Seminar Publications Series 01/2020 ISSN 2704-8969

Deepfake e post-verità: una nuova filiera della disinformazione

Stefano Ricci

International Institute for Global Analyses

Vision & Global Trends. International Institute for Global Analyses
Piazza dei Navigatori 22, 00147 – Rome (Italy)
Seminar Publications Series – N. 1/2020 – November ISSN 2704-8969

First Edition: April 2020

Seminar Publications Series – N. 1/2020
Seminario di studio: *Flussi dell'informazione e processi decisionali.*
Camera dei Deputati, Sala della Regina – 28 gennaio 2020

www.vision-gt.eu

Deepfake e post-verità: una nuova filiera della disinformazione

Buongiorno,

inizio la mia relazione permettendomi di fare una piccola appendice su quelli che sono stati i contenuti degli interventi precedenti, in particolar modo quando si è parlato dell'*overload informativo* cui è sottoposto l'individuo quotidianamente.
Partiamo da alcuni semplici numeri: ogni sessanta secondi, Facebook è attraversato più o meno da trecentocinquanta gigabyte di dati personali; sempre in questi sessanta secondi, i server di Google registrano due milioni di *query* di ricerca vengono quindi effettuate due milioni di ricerche.
Nell' arco di ventiquattro ore, Youtube registra più di quattro miliardi di pagine visualizzate e, sempre in queste ventiquattro ore, Twitter vede nascere duecento milioni di nuovi *tweet*.
Il 90% delle informazioni, dei dati, che sono stati prodotti a partire dall' avvento dell'era informatica è stato prodotto negli ultimi due anni.
Appare allora abbastanza evidente come in questo caos, in questa supremazia del disordine sull'ordine costituito dalla cultura sia facile generare terreno fertile per la diffusione delle *fake news*. Nella maggior parte dei casi, queste *fake news* tendono anche a sviluppare, a sfruttare le innovazioni tecnologiche che vengono nel corso del tempo presentate alla grande opinione pubblica.
Nell' arco degli ultimi anni, si è spesso parlato di apprendimento automatico, di *machine learning*, di sicurezza informatica, di computazione quantica e via discorrendo; ora, però, queste grandi innovazioni tecnologiche, che pur hanno un loro ambito d'applicazione, come possono andare ad alimentare il caos informativo di cui abbiamo appena parlato?
Uno degli esempi più eclatanti è quello del *deepfake*, salito agli onori delle cronache giusto negli ultimi due anni: il *deepfake* non è altro che la commistione di contenuti reali, quindi immagini o video o audio assolutamente reali, con dei contenuti (immagini, video o audio) artefatti al 100%.

