

# UNA NUOVA CATEGORIA CULTURALE NELL'EPOCA DEL DIGITALE

# LA DATACY

DALLA QUALITÀ DEI DATI ALLA POST-VERITÀ

CARLO BATINI

II PARTE

*Nel numero 2/2017 di «Gnosis» abbiamo marginalmente trattato alcuni argomenti che, ripresi nel numero 4/2018, sono sviluppati in questo articolo in modo più organico, nel contesto dell'esame di uno dei più importanti temi connessi ai fenomeni della digitalizzazione: la qualità e veridicità dei dati digitali. Relativamente circoscritto nei sistemi informativi tradizionali che adottano modelli di rappresentazione relazionali strutturati, nel Web, in cui i dati sono creati e scambiati senza nessuna intermediazione e a costi irrisori, assume invece dimensioni molto più ampie, confinando con la categoria chiamata post-verità.*

**La** *Datacy*, introdotta nella prima parte dell'articolo<sup>1</sup>, è una nuova categoria interdisciplinare che fornisce gli strumenti per comprendere il mondo nell'epoca della diffusione dei dati digitali. In questa seconda parte focalizzeremo l'attenzione sul tema relativo alla qualità dei dati, che nel web e nelle reti sociali si trasforma in una categoria investigata da filosofi, umanisti, studiosi di comunicazione: la *verità*.

Nell'articolazione del tema useremo, con significati simili, come del resto frequentemente accade nella letteratura scientifica, i termini «dato» e «informazione»: il primo, prevalentemente riferito all'informazione strutturata (es. tabelle); il secondo, più usato con riguardo a rappresentazioni debolmente strutturate di tipo linguistico (es. messaggi Twitter) o basate sui sensi (es. immagini, suoni).

1. BATINI 2018.





Da sempre l'informazione scambiata tra umani è stata oggetto di deformazione e comunicazione non corretta. Attraverso un'immagine fotografica della zona centrale di Napoli e due mappe tratte da guide del Touring del tempo<sup>2</sup>, si comprende come le carte topografiche possano alterare la realtà, trasformando una fitta rete di edifici separati da strade anguste in una zona di villini unita da ampie strade<sup>3</sup>.

#### QUALITÀ DEI DATI NEI SISTEMI INFORMATIVI TRADIZIONALI

Nelle basi dei sistemi informativi tradizionali i dati sono descritti mediante modelli di rappresentazione strutturati, evolutisi nel tempo nei sistemi di gestione di basi di dati da modelli gerarchici e reticolari verso *modelli relazionali*.

I sistemi informativi che offrono servizi agli utenti – ad esempio un sistema di prenotazione – fanno uso di dati non solo strutturati ma anche *controllati*, nel senso che la loro fonte e il dominio di definizione sono in genere noti, e la natura strutturata del dato, così come i vincoli d'integrità definibili, costituiscono un filtro all'immissione di dati scorretti.

La qualità di un singolo dato (o di un loro insieme, detto anche *dataset*) è la caratteristica che si basa sulla sua capacità di soddisfare necessità e aspettative, esplicite o implicite, dei fruitori. Una *dimensione* di qualità è una specifica proprietà associabile alla qualità, usualmente non misurabile.

Come nel caso di un testo scritto, una dimensione di qualità riguarda la sua comprensibilità, intesa quale ridotto sforzo cognitivo necessario per ricostruire il significato del testo.

Una *metrica* di qualità è una procedura di misurazione di una dimensione di qualità che, partendo dalla dimensione da misurare, associa a essa un valore numerico (es. 0,7) o ordinale (es. alta) in un dominio di valori.

Nel caso della *comprensibilità* di un testo sono state definite varie metriche, tra cui gli indici che misurano la percentuale di parole, sul totale che non fanno parte di un elenco, considerate comprensibili a una persona che abbia raggiunto un determinato titolo di studio (ad esempio, la scuola dell'obbligo)<sup>4</sup>.

