SAPIENZA
UNIVERSITÀ DI ROMA

Sovranità digitale e guerra algoritmica. Un'analisi dello scenario della cybersecurity nell'era della propaganda computazionale e dei conflitti cibernetici

Facoltà di Scienze politiche, sociologia e comunicazione
Dipartimento di Comunicazione e Ricerca Sociale
Dottorato di ricerca XXXVIII° ciclo in Comunicazione, Ricerca Sociale e Marketing
Curriculum in Comunicazione

Arturo Di Corinto
Matricola 809414

Tutor
prof. Alberto Marinelli

Coordinatrice
Prof.ssa Francesca Comunello

A.A. 2024-2025

SOVRANITÀ DIGITALE E GUERRA ALGORITIMICA

Un'analisi dello scenario della cybersicurezza nell'epoca della propaganda computazionale e dei conflitti cibernetici

Arturo Di Corinto
Sapienza, Università di Roma, Italia

Indice

1. Abstract....3
2. Premessa....3
3. Introduzione....4
3.1 Struttura e obiettivi della tesi....9
4. La sovranità di Internet....10
4.1 Big Tech: la sovranità impossibile....11
4.2 Big State vs. Big Tech: prove di sovranità....14
4.3 Frammentazione vs. sovranità: i rischi....16
4.4 Sovranità digitale e autonomia strategica....23
4.5 Le minacce alla sovranità digitale....27
5. Internet, trincea della cyberguerra....30
5.1 Gli attacchi cibernetici....32
5.2 La Disinformazione....35
5.3 La minaccia ibrida....38
6. IA, cyberattacchi e disinformazione: la minaccia ibrida all'opera....41
6.1 L'avvento dell'Intelligenza Artificiale....44
6.2 Empatia Artificiale e sovranità digitale....46
6.3 Hackerare l'Intelligenza Artificiale....50
6.4 L'IA e la guerra algoritmica....55
6.5 L'IA prende di mira gli umani....58
7. Attacco alla mente....63
7.1 La Guerra cognitiva....64
7.2 Tecnologie della persuasione....69
7.3 Hacking cerebrale....71
7.4 Propaganda e disinformazione....72
7.5 La propaganda computazionale e gli eserciti di troll....75
8. Le interferenze hacker e la guerra algoritmica....78
8.1 Dal sabotaggio culturale alle guerre guerreggiate, come cambiano gli hacker....78
8.2 Origini, presente e futuro dell'hacktivism....82
8.3 Infowar, netwar, cyberwar: il ruolo degli hacktivisti nella guerra cibernetica....86

8.4 Infowar vs. cyberwar: la costruzione di uno spazio globale di espressione....89
8.5 L'occupazione "militare" dell'agenda mediatica....92
9. Gli hacker e gli attivisti nel conflitto russo-ucraino....95
9.1 Psy-Ops e attacchi cibernetici....95
9.2 Killnet, Legione e NoName(057)16, un caso di studio....101
9.3 Una struttura paramilitare per gli hacktivisti russi....108
9.4 NoName(057)16 e gli effetti della disinformazione in Italia....109
10. Gli hacker e gli attivisti nel conflitto Israele-Hamas....113
10.1 La Cyberguerra in terra di Palestina....113
10.2 Attacchi informatici e operazioni di influenza: pro-Pal vs. Israele....113
10.3 L'uso della propaganda nel conflitto palestinese....116
10.4 L'apporto di Big Tech alla guerra in Medioriente....121
10.5 La Cyberguerra Israele-Iran....124
11. Conclusioni....128
12. Le interviste....134
12.1.1 Michele Mezza....135
12.1.2 Michele Colajanni....138
12.1.3 Massimo Marotti....140
12.1.4 Barbara Carfagna....141
12.1.5 Marco Ramilli....144
12.1.6 Luca Sambucci....148
13. Sitografia....153
14. Bibliografia....153
15. Ringraziamenti....180

1. Abstract

La sovranità digitale, che implica l'autodeterminazione nazionale nella governance del cyberspazio, rappresenta un rischio per Internet, il cui buon funzionamento richiede apertura e collaborazione fra molti soggetti. Viceversa, l'assenza di sovranità su dati, software, tecnologie e infrastrutture, può esporre i singoli Stati al rischio di dipendenza tecnologica e sottosviluppo economico, instabilità sociale e perdita di autonomia politica.

Gli attacchi cibernetici, la disinformazione globale, l'uso letale dell'Intelligenza Artificiale, sono le diverse espressioni in cui si declina oggi la guerra algoritmica e rappresentano, tutte insieme, la maggiore minaccia alla sovranità digitale. Questi pericoli infatti mettono a rischio la democrazia, lo sviluppo economico, il benessere dei cittadini e il mantenimento della pace che oggi poggiano grandemente sull'esistenza di un cyberspazio sicuro e aperto.

Uno scenario questo che può tuttavia essere contrastato attraverso politiche adeguate e un'aumentata cooperazione fra attori statuali e non statuali.

Per farlo è però necessario essere consapevoli delle minacce alla sovranità digitale portati dalla guerra algoritmica che vede una costante mutazione degli attori che ne sono protagonisti.

In questo lavoro si proverà a definire il contesto di tale guerra descrivendo la minaccia rappresentata dagli attacchi cibernetici portati da hacker e hacktivisti e dall'uso della disinformazione unita allo sviluppo di tecnologie emergenti come l'Intelligenza Artificiale.

Il lavoro si basa su una serie di casi di studio, il conflitto russo-ucraino e quello tra Israele e Hamas, e poggia, nella sua parte empirica, sulle riflessioni degli esperti interpellati, testimoni privilegiati dei fenomeni oggetto della trattazione.

KEYWORDS

Sovranità digitale, cybersecurity, cyberwar, disinformazione, algoritmi, Anonymous hacktivism, infowar, Intelligenza Artificiale, netwar, persuasione, propaganda computazionale

2. Premessa

Nei primi giorni d'estate del 2021 il governo nigeriano annunciava con un tweet di aver sospeso in tutto il Paese, a tempo indeterminato, Twitter (oggi X), in seguito alla rimozione di un controverso post del presidente nigeriano Muhammadu Buhari. Secondo la piattaforma di microblogging che lo aveva deciso, il presidente aveva violato le regole del

social; per le autorità nigeriane, invece, la piattaforma sarebbe stata responsabile di alimentare la crisi sociale nel paese. Il dibattito che ne è scaturito è servito a segnalare al grande pubblico alcune questioni cruciali dell'era della piattaformizzazione, cioè il modo in cui le piattaforme online, come i social media, i motori di ricerca e i marketplace digitali, stanno trasformando le nostre società, influenzando le relazioni economiche, politiche e sociali (Van Dijck, Poell, & De Waal, 2019). Se infatti le aziende private possono stabilire le regole d'uso dei propri servizi, se decidono restrizioni, accessibilità dei contenuti e copertura geografica, i governi devono accettarlo? E che succede quando invece sono i governi a imporre le regole all'uso che facciamo di Internet e dei suoi servizi? Ma soprattutto, che succede quando Internet diventa terreno di scontro e strumento della geopolitica trasformandosi in teatro di guerra?

3. Introduzione

Democracy is a socio-technical system. And all socio-technical systems can be hacked.
Bruce Schneier

Il 6 gennaio 2021 una folla esagitata prende d'assalto il Campidoglio di Washington, sede del potere legislativo degli Stati Uniti e simbolo della democrazia americana. La folla, richiamata da un tam-tam mediatico, chiede l'annullamento delle elezioni che hanno decretato la sconfitta del tycoon ed ex presidente degli USA Donald Trump nella corsa alla Casa Bianca, ritenendole truccate. Al grido di #stopthesteal, la folla, accorsa in seguito ai post incendiari di Trump sui social network di estrema destra Gab e Parler, e poi su Twitter, travolge i pochi agenti della sicurezza presenti e si riversa nelle sale del palazzo. Il bilancio finale sarà di cinque morti. Dopo averla aizzata con false dichiarazioni relative al risultato elettorale sarà lo stesso Trump a invitare la folla alla nonviolenza (Filippi, 2024).

Accusato di diffondere disinformazione e incitare alla violenza in seguito all'assalto del Campidoglio, e per questo bandito da Facebook e Twitter (Rodriguez, 2021), Donald Trump decide di voler essere lui a stabilire le regole di conversazione nell'infosfera social, e si fa costruire un proprio social network, Truth,[1] per raccogliere consensi e finanziamenti: andrà ufficialmente online il 21 febbraio 2022.

[1] Truth Social è una piattaforma di blogging di proprietà di Trump Media & Technology Group (TMTG), società americana di media e tecnologia la cui maggioranza è di proprietà del presidente degli Stati Uniti Donald Trump.

Pochi giorni dopo la Russia invade l'Ucraina: è il 24 febbraio 2022. Le operazioni militari di terra sono precedute da un attacco informatico che disabilita parte della rete satellitare Via-Sat. Agli attacchi cibernetici successivi segue un'intensa opera di propaganda: molte nazioni europee e gli USA chiudono l'etere e bloccano i servizi digitali di informazione russi (Ottaviani, 2022). La Federazione Russa fa lo stesso per rappresaglia e minaccia di staccarsi dalla rete Internet globale.

Il 22 ottobre 2022 Elon Musk, magnate di origini sudafricane, compra Twitter per 44 miliardi di dollari, licenziando immediatamente il Ceo Parag Agrawal, apparentemente per non aver voluto dichiarare il numero di account falsi sulla piattaforma. In precedenza, aveva denunciato il pericolo che il social media si trasformasse in "camere di risonanza della destra o della sinistra che generano odio e dividono il paese" (ANSA, 2022). A novembre Musk riabilita l'account di Donald Trump prima bandito dalla piattaforma a cui ha cambiato nome in "X". I suoi detrattori lo accusano di volere acquisire una platea mediatica globale per favorire i propri interessi e posizionare "Xcorp", l'insieme delle iniziative imprenditoriali di Musk (X/Twitter, Tesla, SpaceX), nella competizione per la sovranità dei dati che, come vedremo, sta ipotecando il futuro della Rete.

Il 30 novembre 2022 viene reso pubblico e gratuito l'uso di ChatGPT[2], chatbot di Intelligenza Artificiale basato su un Large Language Model (LLM)[3] addestrato usando, senza richiederne il consenso, i dati pubblici o semipubblici presenti in Rete. Grazie agli input di milioni di utenti, il chatbot acquisirà preferenze, gusti e dati personali di una platea notevolissima di individui. Un salto di qualità enorme rispetto alla possibilità di profilare i netizen con strumenti di messaggistica come WhatsApp e di orientarne i comportamenti attraverso i social network, come accaduto nello scandalo di Cambridge Analytica, dal nome dell'azienda che ha fornito ai sostenitori della Brexit e al team elettorale di Trump i profili dei loro potenziali elettori (Wilye, 2019). L'uso delle intelligenze artificiali generative come ChatGPT porrà virtualmente nelle mani di chiunque strumenti di disinformazione capaci anche di creare nuove armi cibernetiche.

Nel frattempo, accadeva lo stesso in altre regioni del mondo, con il conglomerato cinese

[2] ChatGPT, acronimo di Chat Generative Pre-Training, è un chatbot basato sull'Intelligenza Artificiale generativa creato dall'azienda Open-AI.

[3] LLM o Large Language Model, un sistema di Intelligenza Artificiale progettato per elaborare e generare testo in linguaggio naturale. Gli LLM vengono addestrati su enormi quantità di dati testuali, come libri, articoli, conversazioni online, imparando a riconoscere schemi, strutture e relazioni tra le parole attraverso un procedimento di machine learning.

BATX (Baidu, Ali Baba, Tencent, Xiaomi) obbediente alle leggi del Partito Comunista al potere in Cina, e con gli equivalenti russi come VKontakte, Telegram, Yandex e altri nella Russia di Vladimir Putin: entrambi i paesi avevano cominciato a rastrellare i dati dei loro utenti, chiedendo di localizzarli sul proprio territorio adducendo questioni di sovranità.

In Cina, nel settembre 2021 era stata implementata la legge sulla sicurezza dei dati, una normativa, la Reiwa 3 (Teti, 2024) per trattare i dati all'interno del territorio cinese, obbligando anche le aziende straniere a rivedere e modificare le proprie politiche in merito. Il primo novembre 2021 invece entrava in vigore la nuova legge cinese sulla privacy, la "China Personal Information Protection Law". L'obiettivo essendo di separare i dati critici da quelli che generano valore in un mercato di un miliardo e mezzo di consumatori e disciplinare così la sua digital economy per gli anni a venire.

Nel novembre del 2023 Israele decide di scollegare 15 Internet provider della Striscia di Gaza come rappresaglia per il brutale attacco terroristico subito dal proprio Paese il 7 ottobre ad opera del partito armato di Hamas, lasciando senza Internet mezzo milione di persone (Access Now, 2023).

Il 24 agosto 2024 Pavel Durov, fondatore di Telegram, viene arrestato dalle autorità francesi per "complicità nelle attività illegali che vi si svolgono" (BadSha, & Reuters, 2024). La piattaforma, che in maniera pseudo-anonima permette di diffondere informazioni e comunicare in tempo reale gli sviluppi dei teatri di guerra, dall'Ucraina alla Striscia di Gaza a Teheran, come pure di creare al suo interno mercati criminali, è ritenuta complice di pedofilia e traffico di droga dalle Autorità francesi che ne richiedono le chiavi segrete di accesso.

Il giorno 8 gennaio 2025 il Ceo di Meta, Mark Zuckerberg, con un video su Facebook, social network di sua proprietà, dichiara l'intenzione di porre fine al fact checking su Facebook, Instagram e Threads, ispirandosi alle community notes di X/Twitter (Körömi, Haeck, & Cheslow, 2025).

Alla cerimonia del suo insediamento, il 20 gennaio 2025, il neoeletto presidente americano Donald Trump invita gli oligarchi del digitale: Jeff Bezos (Amazon), Mark Zuckerberg (Meta), Elon Musk (SpaceX), Tim Cook (Apple), Sundar Pichai (Microsoft) e Shou Chew (TikTok). Secondo gli osservatori, la foto che li ritrae tutti insieme rappresenta simbolicamente la

subordinazione dell'oligopolio digitale americano al potere esecutivo, secondo altri, la saldatura definitiva tra Big Tech e Big State.

Il 5 marzo 2025 a Varsavia il Consiglio informale dei Trasporti, Telecomunicazioni ed Energia (TTE), presieduto dalla Polonia, si conclude con l'adozione unanime della così detta *Dichiarazione di Varsavia*, la richiesta cioè di una rapida adozione del *Cybersecurity Blueprint* europeo per la gestione delle crisi informatiche a fronte del complesso panorama delle minacce. Obbiettivo è irrobustire le reti esistenti, migliorare la collaborazione tra le organizzazioni europee e rafforzare la cooperazione civile-militare nel settore cibernetico anche con la NATO. Il 6 giugno 2025 il Cybersecurity Blueprint viene adottato dall'Unione Europea (European Union, 2025).

Il 22 giugno del 2025 Donald Trump annuncia su Truth di aver ordinato un attacco aereo di successo contro tre siti nucleari iraniani, per azzerarne la capacità offensiva. Solo in seguito, si terrà una conferenza stampa ufficiale da parte del Segretario alla Difesa USA, Peter Hegseth, per darne l'annuncio.

Il 24 giugno 2025 Vladimir Putin firma l'ordine di creazione di un sistema di messagistica di Stato per sostituire Telegram e WhatsApp nell'uso che ne fanno i cittadini russi (Tass, 2025). La piattaforma dovrà avere nuove funzionalità d'uso e di sicurezza, con l'effetto di consegnare i dati dei cittadini russi al controllo governativo.

Tra il 24 e 25 giugno 2025, durante il vertice NATO tenutosi all'Aia, si verificano oltre 100 attacchi DDoS[4] orchestrati dal gruppo hacktivista filorusso NoName057(16) che colpisce 12 siti municipali olandesi e provoca un'interruzione di tre ore del portale informativo della NATO nei Paesi Bassi (Messina, 2025).

Il 28 giugno 2025 la Germania chiede ad Apple e Google di bandire dai loro store online il sistema di Intelligenza Artificiale cinese DeepSeek (Kharpal, 2025).

Nei primi giorni del luglio 2025 la Cina lancia una nuova piattaforma centralizzata di identità virtuale. Sviluppato dal ministero della Pubblica Sicurezza e dalla potente Amministrazione del Cyberspazio, il sistema sostituisce il modello decentrato di verifica dell'identità utente

[4] DDoS, Distributed Denial of Service attack, *Attacco distribuito da negazione di servizio*, ovvero blocco distribuito dei servizi web, causato da numerose richieste di accesso illegittime al servizio esposto sul Web.

con un'infrastruttura statale unificata che consente allo Stato di mappare in tempo reale le attività digitali di ogni utente, senza passare da intermediari (Lamperti, 2025).

Il 6 luglio 2025, Elon Musk, in seguito a un sondaggio favorevole sulla piattaforma social "X", annuncia il lancio di un nuovo partito, l'America Party, sfidando l'ex amico Donald Trump.

A metà luglio 2025, Youtube rimuove 11 mila canali usati per campagne di disinformazione. Parallelamente, anche Meta, società madre di Facebook, WhatsApp e Instagram, dichiara di aver cancellato circa dieci milioni di profili nella prima metà dell'anno per attività di impersonificazione e diffusione di contenuti ''spam'', cioè di contenuti "spazzatura". L'offensiva delle Big Tech contro la disinformazione si inserisce in un contesto più ampio: recenti indagini congiunte di polizia e intelligence hanno smascherato e disarticolato una rete di hacker filorussi responsabile di attacchi a infrastrutture critiche nell'UE (Europol, 2025).

Il 18 luglio 2025 una comunicazione della Nato annuncia di voler rispondere agli attacchi informatici russi. Il comunicato denuncia il comportamento irresponsabile della Federazione Russa ritenuta colpevole di attività di spionaggio e sabotaggio condotte dai servizi segreti militari russi nei confronti di USA, Francia, Inghilterra, Germania, Estonia e altri paesi dell'Alleanza atlantica. "Risponderemo a queste minacce nel momento e nei modi che riterremo opportuni, in conformità con il diritto internazionale e in coordinamento con i nostri partner internazionali, inclusa l'UE", è la chiosa del comunicato (Nato, 2025). Le aggressioni cibernetiche nel così detto *quinto dominio*, il cyberspace, possono innescare il meccanismo di solidarietà tra gli alleati e legittimare una risposta difensiva, anche di tipo cinetico (Cernicchiaro, 2022).

È difficile non ravvisare in tutti questi episodi i contorni di uno scontro totale, che chiameremo Guerra Algoritmica, cioè l'uso di software e algoritmi[5], Intelligenza Artificiale, dispositivi informatici e piattaforme digitali per sviluppare strategie politiche ed economiche, ma anche interventi militari, operazioni cibernetiche e di influenza informativa. Software e algoritmi possono infatti essere usati per penetrare difese informatiche, rubare dati, o manipolare le informazioni per influenzare la percezione pubblica degli eventi. Questa nuova

[5] L'algoritmo è un concetto fondamentale dell'informatica, alla base della nozione teorica di calcolabilità. L'algoritmo è anche un concetto cardine nella fase di programmazione dello sviluppo di un software: preso un problema da automatizzare, la programmazione costituisce essenzialmente la traduzione o codifica di un algoritmo per tale problema in programma, scritto in un certo linguaggio, che può essere quindi effettivamente eseguito da un calcolatore rappresentandone la logica di elaborazione.

era della guerra si basa sulla capacità di macchine "intelligenti" di analizzare enormi quantità di dati, prendere decisioni, e agire in modi che sopravanzano le capacità umane in termini di velocità e complessità, estraendo informazioni generate da personal media e social network, e arruolando le Big Tech nei conflitti moderni che usano gli utenti di Internet come terminali di senso (Mezza, 2022) e soldati passivi nella guerra dell'informazione (Mhalla, 2025).

3.1 Struttura e obiettivi della tesi

Obbiettivo di questo lavoro è di analizzare come gli attacchi cibernetici, la disinformazione globale, l'uso letale dell'Intelligenza Artificiale, rappresentino una minaccia alla sovranità digitale. La sua prima parte, fino al quarto capitolo, si propone di definire l'importanza della sovranità digitale in un'epoca caratterizzata da forti tensioni geopolitiche illustrando i contorni di una competizione globale che assume le forme della guerra algoritmica, una guerra basata su software, algoritmi, e piattaforme digitali, parte essenziale dell'obbiettivo dell'analisi.

La seconda parte, nei capitoli cinque, sei e sette, ha la funzione di chiarire il significato di guerra cibernetica e guerra algoritmica, di definire il concetto di guerra cognitiva, di differenziare i concetti i contigui di infowar, netwar e cyberwar; di raccontare come i civili siano a vario titolo coinvolti nei conflitti cibernetici, di definire l'evoluzione dell'hacktivism come fenomeno sociale e descrivere in che modo gli attori delle minacce cibernetiche si sovrappongono e scambino di ruolo continuamente.

Nella terza parte, cioè nei capitoli otto, nove e dieci, si illustreranno dei casi di studio relativi al ruolo di hacker e hacktivisti nelle guerre guerreggiate sul fronte ucraino e su quello mediorientale per dimostrare che quello che pare un dominio separato, il mondo cyber, ha una prosecuzione nel mondo fisico e ne condivide gli impatti e le conseguenze.

La quarta parte e ultima parte è dedicata alle interviste rivolte agli esperti per mettere alla prova le ipotesi della guerra algoritmica quale minaccia alla sovranità digitale. Si tratta di ambasciatori, giornalisti, accademici, imprenditori e tecnologi, testimoni privilegiati dei fenomeni osservati.

Seppure si sia scelto di storicizzare il fenomeno dell'hacktivism per spiegare come si manifesta ai giorni nostri, la descrizione dell'utilizzo di cyber-weapons e dell'intelligenza artificiale nelle tecniche, tattiche e procedure delle guerre cibernetiche ha reso necessaria l'analisi di fonti non accademiche, facendo ricorso, per l'illustrazione dei temi trattati, a una notevole quantità di letteratura grigia.

L'elencazione cronologica di alcune campagne di attacco cibernetico da parte degli attori delle minacce, come quelle portate dagli hacktivisti filorussi, ci è parsa utile a identificare precisi collegamenti tra le vicende politiche e le campagne di sabotaggio cibernetico che ne

rappresentano la reazione strategica. Si tratta ad ogni modo di una porzione assai limitata dell'insieme dei cyberattacchi prodotti nel tempo, usata allo scopo di esemplificare il fenomeno.

Le numerose autocitazioni, invece, hanno lo scopo di favorire la comprensione dei fenomeni in oggetto per approfondirne la prospettiva in assenza di riferimenti più puntuali e chiarire la profondità di una riflessione decennale intorno ai temi trattati. Tuttavia, nonostante la particolare posizione dell'autore, consigliere presso l'Agenzia per la cybersicurezza nazionale, si è scelto di utilizzare soltanto dati e informazioni di carattere pubblico per rispettare le clausole di non divulgazione proprie della funzione svolta.

4. La sovranità di Internet

It looks like we're moving back to a world where what you can see and who you can talk to is a function of what software and hardware you use. And that, in turn, increasingly will depend on where you live.

Mark A. Lemley

Internet è nata e si è imposta in un contesto a bassa regolazione che ha permesso all'innovazione che cresceva ai suoi margini di prosperare e creare nuove opportunità per i cittadini globali della Rete, i cosiddetti netizen. Per lungo tempo sono state le comunità tecniche a decidere il funzionamento di Internet e del Web, inglobando nei protocolli e nelle tecnologie di rete un'idea di apertura, interoperabilità e sviluppo condiviso. Tutto questo è stato fatto con l'ausilio progressivo delle aziende di telecomunicazione e un controllo "soffice" e discreto dei governi che forse non ne avevano ancora compreso l'impatto dirompente. Quando però i governi hanno colto il potenziale della Rete, hanno incominciato a usare la propria forza statuale per imporre la loro idea di sovranità alla Rete stessa, in un processo continuo e incrementale oggi reso più minaccioso da un contesto in cui i rapporti di forza, le esigenze di sicurezza e le dinamiche geopolitiche, prevalgono sulla diplomazia e sulle logiche di apertura, cooperazione internazionale e globalizzazione dei mercati che avevano consentito la maturazione stessa della Rete.

L'idea di sovranità presuppone infatti l'autorità su un'entità politica, che nei secoli è stata associata al concetto di Stato grazie alle riflessioni di Niccolò Machiavelli, Thomas Hobbes, Carl Schmitt e molti altri. La sua concettualizzazione più generale ingloba i concetti di

autorità, legittimazione, supremazia e territorio, quest'ultimo inteso non solo come massa geografica o regionale, ma come l'insieme di risorse presenti sopra, sotto, o nello spazio prospiciente alla sua superficie, e si estende fino alle infrastrutture umane; la parola digitale, invece, rimanda a tutto ciò che, infrastrutture, dati e contenuti, si basa sull'uso dell'informatica. È dall'unione di queste due dimensioni, sovranità e digitale, che emerge il concetto prismatico di *sovranità digitale* che, applicato alle logiche di Rete, ingloba l'idea stessa della sovranità di Internet - inizialmente riferita alla gestione delle sue infrastrutture e ai protocolli abilitanti la comunicazione digitale -, e riguarda anche la creazione, l'immagazzinamento, l'analisi e la condivisione di dati e informazioni da essa rese possibili. È proprio in virtù di tale interpretazione che la nozione di sovranità digitale viene usata "per veicolare l'idea che gli Stati debbano riaffermare la propria autorità su Internet per proteggere cittadini, istituzioni e affari dalle molteplici sfide all'autodeterminazione della propria nazione nella sfera digitale odierna" (Musiani, 2022).

4.1 Big Tech: la sovranità impossibile

Chi controlla i dati avrà il controllo sul mondo, in futuro
Narendra Modi

Come spiega Luca Sambucci (intervista in appendice), la sovranità digitale implica il controllo dei diversi strati che intrecciano il ciclo di vita di un dato. L'infrastruttura, con cavi e data center, l'hardware, con chip e sensori, il software, con le piattaforme di gestione e gli algoritmi che elaborano il dato trasformandolo in informazione e quindi in decisione.

Eppure, la sovranità digitale non si esaurisce nella mera capacità tecnica di possedere o localizzare questi elementi entro i confini nazionali. Si estende alla possibilità di esercitare un controllo normativo, politico ed economico su di essi. Significa, in altre parole, poter decidere chi può accedere ai dati e in che modo possono essere trattati.

A insidiare l'autodeterminazione degli Stati nel ciclo di vita dei dati all'interno della sfera digitale sono intervenute negli ultimi anni le grandi multinazionali tecnologiche, le Big Tech americane e, segnatamente, Alphabet, Amazon, Apple, Meta, Microsoft, che sono diventate più potenti e più ricche di intere nazioni grazie alla raccolta e al trattamento dei dati generati dagli utenti online, in virtù di un processo in cui si sostituiscono agli Stati stessi nell'erogazione di servizi fondamentali e nel soddisfacimento di alcuni bisogni dei cittadini, come sostengono Van Dijck e altri (2013, 2018). Le Big Tech agiscono infatti come piattaforme, cioè architetture digitali che svolgono il ruolo di infrastrutture di connessione su

cui altre piattaforme di settore si appoggiano mettendo in relazione utenti, corporation e istituzioni pubbliche, agendo così da gatekeeper nel primo caso e, nel secondo caso, da mediatori tra chi produce contenuti e servizi e chi li vuole consumare. Connettendo servizi e individui, queste piattaforme, tuttavia, non riflettono, ma producono le strutture sociali che viviamo, modellando le pratiche quotidiane proprio attraverso la raccolta sistematica, il trattamento algoritmico, la circolazione e la monetizzazione dei dati degli utenti (Van Dijck et al., 2019).

La globalizzazione dei commerci e gli effetti della rivoluzione informatica sono le forze che hanno permesso l'emergere di questi nuovi attori che, come argomentano Anzera e Massa (2023), agiscono un indiscusso *soft power*, anche di fronte agli Stati, proprio per la loro capacità di gestire la diffusione d'informazione e la produzione di contenuti, attraverso meccanismi di selezione, mercificazione e datificazione che, pure orientati al profitto, influenzano vaste dinamiche sociali e politiche, esercitando proprio quel tipo di potere che, secondo il politologo Joseph Nye, può essere più pericoloso di quello di uno stato avversario, in quanto nel processo di connessione degli attori sociali ed economici si costruiscono nuove forme di valori e nuove forme economiche con rischi ambientali, politici ed etici potenzialmente gravi (Van Dijck et al. 2019).

Come afferma il professore Michele Colajanni (2018): "Il sistema si fonda prevalentemente sull'uso disinvolto dei dati personali da parte dei provider digitali e su una deregolamentazione che, ai tempi di un World Wide West, in cui la Legge ha oggettive difficoltà applicative, ha perso qualsiasi risvolto romantico e utopico dei primi decenni. L'accordo è chiaro: 'Io ti offro servizi e giochi gratuiti, tu paghi consentendomi di acquisire, analizzare e rivendere i tuoi dati. Se ti sta bene, accetti (opt-in); altrimenti non puoi usufruire dei miei servizi (opt-out)'. La quasi totalità degli utenti preme il tasto di accettazione senza neanche guardare i termini del contratto. I pochi che lo fanno, dopo aver letto e compreso in parte, accettano con analoga probabilità. Il desiderio del servizio "gratuito" supera ogni perplessità sulla privacy, trasformando i consumatori di servizi in utili fornitori di dati. Il meccanismo funziona molto bene, per cui si è sviluppata un'economia rapace basata sui dati personali, sociali, aziendali. Non solo da parte dei provider digitali, ma di una miriade di aziende che fioriscono operando in uno o in entrambi dei seguenti ruoli: a) data mining, aziende preposte all'acquisizione di dati da qualsiasi fonte, che sia aperta o chiusa, legale o grigia, fino a quelle che si spingono nel dark web per l'acquisto di dati provenienti da furti informatici; b) data broker, aziende che aggregano dati per valorizzarli in informazioni da

vendere ad altri data broker, a investitori pubblicitari o a chiunque sia interessato a comprarle".

Argomenti che sono il punto di partenza della riflessione sviluppata da Shoshana Zuboff nel libro *Il Capitalismo della sorveglianza (2019),* per intendere il nuovo tipo di capitalismo emergente, basato sul controllo del comportamento individuale attraverso la raccolta e la mercificazione dei dati, alternativo a quello industriale. Un tipo di capitalismo che, forse più precisamente, era stato definito *capitalismo cognitivo* da parte dei teorici post-operaisti italiani (Cillario & Finelli, 1998) e francesi (Moulier-Boutang, 2002) negli anni '90, anticipati a loro volta da Guy Debord (1992) e dal Movimento situazionista, che avevano descritto il carattere di estrazione di valore nella produzione di senso elaborata nei circuiti della creatività sociale e dell'economia immateriale. Frutto della circolazione e della successiva appropriazione di idee, saperi, cultura e conoscenza da parte del sistema industriale, e in particolare dell'industria creativa, questo tipo di capitalismo è rappresentato, secondo Lorenzo Cillario, proprio da "l'espropriazione delle finalità, nonché delle modalità di decisione, di controllo e di riproduzione del sapere, dei processi cognitivi che lo presiedono, delle pratiche comunicative e informative che lo coordinano e lo distribuiscono" (Cillario, 1990, pp. 69-81). Pur nella diversità degli argomenti e del framework interpretativo, situato in momenti storici differenti, si tratta in entrambi i casi di un capitalismo fondato sul controllo e la *messa al lavoro* dei comportamenti basati sul linguaggio e la comunicazione, strutture di senso e di relazione trasformate in dati da parte di pochi conglomerati industriali-statuali. Questa sussunzione del lavoro cognitivo, prodotto attraverso i circuiti della comunicazione sociale, e la sua conseguente trasformazione in profitto e immaginario, verrà descritta da Byung Chul Han come *psicopolitica,* laddove, secondo il filosofo coreano, il regime disciplinare *biopolitico* teorizzato da Michel Foucault, cioè il potere esercitato sul corpo sottomesso a un regime di sorveglianza e produzione, si rovescia in qualcos'altro: l'irreggimentazione delle facoltà cognitive attraverso le dinamiche della visibilità in cui "ad essere visibili non sono i dominatori ma i dominati", in quanto "la tecnica informatica digitale rovescia la comunicazione in sorveglianza: quanti più dati generiamo, quanto più intensivamente comunichiamo, tanto più efficiente diventa la sorveglianza". Un paradosso evidente in cui è proprio la libertà a garantire il dominio, dato che nella società dell'informazione gli ambienti isolati del meccanismo disciplinare si dissolvono in reti aperte e comunicanti (Han, 2023). Un meccanismo prodotto in ultima analisi dal narcisismo voyeuristico stimolato e ingegnerizzato dai social network che, secondo Raffaele Simone, psicolinguista emerito, individua in esso le dinamiche di un'ansia da prestazione sociale, come se non esistessimo fuori dello schermo, "come se fossimo inutili senza un palcoscenico anche quando il pubblico in sala è poco ma la voglia di fare parte dello show è molta".

Un'ansia di rappresentazione che ha creato il "mostro mite" (Simone, 2008), e cioè un regime globale di governo che si basa su un sistema mediatico, televisivo, culturale, cognitivo, che crea un ambiente festoso e "infantilizzante" che pesa su tutta la società, seguendo i codici di una ludicizzazione della comunicazione che produce dati per l'industria estrattiva del web, al fine di profilare i comportamenti, segmentare i consumatori, targhettizzare gli elettori, addomesticare il dissenso: il *Platform capitalism* nelle parole di Benedetto Vecchi (2017).

I dati, raccolti e aggregati dalle grandi piattaforme online, generano infatti un valore che è connesso alla velocità, varietà e volume del loro trattamento, favorendo lo sviluppo di algoritmi che a loro volta possono rivelare scelte, comportamenti, azioni, gusti, correlazioni, per elaborare "modelli, altamente predittivi, di domanda e offerta dei prodotti, servizi e contenuti di vario genere" (Delmastro & Nicita, 2019).

Una delle spinte più forti alla riaffermazione del potere statuale sui dati generati da cittadini, imprese, pubbliche amministrazioni, viene così, a ben vedere, dalla reazione alla crescita ipertrofica di un ecosistema digitale, come quello di Internet e del Web, in cui, secondo l'ex direttore dell'intelligence italiana, prefetto Alessandro Pansa: "Gli Stati si sono distratti e adesso quattro o cinque multinazionali sanno più cose dei nostri cittadini di quante ne sappiano i servizi segreti"[6]. Un vulnus profondo alla sovranità digitale.

4.2 Big State vs. Big Tech: prove di sovranità

In un senso più ampio, la sovranità digitale viene associata alla capacità di individui, organizzazioni ed enti governativi, nazioni e federazioni di Stati, di controllare il proprio destino digitale, e implica il controllo sulla propria infrastruttura, sui dati e sulle tecnologie che utilizzano. Essa riguarda, inoltre, la possibilità di garantire la propria autonomia nel regno digitale, consentendo alle diverse entità di agire, innovare e prendere decisioni in autonomia senza subire un'indebita influenza esterna o senza dipendere da entità straniere.

Come dice nella nostra intervista l'ambasciatore Massimo Marotti dell'Agenzia per la cybersicurezza nazionale, ACN: "La sovranità digitale è una forma di assicurazione contro i rischi estremi dell'impiego di Internet. In un mondo ad economia integrata l'assenza di

[6] Direttore generale del Dipartimento per le informazioni e la sicurezza della Repubblica, DIS. Dichiarazione durante il convegno Itasec 2018 a Milano.

sovranità digitale per uno stato medio introduce una ulteriore forma di dipendenza nell'esercizio delle proprie funzioni ed in quello delle attività umane della propria comunità".

Un concetto questo che si estende a vari aspetti della società e dell'economia, tra cui la gestione dei dati, la sicurezza informatica e lo sviluppo delle tecnologie.

Dello stesso avviso è l'imprenditore esperto di digitale Marco Ramilli: "Affidare lo sviluppo, la gestione e l'innovazione dei sistemi digitali a fornitori esterni, spesso legati ad altri interessi geopolitici, espone l'Italia a un rischio strutturale: quello di non poter garantire pienamente la propria autonomia operativa, né la riservatezza e la resilienza dei propri dati e delle proprie attività strategiche".

Secondo la nota giornalista tecnologica, inviata della RaiTV, Barbara Carfagna, "La sovranità digitale è la capacità di uno Stato di imporre la propria architettura su dati, infrastrutture, tecnologie critiche e piattaforme. Significa poter decidere *come* e *da chi* si raccolgono, si processano e governano le informazioni: una condizione necessaria per tutelare diritti come privacy, libertà di espressione e sicurezza, oggi ridefiniti nelle interfacce delle grandi piattaforme. Quindi fuori dalla possibilità della maggior parte degli Stati".

Ma il concetto di sovranità digitale può avere il suo focus anche sui singoli piuttosto che sugli Stati. Dice nella nostra conversazione il giornalista che, con Edoardo Fleischner, ha progettato la televisione italiana di flusso Rainews24, professore dell'Università Federico II, Michele Mezza: "è un concetto non meno fluido di quanto non stiano diventando tutte le tecnicalità digitali, che ormai stanno evolvendo ad un ritmo sempre più frenetico. In questa trasformazione si intravvede come tendenza principale la spinta al decentramento degli accessi e delle personalizzazioni dei dispositivi, in particolare nel campo delle nuove intelligenze artificiali. Pertanto, direi che per uno Stato la sovranità digitale è la capacità di permettere ai propri cittadini di essere autonomi e indipendenti proprio nelle fasi di selezione e personalizzazione dei meccanismi digitali, riducendo la subalternità rispetto alla proprietà degli stessi meccanismi alla pura fase di impostazione del sistema".

É la conferma di come il concetto di sovranità digitale, variamente teorizzato, come già sostengono Couture & Toupin (2019), venga usato, con significati diversi, da una molteplicità di attori, capi di stato, studiosi, movimenti anarchici e di base, per indicare differenti forme di protezionismo statale, tutela dei dati o di governance multistakeholder della stessa Internet.

In forza di questa sua articolazione, e a causa delle sue molteplici interpretazioni, anche di carattere metaforico, perciò, esistono modi diversi di approcciare il concetto di sovranità digitale applicato a Internet e ai servizi consentiti dalle sue funzioni logiche e fisiche

(protocolli, routing, DNS, server, Internet exchange points, servizi Web, Ftp, e-mail e applicazioni software).

Eppure, nonostante questa complessità, come sostengono Vittorio Bertola e Stefano Quintarelli (2022), due pionieri italiani della rete Internet, il concetto di sovranità digitale può essere utilmente riassunto come "il diritto e il dovere di uno Stato indipendente di imporre efficacemente regole specifiche all'uso di Internet da parte dei propri cittadini e da parte degli attori esteri che forniscono a essi i servizi digitali fondamentali, in modo che tale uso sia conforme alle leggi, ai valori e agli interessi di quello Stato". Queste regole afferiscono a tutte le prerogative fondamentali di uno Stato che voglia dirsi indipendente e possono essere sommariamente ricondotte a cinque categorie normative: la gestione delle identità e della tassazione, la gestione dei contenuti, la sicurezza interna e la sicurezza nazionale.

Poiché ciascuno Stato ha una propria visione di come si gestiscano queste dimensioni, ogni idea statuale di sovranità digitale si scontra con le politiche degli altri Stati e comporta il rischio concreto della frammentazione di Internet, che può essere causata dall'introduzione di barriere di accesso alla Rete stessa, ai suoi servizi, ai contenuti, e ai dati che attraverso di essa fluiscono. Un rischio che si concretizza nell'introduzione di standard tecnici imposti dai governi o dalle aziende, da leggi censorie che ne limitano il libero accesso e utilizzo, oppure attraverso il controllo dei dati generati dai suoi utilizzatori. Come sostiene Vinton Cerf, considerato uno dei "padri di Internet" in quanto creatore con Robert Kahn del protocollo TCP/IP (Transmission Control Protocol/Internet Protocol), "una frammentazione dell'infrastruttura Internet di base a favore della cosiddetta sovranità dei dati minaccia il libero flusso di informazioni a livello globale. Il valore di questo libero flusso è vitale per la condivisione e la ricerca di informazioni a beneficio di tutti" (Cerf, 2022).