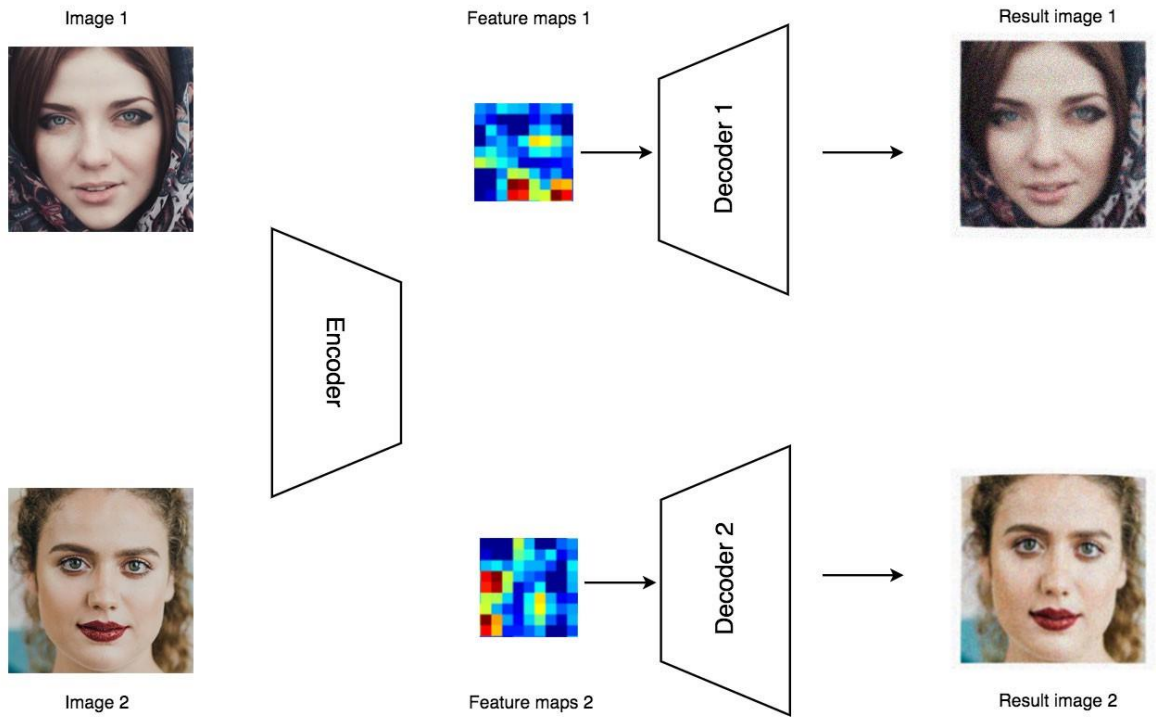

Figura 1: Come funziona il deepfake

Di fatto, questa tecnologia permette di prendere il corpo di un attore, montare sul corpo di quest'attore il viso di una potenziale vittima, della vittima che si vuol colpire tramite la diffusione del contenuto artefatto, e attraverso una serie di algoritmi e sistemi di apprendimento automatico, la macchina è in grado di montare il viso della vittima sul corpo dell'attore, per risultati estremamente realistici.

Il concetto alla base del *deepfake* nacque nel 2014 per mano di un informatico statunitense, Ian Goodfellow; il campo da gioco alla base di questa tecnologia era legato all'apprendimento da parte delle macchine del contenuto di uno specifico video, di una specifica foto: di fatto, si stava insegnando alla macchina a riconoscere i particolari che compongono un'immagine o un video.

Immaginando una fotografia, dunque, saper riconoscere una catena montuosa, piuttosto che l'esposizione solare, piuttosto che il viso dei componenti del contesto fotografico.

Ora, la situazione è cambiata nel 2017, quando sul principale aggregatore di *social news*, Reddit, apparve il contenuto di un utente rimasto ancora anonimo, tale "deepfakes", il quale offrì alla *community* degli strumenti totalmente *open source* per la realizzazione di *deepfakes* direttamente in casa propria.

Ciò cosa ha significato?

Ciò ha significato essenzialmente che la grande mistificazione informativa aveva abbandonato i laboratori delle università o delle grandi industrie tecnologiche, per approdare in quelli un po' più artigianali delle abitazioni private.
Questo, ovviamente, non vuol dire che non siano necessarie risorse informatiche di un certo livello, o capacità tecniche particolari; voglio soltanto dire come basti avere un software *open source*, cioè senza licenza, per poter realizzare il proprio *deepfake*.
Un esempio di *deepfake* l'abbiamo già visto in Italia su *Striscia La Notizia*, con un falso video su Matteo Renzi.

Figura 2: il deepfake realizzato da "Striscia La Notizia" con Matteo Renzi

Il *deepfake* divenne celebre nel 2018 perché venne utilizzato a sfondo satirico da parte di alcune trasmissioni televisive statunitensi; la mistificazione fu talmente perfetta che in molti caddero nella trappola del *deepfake*.
A questo punto potremmo anche iniziare una riflessione più squisitamente geopolitica: nel caso in cui dovesse apparire un *deepfake video* ritraente, ad esempio, il premier israeliano nell'atto di parlare a stretto giro con i suoi collaboratori in merito all' uccisione di un generale iraniano, cosa potrebbe accadere a livello internazionale?
Consideriamo di vivere anche nell' epoca della cosiddetta *post-verità*.

