifferente nella raccolta e analisi dei dati.

4.2. Caso Western Union

Verrà ora presentato nel seguente paragrafo l'applicazione a un caso reale di alcune tecniche di analisi dati mediante l'uso di strumenti idonei al supporto umano nello svolgimento di tale operazione. Nello specifico si tratta di procedere con l'utilizzo delle funzioni di analisi di sentimento, analisi tematica e calcolo della frequenza di parole. Quest'ultima, anche se eseguita senza algoritmi di IA è stata inserita con lo scopo di fornire un quadro più completo delle funzionalità che gli strumenti presentano. Per fare ciò, si farà riferimento sia all'utilizzo di un software sia alla programmazione con Python usando, per quest'ultimo, due diversi metodi: con e senza l'uso di reti neurali. Lo scopo

⁵⁵ Tali reti pre-addestrate sono state create su un set di dati specifico, il che significa che possono non essere ottimizzate per i dati che il ricercatore sta analizzando. Per mitigare questo problema, si può effettuare un ulteriore addestramento, noto come "fine-tuning", dove la rete viene ulteriormente addestrata su un set di dati più idoneo al caso in questione. Questo può migliorare la precisione, ma richiede tempo e risorse aggiuntivi.

è quello di fornire una panoramica delle scelte che un ricercatore, che voglia sperimentare l'analisi con l'uso di algoritmi di IA, può fare tra un software a pagamento e i metodi che possono essere affrontati tramite lo strumento della programmazione, disponibile senza l'impiego di risorse monetarie.

Per l'applicazione del caso si è fatto riferimento al sito di recensioni mondiale comunemente noto come TrustPilot. Tale piattaforma gratuita si presenta aperta a tutti, fondata sulla trasparenza e come afferma la stessa azienda è stata creata al fine di raccogliere recensioni per aiutare i consumatori a fare acquisti in sicurezza e per fornire preziosi approfondimenti alle aziende per aiutarle a migliorare l'esperienza che offrono ai propri clienti. Nello specifico, per l'analisi, è stato selezionato l'insieme di recensioni appartenenti all'azienda Western Union. La scelta è ricaduta su tale azienda per due principali motivi: l'enorme comunità di utenti attivi nel postare recensioni e la provenienza di tale azienda che garantisce la presenza della maggior parte di recensioni in lingua inglese, ideale per gli algoritmi che vengono utilizzati dalle librerie di Python.

Western Union è un'azienda americana leader nel settore dei servizi finanziari, specializzata in trasferimenti di denaro a livello internazionale. Fondata nel 1851 a New York, conta ad oggi circa 300 mila sedi in oltre 200 paesi con un fatturato di 4,5 miliardi di dollari. Il core business di tale azienda risulta essere il trasferimento di denaro da un paese all'altro tramite app mobile oppure di persona tramite sportelli fisici. Oltre ai trasferimenti di denaro, Western Union offre anche altri servizi finanziari, come pagamenti di bollette e servizi di cambio valuta.

La scelta del software su cui effettuare la sperimentazione degli strumenti di analisi con algoritmi di IA è ricaduta sul rinomato software NVivo precedentemente delineato. Ciò si è reso possibile grazie alla presenza di una prova gratuita di 14 giorni e alla semplicità di utilizzo anche per un utente alle prime armi. Una caratteristica importante di NVivo è la necessità di avere già i dati pre-raccolti per l'analisi. Prima di utilizzare NVivo, è quindi essenziale effettuare una raccolta accurata dei dati. Una volta raccolti e salvati i dati in formati compatibili, come documenti di testo, fogli di calcolo, o file audio/video, gli stessi possono essere importati in NVivo per l'analisi qualitativa o mista.

Si è reso pertanto necessario, prima di applicare gli strumenti di analisi, riuscire a "scaricare" le recensioni fornite dagli utenti e salvarle in un apposito file csv. Per

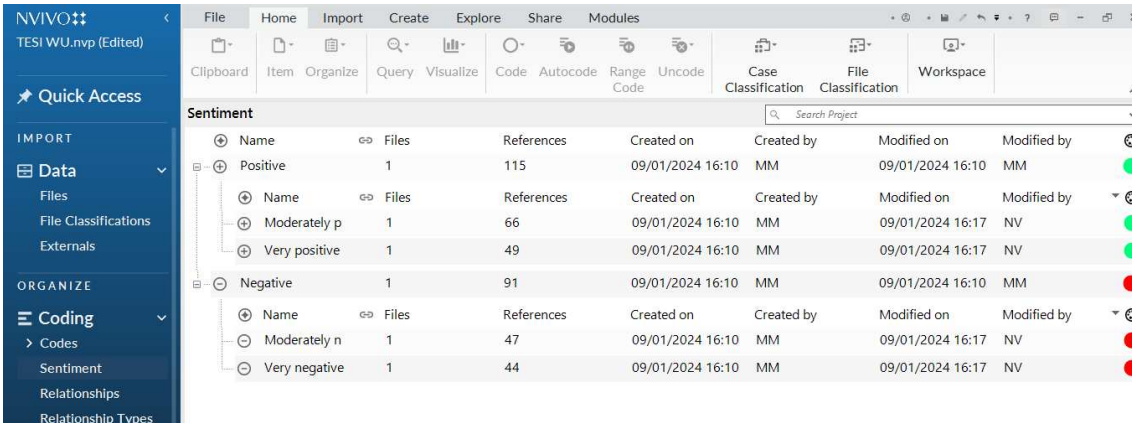
effettuare l'operazione di scraping, si è fatto ricorso alla programmazione tramite Python nel quale è stata utilizzata la libreria “BeautifulSoup”⁵⁶ in grado di scaricare e analizzare il contenuto delle pagine web. È stato dunque necessario specificare l’URL del sito web e la classe HTML in cui erano contenute le recensioni. La funzione “scrape_page” è stata poi impiegata per automatizzare il download e l'analisi del contenuto di ogni pagina, mentre “scrape_pages” ha esteso questo processo a più pagine. Questa parte di codice è stata successivamente mantenuta in comune anche per la scrittura del codice manuale per l’analisi tramite Python. Per garantire l'acquisizione di un insieme di dati ampio e rappresentativo, si è proceduto con il download di venti pagine contenenti recensioni. Tale metodologia è stata adottata con l'intento di raccogliere un volume di informazioni adeguato ad un'analisi sostanziale e per facilitare un'equilibrata sintesi di valutazioni sia positive che negative. Complessivamente, 299 recensioni sono state scaricate e conservate in un file adeguato. È stato successivamente possibile caricare nel software NVivo l’apposito file e avviare la potente funzione “Autocode” la quale, tramite appositi algoritmi basati su IA e l’istruzione di codificare le intere celle di recensioni e non solo singole frasi all’interno delle stesse, ha effettuato l’analisi dei sentimenti e l’analisi tematica. È inoltre importato sottolineare che il software NVivo ha la capacità di eseguire autonomamente la necessaria pre-elaborazione dei dati eseguendo sia la rimozione delle stop-words sia la tokenizzazione.

Rispetto all’analisi dei sentimenti, il software si basa su un algoritmo in grado di esaminare il sentimento delle parole dei testi in modo isolato, non considerando quindi il contesto. Viene a mancare, di conseguenza, il riconoscimento del sarcasmo, dei doppi negativi, delle espressioni dialettali e delle ambiguità. Il processo di analisi del sentimento delle recensioni impiega un lessico, ovvero una raccolta predefinita di parole ognuna delle quali è associata a un valore emotivo determinato. Nell'analisi, ciascuna parola del testo viene confrontata con questo lessico; a seconda della corrispondenza trovata, la parola viene assegnata un punteggio che riflette un sentimento, variando da estremamente negativo a estremamente positivo. Questo metodo si basa sull'assunzione che il significato e il sentimento di una parola siano costanti, il che non è sempre vero nel linguaggio

⁵⁶ BeautifulSoup è un pacchetto Python per l'analisi di documenti HTML e XML. Crea un albero di analisi per le pagine analizzate che può essere utilizzato per estrarre dati da HTML, utile per il web scraping. Tale libreria aiuta quindi a lavorare con i dati che si trovano nei siti web; può leggere e organizzare le informazioni nascoste nel codice delle pagine web, rendendo facile trovare e prendere i dati che servono.

naturale dove il contesto può cambiare il significato di una parola. La precisione, dunque, di questa analisi risulta limitata.

Grazie a questo algoritmo, NVivo è stato in grado di identificare il numero di recensioni positive e negative inserendo nel dettaglio anche la numerazione delle recensioni classificate come “molto positive”, “moderatamente positive”, “molto negative” e “moderatamente negative”. Così facendo il ricercatore si trova ad avere davanti a sé una panoramica generale dell’andamento di cosa pensano gli utilizzatori del servizio. Inoltre, NVivo offre la possibilità di osservare nello specifico le singole recensioni classificate con un tono positivo o negativo consentendo così una maggiore comprensione del contesto e un’analisi più approfondita. In totale nel nostro caso specifico sono state identificate 115 recensioni positive e 91 negative. Di seguito viene presentata la schermata (Fig. 4.2-1) che il software offre, rappresentante le quantità di recensioni positive e negative.