2. BATINI 1984.

3. BATINI 1984 è scaricabile in <<http://hdl.handle.net/10281/97703>> [22-1-2019].

4. Cfr. DE MAURO 2003.

Le più importanti tra le dimensioni di qualità nelle basi di dati sono:

- *accuratezza*, intesa come aderenza del dato al fenomeno osservato;
- *completezza*, cioè l'estensione con cui il dato rappresenta la realtà osservata;
- *tempestività* di aggiornamento, vale a dire rapidità con cui cambiamenti nel fenomeno osservato corrispondono ad aggiornamenti del dato digitale;
- *consistenza*, cioè il rispetto di un insieme di regole logiche definite per rappresentare le proprietà del dato.

Le metodologie per la misurazione e il miglioramento della qualità dei dati nei sistemi informativi tradizionali si basano sul confronto tra i dati stessi e una conoscenza di riferimento, che può essere costituita da insiemi di dati certificati, vincoli logici tra dati, ovvero acquisita mediante ricerche o indagini. Qualora una tabella contenga l'albo delle aziende fornitrici di un'organizzazione, per controllare l'accuratezza dei nomi dei fornitori dobbiamo fare riferimento a un elenco certificato delle aziende, come il Registro delle imprese detenuto in Italia da Unioncamere. Se non ritroviamo il nome di un'azienda possiamo arrivare alla conclusione che esso sia errato, ma per correggerlo possiamo condurre un'indagine, costosa, volta a rintracciare la ditta e a ricostruirne la natura giuridica, ovvero usare un'euristica, sostituendo al nome scorretto quello che nel Registro imprese è meno lontano, secondo un'opportuna funzione di distanza. Sia per valutare la qualità che per migliorarla, facciamo dunque uso di conoscenza supplementare. Una trattazione esauriente del tema della qualità dei dati nelle basi di dati, nelle immagini, nelle mappe geografiche e nei testi non strutturati, nonché una discussione sulle metodologie e tecniche per la valutazione e il miglioramento della loro qualità nei sistemi informativi tradizionali le troviamo in *Data and Information Quality*, del 2016<sup>5</sup>.

#### LA QUALITÀ DEI DATI NEL WEB

L'inarrestabile crescita dei dati prodotti nell'*ecosistema* costituito dal web modifica profondamente il concetto e le metodologie per la qualità dei dati nei sistemi informativi tradizionali, in virtù della grande varietà delle tipologie e per l'impossibilità, spesso, di confrontare il dato con una 'conoscenza di riferimento'.

5. BATINI – SCANNAPIECO 2016.

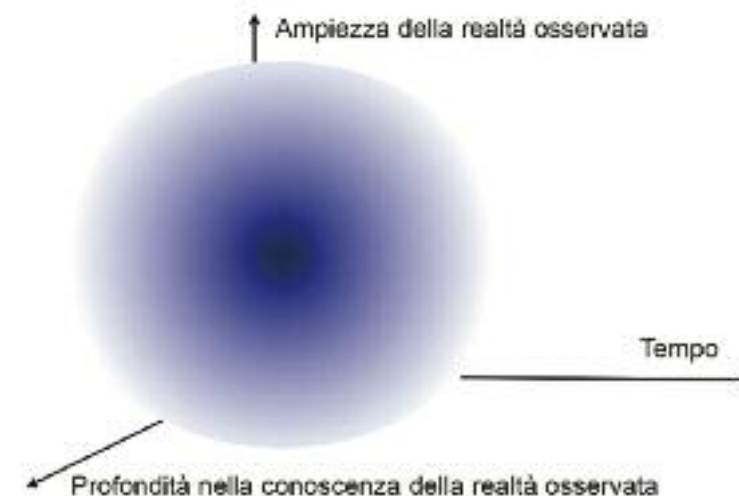




Le differenze tra i due ambiti sono molteplici:

