

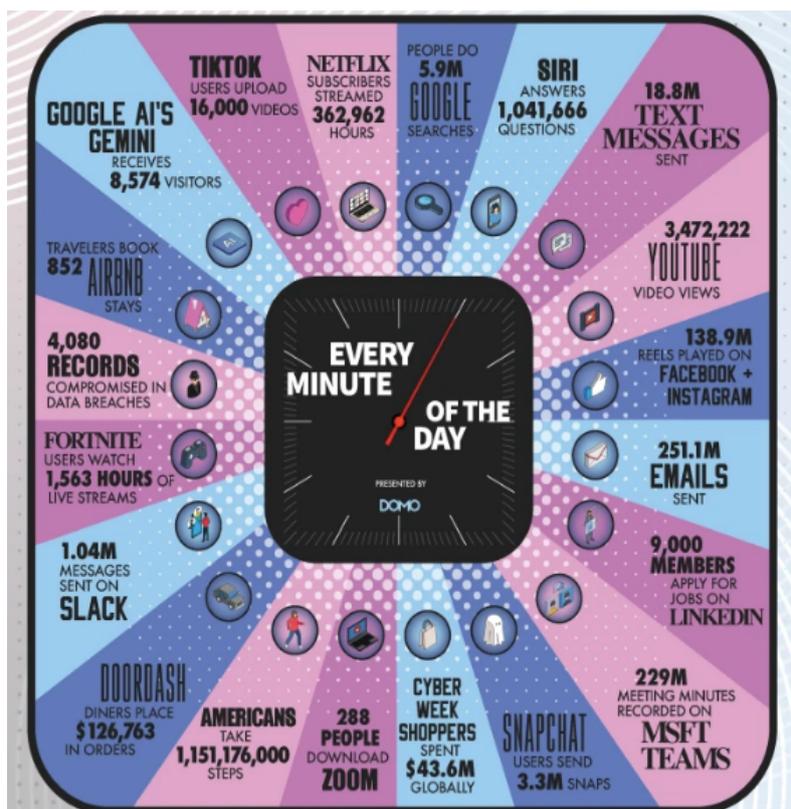
2. Attenzione alle emozioni! Come difendersi dai trabocchetti della AI

Lo abbiamo già visto nella prima attività di approfondimento: la disinformazione e la misinformazione (per comodità, possiamo anche continuare a chiamarle bufale o “fake news”, purché consapevoli di quanto siano limitanti queste definizioni) giocano in maniera molto sottile con le nostre emozioni e i nostri sentimenti, positivi o negativi che siano.

La **reazione emotiva** è quella che ci porta a bypassare la razionalità, ad accelerare i tempi di condivisione, a reagire in maniera istintiva e rilanciare un contenuto senza verificarlo. Se poi a questo uniamo il continuo **bombardamento di informazioni** a cui siamo sottoposti in ogni momento della nostra giornata, ci sarà più facile comprendere come l’infodemia sia un pericolo concreto contro il quale dobbiamo imparare a difenderci.

Ma cos’è l’infodemia? Stando all’Enciclopedia Treccani: “s.f. Circolazione di una quantità eccessiva di informazioni, talvolta non vagliate con accuratezza, che rendono difficile orientarsi su un determinato argomento per la difficoltà di individuare fonti affidabili.” Si tratta di un neologismo, un termine coniato quasi in concomitanza con l’esplosione dell’allarme Covid. "Con il neologismo infodemia l’Organizzazione mondiale della sanità (Oms) ha voluto, in questi giorni in cui la paura del Coronavirus impazza, sottolineare che forse il maggiore pericolo della società globale nell’era dei social media è la deformazione della realtà nel rimbombo degli echi e dei commenti della comunità globale su fatti reali o spesso inventati”, scriveva Leonardo Becchetti su Avvenire, il 5 febbraio 2020.

Un aspetto fondamentale dell’infodemia è la quantità, il flusso di notizie. Che nel tempo della rete sembra essere esondato da ogni argine. Ecco, ad esempio, cosa accade ogni 60 secondi sul World Wide Web, dove “i dati non dormono mai”, come scrivono i ricercatori di Domo (autori ormai da 12 anni consecutivi del visual report “One Internet Minute”):



(La grafica è disponibile in alta risoluzione sul sito <https://www.domo.com/learn/infographic/data-never-sleeps-12>)

Il sovraccarico informativo, sia in termini di quantità che di velocità nella diffusione e condivisione, ci spinge ad abbassare la soglia di attenzione, che è uno degli strumenti principali per intercettare le “fake news” o le verità parziali e distorte.