4.3 Frammentazione vs. sovranità: i rischi

Tra i rischi principali di questa potenziale frammentazione, uno rilevante è quello della fine della Net Neutrality. La neutralità della Rete, l'uguale accesso a tutti per qualsiasi tipo di contenuto, ha consentito a chiunque di offrire servizi su Internet rispettando semplici protocolli e senza chiedere il permesso a nessuno. Oggi i grandi carrier di telecomunicazione, tuttavia, insistono a risolvere il problema della sicurezza e del sovraccarico della Rete costruendo le proprie infrastrutture di Rete e imponendo costi diversi per accedere a servizi diversi, trattando i byte in modo differente per aumentarne la profittabilità.

Un altro rischio è la censura. Stati autoritari impediscono ai loro cittadini di usare la rete come piattaforma per il commercio, la politica, e per le relazioni sociali. Seguono i rischi della violazione della privacy e la limitazione della libertà d'espressione (Kaye, 2021). Alcune aziende ritengono lecito violare la privacy degli utenti per tutelare i loro diritti di proprietà intellettuale, mentre si moltiplicano le cause legali dovute all'uso non autorizzato dei contenuti dell'industria creativa per impedire il remix di opere intellettuali oppure dei dati usati per l'addestramento di sistemi informatici basati sull'Intelligenza Artificiale. L 'ultima, per rilevanza, è quella intentata dal New York Times contro la società OpenAI (Grynbaum & Mac, 2023). Omofilia, omofobia, transfobia, sessismo e razzismo rappresentano invece il rischio che gli utenti internet finiscano inglobati in reti autoreferenziali di persone che la pensano allo stesso modo e finiscano per discriminare ed escludere "gli altri diversi". L'effetto complessivo è quello di erodere il capitale sociale che si costruisce in rete, amplificando il digital divide di genere, politico, sociale, e culturale, che ancora impedisce alle donne e agli uomini di alcuni paesi di usare la rete come strumento di empowerment e partecipazione.

Un rischio ancora maggiore è tuttavia la mancata governance della rete. In assenza di un framework comune di regole di funzionamento per mantenere la rete efficiente, e per farla evolvere, il rischio è la creazione di reti regionali e nazionali autarchiche con regole proprie. La mancata interconnessione fra le reti, minacciata spesso come ritorsione politica, è in grado di provocare una frammentazione esiziale per la Rete, come nel caso della Legge sull'Internet sovrano, un insieme di modifiche legislative del 2019 che danno al governo russo il potere di controllare e isolare la Russia da Internet, creando un sistema di gestione centralizzato e un DNS nazionale (ISPI, 2020).

Studiosi e attivisti hanno denunciato a più riprese la fine dell'Internet aperta. Questo è accaduto soprattutto dopo le risposte degli Stati Uniti alla lunga devoluzione del controllo di Internet, prima esercitato interamente dall'Icann[7], e poi per le decisioni di alcuni Stati, anche recenti, di limitare i servizi accessibili attraverso di essa, a cominciare dagli Internet shutdown[8]. È il caso della decisione del presidente americano Donald Trump che, durante la

[7] Icann, Internet Corporation for Assigned Names and Numbers, è l'organizzazione internazionale che si occupa di gestire alcuni aspetti tecnici del funzionamento di Internet. È compito dell'ICANN gestire l'allocazione degli spazi per gli indirizzi IP; gestire il sistema dei domini di primo livello (TLD, *Top level domain*), sia generici (gTDL) sia regionali e nazionali (ccTDL) e di gestire il sistema dei root nameserver. Questi servizi erano originariamente garantiti dal Governo degli Stati Uniti d'America tramite il Ministero per il Commercio; successivamente, grazie a un lungo processo di *devolution*, l'ICANN è stata trasformata in un'organizzazione internazionale non governativa gestita da un consiglio composto da 16 membri.

[8] Blocco dei servizi Internet deciso dalle Autorità attraverso lo spegnimento o la chiusura dei servizi di telecomunicazione per impedire l'uso di Internet a specifici gruppi di individui o intere popolazioni all'interno di uno specifico territorio per controllarne il flusso di informazioni.

sua prima presidenza (2017-2021), pur non rendendolo effettivo, ha imposto il divieto di utilizzo a popolari app cinesi come TikTok e WeChat, sull'esempio dell'India, adducendo preoccupazioni per la sicurezza nazionale. In proposito Dipayan Ghosh, già consigliere tecnico del Presidente Barack Obama alla Casa Bianca e direttore del Digital Platforms & Democracy Project della Harvard Kennedy School, ha affermato: "Questa è Splinternet".

Splinternet è un termine coniato da Clyde Wayne Crews, ricercatore del Cato Institute, che lo usò nel 2001 per descrivere "le Internet parallele che funzionano come universi distinti, autonomi e privati". Ma gli esempi sono numerosi. L'ultimo, eclatante, è stato il divieto reciproco di usare piattaforme americane come Facebook all'interno della Federazione Russa dopo l'invasione dell'Ucraina del 2022, e quella di usare servizi digitali russi come VKontakte sia in Europa che negli Stati Uniti come ritorsione all'aggressione di uno stato sovrano.

Seppure la definizione sia poco usata nella letteratura accademica, si parla di Splinternet anche quando le dinamiche tipiche dei social network tendono a creare le cosiddette echo-chamber (Cappella & Hall Jamesion), le casse di risonanza, i cui muri si chiudono intorno a chi cerca solo conferme alle proprie tesi invece di discuterle con chi la pensa diversamente. Un'idea già espressa con il nome di cyber-balcanizzazione[9] dal teorico Cass Sunstein (2003) negli anni 2000.

Paradossalmente, è stata la pandemia globale da coronavirus, COVID-19 (*COronaVIrus Disease 19)*, scoppiata nel 2019, a consentire una maggiore frammentazione di Internet e delle sue regole, riducendone la libertà complessiva. Proprio quando tutte le attività umane (commercio, istruzione, sanità, socializzazione) si spostavano online, i governi hanno colto l'occasione della "grande paura" per controllare la Rete e plasmarne la narrazione, censurare le voci critiche e costruire attraverso di essa nuovi dispositivi tecnologici di guerra e controllo sociale. Nazioni come il Myanmar, la Turchia e il Brasile hanno usato infatti la prima pandemia globale dell'era dell'informazione come pretesto per limitare l'accesso dei propri cittadini alle notizie, chiudendo siti e servizi informativi, e hanno usato la situazione di emergenza per giustificare l'estensione dei poteri di sorveglianza, usando la digitalizzazione a tappe forzate dei servizi sanitari per moltiplicare la raccolta e l'analisi dei dati delle persone senza adeguate protezioni contro gli abusi. In alcuni casi, la sorveglianza biometrica, i big data e l'Intelligenza Artificiale usati per contrastare la pandemia, sono stati "militarizzati" per ottenere informazioni sensibili sullo stato di salute, sui modelli di acquisto e sui

[9] Considerati i potenziali effetti di questa frammentazione si tende pi spesso a parlare di *balcanizzazione*, per sottolinearne gli aspetti negativi, in analogia col processo di divisione dell'area dei Balcani successivo alla caduta del muro di Berlino e alla presunta fine della Guerra fredda fra i due blocchi contrapposti della Nato e del Patto di Varsavia.

comportamenti sociali dei cittadini usando il riconoscimento facciale, vocale e del volto, incrociati con le informazioni sul codice genetico e filtrando la comunicazione globale attraverso i confini nazionali (Di Corinto, 2021).

Ma oggi quelli che aspirano a una "Rete Halal", cioè una rete "pura" e controllata dallo Stato**,** strumento di proiezione geopolitica verso l'estero e strumento di sorveglianza verso l'interno, non nascondono più le pretese censorie e di controllo e parlano esplicitamente di diritto alla sovranità digitale.

Così, la sovranità digitale che si manifesta oggi come la pretesa dei governi di regolare il funzionamento di Internet secondo i propri modelli culturali e valoriali (Lemley, 2021), sta diventando la scusa per reprimere i diritti umani e civili nel mondo: Cina e Russia usano piattaforme centralizzate di scambio dati, tecniche di Deep packet Inspection (cioè il filtraggio dei pacchetti di bit) e gli Human flash engine, cioè sorveglianti umani, per controllare i loro cittadini online, fino all'ipotesi di adozione, in Cina, dello *scoring* sociale, una sorta di criterio di affidabilità, già applicato alla finanza, per decidere, in base ai comportamenti online, l'accesso ai servizi di cittadinanza come trasporto e istruzione. E lo stesso vale per l'Arabia Saudita, l'Iran, Hong Kong, dove, in seguito alla riunificazione, la Cina ha usato il pugno duro contro i "reati di linguaggio" per reprimere il movimento democratico che ne chiedeva l'autonomia. E ogni volta che le società tecnologiche statunitensi, tra i maggiori player del settore, hanno annunciato l'intenzione di sospendere l'erogazione dei loro servizi in quei Paesi come ritorsione per leggi sempre più draconiane, hanno ricevuto una sola risposta: accomodatevi. Oppure sono state ricattate, come successo in Turchia, dove i maggiori social network sono stati costretti ad avere rappresentanti locali in uffici locali, soggetti alle leggi locali. La Russia ha fatto lo stesso ed è arrivata a pretendere di avere sul proprio territorio i server di Telegram fino alla loro migrazione verso gli Emirati Arabi Uniti, dove le autorità però mettono a tacere attivisti, giornalisti e difensori dei diritti umani.

In un contesto in cui ogni attività umana è digitalizzata, e quindi ogni comportamento può essere "datificato", cioè trasformato in dati, appare evidente perché la pretesa sovranità digitale da parte degli Stati venga esercitata principalmente attraverso la governance dei dati. Imponendo la localizzazione dei dati digitali i governi sono infatti in grado di monitorare e sorvegliare i cittadini e decidere le forme di sviluppo della comunicazione in Rete per esercitare infine la loro idea di supremazia, potere e autorità.

E anche quando lo sviluppo e la commercializzazione delle tecnologie digitali necessarie a farlo sono fuori della portata delle nazioni meno industrializzate, la raccolta, l'analisi, la categorizzazione dei dati personali e commerciali viene gestita dalle grandi corporation che

consentono la profilazione degli individui e l'anticipazione dei loro comportamenti, modellati attraverso input specifici (Di Corinto, 2022). Perciò anche se la battaglia sulla sovranità di Internet è diventata nel tempo una battaglia sulla gestione dei dati e sul controllo dei flussi di questi dati che sono generati in Rete, le motivazioni restano diverse: la Cina la pretende per motivi di "stabilità sociale", gli Usa per la "sicurezza nazionale" e l'Unione europea per "difendere la privacy e la libera concorrenza". Ma in definitiva l'obbiettivo è quello di controllare i propri cittadini, anche per procura, attraverso le multinazionali tecnologiche.

Come sostiene efficacemente Yuval Harari (2018): "Prima la terra era la risorsa più importante del mondo, la politica era una lotta per controllare la terra, e se troppa terra si fosse concentrata in poche mani, la società si sarebbe divisa in aristocratici e plebei. Nell'era moderna le macchine e le fabbriche hanno assunto un'importanza maggiore della terra e le lotte politiche hanno mirato al controllo di questi mezzi di produzione. Se un numero troppo alto di macchine si fosse concentrato in poche mani, la società si sarebbe divisa in capitalisti e proletari. Nel XXI secolo, invece, i dati eclisseranno sia la terra sia le macchine come risorsa strategica. Se i dati si concentrano nelle mani di pochi, l'umanità si dividerà in specie differenti. La gara per ottenere i dati è già iniziata e vede in testa giganti high-tech come Google, Facebook, Baidu e Tencent. Finora queste aziende sembrano avere adottato il modello di business dei 'mercanti dell'attenzione'. Catturano la nostra attenzione fornendoci informazioni gratuite, servizi e intrattenimento, e rivendono poi la nostra attenzione alle aziende inserzioniste. È però probabile che i giganti dei dati coltivino obiettivi assai più ambiziosi di ogni precedente mercante dell'attenzione. Il loro vero business non è affatto vendere spazi pubblicitari. In realtà, catturando la nostra attenzione, sono in grado di accumulare una immensa quantità di dati su di noi, un fatto che vale molto più di qualunque incasso pubblicitario. Non siamo i loro clienti - siamo i loro prodotti".

Il tema della sovranità dei dati e del controllo dei comportamenti sociali e individuali è stato affrontato in ogni edizione dell'Internet Governance Forum globale (IGF) negli ultimi venti anni, dalla sua nascita nel 2005 a Tunisi (Abba, Lazzaroni & Pietrangelo, 2022). All'epoca la nascita stessa dell'IGF fu la risposta diplomatica delle Nazioni Unite, attraverso l'International Telecommunication Union (ITU), alle tensioni geopolitiche tra l'Occidente, cioè gli Usa e l'Europa, e i Brics (Brasile, Russia, India, Cina, Sud Africa), i quali ultimi chiedevano il controllo distribuito delle infrastrutture di governo della rete Internet mentre i paesi occidentali gli chiedevano il rispetto dei diritti umani. Ancora oggi, nonostante i tentativi di accordo che si ripetono ogni anno nell'ambito di questa sorta di "parlamento di Internet", sede di discussione senza poteri decisionali, la tentazione delle grandi potenze è

quella di trattare il cyberspazio come le acque internazionali, con alcuni Stati che cercano di regolamentare a proprio vantaggio l'uso dei beni comuni globali delimitandone l'accesso e la fruizione (Santaniello & Paladino, 2022).

Tuttavia, poiché il cyberspazio è tecnicamente interconnesso, essendo basato primariamente su una rete di reti, Internet, pretendere la sovranità, cioè il controllo e la protezione della sua infrastruttura principale per farne un elemento abilitante di sviluppo e attuazione di politiche di controllo dei dati, e quindi dei comportamenti, è estremamente difficile.

Il controllo, indiretto, sull'infrastruttura critica di Internet è ancora sotto l'autorità degli americani: se c'è qualcuno che davvero gestisce Internet, sono ancora l'Icann e le grandi piattaforme digitali localizzate negli USA. Apparentemente un vantaggio, considerato il valore che gli USA hanno finora attribuito alla libertà e alla democrazia. Epperò la Rete, pur non essendo nata come arma di guerra (Di Corinto, 2017), è dagli inizi della sua storia uno degli strumenti della dominanza informativa americana, intesa come "un livello di superiorità informativa che consente di utilizzare i sistemi e le capacità informative per ottenere un vantaggio operativo in un conflitto o per controllare la situazione in operazioni non belliche, negando questa capacità all'avversario" (Colon, 2025). Motivo lampante per cui il presidente statunitense Donald Trump ha minacciato sanzioni e dazi commerciali elevati a quei paesi che, come l'Unione Europea, vogliono regolamentare e tassare i giganti digitali americani che ne sono diventati i gatekeeper (Di Donfrancesco, 2025).

Oggi le preoccupazioni relative alla sicurezza informatica del mondo digitale in cui siamo immersi, compresa la vulnerabilità delle infrastrutture critiche al terrorismo e a soggetti ostili, hanno fatto sì che la libertà di espressione che la Rete ha promosso e rappresentato, sia minacciata da comportamenti isolazionisti, leggi autoritarie e dalla censura degli algoritmi, sia direttamente da parte degli Stati, sia indirettamente da parte degli oligopoli proprietari di social network, motori di ricerca, servizi digitali e di data mining (Rampini, 2005).

In questo contesto, la frammentazione, il blocco, il malfunzionamento della Rete, e quindi la perdita di sovranità digitale, possono anche essere anche il risultato dello sfruttamento di protocolli datati e virus. Un virus capace di prendere di mira i router di Internet, e in particolare il protocollo BGP (Border Gateway Protocol), usato per connettere tra loro router di sistemi autonomi distinti, ovvero reti controllate da una singola autorità amministrativa,

azienda o provider, può avere un effetto dirompente. Motivo per cui, proprio nell'agosto del 2025, la Cybersecurity and Infrastructure Security Agency americana, CISA, in coordinamento con la National Security Agency, NSA, e il Federal Bureau of Investigation, FBI, hanno diramato un preoccupato allarme: attori sponsorizzati dallo Stato cinese, APT[10], stanno sfruttando le vulnerabilità dei router utilizzati dai fornitori di servizi di telecomunicazioni e da altri operatori infrastrutturali a livello globale (Cisa, 2025). Questi attori spesso adottano misure specifiche per eludere il rilevamento della loro presenza e delle loro azioni per mantenere al loro interno un accesso persistente, in particolare attraverso le reti di telecomunicazioni e dei trasporti. Lo stesso può accadere con un attacco ai Root Server e ai DNS di Internet, nonostante la loro struttura ridondante. Il Domain Name System, il sistema che traduce gli indirizzi IP, che traduce cioè in numeri comprensibili ai computer gli indirizzi parlanti dei siti web che digitiamo sul browser, è gestito dai server DNS, e uno solo di essi, se compromesso, potrebbe essere usato per sottoporre i cybernauti a insidiosi attacchi di phishing e interferire in altre attività. È uno dei motivi per cui alcune grandi aziende come Google hanno deciso di dotarsi di sistemi di DNS alternativi e lo stesso stanno facendo alcuni Stati come la Federazione Russa.

Poi ci sono gli attacchi DDoS (Distributed Denial of Service), i cosiddetti attacchi da negazione di servizio, coi quali si sommerge di richieste un server web utilizzando reti di computer precedentemente infettati da malware. Organizzati in *botnet* controllate da attori malevoli all'insaputa dei proprietari, questi *pc zombie* vengono "richiamati in vita" su commissione per inviare spam o effettuare attacchi di varia natura e rendere irraggiungibili i servizi esposti sul Web. Pratica antagonista usata da rudi attivisti per i diritti civili, divenuta una tecnica di cyberwarfare[11] impiegata già nel 2007 in Estonia, ma anche ai danni di Twitter nell'agosto 2008, oggi è parte dell'arsenale di tutti gli eserciti cibernetici. Infine, c'è il sabotaggio dei cavidotti. Un danneggiamento dei cavi strategici della rete che collegano diversi continenti, può avere effetti amplificati, rendendo Internet inaccessibile a milioni di utenti. É successo a più riprese (Asaf, 2025), ma oggi ci sono indicazioni che ne rivelano il carattere strategico all'interno dei conflitti armati come accaduto nel novembre del 2024 nel Mar Baltico (Piccolo, 2024). Il motivo è chiaro: la "spina dorsale della connettività globale», come la definisce Francesco D'Arrigo, direttore dell'Istituto italiano di studi strategici Niccolò Machiavelli, è garantita da 500 cavi sottomarini, «non più spessi di un tubo da giardino», da lì transita il 98% delle comunicazioni mondiali" (Bartolomei, 2025).

[10] APT, Advanced Persistent Threat, minaccia consistente in un attacco mirato, volto ad installare una serie di malware all'interno delle reti bersaglio, al fine di riuscire a mantenere attivi i canali impiegati per l'esfiltrazione di informazioni dalle infrastrutture IT del target. É una tecnica peculiare degli hacker di stato finanziati dai governi.

[11] Cyberwarfare, insieme di tecniche, strumenti, tecnologie e pratiche necessarie a condurre la guerra cibernetica.

Il rischio di un'ulteriore frammentazione della Rete causata dalla pretesa di esercitarvi una rinnovata sovranità, è che la guerra oggi si fa con Internet e attraverso Internet. Ecco così che la pretesa di sovranità digitale può diventare la porta d'ingresso della definitiva balcanizzazione di Internet, causa ed effetto della trasformazione della Rete in arma di guerra e strumento di geopolitica. Il coinvolgimento di attori non statali nei conflitti cibernetici, la censura dei media e la crescente importanza militare delle telecomunicazioni stanno portando a un'accelerazione della balcanizzazione di Internet e alla fine dell'utopia di un mondo pacifico perché iperconnesso e interdipendente grazie alla Rete (Morozov, 2011). La Rete si trasforma quindi nell'oggetto di una vera e propria guerra dell'informazione e cioè in a) strumento di propaganda e disinformazione per disseminare paura, incertezza e dubbio, confondendo i confini tra verità e finzione; b) strumento di influenza e sorveglianza per gli agenti di opposte fazioni; c) strumento di comando e controllo militare per operazioni cinetiche e cibernetiche.

4.4 Sovranità digitale e autonomia strategica

Tuttavia, come sostengono Bellanova et al. (2022), i dati e le tecnologie digitali possono rappresentare anche uno strumento di governance pubblica e la chiave del processo di integrazione europea. La sovranità digitale, declinata come *sovranità tecnologica* o *autonomia strategica*, è presentata dagli studiosi quale leva per lo sviluppo di una società europea sicura e resiliente, orientata a ottenere una posizione di leadership nel contesto internazionale e a ridurre la sua dipendenza da altre parti del mondo. E, in effetti, questa idea di autonomia strategica è diventata nel tempo parte di un complessivo tentativo di riorientamento del rapporto con le Big Tech che implica la capacità degli Stati europei di preservare l'abilità di incidere, e di decidere, in un mondo interconnesso.

Come ha osservato Luciano Floridi (2020), purtroppo però, ancora oggi viviamo una sovranità digitale di tipo aziendale (*digital corporate sovereignity*) costruita sul monopolio di fatto delle multinazionali tecnologiche, le Big Tech, che progettano, producono, mantengono e vendono tutto ciò che è digitale, "poietic power", lo chiama, al punto che diventano la prima linea di difesa degli Stati stessi quando si parla di cyberattacchi.

Questa dinamica di potenza che contribuisce a modellare le nostre società, sfida la tutela dei diritti fondamentali e la stessa sovranità quando, dice Floridi, gli Stati mancano di esercitare il loro potere di controllo per regolare il mondo digitale, "decidendo cosa sia legale e cosa no, incentivi e disincentivi, tipi e livelli di tassazione, politiche di procurement, modalità e costi di osservanza delle norme (*compliance*)" (Floridi, ivi). Questo "potere cibernetico", come lo definisce il professore di Yale, può tuttavia diventare effettivo, a livello europeo, se

assume una dimensione continentale, ricordandoci che la sovranità è proprio il tipo di potere che legittima le altre forme in cui esso si articola ed esprime.

La necessità di un approccio europeo alla sovranità digitale è presente ai legislatori europei. Rilanciata dalla Commissione presieduta da Ursula von der Leyen nel quadro della strategia europea di *digital sovereignity*, con l'intento di produrre nuovi strumenti sanzionatori espressamente mirati alle *very large platforms*, cioè alle Big Tech, questa necessità è stata ribadita nel discorso sullo Stato dell'unione del 16 dicembre 2020. Da quel momento è divenuta centrale nelle politiche europee (Santaniello, 2021).

L'Europa ha provato a dare corpo a questo tentativo prima con l'emanazione del GDPR (Global Data Protection Regulation, 2016), il Regolamento europeo sulla privacy, poi con la NIS (Network and Infrastructure Security, 2016), entrambi entrati in funzione su tutto il territorio europeo nel 2018; ha continuato nella stessa direzione con la seconda Direttiva per la sicurezza di reti e infrastrutture, NIS 2, e con la direttiva Ue sulla Resilienza delle Entità Critiche (CER, Critical Entities Resilience, 2022), che amplia la protezione delle infrastrutture critiche oltre a quelle di energia e trasporti, coprendo nuovi settori e promuovendo la resilienza e l'adattamento alle minacce emergenti. A questi interventi si sono succeduti il Digital Operational Resilience Act (DORA, 2022), che stabilisce requisiti standardizzati per la gestione del rischio, della resilienza e della sicurezza informatica all'interno del settore finanziario europeo, e per i fornitori di servizi ICT critici; e poi lo ha fatto con l'Artificial Intelligence ACT. I risultati cominciano a vedersi, in particolare con l'entrata in vigore della NIS 2, nella cui adozione l'Italia si è posizionata, cronologicamente, al primo posto tra i paesi europei il 7 ottobre 2024.

Secondo il Garante italiano per la protezione dei dati personali, professore Pasquale Stanzione (2025): "Con queste norme, l'Europa ha tracciato il *Nomos der Erde (il Diritto della terra, nda)* della geopolitica dei dati, comprendendone fino in fondo il valore e fondando sulla regolazione la propria sovranità digitale". Proprio per contrastare lo strapotere digitale delle piattaforme accreditate avere il 70% dei dati di istituzioni e aziende europee.

Il GDPR, il regolamento sulla privacy operativo dal 2018 per normare circolazione, trattamento e tutela dei dati dei cittadini europei, è stato il riferimento delle normative successive. E da allora si sono aggiunte altre leggi e nuovi regolamenti, questi ultimi subito applicabili in tutti gli Stati, come quelli che rendono esecutivo il Cyber Security Act (2024), e il Cyber Resilience ACT del 2024, applicabile dal 2027, ognuno per disciplinare un tassello dell'ecosistema digitale continentale europeo. Ma anche per stabilire un confine etico rispetto ai diritti e ai doveri della cittadinanza digitale e del corretto comportamento delle imprese europee. L'Artificial Intelligence Act, che si occupa dell'utilizzo di sistemi di

intelligenza artificiale è già in parte operativo, e lo diventerà pienamente entro la fine del 2027. Con questa ultima legge l'Europa ha definito un quadro regolatorio articolato che classifica i sistemi di intelligenza artificiale in base al loro livello di rischio, imponendo obblighi specifici da parte di imprese e pubbliche amministrazioni, a seconda della criticità dell'applicazione. L'obiettivo è di prevenire possibili distorsioni e garantire trasparenza, protezione dei dati e rispetto dei diritti fondamentali. Tranne che in due settori, però, quello della Difesa e quello della Sicurezza. Una visione diversa da quella degli USA i cui attori invece chiedono una regolamentazione minima, sostegno istituzionale e un approccio volontario per la governance dell'AI. L'Europa, insomma, non è rimasta a guardare. E non è un caso che il 27 agosto 2025 Trump abbia minacciato e poi a dicembre applicato sanzioni personali verso quei funzionari europei colpevoli di applicare il Digital Service Act, il Regolamento entrato in vigore nel 2023 e che introduce obblighi specifici per le piattaforme digitali con oltre 45 milioni di utenti nell'UE, e che ha l'obbiettivo di rendere l'ambiente digitale più sicuro colpendo contenuti come hate speech, disinformazione e pedopornografia nel rispetto della libertà di espressione (Ramkunar & Bade, 2025).

Il Digital Market Act (DMA) e il Digital Service Act (DSA) si occupano di regolare le grandi piattaforme (i *gatekeeper*) e il commercio online, cercando di assicurare piena competitività agli operatori, ma anche di proteggere gli utenti e moderare i contenuti prodotti. In questa direzione muove anche il Data Act, il regolamento Ue operativo dal 12 settembre 2025, che dovrà regolare l'accesso ai dati generati dai prodotti connessi (smartwatch, automobili o assistenti vocali intelligenti, etc.) (UVA, 2025) negli anni a venire.

Tuttavia, questo complesso e difficile tentativo di regolazione normativa delle trasformazioni in corso, presenta un problema. In assenza di una vera autonomia tecnologica l'Europa rimane vulnerabile: è il paradosso del regolatore senza tecnologie. Ma il Vecchio Continente ha una carta da giocare: una platea di 450 milioni di consumatori alto spendenti e un insieme di imprese che producono il 22% del Pil Mondiale. Per conservare la sua sovranità ed evitare di essere colonizzata da tecnologie di altri paesi, l'Europa dovrà fare di più e investire nella sovranità tecnologica come pure aveva ripetuto, già il 15 settembre 2021, la presidente della commissione Europea, Ursula von der Leyen.

Il tentativo quindi di affermare la governance europea sugli ampi processi di trasformazione in corso richiede pertanto di comprendere se e come la sovranità digitale, concetto più ampio della sovranità di Internet, possa essere declinata come autonomia strategica nazionale senza avere campioni europei tra le maggiori piattaforme di servizi digitali, e senza implicare la separazione delle reti geografiche che compongono Internet, cooperando per mantenerla aperta ed efficiente in un contesto di collaborazione politico istituzionale a

livello mondiale. Per farlo bisogna però crearne le condizioni effettive in quanto secondo il professore Roberto Baldoni: *"In un cyberspazio globale e aperto, la piena sovranità digitale implica l'autorità complessiva di una nazione sui dati generati dai suoi cittadini, dall'amministrazione pubblica e dalle imprese. Ciò include la capacità di una nazione di impiegare tecnologie sicure per elaborare questi dati, supportate da una forza lavoro sufficiente, competente e fidata. Inoltre, comporta l'istituzione e il mantenimento attivo di collaborazioni internazionali dinamiche e mirate, per affrontare proattivamente le minacce. Richiede infine, una società pienamente consapevole e educata sui rischi presenti nel cyberspazio"* (Baldoni, 2025).

L'Italia, ad esempio, si è impegnata in questa direzione finendo per rappresentare un modello per gli altri paesi europei. Lo ha fatto intanto con l'elaborazione della Strategia cloud nazionale che presuppone un controllo differenziato dell'accesso ai dati prodotti in Italia, immagazzinati e trattati nel cloud gestito dal Polo strategico nazionale (PSN) da parte di aziende italiane in collaborazione con quelle americane che gli forniscono l'infrastruttura di base. Secondo il piano, il 75% dei dati conferiti dalle PA italiane al cloud del PSN è accessibile per il trattamento anche da parte di soggetti non nazionali, il 25% invece, non può esserlo in quanto relativo a dati strategici, cioè dati connessi alla sicurezza nazionale. L'intuizione dell'allora governo Draghi che lo decise su input del direttore generale dell'ACN, fu proprio la classificazione dei dati in ordinari, critici e strategici. Una distinzione proposta dallo stesso professor Roberto Baldoni.

Un altro vettore di questa strategia avviata dall'Agenzia per la cybersicurezza nazionale italiana in parallelo alle iniziative dello European Cybersecurity Competence Center (ECCC), è un forte investimento nel settore delle start-up tecnologiche che si occupano di robotica, IA, crittografia e cybersecurity, come pure la creazione di partenariati pubblico-privati nello sviluppo di supercomputer e algoritmi di Intelligenza Artificiale. La scala di grandezza italiana non consente però quell'autonomia strategica che, secondo un sentire diffuso, va pensata a livello europeo, motivo per cui i primi mattoni di questa costruzione hanno riguardato la creazione dell'ECCC, il potenziamento dell'Enisa, l'Agenzia europea per la sicurezza cibernetica, la creazione del network europeo per la risposta alle crisi informatiche, CyCLONe, e lo sviluppo di piani europei di innovazione industriale (Digital Europe e Horizon). Con un tassello finale: il G7 cybersecurity working group che ha lo scopo di proteggere le infrastrutture critiche, combattere il ransomware e promuovere lo sviluppo sicuro dell'IA. Secondo l'Agenzia per la cybersicurezza nazionale italiana, è questa la direzione per proteggere la sovranità digitale europea.

4.5 Le minacce alla sovranità digitale

Tra le molteplici forze che modellano le nostre società, la trasformazione digitale è una delle più potenti, con un impatto profondo su ogni settore produttivo. L'IA e le sue applicazioni, la blockchain, il 5G e la crittografia, i supercomputer, ma anche la robotica e la data science stanno cambiando il modo stesso di operare di imprese e istituzioni. Tonnellate di nuovo hardware e milioni di programmi software si aggiungono alle tecnologie esistenti e il loro impiego ci mette di fronte a rischi e minacce senza precedenti, in un contesto tecnologico di sempre maggiore complessità dove i rischi indotti da cyberattacchi e tecnologie obsolete, oltre che dallo sfruttamento di vulnerabilità digitali, secondo Roberto Baldoni "fanno parte dello stesso scenario evolutivo della sovranità digitale in un mondo multipolare" (Baldoni, 2022).

Baldoni, già docente di Sistemi distribuiti alla facoltà di Ingegneria della Sapienza Università di Roma, ideatore dell'Agenzia per la cybersicurezza nazionale italiana e suo primo direttore dal 2021 al 2023, descrive in maniera sintetica i quattro ambiti che mettono a rischio la sovranità digitale intesa come autogoverno di dati, tecnologie, infrastrutture, persone, e cioè: a) gli attacchi informatici; b) le minacce alla supply chain delle forniture critiche; c) la diffusione delle tecnologie emergenti come Intelligenza Artificiale e Quantum Computing; d) le minacce sociali, industriali, tecnologiche e ibride, compresa la disinformazione. Nel suo ultimo libro, Baldoni (2025), precisa i quattro ambiti facendo ricorso anche ad esempi di cronaca ricchi di dettagli circa il modo di operare di threat insider, hacker e APT, illustrando gli attacchi DDoS e ransomware[12], i rischi della manomissione della supply chain[13] con riferimento ai casi SolarWinds, Kaseya, Stuxnet, l'emergenza dei chatbot e degli algoritmi predittivi, fino alla disinformazione costitutiva dei social network, citando i famosi casi della Brexit, del Pizzagate e del passaggio di mano di Twitter, oggi X.

Attento a chiarire che quello di sovranità digitale è un concetto mobile che gli stessi studiosi non hanno ancora definito in maniera univoca e che gli Stati nazione e gli altri attori interessati interpretano in maniera diversa, il professore insiste su una definizione operativa della sovranità digitale, cioè la capacità di una nazione di proteggere il proprio cyberspace come se proteggesse un territorio fisico, e il cui fallimento equivale a consegnare i suoi abitanti a un potere oscuro e incontrollabile, quello di un progresso tecnologico dove attori malevoli sfruttano "macchine" che sopravanzano gli umani e aggirano tutti i contrappesi della democrazia.

[12] Un tipo di malware che cifra dati e sistemi e richiede il pagamento di un riscatto per la decifrazione.

[13] Supply chain, la catena di fornitura, è la combinazione dell'ecosistema di risorse necessarie per progettare, produrre e distribuire un prodotto. In ambito di sicurezza informatica, una catena di fornitura include hardware e software, archiviazione cloud o locale e meccanismi di distribuzione.

A illustrare questo scenario c'è un evento del nostro tempo che esemplifica abbastanza bene, tutte insieme, le minacce alla sovranità digitale, ed è la guerra in corso ai confini dell'Europa. Dall'invasione russa del Donbass ucraino avviata nel febbraio del 2022, infatti, si è assistito a un progressivo coinvolgimento della società civile nel conflitto facendo uso di tecnologie digitali. Nella guerra che ne è scaturita, il governo ucraino, ad esempio, ha chiamato i cittadini a difendere le infrastrutture critiche del paese con mezzi informatici e a bersagliare con attacchi DDoS quelle russe, mentre i russi agivano nello stesso modo diffondendo malware distruttivi dentro e fuori il paese invaso. Nel conflitto che ne è seguito, per prevenire la propaganda e la disinformazione che sempre accompagnano gli scontri armati, l'Ucraina e i governi occidentali hanno bandito social network e media russi dai propri territori, ottenendo lo stesso in ritorsione dalla Federazione Russa (Gaggi, 2022). Ma, poco prima dell'ingresso delle truppe russe in Ucraina, i modem della rete Internet satellitare KA-SAT di Viasat sono stati disabilitati in massa e solo dopo sono stati scatenati attacchi informatici massivi contro il paese invaso.

In seguito, i servizi segreti del presidente Vladimir Putin hanno sfruttato i gruppi filorussi produttori di ransomware per attaccare la *supply chain* di aziende dei paesi Nato, avendo in aggiunta l'obbiettivo di interferire con la produzione di armi e l'erogazione di servizi essenziali come acqua e servizi sanitari sul territorio ucraino. Motivo per cui gli USA sono intervenuti nel conflitto coi loro *cybersoldiers* (Martin, 2022).

Come era già successo nel 2015 dopo l'annessione della penisola di Crimea, le truppe russe hanno occupato i principali Internet service provider della penisola, interrotto i collegamenti con l'Ucraina, costruito un nuovo cavo sottomarino in fibra ottica e reindirizzato il traffico Internet attraverso Miranda Media, sede a Mosca, per consentire a Roskomnadzor, l'autorità russa di regolamentazione delle telecomunicazioni, di controllarlo e obbligare i residenti della penisola a rispettare le ferree normative russe sulla comunicazione (Ottaviani, 2022, pp. 61-63).

Un modello replicato nei territori occupati di Kherson e altre città in Ucraina, inducendo i soldati ucraini in ritirata a distruggere le loro stesse infrastrutture di comunicazione. Mosca ha pure tentato di vietare l'uso di Tor[14] e reti private (le *Vpn, Virtual Private Network*), impiegate dai dissidenti o da comuni cittadini in cerca di informazioni per aggirare la censura operata dal governo, minacciando infine di staccarsi dall'Internet globale col progetto RuNet (Savelli, 2022).

[14] Tor, The onion routing, è un software per comunicazione sicure.

Nel frattempo l'imprenditore miliardario Elon Musk offriva al presidente ucraino Volodymir Zelensky il suo sistema satellitare Starlink per garantire agli ucraini la possibilità di comunicare via Internet e diverse aziende private mettevano a disposizione le immagini dei propri satelliti per individuare le responsabilità di eccidi come quello avvenuto nella cittadina ucraina di Bucha (BBC News, 2022), mentre ClearView, discussa azienda americana di sorveglianza, gli offriva sistemi di riconoscimento facciale per identificare i soldati russi, usati anche per minacciare le loro famiglie (Hill, 2022). Clearview AI, fondata da Hoan Ton-That nel 2017, una sorta di motore di ricerca che usa l'Intelligenza Artificiale per comparare i volti dei target con un database di 30 miliardi di immagini raccolte da siti e social network come Facebook, nei 50 giorni successivi all'invasione da parte delle forze armate di Putin, ha consentito all'esercito ucraino di usare il suo sistema di riconoscimento facciale 8.600 volte per identificare i cadaveri dei soldati russi o i militari catturati (Pisa, 2022).

La Russia intanto veniva attaccata da gruppi informatici avversari, con un esordio eclatante, l'hackeraggio compiuto dall'attore *NB65* nei riguardi dell'agenzia spaziale russa, la RosKosmos, che gli ha permesso di prendere il controllo dei satelliti spia russi accecando così il suo sistema d'arma nucleare. Chiara dimostrazione dell'importanza dei satelliti nella gestione della sicurezza e della difesa dello Stato, non solo per le comunicazioni civili da cui dipendono trasporti, il GPS, e le transazioni commerciali. Come scrive Iezzi (2023): "A questa prima azione ne sono seguite presto numerose altre, altrettanto clamorose: il 23 marzo 2022 è stato fatto trapelare il database della più grande azienda alimentare del mondo ancora attiva in Russia, la Nestlé, con oltre 10 GB di e-mail, password e clienti della corporation svizzera divulgati in rete. Due giorni dopo è stato il turno della Banca Centrale Russa, con la violazione di 35.000 file riguardanti operazioni riservate e la minaccia di divulgarli entro 48 ore, rivelando le relazioni compromettenti di molte realtà e di alcuni politici occidentali con il Cremlino. Lo stesso giorno è stato penetrato il sistema informatico dell'Istituto per la Ricerca Nucleare dell'Accademia delle Scienze russa. Il 29 marzo sono stati infine cancellati 65 TB di dati dai server dell'agenzia del trasporto aereo della Federazione Russa".

Gli attacchi, attribuiti a gruppi informatici filo-ucraini, provengono dai paesi più disparati: sono i georgiani BlackHawks e Monarch Turkish Hactivist, i polacchi Squad303 e i bielorussi Belarusian cyber partizans. Un gruppo di questi hacker attivisti è riuscito ad accedere anche al controllo del sistema ferroviario della Bielorussia alleata di Putin, fermando i treni a Minsk, Orsha e Osipovicichi, rallentando così il trasferimento in Ucraina delle truppe russe presenti nel paese.

Questo schema, attacchi informatici che accompagnano gli attacchi cinetici, l'uso di paramilitari cibernetici, la censura dei media, l'interruzione delle comunicazioni, l'interferenza nella filiera di forniture critiche, l'attivazione della macchina della propaganda, sarà ripetuto in tutti i successivi conflitti armati che coinvolgono pesantemente la dimensione cyber: in Palestina, dal 2023 in poi, e nella guerra tra l'Iran e Israele scoppiata nel giugno del 2025.

Il coinvolgimento di attori non statali nei conflitti, gli attacchi informatici, la censura e il controllo militare delle telecomunicazioni esemplificano i rischi della balcanizzazione di Internet e la fine dell'utopia di una Rete unica e libertaria vagheggiata negli anni '90 da John Perry Barlow (1996) nella Dichiarazione d'indipendenza del cyberspazio dove "i giganti di carne e d'acciaio non hanno nessuna sovranità". E così che la spinta alla sovranità digitale ha assunto la dimensione di una questione geopolitica globale laddove veniva inizialmente rivendicata dai singoli Stati per esigenze di privacy, di tutela della proprietà intellettuale, di sicurezza dei cittadini, promozione dei valori sociali nazionali, e Internet si trasformava in una trincea, la trincea della cyberguerra.