- il costo di produzione e trasmissione dei dati è praticamente nullo, a fronte di possibilità di diffusione sempre più ampie. È praticamente gratuito l'inoltro di un'email sia a una persona che a un indirizzario di centinaia o migliaia di destinatari; scrivere un messaggio su Twitter richiede uno sforzo cognitivo molto basso e, allo stesso tempo, su Twitter ci possono essere milioni di follower, e i tweet si diffondono molto più velocemente dei dati recati dalle fonti cartacee tradizionali; infine, leggere un tweet richiede uno sforzo cognitivo significativamente inferiore rispetto a un articolo di giornale;
- quanto al carattere controllato dei flussi di dati nei sistemi informativi tradizionali e alla presenza, spesso, di un soggetto certificatore del dato, nel web si assiste a una sostanziale disintermediazione tra la fonte e il ricevente. Ad esempio, nelle reti sociali non esiste nessun meccanismo di filtro; peraltro, talvolta sono presenti, soprattutto nei siti che adottano il *crowdsourcing*, meccanismi redazionali che segnalano dati e informazioni contraddittori o chiaramente di parte, ovvero atti a intercettare frasi o video che violano regole etiche o fanno riferimento a informazioni di natura pornografica ecc. Il 'rovescio della medaglia' sta nel fatto che queste forme di controllo possono facilmente tramutarsi in censura;
- nel web abbiamo una distinzione tra «fonte» del dato, «mezzo» con cui è trasmesso e «messaggio» che viene trasmesso, e il tema della qualità riguarda tutti e tre i livelli, con intrecci rilevanti. Per esempio, esiste una significativa differenza tra i messaggi inviati via Facebook e quelli via Twitter, derivanti dai vincoli sulla lunghezza del messaggio e il conseguente uso di vocabolari arricchiti con simboli, abbreviazioni e metafore;
- nel web, che include estensivamente l'Internet delle cose, le reti sociali, le tecnologie mobili, i tipi di dati che fanno riferimento alle rappresentazioni linguistiche e a quelle basate sui sensi si differenziano a seconda degli usi e delle intenzioni del produttore del dato;
- non ci sono (né ci possono essere) standard universali per scambiare l'informazione nel web. Essa può essere alterata o creata anonimamente sotto falsa identità o con l'intenzione dell'inganno;
- nei siti che nascono per *crowdsourcing* l'informazione è il risultato di molteplici contributi e versioni, e si perde la conoscenza sulla fonte e sul processo di provenienza;
- le dimensioni di qualità e le metodologie per la valutazione e il miglioramento della qualità, dovendo modellare una grande e diversificata varietà di fenomeni, sono molto meno consolidate rispetto ai sistemi tradizionali. Quelle più citate riguardano il trust<sup>6</sup>, la veracità, la credibilità e l'esperienza.

6. SHERCHAN ET AL. 2013.



Per esprimerci con una metafora, i dati si espandono nel web come una sfera opaca, in cui accanto a quelli di qualità compaiono con sempre maggiore intensità quelli imprecisi, sfocati, incompleti ed eterogenei al loro interno, rendendo più arduo riportarli a coerenza e ricostruirne la validità (vedi figura in alto).

L'immagine della sfera è anche utile per farci capire che, in ragione dei nostri limiti cognitivi, un ampliamento della conoscenza dei fenomeni può portarci a percepire tracce sempre più frammentate, eterogenee e imprecise di una realtà via via più complessa, che vanno ricomposte in una conoscenza comune. Ciò è complicato e costoso. Nella grande sfera opaca il concetto 'minimalista' di qualità del dato, adottato nei sistemi informativi tradizionali, si amplia a dismisura per i vari tipi d'informazioni forniti dall'Internet delle cose, scambiati nel web e nelle reti sociali, acquisiti e condivisi per mezzo dei telefoni cellulari, memorizzati nel Cloud, approdando in ultima analisi al concetto di 'verità' ampiamente studiato da secoli nella filosofia del linguaggio, nella filosofia morale, nella logica e nell'etica, e alle sue sistematizzazioni ed evoluzioni nei concetti d'ipoverità, iperverità e post-verità. L'analisi, le classificazioni e le tecniche proposte si differenziano significativamente a seconda del tipo d'informazione, e si concentrano maggiormente nell'ambito dei messaggi, dei testi in formato libero, degli articoli di giornale, dei *rumours* (voci), opinioni, notizie (news), microblogging, informazione specialistica ecc.