Inoltre, uno degli aspetti principali della disinformazione, il suo vero e proprio punto di forza, sta nella sua capacità di scegliere come bersaglio principale non la razionalità, bensì l’emotività di chi legge. Una “fake news” efficace sa evocare in noi sentimenti istintivi e molto forti, a volte radicati nei nostri pregiudizi e nelle nostre paure, in altri casi nelle nostre speranze e nel nostro senso di appartenenza a un gruppo. Una reazione “di pancia” ci porterà nuovamente a scoprire il fianco e ad agire d’istinto, motivandoci a condividere con i nostri amici e contatti quella notizia che ci ha suscitato tanta rabbia e sdegno, oppure che ci ha entusiasmato e fatto sentire parte di qualcosa di importante. E senza neanche volerlo, eccoci trasformati in un nuovo anello dell’infinita catena della viralità.

Queste due problematiche - **riduzione della soglia di attenzione, aumento dell’impatto emotivo** - vengono ulteriormente esasperate dalla **predominanza dell’aspetto visivo** nella costruzione e diffusione delle notizie. Da molto tempo, ormai, la nostra è una società dell’immagine, o meglio, delle immagini. La notizia è spesso definita dalla disponibilità di “prove visive” del fatto accaduto. Le foto e i video hanno un impatto immediato e una potenza comunicativa molto diversa da quella di testo e parole.

Per questo, oggi, uno degli strumenti più insidiosi per la costruzione di una narrazione collettiva sono proprio le immagini, manipolate, ricontestualizzate, a volte create dal nulla grazie alle potenzialità dell’intelligenza artificiale. “Il volume della disinformazione generata dall’Intelligenza artificiale, in particolare le **immagini deepfake** relative alle elezioni, è

aumentato in media del 130% al mese su X nell'ultimo anno”, lanciava l’allarme RaiNews a marzo 2024. I dati provenivano da uno studio del Center for Countering Digital Hate (Ccdh), un'organizzazione no-profit britannica che si batte contro l’incitamento all’odio online (qui potete leggere l’articolo completo: <https://www.rainews.it/articoli/2024/03/la-disinformazione-generata-dallintelligenza-artificiale-e-aumentata-in-un-anno-del-130-su-x-ea0da9e2-2f48-4104-a93e-b5db174d4418.html>).

Ma i cosiddetti “deepfake” possono entrare in gioco anche con obiettivi più immediati e concreti, come quello di attirare gli utenti in una truffa. Un caso interessante è stato quello dei video circolati su Facebook in cui personaggi noti del mondo dello spettacolo “sponsorizzavano” uno schema di investimenti online. In questo articolo di Fanpage potete visionare sia il video nella sua interezza, sia l’analisi che ne è stata fatta: <https://www.fanpage.it/innovazione/tecnologia/fabio-fazio-promuove-investimenti-online-in-un-video-e-un-deepfake-ed-e-uguale-alloriginale/>

Come difendersi da queste trappole? Nel caso delle immagini manipolate, è importante saper effettuare una **ricerca inversa per immagini** grazie a strumenti come Google Immagini o TinEye, dove caricando il file o inserendo l’URL della foto che ci interessa potremo risalire alla prima volta in cui è stata pubblicata in rete (in questo modo potremo, ad esempio, capire se si tratta di un’immagine decontestualizzata) o ai siti da cui è stata diffusa (da cui si può evincere l’affidabilità delle fonti). Oppure, per chi volesse cimentarsi con strumenti più complessi, potremo ricorrere a FotoForensics per valutare il **livello di manipolazione** dell’immagine stessa.

Per quanto riguarda le immagini generate con l’intelligenza artificiale, possiamo ricorrere agli stessi strumenti che abbiamo appena menzionato, ma prima ancora dobbiamo imparare a focalizzarci sui dettagli (dove spesso l’intelligenza artificiale, che si basa sulla rielaborazione di dati provenienti da foto reali, commette qualche “piccolo” errore, ad esempio inserendo dita in più...) e sull’eccesso di perfezione.

Nelle esercitazioni che seguono, proveremo a trasformarci in “investigatori delle immagini digitali”.

PRIMO ESERCIZIO

Artificial or real?