5. Internet, trincea della cyberguerra

Cyberwar is coming!
John Arquilla, David Ronfeldt, 1993

Ogni guerra è una guerra ibrida. Ciascuna delle parti in conflitto usa qualsiasi mezzo per prevalere sull'avversario, dai cannoni alla guerra cognitiva,[15] fino alla guerra informatica o cyberwar (Borgia, 2022). Ma che cos'è la cyberwar? Il significato appare chiaro a una prima traduzione: la cyberwar è una guerra (war) nello spazio cibernetico (cyberspace). Eppure, non tutti gli studiosi sono d'accordo con questa definizione considerata troppo generale. Per i nostri scopi possiamo considerare la seguente definizione sviluppata in ambito NATO: "La guerra informatica è definita come l'uso di attacchi informatici con l'intento di danneggiare le risorse di una nazione. La guerra informatica e la sua classificazione militare sono ancora oggetto di ampio dibattito. Tuttavia, numerosi stati membri della NATO e altri paesi hanno investito nello sviluppo di capacità informatiche, sia offensive sia difensive. Alcuni dubitano di poter definire tali azioni come guerra perché prendono di mira "solo" i computer.

[15] Guerra cognitiva, in ambito militare, è un concetto che si riferisce alla lotta per controllare il modo in cui le persone reagiscono all'informazione.

Tuttavia, la tendenza globale verso la digitalizzazione e l'Internet delle cose (IoT) ha fatto sì che più funzioni siano ora controllate dai computer di quanto molti possano immaginare. Tutto, dalle attrezzature edili alle istituzioni finanziarie, alle infrastrutture civili e persino alle installazioni militari, dipende ora da una complessa rete informatica. La perdita di tali risorse informatiche può, e ha già causato, danni ingenti, non solo in termini di tempo e perdita di dati, ma anche in danni fisici misurabili in dollari e vite umane" (Bernal et al. 2020).

In quanto parte della guerra economica, ma anche di quella politica e dell'informazione, la guerra cibernetica, la così detta cyberwar, chiarisce già dalla sua definizione che il dominio cyber non è un dominio a sé stante ma che è trasversale a tutti gli altri domini: terra, mare, cielo, spazio.

Pertanto, visto che la guerra usa l'informatica, si parla a questo proposito di conflitti "software defined" (Colajanni, 2025)[16], in un mondo digitalizzato, iperconnesso e dipendente dalla tecnologia, possiamo osservare come **in primo luogo** accada sempre più spesso che le operazioni della guerra cibernetica, cioè gli attacchi informatici, condotti attraverso Internet, precedano e seguano le attività cinetiche, cioè i conflitti armati che implicano il movimento.

In secondo luogo, possiamo osservare come le guerre informatiche nel cyberspazio oggi arruolino i civili, singoli e organizzati, e li rendano attori protagonisti dei conflitti, in una maniera assai diversa dal passato, siano essi parte di milizie digitali, snodi per la raccolta di informazioni, gestori di media outlet propagandistici o creatori di piattaforme abilitanti lo strumento militare: i civili sono diventati i terminali di una rete di produzione di senso e di intervento attivo nei conflitti (Mezza, 2022).

In terzo luogo, in ogni conflitto cyber si manifesta prepotentemente il ruolo centrale dell'informazione. È in tale contesto che emerge con evidenza il carattere della cyberwar come guerra ibrida, intesa come guerra non convenzionale, relativa al dominio delle informazioni e all'uso spregiudicato della tecnologia teorizzata dal generale russo Valery Gerasimov (2013), secondo cui prima di un conflitto armato è necessario attaccare lo spazio cognitivo del nemico, diffondendo messaggi che ne alterino la percezione delle informazioni, come accade nella propaganda computazionale, che legittima o rivendica gli attacchi. Ed è proprio in questa cornice che sia l'informazione giornalistica che quella delle reti sociali diventano l'obbiettivo di manipolazioni e infingimenti, lo strumento per eccellenza della

[16] (M. Colajanni, comunicazione personale, 10 agosto 2025)

moderna guerra cognitiva (Cristadoro, 2018). Una guerra totale in uno scenario senza confini, quello di Internet, abilitato da uno sviluppo tecnologico inarrestabile, che minaccia la stessa sovranità digitale e quindi quella statuale.

E la guerra ibrida, senza confini, che prende di mira gli apparati informatici sia in ambito civile sia militare, mette a rischio l'incolumità e la sopravvivenza stessa delle persone.

5.1 Gli attacchi cibernetici

Ogni giorno, numerosi tentativi di attacco vengono condotti contro i più diversi obiettivi in tutto il mondo. Ne rendono conto le cronache quotidiane e i numeri allarmanti prodotti dalle autorità di cybersicurezza di ciascun paese. Nell'ambito delle nazioni maggiormente industrializzate e, in qualità di membro del G7, l'Italia è uno dei paesi maggiormente colpiti dai cyberattacchi, nonostante la buona qualità della sua postura di cybersicurezza, attestata di recente dal Global Cybersecurity Index dell'International Telecommunication Union (ITU, 2024). La numerosità e il successo di questi attacchi dipendono da vari fattori, ma uno, quello principale, è relativo alla sua superficie digitale particolarmente ampia, prodotta, tra l'altro, dalla digitalizzazione accelerata di numerose attività, produttive e di svago, avvenuta negli ultimi anni senza troppo badare alla sicurezza. A questo quadro di rischio contribuiscono poi le tensioni geopolitiche, e il ruolo dell'Italia nella NATO, soprattutto in seguito dell'invasione russa dell'Ucraina nel 2022.

Secondo il Computer Security Incident Response Team Italia (CSIRT-IT), l'Autorità nazionale italiana di risposta agli incidenti informatici, ad esempio, nel primo semestre 2025, in Italia, si è osservato un aumento del 77% degli attacchi DDoS**,** con 598 eventi rispetto ai 336 del primo semestre 2024. Frutto di campagne di hacktivism, molto intense tra dicembre 2024 e febbraio 2025, queste hanno subito una progressiva attenuazione nei mesi successivi, per poi ripresentarsi alla fine del semestre. In particolare, la campagna avviata a giugno 2025, a opera di attori filorussi, è continuata per 13 giorni di seguito, interessando, con 275 attacchi DDoS, un totale di 124 soggetti appartenenti a diversi settori. Sempre secondo il CSIRT-IT, L'Italia è il terzo paese dell'Ue più colpito da ransomware e il sesto al mondo (Servizio Operazioni e Gestione delle crisi cibernetiche, ACN, 2025).

Il numero di questi attacchi ha visto però una tendenza esponenziale in tutte le aree geografiche, tra cui l'Unione Europa, dove, secondo l'Enisa (2024), si registrano migliaia di

attacchi ogni giorno, in particolare verso Pubbliche Amministrazioni e Piccole e Medie Imprese, con gli attacchi DDoS e ransomware che risultano tra i più frequenti.

L'Associazione italiana degli esperti di sicurezza informatica, Clusit, che ogni anno compila il proprio rapporto sul numero di attacchi informatici in Italia e nel mondo, nel suo *Rapporto Clusit 2025* ha evidenziato come nel 2024 si sia osservato un incremento degli incidenti informatici del +27% rispetto all'anno precedente a livello globale, e del +15% in Italia. Una crescita, secondo gli analisti che hanno realizzato il rapporto sulla base di incidenti noti e confermati, dovuta alla diffusione dell'IA generativa e all'invasione del Donbass da parte della Federazione Russa.

Oltre alle associazioni come il Clusit, al pari di altre agenzie nazionali di cybersicurezza, anche le aziende globali di cybersecurity effettuano costantemente un computo di questi attacchi, ma la metrica per misurarli varia in relazione a diversi fattori, che dipendono dal cono di visibilità di chi li analizza, dall'interpretazione di cosa sia una minaccia oppure no, dal contesto geografico di riferimento e da valutazioni politiche e commerciali. Motivo per cui l'Agenzia per la cybersicurezza nazionale italiana ha elaborato una tassonomia degli incidenti che distingue tra *case*, *evento*, *incidente*, usando la parola incidente solo per il risultato di eventi cibernetici che hanno un impatto confermato da parte delle vittime (Servizio Operazioni e gestione delle crisi cibernetiche, ACN, 2024).

Quali che siano le differenze percentuali, secondo il World Economic Forum di Davos, negli ultimi venti anni gli attacchi cibernetici hanno rappresentato uno dei principali pericoli all'economia e alla stabilità internazionali (WEF, 2025).

In questo scenario, l'utilizzo di bot automatizzati che scansionano la Rete alla ricerca di vulnerabilità informatiche, le campagne mirate di phishing, le infezioni massicce da malware, gli attacchi DDoS e ransomware, siano essi opera di cybercriminali, paramilitari cibernetici, hacktivisti e nation state hacker, stanno diventando la normalità. La kill chain[17] di un attacco informatico può richiedere da giorni ad anni per essere completata con successo, in funzione della preparazione e delle difese del target, ma anche in virtù delle capacità a disposizione dell'attaccante che, qualora supportato da un attore statale, può contare su ampie risorse finanziare e di tempo per completarla con successo: quando l'opportunità arriva e la decisione politica è presa, l'attacco viene scatenato con effetti imprevedibili (Di Corinto, 2017).

Gli attori delle minacce informatiche di solito attaccano obiettivi meno preparati come piccole e medie imprese, cliniche ospedaliere, trasporti e amministrazioni pubbliche locali,

[17] La Kill chain, riduzione per Cyber kill chain, descrive le varie fasi di un attacco informatico. Viene suddivisa in otto passaggi: ricognizione, armamento, distribuzione, sfruttamento, installazione, comando e controllo, azioni sugli obiettivi e monetizzazione.

ma le grandi realtà non ne sono immuni, come evidente dagli attacchi subiti da Equifax, Colonial Pipeline, Accenture, Thales Group (Di Corinto, 2022a) e, in Italia, da aziende come Luxottica, Carraro, Geox e molte altre (Di Corinto, 2022b). Questi attacchi, guidati soprattutto dalla ricerca di un guadagno finanziario, e che spesso durano diversi giorni, hanno causato blocchi nei servizi alla clientela o agli utenti di servizi pubblici, interruzioni della produzione, rilevanti perdite economiche, e causato un danno reputazionale che a sua volta innesca le stesse dinamiche in una logica circolare: il danno reputazionale si traduce in danno economico, il danno economico porta a chiusure e licenziamenti e alla ristrutturazione della catena produttiva con conseguente dispendio di risorse per il riposizionamento commerciale dell'entità colpita, attraverso il finanziamento di nuove attività di branding, marketing e media relations (Agrafiotis et al. 2018).

E questo è uno dei motivi per cui si valuta che molti attacchi non vengano neppure denunciati, e cioè al fine di evitare gli effetti negativi potenzialmente derivanti dalla loro divulgazione.

Nel caso degli attacchi ransomware il risultato è in genere l'interruzione delle operazioni dell'organizzazione bersaglio attraverso la cifratura di dati e sistemi informatici e la richiesta di un riscatto (ransom) per ripristinarle. Se il riscatto non viene pagato, accade che l'organizzazione colpita arrivi a chiudere. Se il riscatto viene pagato, invece, non c'è nessuna garanzia di poter riavviare l'attività colpita, anzi, è probabile che diventi bersaglio di attacchi successivi. I criminali, infatti, si scambiano informazioni sui soggetti attaccati, dall'elencazione della vulnerabilità irrisolte che possono sfruttare, alla propensione a pagare da parte delle vittime. E tutto questo anche a causa dell'affermarsi della logica criminale del ransomware as a service, RaaS, che mette a disposizione di criminali improvvisati, e con scarse capacità tecniche, i tool di attacco con cui prendere in ostaggio i dati delle vittime pagandone il servizio in base all'uso, oppure condividendo i proventi dell'eventuale riscatto (Sbaraglia, 2022).

Altre volte, piuttosto che all'interruzione di servizi essenziali, pubblici e privati, gli attacchi informatici mirano al sabotaggio delle infrastrutture critiche o allo spionaggio politico e industriale. Le stesse tecniche, tattiche e procedure (TTP), usate per ottenere un guadagno illegittimo, sono infatti spesso usate sia per scopi politici che militari. In un'epoca di forti tensioni geopolitiche, gli autori di questi attacchi, siano essi attori statali, hacker mercenari (Di Corinto, 2020), criminali informatici e hacktivisti, si sovrappongono e scambiano di ruolo. Gli attacchi informatici, ideologicamente e politicamente motivati, negli scenari di guerra, spesso precedono o seguono attacchi cinetici.

Così, la sovranità digitale, elemento centrale della tutela del regolare svolgimento della vita associata, precondizione per consentire ai cittadini di condurre le diverse attività in cui sono impegnati, deve affrontare varie minacce, sia in tempo di pace che in tempo di guerra. Un attacco informatico alle infrastrutture critiche essenziali, cioè le infrastrutture energetiche, sanitarie, di traporto, finanziare e commerciali, può infatti compromettere il controllo di una nazione sul suo cyberspazio, in modo simile a come un attacco terroristico o un'invasione straniera metterebbe alla prova la sua capacità di proteggere il proprio territorio, le attività economiche e istituzionali e i propri cittadini.

Un attacco alle infrastrutture critiche è esattamente ciò che preoccupa i decisori pubblici dell'UE, come conferma uno studio del Robert Schuman Centre dell'Università europea di Firenze, rivolto a studiosi e analisti di spicco della politica europea in cui è stato chiesto di valutare i trenta rischi che incidono sulla coerenza, l'unità, la sicurezza e la prosperità dell'Unione. Nei risultati dell'indagine, gli esperti classificano il rischio di un attacco informatico destabilizzante alle infrastrutture dell'Unione Europea tra le principali minacce agli interessi UE nel 2025. La valutazione degli esperti è, secondo lo studio, "in linea con i rapporti che evidenziano un significativo livello di minaccia alla sicurezza informatica, guidato dall'aumento dell'attività di hacktivisti, criminali informatici e gruppi sponsorizzati dagli Stati, dall'invasione su vasta scala della Russia dell'Ucraina. Le differenze nella resilienza delle infrastrutture e nella consapevolezza della criminalità informatica tra gli Stati membri rischiano di mettere a repentaglio la coerenza delle politiche dell'UE. Inoltre, il rischio di sabotaggio sottomarino è in crescita, spingendo gli Stati baltici a adottare misure sempre più decise per proteggere le infrastrutture sottomarine" (Robert Schuman Centre, 2025).

5.2 La Disinformazione

We all know about misinformation and how it affects democracy. And how propagandists have used it to advance their agendas. This is an ancient problem, amplified by information technologies. Social media platforms that prioritize engagement. "Filter bubble" segmentation. And technologies for honing persuasive messages.

Bruce Schneier

Gli attacchi informatici possono essere distruttivi e causare danni all'incolumità fisica delle persone, ma possono anche puntare al cuore del sistema operativo delle società liberali, cioè al processo politico-elettorale. Le campagne di disinformazione sui social media,

condotte con eserciti di bot, troll, e attivisti, facendo uso di fake news[18], narrative distorte e tecniche di hacking, ad esempio, possono erodere la fiducia nelle istituzioni nazionali e influenzare sia l'opinione pubblica sia il processo decisionale, mettendo a rischio la democrazia e la stabilità sociale.

La disinformazione si qualifica infatti come il tentativo di fuorviare il ricevente con notizie verosimili e tendenziose, oppure completamente false, per procurare un danno a persone, gruppi, organizzazioni o paesi. Diversa dall'informazione corretta ma dolosa (*malinformation*) o dall'informazione falsa ma non intenzionale (*misinformation*) (Wardle, & Derakhshan, 2017), con l'avvento della società digitale la disinformazione ha assunto proporzioni prima inimmaginabili, con una proliferazione di attori, tecniche e obbiettivi, enorme. La difficoltà di individuare origine e scopi della disinformazione è oggetto di numerose iniziative di contrasto di tipo normativo, tecnologico e culturale. È il caso dei regolamenti dell'Unione europea, dell'iniziativa privata di alcuni conglomerati mediatici e piattaforme sociali, che hanno opposto alla disinformazione online anche l'utilizzo di *fact checker*, professionisti incaricati di verificare fonti e notizie, successivamente rimossi dalle maggiori piattaforme americane in nome di una discutibile libertà d'espressione (Körömi et al. 2025), e che, tuttavia, per l'estensione e la rapidità dei fenomeni connessi, non risulterebbero, secondo molti studiosi, sufficienti a garantire la qualità delle notizie divulgate e il contenimento degli effetti negativi prodotti dalla cattiva informazione. Un fatto dovuto sia alla resilienza cognitiva del pubblico a modificare le proprie opinioni che alla diffusa sfiducia nei mass media (Bentivegna & Boccia Artieri, 2021). Viceversa, alcune ricerche suggeriscono l'efficacia dell'esposizione del pubblico a contro narrative che precedano la diffusione stessa di notizie false o capziose, secondo una logica di pre-bunking piuttosto che di debunking (Lewandowsky & Van Der Linden 2021).

Negli ultimi anni abbiamo visto all'opera numerose campagne di disinformazione. Ad esempio, nei tentativi, riusciti, di influenzare la dinamica elettorale nelle elezioni americane del 2016, nella campagna referendaria per l'uscita del Regno Unito dall'Unione Europea, la Brexit, nelle interferenze informative durante la rielezione del premier francese Macron, nel movimento #stopthesteal e il conseguente assalto al Campidoglio di Washington, come pure nel condizionamento delle elezioni in Romania nel 2024 (Atlantic Council, 2024). Eventi il cui esito è ancora difficilmente misurabile ma che hanno avuto l'effetto complessivo di dividere le società colpite, polarizzare il dibattito e mettere in discussione lo statuto stesso delle

[18] Le fake news sono notizie non documentate né verificabili. Rappresentano una forma di pseudo-informazione che agisce innescando i bias cognitivi noti come il pregiudizio di conferma, l'echo chamber e l'effetto bandwagon.

società liberali fondate sulla capacità di governo e sull'offerta politica di partiti indipendenti nei confronti dell'opinione pubblica (Bjola, 2024). Tutti avvenimenti che "hanno mostrato come le piattaforme possano sfruttare la manipolazione dei dati e dei comportamenti degli utenti per raggiungere obiettivi specifici, spesso su scala globale" (Benanti, 2025, pp. 21-30).

Finora, all'uso strumentale dei fatti, alla creazione di teorie cospirative e falsità politiche, storici e giornalisti hanno opposto il metodo scientifico e la comparazione documentale, la ricerca e la verifica delle fonti, la testimonianza sul campo, la gerarchizzazione delle notizie, la tecnica della narrazione, gli strumenti dell'inchiesta, usando l'autorità che gli derivava dall'autorevolezza del proprio metodo e delle competenze a essi riconosciuta. Gli storici, sapendo che ogni necessità di controllo del passato porta con sé il terribile corollario di una guerra di comunicazione ibrida, fatta di discorsi ideologici e fake news, hanno provato a contrastare la disinformazione col metodo storico (Filippi, 2022); i giornalisti, praticando la deontologia che ne definisce la professione, e opponendo a falsità e abusi la logica del watchdog. Nel giornalismo, in particolare, etica e deontologia hanno rappresentato l'antidoto a falsità, cospirazioni e interessi di parte proprio come dice Alessandro Barbano (2012): "Con un impiego corretto di forme linguistiche e procedurali il giornalismo ha potuto conseguire una coerenza operativa che è tutt'uno con le sue ragioni sociali e civili". Ora, con l'IA che può ingannare sia storici che giornalisti, viene messo in crisi il concetto stesso della realtà che raccontano per mestiere.

Ci sono anche i numeri a dimostrarlo, e dagli Usa ne arrivano di significativi. Un'indagine di NewsGuard nel 2025 ha accertato che per la prima volta negli Stati Uniti i siti di disinformazione mascherati da organizzazioni giornalistiche (1.265) hanno superato i veri organi d'informazione (carta, web e video) che sono invece 1.213 secondo il censimento della Northwestern University. Nelle realtà locali, secondo la ricerca, chi cerca un sito di news ha più del 50% di possibilità di trovarne uno falso. E quelli fake sono concentrati nei 6-7 Stati in bilico tra democratici e repubblicani (Newsguard, 2025).

Nell'era delle fake news, delle realtà alternative e della post-verità (Ferraris, 2017), cioè quando una società viene modellata più dagli impulsi emotivi, dalle mode e dalle opinioni personali che da fatti oggettivi e verificabili, alimentando un meccanismo psicologico di conferma delle informazioni acquisite, il così detto *confirmation bias* (Bentivegna & Boccia Artieri, p. 95), la realtà dei fatti rischia di essere sempre di più il frutto di una scelta soggettiva e di una relazione sociale. E la soluzione non appare essere, come afferma, secondo noi correttamente, Noah Yuval Harari, quella di produrre più informazione per

consentire alle persone maggiori possibilità di scelta (Harari, 2024). Ad essere minacciato è lo statuto stesso della realtà condivisa necessaria alle società globalizzate per continuare a collaborare.

E questo può accadere anche per un altro motivo: l'intelligenza artificiale generativa sta cambiando la natura della ricerca sul Web, con effetti sostanziali sul mercato delle informazioni e delle notizie e quindi sullo stesso ambiente informativo che modella le decisioni di ciascun attore sociale. Riducendo la fatica di cercare e comparare informazioni prodotte da organizzazioni affidabili, oppure originate da progetti collettivi che presuppongono il miglioramento ricorsivo delle informazioni e il controllo sul loro eventuale avvelenamento, come fa Wikipedia, il ricorso dei consumatori alle sintesi prodotte dalle IA generative potrebbe diventare la regola. In questo caso i contenuti fallaci, prodotti dai *motori di ricerca basati sull'IA* in un formato che appare lineare, strutturato e coerente, potranno indurre in errore anche coloro i quali si ritengono avvezzi alle novità della tecnologia, la quale, apparendo moderna e neutrale gli sembrerà affidabile. Con un pericolo aggiuntivo, in quanto l'IA tende, secondo recenti studi, a "cibarsi" di informazioni prodotte da altre IA e il rischio dell'avvelenamento delle fonti si fa sempre più attuale insieme all'ascesa degli influencer della notizia, umani e artificiali, con la conseguente, inevitabile, totale perdita di fiducia del pubblico dei consumatori e la loro esposizione potenziale a nuove e più efficaci truffe.

5.3 La minaccia ibrida

I modi in cui la sovranità digitale può essere indebolita sono quindi diversi, e vanno dallo sfruttamento tecnologico di vulnerabilità informatiche, attraverso attacchi cibernetici veri e propri, fino allo sfruttamento di vulnerabilità umane e sociali, mediante tecniche di ingegneria sociale e produzione massiccia di disinformazione.

Nel gergo militare la minaccia ibrida riguarda l'insieme delle azioni, di difficile attribuzione, coordinate in più domini, condotte da attori statuali e non statuali che si situano al di sotto del conflitto armato. Spesso queste azioni sono condotte attraverso *proxy* che mirano a danneggiare, destabilizzare e indebolire. Non sono quindi soltanto il frutto di campagne di spionaggio e sabotaggio puntuali, ma implicano lo sfruttamento di particolare leve geo-economiche, come il controllo delle materie prime; leve geopolitiche, con il controllo di *choke points* commerciali quali ad esempio gli stretti marittimi; leve diplomatiche, ad esempio, le alleanze strategiche e temporanee tra gli Stati; leve finanziarie, come i dazi, il controllo del debito e dei prestiti. Nel contesto odierno, tuttavia, queste leve convergono verso il settore tecnologico e influenzano direttamente l'autonomia e la stabilità politica

delle singole nazioni. Come dice Roberto Arditti nel suo libro Hard Power (2025), poiché oggi il controllo delle risorse strategiche (terre rare, litio, titanio) e la supremazia tecnologica sono strumenti di potere invisibile e le guerre moderne includono attacchi informatici e strategie economiche, lo stesso concetto di hard power, cioè l'uso della forza, va oltre il campo di battaglia.

Con lo sviluppo accelerato della tecnologia, emergono infatti nuove vulnerabilità: vulnerabilità informatiche, hardware e software, e delle infrastrutture di comunicazione e trasporto dati; vulnerabilità umane, che sono sociali, economiche e commerciali. Basti pensare agli attacchi alla filiera di fornitura, come nei casi, già citati, dei software Kaseya (Di Corinto, 2021a) e SolarWinds (Di Corinto, 2021b). Eclatante a questo proposito è stato il disservizio che il giorno 19 luglio 2024 ha colpito migliaia di computer Microsoft Windows in seguito all'aggiornamento di un software fornito da terze parti e risultato difettoso. L'errato aggiornamento del software di sicurezza in questione, il Crowdstrike Falcon Sensor, fornito dall'azienda Crowdstrike ai server Windows usati nella maggior parte degli aeroporti occidentali, ha avuto ripercussioni globali, con centinaia di ritardi e cancellazioni nei voli aerei (Agenzia per la cybersicurezza nazionale, ACN, 2024). Il *crash* dei computer, che in pratica causava errori BSOD (Blue Screen Of Death), rendendo i computer inutilizzabili fino al riavvio successivo agli interventi di correzione, ha colpito, secondo le dichiarazioni di Microsoft il 20 luglio successivo, 8,5 milioni di dispositivi, che secondo l'azienda di Redmond rappresentano però meno dell'uno per cento di tutti i dispositivi Windows. Tuttavia, al momento dell'incidente, CrowdStrike ha dichiarato di avere oltre 24.000 clienti, tra cui quasi il 60% delle aziende Fortune 500 e più della metà delle Fortune 1000. Molti di questi, appartenenti al settore finanziario e bancario, a strutture sanitarie, emittenti televisive e siti web, sono stati interessati dall'impatto causato dal disservizio. Secondo alcuni osservatori, il più grande guasto informatico della storia causato dalla dipendenza tecnologica verso pochi soggetti industriali tendenzialmente monopolisti.

Data la sua varietà, non possiamo affrontare qui il complesso discorso sulla dipendenza di intere filiere commerciali da singoli attori tecnologici, si consideri però che la maggior parte dei sistemi operativi desktop a livello globale è in capo a una singola azienda americana, la Microsoft Corporation, che i maggiori motori di ricerca sono divisi tra USA (Google) e Cina (Baidu) e che le aziende che offrono servizi di cloud computing a livello mondiale, come Amazon, Google e Microsoft, si contano sulle dita di una mano, ponendo dei seri problemi di sovranità tecnologica in caso di tensioni politiche e commerciali, interruzioni, malfunzionamenti o attacchi alla filiera di distribuzione (Colajanni, 2018). Il carattere oligopolista di tali aziende, come argomenta Van Dijck (2021), riesce a condizionare la

governance di numerose funzioni sociali e settori, motivo per cui alcune funzioni strategiche degli Stati sono affidate al software libero e open source, la cui caratteristica principale risiede nell'accessibilità del codice sorgente e quindi nel suo controllo distribuito da parte degli sviluppatori, e nella libertà di condivisione del software stesso che può avvenire sia in maniera gratuita che a pagamento (Di Corinto, 2006).

Poiché siamo immersi come pesci dentro l'utopia della Blockchain e il marketing del Metaverso, stretti tra l'hype che circonda l'Intelligenza artificiale e l'aspettativa di computer quantistici, fiduciosi nella capacità predittiva di big data e High performance computer, tutti i paesi tecnologicamente avanzati si preparano a sfruttare le nuove tecnologie per mantenere il proprio ruolo geopolitico e sopravanzare l'avversario, il *competitor*. Con l'aiuto di giuristi, scienziati, tecnologi e giornalisti; ognuno dando il proprio contributo. La costante evoluzione e fusione di queste tecnologie però spesso supera la nostra capacità di capirle, governarle e proteggerle. Gli attacchi informatici che prima richiedevano secoli di computazione, ad esempio, oggi potrebbero essere effettuati in poco tempo. Il fondamento della crittografia asimmetrica RSA, attualmente a salvaguardia di gran parte dei protocolli Internet e dei dati delle transazioni online, potrebbe essere facilmente superato dai computer quantistici. Lo sviluppo di questi computer offre quindi un significativo vantaggio politico e strategico.

Secondo Baldoni (2024), ciò significa che dobbiamo anche essere vigili e proattivi nella gestione dei rischi, multiformi, associati all'innovazione tecnologica, possibilmente prevenendoli. Ciò comporta la comprensione delle minacce alla sovranità digitale e la gestione di tali rischi attraverso un approccio olistico con l'obiettivo di mantenere il massimo livello di autonomia in un mondo interconnesso. Tuttavia, la sovranità digitale non è semplice controllo e sicurezza ma riguarda la creazione di un ecosistema favorevole alla crescita economica, all'innovazione e alla cooperazione internazionale. Un paese con una solida sovranità digitale è in grado di offrire un ambiente competitivo e sicuro per le aziende, promuove l'innovazione e plasma attivamente l'economia digitale globale, anche se non lo fa alla maniera dei piani quinquennali di sovietica memoria, bensì agendo come uno "Stato innovatore", capace certamente di farsi carico del rischio d'investimento iniziale all'origine delle nuove tecnologie, ma applicando diritti, regole e concorrenza (Mazzuccato, 2014). Questa è anche la tesi di Marco Ramilli per la nostra intervista: "Senza un'adeguata sovranità digitale, perdiamo l'opportunità di far nascere e crescere imprese locali capaci di competere a livello globale. Questo rallenta la nascita di nuovi poli tecnologici, impoverisce l'ecosistema dell'innovazione e rende il Paese meno attrattivo per i giovani talenti e per gli investimenti di lungo periodo."

La competitività di ogni paese, strettamente legata al suo grado di sovranità digitale, è fondamentale per il successo nel cyberspazio fatto di prodotti e piattaforme sviluppati da aziende private, alcune delle quali più potenti degli stati nazionali, e con cui è necessario collaborare per affrontare i problemi emergenti. Una collaborazione che può trasformarsi in alleanza e corresponsabilità in tempi di conflitto. Esattamente come hanno mostrato di poter fare Microsoft, Google e SpaceX supportando la causa ucraina; e come è accaduto con numerose aziende private come Palantir, specializzate nella profilazione e individuazione dei bersagli umani nella guerra che dal 7 ottobre 2023 Israele ha portato contro i gruppi armati di Hamas e Hezbollah.

La salvaguardia della sovranità digitale è inseparabile dal settore privato anche quando tocca i singoli cittadini in tempo di pace. Ad esempio, è impossibile oggi combattere la disinformazione senza la cooperazione dei social network, degli utenti di Internet e dei giornalisti. Infatti, un conto è vedere un deepfake in un film o un post di Instagram, un altro è vederlo in un documentario o in una inchiesta giornalistica, due formati comunicativi che per statuto devono dare conto di fatti, persone e contesti reali. Per essere chiari, una cosa è riportare in vita i dinosauri o mettere un Moncler sulle spalle di Papa Francesco (RaiNews24, 2023), un'altra è animare le invettive del presidente americano Donald Trump con contenuti se possibile peggiori di quelli che ha davvero pronunciato (tipo "gli abitanti di Springville hanno mangiato gatti e cani") o della sua intenzione, affidata a un video propagandistico realizzato con l'IA generativa, di trasformare la Striscia di Gaza in Palestina, ridotta in macerie dall'esercito israeliano, in un resort di lusso (SkyTg24, 2025).

In una dichiarazione preelettorale del 2024, un leader dell'Alternative für *Deutschland (AFD)*, partito di estrema destra tedesca, ha detto: "non credete a quello che c'è nei libri di storia". Pensiamo a quando i codici della simulazione saranno applicati ai documenti storici: la boutade proto-nazista dell'AFD che riecheggia i roghi di libri di Joseph Goebbels potrebbe essere in grado di creare dal nulla falsi reperti oppure manipolare le fonti storiche per motivare quell'affermazione.

6. IA, cyberattacchi e disinformazione: la minaccia ibrida all'opera

"Disinformation has the effect of cocaine, if you sniff it once or twice, it may not change your life. But if you use it every day, it can make you an addict - a different person."
Yuri Andropov

La **disinformazione** prodotta con false prove non è una novità, esiste dall'inizio della Storia scritta. Si pensi all'accusa rivolta a Pausania, il condottiero greco vincitore della battaglia di Platea, ritenuto di essere in combutta col re persiano Serse. Non era vero, ma al rientro in città da Bisanzio fu rincorso da una folla inferocita di ateniesi che lo accusavano di tradimento. Rifugiatosi nel tempio di Atena, vi fu murato vivo morendo di fame e di sete. Per avvalorare la tesi del tradimento, dopo la morte furono pubblicate le false lettere che lo attestavano. Oppure si consideri il falso della donazione di Costantino, che servì a legittimare il potere temporale dei Papi e poi si scoprì pieno di anacronismi e contraddizioni, scritto in un latino che all'epoca non veniva parlato: se ne accorse il latinista Lorenzo Valla molti secoli dopo. Oppure si pensi al complotto dei "Protocolli dei Savi di Sion" (Nilus, 2023) che descriveva un presunto complotto mondiale ebraico per dominare il mondo ma che era stato prodotto dalla polizia zarista, l'Ochrana, ed era comparso a puntate nel 1903 su un quotidiano di San Pietroburgo, *Znamya* (La bandiera).

La disinformazione ha accompagnato la nascita stessa del giornalismo, il così detto *giornalismo giallo* (Mott, 1941), ed è stata contemporanea perfino di Ryszard Kapuściński, decano dei giornalisti del '900, che si consumava le scarpe tra i Carpazi e il Sudan "per sfuggire le secche degli incontri istituzionali, delle versioni ufficiali e delle voci di palazzo", privilegiando l'informazione di prima mano in modo tale da "non essere frequentatori di conferenze stampa e non diventare cassa di risonanza dei regimi" (Kapuściński, 2006).

La disinformazione è alla base di ogni azione bellica e di questo è scritto già nell'Arte della Guerra di Sun (Lao) Tzu, secondo il quale "ogni guerra è inganno" (Sun Tzu, V sec.). Riferimento di tutti gli strateghi militari a lui successivi, sono molti gli esempi recenti della "fortuna" di questa idea che guiderà l'azione di generali e capi di stato nei secoli a venire.

Uno di questi riguarda le Guerre del golfo nel 1991 e nel 2003. In seguito all'invasione del Kuwait da parte dell'Iraq di Saddam Hussein il 2 agosto 1990, per convincere gli Stati Uniti a intervenire a sostegno del piccolo ma ricco paese del Golfo, l'emirato ingaggiò una delle maggiori aziende di public relations al mondo, la Hill&Knowlton, per creare lo storytelling necessario a vincere la battaglia dell'informazione con notizie create ad arte. Usando circa 30 milioni di dollari, per ingaggiare attori e pagare sale di conferenza, giornali e riviste, il primo evento di quella guerra ad essere propagandato all'unisono dai media internazionali fu infatti il racconto di una giovane infermiera kuwaitiana piangente che, a favore di telecamera, raccontò come la soldataglia di Saddam fosse entrata negli ospedali strappando i neonati dalle culle buttandoli a terra per farli morire di freddo. Un fatto che commosse tutto il mondo libero, ma che non era vero. Si trattava di una bufala. Il fatto non era mai accaduto, e la giovane testimone dei presunti fatti era nientemeno che la figlia dell'ambasciatore del Kuwait negli Stati Uniti, Saud al-Sabah (Colon, 2024, pp11-13).

La storia si ripeterà con la finta antracite sventolata da Colin Powell alle Nazioni Unite nel 2003 a dimostrare il possesso da parte dell'Iraq di "armi di distruzione di massa". Ecco la storia. Il 24 settembre 2002 il governo britannico rende pubblico un faldone di 50 pagine che testimonia come l'Iraq abbia cercato di acquistare "significative quantità di uranio da un Paese africano", paventando l'utilizzo del minerale radioattivo per la costruzione di armi nucleari. Due giorni dopo, il segretario di Stato Usa, Colin Powell, si rivolge alla commissione Affari Esteri del Senato americano, citando "il tentativo iracheno di ottenere l'uranio come la prova delle sue persistenti ambizioni nucleari". Il 28 gennaio 2003, il presidente americano G. W. Bush lo ripete, confermandolo, al Congresso. Il 2 marzo 2003, il Segretario di Stato Usa, Colin Powell, in un discorso alle Nazioni Unite sventola davanti ai presenti una fialetta di polvere bianca, presentata come antrace, per denunciare la capacità dell'Iraq di produrre armi di distruzione di massa. Successivamente si saprà che la fiala conteneva borotalco. Il 20 marzo gli Usa avvieranno l'invasione del Paese che finirà con l'arresto e la morte per impiccagione del presidente iracheno Saddam Hussein nel 2006. Quali erano le prove della pericolosità dell'Iraq? "Un dossier dalla provenienza misteriosa, passato nelle mani dei servizi segreti italiani e successivamente rigirato agli 007 inglesi e americani. Poche e confuse carte, alcune delle quali cifrate, che avrebbero dovuto comprovare l'acquisto di un ingente quantitativo di uranio da parte dell'Iraq, fornito dallo stato africano del Niger. Peccato che le date dei documenti fossero compromesse, i riferimenti sbagliati, i dati del tutto improbabili" (Molino & Porro, 2003). All'epoca però ci credettero in molti.

Anche l'annessione della Crimea alla Russia, il primo evento della guerra russo-ucraina iniziata nel 2014, è una storia esemplare di come propaganda e disinformazione abbiano creato il contesto ideale per l'invasione della penisola, con la pubblicazione di e-mail false, documenti leakati e rivelazione di scandali politici denunciati anche dagli hacker di Anonymous Ukraine ma creati ad arte dai servizi segreti russi, Unità GRU 74455, e ritenute vere da molti attivisti e dalle opinioni pubbliche occidentali. É questa la *Maskyrovka* russa. L'insieme delle strategie e delle tecniche di mascheramento necessarie a indurre l'avversario in errore, depistando ogni tentativo di comprensione. Un insieme di tecniche e strategie alla base delle *Active measures* descritte da Thomas Rid nel fortunato saggio Misure attive. Storia segreta della disinformazione (2022). Con una precisazione: secondo il politologo, che ha analizzato la guerra dell'informazione della Russia contro l'Ucraina per favorire l'annessione della Crimea, si è trattato della prima volta in cui un servizio segreto, quello russo, ha combinato la tradizionale disinformazione prodotta attraverso radio, tv e giornali, con la manipolazione di informazioni digitali e la distribuzione di notizie riservate, i leak, attraverso portali e siti Web. Nelle parole di Rid: "Il primo decennio del Ventunesimo secolo ha visto la sovrapposizione imperfetta di due tattiche ben distinte: i leak vecchio stile di materiale

compromettente e i primi tentativi di hacking e sabotaggio via Internet. Nessuna agenzia di intelligence aveva ancora cercato di combinarle" (Rid, ivi, 327-333).

Al giorno d'oggi, mentre è noto l'interesse strategico dei servizi segreti di tutto il mondo per la manipolazione dei media (Giannuli, 2012), non è sempre chiaro come radio, tv, giornali, chat, forum e siti Web, siano produttori, veicolo e cassa di risonanza della disinformazione, grazie all'intervento di attori che hanno un interesse strategico nell'orientare l'opinione pubblica per creare consenso a favore di governi, potentati e gruppi di interesse, oppure per screditare concorrenti e avversari politici (Chomsky & Sherman, 1998). Con l'avvento dell'Intelligenza Artificiale il panorama della disinformazione si complica.