Nei giornali il problema è particolarmente critico. Per «The Guardian»<sup>7</sup> stabilire la verità costa; i giornalisti che hanno appreso il loro mestiere nell'epoca pre-digitale non sapevano molto di profitti e perdite, né applicavano ciò che viene chiamato il *business model*. In questo ambito, le notizie grezze, non verificate, sono gratuite e quelle validate molto costose. Questa situazione è aggravata dal fatto che le tecnologie digitali attraverso il fenomeno dei Big Data estendono enormemente per loro natura la capacità di diffondere informazioni false.

Al giornalismo resta il fondamentale compito di fornire una visione realistica, ragionata e completa di tutti gli elementi di giudizio dei fatti e degli eventi che influenzano la vita dei lettori.

Ciò, tuttavia, in un contesto caratterizzato<sup>8</sup> dall'intreccio tra *revenue* derivanti dalla pubblicità, dagli algoritmi basati sui click, dal monopolio sui dati di profilazione detenuto dalle grandi aziende che possiedono le reti sociali e utilizzano l'e-commerce, la decrescita della produzione di giornali cartacei provoca la progressiva perdita di mercato dell'informazione 'meditata' e con alta validità temporale, e in modo inversamente proporzionale cagiona la crescita dell'informazione prodotta sul momento, veloce, emotiva e ad alto decadimento temporale.

Il mutato contesto in cui i dati sono rappresentati, comunicati e percepiti porta a diversificare i paradigmi e le scienze che investigano i concetti che fanno riferimento alla verità. I punti di vista secondo cui è affrontato il problema sono almeno quattro: l'approccio dell'informatica, l'approccio ontologico, l'approccio delle scienze cognitive, l'approccio della filosofia del linguaggio.

Esaminiamoli di seguito.

#### APPROCCIO INFORMATICO

L'approccio informatico cerca di valutare la qualità dei dati, con particolare riguardo ai social media, attraverso l'individuazione di tecniche automatiche per la valutazione delle dimensioni di qualità sul web e, in particolare, della credibilità. Pur in un dominio ristretto ai social media, è diffusa la consapevolezza che le applicazioni sociali coprono un ampio spettro di scopi, in differenti contesti e diversificati tipi di dati.

7. «The Guardian» 2019: <www.theguardian.com.> [22-1-2019].

8. QUINTARELLI 2019.

Le tecniche sono investigate per diverse tipologie di dati, quali le *review*, suddivise in *untruthful*, quando forniscono deliberatamente *review* positive o negative allo scopo di ingannare il fruitore del dato, *review* sul brand piuttosto che sul prodotto o servizio, e *unreview*, che non contengono opinioni e perciò disorientano, messaggi diffusi attraverso il *microblogging* e suddivisi in *conversation item*, che riguardano l'utente e la sua cerchia di amici, e *news item*, che si riferiscono all'informazione più generale, l'informazione medico-scientifica, rumours e altro. Abbiamo voluto dettagliare la classificazione per mostrare concretamente quanto debba diversificarsi la ricerca in virtù della generalità ed estensione del web, e quale sia il divario che si sta ampliando tra il fenomeno dei dati digitali e i risultati della ricerca.

Negli approcci più promettenti<sup>9</sup> (per un *survey* esauriente), la credibilità è vista come una qualità percepita dal ricevente l'informazione, non viene definita ed è considerata composta di molteplici dimensioni che riguardano diverse caratteristiche: fonte, relazioni virtuali tra gli utenti nella rete sociale, messaggio (o combinazione della fonte e del messaggio) e mezzo di trasmissione. La credibilità è attestata in connessione all'expertise (la conoscenza, lo skill e l'esperienza percepita della fonte dell'informazione), e alla *trustworthiness*, la percezione di quanto un'informazione emessa dalla fonte è valida. Anche l'expertise è basata su fattori soggettivi, ma include altresì caratteristiche oggettive della fonte (come le certificazioni); così pure la *trustworthiness* comprende elementi oggettivi quali l'affidabilità misurata attraverso metriche oggettive.