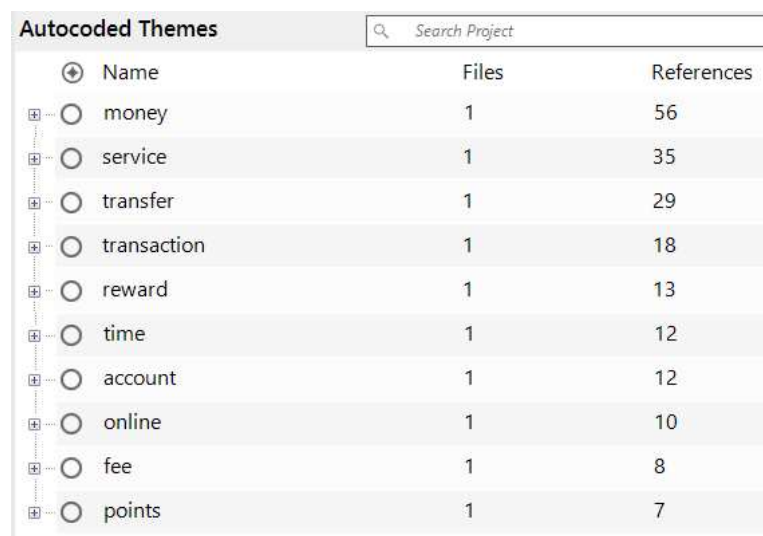
Name	Files	References	Created on	Created by	Modified on	Modified by	
Positive	1	115	09/01/2024 16:10	MM	09/01/2024 16:10	MM	●
Moderately p	1	66	09/01/2024 16:10	MM	09/01/2024 16:17	NV	●
Very positive	1	49	09/01/2024 16:10	MM	09/01/2024 16:17	NV	●
Negative	1	91	09/01/2024 16:10	MM	09/01/2024 16:10	MM	●
Moderately n	1	47	09/01/2024 16:10	MM	09/01/2024 16:17	NV	●
Very negative	1	44	09/01/2024 16:10	MM	09/01/2024 16:17	NV	●

Figura 4.2-1 - Screenshot della schermata del software Nvivo raffigurante l'analisi del sentimento

Per l’analisi tematica invece, è stata restituita una lista dei temi principali che il programma ha identificato. In particolare, NVivo esamina i testi per identificare temi, analizzando sia il contenuto che la struttura delle frasi alla ricerca di parole o frasi che ricorrono frequentemente. Questi elementi ricorrenti vengono considerati come temi. Il software valuta poi l’importanza di questi temi in base alla loro frequenza nei documenti, con i temi più frequenti ritenuti più significativi. I temi identificati vengono poi organizzati in gruppi, ognuno rappresentante un’idea generale. All’interno di ciascun

gruppo, si trovano nodi secondari per ogni tema specifico. I risultati dell'analisi vengono quindi presentati in una matrice di nodi che evidenzia i nodi per ogni idea generale e il numero di riferimenti di codifica da ciascun file. NVivo non include tutti i temi identificati nell'analisi finale, ma si concentra solo su quelli più rilevanti, facilitando così la focalizzazione sulle parti più importanti dei dati. Si offre poi la possibilità di personalizzare l'analisi selezionando quali temi includere come nodi. Inoltre, NVivo raggruppa i temi confrontando le parole con radici comuni. Ad esempio, parole come 'casa', 'case' e 'abitazione' vengono considerate varianti dello stesso tema e raggruppate insieme. Per ciascun gruppo tematico, la parola o frase più frequente viene utilizzata come nome del nodo, rendendo i risultati dell'analisi più chiari e facilmente comprensibili.

Nel caso specifico di Western Union il software ha identificato e raffigurato mediante la seguente tabella (Fig. 4.2-2) i temi principali nominati con la parola più frequente di ciascun nodo. La terza colonna mostra il numero di volte che il tema è stato menzionato o codificato nei dati. Questo dà un'idea della frequenza e dell'importanza di ciascun tema all'interno del file o del set di dati.



Autocoded Themes		Search Project	
Name	Files	References	
money	1	56	
service	1	35	
transfer	1	29	
transaction	1	18	
reward	1	13	
time	1	12	
account	1	12	
online	1	10	
fee	1	8	
points	1	7	

Figura 4.2-2 - Output analisi tematica con il software NVivo

Il tema "money", con 56 riferimenti suggerisce che è un argomento centrale o un concetto molto discusso nei dati. A seguire anche i temi "service" con 35 riferimenti e "transfer" con 29 riferimenti, risultano avere un'importanza significativa. Altri temi come "transaction", "reward", "time", "account", "online", "fee", "points" hanno una frequenza

decescente di riferimenti, suggerendo che potrebbero essere meno centrali rispetto ai temi principali ma comunque presenti nei dati.

Il software è inoltre in grado di offrire un'ulteriore rappresentazione dei risultati tematici in grado di dare al ricercatore una visuale generale dell'importanza dei temi e dei nodi secondari, come si vede qui di seguito (Fig. 4.2-3).

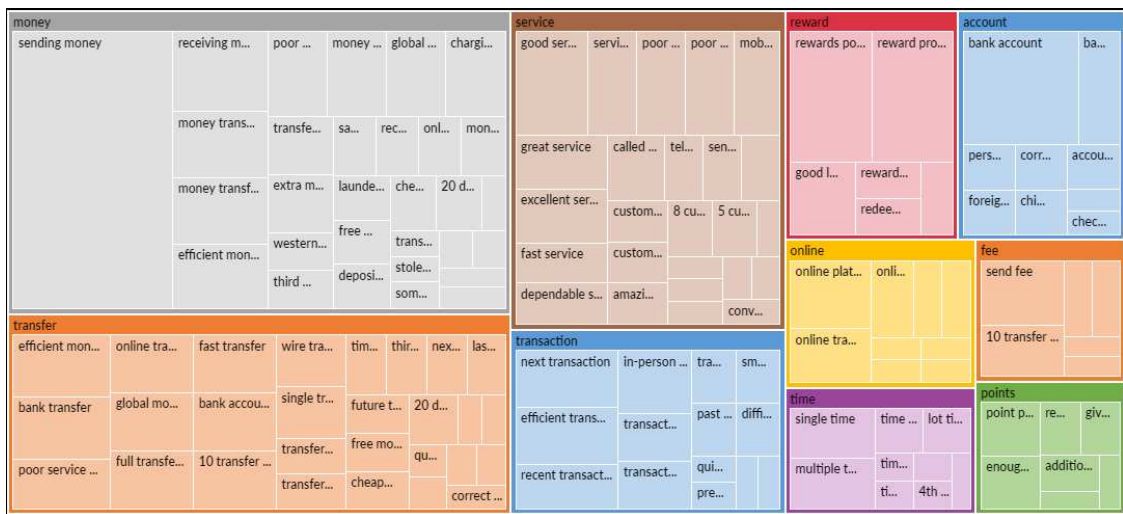


Figura 4.2-3 - Rappresentazione aggiuntiva dell'analisi tematica fornita da NVivo

Infine, oltre all'analisi tematica e del sentimento è stato possibile accedere alla funzione del calcolo della frequenza di parole. Il programma è stato in grado di fornire come output il grafico delle nuvole di parole identificando le dieci parole maggiormente citate. (Fig. 4.2-4).

Word	Length	Count
money	5	164
easy	4	71
send	4	66
transfer	8	62
service	7	55
western	7	55
time	4	51
union	5	50
use	3	49
transaction	11	43



Figura 4.2-4- Conteggio e nuvola delle 10 parole più citate ottenute tramite NVivo

Affrontando, invece, l'applicazione delle tecniche di analisi dati tramite Python sono state prese in considerazione due casistiche di processamento, entrambe con un output in grado di fornire l'analisi del sentimento, il topic modeling e il calcolo della frequenza di parole. La scelta è stata quella, infatti, di presentare due dei differenti metodi nel quale un ricercatore può cimentarsi quando decide di programmare il suo codice per l'analisi tramite il linguaggio Python.

Nel primo programma sono state impiegate librerie che utilizzano algoritmi basati su tecniche di modellazione statistica, regole e lessici per scoprire strutture nascoste nei dati. Questi algoritmi, pur incorporando alcuni principi di apprendimento automatico, non si basano sull'uso di reti neurali. Nel secondo invece si è provato ad utilizzare le reti neurali pre-addestrate per l'analisi del sentimento e per il topic modeling. Il tutto eseguito per fornire una sorta di panoramica delle alternative disponibili per l'esecuzione di una corretta analisi dei dati raccolti supportata da strumenti digitali di elevata potenza di calcolo.

Nella prima casistica si è scelto, in base alle caratteristiche degli strumenti disponibili, di procedere attraverso l'implementazione di VADER per l'analisi dei sentimenti e dell'algoritmo LDA, descritto in precedenza, per la funzione di topic modeling.

VADER (Valence Aware Dictionary and Sentiment Reasoner), presente nella libreria NLTK è uno strumento di analisi del sentimento pre-addestrato, basato su lessico e regole che non incorpora l'apprendimento automatico. Esso utilizza una forma di intelligenza simbolica per elaborare e interpretare il linguaggio umano. Una particolarità di questo

algoritmo è il lessico su cui si basa, il quale è specificamente in sintonia con i sentimenti espressi nei social media" (Hutto & Gilbert, 2014). Esso, infatti, funziona tramite un lessico di sentimenti, che consiste in un elenco completo di caratteristiche lessicali, come parole, etichettate secondo la loro orientazione semantica come positiva o negativa. Quando VADER analizza un testo, cerca la presenza di parole contenute nel suo lessico di sentimenti. Questo metodo consente a VADER di identificare elementi nel testo che indicano il tono emotivo dello scrittore, sia esso positivo o negativo. Inoltre, VADER è attrezzato per gestire il linguaggio informale, multipli segni di punteggiatura, acronimi, persino emoticon nella sua analisi. VADER è uno strumento open source, il che lo rende facilmente integrabile in varie applicazioni e linguaggi di programmazione. Tuttavia, potrebbe non essere efficace su testi formali o con sentimenti complessi e sfumati, e il suo lessico potrebbe richiedere aggiornamenti occasionali (Chadha et. al, 2023).