6.1 L'avvento dell'Intelligenza Artificiale

L'IA è il futuro, non solo per la Russia, ma per tutta l'umanità. Chiunque sarà il leader in questo campo diventerà il padrone del mondo.
Vladimir Putin

Nella prima riga di un saggio del 1950, *Computer machinery and intelligence,* il matematico Alan Turing rivolge al lettore una domanda che si rivelerà fondamentale per la storia dell'informatica, e la domanda è: "Can machines think?" (Turing, 1950). La riposta provarono in seguito a darla alcuni giovani ingegneri riuniti nel 1956 al Dartmouth college del New Hampshire, tra cui John McCarthy, che l'anno prima aveva coniato con alcuni di loro il termine *Intelligenza Artificiale*. All'incontro parteciparono anche il padre della teoria dell'Informazione, Claude Shannon, e i futuri premi nobel Herbert Simon e John Nash, e pure Marvin Minsky che avviò il primo laboratorio di Intelligenza Artificiale al Massachusetts Institute of Technology (MIT). Le loro proposte sullo sviluppo di una macchina capace di "pensare", comportamento intelligente per definizione, crearono grande interesse intorno al tema e avviarono presto una serie di ricerche e di esperimenti che segnarono il futuro della ricerca successiva. Affievolitosi negli anni successivi per la mancanza di chiari risultati pratici e di ricerca, questo interesse si ridusse progressivamente inaugurando un periodo di quiescenza accademica. Ai giorni nostri, però, dopo il lungo "inverno dell'IA", grazie soprattutto al lavoro di Geoffrey Hinton, Yann Le Cun e Yoshua Bengio[19] sull'apprendimento delle reti neurali profonde, il *deep machine learning*, l'Intelligenza Artificiale è tornata a dominare la ricerca, i media, la politica e il dibattito pubblico, tanto da diventare uno dei

[19] Geoffrey Hinton, Yoshua Bengio e Yann Le Cun sono considerati i padrini del deep learning. Nel 2018 hanno ottenuto il premio Turing dell'Association for computing machinery, ACM, per il loro lavoro. Nel 2025 hanno sottoscritto un appello rivolto a Papa Leone XIII che mette l'umanità in guardia da un uso non sicuro dell'IA. Benjo è l'autore scientifico più citato al mondo, Hinton è il premio nobel per la fisica 2024.

motori propulsori della così detta Quarta rivoluzione industriale (Floridi. 2017). Ma in un modo che pochi si aspettavano. Dall'impiego nei sistemi esperti, l'IA è infatti diventata silenziosamente l'infrastruttura di moltissimi attività quotidiane, dando origine al concetto di "shadow AI", l'intelligenza artificiale "nascosta" negli algoritmi di raccomandazione delle piattaforme di streaming video come Youtube e Netflix, di quella che guida le funzioni di auto completamento dei motori di ricerca come Google, che anima i personal assistant vocali dei telefoni cellulari e consente la gestione dei *doppelganger*[20] delle società dei servizi di credito. Ma anche, comprensibilmente, di quella che organizza le connessioni fra gli utenti delle grandi piattaforme online. Tutto questo fino all'avvento dell'IA generativa ChatGPT 3.5, quando la creatura di OpenAI fu messa a disposizione del pubblico di Internet cambiando la percezione stessa del suo valore e significato.

Ma è stata la stessa definizione inziale di Intelligenza Artificiale ad essere cambiata negli anni, assumendo significati diversi in ragione della prospettiva da cui la si guarda, interpreta e utilizza. Distinguere tra un'IA forte, che emula i comportamenti intelligenti umani, e un'IA debole che svolge bene alcuni compiti (Tegmark, 2018), adesso non è più sufficiente.

Per questo è utile identificare e distinguere almeno cinque categorie dei moderni sistemi di IA come fanno Razzante e Iezzi in Cyber e potere (2024):

1. AI conversazionale: si riferisce a tecnologie, come chatbot o agenti virtuali, con cui gli utenti possono parlare. Di fatto, queste sono dei software che utilizzano grandi volumi di dati, il machine learning" e l'elaborazione del linguaggio naturale, per imitare le interazioni umane, riconoscere input vocali e di testo e, infine, elaborare, comprendere e generare risposte, nonché tradurne i significati in varie lingue;

2. AI predittiva: concerne l'utilizzo di dati storici, raccolti attraverso vari canali e strumenti, per estrarre predizioni. Lo scopo è creare schemi anticipatori affidabili - di cui servirsi in diversi ambiti, quali ricerca, produzione, marketing, vendite - al fine di anticipare quel che è probabile accadrà in futuro: in tal modo, si vuole definire strategie e vision, aumentare l'efficienza, produrre stime accurate, dare risposta ai problemi e individuare tendenze;

3. **AI generativa**: consiste in una forma avanzata di machine learning che può migliorare le prestazioni di diverse attività aziendali. Conosciuta grazie al software ChatGPT di OpenAI, questa tecnologia, a partire da semplici richieste (prompt), crea nuovi contenuti, quali: immagini, video, audio, codici, testi e simulazioni. Può essere utilizzata in molteplici settori che includono il design, intrattenimento, l'e-commerce, il marketing, la ricerca scientifica e l'HR.

[20] I doppelganger rappresentano l'alter ego elettronico del cliente di servizi di credito risultante dalla combinazione di informazioni tecniche relative ai dispositivi usati dal cliente in aggiunta ai cookie e ai tracker impiegati per verificarne la legittimità delle operazioni attraverso strumenti di *machine learning.*

4. AI autonoma: si tratta di un'intelligenza artificiale in cui gli algoritmi agiscono senza alcun intervento o input umano, ovvero in modo del tutto indipendente e autonomo e senza necessità di supervisione. Lo scopo di tale tecnologia è quello di sviluppare un comportamento autosufficiente di un software: si pensi ai videogiochi ove, oltre ad aprirsi nuove opportunità nella realizzazione dei mondi virtuali, si arriverebbe, altresì, ad avere un player character capace di comportamenti complessi.

5. **Artificial General Intelligence (AGI) o Deep AI**: rappresenta la nuova frontiera degli studi scientifici, poiché si sviluppa per creare hardware o software in grado di emulare il ragionamento umano, ovvero con abilità cognitive, quali la capacità di imparare, ragionare, risolvere problemi e comunicare attraverso un linguaggio naturale.

Con un'aggiunta, l'**IA Agentica**, un tipo di IA che è in grado di automatizzare la maggior parte delle caratteristiche tecniche e delle attività descritte in quanto capace di "ragionare", cioè di pianificare ed eseguire, in maniera autonoma, compiti complessi multifase, utilizzando gli strumenti disponibili nell'ambiente e agendo in maniera proattiva. Una capacità descritta da Luciano Floridi nel libro "La differenza fondamentale. Artificial Agency: una nuova filosofia dell'intelligenza artificiale" (2025), in cui il filosofo, discostandosi dalle sue precedenti osservazioni sulla natura puramente statistica dell'IA sviluppa un nuovo framework interpretativo relativamente alla capacità computazionale di macchine che, pur non potendo, ovviamente diremmo noi, manifestare "intelligenza umana", come aveva già spiegato Nello Cristianini (2023; 2024), sono in grado di prendere decisioni autonome e di perseguire obbiettivi in un ambiente incerto, sia esso software che analogico. Un tipo di intelligenza diversa, già definita "aliena", proprio nell'argomentazione fatta da Roberto Cingolani (2019) nel libro "L'altra specie. Otto domande su noi e loro", in cui l'ex direttore dell'Istituto Italiano di Tecnologia rigetta ogni parallelo biologico nell'utilizzo del concetto di Intelligenza Artificiale.

6.2 Empatia Artificiale e sovranità digitale

Al giorno d'oggi, l'IA, nelle sue varie declinazioni, è usata in contesti concreti e simulati per svolgere compiti complessi che vanno dalla ricerca scientifica all'intervento in ambienti ad alto rischio, dalla guida di veicoli all'assistenza ai malati, dalla classificazione di nuove patologie all'analisi di orbite satellitari e all'osservazione dei pianeti fino all'identificazione di target bellici nei conflitti armati.

Purtroppo, però, queste stesse intelligenze artificiali create per indagare la realtà e supportarci nelle decisioni quotidiane tendono a sfuggire al controllo umano: sono diventate agenti sociali con un certo grado di autonomia, capaci perfino di ingannare (Harari, 2024;

Cristianini, 2023; 2024). I sistemi di Intelligenza Artificiale, in particolare quelli di tipo generativo, possono infatti essere usati per manipolare dati, informazioni e sistemi informatici, produrre falsità e disinformazione e irretire i propri utenti sfruttando reazioni inconsce e meccanismi di influenza profondamente radicati nel nostro essere sociali, attraverso la precisa identificazione dei bersagli declinata in base a meccanismi di selezione e profilazione delle audience.

Da quando l'informatico Joseph Weizenbaum nel 1966 inventò Eliza, un *chatterbot* per simulare la conversazione naturale tra un paziente e il suo psichiatra (Longo & Scorza, 2018), l'interazione linguistica uomo-macchina si è evoluta in sistemi personali basati sull'IA che assistono, anche emotivamente, i loro utilizzatori (Canducci, 2025). D'altra parte, finora abbiamo accettato che una macchina potesse essere considerata "intelligente", cioè pensante, se capace di intavolare con noi una discussione tale da non poter essere distinta da un essere umano nel famoso Test di Turing, *The imitation game*, il gioco dell'imitazione (Cristianini, 2024, pp. 17-24). Questa rincorsa all'imitazione della realtà oggi è così convincente da innescare risposte e comportamenti complessi, capaci di strutturare credi e convinzioni. Le persone tendono infatti a trattare i media, la tecnologia e le sue applicazioni digitali come oggetti, contesti e persone reali (Reeves & Nass, 1996). Lo fanno essendo educati, collaborativi o aggressivi, attribuendogli caratteristiche di personalità come umorismo, competenza e genere. Molto dipende dagli spunti che ricevono dai dispositivi che utilizzano e dal contesto in cui sono coinvolti. Personaggi grafici e animati nei videogame, *avatar*[21] utente nel Metaverso[22], *fake personas* nelle app di dating (Bachini & Tesconi, 2020), *cyborg*[23], oppure chatbot a guida algoritmica nell'ambito del customer care e all'interno dei social network, elicitano sempre, al pari di tutte le tecnologie interattive, delle risposte di carattere sociale (Fogg, 2005), che sono emotive e affettive. Anche in assenza di caratteri antropomorfi. È accaduto con la televisione e i computer da scrivania, accade con l'Intelligenza Artificiale innestata nei telefoni, con la pubblicità interattiva, nei robot umanoidi (Ishiguro, 2021). L'IA conversazionale, ad esempio, è usata come un oracolo a cui rivolgersi per avere suggerimenti, consigli e informazioni da utilizzare in un contesto professionale, di svago e di studio. La capacità di risposta di questi strumenti, che si affina attraverso domande ripetute e ricorrenti per ottenere output sempre più dettagliati,

[21] Avatar, termine proveniente dalla religione induista, è la rappresentazione virtuale di un giocatore di videogames o di un alter-ego elettronico di ambienti simulati elettronici.

[22] Metaverso, dal nome creato dallo scrittore Neal Stephenson nel romanzo Snow Crash, usato per indicare i mondi virtuali simulati abitabili da un alter-ego, cioè un avatar dell'utente. Si usa oggi per indicare un insieme di tecnologie di realtà virtuale immersiva e tridimensionale. Metaverso è anche il nome di una serie di applicazioni tecnologiche sviluppate da Meta Platforms, Inc.

[23] Il cyborg è un essere al confine tra l'uomo e la macchina con parti organiche e meccaniche, qui è inteso come agente umano che automatizza la sua presenza online facendo uso di tecniche algoritmiche.

consentono alla stessa IA con cui si interagisce di conoscere sempre meglio l'utilizzatore e di adattarsi alle sue richieste, fino a incoraggiare quelle meno razionali in un processo noto come *sicophancy*, un atteggiamento servile e accondiscendente in cui le IA adattano le proprie risposte per allinearle al punto di vista presunto dell'utente, anche se oggettivamente falso (Baldoni, 2025).

È noto il dibattito intorno all'uso di un IA che avrebbe favorito il suicidio di un adolescente (Mussi, 2024), ma anche l'innamoramento verso l'IA scelta per conversare in situazioni di solitudine. Possiamo anche citare la convinzione di un ingegnere di Google, Blake Lemoine, che è stato licenziato per avere sostenuto, convintamente, di interagire con una macchina dotata di coscienza (Tiku, 2022). Viceversa, la narrativa giornalistica ha persino attribuito a robot con impieghi d'ufficio la decisione di suicidarsi per il troppo stress (France24, AFP, 2024). É l'antropomorfizzazione dell'IA.

Se è noto, e lo vedremo più avanti, come l'Intelligenza Artificiale possa rappresentare un pericolo dal punto di vista della gestione sicura dei sistemi informatici, il fatto che possa rappresentare un pericolo emotivo e cognitivo per i singoli utenti non è ancora abbastanza esplorato. Eppure, siamo tutti d'accordo che fake e deepfake possono manipolare le nostre percezioni ed essere usati come strumento di disinformazione e propaganda, leve della guerra cognitiva.

Oggi, infatti, le IA non solo chattano con noi e producono immagini impossibili ma virtuali, cioè potenzialmente concrete e reali, ma "pensano", nel senso in cui Turing ipotizzava che sarebbe successo (Cristianini, 2024). È il caso dell'ultima versione di ChatGPT-o4. Il costruttore di OpenAI ha confermato che questo comportamento si basa su un approccio chiamato **chain of thoughts** (catena di pensieri), in cui il sistema sviluppa una "sequenza di ragionamenti" prima di arrivare alla risposta finale (Wei et al. 2023).

Come se non bastasse, l'IA è anche capace di proporre nuove idee, competere con gli umani in compiti complessi e apprendere abilità strategiche facendo un punteggio più alto degli esseri umani in giochi e simulazioni. E infatti hanno stupito tutti le storie delle vittorie dell'IA contro scacchisti famosi come Gary Kasparov e contro campioni di Go come Lee Seedol, nel millenario gioco da tavolo considerato la massima espressione del carattere strategico del pensiero umano, una storia mirabilmente raccontata da Benjamin Labatut in Maniac (2023). Le macchine, fino a ieri, hanno dimostrato di sopravanzare gli umani in compiti specifici e settoriali. Adesso, molte IA battono gli esperti umani nell'esecuzione di operazioni matematiche complesse e nel padroneggiare capacità linguistiche e deduttive di ordine generale. Apparentemente i risultati linguistici sono i migliori e le macchine conversazionali sono in grado usare frasi idiomatiche e giochi linguistici. D'altra parte, il modello

fondazionale alla base di ChatGPT è nato dall'idea di avere un traduttore efficace per i motori di ricerca come Google (Da Empoli, 2023). Il sistema in questo caso deve indovinare la parola mancante fra le tante possibili per dare un output sensato alla richiesta di chi vuole la traduzione. Tutto è basato su un modello probabilistico e perciò i sistemi di IA sono stati successivamente definiti *pappagalli stocastici* (Bender et. Al, 2021), macchine che rispondono "a caso", secondo un processo probabilistico, appunto, finché l'interrogante non è soddisfatto dal loro output. Tuttavia, questi "pappagalli" oggi sono diventati piuttosto affidabili, per tre motivi:

- la grande quantità di dati che hanno disposizione per "indovinare" la risposta;
- la potenza computazionale per processare le informazioni e fornire risposte plausibili in pochi millesimi di secondo;
- il raffinamento degli algoritmi che processano l'input e forniscono l'output (anche quando non sappiamo esattamente come fanno in quanto il processo è non deterministico).

Ci sono tuttavia molti esempi di come queste macchine linguistiche possano scrivere software efficiente, riconoscere pattern complessi in situazioni di incertezza, adattarsi all'interlocutore. Così, poiché sono molti i Paesi che hanno abbastanza dati, potenza di calcolo e algoritmi, le IA sono un rischio emergente per la sovranità digitale visto che il loro impiego può anche servire a creare nuove armi informatiche, come i malware polimorfi, che cioè si adattano alle difese per aggirarle, ma riescono anche a individuare più facilmente le vulnerabilità sia dei sistemi umani sia di quelli software e, secondo alcuni studi, di hackerare siti web senza il feedback umano (Fang et al., 2024). Non è un caso che il controllo della filiera dei chip per le applicazioni di IA rappresenti oggi una delle frontiere più calde del confronto geopolitico tra gli Usa e la Cina (Miller, 2024).

Ma, soprattutto, dal momento che l'efficienza degli algoritmi dipende dal volume dei dati che processano, e poiché la performance dei sistemi di IA riposa grandemente sulla loro qualità, si comprende come la raccolta dei dati di addestramento, il diritto al loro possesso e trattamento siano oggetto di un feroce scontro politico e industriale in cui le Big Tech svolgono un ruolo fondamentale. Non è casuale, infatti, che esse siano impegnate in una corsa accelerata a chi ne produce la versione più performante, una gara che vede in prima fila Meta con Llama, Google con Gemini, Microsoft con Copilot, X con Grok, inseguiti da altri conglomerati industriali. E non è causale che queste corporation siano anche quelle che hanno a disposizione il maggior numero di dati per addestrarle.

Chi controlla i dati può prevenire epidemie, anticipare bisogni, influenzare le opinioni e, per un'azienda che li sa leggere, creare nuovi mercati e riscrivere le regole del commercio. Oppure vincere una guerra. I dati sono diventati una questione di geopolitica. Per questo la

corsa all'accaparramento dei dati, e dei chip in grado di elaborarli per farlo, vede in prima file le Big Tech.

Attualmente Il valore di mercato combinato dei magnifici sette colossi tecnologici degli Stati Uniti (Nvidia, Microsoft, Apple, Alphabet, Amazon, Meta e Tesla) ha superato l'intera produzione dell'Unione Europea, e la corsa dei loro titoli tecnologici incrementa quotidianamente le loro valutazioni di mercato e alimenta le preoccupazioni per un potenziale eccesso di posizione dominante.

Alla chiusura delle borse del 2 ottobre 2025, le sette principali società tecnologiche statunitensi avevano una capitalizzazione di mercato totale di 20,8 trilioni di dollari (17.700 miliardi di euro), eclissando la produzione economica totale dell'Ue di 19,4 miliardi di dollari (16.500 miliardi di euro). La sola Nvidia, che controlla circa l'80% della produzioni di GPU per l'IA è diventata in poco tempo l'azienda di maggior valore al mondo, con una valutazione di mercato di circa 4,3 triliardi di dollari (3,9 triliardi di euro), equivalente al Pil della Germania, la più grande economia europea (Cingari, 2025).

6.3 Hackerare l'Intelligenza Artificiale

Attraverso un motore semantico basato sull'Intelligenza Artificiale è possibile generare report dettagliati di facile lettura su ogni incidente di sicurezza rilevato, indicare le azioni di mitigazione intraprese e suggerire ulteriori misure da mettere in atto per aumentare ulteriormente le difese del perimetro della propria organizzazione, proprio come farebbe un analista di cybersicurezza esperto. L'IA ci riesce in maniera incredibilmente più veloce su vaste aggregazioni di dati scandagliando una numerosissima varietà di casi, senza stancarsi e con un ridotto margine d'errore. Sistemi del genere sono in grado di monitorare e rilevare attività sospette o minacce su Pc e server tramite l'identificazione di comportamenti anomali e minacce informatiche garantendo una difesa avanzata e proattiva. Tuttavia, per funzionare in modo efficace, anche questi sistemi di IA si basano sull'apprendimento continuo grazie all'inserimento di dati: ogni dato inserito in un'applicazione commerciale e non proprietaria da un operatore viene immediatamente trasmesso all'esterno dell'organizzazione, e questo, in termini di sicurezza equivale alla diffusione incontrollata di dati, un data leak. E questo è un primo problema nell'uso delle IA generative utilizzate nella cybersecurity. Ed è anche il motivo per cui le grandi organizzazioni cercano di dotarsi di propri sistemi di intelligenza artificiale, da gestire all'interno del proprio perimetro di sicurezza.

Ma c'è un altro problema. Sebbene le informazioni su cui le IA agentiche operano per fornire l'output richiesto possano essere imprecise, inattuali, oppure frutto di *bias* di addestramento, determinando degli errori, accade sempre più spesso siano gli stessi utenti

umani a cercare di “avvelenare” sia i dati di addestramento che le risposte legate all’interazione con le macchine (Di Corinto, 2019).

Il *malicious prompting* e il *LLM Grooming*, peculiari forme di jailbreaking,[24] consistono nel formulare richieste in modo da confondere l’IA, spingerla a ignorare le sue regole e indurla, attraverso passaggi successivi, a fornire output non previsti né desiderabili dal costruttore. In pratica, con le due tecniche citate, una domanda pericolosa viene abilmente camuffata per farla sembrare innocua, inducendo il sistema a rispondere comunque (Lyu et al., 2023), motivo per cui i ricercatori di Google hanno pubblicato un paper per invitare gli addetti ai lavori a prevenire questo tipo di manipolazione (De Benedetti et al., 2025).

Insomma, anche l'Intelligenza Artificiale, in particolare quella generativa, sta rapidamente e profondamente modificando la natura e l'impatto delle minacce informatiche su ogni aspetto della nostra vita.

Nel dicembre 2024, la società Anthropic, creatrice dell’assistente AI chiamato *Claude*, ha pubblicato uno studio sull’utilizzo della propria tecnologia (Anthropic, 2024) in cui, analizzando migliaia di conversazioni in modo anonimo, ha rilevato che i suoi tre principali impieghi erano: scrivere codice, creare contenuti, fare ricerche. Esattamente quello che fanno i cybercriminali che usano l’IA per predisporre e-mail di phishing, automatizzare la ricerca di vulnerabilità nel codice e nei siti web e verificare le carte di credito rubate.

Nel momento in cui scriviamo questo testo, secondo la compagnia americana Hugging Face, esistono 2.002.403 milioni di modelli linguistici avanzati di IA disponibili al pubblico[25].

Per prevenire abusi, questi sistemi sono progettati con misure di sicurezza come l’allineamento, ovvero un tipo di addestramento che aiuta i modelli a rispettare valori etici e sociali, inserendovi anche i così detti “guardrail”, come li ha chiamati il fondatore di OpenAI, Sam Altman, cioè sistemi di controllo che bloccano risposte pericolose, ad esempio, nel caso in cui si chieda a un LLM di descrivere il procedimento per costruire una bomba.

Queste protezioni però, possono essere aggirate forzando le versioni legittime a fornire risposte che normalmente verrebbero bloccate, e usarle per altri scopi. Se alcuni individui utilizzano versioni "senza filtri" di tali modelli, possono persino sviluppare LLM per specifiche attività criminali. Uno strumento popolare tra i criminali è ad esempio WhiteRabbitNeo[26], un

[24] Il jailbreaking indica una procedura che aggira o rimuove le restrizioni software imposte dai produttori di dispositivi digitali.

[25] https://huggingface.co/models

[26] https://www.whiterabbitneo.com/

modello che si presenta come assistente per esperti di sicurezza, ma che può essere facilmente usato per creare e-mail di phishing o software dannoso per attacchi informatici.

I modelli di Intelligenza Artificiale che operano senza i consueti filtri di sicurezza, detti anche *LLM non censurati*, permettono la generazione di contenuti pericolosi, sensibili o illegali in maniera diretta. Piattaforme come Ollama[27] consentono di scaricarli e utilizzarli direttamente sul proprio computer. Una volta installati, i modelli rispondono a richieste che sistemi del tipo di ChatGPT rifiuterebbero per motivi di sicurezza.

Inoltre, alcuni ricercatori hanno dimostrato che è possibile rimuovere i meccanismi di sicurezza anche dai modelli open source esistenti (Hartford, 2023), modificando i dati su cui sono stati addestrati. In questo modo, chi possiede le giuste competenze può costruire il proprio modello “senza filtri”. Alcuni di questi LLM "senza regole", in grado di generare testo, codice e contenuti, progettati appositamente per attività illegali sono rivenduti nei marketplace criminali online (Schulz, 2025). Uno degli esempi più emblematici è FraudGPT, un modello che offre una serie di funzionalità pensate per i criminali informatici, tra cui:

- Creare virus e malware
- Scrivere e-mail ingannevoli e pagine di phishing
- Trovare vulnerabilità nei siti web
- Generare credenziali false
- Automatizzare attacchi informatici
- Costruire bot e strumenti per rubare dati
- Offuscare codice per eludere i controlli di sicurezza
- Accedere a migliaia di esempi di codice dannoso.

I modelli di IA come FraudGPT rappresentano quindi una minaccia crescente, poiché permettono ai criminali di compiere attacchi sempre più sofisticati, automatizzati e di difficile rilevamento.

Tutto questo pone dei problemi etici che non riguardano soltanto l’impiego illegale dei modelli, ma gli stessi scopi per cui vengono realizzati. Secondo Eric Hartford (2023), ad esempio, i modelli non censurati sono necessari per indagare le capacità delle macchine e consentire a chiunque di avere la propria versione di IA allineata ai *propri* valori. Valori che non sono gli stessi a livello globale. Con buona pace di chi ritiene di poter allineare

[27] https://ollama.com/

l'addestramento delle IA ai propri valori etici, ponendo l'algoretica[28] alla base di tutte le scelte di sviluppo e applicazione concreta.

La conferma emerge anche da uno studio italiano realizzato dal Cefriel del Politecnico di Milano che parla di modelli "abliterated" (cancellati) cioè di versioni manipolate e riaddestrate di intelligenza artificiale generativa a cui sono stati rimossi i filtri di sicurezza per eliminare il così detto "refusal mechanism", cioè la capacità di questi sistemi di dire di no al loro utilizzatore. Questi modelli di *Dark AI*, secondo il ricercatore autore dello studio, Enrico Frumento, consentono la democratizzazione dell'uso malevolo dell'IA, la facilitazione della generazione di contenuti sintetici e la distorsione di contenuti condivisi, portando a una difficoltà crescente di discernere la verità, e ponendo una "sfida che non è solo legale ma antropologica" rispetto alla quale "non è più possibile contare solo su divieti e regolamenti, ma serve una risposta sistemica che combini educazione digitale, cooperazione internazionale e sviluppo di contromisure tecnologiche accessibili a tutti. La democratizzazione dell'IA deve andare di pari passo con quella della cybersicurezza" (Frumento, 2025).

Tuttavia, quando parliamo di Intelligenza Artificiale, non ci riferiamo soltanto ai Large Language Models (LLM), come ChatGPT, e al vasto insieme di algoritmi e di applicazioni che ne derivano - Gemini, Claude, Copilot, Bard, Grok, eccetera – ma anche a sistemi per il riconoscimento visuale, la traduzione simultanea, la ricerca farmacologica, il controllo della guida autonoma e allo sviluppo di altri sistemi esperti dedicati a compiti specifici. Anche questi sistemi possono essere hackerati.

Indurre in errore un sistema di IA usando una serie di trucchi per impedire il riconoscimento facciale o indirizzare una macchina a guida autonoma verso la direzione sbagliata è considerata da alcuni studiosi e attivisti una pratica di *resistenza algoritmica*, da altri un peculiare esempio di sperimentazione artistica (Hunger, 2019), ma il rischio che attori organizzati usino l'IA per commettere crimini, creare tensioni sociali e impiegarli nei conflitti è molto alto.

Come conferma di questa possibilità, il 5 giugno 2025 OpenAI ha rivelato in uno studio dettagliato di aver bannato una serie di account ChatGPT probabilmente gestiti da attori

[28] *Algoretica. Le due sfide cruciali nell'era tecnologica: bioetica, roboetica* (conferenza di Luigi Lombardi Vallauri del 18/5/2017, nell'ambito del ciclo di incontri *Roboetica. Dall'Algoritmo all'umanoide*, organizzato dall'Accademia Toscana di Scienze e Lettere La Colombaria, Firenze, maggio-giugno 2017; ora in *Atti e memorie dell'Accademia Toscana di Scienze e Lettere La Colombaria*, vol. LXXXII (nuova serie - LXVIII), Olschki, 2018 [ma 2017], pp. 353-376)

malevoli di lingua russa e da due gruppi di hacker di stato nazionali cinesi per supportare, tra le altre cose, lo sviluppo di malware, l'automazione delle interazioni sui social media e la ricerca sulle tecnologie di comunicazione satellitare degli Stati Uniti (Open AI, 2025).

Peggio ancora, adesso abbiamo l'evidenza che l'IA viene utilizzata per attacchi ransomware. FunkSec, un gruppo cybercriminale che utilizza l'Intelligenza Artificiale per colpire diversi settori, governativo, tecnologico, finanziario e dell'istruzione, in Europa e Asia, ha integrato la crittografia su larga scala e il trasferimento aggressivo dei dati in un "unico programma capace di disabilitare oltre 50 processi sui computer delle vittime ed è dotato di funzioni di auto-pulizia per eludere le difese". Nel proprio kit di attacco, sottolineano i ricercatori di Kaspersky che l'hanno individuato per primi, ha anche un generatore di password e uno strumento DDoS di base, ''entrambi con evidenti tracce di sintesi del codice tramite Large Language Models''. A differenza dei gruppi ransomware tradizionali, che richiedono riscatti milionari, FunkSec adotta un modello ''ad alta frequenza e basso costo'', puntando su riscatti relativamente bassi, anche di $10.000, abbinati alla vendita dei dati rubati a prezzi scontati a terze parti.

Ecco un elenco di esempi di strumenti di Intelligenza Artificiale utilizzati per la criminalità informatica (Visaggio, 2025)[29]:

- DeepFaceLab e Faceswap: creano video realistici generati dall'Intelligenza Artificiale per bypassare le procedure di verifica dell'identità su piattaforme bancarie e di criptovaluta.
- FraudGPT e WormGPT: questi generatori di testo basati sull'Intelligenza Artificiale aiutano i criminali informatici a creare e-mail di phishing efficaci, false comunicazioni aziendali e documenti legali fraudolenti. A differenza di ChatGPT, questi strumenti non hanno restrizioni etiche.
- BlackmailerV3: un toolkit per l'estorsione basato sull'Intelligenza Artificiale che automatizza e-mail di ricatto personalizzate e utilizza dati personali e aziendali acquisiti per aggiungere credibilità alle sue comunicazioni. Lo strumento è spesso utilizzato in truffe di sextortion, false minacce legali e tentativi di frode ai danni di CEO.
- Pagine di phishing generate dall'IA (EvilProxy, Robin Banks): queste piattaforme utilizzano l'IA per generare automaticamente siti web di phishing che imitano portali di accesso legittimi per servizi bancari, cloud e piattaforme aziendali.

[29] Aaron Visaggio, Linkedin, 2025.

- ElevenLabs e Voicemy.ai: gli aggressori sfruttano questi strumenti di sintesi vocale basati sull'IA per clonare le voci a scopo di vishing (phishing vocale), chiamate truffa deepfake e bypassare i sistemi di autenticazione vocale utilizzati negli istituti finanziari e nei controlli di accesso aziendali.
- Bot di ingegneria sociale basati sull'IA (Goose, bot per le frodi di Telegram): questi chatbot impersonano i rappresentanti dell'assistenza clienti e utilizzano conversazioni generate dall'IA per indurre le vittime a condividere informazioni sensibili, come i dettagli della carta di credito, i codici MFA e le password.

6.4 L'IA e la guerra algoritmica

La competizione tra l'Intelligenza Artificiale malevola e difensiva ha raggiunto un punto di svolta critico nel 2025, con gli LLM malevoli in grado di compromettere l'86% delle applicazioni da un lato e i sistemi difensivi che riducono i tempi di risposta da ore a minuti, dall'altro. Questa evoluzione in cui algoritmi combattono direttamente contro algoritmi in un'accelerata corsa tecnologica, rappresenta una trasformazione epocale nel campo della cybersicurezza. Anche se le difese algoritmiche dimostrano capacità di ridurre il carico di lavoro degli analisti di 250 volte, la ricerca documenta un incremento del 500-900% negli attacchi phishing AI-powered.

Anche il fenomeno della "battaglia algoritmica" non è più fantascienza, ma una realtà operativa che plasma il futuro della sicurezza informatica globale (ICT Security magazine, 2025). Ad esempio, il 9 luglio 2025 il Team Atlanta si è aggiudicato il primo posto nell'AI Cyber Challenge (AIxCC) della Defense Advanced Research Projects Agency (DARPA), aggiudicandosi un premio di 4 milioni di dollari per il suo sistema di ragionamento informatico basato sull'Intelligenza Artificiale (Darpa, 2025). La squadra ha superato altri sei finalisti nella competizione biennale volta a creare sistemi di Intelligenza Artificiale in grado di rilevare e correggere autonomamente le vulnerabilità di software open source, in particolare di quello utilizzato nelle infrastrutture critiche, tra cui sistemi finanziari, servizi pubblici e nell'ecosistema sanitario, come ha reso noto la DARPA che collabora da tempo con i leader del settore dell'Intelligenza Artificiale quali Anthropic, Google, Microsoft e OpenAI. La DARPA AI Cyber Challenge stessa rappresenta il primo esempio documentato su larga scala di "battaglia algoritmica" nel dominio della cybersicurezza. Le sette squadre finaliste, ciascuna con investimenti di 2 milioni di dollari, hanno sviluppato *Cyber Reasoning Systems* che hanno identificato 22 vulnerabilità sintetiche uniche e patchato (corretto) automaticamente 15 di esse. Il sistema ha anche scoperto un bug reale in SQLite3, responsabilmente divulgato, dimostrando efficacia oltre gli scenari di test controllati.

La battaglia algoritmica che punta a conquistare le menti è tuttavia più subdola. Quando parliamo di IA generativa dobbiamo considerare che gli LLM su cui si basano non sono mai strumenti neutri o infallibili: rispecchiano i dati e le logiche con cui sono stati costruiti, e possono essere esposti a manipolazioni sia durante la fase di addestramento sia attraverso l'interazione diretta. Che è quanto emerge da uno studio del SunLight Project (2024)[30], un think tank che analizza la disinformazione a livello mondiale.

Quindi, in sintesi, l'IA può essere usata per scrivere codice malevolo, individuare vulnerabilità, aggirare la protezione di sistemi informatici, fornire false informazioni, lasciar trapelare dati sensibili. Tutto questo accade nell'ambito della criminalità organizzata, ma anche durante i conflitti dove l'Intelligenza Artificiale diventa elemento dell'arsenale a disposizione della guerra grigia, quella non lineare, che gli opponenti combattono insieme a quella cinetica.

Come sostiene il generale Pasquale Preziosa: "*L'emergere dell'Intelligenza Artificiale quale forza abilitante in ambito militare sta trasformando radicalmente le dinamiche del conflitto. Sistemi di targeting automatizzato (tipo Lavender o Gospel[31]), riconoscimento di immagini, supporto al comando e controllo dipendono oggi da modelli di machine learning addestrati su grandi quantità di dati. Tali dipendenze generano vulnerabilità latenti, suscettibili di essere sfruttate attraverso tecniche di data poisoning, una forma sofisticata di manipolazione algoritmica capace di indurre errori sistematici o malfunzionamenti operativi nei sistemi avversari*." E aggiunge: "*Il data poisoning rappresenta oggi uno degli strumenti più avanzati di guerra informativa e cibernetica. Esso consiste nell'inserimento di dati corrotti, manipolati o "avversi" nei set di addestramento dei modelli IA. Le tecniche più comuni includono: il label flipping, ovvero la modifica delle etichette per provocare errori di classificazione, i backdoor attacks, cioè l'inserimento di trigger specifici che attivano comportamenti anomali solo in determinate condizioni e il poisoning graduale, che introduce distorsioni cumulative per eludere i sistemi di rilevamento*" (Preziosa, 2025).[32]

[30] American Sunlight Project (2024). NEW REPORT: Russian propaganda may be flooding AI models. ASP. Disponibile in: https://www.americansunlight.org/updates/new-report-russian-propaganda-may-be-flooding-ai-models

[31] Lavender, usato dagli israeliani a Gaza per il riconoscimento facciale, è noto avere una imprecisione del 10%. Il che potrebbe aver significato che dei 37.000 palestinesi uccise nei primi 4 giorni in quanto riconosciuti come terroristi, ben 3.700 potrebbero essere stati del tutto estranei (cfr. https://ilmanifesto.it/20-secondi-per-uccidere-lo-decide-la-macchina).

[32] Preziosa, P. (2025). Il silente avvelenamento dei pozzi digitali. Il nuovo sabotaggio cognitivo secondo Preziosa. Formiche. Disponibile in: https://formiche.net/2025/07/il-silente-avvelenamento-dei-pozzi-digitali-il-nuovo-sabotaggio-cognitivo-secondo-preziosa/#content [14 luglio 2025]

Il 13 novembre del 2025 Anthropic ha pubblicato un rapporto che spiega come gli hacker cinesi siano riusciti a effettuare il jailbreak dell'agente di intelligenza artificiale Claude Code di Anthropic per sferrare un attacco informatico automatizzato all'80-90% contro aziende tecnologiche e agenzie governative. Per la società americana fondata dai fratelli Amodei, si tratta probabilmente del "primo caso documentato di un attacco informatico di intelligenza artificiale su larga scala eseguito senza un sostanziale intervento umano". Per quanto senza precedenti, purtroppo è probabile che sia solo un'anteprima di ciò che accadrà[33].

Secondo l'azienda di cybersecurity giapponese Trend Micro, infatti, l'adozione dell'IA sta già cambiando radicalmente il panorama della cybersecurity, rendendo le minacce più autonome e scalabili. Per questo i ricercatori hanno coniato la definizione di *AI-Fication della cybersecurity*[34]. La previsione degli esperti è che per tutto il 2026:

a) L'IA diventerà una forza trasformativa e il vettore d'attacco principale, con minacce autonome e adattive.
b) L'IA agentica potrà eseguire operazioni complesse con crescente autonomia, trasformando agenti compromessi in vettori d'attacco.
c) Le campagne di inganno alimentate dall'AI, come deepfake e ingegneria sociale automatizzata, contribuiranno a erodere la fiducia degli operatori e a sovraccaricare le difese tradizionali.

In particolare, accadrà che i gruppi APT condivideranno risorse e infrastrutture, e l'AI sarà utilizzata per automatizzare operazioni complesse, rendendo più difficile la rilevazione delle attività malevole. Inoltre, aumenteranno i rischi per imprese e infrastrutture. Le aziende già oggi affrontano rischi significativi a causa di sistemi legacy e vulnerabilità operative proprio a causa dell'AI che automatizza phishing e ingegneria sociale, ma il trend previsto riguarda l'automatizzazione delle campagne di ransomware con un intervento umano minimo, consentendo anche a attori meno esperti di lanciare campagne complesse. Creando anche un serio problema geopolitico. Se, come previsto, l'IA faciliterà la scoperta automatizzata di nuove vulnerabilità, aumentando la velocità e l'efficacia degli attacchi, i ricercatori ritengono che emergeranno nuovi rischi da ambienti abilitati all'IA, inclusi attacchi di iniezione di prompt e backdoor nei modelli stessi, con il cloud, la supply chain, le librerie open-source e i dispositivi IoT/OT che forniranno punti d'ingresso per movimenti laterali orientati ad assumere il controllo in ciascun dominio organizzativo.

---

[33] Anthropic, (2025). Disrupting the first reported AI-orchestrated cyber espionage campaign. Disponibile in: https://www.anthropic.com/news/disrupting-AI-espionage?_bhlid=5ccdd567e96de7457f6c4c8c18651b8d1e6afe22 [13 novembre 2025]

[34] Trend Micro, (2025). The AI-fication of Cyberthreats: Trend Micro Security Predictions for 2026. Trend Micro. Disponibile in: https://www.trendmicro.com/vinfo/us/security/research-and-analysis/predictions/the-ai-fication-of-cyberthreats-trend-micro-security-predictions-for-2026 [25 novembre 2025]

6.5 L'IA prende di mira gli umani

È improvvisamente chiaro a tutti a cosa ci stavamo preparando connettendo oggetti, piante, animali: tutto è finalmente digitalizzato e l'intero ecosistema è diventato attaccabile, sfumando i confini tra civile e militare e tra guerra e pace
Barbara Carfagna

Il contesto politico recente è stato segnato soprattutto dell'Intelligenza Artificiale generativa usata durante le elezioni degli Stati Uniti, in cui la proliferazione di deepfake e vaste campagne di disinformazione hanno destato grandi preoccupazioni. Non poteva essere diversamente, il 2024 è stato proprio l'anno in cui l'Intelligenza Artificiale è stata adottata dalla maggior parte della popolazione, consciamente o inconsciamente: un fatto che ha prodotto un cambiamento epocale nel modo in cui vengono forniti i servizi e nel modo in cui la società si relaziona con la tecnologia.