Le metodologie di base utilizzate sono: *data-driven*, quando utilizzano tecniche di machine learning per identificare l'informazione falsa; *model-driven*, quando si focalizzano su schemi di aggregazione; *graph-based*, quando sfruttano la struttura delle entità connesse nel grafo costituito dalla rete sociale, estendendosi, tra le entità, alla struttura di connettività degli utenti, dei prodotti / servizi, dei messaggi. Sia i modelli data-driven che quelli model-driven si focalizzano su un certo numero di caratteristiche (*feature*), che possono essere estratte sia dal messaggio che dalla fonte; inoltre, le *feature* possono essere sia linguistiche che facenti riferimento a meta-dati. Ovvero, possono essere riconducibili al comportamento del soggetto-fonte, alla struttura delle sue relazioni, alla natura del prodotto / servizio riferito nel messaggio.

9. VIVIANI 2017.





Tra le feature linguistiche distinguiamo quelle lessicali, quelle stilistiche e le inconsistenze semantiche. Tra le feature comportamentali si distinguono i dati pubblici disponibili sui siti web, dati privati quali gli indirizzi IP, il tempo intercorso per postare un testo, la locazione fisica del fornitore dei contenuti ecc. Una categorizzazione esaustiva degli approcci alla valutazione di credibilità compare in Viviani<sup>10</sup>.

#### APPROCCIO ONTOLOGICO

L'approccio ontologico qui descritto, a partire da Ferraris, fornisce modelli classificatori e interpretativi, più che metodi e tecniche: l'ipoverità è ciò che è creduto tale da una comunità e che viene corroborato da un insieme di procedure<sup>11</sup>. L'iper verità caratterizza il contesto per cui la proposizione «la neve è bianca» sarebbe vera anche se non ci fosse mai stato un essere umano sulla terra in grado di formularla. La verità è il risultato tecnologico del rapporto tra ontologia (ciò che esiste) e epistemologia (ciò che conosciamo).

Le tre categorie possono essere esemplificate come segue:

- ontologia – in questo barattolo ci sono 12 fagioli. L'approccio ontologico fa riferimento a ciò che è;
- epistemologia – enuncio la frase «in questo barattolo ci sono 12 fagioli». L'epistemologia fa riferimento a ciò che sappiamo;
- tecnologia – io conto 12 fagioli. L'approccio tecnologico si riferisce a ciò che possiamo fare, misurare.

E, ancora: «il barattolo ha un certo peso» fa riferimento all'ontologia; «lo metto su una bilancia per misurarne il peso» fa riferimento alla tecnologia; la frase «il barattolo pesa 100 grammi», così come la frase «il barattolo pesa tre onces e mezza» fanno riferimento all'epistemologia.

DISCIPLINA	RIGUARDA LA...	OSSERVA...
Ontologia	Realtà	Oggetti
Epistemologia	Verità	Concetti
Tecnologia	Interpretazione	Fatti

In conclusione, la verità «è *relativa* rispetto agli strumenti tecnici di verifica, ma assoluta rispetto alla sfera ontologica a cui fa riferimento e all'esigenza epistemologica a cui risponde».

10. *Ibidem*

11. FERRARIS 2017.

#### APPROCCIO COGNITIVO

Quest'approccio parte dall'osservazione<sup>12</sup> che l'attività d'interpretazione del dato richiede uno sforzo cognitivo, che trova i propri limiti nell'idea della razionalità limitata<sup>13</sup>. Sebbene il web abbia ridotto alcuni costi connessi alla ricerca dei dati accrescendone l'accessibilità, rimangono nella loro interpretazione, e sono relativamente incompressibili, oneri significativi d'interazione in virtù della vastissima area di dati disponibili, della sua eterogeneità e opacità.

Dalla psicologia cognitiva non arrivano dunque buone notizie; riconoscere la cattiva informazione richiede processi cognitivi complessi. Un semplice mito è più attrattivo cognitivamente di una complicata correzione. Per coloro che sono fortemente convinti delle proprie idee, gli argomenti contrari possono rafforzare le loro convinzioni; pertanto, non è tanto rilevante ciò che la gente pensa ma come pensa. Per questi motivi la ricerca si è soprattutto soffermata sulle euristiche utilizzate nella valutazione della credibilità e basate su:

- reputazione, che porta a privilegiare alternative riconoscibili rispetto a quelle meno familiari;
- endorsement (sostegno, appoggio), fondato sulle valutazioni espresse da altri, cui affidiamo la nostra;
- consistenza, che riguarda il confronto tra fonti per evidenziarne le differenze. Nel caso d'inconsistenze, sono proposte varie tecniche per la scelta tra le alternative;
- auto-conferma, che misura la credibilità sulla base della conferma delle precedenti credenze; rientrano in questa tematica le analisi documentate sulle camere dell'eco<sup>14</sup>.