Per una corretta applicazione di tali strumenti è stato necessario, dopo aver eseguito, come nel caso del software NVivo, la raccolta automatizzata tramite il web scraping, procedere con la propedeutica preelaborazione del testo al fine di ottenere una maggiore precisione nella procedura di analisi. La sequenza eseguita è stata la seguente:

- Conversione in minuscolo: questo è tipicamente il primo passo nella preelaborazione del testo. Convertire tutto il testo in minuscolo aiuta a ridurre la complessità, assicurando che parole identiche non vengano considerate diverse solo a causa delle differenze nella capitalizzazione. Attraverso la funzione “.lower()” è stato possibile eseguire tale operazione.
- Rimozione di caratteri alfanumerici: questo passo aiuta a pulire il testo da tutti i caratteri che potrebbero non contribuire al significato, come la punteggiatura e i caratteri speciali.
- Tokenizzazione: il processo di tokenizzazione trasforma una stringa di testo in una lista di token. Un token è una singola parola o simbolo. Questo passaggio è fondamentale perché permette di analizzare il testo parola per parola.
- Rimozione delle Stop-words: Fase che permette di eliminare tutte le parole comuni che non contribuiscono al significato del testo come per esempio le preposizioni.

Successivamente è stato dunque possibile procedere con l'analisi del sentimento tramite l'algoritmo VADER. Tramite il modulo `SentimentIntensityAnalyzer` viene fornito il punteggio composto⁵⁷, che varia tra -1 (estremamente negativo) e +1 (estremamente positivo), calcolato analizzando la polarità di ogni parola nel testo e sommando questi valori. Questo approccio permette di ottenere una visione generale del sentimento espresso nella totalità delle recensioni. L'analisi del sentimento è stata visualizzata utilizzando le librerie Python "seaborn" e "matplotlib", strumenti di riferimento nel campo della visualizzazione dati. Nello specifico matplotlib fornisce la base per il plotting, mentre seaborn, che è costruita su matplotlib, offre un'interfaccia più semplice e stili grafici migliorati. Questa combinazione ha permesso di generare un istogramma dettagliato, che rappresenta efficacemente la distribuzione dei punteggi di sentimento derivati dall'analisi.

L'istogramma è stato configurato con l'asse orizzontale (asse X) rappresentante l'intervallo dei punteggi di sentimento, estendendosi da valori fortemente negativi a fortemente positivi. L'asse verticale (asse Y) quantifica la frequenza di ogni punteggio di sentimento all'interno del corpus analizzato. Questo approccio grafico fornisce una visione immediata e quantitativa della prevalenza di sentimenti positivi, negativi o neutri nelle recensioni esaminate, come mostra il seguente grafico.

⁵⁷ Tale punteggio è un indicatore che riassume il sentimento generale di un testo. È calcolato come una media ponderata dei punteggi assegnati a ciascuna parola nel testo, tenendo conto non solo delle parole singole, ma anche del loro contesto

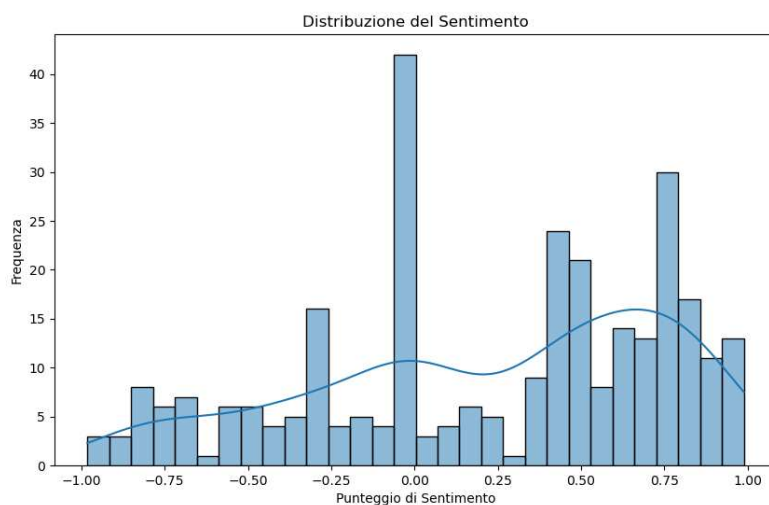


Figura 4.2-5 - Istogramma raffigurante la distribuzione del sentimento calcolata tramite l'algoritmo VADER

Inoltre, per ottenere un'analisi più dettagliata ogni recensione è stata analizzata separatamente per determinare il singolo sentimento associato a ogni recensione. La funzione classifica le recensioni positive o negative in base alla somma del loro punteggio del lessico di VADER. Se superiore a 0.05 viene considerata positiva, se inferiore a -0.05 negative. Tali recensioni sono state poi salvate in due file csv distinti con l'obiettivo di fornire al ricercatore un file per leggere tali recensioni e capirne meglio il contesto e i punti di forza e debolezza del proprio servizio. In totale si sono ottenute 175 recensioni positive, 81 recensioni negative e 43 recensioni considerate neutre. L'approccio usato si concentra sull'analisi di ogni recensione nel suo complesso, piuttosto che dividere la recensione in singole frasi per analizzarle separatamente. Il vantaggio di questo metodo è che considera il contesto globale della recensione, così concentrando il focus su un aspetto che è spesso fondamentale per comprendere appieno il sentimento espresso dall'autore. Tuttavia, analizzando l'intera recensione come un'unità unica, è possibile che un numero maggiore di recensioni venga classificato come neutro. Questo accade perché, in una recensione completa, possono coesistere espressioni positive e negative che, sommate, risultano in un bilanciamento del sentimento generale, portando a una classificazione neutra.

Per il topic modeling si è fatto ricorso alla libreria “gensim” per poter implementare l’algoritmo, di apprendimento automatico non supervisionato, LDA, uno dei più popolari per il topic modeling.

Per implementare efficacemente il modello Latent Dirichlet Allocation (LDA), è stato essenziale non solo pre-elaborare accuratamente i dati, rimuovendo elementi superflui e standardizzando il testo, ma anche sviluppare un dizionario specifico, “dictionary”. Questo dizionario assegna a ciascuna parola del nostro dataset un identificativo numerico unico, facilitando così l’analisi computazionale. Inoltre, il set di dati è stato trasformato in un corpus nel formato “bag of words”⁵⁸. Così facendo i dati sono rappresentati come una serie di coppie (ID parola, frequenza parola), permettendo al modello di analizzare la frequenza delle parole nel dataset. Infine, il modello LDA è stato addestrato sul corpus preparato. In questo passaggio si è costruito un modello basato su cinque temi in cui ogni tema identificato dal modello è una distribuzione di probabilità su tutte le parole nel dizionario, e le parole chiave all’interno di un tema contribuiscono a definirlo con un certo peso. Le parole con la probabilità più alta in un tema sono considerate le più rappresentative di quel tema.

Il modello LDA, nel nostro caso, ha restituito i seguenti cinque temi con le rispettive parole classificate in ordine di rilevanza.

Tema 1	Tema 2	Tema 3	Tema 4	Tema 5
Money	Money	Money	Money	Money
Transfer	Transfer	Time	Transfer	Transfer
Easy	Easy	Send	Service	Send
Service	Time	Transfer	Easy	Time
Send	Send	Easy	Time	Transaction
Time	Service	Use	Send	Easy
Point	Transaction	Transaction	Always	Service

Tabella 4.2-1 - Temi individuati con LDA

⁵⁸ Formato in cui un testo viene visualizzato come un insieme di parole. Tale formato ignora l’ordine delle parole ma cattura la molteplicità.

Come si nota ogni tema è caratterizzato dalle parole chiave che meglio lo descrivono con tali parole ordinate con peso decrescente per lo specifico tema. È facilmente osservabile come i temi identificati siano abbastanza simili tra di loro. Questo può succedere quando i dati analizzati sono molto simili o si concentrano sugli stessi argomenti e di conseguenza i modelli di topic modeling potrebbero avere difficoltà a distinguere i temi. Inoltre, l'algoritmo ha assegnato molte delle seguenti parole con pesi simili dando quindi un'importanza quasi uguale tra tutte le parole dei temi.

Nello specifico tale analisi mostra in particolare come tutti i cinque temi siano indirizzati nel segnalare la facilità del trasferimento dei soldi, la velocità di utilizzo e il tempo necessario per le operazioni di trasferimento e la qualità del servizio.

Per completezza, infine, come per il software NVivo si è eseguito e raffigurato tramite il grafico, il calcolo delle dieci più comuni parole citate nelle recensioni. Questo è stato possibile grazie alla libreria "collections" e alla classe "counter" utile per contare gli oggetti. Mentre per la visualizzazione tramite grafico si fatto ricorso alle librerie seaborn e matplotlib citate in precedenza. In questo caso sull'asse delle ascisse viene rappresentato il conteggio delle parole, mentre sull'asse delle ordinate le parole in oggetto.