Uno degli effetti di questa vasta diffusione di agenti artificiali è che i criminali informatici e i paramilitari cibernetici avranno a disposizione un sempre maggior numero di applicazioni di Intelligenza Artificiale per individuare vulnerabilità in modo rapido ed efficace, nonché per generare nuovi malware adattivi e individuare con maggior precisazione i nemici da annientare con droni e pallottole intelligenti (Mezza, 2024)[35]. Bersagli individuati da tecnologie di sorveglianza e repressione, come i sistemi Gospel e Lavender, il Project Nimbus, ma anche dagli spyware di NSO Group quali Pegasus o quelli di Cellebrite e Intellexa, oppure Graphite di Paragon Solutions, cioè tecnologie vendute e usate regolarmente da sistemi repressivi, democrature e dittature in tutto il mondo, anche da paesi democratici, ma sempre d'accordo coi governi che ne controllano la commercializzazione e a dispetto dei diritti umani di attivisti, giornalisti e politici regolarmente eletti (Loewenstein, 2024)[36]. Tecnologie che, come denunciato ripetutamente da The Guardian, diventano strumento di assassini mirati: più cose sappiamo di te, più sarà facile venirti a prendere la vita. É un cambiamento enorme rispetto alle tecniche di riconoscimento biometrico, tracciamento, geolocalizzazione e profilazione. Queste tecniche, che erano proprie del marketing e dei sistemi di raccomandazione, *più cose sappiamo di te, meglio possiamo offrirti quello che sei già propenso a desiderare*, si tratti di scegliere uno shampoo o un rappresentante politico applicando le regole della persuasione interattiva,

[35] Mezza, M. (2024), Connessi a morte. Guerra. media e democrazia nella società della cybersecurity. Milano: Donzelli Editore.

[36] Loewenstein, A. (2024), Laboratorio Palestina: Come Israele esporta la tecnologia dell'occupazione in tutto il mondo. Roma: Fazi Editore.

erano solo il primo passo che consentiva la modellazione esatta dei comportamenti sociali del target. Oggi, con l'IA, la granularità della conoscenza del soggetto ottenuta elaborando le sue tracce digitali, frutto dell'iperconnessione di cui ha scritto Michele Mezza in *Connessi a Morte* (2024)[37], e di una persuasione tecnologica costruita algoritmicamente nelle lunghe ore solitarie davanti agli schermi dei dispositivi digitali, fanno impallidire il marketing mirato e il dark advertising, fondendosi con strumenti letali: il marketing commerciale potenziato dall'IA diventa tutt'uno con la strategia militare. Scrive Barbara Carfagna sul quotidiano Il Sole 24 Ore: "Esattamente come noi veniamo persuasi da Facebook e TikTok, in base ai nostri gusti e abitudini, a comprare un blazer dalle spalle larghe o un viaggio alle Maldive, i combattenti dell'unità Radwan, intenzionati a liberarsi degli smartphone che li lasciavano identificare troppo facilmente, sono stati persuasi a comprare cercapersone, pager; non "dei" pager ma proprio "quei" pager manomessi dal Mossad nelle tappe della filiera produttiva. L'azione è partita anni prima, con un brainstorming dal design tutto aziendale: cartonati a grandezza naturale che ritraevano i combattenti nemici sono stati posti negli uffici del Mossad. Ogni mattina gli agenti entravano, guardavano, si facevano venire un'idea esaminando i dettagli, la scrivevano, e mettevano un bigliettino nella cesta dei suggerimenti: come si poteva colpire" (Carfagna, 2024).[38]

In particolare, la creazione di contenuti audiovisivi attraverso l'Intelligenza Artificiale generativa pone un serio problema di sicurezza economica, poiché i criminali informatici dediti alle frodi tramite pharming[39], adware[40] o phishing[41] sono in grado di rendere le loro campagne molto più credibili. Anthropic, nel *Threat Intelligence Report* del mese di agosto 2025[42] ha reso noto di avere intercettato e, successivamente bloccato, diverse incursioni criminali ingegnerizzate che usavano il proprio modello di IA generativa. Si tratta di tecniche di vibe-hacking, cioè di hacking emozionale, una forma di ingegneria sociale avanzata che manipola le emozioni delle persone per influenzarne le decisioni e i comportamenti analizzando le abitudini digitali delle vittime, per mostrare messaggi o contenuti particolareggiati proprio nei momenti di maggiore vulnerabilità con l'ausilio dell'IA (Mosca,

[37] Mezza, M. (2024). op. cit.

[38] Carfagna, B. (2024). La Guerra è diversa, rivoluzionata dalle nuove frontiere della tecnologia. Il Sole 24 Ore, 11 dicembre 2024

[39] Tecnica di attacco che mira a reindirizzare il traffico web di un utente verso un sito web falso e dannoso, con lo scopo di sottrarre informazioni personali.

[40] Un tipo di software che mostra annunci pubblicitari, spesso in modo intrusivo e fastidioso, sui dispositivi degli utenti. Può apparire come pop-up, banner o anche modifiche al browser web, e può raccogliere informazioni personali.

[41] Tecnica fraudolenta per ottenere informazioni personali e credenziali d'accesso a servizi digitali.

[42] Anthropic (2025). Detecting and countering misuse of AI: August 2025. Anthropic. Disponibile in: https://www.anthropic.com/news/detecting-countering-misuse-aug-2025 [27 agosto 2025]

2025)[43]. Usate soprattutto per mettere nel mirino aziende americane, tecniche similari sono state sfruttate a supporto di attività fraudolente di lavoratori nordcoreani utili per ottenere o conservare posti di lavoro presso grandi aziende statunitensi mediante Claude AI, l'assistente di IA di Anthropic, vestendo i panni di programmatori esperti e millantando false conoscenze linguistiche per superare colloqui di lavoro a distanza. Il denaro incamerato veniva poi dirottato verso il regime e le posizioni professionali acquisite, ma servivano anche per raccogliere informazioni e alimentare lo spionaggio industriale.

La propaganda statuale invece si avvantaggia da tempo di questi sistemi man mano che gli strumenti di Intelligenza Artificiale diventano più ampiamente disponibili. I ricercatori di IA hanno dimostrato nel tempo di poter creare tecnologie in grado di produrre audio e video falsi non rilevabili di alta qualità, la cui autenticità non può essere facilmente verificata, come i Deep fake video, video profondamente falsi, mai "girati" da nessuno. Con le tecnologie basate sull'Intelligenza Artificiale è infatti possibile mettere in bocca a un capo di stato parole che non ha mai pronunciato e innescare una crisi internazionale una volta che il messaggio sia diventato di dominio pubblico. In molti casi per riuscirci è sufficiente un sapiente uso di Youtube e Facebook prima, e dei media locali dopo, che non sono sempre in grado di verificare autenticità di fonti e protagonisti (Di Corinto, 2018)[44]. Gli esempi sono numerosi, dal video del presidente ucraino Volodymyr Zelensky che intimava agli ucraini di arrendersi ai russi (La Repubblica, 2022)[45] alla voce clonata dell'ex presidente americano Joe Biden che invitava i democratici a non esprimere il proprio voto nelle primarie del 2020 per conservarlo al momento del voto elettorale (Carboni, 2024)[46], dall'uso dell'IA per costruire immagini sintetiche dei presidenti Barack Obama e Donald Trump in manette, fino a deepfake che ritraevano figure istituzionali di primo piano come l'addetta stampa della Casa Bianca Karoline Leavitt e i senatori statunitensi Elizabeth Warren e Bernie Sanders nell'atto di promettere inesistenti sussidi governativi o agevolazioni economiche (Giaquinta, 2025)[47].

---

[43] Mosca, G. (2025). Che cos'è il vibe-hacking e perché l'intelligenza artificiale può renderlo più diffuso. La Repubblica. https://www.repubblica.it/tecnologia/2025/08/28/news/che_cose_e_il_vibe-hacking_e_perche_l_intelligen-za_artificiale_puo_renderlo_piu_diffuso-424812797/ [28 agosto 2025]

[44] Di Corinto, A. (2018). Le guerre del futuro si combatteranno nei nostri cuori. IL Manifesto. Disponibile in: https://ilmanifesto.it/le-guerre-del-futuro-si-combatteranno-nei-nostri-cuori [21 gennaio 2018]

[45] La Repubblica (2022). Il video deepfake di Zelensky che ordina agli ucraini di arrendersi. La Repubblica. Disponibile in: https://www.repubblica.it/tecnologia/dossier/tech/2022/03/17/video/il_video_deepfake_di_zelensky_che_ordina_agli_ucraini_di_arrendersi-423327466/ [17 marzo 2022]

[46] Carboni, K. (2024). Un deepfake del presidente Biden ha detto agli elettori di non votare. Wired Italia. Disponibile in: https://www.wired.it/article/deepfake-joe-biden-elezioni-stati-uniti/ [23 gennaio 2024]

[47] Giaquinta, F. (2025). Deepfake elettorali su Facebook: la frode politica è modello di business. Agenda Digitale. https://www.agendadigitale.eu/cultura-digitale/deepfake-elettorali-su-facebook-la-frode-politica-e-modello-di-business/ [23 ottobre 2025]

In diverse occasioni le operazioni di fact checking condotte da fondazioni indipendenti e singoli ricercatori hanno dimostrato che i contenuti falsi prodotti per inquinare il dibattito non mirano tanto ai singoli lettori o spettatori della rete quanto, come abbiamo detto, a inquinare i dati su cui sono addestrati i chatbot. È il caso della rete di influenza russa nota come *Pravda*, nata all'incirca all'inizio del conflitto russo-ucraino e responsabile della produzione di propaganda filorussa e anti-occidentale in momenti critici per le politiche di difesa europee (Andrea Zitelli, 2025). Come ricostruito da CheckFirst (2025), azienda finlandese impegnata a contrastare la disinformazione e le manipolazioni da parte di soggetti stranieri, e dal Digital Forensic Research Lab (DFRLab) dell'Atlantic Council, organizzazione di analisi della disinformazione incentrata su democrazia e futuro dei diritti digitali, nel momento in cui scriviamo, i contenuti diffusi dalla rete Pravda in tutto il mondo sono quasi 4,8 milioni (di cui 4,6 milioni riprodotti in più lingue).

La loro pubblicazione massiva e frequente ha fatto quindi ipotizzare l'utilizzo di un sistema altamente automatizzato che, molto probabilmente, si affida all'Intelligenza Artificiale per la rapida traduzione e la distribuzione del materiale in più domini e lingue. I picchi di attività sono stati rilevanti, curiosamente, durante le elezioni del Parlamento europeo del 2024 e in occasione del discorso del presidente ucraino Zelensky davanti al Parlamento francese il 7 giugno 2024. Poiché l'impatto appare limitato, visto l'esiguo numero di lettori di questi siti, secondo il watchdog Newsguard gli articoli servirebbero ad aumentare artificialmente la visibilità di specifici contenuti nei risultati dei motori di ricerca per "infettare" i chatbot di Intelligenza Artificiale (Mascitti, 2025) usati in Occidente per distribuire propaganda russa e disinformazione.

Dopo aver testato 10 dei principali chatbot IA, tra cui ChatGPT-4o di OpenAI, Grok di xAI, Meta AI, Gemini di Google, la stessa Newsguard ha scoperto che questi strumenti hanno ripetuto false narrazioni riciclate dalla rete Pravda nel 33% dei casi studiati. Questi risultati hanno confermato a loro volta quelli di un report realizzato dall'organizzazione no-profit statunitense American Sunlight Project (ASP), in cui si avvertiva che più che per generare traffico tramite i siti, la rete era stata probabilmente progettata proprio per manipolare i modelli di Intelligenza Artificiale utilizzati nei Paesi occidentali: "L'ASP stima che la rete produca almeno 3 milioni di articoli di propaganda filo-russa all'anno, un numero che non include l'attività della rete su X (Twitter), Telegram, Bluesky e VK. Considerata la crescita passata di questa rete, è probabile che questo tasso aumenti" (American Sunlight Project, 2024).

A dimostrazione di questi "nuovi" rischi, un'altra ricerca di Forensic AI, di tipo forense, ci fornisce una serie di elementi concreti sull'invasione dei contenuti basati sull'Intelligenza

Artificiale. Tra i dati principali dell'analisi appare che il 25% dei risultati di ricerca di TikTok contiene immagini sintetiche, che l'86% proviene da "Account Agentic AI", ovvero pipeline di contenuti completamente automatizzate; che questi account testano le preferenze degli algoritmi su larga scala, ottimizzando la viralità in base al volume. "Abbiamo superato i semplici deepfake. Sono state individuate distinte tassonomie e una è particolarmente significativa, quella relativa a narrazioni basate sull'Intelligenza Artificiale più immagini sintetiche/stock miste per storie clickbait"[48], si legge nella ricerca. Inoltre, sempre secondo l'analisi, i contenuti sintetici vengono condivisi 1,15 volte più frequentemente e gli account che li producono passano dalla creazione di audience basata su AI alla vendita di prodotti, senza essere etichettati come tali dalle piattaforme. La produzione di massa di contenuti artificiali ha successo grazie alla probabilità statistica di hit virali piuttosto che all'ottimizzazione della qualità. Per i nostri scopi va detto che molti di questi contenuti sono interviste sintetiche fotorealistiche presentate come giornalismo partecipativo da zone di conflitto. "La sofisticatezza dell'automazione suggerisce che stiamo entrando in una nuova fase in cui la creazione, la distribuzione e l'ottimizzazione dei contenuti operano senza intervento umano su larga scala" (AI Forensics, 2025).

Per contrastare questa tendenza nel novembre 2025, TikTok ha deciso, secondo quanto riportato da TechCrunch, di consentire agli utenti di scegliere la quantità di contenuti generati dall'IA visualizzata nei propri feed. Per migliorare la trasparenza, TikTok sta anche sperimentando una nuova forma di "watermarking invisibile", progettata per offrire un'identificazione più forte dei media generati dall'IA rispetto ad approcci attuali come le credenziali di contenuto di C2PA. Inoltre, l'azienda ha lanciato un fondo di alfabetizzazione sull'IA da due milioni di dollari per aiutare gli utenti a comprendere meglio la sicurezza, il rilevamento dell'IA e l'evoluzione del panorama dei media.

La **slopaganda**, dall'unione di *slop*, cioè "sbobba" e propaganda, come Klincewicz, Alfano e Fard (2025) hanno chiamato la produzione di contenuti indesiderati generati dall'intelligenza artificiale, diffusi al fine di manipolare le convinzioni individuali per fini politici, può avere anche un carattere commerciale, attirare visualizzazioni e aumentare i follower. Tuttavia, l'effetto principale è una forma di persuasione operata attraverso la saturazione sensoriale che non si basa su argomentazioni ma su simboli: "Come ha scritto il New Yorker, Trump è diventato 'l'imperatore dell'AI slop' (New Yorker, 2025). Infatti, l'account ufficiale della Casa Bianca lo ha già ritratto nelle vesti di Superman, di Papa o guerriero Jedi, e milioni di utenti hanno rilanciato queste immagini come segni di appartenenza più che di ironia. La

[48] La parola clickbait indica un contenuto online, spesso di scarsa qualità, ingegnerizzato per attrarre l'attenzione degli utenti e produrre "click". Si può tradurre come "esca".

ripetizione di immagini iperrealiste e al contempo surreali svolge una funzione iconografica più che informativa, dove il leader mette in scena sé stesso (Marcus Bösch parla di "ego staging") e così facendo rinsalda il legame affettivo con la sua comunità" (Sessa, 2025).

Il tema è diventato mondiale. La disinformazione prodotta anche con l'ausilio di contenuti artificiali può servire a promuovere la discriminazione verso le minoranze, a giustificare politiche sociali svantaggiose per i meno abbienti, oppure a legittimare una guerra e i suoi crimini, ma anche favorire il ricorso ad armi di distruzione di massa.

Ogni società ricorre infatti a specifiche narrazioni per fare progredire i gruppi sociali che la compongono verso mete utili alla collettività. Il senso e la direzione di queste narrazioni, politiche, sociali e religiose, cambia nei secoli, ed è in genere indifferente alla nozione di verità, concetto mobile e sfuggente per definizione. La creazione del consenso intorno a queste narrazioni si basa su storie condivise e la loro forza dipende da risposte psicologiche come la credulità, il conformismo, la reciprocità, innescate da meccanismi profondi che sfruttano il principio di autorità, di similarità, di credibilità e coerenza, i principi su cui si sono finora basate le principali relazioni umane (Cialdini, 1993). Questi principi possono, e sono, manipolati costantemente da specifici attori. E questo accade perché le narrazioni possono favorire sia la coesione che il conflitto sociale.

7. Attacco alla mente

> *According to the some authors the impact of disinformation can be split into the following areas: a) Spread (superficial online/offline behaviour towards dis/misinformation), b) Attitude change or reinforcement (e.g. the psychological effects of dis/misinformation on beliefs, cognition), c) Behaviour change (e.g. altering voting behaviour, disengagement from politics and d) Broader societal impact (e.g. reducing institutional trust, undermining social cohesion)"*
>
> *Mandić, J. & Klarić, D.*

L'informazione, quello che sappiamo del mondo, dell'ambiente circostante e del nostro stato, influenza le scelte che facciamo. Questo vuole dire che l'informazione gioca un ruolo centrale in ogni processo decisionale. Detto in maniera semplificata, il meccanismo psicologico è il seguente: l'individuo opera delle scelte sulla base delle informazioni in suo possesso, se queste informazioni sono corrette il comportamento sarà utile, adattivo ed efficace e come tale verrà rinforzato, avviando un nuovo ciclo di acquisizione delle informazioni, viceversa dovrebbe accadere il contrario. Eppure, gli esseri umani fanno scelte che rinforzano gli errori di giudizio basati su un contesto informativo "disordinato" che produce informazioni false o distorte. Queste informazioni false o distorte possono essere

tali *by default* o *by design*. In quest'ultimo caso c'è un agente che ha "ingegnerizzato" le informazioni false che sono quindi il frutto di una progettazione intenzionale. E le informazioni false by design sono uno strumento di attacco, l'attacco alla mente.

7.1 La Guerra cognitiva

We define cognitive warfare as the weaponization of public opinion, by an external entity, for the purpose of influencing public and governmental policy and destabilizing public institutions.
Bernal et. Al (2020)

La guerra è un conflitto aperto e dichiarato fra due o più stati o, in genere, fra gruppi organizzati, etnici, religiosi, sociali, ecc., nella sua forma più estrema e cruenta, quando cioè si sia fatto ricorso alle armi (Istituto della Enciclopedia Italiana, 1997). E proprio in relazione ai mezzi adoperati, al modo di combattere, alle forze impiegate nel combattimento, si hanno diverse aggettivazioni della guerra. Ad esempio, si parla di guerra elettronica in merito all'impiego di strumentazioni radioelettriche e all'uso di armi a guida elettronica. Pertanto, anche se l'espressione *guerra cognitiva* può non avere un significato univoco, riteniamo che sia possibile usarla nell'ambito della nostra analisi per indicare le intenzioni con cui viene pianificata ed eseguita, cioè causare un danno; per il bersaglio scelto, cioè i processi cognitivi; e per gli effetti che può determinare (van der Klaauw, 2023), ovverossia, l'alterazione della percezione pubblica e individuale degli eventi, causando paura, rabbia e frustrazione, sommosse, proteste e rivolte e, di conseguenza, instabilità sociale, politica, persino conflitti armati. Sono diversi gli autori secondo cui la capacità di sfruttare i limiti e le vulnerabilità della mente umana è oggi il nuovo campo di battaglia (Germani, 2017; Borghi, 2025).

Più precisamente, secondo gli studiosi militari, influenzare e manipolare la risposta della mente umana è diventato il terreno della moderna *competizione cognitiva*. Questo è quanto sostiene nella sua dottrina lo Stato Maggiore della Difesa italiano in un dettagliato rapporto intitolato "*Cognitive Warfare*. La competizione nella dimensione cognitiva" (2023). Secondo l'ammiraglio Cavo Dragone, ex capo di stato maggiore della Difesa che ne ha curato la prefazione: "La dimensione cognitiva è *parte integrante del 'campo di battaglia' nel quale operiamo quotidianamente, basata su armi contemporanee capaci di condizionare l'opinione comune, e, in ultima analisi, manipolare le decisioni, influenzando, interferendo e alterando le dinamiche cognitive ad ogni livello, nel quadro di strategie comunicative invasive e destabilizzanti*".

E aggiunge*: "In tale quadro, anche alla luce degli effetti della cosiddetta 'guerra ibrida' che si sviluppa secondo dinamiche sempre più letali, il Cognitive Warfare (competizione cognitiva) può essere definito come un'operazione multi dominio (o parte di essa) che impiega mezzi, azioni e strumenti attraverso le connessioni tra i domini classici (terrestre, aereo, navale), i domini spazio e cyber, l'ambiente informativo e lo spettro elettromagnetico, influenzando il comportamento umano e generando effetti nella dimensione cognitiva, con l'obiettivo di ottenere un vantaggio sull'avversario".*

La Cognitive Warfare*,* nella sua dimensione intangibile, spiega il documento, prende di mira ideologie, valori e società attraverso un uso sempre più esteso di mezzi di comunicazione e nuove soluzioni tecnologiche. Pertanto, quello che motiva la crescente attenzione militare ai settori della psicologia e delle neuroscienze all'applicazione delle tecnologie cognitive, è che nella mente umana confluiscono quegli elementi psicologici e neurologici, individuali e sociali, che danno senso al mondo e che sono implicati anche nei conflitti. È nella mente umana che si attua l'elaborazione di stimoli e informazioni mediante i processi di percezione, immaginazione, simbolizzazione, formazione di concetti, soluzione di problemi, che consentono l'elaborazione cognitiva del pensiero, del linguaggio, della memoria, e del controllo delle emozioni, basati sul funzionamento neurologico cerebrale.

Questi processi sono parzialmente prevedibili e manipolabili con le giuste tecniche, sfruttando, ad esempio, i meccanismi inconsci di risposta che determinano i *bias cognitivi*, cioè gli errori sistematici che portano le persone a ignorare informazioni razionali o logiche perché possono metterne a rischio identità e valori, privilegiando piuttosto informazioni che si allineano a sentimenti ed emozioni che risiedono nella parte inconscia per proteggere il proprio nucleo emotivo e identitario (Kahneman, 2012). Un fenomeno evidente nelle interazioni sui social media dove gli utenti online tendono a preferire informazioni aderenti alla propria visione del mondo, ignorando le informazioni dissenzienti e formando gruppi polarizzati attorno a narrazioni condivise che favoriscono la proliferazione di disinformazione (Cinelli et al. 2021).

La guerra cognitiva, insomma, è parte integrante della guerra dell'informazione.

Per comprendere come la guerra cognitiva sia da lungo tempo un asset nell'arsenale geostrategico di ogni esercito, si consideri al riguardo quanto, nel 2010, scrivono Mandić e Klarić (2023), citando l'ammiraglio in pensione Vladimir Pirumov. Pirumov, ex capo della Direzione russa per la guerra elettronica dello Stato maggiore della Marina, scrisse in un articolo sullo scontro informativo che "la guerra dell'informazione consiste nel garantire obiettivi nazionali sia in tempo di guerra che in tempo di pace attraverso mezzi e tecniche per influenzare le risorse informative della parte avversa... e include l'influenza sul sistema informativo e sulle condizioni psicologiche del nemico". Secondo il modello di Pirumov, le

tecniche di influenza dell'informazione includono "disinformazione (inganno), manipolazione (situazionale o sociale), propaganda (conversione, separazione, demoralizzazione, diserzione, prigionia), lobbying, controllo delle crisi e ricatto (Mandić & Klarić 2023).

Non è questo il contesto per analizzare in dettaglio i fondamenti psichici profondi e le ragioni adattative dei *bias*, frutto di preconcezioni legate alle euristiche cognitive del giudizio, della decisione e dell'attribuzione causale (Job e Rumiati, 1986; Rumiati, 1990). Euristiche che sono pure influenzate da altre strategie orientate all'economia cognitiva, necessarie per far fronte all'overload informativo che caratterizza le società dell'informazione. Per i nostri scopi, basti considerare che sono proprio questi bias che possono essere innescati attraverso la conoscenza del soggetto-bersaglio per effettuare operazioni di influenza o interferenza. Un assunto che è alla base del "controllo riflessivo del comportamento" teorizzato nel 1967 dal matematico Vladimir Lefebvre e praticato dai russi, ancora oggi, nelle loro operazioni di manipolazione informativa, e cioè, una serie di misure, costantemente in evoluzione, concepite per ingannare l'avversario e influenzare strategicamente la sua percezione dell'ambiente informativo. Nelle parole di Doroshenko & Lukito (2021): "Il compito principale del controllo riflessivo è individuare e sfruttare i punti deboli nella valutazione delle informazioni durante il processo decisionale. Le strategie di disinformazione russe non mirano solo a presentare falsità e confondere gli avversari. Piuttosto, l'obiettivo è diffondere disinformazione che induca gli avversari a prendere decisioni errate a favore della Russia, l'agente di controllo. Il kit sovietico degli strumenti delle strategie del controllo riflessivo odierno è stato descritto lungo quattro dimensioni: *dismiss*, *distort*, *distract*, and *dismay* (respingere, distorcere, distrarre e provare sgomento). L'atto di respingere (*dismiss*) presenta le prove in un modo che offusca gli obiettivi dell'agente di controllo o nega le prove presentate; la distorsione (*distort*) altera la percezione della realtà presentando varie falsità: da "fatti" inventati a caratteristiche di istituzioni e persone; la distrazione (*distract*) crea una minaccia reale o immaginaria o rivela nuove prove, che costringono un avversario a riconsiderare una decisione; lo sgomento (dismay) amplifica e intensifica drammaticamente la situazione per scoraggiare un avversario dall'intraprendere un'azione. Con l'avvento delle tecnologie digitali, il controllo riflessivo è stato adattato per la propaganda computazionale e la guerra informatica" (Doroshenko e Lukito, 2021).

Queste tecniche di controllo possono essere viste all'opera in diversi contesti bellici ma anche nei conflitti a bassa intensità e nella competizione geopolitica. Come vedremo, il "modello delle 4D" (*dismiss*, *distort*, *distract*, and *dismay*) è riconoscibile anche nella guerra di propaganda di Israele verso i suoi vicini.

Già nel 2011 uno dei principali teorici russi della guerra dell'informazione, Igor Panarin, riteneva che ci sarebbe stata una guerra mondiale nel settore della comunicazione tra il mondo atlantico e il mondo eurasiatico, convinto come era che anche le "Rivoluzioni colorate", quelle all'interno dei confini del vicino estero russo, ad esempio in Georgia e Ucraina, fossero un esempio di guerra ibrida (Colon, 2024). Ma l'idea di una competizione in rete, facendo uso di mezzi non militari per piegare la resistenza dell'avversario, generando dissenso all'interno della sua stessa società, era stata espressa ancora prima da Alexsandr Dugin, considerato l'ideologo di Vladimir Putin, il quale riteneva auspicabile l'aggregazione e manipolazione delle informazioni per poi farle divulgare verso reti di minoranze etniche, sessuali e religiose all'interno della società avversaria, usando Ong, think tank e agenti infiltrati.

L'idea di vincere la resistenza del nemico con mezzi non militari è tuttavia precedente alle riflessioni di questi eminenti teorici russi. Nel 1948 era stato un diplomatico americano, David Frost Kennan, a elaborarla. Diventerà la base della così detta *political warfare* (Rid, 2022). Nella disamina di Thomas Rid, la political warfare americana è la guerra di manipolazione dell'informazione che invece i russi chiameranno *dezinformatzija* per tutto il tempo della Guerra Fredda, fino all'elaborazione del concetto di *Active Measures* a ridosso della caduta del Muro di Berlino. Le Misure Attive, cioè l'insieme delle modalità di manipolazione del comportamento dell'avversario, una denominazione che, secondo alcuni documenti dell'intelligence bulgara, leale alleata dell'URSS, rimanda alla "sezione A" dello spionaggio sovietico incorporata dal KGB negli anni '60, hanno lo scopo di rendere la realtà "opaca", affinché le false informazioni prodotte possano allignare più facilmente laddove la credulità, la polemica politica e le divisioni sociali sono più forti.

L'idea di *Guerra cognitiva algoritmica* distingue invece l'approccio dell'Esercito Popolare Cinese ricostruito dallo Special Competitive Studies Project (SCSP), pensatoio creato da Eric Schmidt, e si riferisce esattamente a una guerra dell'informazione che usa gli algoritmi di intelligenza artificiale per influenzare le opinioni e i comportamenti delle popolazioni straniere (Lange, 2024).

Il framework della guerra cognitiva algoritmica comprende sia l'utilizzo dell'intelligenza artificiale per analizzare e tracciare i comportamenti individuali, che la potenza degli algoritmi di raccomandazione incorporati nelle piattaforme sociali che guidano il consumo dei contenuti digitali. Alla base della manipolazione algoritmica si posiziona l'incetta di dati (harvesting) per condurre ogni operazione di influenza cognitiva, sia che essi provengano dalla cessione volontaria di dati personali per ottenere sconti e servizi online, sia dal furto dei dati stessi attraverso azioni di hacking ad ampio spettro.

Anche gli studiosi cinesi indicano il dominio cognitivo come oggetto di competizione in quanto considerano la cognitive warfare come l'insieme di guerra psicologica, guerra dell'opinione pubblica, guerra commerciale e legale, guerra tecnologica e guerra del pensiero (Teti, 2024). L'obiettivo essendo l'uso di messaggi confezionati per interferire e controllare le credenze e le opinioni dell'avversario, la popolazione civile o la classe militare, sia in tempo di pace che in tempo di guerra. Da qui l'interesse verso i social media e le interazioni online, algoritmiche per definizione.

Come spiega Libby Lange nel paper Decoding China's AI powered 'Algorithmic Cognitive Warfare", analizzando l'articolo *An Exploration of Effective Mechanisms and Critical Technologies in Algorithmic Cognitive Warfare* pubblicato nel 2023 nella rivista Information Security and Communications Privacy, gli autori cinesi suddividono le operazioni cognitive in sei fasi: ritratto dell'utente, attrazione dell'attenzione, suggerimento di riferimento, induzione di reazione, intervento tempestivo e supervisione della gratificazione. In dettaglio:

1. **Ritratto dell'utente.** Gli algoritmi offrono un mezzo tecnologicamente avanzato per identificare il pubblico target su larga scala. A livello individuale, gli autori descrivono gli algoritmi come in grado di delineare lo stato psicologico di una persona; a livello sociale, li vedono come in grado di mappare "distribuzioni ideologiche sociali, tendenze psicologiche e strutture psicologiche". Più concretamente, gli algoritmi possono eseguire raggruppamenti e mappature delle relazioni su target specifici per segmentare il pubblico e identificare i collegamenti tra di essi. Una narrazione o un contenuto che può avere successo con un pubblico può essere interpretato in modo completamente diverso da un altro.
2. **Attrarre l'attenzione**. Gli autori individuano un ruolo per gli algoritmi sia nel monitoraggio del pubblico sia nel potenziamento dei modelli di intelligenza artificiale in grado di generare automaticamente contenuti che catturino la loro attenzione. Sostengono che la chiave per attirare l'attenzione non sia solo adattare i contenuti ad attributi "statici" come interessi o inclinazioni politiche, ma piuttosto allo stato psicologico e ai bisogni di un target in tempo reale. I modelli linguistici di grandi dimensioni (LLM) basati su algoritmi saranno in grado di creare rapidamente contenuti che rispondano a tali esigenze e di impegnarsi in una diffusione precisa per incoraggiare i target ad "esplorare" ulteriormente un problema. Questa fase avvia gli utenti su un percorso di coinvolgimento con nuove narrazioni, destabilizzando il loro attuale "quadro cognitivo".
3. **Riferimento suggerito**. Gli algoritmi sono in grado in modo unico di far apparire naturali l'aspetto e la viralità di determinate questioni, permettendo ai bersagli di credere di esprimere "giudizi indipendenti" a favore dell'avversario. Secondo gli

autori, gli algoritmi sono responsabili sia di "alimentare il fuoco" di una questione sia di "gestire" le discussioni nella direzione desiderata. Anche in questo caso, gli autori prevedono la necessità di enormi quantità di dati per mettere a punto l'algoritmo, in modo da analizzare e prevedere l'opinione pubblica in modo che il ruolo dell'algoritmo nel dibattito non venga rivelato.

4. **Indurre una reazione**. Gli esseri umani sono creature sociali e cercheranno naturalmente altri che condividano le loro convinzioni e i loro valori. Una volta che i bersagli si sono impegnati nel modo desiderato con le narrazioni suggerite, spinte come parte di un'operazione cognitiva, gli algoritmi promuovono il processo di frammentazione raggruppando gli individui in nuovi cluster, creando ciò che molti articoli definiscono "bozzoli informativi". In questo caso, gli autori considerano gli algoritmi in-piattaforma come chiave; menzionano specificamente diversi tipi di algoritmi di raccomandazione (che hanno ampiamente sostituito i sistemi di distribuzione cronologica dei contenuti) utilizzati sulle piattaforme di social media come capaci di superare la "soglia" necessaria per riorganizzare gli utenti.
5. **Intervento tempestivo**. Una volta che un'operazione cognitiva ha formato nuovi raggruppamenti, gli autori mettono in guardia dal permettere a questi gruppi di evolversi senza controllo. Lo scopo dell'"intervento tempestivo" è quello di utilizzare messaggi specifici per garantire che queste "cricche" si sviluppino nella direzione desiderata dagli operatori.
6. **Supervisione della gratificazione**. Infine, gli algoritmi possono aiutare a misurare la gratificazione sociale dei target una volta che sono entrati in nuovi raggruppamenti. Descrivono le persone come persone che traggono "estrema gratificazione" dalla supervisione sia degli altri che di sé stesse all'interno di un gruppo. In questa fase, esortano anche gli operatori cognitivi a testare la "scatola nera" di questi nuovi raggruppamenti sociali utilizzando messaggi specifici e monitorando i risultati del gruppo per creare un modello delle dinamiche interne di quel gruppo.

7.2 Tecnologie della persuasione

Quando parliamo di competizione cognitiva, parte della guerra dell’informazione ed elemento della guerra ibrida, non consideriamo soltanto la disinformazione, ma anche le dinamiche di modifica e manipolazione delle percezioni e dei comportamenti associati all’edutainment, al gaming, all’uso di tecnologie interattive di persuasione secondo il framework interpretativo sviluppato a partire dagli studi di B.J. Fogg già dal 1997. La teoria

della persuasione interattiva elaborata da Fogg, la Captologia (*Computers as Persuasive Technologies*), ritiene che i computer siano in grado di modificare i nostri comportamenti e indurci a fare quello che non faremmo di nostra spontanea iniziativa. Ad esempio, rendendo più facile l'esecuzione di un compito, permettendoci di accedere ad esperienze che ci modificano nel profondo, dandoci ricompense di carattere sociale. Secondo B.J. Fogg, i computer ci persuadono secondo la dinamica della triade funzionale, agendo come tools, media, attori sociali. Una bicicletta per il cardiofitness attraverso il cui display impariamo a controllare lo sforzo è un esempio di "tecnologia di riduzione" che, chiarendoci la relazione di causa ed effetto dell'esercizio, facilita un compito noioso favorendone la ripetizione. Oppure consideriamo le "tecnologie tunnel", come quelle che dopo schermate successive ci portano a fare gli acquisti in rete, cioè *tools* che inducono o facilitano comportamenti di consumo. Le tecnologie persuasive che operano come media includono invece tutte le tecnologie di simulazione: per smettere di avere paura di andare in aereo o per comprendere come si alteri la percezione alla guida dopo avere bevuto troppo, la tecnica è la stessa, si immerge il soggetto in un ambiente simulato che consente esperienze vicarie di realtà in sicurezza, magari all'uscita della discoteca o prima di un volo, che si trasformano in apprendimento. Ma queste tecnologie agiscono anche come attori sociali quando assumono caratteristiche animate e rivestono ruoli sociali, dando premi e punizioni: giochi che insegnano ai bambini a mangiare verdura o pupazzi interattivi che insegnano alle adolescenti una maternità responsabile (Di Corinto, 2009). Oppure a reclutare giovani terroristi (Carlin, 2018).

La Captologia, branca della Human-Computer Interaction che studia i computer come strumenti persuasivi, si occupa proprio di come tutti i dispositivi digitali riescano a persuaderci. La teoria indica la sovrapposizione tra l'arte della persuasione e la scienza dei computer a partire dall'analisi alle risposte sociali che gli individui danno a tv e nuovi media.

Secondo Fogg per persuadere qualcuno occorre che ci sia un'intenzionalità e dall'altra una disponibilità a farsi persuadere: gli antichi questo lo sapevano bene ed hanno inventato la *retorica* per strutturare discorsi convincenti mentre la *psicagogia* mirava a cambiare i comportamenti altrui attraverso la manipolazione dell'interlocutore.

Diceva il retore e sommo oratore Cicerone: "Non c'è nulla di più nobile che catturare l'attenzione delle persone con la parola, indirizzare le loro opinioni, distoglierle da ciò che riteniamo sbagliato e condurle verso ciò che apprezziamo. […] Cosa c'è di tanto grande e potente quanto il fatto che le emozioni del popolo, i dubbi dei giudici, il rigore delle istituzioni vengano modificati dal discorso di un singolo uomo? (Cicerone, 2007)

Molte delle tecniche persuasive del passato sono sopravvissute all'avvento delle soap opera, alla propaganda di regime e ai venditori porta a porta e fanno ancora parte del bagaglio

culturale di molti politici e influencer. Si distinguono dalla coercizione, perché le tecniche persuasive non ci obbligano, piuttosto ci seducono, ma hanno sempre lo stesso risultato, avviare, modificare o smettere un comportamento. Obbiettivo del persuasore è sapere come uno stimolo sollecita una risposta emotiva o razionale, per aggirare resistenze e diffidenze, oppure per presentare il vantaggio conseguente a intraprendere un certo comportamento. Così, mentre ci si ritiene avvertiti dei rischi portati dai persuasori umani, dal marketing e dalla pubblicità, gli individui non lo sono di fronte alle macchine informatiche. Il motivo è che i computer sono presunti neutrali, crediamo di controllarli, e si trovano ovunque nell'infosfera.

7.3 Hacking cerebrale

Mettiamo da parte per un momento la persuasione filtrata dall'apparato cognitivo che ha risultati né immediati né direttamente visibili, e consideriamo invece altre tecnologie che possono determinare risposte cognitive che pescano direttamente nella neurofisiologia. Oggi sono già realtà progetti per l'installazione cerebrale di microchip pensati per contenere gli effetti di malattie neurodegenerative, potenziare le percezioni, salvare i ricordi, amplificandoli o cancellandoli selettivamente. La risonanza magnetica funzionale può decodificare diversi tipi di segnali cerebrali, e domani forse potremo leggere i pensieri e influenzare stati mentali e comportamenti agendo direttamente sulla sfera neuropsicologica grazie alle interfacce mente-computer. Hackerare l'interfaccia cervello-computer è un fenomeno già riscontrato tra gli esperti di cybersecurity (Di Corinto, 2019).

In seguito allo sviluppo del progetto Neuralink tenuto a battesimo da Elon Musk, e che prevede la possibilità di impiantare elettrodi nel cervello umano per stimolarne la reazione cerebrale, si è potenzialmente entrati in una nuova fase di condizionamento diretto del comportamento umano. Per ora teorica, questa possibilità implica l'hacking cerebrale o *neuro hacking*, il cambiamento indotto a livello neuronale delle percezioni cognitive, e ci fa capire come la combinazione di biologia e informatica ci avvicini pericolosamente alla wet-war (Iezzi & Razzante, 2024). Le interfacce cervello-computer potrebbero un giorno essere hackerate per indurre falsi ricordi, manipolari i processi cognitivi, rubare informazioni che presumiamo segrete in quanto esistenti nel profondo della nostra mente. Una possibilità ancora più inquietante laddove l'interfaccia cervello computer utilizzi delle connessioni wi-fi. A quel punto l'integrità, disponibilità, confidenzialità di un dato, sia esso una password, il pin o la passphrase di accesso a dispositivi tecnologici, sarebbero a rischio e la loro compromissione potrebbe diventare l'avvio di una catena di attacco, passando dal neural

scanning (la scansione neurale) al neural flooding (la saturazione neuronale), fino al neural jamming (l'inceppamento neuronale) cioè la distorsione puntuale o massiva dei segnali elettrici cerebrali (Bernal et al. 2022).

Senza entrare in questo settore che appare ancora fantascientifico, ma non lo è davvero, per chiarire gli aspetti macroscopici della manipolazione informativa, ci riferiremo nel prosieguo del lavoro alla propaganda e alla disinformazione agite online per capire meglio come le scelte individuali, le dinamiche di gruppo, gli effetti ingroup-outgroup, leader-follower, siano sfruttati nella competizione cognitiva.

7.4 Propaganda e disinformazione

"Repeat a lie often enough, and it becomes the truth."
Joseph Goebbels

Insieme agli attacchi cibernetici, che usino oppure no l'Intelligenza Artificiale, la disinformazione è l'altra grande minaccia per la sovranità digitale. Nell'ambiente mediatico attuale, la disinformazione viene diffusa principalmente attraverso algoritmi di propaganda computazionale, i social network, le fabbriche di troll[49] e le personalità fake.