Figura 3: anche il Presidente degli Stati Uniti, Donald Trump, è finito "vittima" del deepfake

Grazie anche a questi 350 gigabyte di dati che vanno ad attraversare i server di Facebook o queste quattro miliardi di visualizzazioni da parte di YouTube, viviamo in un'epoca in cui la verità è passata in secondo piano rispetto a quella che è l'opinione generata su un determinato tema.
Quindi, a salire in sul podio, non è tanto la verità dell'informazione che sto cercando di veicolare, quanto soprattutto il responso che riesco a generare nel pubblico che va a leggere quella data notizia.
Diciamo che le tecniche di disinformazione a questo punto hanno trovato un validissimo alleato proprio nel *deepfake* perché va a minare uno dei concetti cardine del giornalismo, quello della verificabilità dell'informazione.
Non devo più semplicemente andare a ricercare l'autore di quel determinato articolo, verificare la fonte: [con il *deepfake*] io sono messo nella condizione di vedere qualcosa che di fatto non è reale.
Quanto tempo servirà affinché quel video venga bollato come non veritiero, come inaffidabile?
Circolano ancora moltissimi video fasulli, sui principali *social network*, che vengono invece ritenuti affidabili e che ritraggono Putin parlare in un certo modo di Donald Trump, ad esempio: il tempo di gestazione del video, dell'informazione sui canali digitali è molto elastico.
Consideriamo ancora un'altra cosa.
Il cellulare, lo smartphone costituisce per noi, nel bene o nel male, la principale porta di accesso al mondo esterno; quando sfogliamo l'home

page, ad esempio, di Facebook abbiamo circa due o tre secondi per accettare come vera o falsa una determinata informazione … perché tanto passiamo subito all'informazione successiva!
Quindi come riesco a verificare che quell'informazione che ho appena letto, o che quel video che ho appena appena visto, sia falso o sia veritiero?
Le soluzioni offerte, ad esempio, da *First Draft*, un sito di *debunking* legato all'Università di Harvard, o le istruzioni che sono state fornite da Facebook ai propri utenti, al momento appaiono un po' come superficiali e inutili: si parla di verificare l'autore [di un articolo], piuttosto che di andare a ricercare la stessa informazione in altre pagine, su altre testate giornalistiche.
Ricordiamoci però che tutta la discussione sui principali temi d'attualità avviene sullo smartphone, che non è propriamente la piattaforma ideale per poter copia/incollare un contenuto e andarlo a ricercare su un'altra piattaforma, rileggere la stessa informazione per vedere se ho perso o frainteso qualcosa, ricercare l'affidabilità di quel *deepfake* video e così via discorrendo.
Potremmo quasi definire il *deepfake* come uno dei principali strumenti che andranno a costituire il bagaglio, o meglio ancora, l'arsenale informatico della futura guerra informatica.
Anche sul tema della guerra informatica bisogna fare una piccola chiarezza: guerra informatica non vuol dire soltanto che da questo computer riesco ad accedere ad un altro computer, non è soltanto l'intrusione informatica.
L'intrusione informatica non è altro che una delle espressioni della guerra informatica.
La *cyber-warfare* di cui si parla spesso, in realtà, non è altro che la conquista e il dominio dello spazio virtuale del proprio avversario o dei propri avversari.
Quindi, se da un lato parliamo di dominio del singolo computer, ancor più sofisticato e fondamentale è il dominio dello spazio virtuale degli utenti che compongono quella determinata fetta sociale o di quel determinato Paese che voglio andare a colpire.
Non dimentichiamoci che tutte le cosiddette primavere arabe si sono diffuse attraverso *social media* e nella maggior parte dei casi, cosa di cui spesso tendiamo a dimenticare, nonostante i regimi avessero chiuso di fatto i rubinetti di internet; gli attivisti presenti nei Paesi vicini, o gli attivisti europei, hanno semplicemente acquistato degli strumenti da ventisette dollari, i Raspberry, cioè dei piccoli computer a singola scheda, e li hanno