Nel lavoro viene presentata un'analisi effettuata su circa 1000 agenzie di stampa e 400 milioni di utenti, in cui è stata esplorata l'anatomia del consumo di notizie su Facebook su scala globale. La conclusione che si trae è che gli utenti, quando accedono al web per fini informativi, tendono a focalizzare la loro attenzione su un numero limitato di pagine, andando a selezionare un gruppo ristretto di media da cui attingere informazioni e rafforzando così le proprie opinioni, senza mai metterle in discussione.

Di fatto, si chiudono nella loro bolla;

- violazione delle aspettative, che assume una fonte come non credibile ove abbia violato le aspettative in precedenti circostanze;
- intento persuasivo, che tende a non considerare credibile il dato che viene percepito come affetto da un pregiudizio (bias); questa euristica è tipica dell'informazione commerciale.

12. METZGER – FLANAGIN 2013.

13. SIMON 1955.

14. Cfr., ad esempio, QUATTROCIOCCHI – VICINI 2016.





Se la focalizzazione delle indicate euristiche è un rilevante passo avanti nella comprensione della qualità dei dati sul web, la ricerca è ancora aperta per l'individuazione della correlazione con il profilo utente e dell'influenza delle euristiche in contesti di comunicazione mediata dall'automazione.

#### APPROCCIO DELLA FILOSOFIA DEL LINGUAGGIO

Nel testo di Anna Maria Lorusso<sup>15</sup>, cui ci ispiriamo per descrivere tale approccio<sup>16</sup>, viene richiamata la definizione dell'*Oxford Dictionary* secondo cui la post-verità viene indicata come argomentazione che fa riferimento a circostanze in cui l'oggettività dei fatti è meno influente nel formare la pubblica opinione rispetto all'emotività e alle credenze personali. Indica, perciò, una strategia retorico-persuasiva in cui la componente soggettiva e passionale è prevalente su quella referenziale.

«I media non rappresentano un reale già fatto, che sta da qualche parte nel mondo, i media costruiscono il reale, lo modellano. Gli spazi mediatici sono luoghi di costruzione del reale perché sono i luoghi in cui elaboriamo i modelli con cui poi classifichiamo il mondo e ci muoviamo in esso; da qui l'affermazione: è vero, o è reale solo ciò che passa dalla televisione, affermazione ormai datata, aggiornandola a partire dai nuovi media e reti sociali»<sup>17</sup>.

Un tempo c'erano (solo) i giornali; i social media sono la generalizzazione (qualcuno dice la democratizzazione) delle agenzie di verità. Oggi chiunque sembra autorizzato a produrre non la sua versione del mondo, ma una versione del mondo che pretende di essere vera, che legittima solo la logica esclusiva del vero / falso. Non esiste più la coesistenza tollerante delle sfumature.

Lorusso analizza criticamente anche l'illusione del *fact checking*. In un mondo di verità moltiplicate la possibilità statistica dell'errore cresce a dismisura; se gli agenti dell'informazione si moltiplicano e possono essere ovunque, allora il controllo diventa infinitamente più difficile.

Più che di verità assolutizzata, ha senso parlare di *livelli di verità*. In <politifact.com> esistono diversi livelli: *true*, notizie accurate e complete; *mostly true*, notizie che richiedono alcune integrazioni

15. Lorusso 2018.

16. SEARLE 2000.

17. LORUSSO 2018.

e chiarimenti; *half true*, notizie che trascurano dettagli importanti e decontestualizzano l'informazione; *mostly false*, notizie che ignorano punti di vista cui corrisponde un'altra lettura dei fatti; *false*, notizie non accurate; *pants on fire*, notizie che sostengono tesi ridicole.