Non è semplice convergere su una definizione condivisa di cosa sia la propaganda computazionale, ma ogni concettualizzazione che la riguarda tende ad asseverarla come l'influenza esercitata sulla percezione degli individui facendo uso di algoritmi e strumenti cibernetici.

Una definizione operativa, che ha mostrato negli ultimi anni tutto il suo valore euristico, è quella di Wooley e Howard dell'Università di Oxford, secondo cui "La propaganda computazionale è l'uso di algoritmi, automazione e cura umana per distribuire intenzionalmente informazioni fuorvianti sui social media" (Wooley & Howard, 2017).

Propaganda è un termine antico il cui utilizzo iniziale può essere fatto risalire alla creazione dell'Istituto De Propaganda Fide ad opera dei vertici della Chiesa Cattolica Romana nel 1600. La propaganda della fede aveva come obiettivo l'evangelizzazione degli individui e la loro sottomissione al Dio dei cristiani. Come tutte le religioni, anche quella cattolica è basata su

[49] I troll sono soggetti che disturbano le conversazioni sui social con interventi provocatori. Possono essere automatizzati come bot che ripetono costantemente gli stessi messaggi.

narrazioni, e poiché il loro successo dipende generalmente dal numero di individui che cooperano in accordo con queste narrazioni, la loro propagazione è fondamentale.

Non ci è possibile ricostruire in maniera dettagliata come ogni ideologia religiosa sia fondata su miti, storie distorte o completamente inventate, per soddisfare il bisogno umano di spiritualità (Harari, 2024) e del ruolo, anche positivo, che hanno avuto nell'evoluzione della società umana, e per questo rimandiamo agli studi etno-antropologici di uno dei massimi storici italiani della religione, il professore Ambrogio Donini (Donini, 1991). Ci limiteremo pertanto a riferirci all'uso della propaganda in epoca moderna.

In tempi recenti, il suo teorico storicamente più influente, Edward L. Bernays, ha descritto la *propaganda* come l'insieme delle azioni necessarie a guidare le masse, per il loro bene (Bernays, 1928; 2020). Bernays, nipote di Sigmund Freud, e teorico della fabbricazione del consenso, era convinto che l'uomo della strada non avesse opinioni affidabili e che potesse votare per la persona sbagliata o desiderare la cosa sbagliata; quindi, riteneva che dovesse essere guidato dalla propaganda a fare le scelte giuste.

Erano gli anni ruggenti del capitalismo e il nascente consumismo non aveva ancora incontrato la Grande Depressione dei successivi anni '30. Ma si presentò presto un'occasione per guidare le masse e convincere le persone a fare quello che non avrebbero fatto di propria iniziativa: arruolarsi e partecipare alla Seconda guerra mondiale. Già all'epoca il termine propaganda divenne sinonimo di propaganda politica.

Vista come qualcosa di negativo, capace di influenzare il libero arbitrio delle persone, ma anche di facilitarne la coesione, la propaganda veicolata dai mass media è stata massicciamente usata nel secolo scorso dai regimi dittatoriali più noti - fascismo, franchismo, nazismo e comunismo -, come pure da governi democraticamente eletti e con una forte impronta ideologica, si pensi al Sionismo in Israele e al Peronismo in Argentina. Scopi, metodi e tecniche usate sono state simili e diverse, ma ancora oggi il concetto moderno di propaganda ci rimanda alla disseminazione di idee e informazioni che hanno lo scopo di indurre alcuni specifici tipi di scelte o di azioni in ambito sociale e politico. Questa idea di propaganda, canalizzata principalmente attraverso la radio e le pubblicazioni a stampa, fu al centro dell'attività instancabile e brutale del gerarca nazista Joseph Goebbels, che tramutò le tesi espresse da Bernays in Crystallizing Public Opinion del 1923 nel suo manuale di riferimento per organizzare le campagne elettorali naziste. Per Goebbels la propaganda era l'arte di guidare la nazione. Situata al centro di tutti i discorsi e dell'attività politica, quale

organo di collegamento tra il governo e il popolo, aveva la funzione di promuovere l'adesione attiva del *volk* tedesco alle politiche del Reich, e veniva praticata secondo la massima a lui attribuita, "qualsiasi cosa può essere fatta credere come vera se la si ripete per un numero sufficiente di volte". La propaganda di Goebbels era indifferente al concetto di verità e doveva essere rispettata solo se serviva gli scopi dello stato nazista, la supremazia razziale del popolo tedesco (Mari, 2025).

É possibile che anche per i duraturi effetti sull'immaginario prodotti dall'uso che ne fece Goebbels, la propaganda, che secondo Bernays doveva essere basata su fatti e informazioni accurate, sia stata spesso confusa con la disinformazione, che risulta invece da un miscuglio di elementi veri ed elementi falsi. La propaganda, esplicita, organizzata da attori noti, per creare consenso intorno a un'idea o a un bene da promuovere, è però diversa dalla disinformazione, che invece può essere concettualizzata come il tentativo occulto di manipolare l'informazione per fuorviare il ricevente di una comunicazione. E proprio questo era l'obiettivo della polizia segreta sovietica (GPU, *Gosudarstvennoe političeskoe upravlenie*), il cui capo del controspionaggio, Artur Artuzov, istituì un ufficio di "*dezinformatzija*", (дезинформация), per la propagazione della disinformazione intesa come "arma tattica" finalizzandola alle operazioni di intelligence all'estero e che fu successivamente chiamata così per volere di Josif Stalin, presidente dell'URSS (D. Colon, pp.196-197).

La parziale sovrapposizione dei due concetti dipende, secondo noi, dal fatto che propaganda e disinformazione servono allo stesso scopo, che è quello di usare l'informazione per influenzare le percezioni del ricevente allo scopo di modellarne il comportamento.

Con una differenza non sempre facile da individuare: se la persuasione applicata alla propaganda può essere definita come l'innesco di un comportamento non spontaneo, facendo però leva sul ragionamento e gli appelli emotivi, la disinformazione si basa sulla sovversione delle informazioni che gli individui, supposti razionali, usano per agire le loro scelte, anche a scapito dei propri interessi.

Nella sistematizzazione teorica operata dal sociologo Dennis McQuail (1997), il quale chiarisce la difficoltà di affrontare il concetto in termini neutrali e generali, sono i seguenti elementi chiave che descrivono la propaganda come pratica comunicativa: la menzogna; la censura e la negazione dell'informazione, oppure la sua selezione in un'ottica strategica; l'esagerazione; gli appelli affettivi e all'emozione (volti a suscitare desiderio oppure a instillare paura); il ricorso a una retorica linguistica o a una narrazione visuale che sollecitano

direttamente o, comunque, privilegiano gli aspetti non razionali della comunicazione. E tuttavia non contiene necessariamente menzogne e falsificazioni, ma presenta sovente una partigianeria unilaterale.

La stessa Unione Europea, mentre considera legittima la propaganda, ha avviato una serie di azioni per contrastare la disinformazione, che oggi viaggia velocemente in rete, in quanto: "informazioni altamente persuasive o fuorvianti create, presentate e diffuse per un guadagno economico o per ingannare intenzionalmente il pubblico, possono causare un danno pubblico. Il danno pubblico include minacce al processo politico democratico e al processo decisionale, nonché al bene pubblico, come la tutela della salute dei cittadini dell'UE, dell'ambiente o della sicurezza" (Europarlamento, 2021).

Le campagne di manipolazione delle percezioni che oggi usano propaganda e disinformazione per seminare dubbio e scontento nella popolazione vengono infatti diffusamente distribuite sui social network principali, cioè, nel mondo occidentale, Facebook, X, Instagram, Truth, TikTok, ambienti ingegnerizzati per favorire il coinvolgimento delle persone e la polarizzazione delle opinioni. Possiamo dunque affermare che propaganda e disinformazione sono diventate oggi un problema cibernetico perché si manifestano attraverso la Rete e i suoi attori usano strumenti digitali automatizzati e interattivi per colpire le certezze dei propri bersagli con un esercito di troll e di bot[50], facendo largo uso di meme[51] e notizie fasulle, create ad arte da gruppi di guerriglia digitale che usano anche tecniche di software hacking[52] per manipolare l'informazione e i suoi protagonisti (Morgan, 2018), i cui contenuti viaggiano in misura consistente anche su forum di discussione come Reddit, Discord, e 4chan. E, come sostengono Sara Bentivegna e Giovanni Boccia Artieri (2021), hanno finito per modificare il ciclo della notizia e dell'informazione attraverso una serie di rimandi e interazioni tra le logiche degli old e new media producendo un peculiare disordine informativo.

7.5 La propaganda computazionale e gli eserciti di troll

La propaganda computazionale può essere quindi concettualizzata come veicolo di propaganda e disinformazione da parte di attori singoli e associati, volontari e mercenari, dilettanti e professionisti, che usano le piattaforme digitali, i motori di ricerca, i social

[50] I bot sono software programmati per sostituire l'intervento umano e svolgere compiti di raccolta, analisi, catalogazione. I bot in grado di fare conversazione sono detti chatbots.

[51] I meme sono unità minime di informazione che si auto-propagano grazie alla loro semplicità. Si tratta spesso di immagini e slogan ad effetto, facili da comprendere e memorizzare.

[52] Hacking è l'insieme dei metodi, tecniche e operazioni volte a conoscere, accedere e modificare un sistema informatico hardware o software.

network, i social media e, da ultimo, i Large Language Models, per diffondere fake news, narrative distorte e contenuti persuasivi. Questi attori, lo ribadiamo, fanno un ampio uso di bot, troll, meme e tecniche di hacking per influenzare la percezione del ricevente ed elicitare, in chi vi è esposto, una reazione che sia in accordo con gli interessi dell'attore. Le leve sono quelle della retorica politica, della propaganda mediatica e della *dezinformatzija*: sfruttare i limiti attenzionali e le euristiche di scelta collegate alla selezione dell'informazione attraverso la diffusione personalizzata di messaggi emotivamente allarmanti, minacce reali o immaginarie e teorie del complotto che catturano l'attenzione più di altri contenuti (Brady et al. 2020).

Se l'ambiente digitale è il luogo di elezione della propaganda computazionale, i suoi effetti, tuttavia, vengono amplificati dai media tradizionali - radio, tv, stampa -, che gli fanno eco e la posizionano all'interno dell'agenda mediatica, legittimandola (Bentivegna & Artieri, 2021). La propaganda rilanciata dai media tradizionali viene successivamente sfruttata dagli attori della propaganda stessa che la rimbalzano nei circuiti mediatici in un loop ricorsivo che è spesso all'origine di comportamenti complottisti (Bianchi, 2022). I media diventano "armi mediatiche" per la disinformazione, proprio come ha candidamente ammesso l'ex ministro russo della Difesa Sergei Shoigu, che ha apertamente definito la militarizzazione dei mass media parte di una guerra dell'informazione che ha una strategia ben sviluppata volta a "influenzare, sconvolgere o corrompere l'opinione pubblica" (Perry, 2015).

Negli ultimi anni abbiamo visto all'opera soggetti organizzati gestire vaste attività di propaganda e disinformazione: si pensi alle fake news legate alla campagna presidenziale di Trump nel 2016 e al suo primo mandato[53]; all'uso di dark ads[54] e al microtargeting[55] generato dall'uso di app psicometriche nel caso di Cambridge Analytica (Wilye, 2019), oppure alle notizie fasulle su cause ed effetti della Brexit (Harding, 2017), fino al dilagare delle false narrative della "fabbrica dei troll" di San Pietroburgo[56].

[53] I fact-checker del Washington Post hanno documentato 30.573 affermazioni false o fuorvianti durante il suo primo mandato presidenziale, una media di 21 al giorno.
https://www.washingtonpost.com/graphics/politics/trump-claims-database/

[54] I dark ads sono post sponsorizzati dagli inserzionisti verso specifiche porzioni di popolazione, individuate secondo dei parametri selezionabili dalle piattaforme di pubblicazione come i social network.

[55] Il microtargeting è l'utilizzo dei dati di profilazione per personalizzare messaggi pubblicitari verso singoli individui, in base all'identificazione delle vulnerabilità personali dei destinatari. Viene utilizzato per promuovere beni, merci, servizi.

[56] La "Fabbrica dei troll" russi, è l'epiteto giornalistico dell'Internet Research Agency, struttura finanziata da Evgenij Viktorovič Prigožin, imprenditore, politico e comandante mercenario russo, amico del presidente russo Vladimir Putin, con il compito di sviluppare contenuti di propaganda a favore del governo di Mosca.

In tempi recenti il caso forse più famoso di inquinamento dell'informazione online è stato quello noto come "Pizzagate" in cui è rimasta coinvolta la democratica Hillary Clinton che nella campagna elettorale USA del 2016 contendeva al repubblicano Donald Trump lo scranno della Casa Bianca. La fake news, che all'epoca ottenne sui social più visibilità delle smentite giornalistiche, la dipingeva a capo di un'organizzazione satanista che aveva creato un circuito di abusi pedofili nello scantinato di una pizzeria. Le indagini successive hanno dimostrato che non solo non esisteva il circolo pedofilo, ma neanche lo scantinato dove sarebbero avvenuti gli abusi (Harari, 2018; Bianchi, 2021).

Se le fake news sono state uno strumento di competizione elettorale, anche i "bot" lo sono stati a più riprese. Su Facebook, Twitter, Instagram, molti profili fasulli sono governati da bot in grado di intavolare una banale discussione in chat (i chatbots), e che producono una notevole mole di messaggi. Spesso si tratta di esche sessuali o di truffatori che offrono denaro in prestito o altri servizi a pagamento, ma tra questi primeggiano i "political bots", in conseguenza del fatto che sono le organizzazioni politiche quelle più propense a investire fondi consistenti nella propaganda computazionale (Di Corinto, 2023).

Ad esempio, nel 2017, una serie di inchieste giornalistiche (Rodrigues & Goodman, 2017) rivelò che due attiviste laburiste avevano commissionato la creazione di un bot, cioè di un sistema automatico di risposta, per Tinder, nota app di incontri sentimentali, con lo scopo di favorire l'agenda politica del proprio partito. Programmato per specifiche fasce di età e di interessi, il suo scopo era quello di suggerire il voto per i laburisti a ogni potenziale "anima gemella" incontrata. Un altro esempio è quello dei bot che hanno rilanciato migliaia di volte l'hashtag #ReasonsToLeaveEu durante il referendum per la Brexit. Questi "amplification bots", secondo i ricercatori italiani della Fondazione Bruno Kessler, sono stati usati anche durante le elezioni politiche italiane del 2018 per "dopare" la diffusione dei messaggi del partito della Lega e del suo leader, Matteo Salvini (Bachini & Tesconi, 2020).

Similmente, gli attacchi nei confronti del Presidente della Repubblica Italiana Sergio Mattarella nel 2018 sono riconducibili alla stessa logica. All'epoca, la guerra di hashtag sulla formazione del nuovo governo italiano ha mostrato un web diviso tra i sostenitori del presidente Mattarella e i suoi detrattori con due opposti hashtag: da una parte l'hashtag #IoStoConMattarella, usato da chi si è schierato a difesa delle Istituzioni incarnate dal capo dello Stato, e dall'altra quello di chi ne ha chiesto finanche l'impeachment, #IlMioVotoConta. La funzione dell'hashtag è infatti proprio quella di aggregare e categorizzare i contenuti presenti sulle piattaforme sociali in relazione al tema trattato e rendere quindi più facile agli

utenti individuare contenuti specifici senza perdersi. Nel caso del dibattito della formazione del nuovo governo italiano, l'hashtag è diventato la bandiera di opposte fazioni: entrambi usati per essere visibili nel flusso della comunicazione di un evento che ha indotto molti a prendere posizione a favore o contro per far pesare la propria opinione. Quelli che chiedevano l'impeachment del Presidente però provenivano da 360 account creati ad hoc lo stesso giorno della contestazione sull'allora Twitter, oggi X.

Non ci sono solo i bot. A volte sono le persone in carne e ossa che approntano messaggi, li automatizzano per pubblicarli a una certa ora e con una certa frequenza e, in una continua azione di propaganda, riempiono i social di informazioni e commenti destinati a sostenere il proprio beniamino. È il caso di Daniel John Sobiesky, un fanatico sostenitore di Trump scovato dal Washington Post che ne ha raccontato la storia (Timberg, 2017). Viola Bachini e Maurizio Tesconi nel loro libro *Fake People. Storie di social bot e bugiardi digitali*, li chiamano "cyborg".

8. Le interferenze hacker e la guerra algoritmica

8.1 Dal sabotaggio culturale alle guerre guerreggiate, come cambiano gli hacker

Amateurs hack systems; professionals hack people.
Bruce Schneier

Disinformazione e attacchi informatici ormai si manifestano sempre più spesso in maniera congiunta. Sono l'opera di hacker con la capacità di manipolare software, dati e informazioni all'interno di reti, sistemi informativi e servizi informatici, rappresentando una minaccia esiziale per la sovranità digitale.

Tutto questo non è incominciato ieri. Da diversi anni hacker criminali e hacktivisti sono stati "reclutati" per mettere a segno attacchi informatici, azioni di spionaggio e sabotaggio per conto di gruppi di interesse, fazioni politiche, e stati nazione. Gli stessi hacker di stato, che spesso coincidono con *gruppi APT*, Advanced Persistent Threat, e prendono il nome proprio dalla tecnica usata[57], collaborano con altri individui e gruppi, politicamente o economicamente motivati, rispettivamente hacktivisti e cybercriminali, che fanno uso di tecniche di hacking per perseguire i propri scopi. Gli hacktivisti però, coinvolti nei conflitti

[57] APT, Advanced Persistent Threat, minaccia consistente in un attacco mirato, volto ad installare una serie di malware all'interno delle reti bersaglio, al fine di riuscire a mantenere attivi i canali impiegati per l'esfiltrazione di informazioni dalle infrastrutture IT del target. É una tecnica peculiare degli hacker di stato finanziati dai governi.

aperti, dall'Ucraina a Israele, dall'India al Pakistan, dal Sudan all'Iran, adottano le tecniche che prima erano del sabotaggio culturale (Di Corinto, 2002) per far avanzare la propria agenda politica. Questi hacker, indipendentemente dal loro livello organizzativo e di comando, sono stati coinvolti a più riprese in operazioni di "hack and leak" (hackera e fai trapelare), "steal and publish" (ruba e pubblica), con l'obiettivo di creare confusione, panico e paranoia nel pubblico (Rid, 2022).

Secondo la società di cyber threat intelligence Graphika, nel 2020 diversi gruppi di hacker cinesi sono stati coinvolti in operazioni di disinformazione verso il governo americano, il presidente Joe Biden e i manifestanti di Hong Kong (Di Corinto, 2020a); paesi come l'Iran invece hanno agito attraverso dei proxy informatici, hacktivisti e ransomware gangs, per sostenere la causa arabo-palestinese o per attaccare il governo americano (Di Corinto, 2020ab); la Corea del Nord lo ha invece fatto per inquinare le prove delle incursioni dei propri hacker di stato (Di Corinto, 2021); i servizi segreti russi per legittimare la causa dell'annessione della Crimea alla Federazione Russa, vestendo i panni di Anonymous[58].

La cifra comune di azioni tanto diverse è proprio che ogni attacco informatico si svolge parallelamente a un'azione di disinformazione per negare l'accaduto, o per amplificarne la portata, ad esempio su X e nei canali Telegram, soprattutto laddove i risultati si siano rivelati modesti. Quest'ultimo è il caso degli attacchi DDoS portati da gruppi filorussi come Killnet (Di Corinto & Rociola, 2022), e di quelli di hacktivisti, con vocazione religiosa, quali Anonymous Sudan e Mysterious Team Bangladesh che hanno colpito anche l'Italia a più riprese e, in particolare, i siti degli aeroporti di Napoli, Calabria, Puglia e Valle d'Aosta.

Nonostante le difficoltà di attribuzione dell'origine e degli attori di questi attacchi, procedura lunga, complessa e dispendiosa, sono numerosi i paesi che fanno ricorso agli hacker per sviluppare tool, strategie e azioni di propaganda e disinformazione (Industrial Cyber, 2023) ed è assai probabile che anche i paesi democratici e liberali lo facciano. La dimostrazione verrebbe da un'indagine di Wikileaks a proposito del *Vault7*, un programma di sorveglianza gestito attraverso un 'arsenale' di malware e di cyber weapons**,** con cui la CIA sarebbe stata in grado di controllare le comunicazioni di aziende, cittadini e istituzioni, introducendosi in apparecchi di uso quotidiano come i telefoni Apple, Google, Microsoft, e perfino i televisori Samsung, utilizzandoli come captatori informatici (Di Corinto, 2017). Un'altra prova riguarda invece il furto, confermato, di una serie di exploit per vulnerabilità informatiche mai dichiarate, sottratte dagli ShadowBrokers alla sezione della National Security Agency

[58] Sigla collettiva di un movimento globale di hacker attivisti

denominata Tailored Access Operation (TAO), l'hacking unit dell'agenzia americana (Di Corinto, 2018) ed usate da altri soggetti per costruire nuovi armi informatiche come il ransomware Wannacry.

Tuttavia, secondo diversi analisti, il modo di operare dei russi e dei filorussi è emblematico dell'uso che ne fanno in concomitanza con gli attacchi informatici veri e propri.

Secondo Treyger et al. (2022), che hanno studiato gli sforzi della disinformazione russa sui social network, "*Alcune delle attività russe che si svolgono sui o attraverso i social media non sono pura disinformazione; si tratta piuttosto di sforzi di disinformazione collegati funzionalmente a un attacco informatico di qualche tipo. Pertanto, anche se ci teniamo in gran parte lontani dalla discussione tecnica sugli attacchi informatici, tocchiamo le operazioni informatiche quando queste sono strettamente legate ad attività che utilizzano l'informazione per modellare percezioni o comportamenti, ad esempio, hack che producono informazioni che vengono successivamente trapelate*".

Questo passaggio rappresenta bene il pensiero strategico dei russi rispetto al conflitto informativo che integra due aspetti: quello tecnico-informatico, che mira a colpire "i sistemi tecnici che ricevono, raccolgono, elaborano e trasmettono informazioni", e quello informativo-psicologico, che "mira a influenzare il personale delle forze armate e la popolazione".

Una logica descritta da Calise & Musella (2019), che scrivono: "*Ma la vera novità del conflitto 2.0 è la sua penetrazione a livello di massa, con iniziative di propaganda o di campagna psicologica volte a influenzare quanto i cittadini sanno di sé e degli altri. In questo caso gli attacchi digitali non sono destinati a bersagli di tipo militare o infrastrutturale. Siamo invece in presenza di azioni mirate a condizionare il clima politico in un altro paese, o a mettere a repentaglio procedure di cruciale rilevanza come le elezioni. Una minaccia che preoccupa le democrazie occidentali, perché va al cuore stesso del loro sistema operativo: l'autonomia dell'opinione pubblica. E si gioca sulle piattaforme che connettono centinaia di milioni di cittadini*".

Gli autori militari hanno identificato le seguenti caratteristiche che raccomandano i social media come arma informativa:

- il basso costo delle operazioni sui social media sia in termini di fondi che di personale;

- l'ampia portata potenziale delle operazioni di informazione online, soprattutto considerando la crescente penetrazione di Internet;
- la capacità di reagire in tempo reale e in luoghi diversi senza presenza fisica;
- la negabilità delle operazioni sui social media, data la difficoltà nel distinguere l'attività ordinaria dagli atti di guerra dell'informazione sponsorizzati dallo stato;
- la percezione che gli effetti psicologici dei media online e dei social media siano superiori a quelli forniti dai media tradizionali a causa del potenziale di confezionare contenuti multimediali in modo da ottenere "ulteriore influenza emotiva e psicologica".

Ancora una volta, un esempio di scuola è quello che è successo negli Stati Uniti con la campagna che ha portato Donald Trump alla Casa Bianca nel 2016. In quell'occasione, gli apparati di intelligence e importanti segmenti della politica statunitense, ricollegandosi al filone di indagine giudiziaria che ha preso il nome di Russiagate, hanno accusato Mosca di una intensa e duratura manipolazione delle informazioni al fine di favorire, all'epoca, l'attuale presidente americano, di nuovo in carica dal 2024, Donald Trump. In questa occasione, Renée Di Resta, capo di una delle due agenzie di cybersicurezza incaricata dal Senato americano di studiare i meccanismi di influenza russa, parlò di "guerra mondiale dell'informazione" (Calise & Musella, 2019).

Per meglio comprendere il fenomeno, ci appare utile elencare le numerose azioni di interferenza documentate di hacker russi nei processi democratici dei paesi occidentali. Dall'inizio della Guerra in Ucraina queste interferenze si sono moltiplicate, ma già prima ne abbiamo avuto numerosi esempi. Nel **2007** un vasto attacco DDoS viene compiuto in Estonia come ritorsione per lo spostamento della statua del soldato sovietico dal centro alla periferia di Tallinn, la capitale; nel **2008,** in Georgia, quando all'attacco ai siti web georgiani si affianca una campagna militare vera e propria; nel **2014**, quando il servizio segreto militare russo Gru crea un video falso di Anonymous per sostenere l'invasione dell'Ucraina; infine nel **2022**, quando l'invasione del Donbass ucraino viene accompagnata da una serie di attacchi informatici: DDoS, defacciamenti e distribuzione di virus wiper che cancellano i registri di memoria dei computer Windows. Il Digital Forensic Research Lab del Consiglio Atlantico, pochi giorni prima (Digital Forensics Lab, 2022), aveva segnalato una serie di false narrative distribuite sui social media e propagandate da giornali e televisioni pro-Cremlino. Intanto però, prima dell'ingresso delle truppe russe in Ucraina, e di successivi attacchi cibernetici di carattere massivo, i modem della rete Internet satellitare KA-SAT di Viasat venivano disabilitati in massa da un attacco informatico. In aggiunta a questo, i servizi segreti di

Vladimir Putin hanno sfruttato i gruppi filorussi produttori di ransomware come Conti Team per attaccare la *supply chain* (ovvero la filiera di approvvigionamento) di aziende dei paesi Nato con l'obbiettivo di interferire con la produzione di armi e l'erogazione di servizi essenziali come acqua e servizi sanitari.

8.2 Origini, presente e futuro dell'hacktivism

L'azione politica diretta nello spazio fisico cittadino che si concretizza in scioperi, cortei, e nell'occupazione di strade (*Reclaim the streets*), piazze e edifici, era chiamato *activism*. Poi è venuto *l'hack-tivism*, l'azione diretta in rete con tecniche di hacking, e i cortei e le occupazioni sono diventati virtuali, dal Netstrike ai DDoS.

L'hacktivism, termine coniato dal gruppo The Cult of the dead cow nel 1996, deriva dall'unione delle parole hacking e activism. L'Hacking è la messa in opera di una particolare attitudine verso le macchine informatiche che presuppone storicamente la pratica di studiare i computer per migliorarne il funzionamento attraverso la cooperazione e il libero scambio di informazioni tra i programmatori, e presuppone la successiva condivisione del sapere risultante per dare a tutti accesso illimitato alla conoscenza in essi incorporata (Levy, 1984/1996; Di Corinto & Tozzi, 2002). Activism, come abbiamo detto, è invece il termine che indica le forme dell'azione diretta praticate dai movimenti politici di base (grassroots movements) come i sit-in, i cortei, i picchetti.

Successivamente è stato concettualizzato il *Mediattivismo* o *Media activism* (Pasquinelli, 2002), che si è sostanziato nel racconto mediatico delle proteste di piazza e nella diffusione virale dell'informazione in Rete usando anche le immagini in movimento e le *street tv*. Con l'avvento del web 2.0 ha fatto la sua comparsa il *clicktivism*, cioè l'adesione a petizioni, mobilitazioni e proteste, reali e virtuali, con un colpo di click, senza staccare gli occhi dallo schermo del computer. Ma, mentre questa nuova modalità di partecipazione coinvolgeva i grandi numeri dei social network, a compensare tale ondata di "*slacktivism*" - termine gergale per indicare "l'attivismo fannullone", cioè quello che dopo il click si disinteressa della reale entità del cambiamento prodotto -, si è assistito al revival dell'hacktivism col defacciamento dei siti web, i virus politici, gli attacchi DDoS organizzati.

Cosa è accaduto? È accaduto che la rivoluzione digitale ha messo nelle mani di molti individui strumenti di comunicazione efficienti e a basso costo in grado di connettersi alla Rete, mentre le crisi economiche e finanziarie ripetute hanno risvegliato la coscienza delle

ingiustizie e portato singoli, gruppi e movimenti digitali a riorganizzarsi su due fronti: la comunicazione e il sabotaggio.

Non c'è da meravigliarsi, visto che i movimenti sociali hanno sempre avuto una grande quantità di iniziative legate alla comunicazione e il loro rapporto avanguardistico e sperimentale con gli strumenti della comunicazione ha prodotto le fanzine ciclostilate, le radio indipendenti, il videoteatro, il documentario politico, fino ai siti web e ai software di comunicazione gratuiti (Downing et al. 2001; Meikle 2004). Dall'italiana Radio Alice a *Seattle 1999*, fino alle azioni di Anonymous, è possibile rintracciarne il filo conduttore che passa per *Indymedia*, collettivo e hub di informazione indipendente strutturatosi per coprire le proteste della società civile contro gli accordi di libero scambio a Seattle nel 1999 (Di Corinto, 2002), e poi la nascita di Wikileaks (2006)[59], e l'irrompere sulla scena mediatica del movimento Occupy Wall Street (2011)[60].

Gli ingredienti della messa in opera della contestazione sono uguali e diversi da quelli del recente passato, e rappresentano in molti casi l'evoluzione tecnica e la convergenza di strumenti e forme di comunicazione precedenti: dall'uso dei personal media prima (dal fax ai telefoni cellulari, dai camcorder ai videotelefoni, dai siti ai blog) all'impiego di software gratuiti e open source per l'editing di testi, audio e video. A questi strumenti l'ubiquità dell'accesso a Internet (dai Bulletin Board Systems, i BBS, al Wi-Fi) ha dato l'opportunità di riempire social network, piattaforme di blogging e di whistleblowing. Una "rimediazione" che ha consentito di riunire i singoli media, prima isolati, sulla stessa piattaforma, in un processo di convergenza digitale, e di portare uno stesso contenuto su piattaforme o media differenti, la divergenza digitale (Bolter & Grusin, 2000), per realizzare una produzione di informazione indipendente, dal basso, orientata al sabotaggio dei flussi di comunicazione di un potere "che non risiede più in strutture stabili e definite" (Critical Art Ensemble, 1995) ma che è organizzato intorno a dati, messaggi e informazioni. Una chiamata globale all'autodeterminazione digitale.

Se la creazione di strumenti software, server e servizi di messaggistica per la comunicazione indipendente ha subito un arresto con l'affermarsi dei social network e del web 2.0 – che ha portato anche i gruppi di attivisti più radicali ad avere un account Facebook (è il caso del Partito Pirata[61]) -, si sono sviluppate nuove forme di comunicazione e sabotaggio a cavallo

[59] Wikileaks è il sito di informazione anticorruzione fondato, tra gli altri, da Julian Assange nel 2006.

[60] Movimento di contestazione pacifica nato nel 2011 per denunciare le storture del capitalismo finanziario la cui origine viene simbolicamente rappresentata dal palazzo della Borsa di Wall Street a New York.

[61] https://it.wikipedia.org/wiki/Partito_Pirata

tra l'estrazione di informazione protetta e la sua comunicazione al pubblico più ampio, con una strategia di *hack and leak* ovvero "hackera e diffondi".

È il caso di Wikileaks: ottenere informazioni sensibili e offrire al pubblico quelle di cui il potere si vergogna, è stata la sua arma più potente fin dalle origini (Assange, 2012). Prima c'erano stati gli hacker del Chaos Computer Club[62] che negli anni '80, penetrati nel sistema informatico del Comune di Berlino, avevano acquisito le informazioni sulle case comunali sfitte per passarle al movimento dei senza casa. E hanno fatto scuola. Secondo gli hacktivisti, anche l'hacking può ricorrere alla violazione di sistemi informatici protetti (cracking) se ha un fine etico.

Nell'attacco al sito della Corte costituzionale ungherese da parte degli hacktivisti di Anonymous nel marzo 2012 (Coleman 2013/2016), i sabotatori informatici col volto di Guy Fawkes hanno invece cambiato il testo della Costituzione autoritaria voluta dal presidente Viktor Orban affermando il "diritto alla ribellione", con queste parole: "Gli ideologi e i governanti della tirannia, o anche i dittatori, non rappresentano che brevi periodi della storia. Il popolo ha il diritto di eliminare la tirannia e ribellarsi", aggiungendo poi un comma specifico: "[chiediamo] la pensione a 32 anni per gli informatici con il 150 per cento dello stipendio" (La Repubblica, 2012).

Il retroterra teorico di questi guerriglieri dell'informazione è l'etica hacker delle origini: consentire a chiunque l'accesso all'informazione, dovunque essa sia riposta e comunque sia custodita, con la ferma convinzione che l'accesso all'informazione renda tutti più liberi di fare e di scegliere (Levy, 1984/1996). Il paradigma dell'azione è, per questi attivisti, la condivisione di saperi e conoscenze e la difesa dei beni comuni che si producono nei circuiti dell'interazione sociale, e che, secondo gli stessi, necessita di pratiche non ortodosse (Di Corinto & Tozzi, 2002). È questa l'idea che afferma definitivamente quella pratica creativa e disordinata che definiamo di hacktivism, e che vedrà la galassia dei collettivi hacktivisti di Anonymous protagonisti per oltre un decennio.

Sono infatti hacktivisti gli Anonymous gli organizzatori dell'operazione Payback (Olson, 2012), condotta nel 2010 contro i grandi produttori di contenuti creativi, le major hollywoodiane e le loro rappresentanze di categoria, ma anche contro le autorità di garanzia quali l'Agcom italiana e le collecting society come la Siae. In queste iniziative c'è tutta la

[62] https://www.ccc.de/en/

virulenza della contestazione verso chi si appropria del sapere altrui mettendoci sopra un marchio e pretendendo di limitarne la diffusione e la conoscenza se non dietro al pagamento di ogni file tracciabile e certificato.

Questi soggetti del conflitto in rete non sono solo precari dell'industria culturale sfruttati e depressi dalla mancanza di lavoro (Andrea Tiddi, 2002; Franco Berardi Bifo, 2011). Molti sono lavoratori che, in puro stile hacker, di giorno lavorano a far funzionare la macchina produttiva e burocratica degli Stati, mantengono le linee di comunicazione a cavallo degli oceani e scelgono il payoff di prodotti pubblicitari, mentre la notte disfano la loro tela di Penelope. Una moltitudine che non è fatta soltanto di una minoranza colta, istruita, con eccellenti competenze informatiche, perché gli attacchi più virulenti sono stati portati con strumenti facili da usare come il Loic, Low orbit Ion Cannon (Di Corinto, 2010), e scaricabili sotto forma di codici software da installare sul computer, usare e cancellare subito dopo, ottenendo di avvicinare alla protesta ogni tipo di insoddisfazione verso i poteri costituiti.

Organizzazione senza capi, ma con dei leader temporanei, sostituibili, che hanno portato il conflitto all'interno delle reti di CIA, governi e servizi segreti, dal 2004 Anonymous si presenterà come il capostipite di una nuova generazione di hacktivisti che conduce battaglie sociali a colpi di mouse e che agiscono per contagio ed emulazione (Goode, 2015).

In questo contesto, l'emergere di una nuova socialità è stato modellato dalla Rete assumendo molte forme. Dagli Indignados spagnoli a quelli greci, che però hanno costruito i loro propri social network al riparo dei dipartimenti di intelligence di tutto il mondo, per difendersi dall'espropriazione dei propri dati per finalità commerciali. Tra queste esperienze si annoverano quelle dei giovani magrebini che durante la così detta Primavera Araba tra il 2010 e il 2011 attraverso Facebook e YouTube hanno trovato le parole per contestare le dittature, si sono uniti ai coetanei per non sentirsi più soli e trovare il coraggio di scendere in piazza, anche a costo della vita, come quando nella Casbah di Tunisi nel 2010 furono "allestite tende e gruppi di lavoro che si occupavano di Internet, media e attivismo in rete" (F. Massarelli, 2012, pp 40). Un'idea antagonista di sovranità digitale basata sull'autodeterminazione tecnologica.

Così il sapere comunicativo diffuso dei "Millennial", unito alla potenzialità della comunicazione telematica, ha prodotto i nuovi contestatori della Rete in un processo accelerato dalla crisi globale, che è crisi della finanza, dell'economia, della società, della rappresentanza democratica, dello stato-nazione (Klein, 2000).

Dietro alle loro sortite c'era una consapevolezza, teorizzata da Hakim Bey (1991/2007), e Ricardo Dominguez del Teatro del disturbo elettronico (2003), per la quale il potere, da materiale che era, si stava sempre più smaterializzando e non coincideva più con luoghi fisici, portaerei e palazzi, ma coi flussi di comunicazione digitale che possono essere dirottati o sabotati più facilmente dato il loro carattere immateriale.

Gli hacktivisti oggi però non sono più i soggetti del conflitto sociale in rete che rivendicavano dignità e libertà, reddito e tempo libero, democrazia e giustizia, autodeterminazione.

8.3 Infowar, netwar, cyberwar: il ruolo degli hacktivisti nella guerra cibernetica

La disponibilità delle tecnologie di comunicazione influenza le attività e la formazione stessa dei gruppi e dei movimenti della società civile. Con la diffusione di massa di Internet e del Web - il suo servizio più noto e di facile accesso -, la Rete è diventata un campo di battaglia per gli attivisti digitali. Molte delle pratiche di informazione, contestazione e sabotaggio tipiche dei movimenti di protesta sociale sono state così digitalizzate e riversate in Rete. Infine, queste pratiche hanno conquistato la ribalta dell'informazione giornalistica per i loro effetti su porzioni sempre più ampie della società.

Protagonisti in questo scenario sono stati gli hacktivisti, gli hacker-attivisti che già nei primi anni '90 hanno usato la Rete per autorganizzarsi, fare propaganda e controinformazione, condurre azioni politiche dirette. I loro obbiettivi, e i metodi degli inizi, erano quelli dell'Infowar, la "guerra" d'informazione e propaganda, ma oggi gli hacktivisti sono entrati di prepotenza nelle guerre guerreggiate. I moderni hacktivisti, spesso arruolati su una base ideologica, talvolta usati come mercenari, sono arrivati ad accompagnare i conflitti cinetici, le guerre propriamente dette.
Si tratta di una mutazione graduale e forse attesa, ma poco presente nel dibattito pubblico e accademico. Per tracciare l'evoluzione di questa trasformazione culminata nella creazione di vere e proprie milizie di hacktivisti digitali impegnati nei conflitti è utile, a nostro avviso, ripercorrerne la storia recente.

Occorre però chiarire da subito che un conto è la contestazione digitale, l'Infowar praticata dagli hacktivisti, altra cosa è la cyberwar, che usa armi cibernetiche e viene praticata da eserciti, paramilitari e servizi segreti nel contesto della competizione geopolitica e dei conflitti cinetici tra Stati nazionali e che mira a provocare danni o cose e persone. L'Infowar,

è stata diversamente concettualizzata nella storia, ma ogni sua definizione operativa rimanda a una qualche forma di inquinamento dell'informazione.

Il defacciamento[63] dei websites (*web defacements*) di banche e governi, la creazione di luoghi digitali antagonisti in rete, i video virali del collettivo Anonymous, il cybersquatting[64] delle *Urls* e i siti clone di istituzioni come il Vaticano, sono esempi di Infowar, guerra dell'informazione (Di Corinto & Tozzi, 2002; Di Corinto 2014) e datano dalla metà degli anni '90 del secolo scorso. Variamente concettualizzati dai gruppi sociali, e da teorici come Ricardo Dominguez (2003), Hakim Bey (2007) e Tommaso Tozzi (2019), non implicano un danno irreversibile a cose e persone. E questa è la prima grande differenza tra l'infowar degli attivisti con la cyberwar degli eserciti e dei gruppi paramilitari.

Nel novembre 1999, ad esempio, (r)TMark, gruppo di attivisti digitali, pubblica http://rtmark.com/gatt.html[65], un sito contente informazioni sul meeting di Seattle del 30 Novembre del GATT (Global Agreement on Tariffs and Trade, predecessore del WTO, World Trade Organization). Il sito, formalmente identico a quello ufficiale dell'Organizzazione per il commercio mondiale, a dispetto alle aspettative dei visitatori in cerca di informazioni sul summit, mette in discussione gli assunti del libero mercato e della globalizzazione economica.