tramutati in server, posizionandoli lungo il confine dei principali paesi colpiti da queste rivoluzioni.
Nel frattempo, proprio i *social network* hanno fornito lo strumento per organizzare queste mobilitazioni.
Gli attivisti, poi, utilizzavano piattaforme come lo stesso *Reddit*, dov'è apparso per la prima volta il *deepfake*, per organizzarsi.
Quindi, nel bene o nel male, e lo ripeto, i principali dibattiti d'attualità hanno tutti luogo in uno spazio che è privo di mediazione - lo smartphone - e all'interno di un contesto legislativo internazionale in cui è di fatto impossibile andare a normare la rete.
L'unico modo per controllare la rete, per controllare internet, è spegnerlo definitivamente; esso non può essere "controllato".
Lo vediamo anche con i servizi di *streaming*: se ne chiude uno, attraverso le piattaforme "idra", ne appaiono subito altri nove.
Quindi, internet non può essere controllato; persino in Cina, internet non può essere controllato appieno.
Google ha recentemente messo in guardia, seppur molto bonariamente, sul rischio che questi *deepfake* vadano poi a sfruttare la prossima, futura tecnologia di computazione quantistica, anche se la grande svolta commerciale avverrà con l'avvento del 6G, quindi fra circa sette-otto anni.
Nel caso in cui dovessimo utilizzare la computazione quantistica con il deepfake per disinformazione, a quali effetti potremmo assistere?
Una semplice riflessione può esser quella di immaginarci come tutti al cellulare in questo preciso istante, accedere tutti al medesimo link (magari riportante proprio la notizia relativa a questo convegno) e tutti noi assisteremmo semplicemente a notizie differenti: io avrei una notizia, i miei colleghi qui sul palco leggerebbero un'altra notizia prodotta semplicemente sulla base della profilazione della navigazione che si è svolta attraverso il loro cellulare.

Lo ripeto ancora una volta: parliamo di una realtà in cui YouTube è visualizzato per quattro miliardi di volte ogni giorno, e parliamo di visualizzazioni uniche; quindi, semplicemente immaginando che io vada su YouTube, veda un video e lo chiuda.

Come è stato già detto precedentemente, ci sono persone, soprattutto bambini, che passano diverse ore su YouTube e costituisce quella un'ottima

piattaforma di disinformazione o meglio di, come possiamo definirla, di alterazione della fase cognitiva.

Grazie.

Dott. Stefano Ricci
Sala della Regina - Palazzo Montecitorio - 28 Gennaio 2020

Relazione presentata al

Convegno

Flussi dell'informazione e processi decisionali

Martedì 28 gennaio 2020
Orario: 9:30 – 14:00

Sala della Regina – Camera dei Deputati – Roma

Ingresso: Piazza del Parlamento 25

L'interrelazione tra processi decisionali e flussi di informazione e comunicazione è una questione decisiva per il futuro della civiltà in una molteplicità di situazioni su scala planetaria (dalle grandi tematiche dei cambiamenti climatici globali a quelle della globalizzazione, a quella delle migrazioni) con la quale si confronta l'umanità intera.

Questa questione rileva anche a livelli meno macroscopici, ma non meno importanti relativi alla quotidianità sul fronte dei consumi, in particolare per quanto attiene ai comportamenti alimentari, e alla salute settore nel quale si riscontrano sia ingiustificate recriminazioni nei confronti del sistema sanitario nazionale, sia casi di preoccupante disinformazione come quello sulla presunta diffusione della meningite in Italia e addirittura diffidenza verso i risultati della ricerca scientifica e del rifiuto della responsabilità sociale come nel caso della contestazione dei vaccini.

I tempi attuali – che sembra siano l'era della **post-verità –** vedono alcune difficoltà particolari al riguardo.

Tralasciamo il punto di vista di chi sostiene che la conoscenza approfondita non sia un requisito indispensabile per assumere decisioni e consiglia di affidarsi a sensazioni da assecondare, intuizioni da seguire, opportunità da cogliere, relegando di fatto in questo modo, la razionalità ai margini. Anche per i sostenitori del tradizionale precetto "**conoscere per decidere**" si presentano domande difficili su quanto siano affidabili e non strumentali i contenuti e i processi che portano a costruire i convincimenti, sia degli individui sia dei gruppi sociali.