L'intuizione più rilevante in Lorusso<sup>18</sup> riguarda la caratterizzazione della verità, che non è immagine della realtà, ma costruzione, e dunque un processo. Un itinerario verso una consapevolezza della verità insita nei fatti individua tre loro proprietà:

- *completezza* – nei fatti ci devono essere tutti gli elementi fondamentali;
- *contestualizzazione* – i fatti devono essere messi in relazione con gli elementi cui sono connessi;
- *tenuta* – i fatti devono aver tenuto in conto i punti di vista critici che potrebbero dare un'altra lettura.

Si può affermare che problematizzare la verità, vederne la molteplicità e relativizzarne la natura non significa che tutto vada bene, che tutto è verità, ma vuol dire recuperare il concetto di prova: non tutte le verità hanno la stessa tenuta. I discorsi hanno formazioni, forza, autorevolezza, raggi di applicazione; guardare a questi elementi significa 'provare' la tenuta dei discorsi.

La verità «non può essere raggiunta ma avvicinata», mettendo alla prova le diverse verità espresse in ragione d'interessi diversi, e via via facendo la realtà. Tenendo conto dell'intersoggettività e dell'etica dell'interpretazione e della comunicazione.

Come evidenzia Umberto Eco<sup>19</sup>, non c'è dunque una verifica che basti da sola ad assicurare la veridicità di qualcosa; al massimo ci possono essere verifiche incrociate che possono escludere interpretazioni false o errate. Questa posizione è simile alla congettura secondo cui eseguendo un programma non si può mai provare che è corretto, ma si possono solo trovare nuovi dati in ingresso al programma che dimostrano l'esistenza di errori.

Nel mondo contemporaneo sta accadendo qualcosa di più rilevante di quanto trattato finora: sempre di più la costruzione del senso si dà per *via narrativa*. La notizia è sempre meno pensata come documento e sempre più come racconto. A prevalere non è quindi un criterio di attendibilità, ma di efficacia narrativa definita «credibilità»<sup>20</sup>. C'è una profonda relazione tra fatti, favole, fole, bugie; la forza dei nuovi soggetti d'informazione si misura più sulla

18. *Ibidem*

19. Eco 1990.

20. LORUSSO 2018.



capacità di riuso di frame narrativi consolidati che sull'attendibilità dell'informazione. Sembrerà esagerato, ma la dinamica è la stessa; quando leggo una favola, ho delle esigenze che mi fanno apprezzare quella favola e fanno sì che ci creda e mi appassioni.

L'impressione è che le verità si siano più che altro parcellizzate e, nel parcellizzarsi, si siano moltiplicate in un gioco di specchi e rimbalzi. Attraverso la condivisione sempre più estesa hanno acquisito credibilità. La logica di Facebook rafforza l'idea che le condivisioni siano indice d'identificazione e di adesione, dunque di credibilità.

In sintesi, «oggi i soggetti dell'informazione siamo noi, persone comuni, dalle competenze comuni, il più delle volte con una nostra esperienza da condividere, con un nodo di emozioni da esprimere. In un mondo di verità moltiplicate, assottigliate l'una rispetto all'altra e isolate entro comunità chiuse, la cosa più seria che si perde non è il vero ma il legame sociale. Saper discriminare la verità significa condividere saperi. E condividere saperi significa essere parte della stessa comunità»<sup>21</sup>.


## CONCLUSIONI

La conclusione più matura che possiamo trarre dalla discussione sin qui sviluppata è che il web è un'immensa prateria, in cui è difficile individuare elementi di riferimento per trovare la verità, che può essere semmai avvicinata, non raggiunta. Possiamo dire che sia un dovere etico cercare di avvicinare la verità, ma per farlo c'è bisogno di sforzo cognitivo e mente libera da convinzioni stratificate nel tempo e da ideologie. Nel gennaio 2019 la neo deputata democratica Ocasio Cortez, in risposta all'osservazione di un giornalista del «The Washington Post» che le contestava un'affermazione imprecisa su un indicatore economico, per forzarlo verso la propria tesi ha risposto: «I think that there's a lot of people more concerned about being precisely, factually, and semantically correct than about being morally right». Ma, nel mondo d'oggi, quale partito o movimento si può arrogare il diritto di esprimere una superiorità morale?