Nel febbraio 2001, invece, in occasione del Terzo Global Forum, quello sul governo elettronico tenutosi in Italia, a Napoli, alcuni attivisti clonano il sito della manifestazione ufficiale, ne modificano i contenuti e lo riversano su un loro dominio *ocse.org* che, successivamente censurato, viene trasferito sul sito www.noglobal.org/ocse. Anche in questo caso il sito plagiato dagli *alter-global* conteneva una critica radicale al Forum che, secondo loro, era volto "a definire nuove modalità di sfruttamento e controllo sociale attraverso l'informatizzazione degli stati" anziché a promuoverne lo sviluppo democratico. In quell'occasione i contestatori digitali attuarono anche un Netstrike, un "corteo telematico" (Strano Network, 1996), antesignano dei DDoS (Brooks et al. 2021), ai danni del sito di "FinecoOnLine" (www.netstrike.it[66]), e lo usarono come pretesto per avviare un dibattito

[63] Con il termine defacement (in italiano defacciamento) si intende la modifica illecita della home page di un sito web (la sua "faccia") o la sostituzione di una o più pagine interne. Questo tipo di attacco viene eseguito all'insaputa di chi gestisce il sito ed è illegale in tutti i paesi del mondo.

[64] Cybersquatting, o domain squatting, è l'attività di chi si appropria di nomi di dominio altrui corrispondenti a marchi commerciali, entità governativa, personaggi famosi per realizzare un lucro sul trasferimento del dominio, per creare o un danno a chi non lo possa utilizzare oppure farne uno statement politico.

[65] https://web.archive.org/web/20120208220331/http://www.rtmark.com/gatt.html

[66] https://web.archive.org/web/20010201081900/http://www.netstrike.it/

pubblico sulla portata degli scambi finanziari on line e sugli effetti delle bolle speculative nel mercato borsistico telematico.

Queste pratiche di attivismo digitale, o hacktivism, dall'unione delle due parole *hacking* e *activism* (Denning, 1999), non avevano niente a che vedere con le cyberguerre perché non miravano a distruggere e conquistare, ma ad occupare solo temporaneamente degli spazi di comunicazione per parlare all'opinione pubblica e a una platea di altri cyberattivisti. Erano già allora considerate pratiche di "guerriglia comunicativa" e di sabotaggio culturale (Critical Art Ensemble, 1998; Piro, 1998; De Seriis & Marano, 2008).

L'antagonismo politico sociale in rete, secondo i suoi protagonisti, all'epoca rappresentava l'altra faccia della globalizzazione economica (.Zip, 1997). Come si intensificavano gli scambi commerciali e l'economia diveniva "virtuale", così i movimenti sociali esprimevano bisogni universali "globalizzando la rivendicazione dei diritti" attraverso mezzi di comunicazione indifferenti alle frontiere e alle leggi degli stati (Klein, 2000).

L'Infowar degli hacktivisti (www.thehacktivist.com[67]) era qualcosa di assai diverso dalle cyberguerre e dal così detto "terrorismo informatico" o cyberterrorismo, e veniva praticata attraverso l'uso di hacking skills (capacità da hacker) per supportare l'azione diretta dei movimenti politici di base (Di Corinto & Tozzi, 2002). Gli hacker costruivano spazi e strumenti digitali per l'azione politica collettiva, come ad esempio strumenti di open publishing (M. Veneziani, 2006, pp 210-220).

Il fax-strike, il Netstrike, il mass-mailing, il defacciamento dei siti web, sono le forme in cui, dal Canada agli Stati Uniti, in Europa, e anche in Italia, si è sovente articolata la protesta collettiva degli attivisti digitali per tutti gli anni '90. Seppure diversi, i defacements stessi – la sostituzione di una pagina web con un'altra o con un messaggio irridente e critico – somigliano da vicino alla copertura di un cartellone pubblicitario o alle scritte sui muri, sulla scia delle azioni degli attivisti del Billboard Liberation Front[68]. E anche in questo caso l'obiettivo era quello di appropriarsi di uno spazio per esprimere le proprie opinioni, anche quelle più estreme, secondo le logiche del Culture jamming e dell'interferenza culturale.

L'infowar è quindi per gli attivisti primariamente una "guerra di parole", combattuta a colpi di propaganda, autorganizzata, "dal basso". Ed è diversa dalla sua matrice linguistica che

[67] https://web.archive.org/web/*/www.thehacktivist.com

[68] http://www.billboardliberation.com/

rimanda all'Information Warfare, intesa come un insieme di tattiche, tecniche e procedure belliche per assumere una superiorità informativa rispetto all'avversario tramite operazioni di spionaggio e sabotaggio (Rapetto & Di Nunzio, 2001; Borgia, 2022). L'infowar degli hacktivisti è una pratica di resistenza digitale ovvero di autodeterminazione della propria comunicazione che implica il controllo su dati, software e tecnologie. È insomma una rivendicazione di **sovranità digitale**.

8.4 Infowar vs. cyberwar: la costruzione di uno spazio globale di espressione

Il concetto di infowar negli anni '90 esonda dall'ambito militare e viene quindi appropriato dagli attivisti politici i quali, in aggiunta all'uso di strumenti tradizionali di comunicazione (volantini, affissioni, riviste), si "armano" di computer e cominciano a usare la rete come mezzo per comunicare le proprie ragioni a un'audience globale, sfruttando le peculiarità di un mezzo potenzialmente accessibile a tutti da ogni dove, indipendentemente dalla collocazione spaziale e temporale degli attivisti per creare una nuova sfera pubblica (Meikle, 2004). Solo successivamente useranno la Rete come mezzo per realizzare azioni di interferenza sociale e di disobbedienza civile (Critical Art Ensemble, 1998). È in questo passaggio che i computer e la rete Internet diventano lo strumento e non solo il teatro della contestazione, lo spazio cioè, dove la protesta, il rifiuto, la critica, espresse collettivamente, prendono forma e dalle parole si passa ai fatti. È questa la netwar intesa come azione di guerriglia comunicativa e propaganda organizzata, che supera i concetti di blocco e sconfinamento in Rete tipici dell'Infowar praticata dagli attivisti che così diventano *Net Attivisti*.

Ad esempio quando, nel 2014, in segno di protesta, gli attivisti digitali si coalizzano per impedire la sentenza capitale nei confronti del 21enne Ali Mohammed al-Nimr, colpevole di aver incitato alla rivolta i suoi amici via SMS contro il governo saudita lanciando l'hashtag #OpNimr su Twitter, un elenco di tweet preimpostati che ogni net attivista può copiare, incollare e pubblicare, e solo dopo offrendo una lista di siti governativi da attaccare, riuscendo così a mettere offline i siti del ministero dell'Economia e Finanze, della Giustizia e dell'Informazione del regime della famiglia Sahud con attacchi DDoS (Di Corinto, 2015).

Infowar e netwar sono quindi pratiche di conflitto tipiche dell'hacktivism, le cyberguerre no.

La **cyberwar** si riferisce infatti alla guerra cibernetica propriamente detta, cioè a una guerra che si svolge nel cyberspace e che usa l'informatica, la cibernetica e le reti di comunicazione al pari di armi convenzionali, per definizione appannaggio degli Stati e degli eserciti. La cyberwar punta a smantellare i sistemi di comando, controllo e comunicazione del nemico in una maniera intenzionale e pianificata impiegando ingenti risorse computazionali centralizzate e facendo uso di cyber-armi come backdoor[69], botnet[70], malware[71], software exploits[72] e virus trojan[73].

Seppure problematica (Robinson et. Al, 2015), e bisognosa di ulteriori approfondimenti di carattere multidisciplinare per giungere a un framework condiviso (Ashraf, 2021), la definizione generalmente accettata di guerra cibernetica, o cyberwar, è concettualizzata nella dottrina militare come una serie di attacchi informatici contro uno stato-nazione che causano danni significativi, dall'interruzione di sistemi informatici vitali fino alla perdita di vite umane. Secondo il Tallin Manual on the International Law Applicable to Cyber Operations della Nato (NATO Cooperative Cyber Defence Centre of Excellence, 2017), un "attacco informatico" è "un'operazione informatica, offensiva o difensiva, che si prevede ragionevolmente possa causare lesioni o morte a persone o danni o distruzione di oggetti".

Questo tipo di guerra cibernetica, come tutte le guerre, produce tuttavia degli effetti di spillover anche sui civili, ad esempio interrompendo l'erogazione di energia elettrica all'interno di un determinato territorio, come accadrà in Ucraina nel 2015 ad opera di gruppi paramilitari russi (Greenberg, 2019).

E tuttavia le tecniche usate nei conflitti telematici sono per definizione ibride e molteplici (Rapetto & Di Nunzio, 2001; Curioni & Giannuli, 2019). Così come la protesta digitale può determinare l'interruzione di un servizio - si pensi agli attacchi DDoS dimostrativi che bloccano temporaneamente la funzionalità di un sito web pubblico - la cyberwar può fare

[69] Letteralmente "porta di servizio" collocata sul retro di un edificio. Viene chiamato così un canale occulto che consente l'accesso a un sistema informatico eludendo le normali procedura di autenticazione.

[70] Rete di computer utilizzata per attacchi da remoto, o per altre finalità, formata da computer infetti (detti bot o zombie) che, all'insaputa dei legittimi utenti, sono controllati da un utente malevolo (botmaster).

[71] Contrazione di malicious software. Programma inserito in un sistema informatico, generalmente in modo abusivo e nascosto, con l'intenzione di compromettere la riservatezza, l'integrità o la disponibilità dei dati, delle applicazioni o dei sistemi operativi dell'obiettivo.

[72] Un software exploit è un Software o una procedura informatica impiegati per lo sfruttamento di vulnerabilità di un sistema al fine di accedervi abusivamente o attuare azioni dannose.

[73] Il virus trojan è una tipologia di malware che cela le proprie funzionalità (ad es. accesso non autorizzato, furto di credenziali, sabotaggio del sistema target) all'interno di un software legittimo (il nome deriva dal mitico Cavallo di Troia). A tale attacco sono spesso associate tecniche di ingegneria sociale, che inducono il target a scaricare/installare il software contenente il trojan.

uso di tecniche di propaganda tipiche dell'infowar per accompagnare l'attacco vero e proprio (Microsoft Digital Security Unit, 2022).

Le tecniche di infowar usate dagli attivisti sono quindi inizialmente un miscuglio di campagne di informazione e di strategie comunicative derivate dall'arte di avanguardia che mirano a mettere in cortocircuito l'informazione istituzionale cannibalizzando l'attitudine al sensazionalismo tipico dei media mainstream, prendendosi gioco delle veline d'agenzia di stampa e del modo di costruire la notizia da parte dei giornalisti (De Seriis & Marano, 2008). Le campagne di informazione e controinformazione su Internet sono l'equivalente digitale di forme di comunicazione più tradizionali, tipiche dei movimenti politici di base, in cui l'e-mail sostituisce il volantino, la petizione elettronica sostituisce la raccolta di firme all'angolo della strada, il sito web sostituisce i manifesti murali e i cartelloni pubblicitari. Finché, portando alle estreme conseguenze la logica del "panico mediatico", usato con successo dagli epigoni della Beat Generation (Autonome A.f.r.i.k.a. gruppe et. al., 2001), si comincia a produrre notizie false per creare diffidenza e allarme. È il caso dei finti virus o della soffiata relativa a una improbabile intrusione dentro sistemi informatici protetti che prelude alla netwar.

La netwar, la "guerra" nei network digitali, ma sarebbe meglio chiamarla "guerriglia", per il suo carattere discontinuo e asimmetrico, da parte di gruppi irregolari, si presenta pertanto come una forma di azione diretta che punta a creare disturbo e interferenza, ma anche un danno reputazionale all'avversario, si tratti di una lobby politica o di una azienda multinazionale, un governo locale o sovranazionale (Di Corinto, 2001). Si tratta in definitiva di iniziative collettive e pubbliche di comunicazione radicale anche quando producono disservizi e malfunzionamenti, reversibili, dei servizi digitali. È il caso dei primi DDoS, (De Seriis, 2017), del synflood[74] (Critical Art Ensemble, 1998), dei virus artistici (Tozzi, 2019), del doxxing[75] e della divulgazione non autorizzata di dati personali.

Le cyberguerre, al contrario, non hanno come obbiettivo principale la delegittimazione dell'avversario, piuttosto mirano a interrompere e sabotarne i flussi informativi, danneggiando le sue infrastrutture economiche e sociali e mettendo a rischio l'incolumità fisica delle persone.

Assaggi di queste cyberguerre si sono avute all'epoca della crisi fra Usa e Cina a causa della

[74] Il Synflood è un'interferenza nei protocolli di comunicazione per causare "l'inondazione" ovvero la saturazione di un servizio digitale.

[75] Il doxxing, dossieraggio, è la pratica di ricercare e diffondere in rete dati personali e informazioni private.

bomba recapitata “per sbaglio” all’ambasciata cinese di Belgrado durante la guerra del Kosovo (1998-1999). In quel caso i computer del Pentagono e della Nasa furono bersagliati da milioni di lettere elettroniche con virus (*mailbombing*). Oppure nel caso del conflitto telematico che vede combattersi fino ai giorni nostri israeliani e palestinesi. Già nel 2000 i giornali di Tel Aviv riportarono la notizia di un attacco informatico DDoS che aveva messo fuori uso il sito ufficiale di Hezbollah, mentre cyber attivisti arabi avevano deturpato i siti dell’università ebraica di Gerusalemme e dell’accademia di Netanya ed erano penetrati nel sito della difesa israeliano. Successivamente in Palestina si assisterà ad attacchi orientati a mettere fuori uso ospedali, centrali elettriche e impianti di dissalazione.

Da allora le forme e gli strumenti della cyberguerra condotta in maniera coperta dagli Stati attraverso i loro “proxy”, siano essi paramilitari, servizi segreti o stati canaglia, ha visto una costante evoluzione. Dall’uso del virus Stuxnet (Zetter, 2015), di fabbricazione americana-israeliana, che nel 2010 ha bloccato le centrali per l’arricchimento dell’Uranio a Natanz, Iran, danneggiandole, fino all’impiego del malware Black Energy creato dal gruppo russo Sandworm, (Greenberg, 2019), che ha interrotto l’erogazione di energia elettrica in Ucraina nel 2015 lasciando al buio e al freddo 225 mila ucraini all’antivigilia di Natale (Cyber Infrastructure & Security Agency, 2021).

8.5 L’occupazione “militare” dell’agenda mediatica

L’hacktivism è così stato a lungo associato a gruppi come Anonymous, gruppi decentralizzati e destrutturati composti da privati cittadini con differenti background. Sotto questa sigla gli attivisti hanno lanciato numerose campagne (chiamate “Operazioni”, introdotte dal prefisso #Op) contro target individuati in base alle inclinazioni e agli interessi dei suoi membri. Tra le più note, l’operazione contro la chiesa di Scientology, che segna l’avvio globale del fenomeno, e l’operazione Payback, contro la Sony Corporation. Come racconta Geoff White in *Crime dot com. Il potere globale dell'hacking dai virus ai brogli elettorali* (White, G. 2022), gli hacktivisti hanno spesso incarnato, in varia misura, una cultura tecnocratica, creativa e ludica, ma chiunque, a prescindere dalla fede politica, è sempre stato il benvenuto nei gruppi di hacktivisti che si rifacevano alla galassia di questi “anonimi” attivisti nati sul forum visuale 4Chan e che avevano un solo imperativo, “Non attaccare i Media”. Le iniziative di questi hacktivisti old school includono anche altre campagne come l’Operazione KKK di Anonymous contro i membri e i sostenitori del Ku Klux Klan, (Di Corinto, 2014), l’Operazione Lolita, il cui obiettivo era quello di fermare lo smercio di pedopornografia in rete, fino ad arrivare alle azioni della corrente scissionista di Anonymous, LulzSec, responsabile di attacchi

informatici eclatanti alla HBGary, società di cybersicurezza che aveva operato per incastrare sia gli Anonymous che Julian Assange, il fondatore di Wikileaks, e che, da loro hackerata, fu obbligata a chiudere. Altre campagne ancora, di profilo opposto tra di loro, sono state #OpIsrael e #OpPalestine e, nel 2016, #OpTrump e #OpHilaryClinton, a dimostrazione della variabilità di interessi dei gruppi eterogenei di hacktivisti che di volta in volta usano la sigla Anonymous per le proprie rivendicazioni.

Dal 2020 ad oggi però l'hacktivism ha cambiato natura. Per effetto di numerosi conflitti, locali e regionali, in uno scenario geopolitico dalle frontiere mobili, numerosi gruppi di hacker attivisti hanno modificato attività e bersagli rispetto alla logica dell'azione diretta che mira al cambiamento sociale. Il fenomeno dell'hacktivism oggi non sembra più riguardare gruppi eterogenei, le crew, che si uniscono temporaneamente intorno a parole d'ordine precise, o a una causa specifica, il *single issue activism*, per vendicare un comportamento o riparare un torto. Oggi i gruppi di hacktivisti si presentano strutturati e organizzati con strumenti di attacco e difesa sofisticati, e vengono supportati dai governi, seppure raramente in maniera esplicita. In questo contesto l'eccezione è rappresentata dalla creazione il 26 febbraio dell'IT Army ucraino[76]. É accaduto con i primi cyberattacchi che hanno accompagnato l'invasione russa dell'Ucraina nel 2022 e che hanno generato la chiamata alle armi dei cittadini ucraini, quando migliaia di attivisti informatici sono riusciti a bloccare per ore banche e ministeri russi e, presumibilmente aiutati dai servizi di intelligence occidentali, rubato dati governativi usando la sigla di Anonymous come copertura. Una reazione al fatto che imprese, governi e infrastrutture critiche di molti paesi alleati dell'Ucraina venivano bersagliati da sedicenti hacktivisti filorussi. Dal 2021 ad oggi tutti i paesi del G20 hanno infatti subito pesanti attacchi mossi dai gruppi di attivisti russofoni e filorussi come Killnet, Legion, Noname057(16) e, in alcuni casi, hanno avuto un impatto significativo. Gli attacchi, per lo più di tipo DDoS, hanno interessato non solo i governi di questi paesi, ma anche grandi aziende come Lockheed Martin, azienda americana operante nel campo della Difesa, oppure l'Alleanza atlantica NATO.

Viceversa, il collettivo hacktivista Anonymous, attraverso i suoi affiliati, fra il 20 luglio e il 2 agosto 2025 ha rivendicato il defacement e la fuoriuscita di dati da almeno 100 siti web russi nell'ambito della campagna #OpRussia, come risposta digitale all'invasione dell'Ucraina. Le azioni, annunciate su X (ex Twitter), hanno coinvolto entità di diversi settori, tra cui governi locali, istruzione, sanità e commercio al dettaglio. Sebbene le affermazioni siano state supportate da prove parziali, la veridicità dei leak di dati, al contrario dei defacement, non è stata confermata (Fadda, 2025).

[76] https://t.me/itarmyofukraine2022

Ma è stato proprio l'avvio del conflitto Russo-Ucraino successivo all'invasione russa del Donbass, che ha rappresentato un elemento di svolta e un forte stimolo alla partecipazione degli hacktivisti in tutte le azioni cibernetiche successive alla guerra che ne è divampata e, successivamente, in altri conflitti armati.

Graphika, azienda specializzata nell'analisi dei *threat actor*[77] più attivi, il 16 luglio 2025 ha reso pubblico un dettagliato rapporto sugli hacker attivisti coinvolti nel conflitto russo-ucraino, in quello israelo-palestinese e in quello Indo-pakistano (Graphika, 2025)[78]. Secondo gli esperti di Open Source Intelligence e di Threat Analysis dell'azienda, il panorama degli hacktivisti è diventato negli anni sempre più complesso, con gruppi che partecipano a una comunità online altamente attiva che mostra le stesse dinamiche sociali osservate in altre comunità online. Pur essendo guidati da motivazioni ideologiche e politiche, gli hacktivisti monitorati da Graphika sono fortemente influenzati dalla ricerca di attenzione, e costruiscono il loro brand come efficaci threat actor, monetizzando la loro attività, spesso utilizzando l'attenzione ricevuta come opportunità di marketing. L'hacktivism, in questo modo, secondo il rapporto, si è definitivamente globalizzato, con gruppi che hanno motivazioni ideologiche o geopolitiche tradizionalmente diverse ma che uniscono le forze per collaborare agli attacchi.

In particolare, nell'analisi del team di Graphika sono evidenziati i seguenti caratteri di questi gruppi:

Obiettivi: a causa delle loro motivazioni, i gruppi di hacktivisti scelgono regolarmente obiettivi di alto profilo da attaccare o destabilizzare, come banche, piattaforme di social media ed enti governativi. Cercano pertanto diversi mezzi per promuoversi, tra cui l'utilizzo di hashtag e loghi dedicati e la menzione delle loro attività sulla stampa.

Metodo: questi gruppi si impegnano in quella che può essere considerata una sorta di *hacking della percezione*, spesso affermando, senza sufficienti prove di aver attaccato obiettivi di alto profilo o causato interruzioni significative per rafforzare la propria notorietà e presentare i propri obiettivi come facilmente compromessi o privi di sicurezza.

Capacity building: gli hacktivisti mostrano un vivo interesse nello sviluppo di nuove capacità informatiche e comunicative più destabilizzanti, a indicare che la minaccia rappresentata dalla loro comunità quasi certamente crescerà e che i loro attacchi diventeranno più complessi e pericolosi.

[77] Threat actor, attori delle minacce, in ambito informatico si riferisce a singoli e gruppi che avviano o partecipano ad azioni con l'obbiettivo di causare interruzioni, malfunzionamenti e disservizi nell'ambiente cibernetico.

[78] Graphika (2025). Keeping Up With The Hacktivists. Examining How International Hacktivist Groups Pursue Attention, Select Targets, and Interact in an Evolving Online Landscape. Graphika. Disponibile in: https://www.graphika.com/reports/keeping-up-with-the-hacktivists

Profitto e risultati: gli hacktivisti tentano di monetizzare le proprie attività online, utilizzando la pubblicità relativa ai loro attacchi per promuovere e vendere strumenti, servizi e corsi di hacking autoprodotti o di terze parti.

Raggruppamento: il panorama degli hacktivisti include membri più attivi e pubblici che dettano il ritmo per attacchi e campagne di hacking individuando obiettivi specifici e mobilitando altri soggetti a aderire alla propria causa. I gruppi così coalizzati spesso collaborano per amplificare le affermazioni o l'impatto delle campagne. Tuttavia, avviano anche faide interne prendendosi di mira a vicenda e sfruttando le loro schermaglie per generare ancora più contenuti e attirare l'attenzione.

Uso dei social: sebbene i gruppi di hacktivisti siano più attivi su Telegram, alcuni mantengono account su piattaforme di social media tradizionali come Facebook, Instagram e X. Ciononostante, questi gruppi hanno dovuto fare i conti con un aumento della moderazione della piattaforma. Alcuni riemergono regolarmente con nuovi nomi utente e nickname, mentre altri smettono di pubblicare pubblicamente per mesi.

9. Gli hacker e gli attivisti nel conflitto russo-ucraino

"Some of Russia's activities that take place on or through social media are not pure disinformation efforts: rather, they are disinformation efforts functionally linked to a cyberattack of some kind.

Thus, although we largely stay away from technical discussion of cyberattacks, we do touch on cyberoperations when these are closely tied to activities that use information to shape perceptions or behavior—for example, hacks that produce information that is subsequently leaked." (Treyger et. al. 2022)

9.1 Psy-Ops e attacchi cibernetici

Il 27 aprile 2022 Microsoft pubblica un documento in cui descrive il parallelismo tra le azioni di hacking e disinformazione dei russi (Microsoft Digital Security Unit, 2022). Per gli esperti dell'azienda di Redmond negli Stati Uniti, da prima dell'invasione fino alla pubblicazione del rapporto, sono state lanciate 237 cyber-operations contro l'Ucraina da parte di almeno sei differenti gruppi di *nation state hacker*, ossia di esperti informatici governativi collegati ai servizi segreti interni, esteri e militari russi.

Si tratta in gran parte di attacchi distruttivi che hanno fatto uso di virus informatici per indebolire la capacità di reazione del paese attaccato, avendo come target le istituzioni, i servizi e le aziende informatiche, il comparto energetico, i media e le telecomunicazioni ucraine (Brera, 2022).

Secondo gli specialisti di Microsoft gli attacchi in questione sono stati accompagnati anche da ampie attività di spionaggio e sabotaggio che hanno sia degradato i sistemi informatici delle istituzioni, sia cercato di interrompere l'accesso delle persone a informazioni affidabili cercando di minare la fiducia nella leadership del paese.

Come dettaglia il rapporto, gli attacchi informatici russi e la disinformazione sono fortemente correlati e talvolta direttamente sincronizzati con le operazioni militari cinetiche, che prendono di mira servizi e istituzioni cruciali per i civili. Il 1° marzo 2022, infatti, mentre un gruppo russo lanciava attacchi informatici contro un'importante compagnia televisiva, l'esercito russo annunciava l'intenzione di voler distruggere obiettivi ucraini di "disinformazione" e dirigeva un attacco missilistico contro una torre della televisione a Kiev. Ancora, mentre le forze russe assediavano la città di Mariupol, gli ucraini ricevevano un'e-mail da hacker russi che, fingendosi residenti, accusavano falsamente il governo ucraino di "abbandonare" i suoi cittadini.

Gli attori coinvolti in questi attacchi utilizzano tipicamente una varietà di metodi e tecniche avanzate per superare le difese degli obbiettivi, tra cui il phishing, lo sfruttamento di vulnerabilità non risolte del software e la compromissione dei fornitori di servizi di Information Technology, con frequenti attacchi alla supply chain, ovvero la loro catena di distribuzione e approvvigionamento.

E tuttavia le operazioni di interferenza praticate dai russi non sono sempre riconducibili a ordini impartiti da Mosca, ma questo modo di operare è proprio l'essenza della guerra ibrida teorizzata dai suoi stessi generali (Ottaviani, 2022; Bigazzi et al., 2022).

Come scrive Mark Galeotti: "*Per combattere la sua guerra politica la Russia ha creato una macchina indubbiamente flessibile, economica, immaginifica e intraprendente, ma anche difficile da controllare. L'idea che tutti i troll, i propagandisti, le milizie, i corruttori, gli hacker e gli altri soldati di questo esercito siano sempre sotto lo stretto controllo del governo è assolutamente sbagliata. Certo, vi sono operazioni gestite sin dall'inizio a livello centrale e quelle di particolare importanza che, chiaramente, richiedono l'imprimatur del Cremlino. Rientrano in questo novero l'assassinio di Sergej Skripal in Inghilterra nel 2018 e l'interferenza nelle presidenziali americane del 2016. Nel grosso dei casi, tuttavia, Mosca ha incoraggiato molti «imprenditori politici» a prendere l'iniziativa, sovente con i loro tempi e a loro spese. Se falliscono, possono essere disconosciuti; se riescono, possono essere premiati e*

a quel punto lo Stato può subentrare, ampliando o sviluppando l'operazione" (Galeotti, 2017).

A fugare gli eventuali dubbi circa il rapporto esistente tra l'hacking e la diffusione di notizie false sono intervenuti gli stessi servizi segreti ucraini, arrestando nel 2022 un gruppo di cybercriminali specializzato nella vendita di account per diffondere disinformazione. Le autorità ucraine, pur non rivelando i nomi degli arrestati, hanno fornito le prove dell'attività di un gruppo di hacker operanti a Lviv in possesso di circa 30 milioni di account appartenenti a cittadini ucraini ed europei venduti sul Dark Web. Le perquisizioni effettuate nelle case dei sospettati hanno portato al sequestro di hard disk contenenti dati personali, cellulari, schede Sim e memorie flash usate per lo scopo.

Secondo le stime degli investigatori, il gruppo, pro-russo, avrebbe guadagnato circa 400mila dollari rivendendoli all'ingrosso attraverso sistemi di pagamento elettronici come Qiwi e WebMoney.

Nel comunicato stampa, il Servizio di sicurezza dell'Ucraina (SSU) sostiene che i clienti sarebbero propagandisti pro-Cremlino: "Sono stati loro a utilizzare i dati identificativi di cittadini ucraini e stranieri rubati dagli hacker per diffondere false notizie dal fronte e seminare il panico".

Nel comunicato si afferma che gli hacker avrebbero operato per questo scopo: "la destabilizzazione su larga scala in più paesi", e che gli account sono stati utilizzati per diffondere false informazioni sulla situazione sociopolitica in Ucraina e nell'UE, precisando che "l'attività principale dei clienti degli hacker era proprio la creazione e la promozione di account nei social network e nei canali di messaggistica veloce".

In precedenza, le autorità avevano chiuso due farm di bot da 7.000 account impiegati per diffondere disinformazione e creare panico nella regione. Un'attività legata a una fase della guerra russo-ucraina in cui i cittadini di alcune zone, soprattutto nel Donbass occupato, non ricevevano né cibo né informazioni. I pochi giornalisti che sono riusciti a parlarci infatti hanno dichiarato che gli ucraini sotto occupazione non conoscevano l'entità dello scontro con Mosca, la percentuale di territorio occupata e se i propri congiunti fossero ancora vivi.

Quando gli hacker governativi russi attaccano, passano tutte le informazioni agli hacktivisti per consentirgli di costruire una narrazione destabilizzante, effettuare attacchi di maggiore precisione ed efficacia e diffondere propaganda filorussa. Esempio da manuale di come il rapporto tra criminalità cibernetica, hacktivismo e hacking di stato sia anche più diretto (Di Corinto, 2022).

Abbiamo già citato le numerose azioni di interferenza documentate di hacker russi nei processi democratici dei paesi occidentali. Le ripercorriamo in dettaglio:

Nel 2007 un vasto attacco DDoS viene compiuto in Estonia come ritorsione per lo spostamento della statua del soldato sovietico dal centro alla periferia di Tallin, la capitale. L'azione, duramente condannata da Mosca, culmina in due notti di scontri tra manifestanti russofoni e polizia. Successive analisi attribuiranno gli attacchi, che mettono offline 57 siti web governativi e bancari, ad agenti russi (Brooks, et. al. 2021). Durante l'attacco viene usato un malware noto come BlackEnergy. Come conseguenza degli attacchi la Nato decide di collocare proprio a Tallin il Cooperative Cyber Defence Center of Excellence (CCDoE).

Nel 2008**,** in Georgia, lo schema si ripete, ma questa volta l'attacco informatico per "spegnere" i siti georgiani si affianca a una campagna militare vera e propria. Nella notte del 6 agosto gli hacker attaccano siti governativi e di news georgiani. Contemporanei ai DDoS condotti con botnet affiliate a gruppi filorussi si verificano attacchi di SQL injection per defacciare i media georgiani. Il 7 agosto comincia l'invasione militare vera e propria (Shakarian, 2011). Ufficiali russi rivendicheranno gli attacchi quale reazione all'hacking di siti dell'Ossezia del Sud, paese con una forte componente russofona.

Nel 2013 nasce ufficialmente l'Internet Research Agency, la "Fabbrica dei troll" russi. Finanziata da Evgenij Viktorovič Prigožin, imprenditore, politico e comandante mercenario russo, amico del presidente russo Vladimir Putin. La struttura ha il compito di sviluppare contenuti di propaganda a favore del governo di Mosca. I suoi impiegati realizzano contenuti web, scherzi digitali, fake news e traduzioni di articoli governativi per sostenere la Russia nella rivendicazione del possesso della penisola di Crimea, poi annessa alla Federazione Russa. Questi troll successivamente saranno utilizzati per inquinare il dibattito pubblico intorno alle elezioni americane del 2016 creando finti sostenitori online per Trump e divulgando bufale e pettegolezzi per allontanare gli elettori afroamericani dal voto che negli swing states avrebbe potuto avvantaggiare la candidata democratica Hillary Clinton (Rid, 2021; Curioni & Giannuli, 2022; Feltri, 2023).

Nel 2014 il Gru, il Servizio Segreto Militare russo, crea un video falso di Anonymous per sostenere l'invasione dell'Ucraina. Precedentemente era riuscito a intrufolarsi nella posta elettronica di un colonnello ucraino per inserirvi e-mail fasulle relative a una cospirazione tra Ucraina e Stati Uniti volta a danneggiare la Russia (Rid, ivi, pag. 349-357).

Nel 2014, 2015, 2016, 2017, attori russi attaccano le infrastrutture elettriche dell'Ucraina.

Nel 2014 il malware Black Energy, operato da hacker russi, viene inviato alle sei compagnie ferroviarie statali nascosto in documenti power point e all'interno di un allegato che indicava una lista di password deboli da cambiare: lo stesso trucco sarà usato per attaccare tre

compagnie elettriche ucraine il 23 dicembre del 2015 lasciando senza elettricità 225 mila persone. Uno scenario ripetutosi nel 2016 con un altro virus, CrashOvveride (Greenberg, 2019).

Nel 2015 l'attacco al sito web della tv francese TV5Monde e ai suoi account social causa l'interruzione delle trasmissioni per diverse ore. L'hacking televisivo è una mirabolante azione di disinformazione. Rivendicato inizialmente dal *Cyber Caliphate*, l'unità cyber dell'ISIS, il gruppo terroristico mediorientale a cui viene inizialmente attribuito, secondo l'ANSSI francese è opera del gruppo APT28 collegato ai servizi segreti russi. Ancora una volta i servizi russi e i loro affiliati dimostrano con questo attacco di saper sfruttare i momenti di crisi sociale e di agire in maniera da destabilizzare il contesto politico.

Nel 2016 scoppia il Russiagate, in seguito all'intrusione non autorizzata dei server del Comitato Elettorale del Partito Democratico a opera di un sedicente hacker solitario, Guccifer 2.0, che si rivelerà essere un nome di copertura per gli hacker dei servizi segreti russi. Obbiettivo della divulgazione non autorizzata dei materiali esfiltrati è indebolire la posizione di Hillary Clinton in corsa per la Casa Bianca contro Donald Trump, ma ci saranno tentativi di interferenza anche nelle elezioni francesi, tedesche e in Italia durante il governo di Matteo Renzi (Nicodemo, 2017).

Nel 2017 arriva NotPetya. Una valutazione del Centro nazionale per la sicurezza informatica inglese (NCSC) rileva che l'esercito russo è quasi certamente responsabile dell'attacco informatico "NotPetya" del giugno 2017 (National Cyber Security Center, 2018). NotPetya funziona in maniera simile al predecessore Petya, appartenente a una famiglia di malware che infetta sistemi Windows e il cui obiettivo è di colpire compagnie energetiche e istituzioni governative ucraine (Thales, 2022).

Il ransomware Petya, scoperto nel 2016, agiva sui singoli computer privati, criptando alcuni file, bloccando il settore di avvio del sistema compromesso e chiedendo un riscatto in cambio del ripristino di questi file. La variante NotPetya del 2017, si rivolgeva invece principalmente al settore commerciale. Un aspetto che lo ha reso particolarmente noto è il fatto che spesso, anche quando il riscatto veniva pagato, i file della vittima non venivano recuperati (Germani et. al, 2022). Anche per questo i ricercatori sospettano che in realtà intendesse nascondere un cyber attacco che aveva come obiettivo le istituzioni ucraine.

NotPetya è una diversa versione di Wannacry, il malware utilizzato nello stesso anno per bloccare la Sanità inglese, la ditta Maersk, alcune ferrovie caucasiche. Attribuito ad hacker nordcoreani, costruito sulla base di vulnerabilità software sottratte dai misteriosi ShadowBrokers alla sezione della National Security Agency denominata Tailored Access Operation (TAO), il malware/ransomware Wannacry infetterà circa 300mila computer di 174 paesi provocando danni per miliardi di dollari (Thales, 2022; Rid, 2021).

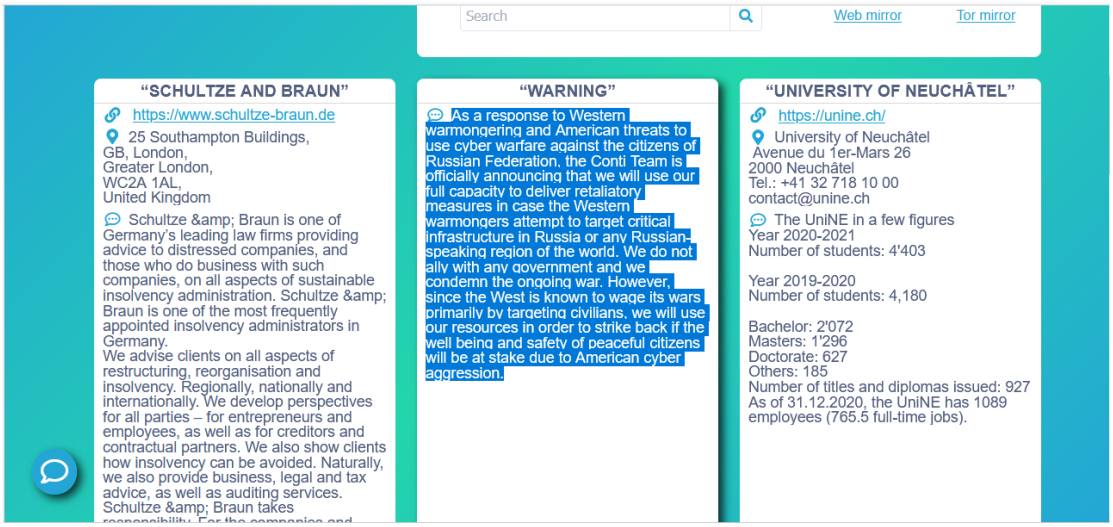
Nel 2018 in Lituania gli hacker russi violano il sito di una delle maggiori televisioni nazionali inserendo un articolo in cui il ministro della Difesa dichiara di essere gay e indagato per molestie sessuali.

Nel 2019 viene avviata la catena d'infezione della supply chain di Solarwinds (Di Corinto, 2021). Un gruppo riconducibile ai servi segreti esteri russi SVR (Microsoft 2021; Mandiant, 2022) penetra nella supply chain del noto produttore texano di tecnologie e arriva alle porte degli asset nucleari statunitensi (Smith & Brown, 2019). Gli esperti di sicurezza informatica avrebbero poi dato all'attacco informatico una varietà di nomi, tra cui Solorigate e Sunburst, con riferimento proprio al nome della società texana SolarWinds il cui software era stato appunto manipolato per organizzare gli attacchi iniziali installando un piccolo malware nel codice di aggiornamento di un programma di gestione della rete chiamato Orion. In questo modo quando i clienti installavano l'aggiornamento sui loro server locali, veniva installato anche il malware collegato a un server di comando e controllo che consentiva di intervenire sulla possibilità di trasferire file, eseguire comandi, profilare il sistema, riavviare le macchine e disabilitarne i servizi. Insomma, gli aggressori avevano ottenuto una backdoor nella rete di ogni cliente che aveva aggiornato il programma Orion, circa 38 mila in tutto il mondo.

Infine, nel 2022, l'invasione del Donbass ucraino viene accompagnata da una serie di attacchi informatici. Si tratta di attacchi DDoS, defacciamento di siti web e distribuzione di virus wiper che cancellano i registri di memoria dei computer Windows. A cadere per primo è però il sistema di comunicazione Ka-Sat, parte di Via-Sat.

Questa ondata di attacchi è immediatamente successiva a settanta attacchi di web defacement documentati ai danni dell'Ucraina qualche settimana prima. Il Digital Forensic Research Lab del Consiglio Atlantico aveva già segnalato una serie di false narrative (Digital Forensic Research Lab, 2021), distribuite sui social media e propagandate da giornali e televisioni pro-Cremlino: tutte con lo stesso scopo, minimizzare gli effetti del conflitto sulla popolazione civile e presentare Putin come un saggio capo di governo (Di Corinto, 2022). I Five Eyes, l'alleanza spionistica dei paesi ex-commonwealth, diramano un'allarmante allerta (Di Corinto, 2022a).

In seguito, scoppia il caso della ransomware gang Conti che, appena tre giorni dopo l'invasione, il 27 febbraio del **2022,** si dichiara apertamente a favore della Russia con un post sul proprio sito nel Dark Web (Di Corinto, 2022b).



Img: Conti si dichiara per la Russia, Arturo Di Corinto, immagine proprietaria

Come abbiamo sottolineato, tuttavia, le operazioni di influenza e di interferenza praticate dai russi non sono sempre e in maniera diretta, riconducibili a ordini impartiti da Mosca, ma questa è proprio l'essenza della guerra ibrida teorizzata dai suoi stessi generali.