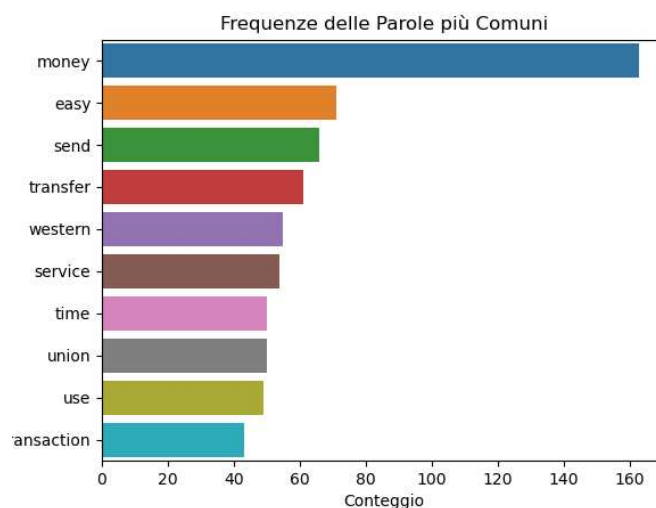


Figura 4.2-6 - Grafico del codice Python delle dieci parole maggiormente citate

Come si può osservare l'output risulta essere perfettamente in linea con quello generato dal software NVivo.

Per la seconda applicazione delle tecniche di analisi dati tramite il linguaggio Python, come anticipato precedentemente, si è fatto ricorso all'utilizzo di reti neurali per il calcolo del sentimento e per il topic modeling. Il dataset di riferimento è rimasto naturalmente invariato da quello utilizzato per il software NVivo e per la prima applicazione di Python. Dopo una breve ricerca sulle possibili reti neurali da utilizzare, la scelta è ricaduta sulla rete pre-addestrata BERT per l'analisi del sentimento e sull'algoritmo BERTopic, che nel suo processo utilizza reti pre-addestrate, per il topic modeling.

BERT (Bidirectional Encoder Representations from Transformers) è una rete neurale pre-addestrata sviluppata da Google. Questa rete si avvale di un'architettura basata sui transformer, che le consente di analizzare il contesto di ogni parola in modo bidirezionale. Ciò significa che BERT è capace di considerare il contesto che precede e segue ogni parola analizzata. Tale capacità permette al modello di catturare tutte le possibili combinazioni di contesto che possono emergere in un testo, rendendolo estremamente efficace nell'eseguire analisi dettagliate. Differenziandosi dalle architetture unidirezionali delle reti neurali ricorrenti (RNN) o convoluzionali (CNN), BERT rappresenta una svolta significativa nell'elaborazione del linguaggio naturale. Questo si traduce anche in una maggiore precisione nell'analisi del sentimento, specialmente in testi lunghi e complessi. Una delle competenze chiave di questa rete è la capacità di gestire parole polisemiche, ovvero termini con più significati. Inoltre, BERT è stato pre-addestrato su un vasto corpus di testo non etichettato, che include l'intera Wikipedia in lingua inglese (circa 2,5 miliardi di parole) e il BookCorpus (800 milioni di parole). L'addestramento di tale rete è avvenuto su due compiti specifici: il modello di linguaggio mascherato⁵⁹ e la previsione della frase successiva⁶⁰. Tuttavia, anche se BERT non è stato direttamente addestrato per l'analisi del sentimento durante il suo pre-addestramento iniziale, la sua comprensione approfondita del linguaggio lo rende estremamente adatto anche per questo compito. Infatti, grazie a un possibile "fine-tuning" che prevede un ulteriore addestramento finale utilizzando un dataset specifico con esempi di testo per un preciso compito, la rete risulta essere in grado di eseguire precise analisi dei sentimenti.

⁵⁹ Capacità di prevedere una parola mancante in qualsiasi posizione della frase

⁶⁰ In questo compito, BERT viene addestrato a prevedere se una frase B è il seguito logico di una frase A

Nel nostro caso è stato eseguito un fine-tuned sulla rete BERT con un dataset specifico di 150mila recensioni in lingua inglese per l'analisi del sentiment, fornito da NLP town⁶¹. Il modello utilizzato per l'ulteriore addestramento prevede di categorizzare recensioni con un punteggio che varia da 1(recensioni estremamente negative) a 5(recensioni estremamente positive). Dunque, dopo aver importato le librerie tra cui NLTK per la preelaborazione del testo effettuando le medesime azioni eseguite in precedenza, e Pytorch per poter caricare il modello BERT è stato possibile procedere con il caricamento del modello di fine-tuning.

Si è poi definita la funzione “analyze_sentiment_bert” in grado di leggere ed elaborare con BERT il file csv contenente le recensioni raccolte. In questo caso, usando il particolare modello fornito da NLP town, l'output restituisce un rating in stelle da 1 a 5 basato sulle probabilità calcolate dal modello. Anche in questo per favorire il ricercatore nell'analisi approfondita delle recensioni aiutandolo a identificare il contesto di alcuni pensieri, si è deciso di salvare le recensioni in due file csv separati: Recensioni positive.csv e Recensioni negative.csv. Nello specifico indirizzando le recensioni con un punteggio di 4 e 5 stelle nel primo file e quelle con punteggio di 1 e 2 stelle nell'altro. In totale si sono ottenute 150 recensioni classificate come positive e 120 classificate come negative.

Anche in questo caso, come si vede nel grafico qui presentato (Fig.4.2-7), creato utilizzando le librerie matplotlib e seaborn, si evidenziano recensioni con punteggio neutro (valore stella 3).

⁶¹ NLP town è una comunità o un'organizzazione che si occupa di progetti e sviluppi nel campo dell'elaborazione del linguaggio naturale (NLP). Nel nostro caso NLP town ci ha fornito il modello “bert-base-multilingual-uncased-sentiment” da utilizzare per l'analisi dei sentimenti.

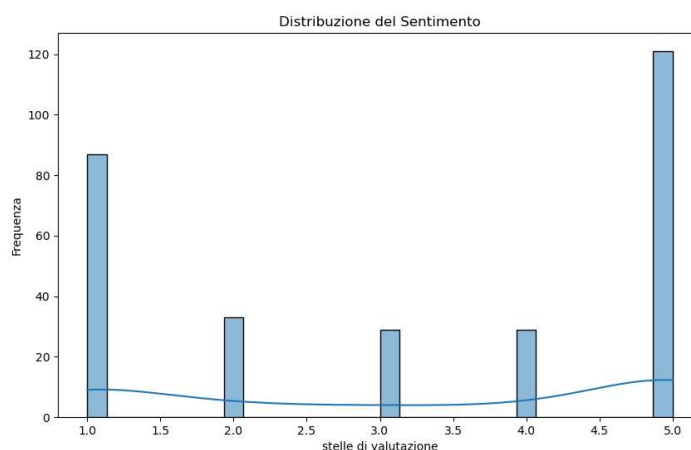


Figura 4.2-7 - Istogramma raffigurante il sentimento calcolato tramite la rete BERT delle recensioni raccolte

Nel contesto dell'analisi dei topic, è stato impiegato l'algoritmo BERTopic, il quale si avvale di reti neurali pre-addestrate, non per l'analisi diretta dei temi, ma per trasformare il testo in rappresentazioni vettoriali. BERTopic è una tecnica di modellazione avanzata degli argomenti che utilizza queste rappresentazioni vettoriali per creare cluster densi e coerenti, facilitando così l'interpretazione degli argomenti. Questi cluster mantengono le parole chiave rilevanti per ogni argomento, rendendo le descrizioni degli stessi facilmente comprensibili. La forza di BERTopic risiede nella sua capacità di sfruttare il potere rappresentativo dei modelli di trasformatori, come BERT, per catturare le relazioni semantiche tra le parole. Questo approccio differisce significativamente dai metodi di analisi dei topic tradizionali, come la Latent Dirichlet Allocation (LDA), che si basano principalmente sul conteggio delle parole e sulla loro distribuzione. Invece, BERTopic si avvale della comprensione semantica profonda fornita dai trasformatori, consentendo di raggruppare i testi in argomenti che non solo condividono termini simili, ma che sono anche semanticamente correlati. Questo risultato si traduce in argomenti più accurati e significativi, che riflettono con maggiore precisione i temi intrinseci nei testi analizzati.

BERTopic costituisce quindi un avanzamento significativo nel campo dell'analisi dei topic, offrendo un metodo più sofisticato e preciso per organizzare e interpretare grandi insiemi di dati testuali.

Nel nostro caso è stato dunque necessario importare la libreria “bertopic” contenente l’algoritmo BERTopic per procedere con l’attuazione del modello. E’ stata definita la funzione “fit.transform” la quale addestra il modello sui dati da noi caricati nel file csv (in questo caso le recensioni) e trasforma tali dati in topic. È stata poi implementata la funzione “get_topic()” che consente di vedere le principali parole chiave con relativo peso che costituiscono ciascun tema.

L’algoritmo ha restituito i seguenti cinque temi con le relative parole chiave indicate in ordine di rilevanza:

Tema 1	Tema 2	Tema 3	Tema 4	Tema 5
Fast	Money	Points	To	Transfers
And	To	Rewards	The	The
Great	Easy	My	My	Days
Service	Send	It	Money	Take
Easy	And	To	They	Instantly
Very	My	You	This	Full
To	Quick	Use	In	Always

Tabella 4.2-2 - Temi individuati con BERTopic

In questo caso come si nota dalle parole chiave elencate, i temi risultano sempre molto simili tra di loro. Questo come è successo anche per l’LDA succede quando i dati tendono a concentrarsi spesso sullo stesso argomento. Tuttavia, a differenza dell’algoritmo LDA sembra che BERTopic sia riuscito a differenziare nel complesso le parole chiave caratterizzanti ciascun tema. Infatti, BERTopic, essendo basato su BERT può catturare meglio i contesti e le sfumature linguistiche delle parole all’interno di set di dati. Questo permette una più accurata differenziazione dei temi, anche quando le parole chiave sono simili tra i vari temi. Inoltre, BERTopic identifica i temi non solo basandosi sulla frequenza di parole, ma anche su come queste vengono utilizzate nel contesto. Di conseguenza, parole funzionali come "to", "my", "they" e "and" appaiono come parte dei temi perché il modello riconosce la loro importanza nel contesto delle frasi.