La diffusione di Internet, del web e dei social e anche più in generale la digitalizzazione e la stessa globalizzazione hanno reso più impegnativa la **selezione e la "pesatura" delle fonti** e contribuito a creare rischi di dissociazione dalla realtà e di circuiti chiusi e distinti. L'aumentata partecipazione, fenomeno in sé indubbiamente positivo, impone di converso che le conoscenze siano non solo accessibili, ma anche valutabili da parte di una vastissima platea di soggetti, anche quando la complessità dei temi da affrontare è proibitiva.

Inevitabilmente, tutti i soggetti sono portatori di interessi, ma alcuni sono più determinati di altri e hanno mezzi migliori per farli valere. La posta in gioco motiva una guerra senza esclusione di colpi da parte dei portatori di interesse economici o di supremazia a fini di potere e il mondo della politica non è certo estraneo alle dinamiche che si producono. Se si prende atto che non è lontana dal vero l'affermazione che i soggetti deboli (individui e comunità di cittadini) sono condizionati e strumentalizzati (in sintesi" usati) si pone allora l'esigenza di **approfondire i meccanismi** attraverso i quali ciò avviene e di provare a contenere la loro efficacia prendendo le opportune contro misure nella misura del possibile, cercando di evitare i due estremi di una rinuncia a priori associata ad oblio da difesa da una parte e di velleità rivoluzionarie non corredate di realistica progettualità alternativa, dall'altra.

In definitiva **la sfida è quella di comprendere e tentare di governare la complessità**, complessità dei contenuti e complessità dei processi, quest'ultima accentuata dall'essenza della democrazia che domanda come diritto-dovere una partecipazione informata e quindi studio, approfondimento confronto di opinioni diversificate.

Il Convegno è organizzato da Vision & Global Trends. International Institute for Global Analyses - si inserisce nel quadro degli studi dedicati ai temi della informazione e comunicazione nella cosiddetta era della post-verità.

Programma
Saluti istituzionali
On. Edmondo Cirielli – Camera dei Deputati, Questore
Dott. Tiberio Graziani –Vision & Global Trends, Presidente

Relazione introduttiva
Prof. Fabio Pistella – ENR Consigliere di Amministrazione – già Presidente Cnr e Direttore generale Enea
Si può governare la complessità dell'intreccio tra flussi informativi e processi decisionali?

Relazioni
C. Amm. Fabio Agostini – Capo Dipartimento Pubblica informazione e comunicazione presso Stato Maggiore della Difesa

Dott. Fabrizio Noli – Caposervizio Esteri Radio Rai
Per una storia dell'opinione pubblica

Prof. Mario Morcellini – Commissario dell'Autorità per le Garanzie nelle Comunicazioni e Consigliere alla Comunicazione Sapienza Università di Roma
Informazione vs decisione nello spazio pubblico mediatizzato

Dott. Stefano Ricci – Esperto di sicurezza cibernetica, saggista
Deepfake e postverità: una nuova filiera della disinformazione

Prof. Angelo Maria Petroni – Segretario generale, Aspen Institute Italia
Il ruolo delle teorie nell'era dei big data

Prof. Roberto Guarasci – UNICAL, Direttore del Dipartimento Culture educazione e società
*Relazioni e Terminologia nella costruzione dell'*affaire *Dominique Strauss-Kahn*

Dott. Giovanni Anzidei – Fondazione Igea Onlus, Vice Presidente – Giornalista scientifico
Dispersione nella comunicazione e disinformazione scientifica

Dott. Andrea Sonnino – Presidente Fidaf e Coordinatore dell'Osservatorio sul Dialogo nell'Agroalimentare
Il metodo partecipativo per l'assunzione di decisioni condivise nell'agroalimentare

Prof. Roberto Setola – Direttore Master Homeland Security, Università Campus Bio-Medico di Roma
Interconnessioni ed interdipendenze

Conclusioni
Dott. Filippo Romeo, Vision & Global Trends, Analista senior