Per accostarsi alla verità sono necessarie pazienza e perseveranza nell'indagine: esattamente il contrario di quanto ci sollecita spesso la Rete, e il principale indizio di verità è la sua tenuta nel tempo<sup>22</sup>.

21. *Ibidem*

22. Eco 1990.

Possiamo dunque tornare all'immagine comparativa della cerimonia d'inaugurazione a Washington delle presidenze di Obama e di Trump, di cui all'articolo del 2017<sup>23</sup>, per suffragare la policy che abbiamo sintetizzato con l'espressione «i dati sono cocciuti». Nella visione che abbiamo sviluppato in precedenza, non potremo mai arrivare all'assoluta certezza che i partecipanti all'inauguration day di Obama fossero più numerosi di quelli presenti nell'analoga circostanza che ha riguardato Trump, ma possiamo certamente contestare l'affermazione della collaboratrice del presidente subentrante su quelli che definì gli «alternative facts»; ciò può essere fatto acquisendo conoscenza ulteriore sulle due immagini, come l'ora di ripresa delle foto, l'intervallo temporale tra l'ora di ripresa e l'inizio della cerimonia, il numero di biglietti validati sulla metropolitana di Washington in quei due giorni, il numero di auto in sosta nei parcheggi limitrofi all'area della cerimonia, avvicinando sempre più la verità fino, per così dire, a toccarla 

23. BATINI 2017, p. 48.

## BIBLIOGRAFIA

- C. BATINI, *Le basi dell'informatica*, Editori Riuniti, Roma 1984.
- C. BATINI, *Big Data. Big Challenges and Big Concerns*, «Gnosis» XXIII (2017) 2, pp. 41-49.
- C. BATINI, *Una nuova categoria culturale nell'epoca del digitale: la Datacy (parte I)*, «Gnosis» XXIV (2018) 4, pp. 64-73.
- C. BATINI – M. SCANNAPIECO, *Data and Information Quality*, Springer, New York 2016.
- T. DE MAURO, *Guida all'uso delle parole*, Editori Riuniti, Roma 2003.
- U. ECO, *I limiti dell'interpretazione*, Bompiani, Milano 1990.
- M. FERRARIS, *Postverità e altri enigmi*, il Mulino, Bologna 2017.
- L. FLORIDI, *La quarta rivoluzione. Come l'infosfera sta trasformando il mondo*, Raffaello Cortina, Milano 2017.
- A.M. LORUSSO, *Postverità*, Laterza, Roma-Bari 2018.
- T. LUKOIANOVA – V.L. RUBIN, *Veracity Roadmap: Is Big Data Objective, Truthful and Credible?*, «Advances in Classification Research Online» XXIV (2014) 1, pp. 4-15.
- M.J. METZGER – A.J. FLANAGIN, *Credibility and trust of information in online environments: The use of cognitive heuristics*, «Journal of Pragmatics» LIX (2013), pp. 210-220.
- W. QUATTROCIOCHI – A. VICINI, *Misinformation*, Franco Angeli, Roma 2016.
- S. QUINTARELLI, *Perché Internet è tutto un articolo acchiappa click? Follow the money*, «Il Foglio» (13 gennaio 2019).
- J.R. SEARLE, *Mente, linguaggio, società*, Raffaello Cortina Editore, Milano 2000.
- W. SHERCHAN ET AL., *A Survey of Trust in Social Networks*, «ACM Computing Surveys» XLV (2013) 4, pp. 1-133.
- H.A. SIMON, *A Behavioural Model of Rational Choice*, «Quarterly Journal of Economics» LXIX (1955) 1, pp. 99-118.
- M. VIVIANI – G. PASI, *Credibility in Social Media: Opinions, News, and Health Information: A Survey*, «Wiley Interdisciplinary Reviews. Data Mining and Knowledge Discovery» VII (2017) 5, pp. 1-25.
- A. ZUBIAGA ET AL., *Detection and Resolution of Rumours in Social Media: A Survey*, «ACM Computing Surveys» LI (2018) 2, pp. 2-36.