Nello specifico, il tema 1 suggerisce velocità, qualità e facilità del servizio. Il tema 2 la facilità e velocità dell’invio di denaro. Il tema 3 l’utilizzo dei punti premio. Il tema 4

suggerisce sempre il denaro come fulcro importante con qualche possibile iterazione di lamentela o domanda dell'utente con il servizio (my, this, they). Il tema 5 invece tratta sempre dei trasferimenti di denaro e della velocità.

Naturalmente l'analisi illustrata è stata condotta a titolo semplificativo e dunque con il solo scopo di fornire un esempio di applicazione di alcuni strumenti disponibili per il supporto dell'analisi dati con intelligenza artificiale in una ricerca netnografica. Né si può dimenticare che tali strumenti forniscono solo una piccola automazione nella analisi dei dati. I risultati devono infatti essere rivisti e interpretati dall'essere umano, l'unico in grado di scovare, a partire da questi dati, importanti insight sui consumatori.

4.3. Vantaggi e limitazioni tra software e programmazione

Eseguita l'applicazione del caso delle recensioni di Western Union utilizzando i due strumenti, software e programmazione, e fornendo un'ampia panoramica delle scelte (software, programmazione con algoritmi semplici di IA, programmazione con reti neurali) che un ricercatore può effettuare quando decide di utilizzare l'IA come supporto alla sua attività di analisi dati per una ricerca netnografica, è ora possibile presentare le principali differenze ottenute nei risultati e affermare quali sono i principali vantaggi e limitazioni che si presentano usando tali strumenti. La seguente tabella riassume in breve i risultati, dell'analisi dei sentimenti e dell'analisi tematica, ottenuti dai 3 metodi indicando le principali differenze emerse nell'eseguire l'analisi.

	Software NVivo	Python con VADER e LDA	Python con BERT e BERTopic
Analisi dei sentimenti	Analisi basata su lessico 115 rec. positive 91 rec. negative 93 rec. neutre	Analisi basata su lessico. Algoritmo in grado gestire il linguaggio informale, segni di punteggiatura, acronimi ed emoticon. 175 rec. positive 81 rec. negative 43 rec. neutre	Analisi basata su rete pre-addestrata in grado di catturare il contesto della parola analizzata. 150 rec. positive 120 rec. negative 29 rec. neutre
Analisi tematica	L'autocode ha identificato dieci temi principali caratterizzati da diverse parole chiave. I temi più importanti, in base al numero di occorrenza nei dati sono stati: Money, Service, Transfer e Rewards. Il SW ha offerto la possibilità di visualizzare le singole parole e frasi maggiormente ripetute in ciascun tema consentendo così di visualizzare meglio il contesto. Nell'eseguire tale analisi si è rivelato molto preciso nel distinguere i diversi temi con una chiara definizione delle parole chiave per ciascun argomento.	L'algoritmo LDA ha identificato cinque temi molto simili tra loro, tutti caratterizzati da parole come Money, Transfer, Easy, Send e Time. In questo caso l'algoritmo non è stato in grado di differenziare bene i temi e di assegnare in modo distintivo i pesi di ciascuna parola chiave. La maggior parte delle parole chiave del tema è costituita dallo stesso peso. L'analisi manca anche di parole funzionali, importanti per il contesto come "my", "they", "and". Tuttavia, i risultati pongono l'accento sul trasferimento di soldi, sul servizio, sul tempo e sui punti proprio come nel software.	L'algoritmo BERTopic ha identificato cinque temi principali distinguendosi per l'inclusione di parole chiave diverse rispetto a quelle selezionate da LDA. Anche in questo caso però, tutti i temi condividono un focus sulla velocità del servizio, sull'invio dei soldi, trasferimenti, punti e ricompense. Particolarmente evidenti sono termini come Money, Fast, Transfers, Points e Rewards. Inoltre, BERTopic ha identificato anche parole come "my", "and", "to", "they" permettendo così di comprendere come le parole chiave si relazionano all'interno del contesto. L'algoritmo è stato dunque più preciso rimanendo coerente ai temi identificati sia dal software sia da LDA.

Tabella 4.3-1 - Confronto risultati delle due analisi con strumenti e metodi differenti

In conclusione, considerando i dati riportati nella tabella, si può affermare che tutti gli strumenti utilizzati riportano, nell'analisi del sentimento, una prevalenza di recensioni positive, evidenziando tuttavia anche una cospicua percentuale di recensioni negative.

In particolare, il software NVivo ha identificato un numero maggiore di recensioni classificate come neutre rispetto a quanto rilevato dagli algoritmi utilizzati nei frameworks di Python. Questo molto probabilmente dovuto alla mancanza di comprensione del contesto e della gestione del linguaggio informale e punteggiatura. Nonostante ciò, un'analisi comparativa delle singole valutazioni positive e negative tra i due metodi ha rivelato una sostanziale similitudine nei risultati. Le variazioni rilevate possono essere attribuite alle specifiche metodologie di elaborazione intrinseche degli algoritmi impiegati, che influenzano la classificazione dei sentimenti. In termini di precisione, la rete BERT risulta essere più accurata, in quanto in grado di comprendere meglio il contesto e le relazioni tra le parole. VADER segue con una precisione moderata-alta, soprattutto per testi informali e linguaggio colloquiale. NVivo, sebbene utile, può essere meno preciso degli altri metodi a causa della sua dipendenza da un lessico statico e dalla mancanza di analisi contestuale.

Per quanto concerne l'analisi tematica invece, nonostante le notevoli differenze in termini di precisione di analisi, vi è stata, una focalizzazione sui temi riguardanti il trasferimento dei soldi, il servizio, la velocità e i punti/ricompense. Questo ha indicato una coerenza tra gli strumenti nell'identificare i temi principali dei dati analizzati. Nello specifico il software NVivo e l'uso dell'algoritmo BERTopic si sono rivelati maggiormente precisi nell'identificare il contesto e nel differenziare le parole chiave dei diversi temi ottenuti. L'algoritmo LDA invece concentrandosi di più sulle parole con alta frequenza, ha ignorato il contesto più ampio in cui le parole appaiono. Così facendo nelle parole caratterizzanti i temi sono venute a mancare le parole funzionali "my", "they", "and" in quanto considerate "rumore" piuttosto che come parti significative dei temi.

È possibile infine affermare che tutti e tre i metodi, oltre a soddisfare le diverse situazioni cui si possono trovare i ricercatori in termini di dati, obiettivi, risorse e tempo, si sono rivelati molto efficaci e coerenti tra di loro con una difficoltà variabile nell'interpretare i risultati a causa della differenza di precisione presente.

Da questi risultati ottenuti è quindi possibile concludere che il servizio fornito da Western Union è considerato dalla maggior parte degli utilizzatori ben strutturato e di semplice utilizzo. Tuttavia, il numero consistente di recensioni negative evidenzia la presenza di possibili aspetti critici, non indifferenti, individuati dagli utenti. A fronte di tali sentimenti negativi, tale analisi potrebbe dunque indirizzare i manager di Western Union ad effettuare specifiche azioni/strategie di miglioramento focalizzate negli argomenti/temi individuati.

Si rende ora necessario analizzare i principali vantaggi e svantaggi nell'utilizzare i due strumenti: il software e la programmazione.

Il principale vantaggio nell'utilizzare il software è la facilità di utilizzo. Infatti, esso è progettato anche per chi non ha competenze tecniche di programmazione ed è spesso progettato appositamente per la ricerca qualitativa, offrendo anche un'interfaccia utente grafica intuitiva, che permette agli utenti di condurre analisi senza la necessità di scrivere codice. Tale interfaccia garantisce inoltre velocità e accuratezza nell'usare gli strumenti disponibili. I software inoltre supportano il caricamento agevolato di file multimediali come video, audio e immagini oltre che testuali permettendo una codifica, una organizzazione e una categorizzazione veloce di grandi set di dati.

Alcuni software eseguono automaticamente, dato un testo, la ricerca di specifiche parole o categorie di parole come il prezzo, il luogo, i nomi etc. ed evidenziano con colori differenti tali parole. Si tratta di un metodo visuale per semplificare la codifica e l'analisi dei testi, consentendo agli analisti di identificare rapidamente le informazioni rilevanti.

Un ulteriore vantaggio nell'usare i software è l'analisi interattiva in modo semplice che il ricercatore può fare esplorando i dati visualizzati nei grafici. È infatti possibile filtrare dati, cambiare tipologia di grafico con un semplice clic e con estrema velocità senza dover agire sul codice di programmazione. Un esempio di queste rappresentazioni visive è stato fornito dal software NVivo nel quale ci ha mostrato oltre alla nuvola di parole anche un grafico a quadri contenente i temi principali individuati con la dimensione dei quadri differente in base all'importanza del tema. Il tutto effettuato tramite l'auto codifica tematica che ha consentito in pochi minuti di ottenere il risultato. In aggiunta tali rappresentazioni visive possono essere utilizzate per presentare i risultati della ricerca a un pubblico più ampio, inclusi membri del team di ricerca non tecnici o stakeholder esterni.