Per alcuni mesi il sito della BBC - la British Broadcasting Corporation, fondata a Londra nel 1922, la più antica emittente nazionale del pianeta e uno dei media più autorevoli al mondo - ha pubblicato un test intitolato, appunto, “artificiale o reale”, in cui gli utenti vengono sfidati in un quiz per immagini su celebrities e argomenti del momento.

Prima di effettuare il quiz, leggiamo il vademecum anti deepfake proposto dall'emittente britannica “How to spot AI images on social media”:

<https://www.bbc.co.uk/bitesize/articles/z6s4239#:~:text=Be%20wary%20of%20perfection&text=Another%20sure%2Dfire%20way%20of,having%20an%20'airbrushed'%20look.>

Potremo poi misurarci con i quiz, dividendoci a squadre o individualmente, e confrontando i risultati finali:

- <https://www.bbc.co.uk/bitesize/articles/znm9qyc>
- <https://www.bbc.co.uk/bitesize/articles/zqnwxg8>
- <https://www.bbc.co.uk/bitesize/articles/zg78239>
- <https://www.bbc.co.uk/bitesize/articles/zqbrqyc>
- <https://www.bbc.co.uk/bitesize/articles/zyjbwsg>
- <https://www.bbc.co.uk/bitesize/articles/zgg86rd>
- <https://www.bbc.co.uk/bitesize/articles/zg2d6rd>
- <https://www.bbc.co.uk/bitesize/articles/zr9y9ty>

- <https://www.bbc.co.uk/bitesize/articles/zqdxxywx>
- <https://www.bbc.co.uk/bitesize/articles/zqqxywx>

Analisi di classe

Il docente inviterà la classe a condividere le riflessioni emerse dalla sfida dei quiz, le difficoltà affrontate e soprattutto i ragionamenti seguiti nel tentativo di individuare la natura delle immagini analizzate.

SECONDO ESERCIZIO



La musica che gira... in Corea

Kim Jong-un, leader della Corea del Nord, è diventato un fan del K-pop... oppure no?

1. **Cosa sappiamo della persona ritratta? È plausibile che venga fotografata in questo contesto?**
2. **Cosa possiamo dedurre dal dettaglio del cd? Chi è l'artista? A che cosa possiamo collegarlo?**

3. Che dettagli della foto ci possono insospettire?

Come abbiamo fatto nelle attività precedenti, proviamo a rispondere a queste domande facendo "brainstorming" di gruppo. Dopo aver condiviso le nostre riflessioni, utilizziamo la foto stessa per una **ricerca inversa**, caricando il file (disponibile in allegato) su Google Immagini e TinEye.

Facciamo attenzione alla tipologia dei siti che condividono questa immagine (in nessun caso comparirà una testata giornalistica, ad esempio), e alle altre ricorrenze proposte dal motore di ricerca (grazie alle quali potremo anche risalire all'immagine originale, non modificata).

Nel caso di dubbi su come utilizzare Google Immagini, ecco la guida (in italiano) proposta da Google News Initiative:

<https://newsinitiative.withgoogle.com/it-it/resources/trainings/reverse-image-search-verifying-photos/>

Per quanto riguarda TinEye, il sito è disponibile solo in inglese, ma è molto intuitivo (<https://tineye.com/>) e ci permette di selezionare varie opzioni, tra cui quella cronologica (prime ricorrenze, ultime ricorrenze).

Proviamo anche a cimentarci con uno strumento più complesso: FotoForensics (<https://fotoforensics.com>). In particolare, una tipologia di analisi offerta da FotoForensics ci è utilissima per capire **se e dove una foto è stata manipolata**: si tratta del parametro ELA, e cioè Error Level Analysis.

Questa griglia permette di verificare l'originalità di un'immagine/fotografia (JPG) in base al numero di errori che contiene. Gli "errori" sono visualizzati come la parte più chiara, devono essere distribuiti più o meno nello stesso modo su tutta l'immagine. Nel caso in cui compaiano dei pixel di altro colore - blu o rosso, ad esempio - la possibilità di manipolazione è molto alta.

Anche in questo caso, possiamo avvalerci dei tutorial di FotoForensics (un'avvertenza: si tratta di uno strumento usato da giornalisti investigativi e analisti, quindi il livello di complessità è alto!). Ma basterà caricare l'immagine di Kim Jong-un e selezionare il fattore analitico ELA per avere risultati interessanti...