Rispetto alla programmazione Python però, alcuni software hanno meno flessibilità in termini di personalizzazione dell'analisi e i ricercatori devono così adeguarsi alle sole funzioni offerte.

Infine, è utile ricordare che i software richiedono un piano a pagamento, spesso molto costoso; fattore che può rivelarsi un ostacolo per alcuni ricercatori o istituzioni.

La programmazione tramite Python risulta essere, contrariamente ai software, uno strumento gratuito che offre ai ricercatori elevata flessibilità e personalizzazione di algoritmi e strumenti durante l'analisi dati adeguandosi così perfettamente a tutte le esigenze dello studio. Un ricercatore può infatti scegliere quali funzioni integrare personalizzando così completamente l'analisi senza alcuna limitazione. Python inoltre consente la possibilità di scegliere tra un'ampia varietà di approcci e strumenti, dai semplici algoritmi che utilizzano lessici e regole fino alle più avanzate reti neurali. Questa varietà offre agli utenti la libertà di scegliere la soluzione più adatta alle loro competenze e ai tipi di dati con cui lavorano, permettendo così un alto grado di personalizzazione e precisione nell'analisi dei dati.

L'ampia comunità online e l'abbondanza di risorse didattiche, che vanno dai tutorial pratici a corsi online approfonditi, facilitano l'apprendimento e l'approfondimento delle competenze in Python, rendendolo un linguaggio accessibile a un vasto pubblico e costantemente aggiornato. La natura open source e la sua continua evoluzione, guidata da una comunità globale di sviluppatori, ne fanno una scelta ideale per ambienti accademici e professionali, permettendo una personalizzazione senza precedenti e una scalabilità adatta a progetti di qualsiasi dimensione.

L'utilizzo di Python, sebbene estremamente versatile e potente, presenta alcune sfide non trascurabili. Come prima cosa necessita infatti di competenze tecniche di programmazione e la curva di apprendimento per tale compito può essere particolarmente ripida. Questo aspetto è dovuto alla natura del linguaggio, che richiede una comprensione delle logiche di programmazione e dei principi di base dell'analisi dei dati.

Certamente la configurazione iniziale di un ambiente di sviluppo Python può essere più complessa rispetto all'utilizzo di software specifici come NVivo. Python, infatti, richiede

una configurazione manuale dell'ambiente di sviluppo, l'installazione di librerie pertinenti e una buona comprensione del funzionamento dei vari strumenti e pacchetti.

Risulta importante sottolineare poi che, pur essendo Python un linguaggio di programmazione privo di un'interfaccia grafica dedicata all'analisi dei dati, offre la possibilità di creare visualizzazioni di dati avanzate tramite l'uso di librerie esterne come “matplotlib” e “seaborn”. Queste librerie permettono agli utenti di generare una vasta gamma di grafici e visualizzazioni interattive, compensando così la mancanza di un'interfaccia grafica nativa.

Un altro aspetto da considerare è la personalizzazione degli algoritmi in Python, che richiede competenze tecniche avanzate. La creazione e l'ottimizzazione di algoritmi, specialmente quelli basati su reti neurali e tecniche di machine learning, possono essere complesse e richiedono una conoscenza approfondita sia del linguaggio di programmazione che dei principi sottostanti l'apprendimento automatico. Queste tecniche possono anche necessitare di risorse computazionali significative, come potenza di calcolo e memoria, soprattutto per l'elaborazione di grandi dataset o per l'addestramento di modelli complessi. In conclusione, mentre Python offre un'incredibile flessibilità e una vasta gamma di strumenti per l'analisi dei dati, il suo utilizzo richiede un impegno significativo in termini di apprendimento e sviluppo delle competenze tecniche. Tuttavia, grazie alla sua vasta comunità e alle numerose risorse didattiche disponibili, anche i principianti possono gradualmente acquisire le competenze necessarie per sfruttare appieno il potenziale di questo potente strumento di programmazione.

Conclusioni

Questo viaggio all'interno delle principali metodologie di comprensione dei consumatori nell'era del digitale volge ora alla conclusione. Per poter trarre alcune considerazioni ulteriori, è utile ripercorrere le principali tappe che sono state affrontate. Nello specifico si sono analizzate due metodologie appartenenti alla etnografia digitale: la netnografia e l'etnografia mobile.

Il primo capitolo, oltre ad esplorare il metodo della netnografia si occupa di fornire una breve presentazione del metodo predecessore ad esso, l'etnografia, e a fornire di conseguenza le principali motivazioni che hanno consentito l'evoluzione di tale metodo. Il WEB 2.0, il quale ha permesso l'introduzione della comunicazione mediata da computer (CMC) ha trasformato completamente l'interazione online, consentendo agli utenti di partecipare attivamente alla creazione e alla condivisione di contenuti. Sono nate così le rinomate comunità online ovvero gruppi di persone con interessi comuni che condividono le loro esperienze e pensieri in spazi digitali quali blog, forum, social media. A fronte di questi nuovi gruppi digitali il metodo della netnografia prevede di utilizzare tecniche di osservazione naturalistiche non intrusive al fine di ottenere importanti informazioni sui consumatori a partire dalle esperienze condivise da loro stessi. L'impiego di queste ricerche risiede dunque, nel bisogno di identificare preziosi insight sui consumatori, che risultano essenziali nel mondo aziendale per guidare strategie di innovazione prodotto/servizio e per le attività di marketing. Tale capitolo si è soffermato inoltre nel presentare i sei step principali necessari ad una corretta applicazione del metodo secondo la metodologia elaborata dal professore Robert Kozinets. Si è dunque affrontata la fase di pianificazione, caratterizzata dalla definizione degli obiettivi e degli argomenti chiave della ricerca al fine di ottenere un ingresso adeguato nella comunità; la fase di entrée volta all'identificazione del luogo di analisi della ricerca; la fase di raccolta dati presentando le due metodologie disponibili per la raccolta (manuale o automatica) e

classificando le diverse tipologie di utenti che condividono informazioni; la fase di analisi e interpretazione dati indicando alcune tipologie di analisi, l'importanza della triangolazione con metodi etnografici tradizionali e l'importanza di una comprensione profonda da parte del ricercatore nei dati ottenuti per la generazione di insight rivoluzionari. Successivamente è stata presentata la fase riguardante il rispetto degli standard etici e del confronto con gli utenti delle comunità dei risultati finali ottenuti al fine di effettuare un doppio controllo e validare così gli esiti. Infine, si sono analizzati in dettaglio i principali limiti e vantaggi che caratterizzano una ricerca netnografica.

L'obiettivo è stato, dunque, fornire una sorta di presentazione e guida operativa per chiunque volesse mettere in atto un processo di ricerca netnografica.

Il secondo capitolo identifica invece come gli strumenti basati sull'intelligenza artificiale possano fornire un valido supporto nel processo di ricerca netnografica. Dopo l'analisi delle principali motivazioni che hanno portato e consentito l'uso di questa tecnologia nella netnografia, al fine di una più completa comprensione, si è proceduto a considerare i principali approcci esistenti di intelligenza artificiale utili a tale metodologia antropologica. L'apprendimento automatico, l'apprendimento profondo con reti neurali e l'elaborazione del linguaggio naturale (NLP) si sono rivelati fondamentali, offrendo algoritmi e reti neurali utili ad analizzare e interpretare grandi quantità di dati online. È necessario precisare che la presentazione di tali approcci è rimasta ad un primo livello di approfondimento, trattandosi di materia molto ampia e articolata. Qui sono stati messi in evidenza gli elementi salienti, tali da arricchire la prospettiva e l'osservazione a 360° del tema in oggetto. Si è successivamente passati ad una maggiore concretezza di applicazione, presentando le fasi netnografiche maggiormente supportate da tale tecnologia e identificando alcuni dei principali algoritmi. Analizzando le diverse fasi si è giunti alla conclusione di quanto l'intelligenza artificiale risulti molto utile e, anzi, ormai fondamentale nella fase di raccolta dati e soprattutto nelle diverse analisi dei dati. Si è quindi introdotto lo strumento di text mining, che si fonda su una fase iniziale di pre-elaborazione dei testi mediante tecniche di elaborazione del linguaggio naturale. Questo processo consente di ottenere dati più strutturati, migliorando significativamente la qualità, l'efficienza e l'efficacia della raccolta dati. La procedura di text mining prosegue con l'impiego di algoritmi che combinano sia le tecniche di elaborazione del linguaggio naturale (NLP) che l'apprendimento automatico, incluso quello profondo. Questo

approccio permette di condurre analisi dettagliate come l'analisi del sentimento, l'analisi tematica e persino l'analisi visiva di immagini o video. Infine, il capitolo ha identificato le principali sfide e considerazioni etiche associate all'uso dell'IA. Si è dunque fatto riferimento alle problematiche dovute all'applicazione del web scraping e ai problemi etici dovuti ai bias che gli algoritmi di IA possono riportare durante il proprio addestramento. Si è inoltre posto l'accento sull'incapacità delle macchine di generare intuizioni rivoluzionare a partire dai comportamenti e dalle preferenze dei consumatori e si è indicata l'importanza di considerare questi strumenti basati sull'IA solo un utile supporto all'azione umana.

Lo step successivo si è concretizzato nell'affrontare, nel terzo capitolo, l'etnografia mobile. Si tratta di un ulteriore metodo avanzato per comprendere al meglio i consumatori, emerso in risposta all'evoluzione tecnologica della società, permettendo di osservare e analizzare i comportamenti dei consumatori in contesti dinamici e reali. Basata sull'uso di dispositivi mobili come smartphone, tablet e laptop tale tecnica consente di eseguire una ricerca etnografica a distanza senza l'intrusione eccessiva da parte del ricercatore. Elemento distintivo di tale metodo è il ruolo che ottiene l'utente intervistato, il quale dovendo raccontare, tramite la pubblicazione di post contenenti testo, immagini, video o audio, o tramite sondaggi specifici, le proprie esperienze di prodotto o servizio, diventa lui stesso un co-ricercatore. In questo caso il ricercatore ha dunque il fondamentale compito di generare task specifici per gli intervistati e preparare domande adeguate alla tipologia di utenti che andrà ad intervistare. Anche per questa tecnica è stata presentata la metodologia di applicazione evidenziando i principali step che un ricercatore si trova ad effettuare nel momento in cui decide di procedere con tale ricerca. Partendo dalla definizione dell'obiettivo si è passati alla fase più importante della ricerca ovvero la selezione dei partecipanti (o co-ricercatori). In questo contesto, è stata data particolare attenzione anche agli aspetti legali, assicurandosi il pieno rispetto delle normative vigenti. È seguita poi la fase di selezione dell'applicazione da utilizzare con l'identificazione delle specifiche da considerare nella scelta. Concluse le fasi di preparazione della ricerca, si è descritto il processo di raccolta e la successiva analisi. Anche in questo caso si è visto che, per una maggiore accuratezza dei risultati, è possibile procedere con una triangolazione dei dati mediante interviste di follow-up. Per l'analisi, proprio come nel caso della netnografia, è possibile far fronte a strumenti sia manuali che automatici

supportati anche dall'intelligenza artificiale. A supporto di tale attività molte applicazioni offrono strumenti di analisi integrati che facilitano tale operazione. Il capitolo si è poi concentrato nel definire i principali punti di forza e limiti del metodo con la successiva presentazione di due dei principali fornitori di applicazioni mobili disponibili, con l'obiettivo di fornire una panoramica completa del metodo e delle soluzioni a cui in ricercatore può attingere per lo svolgimento di tale ricerca.

Infine, con le informazioni teoriche acquisite, è stato possibile, nel capitolo quattro, fornire una dimostrazione pratica di alcuni strumenti supportati dall'IA applicati a un caso reale. Lo scopo è stato quello di rendere noti strumenti a pagamento e non che un ricercatore può utilizzare quando decide di integrare l'intelligenza artificiale come supporto a una ricerca netnografica. Dopo una presentazione di alcuni software e del linguaggio di programmazione Python come soluzione gratuita si è passati alla sperimentazione pratica. Il software NVivo ha fatto da referente per l'opzione a pagamento mentre il linguaggio di programmazione con Python si è focalizzato nell'eseguire l'analisi sia tramite algoritmi tradizionali basati su lessico e regole e algoritmi di apprendimento automatico sia utilizzando reti neurali pre-addestrate. Si sono analizzate così le recensioni della Western Union fornite dal sito mondiale di recensioni TrustPilot. L'intento è stato quello di fornire una panoramica completa delle possibilità di analisi disponibili per il ricercatore in base alla propria situazione di dati e risorse in cui si trova. Per entrambi i metodi si è eseguita l'analisi del sentimento, l'analisi tematica e per completezza anche l'analisi della frequenza di parole.

Tale elaborato ha permesso di delineare le principali differenze che la netnografia e l'etnografia mobile presentano e l'influenza dell'IA nell'analisi netnografica. Nonostante possano sembrare molto simili, quasi "gemelle", analizzando in modo approfondito il funzionamento di ognuna, emerge come siano solamente "sorelle", avendo caratteristiche strutturali parzialmente diverse.

Infatti, il primo punto di discriminazione è l'approccio alla ricerca. La netnografia è un metodo di ricerca qualitativa che consente una comprensione approfondita e contestuale dei dati derivati da ambienti virtuali quali siti web, blog, forum e social media. In tale metodologia quindi il focus sono le piattaforme online e si utilizzano i dati disponibili in quelle sedi per analizzare e comprendere al meglio le culture, le opinioni e i comportamenti dei

consumatori. Dunque, è giustamente categorizzata come un'estensione dell'etnografia, adattata al contesto online. L'etnografia mobile invece, nonostante mantenga, come la netnografia, l'obiettivo di identificare i comportamenti e i pensieri dei consumatori, presenta un campo di applicazione parzialmente diverso. Infatti, tale metodo attraverso l'impiego dei dispositivi mobili, utilizzati da ormai la gran parte della popolazione, si focalizza nell'osservare e interagire con i consumatori direttamente nel proprio mondo fisico. Questa osservazione e interazione consente di seguire i consumatori nella loro vita quotidiana, scoprendo con una maggiore comprensione del contesto e delle motivazioni personali, come essi interagiscono con i prodotti, i servizi e come vivono le esperienze in diversi contesti. Proprio per la sua attinenza al modo tangibile, reale, se sono previste frequenti interazioni con gli intervistati tale metodo tende ad essere più invasivo di quello netnografico. Anche se gli utenti selezionati acconsentono a essere osservati durante il periodo di ricerca, ricevendo anche un incentivo monetario per la loro partecipazione, la loro consapevolezza di essere osservati potrebbe influenzare involontariamente i propri comportamenti, rendendoli in parte diversi dalle loro abitudini e opinioni quotidiane. Questo rischio appare meno presente nella ricerca netnografica in cui, in particolar modo se si sceglie un approccio passivo, i ricercatori non intervengono, diminuendo così il rischio di ottenere informazioni meno veritiere e sotto pressione. È dunque possibile affermare che la differenza sostanziale tra le due tecniche risiede nella metodologia di raccolta e del contesto in cui tali metodi vengono applicati. Mentre l'etnografia mobile implica una partecipazione attiva e consapevole dei soggetti nello studio, la netnografia si basa su osservazioni più discrete e spesso anonime in ambienti online.

Sorge a questo punto spontanea la domanda su quale sia la modalità preferibile. Per parte mia ritengo corretto affermare che la risposta non possa che essere legata allo scopo per cui la ricerca viene apprestata. Così, per esempio, nel contesto dello sviluppo di un nuovo prodotto, l'etnografia mobile si rivela essere la scelta più adatta per fornire insight mirati sulle opinioni dei consumatori. Questo metodo, infatti, permette di ottenere risultati specifici e contestualizzati sottoponendo un gruppo target, accuratamente selezionato, di intervistati a ricerche mirate eseguite nel loro ambiente naturale quotidiano. Ancora, l'etnografia mobile si dimostrerà più efficace della netnografia nell'identificare specifiche ideologie dei consumatori, poiché basata sull'acquisizione diretta di feedback mirati da parte dei consumatori stessi. Pur essendo possibile anche nella netnografia interagire

direttamente con i consumatori richiedendo loro di esprimere pensieri o valutazioni specifiche, si incontra spesso il limite della loro disponibilità. Questo è dovuto alla natura pubblica dei forum online, alla riluttanza a condividere opinioni dettagliate in ambienti aperti, o alla semplice mancanza di interesse o tempo.

D'altra parte, nell'identificare opinioni di determinati servizi o prodotti, o pensieri che i consumatori possono avere riguardo un brand, la netnografia, a mio parere, emerge come strumento particolarmente efficace. Il suo punto di forza risiede nella capacità di raccogliere commenti autentici che emergono spontaneamente nelle conversazioni online, senza il bisogno di una sollecitazione da parte del ricercatore. Il vantaggio principale della netnografia risiede nel fatto che i consumatori spesso si sentono più a loro agio nell'esprimersi liberamente in ambienti online, dove la conversazione avviene tra pari, piuttosto che in presenza diretta di ricercatori. Inoltre, negli spazi virtuali di oggi quasi ogni prodotto, servizio o brand viene commentato e discusso all'interno di varie comunità online, rendendo relativamente semplice per i ricercatori trovare le informazioni di cui hanno bisogno. Importante, inoltre, è il ruolo che ricoprono i social media di oggi i quali attraverso influencer e non, hanno ampliato e facilitato ancora di più la possibilità di identificare opinioni e pensieri su temi specifici.

In sintesi, mentre l'etnografia mobile è ideale per ottenere una comprensione profonda del comportamento dei consumatori in contesti specifici e fisici, la netnografia è più adatta per esplorare ampie tendenze e opinioni in ambienti digitali.

In un'epoca in cui la vita digitale è sempre più pervasiva e complessa è stato necessario rendere noto anche come l'intelligenza artificiale possa essere utilizzata per comprendere meglio le comunità online e i comportamenti digitali fornendo un contributo importante alla netnografia. Si tratta, come è stato anche dimostrato nel caso della Western Union, di applicare potenti algoritmi di raccolta e analisi dati in grado di decifrare automaticamente e con estrema velocità alcune informazioni all'interno di grandi dataset di dati. Dall'applicazione e confronto eseguiti tra il software e la programmazione dei due codici differenti è stato possibile affermare che la scelta ideale dipende da diversi fattori quali la tipologia di dati da analizzare, lo scopo della ricerca, le capacità di programmazione del ricercatore, le capacità computazionali disponibili e il tempo disponibile. Per approfondire la panoramica degli strumenti di analisi disponibili, è ora possibile

riassumere le diverse situazioni nel quale un ricercatore può trovarsi e suggerire la scelta migliore.

In un contesto in cui il ricercatore si trovi ad analizzare dati di grande complessità e dimensione con solide competenze in programmazione e avendo accesso a risorse computazionali avanzate, l'impiego di un codice Python basato su reti neurali rappresenta l'opzione privilegiata.

D'altra parte, qualora l'oggetto di studio fosse costituito principalmente da dati testuali e si presentasse la necessità di un controllo più diretto e dettagliato del processo analitico, o in situazioni dove le risorse computazionali fossero più limitate, l'adozione di un codice Python che integra algoritmi basati su lessico e regole risulta la scelta più adeguata.

Invece, in presenza di scenari, in cui le competenze di programmazione siano meno avanzate o in cui il tempo sia un fattore decisivo, l'opzione preferibile sarebbe l'utilizzo di un software a pagamento. Questi strumenti offrono una soluzione pratica e immediata, con l'ulteriore vantaggio di fornire supporto tecnico e documentazione, facilitando l'analisi anche a chi non possiede una profonda conoscenza tecnica.

Tale confronto, tra metodi di analisi con IA, oltre ad offrire una sorta di guida ha anche delineato la potenza di tali strumenti basati sull'intelligenza artificiale e i limiti che tale tecnologia comporta. Infatti, se da un lato si rivela potente e sicura nell'identificare contenuti nei testi, dall'altro non riesce ancora a svolgere l'importante azione di individuazione di insight sui consumatori, nonché il cruciale obiettivo della netnografia. L'intervento umano, dunque, rimane centrale e fondamentale per interpretare e comprendere pienamente i comportamenti e le culture digitali. Pertanto, l'IA raffigura come uno strumento di supporto e non come sostituto all'uomo. Da non dimenticare infine è il tema della privacy nell'uso dell'IA, che rappresenta una delle più interessanti sfide future per la netnografia e, in generale, per tutto ciò che ci circonda in un mondo in continua evoluzione. Con l'aumento dei dati raccolti, diventa infatti cruciale considerare le implicazioni etiche e legali legate alla privacy degli utenti. È dunque fondamentale che gli strumenti di IA siano ulteriormente perfezionati al fine di garantire il rispetto delle normative sulla protezione dei dati e della privacy individuale, garantendo che l'analisi dei comportamenti online non violi i diritti degli utenti.

Si può legittimamente concludere, allora, che la sfida futura consisterà nel bilanciare l'efficienza degli strumenti di intelligenza artificiale e il rispetto della privacy degli utenti con la comprensione intuitiva e contestualizzata della mente umana. Questo equilibrio sarà fondamentale per sfruttare al meglio le capacità dell'IA, mantenendo al contempo l'eticità e l'integrità della ricerca.

Bibliografia

- Babbie, E., & Buzzi, C. (2010). *Ricerca sociale*. Roma, Italy: Apogeo.
- Birenboim, A., Reinau, K. H., Shoval, N., & Harder, H. (2015). High-resolution measurement and analysis of visitor experiences in time and space: The case of Aalborg zoo in Denmark. *The Professional Geographer*, 67(4), 620-629.
- Bosio, B., Rainer, K., & Stickdorn, M. (2017). Customer experience research with mobile ethnography: A case study of the Alpine destination Serfaus-Fiss-Ladis. In *Qualitative consumer research* (pp. 111-137). Emerald Publishing Limited.
- Burgess, A. (2017). *The Executive Guide to Artificial Intelligence: How to identify and implement applications for AI in your organization*. Springer.
- Canvas8, (2020). The Critical Role of Reviews in Internet Trust.
- Chadha, R. & Chaudhary, A. (2023). A Study Analyzing an Innovative Approach to Sentiment Analysis with VADER. *Journal of Engineering Design and Analysis*. 6. 23-27.
- Corbetta, P. (2014). *Metodologia e tecniche della ricerca sociale* (pp. 283-316). Bologna: il Mulino.
- Cribbett S. (2015). Using mobile ethnography to understand what people do, think and feel.
- De Biase L. (2023). L'Intelligenza artificiale e i nodi dell'approvazione di un regolamento europeo. *Il sole 24ore*.
- Delua, J. (2021). Supervised vs. unsupervised learning: What's the difference. *IBM Analytics*, 12.
- Fornari, D., & Ghezzi, A. I. (2023). *Retailing & Trade Marketing: Le sfide commerciali nell'epoca della complessità*. EGEA spa.
- Gent, E. (2015). AI: Fears of playing God!. *Engineering & Technology*, 10(2), 76-79.

- Hasan, M. R., Maliha, M., & Arifuzzaman, M. (2019, July). Sentiment analysis with NLP on Twitter data. In *2019 international conference on computer, communication, chemical, materials and electronic engineering (IC4ME2)* (pp. 1-4). IEEE.
- Hein, W., O'Donohoe, S., & Ryan, A. (2011). Mobile phones as an extension of the participant observer's self: Reflections on the emergent role of an emergent technology. *Qualitative Market Research: An International Journal*, *14*(3), 258-273.
- Heinonen, K., & Medberg, G. (2018). Netnography as a tool for understanding customers: implications for service research and practice. *Journal of Services Marketing*, *32*(6), 657-679.
- Hobbs, D., & Wright, R. (2006). *The SAGE handbook of fieldwork*. SAGE Publications Ltd, <https://doi.org/10.4135/9781848608085>
- Howard, R. (1993). *The virtual community: Homesteading on the electronic frontier*. Mass: Addison Wesley, 1, 993.
- Kamboj, S., & Rahman, Z. (2017). Understanding customer participation in online brand communities: Literature review and future research agenda. *Qualitative Market Research: An International Journal*, *20*(3), 306-334.
- Kozinets, R. V. (1998). On netnography: Initial reflections on consumer research investigations of cyberculture. *ACR North American Advances*.
- Kozinets, R. V. (1999). E-tribalized marketing?: The strategic implications of virtual communities of consumption. *European management journal*, *17*(3), 252-264.
- Kozinets, R. V. (2002). The field behind the screen: Using netnography for marketing research in online communities. *Journal of marketing research*, *39*(1), 61-72.
- Kozinets, R. V. (2010). *Netnography: Ethnographic Research Online*. Sage Publications Limited.
- Kozinets, R. V. (2015). *Netnography: redefined*. Sage.
- Kozinets, R. V. (2018, January 10). Invasion of the PR-bots. USC Relevance Report. Retrieved from <https://annenbergl.usc.edu/research/center-public-relations/usc-annenbergl-relevance-report/invasion-pr-bots>

- Kozinets, R. V., & Gambetti, R. (Eds.). (2020). *Netnography unlimited: Understanding technoculture using qualitative social media research*. Routledge.
- Kozinets, R. V., & Vagni, T. (2016). *Il culto di "Star trek": media, fan e netnografia*. F. Angeli.
- Krotov, V., & Silva, L. (2018). Legality and Ethics of Web Scraping. the Twenty-fourth Americas Conference on Information Systems. *New Orleans*.
- LeCun, Y., Bengio, Y., & Hinton, G. (2015). Deep learning. *nature*, 521(7553), 436-444.
- Lee, K. F., & Qiufan, C. (2023). *AI 2041: scenari dal futuro dell'intelligenza artificiale*. Luiss University Press.
- Loh, C. E., Sun, B., & Lim, F. V. (2023). 'Because I'm always moving': a mobile ethnography study of adolescent girls' everyday print and digital reading practices. *Learning, Media and Technology*, 1-20.
- Mccarthy K. D. (2022), From Traditional Research to Mobile Ethnography. *Dscout*
- Meo, A. I. (2010). Picturing students' habitus: The advantages and limitations of photo-elicitation interviewing in a qualitative study in the city of Buenos Aires. *International Journal of Qualitative Methods*, 9(2), 149-171.
- Milligan, C., & Bartlett, R. (2015). *What is Diary Method?*. Bloomsbury Academic.
- Moisander, J., Närvänen, E., & Valtonen, A. (2020). Interpretive marketing research: Using ethnography in strategic market development.
- Murphy E. (2020), Your First Mobile Ethnography Project - A Definitive Guide. *Indeemo*
- Řehůřek, R., & Sojka, P. (2010). Software framework for topic modelling with large corpora.
- Russell, S. J., & Norvig, P. (2010). *Artificial intelligence a modern approach*. London.
- Stanoevska-Slabeva, K., & Schmid, B. F. (2001, January). A typology of online communities and community supporting platforms. In *Proceedings of the 34th Annual Hawaii International Conference on System Sciences* (pp. 10-pp). IEEE.

- Stickdorn, M., & Frischhut, B. (Eds.). (2012). *Service design and tourism: Case studies of applied research projects on mobile ethnography for tourism destinations*. BoD–Books on Demand.
- Strandvik, T., Holmlund, M., & Edvardsson, B. (2012). Customer needing: a challenge for the seller offering. *Journal of Business & Industrial Marketing*, 27(2), 132-141.
- Taylor, T. L., Boellstorff, T., Nardi, B., & Pearce, C. (2013). *Ethnography and virtual worlds: A handbook of method*. Princeton university press.
- Zanzotto, F. M. (2019). Human-in-the-loop artificial intelligence. *Journal of Artificial Intelligence Research*, 64, 243-252.
- Zhai, X., & Krajcik, J. (2022). Pseudo AI bias. *arXiv preprint arXiv:2210.08141*.
- Zhang, E. Y., Wang, D., & Lei, S. I. (2021). Mobile Ethnography in Tourism and Hospitality: Concept, Tools, and Applications. In *Handbook of e-Tourism* (pp. 1-21). Cham: Springer International Publishing.
- Zhang, L., Wang, S., & Liu, B. (2018). Deep learning for sentiment analysis: A survey. *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery*, 8(4), e1253.