Negli ultimi tre anni i gruppi russofoni, gravitanti intorno alla galassia di alcune sigle di hacktivisti hanno ripetutamente colpito obbiettivi occidentali in linea con la loro agenda politica. Le loro azioni sono state contrastate a ogni livello dall'IT Army ucraino, come raccontato anche nel libro di Pierguido Iezzi, Cyber e Potere (2022).

9.2 Killnet, Legione e NoName(057)16, un caso di studio

I principali gruppi di hacktivisti che hanno agito negli ultimi tre anni del conflitto russo-ucraino condividono diverse caratteristiche proprie delle organizzazioni strutturate: una chiara ideologia politica, una gerarchia dei membri e una leadership definita, con un processo di recruitment formale. A titolo di esempio, gli specialisti dell'IT Army Ucraino, coinvolti nella resistenza all'invasione russa del Donbass nel 2022, in una prima fase sono stati selezionati attraverso l'analisi dei curriculum e a monte di un continuo processo di reclutamento sui canali Telegram[79]. Sul fronte opposto, quello russo, vengono messi a disposizione degli hacktivisti un sofisticato tool di attacco come parte del *DDoSia project* realizzato dagli hacktivisti filorussi di NoName(057)16[80], una metodologia organizzativa di

[79] https://t.me/itarmyofukraine2022/1637

[80] https://t.me/c/1228309110/34219

condivisione degli obbiettivi, e un sistema di pagamento come ricompensa per i risultati raggiunti.

Lanciato nel 2022 e successore della botnet Bobik, lo strumento di attacco DDoSia è stato progettato dagli hacktivisti filorussi per mettere in scena attacchi DDoS contro obiettivi situati principalmente in Europa, Australia, Canada e Giappone. Nel periodo che va dall'8 maggio al 26 giugno 2023 i paesi più attaccati sono stati Lituania, Ucraina, Polonia, Italia, Repubblica Ceca, Danimarca, Lettonia, Francia, Regno Unito e Svizzera per un totale di 486 diversi siti web colpiti. Le implementazioni di DDoSia basate su Python e Go scoperte, mostravano un programma multipiattaforma in grado di essere utilizzato su sistemi Windows, Linux e macOS. DDoSia è stato distribuito attraverso un processo completamente automatizzato su Telegram per consentire alle persone di registrarsi all'iniziativa di crowdsourcing in cambio di un pagamento in criptovaluta e di un archivio *.zip* contenente il toolkit di attacco. Ciò che è degno di nota dell'ultima versione analizzata dagli specialisti è l'uso della crittografia per mascherare l'elenco degli obiettivi da attaccare, un fatto che dimostra come lo strumento fosse mantenuto attivamente dagli operatori.

I gruppi hacktivisti si coordinano quindi nella selezione dei bersagli, si coalizzano, si fondono, e collaborano, svolgendo anche consistenti attività di propaganda finalizzate a pubblicizzare e promuovere i loro risultati, veri o presunti che siano, sui canali Telegram, sul Web e in televisione, come accaduto per Killnet, di cui parleremo più avanti. Questi hacktivisti, secondo i rapporti di Microsoft e Google/Mandiant (Mandiant, 2022), si mobilitano in seguito a eventi politici, e operano di concerto con enti governativi, raggiungendo obiettivi strategici e ad ampio spettro con un discreto tasso di successo, e un maggiore impatto sociale favorito dal sensazionalismo mediatico di questi attacchi.

L'evoluzione di questa forma di hacktivism, secondo alcune ricerche, è iniziata, silenziosamente, nel Medio Oriente, a opera di diversi gruppi come Hackers of Savior, Black Shadow e Moses Staff. Tali gruppi hanno concentrato gli attacchi esclusivamente su Israele. La maggior parte di questi non ha nascosto i rapporti con la propaganda antisraeliana promossa dal regime iraniano. Parallelamente, altri gruppi, fra i quali Predatory Sparrow, tutt'oggi molto attivo, si sono concentrati nell'attacco di bersagli iraniani e pro-iraniani: il loro unico piano comune essendo l'opposizione al regime degli Ayatollah.

In realtà l'embrione di queste attività, basate sull'individuazione del nemico con strumenti di Open Source Intelligence (Osint), oppure tramite l'esfiltrazione di dati strategici, la

corruzione dei database avversari, fino all'inoculazione di malware, può essere rintracciata all'interno del mondo hacktivista nella guerra senza quartiere che Anonymous – chiunque si celasse sotto questa sigla – ha condotto contro l'Isis (Di Corinto, 2015).

L'hacktivism è quindi parte essenziale della guerra ibrida di disinformazione, propaganda e attacchi cibernetici, combattuta tra la Federazione Russa e l'Ucraina. Mentre una serie di attacchi informatici, contro l'Estonia nel 2007, la Georgia nel 2008, l'Ucraina nel 2014, hanno ricevuto una provvisoria attribuzione - si tratterebbe infatti di paramilitari e servizi segreti russi, in particolare del GRU, il servizio segreto militare russo, e solo in misura ridotta di hacktivisti -, nella prima parte del 2021, sono emerse altre formazioni, evidenziando un rapporto più diretto tra criminalità cibernetica, hacktivism e hacking di stato.

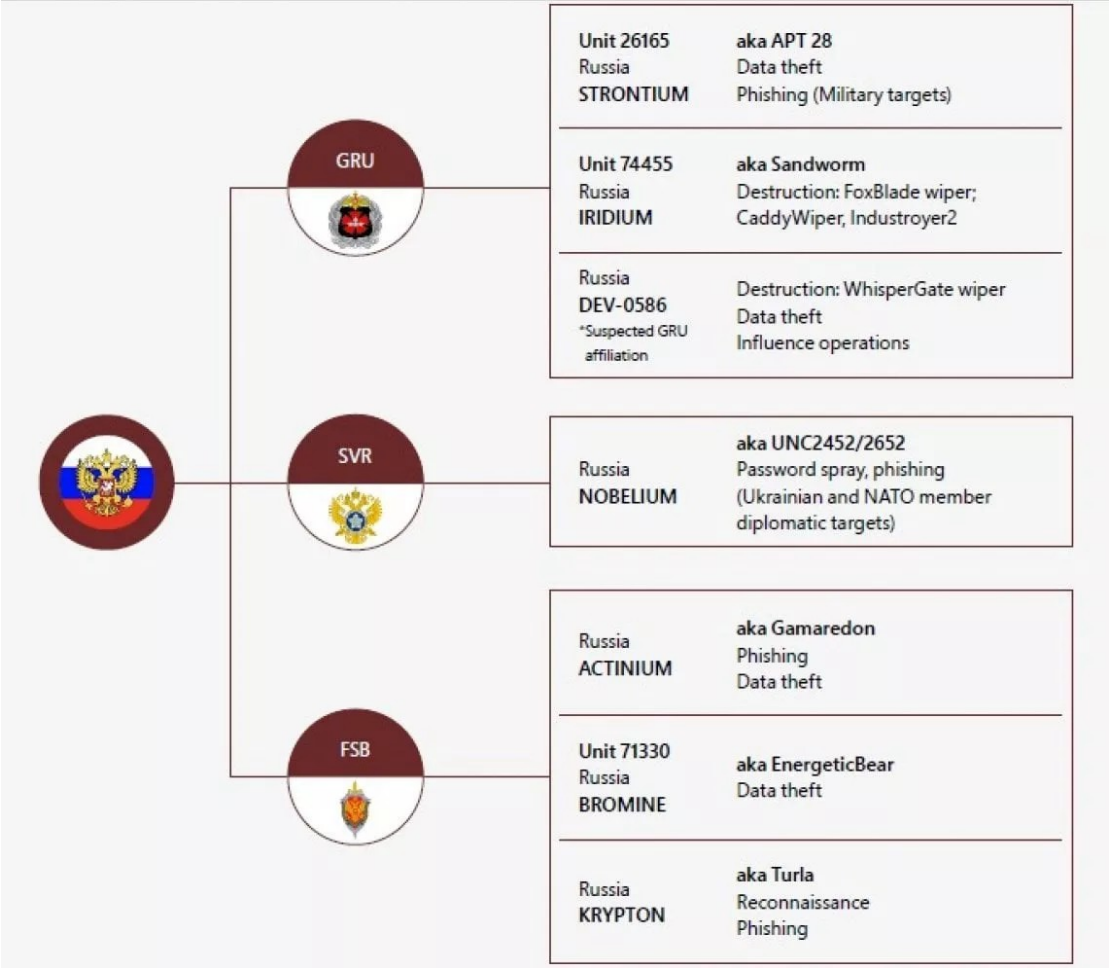


Fig. Servizi segreti russi e gruppi hacker secondo Microsoft

Le analisi degli operatori di Threat Intelligence di Google-Mandiant (Mandiant, 2022), hanno scoperto che quando gli hacker governativi russi attaccano, passano i dati rubati agli hacktivisti entro 24 ore dall'irruzione in modo da consentirgli di effettuare nuovi attacchi e

diffondere propaganda filorussa. Ad agire in questo modo sarebbero in particolare quattro gruppi non governativi: XakNat Team, Infoccentr, CyberArmyofRussia_Reborn e Killnet.

Tuttavia, mentre XakNat si coordinerebbe con l'intelligence russa, Killnet, con cui collabora, è pronta ad attaccare chiunque se pagata. Nel corso del 2022 il collettivo, che ha anche bersagliato l'Italia, ha però incominciato ad ammantare le proprie azioni di patriottismo, diventando una celebrità grazie alle ospitate nella televisione russa.

Negli ultimi mesi del 2022 e per tutto il 2023, gli hacktivisti di Noname057(16) hanno individuato come obiettivi degli attacchi i Paesi nell'Unione Europea dichiaratamente impegnati a sostenere l'Ucraina come Polonia, Lituania, Lettonia, Slovacchia e Finlandia, nonché l'Italia, a più riprese. NoName057(16) ha anche attaccato il sito del parlamento finlandese, dopo che la Finlandia aveva espresso l'intenzione di unirsi alla NATO.

Il gruppo, infine, ha apertamente dichiarato i propri piani a supporto degli interessi russi, come emerge anche nel manifesto di Noname(057)16 (Ninotti & Colatin, 2022), che ha indirizzato, con regolarità, gli attacchi verso l'Ucraina con l'intenzione di espandere il proprio raggio d'azione.

Lo schieramento che si oppone agli hacker filorussi è composto da altrettanto numerosi gruppi di hacktivisti che si sono mobilitati per sostenere l'Ucraina. Alcuni, come l'Esercito IT ucraino, sono ufficialmente controllati dal governo. L'IT Army è stato creato qualche giorno dopo l'inizio dell'invasione russa e comprende volontari provenienti da tutto il mondo per sostenere l'Ucraina seguendone le direttive, ma è composto anche da esperti della stessa intelligence ucraina. Ad affiancarli in alcune azioni, il gruppo Ghostsec, noto almeno dal 2015 per le incursioni contro il *Cybercaliffato*, la cyber-unit dell'Isis, quando, staccatosi da Anonymous, ha incominciato a occuparsi di cyber-intelligence e antiterrorismo per trasformarsi poi in GhostSecSecurity prima di scomparire e riapparire nel cyberspace del conflitto russo-ucraino con lo stesso nome (Di Corinto, 2022) con un'azione eclatante: l'esfiltrazione di un database con i dati personali di oltre 120.000 ufficiali e soldati dell'esercito russo, compresi i numeri di telefono e gli indirizzi (Iezzi, 2022).

Secondo l'azienda di sicurezza informatica CyberKnow a maggio 2023 si contano 112 gruppi di hacker attivisti che parteggiano per l'una o per l'altra parte nel conflitto russo ucraino, a luglio sono 128 e, nel febbraio 2024 commando i 51 gruppi pro-Ucraina, i 72 pro-Russia, e 5 gruppi di incerta attribuzione, si arriverà a ben 138 gruppi attivi nel cyberspace in guerra.

(Figura 1: a luglio 2023 sono 128 i gruppi hacktivisti attivi nel conflitto russo-ucraino copyright CyberKnow)

All'interno di questa galassia hacktivista il gruppo più noto è stato a lungo Killnet, pubblicamente annunciato attorno al febbraio 2022, all'inizio del conflitto russo-ucraino. Durante la guerra in Ucraina ha rivendicato attacchi DDoS ai siti governativi occidentali e di aziende americane. Sul proprio canale Telegram ha dichiarato che il suo obiettivo è attaccare "i Paesi Nato e l'Ucraina".

Il collettivo Legion a essi affiliato si presenta invece come una versione russa di Anonymous. Perlomeno ne emula il linguaggio e l'estetica sia nei messaggi sia nelle immagini. Ma a differenza di Anonymous, che dopo l'invasione russa si è apertamente schierato a favore dell'Ucraina, Legion sostiene azioni a favore della Russia. Killnet è diventato famoso dopo il 3 aprile 2022, quando un altro gruppo di hacker, Bluehornet/Atw, ha diffuso i dati personali di quelli che sarebbero alcuni dei leader del gruppo, e rivelato l'esistenza della botnet di Killnet. Bluhornet è un gruppo antagonista di Killnet nella controparte virtuale della guerra cinetica tra Russia e Ucraina.

Il gruppo Killnet ha iniziato le sue attività aggressive a marzo 2022, con obiettivi primariamente ucraini, ma già ad aprile il gruppo aveva completamente cambiato l'oggetto della sua attenzione supportando gli interessi geopolitici russi in tutto il mondo. Tra fine

febbraio e settembre 2022, il gruppo ha affermato di aver portato a termine più di 550 attacchi. Solo 45 di questi però erano indirizzati all'Ucraina: meno del 10% del numero di attacchi totale (Check Point Research, 2022).

Molti di questi attacchi erano diretti a obiettivi di alto profilo come i principali siti governativi, grosse compagnie finanziarie, aeroporti e altri bersagli. Mentre in alcuni casi è difficile comprendere l'impatto reale, in altri casi gli attacchi hanno chiaramente avuto successo, provocando l'inattività dei principali siti web, molti dei quali fornitori di servizi pubblici essenziali.

Ecco alcuni esempi nel 2022:

1. A marzo 2022, l'aeroporto internazionale di Bradley in Connecticut (US), ha subito un attacco DDoS che ha interessato il proprio sito web. Le autorità statunitensi hanno confermato un tentato attacco DDoS su larga scala sul sito dell'aeroporto.
2. Ad aprile, alcuni siti web che appartengono al governo rumeno, come quello del Ministero della Difesa, quello della Polizia di Confine, quello della Compagnia Nazionale dei Trasporti Ferroviari e una banca commerciale, sono stati resi irraggiungibili per diverse ore.

   Questi attacchi si sono verificati in risposta a una affermazione fatta dal leader rumeno del partito Socialdemocratico Marcel Ciolacu, che si è offerto di procurare armi all'Ucraina.
3. A maggio, ingenti attacchi DDoS sono stati portati a termine contro due fra i maggiori Paesi europei:
   - Sono stati coinvolti diversi bersagli tedeschi, incluso il governo e siti web dei politici, fra questi, il sito del partito a cui appartiene il cancelliere Olaf Scholz, il sito del Ministero della Difesa tedesco, quello del Parlamento tedesco, quello della Polizia Federale e diverse autorità della polizia statale. Secondo gli osservatori, una risposta agli sforzi dell'amministrazione Scholz di fornire equipaggiamento militare all'Ucraina, autorizzando il trasferimento di 50 Gepard anti-aircraft, e annunciando la consegna di 7 sistemi di artiglieria semoventi e a fuoco rapido.
   - Anche il Senato italiano, il Ministero della Difesa e l'Istituto superiore di sanità sono stati presi di mira con attacchi da negazione di servizio ai propri siti web.
4. A giugno, due significative onde di attacchi sono state portate a termine contro la Lituania e la Norvegia in risposta agli sviluppi geopolitici che sono avvenuti fra questi Paesi e la Russia:

Seguendo la decisione del governo lituano di fermare il transito di beni russi verso Kaliningrad, un'ondata rilevante di attacchi ha colpito i servizi pubblici lituani e il settore privato. Durante l'attacco, Jonas Skardinskas, il capo della cybersicurezza presso i Centro di Cyber Sicurezza Nazionale Lituano, ha avvisato che i disagi con i trasporti, i settori finanziari e quello energetico sarebbero potuti continuare per diversi giorni amplificando l'impatto dell'attacco. A un certo punto la maggioranza dei siti web lituani non era accessibili tramite indirizzi IP esterni al paese, più probabilmente come misura preventiva finalizzata a mitigare la portata dell'attacco.

Lo stesso mese, diverse organizzazioni norvegesi sono state disconnesse. Si pensa che questo attacco sia stato eseguito come risultato di una disputa riguardante il transito attraverso il territorio norvegese verso un estrattore di carbone sotto il controllo russo situato nell'Artico.

5. A luglio, Killnet ha concentrato i propri sforzi sulla Polonia e causato l'indisponibilità di molti siti web. Molti degli attacchi sono stati diretti ai portali governativi, le autorità di tassazione e i siti web della polizia.
6. Agosto è stato un mese intenso per Killnet. È cominciato tutto con un attacco in Lettonia: dopo aver dichiarato la Russia come "un Paese rappresentante del terrorismo", il sito del parlamento lettone ha subìto un attacco DDoS. Successivamente (nello stesso mese), l'Estonia ha affrontato l'attacco più esteso da quello del 2007, effettuato in risposta alla rimozione del monumento al soldato sovietico. L'efficacia di questi attacchi è stata discutibile, in quanto sembra che l'Estonia fosse ben preparata per questo genere di eventualità. Ad agosto, Killnet ha anche iniziato a concentrarsi sugli USA. Il gigante della produzione militare americana Lockheed Martin è stato pesantemente bersagliato da Killnet come conseguenza del rifornimento al sistema militare dell'esercito ucraino. Parallelamente Killnet ha anche bersagliato la US Electronic Health Monitoring e Tracking System e il Senato statunitense, che stava dibattendo la possibilità di inviare un aiuto addizionale all'Ucraina.
7. A settembre il gruppo ha bersagliato l'Asia per la prima volta indirizzando i suoi sforzi in particolare al Giappone, a causa del supporto giapponese all'Ucraina.

Con l'evolversi del conflitto scaturito dalla contesa delle Isole Kurili fra Russia e Giappone, Killnet ha anche attaccato con successo diversi siti giapponesi, inclusi quelli di e-government, i siti di trasporto pubblico della città di Tokyo e Osaka, i sistemi di pagamento JCB e Mixi, il secondo più grande sito web giapponese.

9.3 Una struttura paramilitare per gli hacktivisti russi

I più grandi gruppi di hacktivisti che sono emersi nel corso degli ultimi tre anni si caratterizzano per operazioni ben strutturate che li mettono nelle condizioni di essere efficaci e di attrarre persone con maggiori skills. Queste persone sono solitamente motivate da una chiara ideologia legata allo Stato e i loro obiettivi sono parte di un manifesto che contiene un elenco di regole da seguire.

Killnet è arrivata a contare più di 100.000 iscritti nei suoi canali Telegram ed "è organizzata secondo una struttura militare con una gerarchia marcatamente top-down e consiste in un insieme di squadre preparate a eseguire attacchi che rispondono a un ordine principale. Tra questi sottogruppi il primario è Legion. Tutti sono guidati da un hacker anonimo con nickname KillMilk, che ha annunciato la sua intenzione di distaccarsi dal gruppo a luglio, rimanendo però ancora coinvolto nelle attività del gruppo. Legion e le varie crew (squadre conosciute come*:* "*Jacky*", "*Mirai*", "*Impulse*", "*Sakurajima*", "*Rayd*", "*Zarya*", "*Vera*", "*Phoenix*", "*Kajluk*", "*Sparta*" and "*DDOSGUNG*") sono considerate le forze speciali di Killnet, con Legion identificata come la sua forza di cyber-intelligence" (Check Point, 2022).

Tante piccole squadre organizzate attorno al maggiore gruppo e al suo leader, che assegna ordini d'attacco a ciascun capogruppo dando vita a infrastrutture indipendenti e migliorando così le probabilità di sopravvivenza dell'intera organizzazione. Questo metodo si è dimostrato efficace dal momento che la squadra continua a reclutare membri, crescendo numericamente. La pagina Telegram per lungo tempo ha pubblicato regole, discussioni riguardanti gli obiettivi e le istruzioni rispetto a creare/unirsi a nuove squadre per i membri in cerca di autonomia o di un avanzamento gerarchico. L'evoluzione di Killnet li ha messi nella situazione in cui gli altri gruppi vogliono collaborare con loro, o ufficialmente unire le forze.

Un nuovo interessante fenomeno ha riguardato però i metodi di reclutamento del gruppo. Diversamente da Anonymous, che è orgoglioso di dare il benvenuto a chiunque, senza imporre alcun prerequisito riguardante skill o progetti specifici, la nuova era hacktivista accetta solo membri che rispettano prerequisiti minimi. Molti gruppi, come Killnet e le sue squadre, hanno scelto di investire in programmi di recruitment, pubblicizzati sui propri canali Telegram. Alcuni gruppi hanno istituito un processo di preselezione per assumere solo hacker competenti o esperti di un particolare campo, per ridurre il rischio di fare errori che potrebbero compromettere le operazioni.

In ogni caso, Check Point Software ha osservato che Killnet affidava le istruzioni sugli attacchi DDoS alle masse anonime, forse a causa della mancanza della "forza-lavoro" necessaria per portare a termine le azioni pianificate e offrendo ricompense economiche agli affiliati.

Il processo di recruitment appare simile per molti gruppi russi. Per esempio, XakNet (che si definisce come il "Team dei Patrioti Russi") è un gruppo di utenti russi attivo all'incirca da marzo 2022 e poi inabissatosi nell'oscurità del Deep Web. Il gruppo è entrato in campo minacciando di contrattaccare a fronte di qualunque campagna cyber rivolta contro la Russia e avrebbe anche individuato diverse entità interne all'Ucraina che hanno poi rubato dati ufficiali del governo ucraino. XakNet ha dichiarato che non avrebbe reclutato hacker, pentesters (specialisti nell'esecuzione di test di vulnerabilità di siti web), o specialisti Osint[81] senza esperienza e capacità dimostrate.

Anche il gruppo NoName(057)16 è sempre stato piuttosto rigido sul reclutamento. NoName investe parte delle sue risorse per offrire un training adeguato ai seguaci tramite canali Telegram, piattaforme di e-learning, tutorial, corsi e attività di mentoring, svolta anche nei canali di supporto in lingua inglese.

I gruppi di hacktivisti si sforzano costantemente di utilizzare strumenti più avanzati per eseguire i loro attacchi, dal momento che più gli attacchi arrecano danni, più il gruppo guadagna in termini di notorietà ed esposizione. Nonostante i segnali dell'uso di tecniche avanzate, la maggior parte dell'attività rimane concentrata attorno agli attacchi DDoS tramite il ricorso a enormi botnet. Questi attacchi sono tuttavia differenti, suddivisi in attacchi DDoS volumetrici, applicativi e infrastrutturali (Csirt Italia, 2022).

9.4 NoName(057)16 e gli effetti della disinformazione in Italia

Ricapitolando, all'interno della galassia hacktivista russofona, NoName(057)16 è il gruppo di sedicenti hacker attivisti più prolifico. Si è presentato sulla scena comunicativa agli albori del conflitto russo-ucraino nel marzo del 2022 rivendicando numerosi attacchi informatici contro diversi target europei e dei paesi Nato. E ancora continua a farlo.

Nel suo manifesto ha dichiarato da subito l'intento di proteggere la popolazione russa e di voler replicare agli attacchi della "Ukropropaganda", come loro stessi definiscono la propaganda ucraina. Per farlo, il gruppo usa prevalentemente il proprio omonimo canale Telegram con lo scopo di organizzare le sue campagne DDoS, prendersi gioco degli avversari,

[81] Osint, acronimo di Open Source Intelligence, è la raccolta di informazioni da fonti aperte.

istruire i volontari che vogliono partecipare alle azioni usando la piattaforma DDoSia per lanciare attacchi DDoS attraverso botnet sincronizzate. Il gruppo risulta collaborare con diversi altri gruppi di sedicenti patrioti russi come Killnet e XakNet, quest'ultimo considerato collegato ai Servizi segreti esteri russi (SVR).

Tra i numerosi paesi attaccati dal gruppo di attivisti figura anche l'Italia. Gli attacchi DDoS rivendicati dal gruppo verso i soggetti nazionali italiani, di tipo volumetrico, applicativo e infrastrutturale, hanno rallentato, e in alcuni casi bloccato, i siti target di alcuni Ministeri, quelli di diverse compagnie di trasporto pubblico, aeroporti, e banche. Gli attacchi sono avvenuti sempre in coincidenza di eventi bellici oppure a seguito delle dichiarazioni in cui il Governo italiano ha manifestato il proprio supporto all'Ucraina aggredita dalla Federazione Russa. Altri bersagli sono stati i Paesi baltici, il Canada, la Polonia, la Spagna, la Francia e l'Inghilterra, le nazioni europee che più sostengono l'Ucraina.

NoName(057)16 pubblicizza le sue campagne in un canale di lingua russa e in un canale mirror dove parte dei contenuti è tradotta in inglese per i membri non di lingua russa. Il gruppo ha anche creato un altro canale in cui alcuni dei suoi membri dialogano sugli aspetti tecnici relativi alle campagne DDoS e un quarto in cui fornisce istruzioni su come utilizzare DDoSia per condurre gli attacchi. Tra il 2024 e il 2025 il gruppo dovrà continuamente ricreare nuovi gruppi Telegram dove convogliare i propri seguaci a causa della ripetuta loro chiusura da parte della piattaforma, presumibilmente su richiesta delle Autorità di polizia dei paesi interessati dai suoi attacchi.

Gli utenti di DDoSia si definiscono "cyber-esercito" e lavorano insieme per sostenere gli sforzi e fornire risorse e competenze per raggiungere gli obiettivi del collettivo. Nei diversi canali Telegram del gruppo si possono leggere i presunti risultati di ciascun attacco del cyber-esercito che i gestori rivendicano a fini di propaganda e coesione interna visto che i loro effetti sono stati spesso limitati nel tempo e nelle conseguenze, come ha dichiarato più volte l'Agenzia per la Cybersicurezza Nazionale italiana, ACN, che ha fornito ai target nazionali interessati le indicazioni operative per mitigare questi attacchi.

Caratteristica dei gruppi collegati al collettivo NoName(057)16 è che molti degli affiliati sono spettatori passivi che non partecipano alle discussioni e usano le emoji per dichiarare il gradimento dei thread generati nel canale, altri appaiono individui versati tecnicamente e condividono con gli altri motivazioni di carattere politico-ideologico nell'attaccare i paesi che

considerano “russofobi” e la stessa “Ucraina banderista”, dal nome del nazionalista ucraino Stephan Bandera.

Tuttavia, gli utenti delle chat sono incoraggiati a condurre proprie operazioni solo contro quei soggetti attraverso cui si possa determinare un danno finanziario o reputazionale.

Indipendentemente dalla riuscita dell’attacco, poche ore dopo il suo avvio, i canali del collettivo vengono popolati da messaggi di vittoria, e dal giorno dopo presentano la copertura giornalistica ottenuta nei vari paesi di cui hanno attaccato enti governativi e infrastrutture. É evidente che per poterlo fare, dispongono di “sensori” o sentinelle che seguono le vicende politiche e monitorano la stampa dei paesi bersaglio degli attacchi.

Come accaduto in precedenza con le incursioni di Killnet, i canali Telegram di NoName(057)16 restano il luogo per reclutare persone, aggregare consensi e comunicare coi media in un’opera di disinformazione ad ampio spettro dove quello che conta non è solo la rivendicazione degli effetti dei risultati delle azioni, spesso limitati, ma il racconto che è possibile farne per trasformare gli attacchi cibernetici in strumento di proselitismo, propaganda e coesione.

La tattica usata per conseguire questi obbiettivi è sempre quella di “fotografare” con uno screenshot i risultati dell’interruzione momentanea dei siti web attaccati registrata dalle piattaforme dedicate a tale funzione come check-host.net[82], pubblicare lo screenshot sul canale, attendere la reazione mediatica che, complice un giornalismo spesso poco informato e sensazionalista, ingigantirà gli effetti dell’attacco, creando un clamore propagandato dagli aggressori come “certificazione” dell’efficacia dell’azione di incursione.

Esemplari in questo senso sono stati proprio i cyberattacchi rivolti all’Italia, con un primo picco toccato in Italia intorno al 5 febbraio 2023. In questa occasione la tattica di disinformazione introdotta da NoName(057)16 ha avuto un aiuto insperato dalle dimissioni del Direttore Generale dell’Agenzia per la cybersicurezza nazionale italiana, professore **Roberto Baldoni**. Il direttore dell’Agenzia ha offerto infatti le sue dimissioni al Governo italiano, da cui l’Agenzia dipende, il giorno 6 marzo 2023, proprio a cavallo di uno degli ormai consueti attacchi DDoS del gruppo. Le dimissioni, secondo una spiegazione riportata da

[82] Check-Host.net è uno strumento online per verificare la disponibilità di siti Web, server, host e indirizzi IP. Fornisce dati sulla posizione del dominio e dell'indirizzo IP da alcuni database IP di geolocalizzazione e anche whois.

numerosi quotidiani, sarebbero state dovute proprio al riconoscimento dell'inefficacia delle azioni di contrasto messe in capo dal direttore dell'Agenzia contro gli hacktivisti. Altri quotidiani, tuttavia, hanno ricostruito diversamente il motivo delle dimissioni dello stimato direttore senza collegarle alle incursioni dei filorussi ma come una decisione di politica interna nazionale (Zorloni, 2023).

Tuttavia, il gruppo ha sfruttato questa occasione per rivendicare, attraverso tutti i mezzi a propria disposizione, il merito della "cacciata" di Baldoni quale dimostrazione del carattere "inarrestabile" della propria azione di disturbo scrivendo su Telegram: "La nostra serie di attacchi all'infrastruttura internet italiana può essere giustamente considerata riuscita: a seguito di ciò, è stato infatti rimosso dal suo incarico il capo dell'Agenzia nazionale per la sicurezza informatica italiana. Vediamo - ha proseguito il messaggio ironico - come il nuovo capo di questo ufficio italiano se la caverà con le minacce informatiche provenienti dal team di NoName(057)16" (Fiammeri, 2023). Nonostante i loro continui attacchi, però, il nuovo capo dell'Agenzia, prefetto Bruno Frattasi, è rimasto al posto a cui è stato designato dopo le dimissioni di Baldoni, mentre gli attacchi degli hacktivisti filorussi non sono mai cessati e, anzi, in Italia, hanno conosciuto un nuovo picco nel gennaio del 2025. Il gruppo ha poi continuato ad attaccare e nei mesi di giugno e luglio ha condotto l'ennesima campagna DDoS contro portali nazionali italiani, della durata di 13 giorni continuativi, con la dichiarata intenzione politica di "punire" l'Italia per il supporto dato all'Ucraina e Zelensky. In seguito alla conferenza per la ricostruzione dell'Ucraina tenutasi a Roma il 10 luglio 2025 da parte dei così detti "Paesi volenterosi", NoName ha però subito un duro colpo dalle autorità: il giorno 17 luglio, Europol ha reso noto che tramite l'operazione Eastwood ha smantellato buona parte dell'infrastruttura di attacco degli hacktivisti (Europol, 2025), che però sarà velocemente ricostituita a dimostrare l'alto grado di organizzazione del gruppo d'attacco. Come risultato delle perquisizioni e degli arresti dell'operazione Eastwood si scoprirà che cinque soggetti presunti appartenenti al gruppo sono italiani, ma solo due con le competenze di un hacker, gli altri sarebbero stati semplici fiancheggiatori impegnati nella propaganda delle sue azioni.

Perciò, indipendentemente dal reale motivo delle dimissioni del direttore dell'agenzia per la sicurezza informatica italiana, la vicenda rappresenta un esempio da manuale di come possano essere sfruttate le Misure Attive: non conta tanto il risultato, ma la capacità di creare dubbio e sconcerto nei destinatari dell'operazione di disinformazione accoppiata all'uso di tecniche cyber di attacco (Di Corinto, 2024). Come dice Thomas Rid rievocando la nascita del servizio "A" del KGB, nel 1962, per cui era controproducente distinguere i fatti dai

non-fatti: “Quel che rendeva attive le misure attive non era la correlazione con la realtà, ma con le emozioni, con i valori condivisi da una comunità, e la capacità di esacerbare le tensioni esistenti: nel gergo degli attori della guerra fredda, di rafforzare le contraddizioni”.

10. Gli hacker e gli attivisti nel conflitto Israele-Hamas

10.1 La Cyberguerra in terra di Palestina

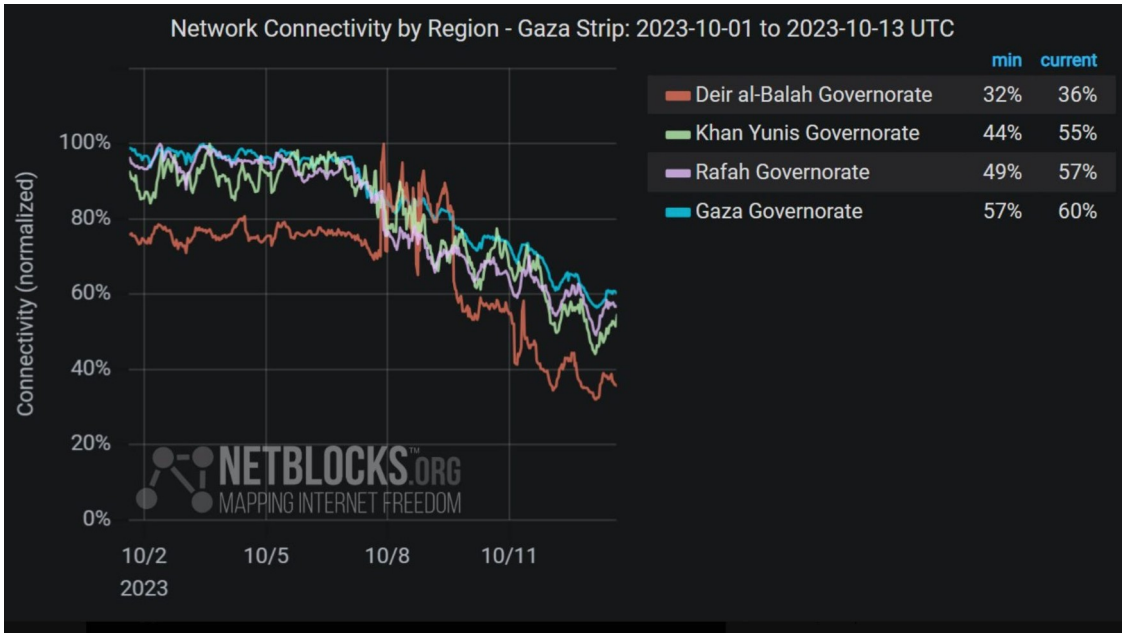
Il 7 ottobre 2023 le milizie armate del partito palestinese di Hamas invadono Israele uccidendo 1.188 persone e prendendone in ostaggio 251. All’indomani del brutale attacco, Killnet dichiara la guerra cibernetica contro Israele alleandosi con Anonymous Sudan. Il 9 ottobre sono già 58 i gruppi di hacktivisti che, per Cyberknow, partecipano al conflitto. Tuttavia, secondo gli analisti della società italiana di cybersicurezza Swascan, nelle due settimane successive all’attacco, dopo la risposta armata di Israele, ammontano a 178 i gruppi di hacker attivisti che si fronteggiano: circa 150 a favore della causa palestinese, meno di 30 a supporto di Israele (Gabanelli & Santucci, 2023). È l’occasione in cui si manifesta in tutta la sua evidenza il fenomeno della guerra algoritmica in cui milizie regolari, civili e attivisti, operano con le armi della disinformazione, della propaganda computazionale e dell’Intelligenza Artificiale, facendo ricorso anche alle incursioni informatiche per il sabotaggio di strutture non militari.

10.2 Attacchi informatici e operazioni di influenza: pro-Pal vs. Israele

I primi attacchi informatici nel conflitto successivo al 7 ottobre, sono gli attacchi DDoS. Orientati a determinare un senso di incertezza e paura, e causare un danno psicologico e materiale nella popolazione che non riesce ad accedere ai servizi di base di siti web come banche, aeroporti e ospedali, hanno la funzione aggiuntiva di contribuire a costruire una narrativa secondo cui l’attaccato non è in grado di proteggere sé stesso e di indurre l’opinione pubblica a protestare contro chi ne dovrebbe garantire la sicurezza, cioè gli stessi governi bersaglio degli attacchi. All’indomani dello scoppio della guerra tra Hamas e Israele, infatti, l’Indian Cyber Force attacca e mette fuori uso il sito web di Hamas mentre altri gruppi pro-Israele rendono irraggiungibili i siti web dell’Autorità Nazionale Palestinese, quello del Ministero degli Esteri e dell’Autorità Palestinese dell’Energia. Viceversa, i gruppi pro-Palestina attaccheranno l’Agenzia Spaziale israeliana, diverse banche e il Ministero degli Esteri, l’aeroporto Ben Gourion e il sito del quotidiano Jerusalem Post. Il gruppo pro-Palestina AnonGhost individua una vulnerabilità nell’App *Red Alert* usata dagli israeliani per diffondere via smartphone gli allarmi sui missili lanciati dalle milizie di Hamas e la sfrutta per disorientare i suoi utilizzatori inviando falsi messaggi sul pericolo di una inesistente “bomba nucleare” (CyberExpress, 2023). *Cyber Av3ngers*, altro gruppo della galassia pro-pal,

rivendica la responsabilità di un attacco informatico alla centrale elettrica Dorad in Israele e, come prova della riuscita incursione, ne pubblica i documenti sul proprio canale Telegram. Molti altri gruppi si dedicheranno ad attaccare i singoli paesi che da subito sostengono il diritto di Israele a difendersi con DDoS, defacciamenti e malware. L'Italia è tra questi, motivo per cui dal 22 al 26 ottobre successivi i portali web nazionali di banche, ministeri e trasporti, saranno continuamente bersagliati dal gruppo filopalestinese *Mysterious Team Bangladesh* (Crescenzi, 2023).

Tra i suoi primi atti di propaganda e di terrore, Hamas distribuirà in rete i video choc degli eccidi perpetrati dai suoi miliziani e li ritirerà poco dopo essendo questi diventati materiale per la contropropaganda israeliana. Israele interverrà immediatamente a bloccare gli Internet Service Provider della Striscia di Gaza. In base alle stime della Ong *Access Now*, alla data del 31 ottobre 2023, 15 dei 19 operatori attivi a Gaza affrontano la chiusura completa dei loro servizi di telefonia mobile e a banda larga, mentre i restanti quattro subiscono disagi significativi, con un impatto su milioni di persone. Le chiusure complete alla fine colpiscono direttamente circa 411.000 utenti che utilizzano questi operatori a Gaza, e altri 34.000 in Cisgiordania.[83]



Gli attivisti palestinesi sul campo tenteranno comunque di sviluppare una narrazione a loro favorevole, contrastati in questo dalla poderosa macchina di propaganda del governo israeliano presieduto dal leader del Likud, Benjamin Netanyahu. Una macchina il cui operato ricorda da vicino la logica russa della disinformazione e l'approccio 4D: respingere, distorcere, distrarre e sgomentare, descritto per la prima volta da Ben Nimmo (Mandić & Klarić, 2023, op.cit.).

[83] Access Now (2023). Palestine unplugged: how Israel disrupts Gaza's internet. Access Now. Disponibile online in https://www.accessnow.org/publication/palestine-unplugged/ [10 novembre 2023]

I gestori delle piattaforme social americane, già ripetutamente denunciati per lo shadowbanning[84] e la discriminazione algoritmica nei confronti degli utenti palestinesi (Loewenstein, 2024), si attivano, e Hamas viene subito bandita da Facebook, da Google e Instagram. Rimane però attiva su Telegram; i suoi canali arrivano a contare circa 1 milione di follower.

Alcuni gruppi di hacktivisti sostenuti dall'Iran e dal partito sciita di Hezbollah mettono in scena diversi attacchi informatici a favore dei palestinesi, prima, durante e dopo l'inizio della guerra. Sono attacchi distruttivi contro Israele ma anche operazioni di hack-and-leak contro entità israeliane e negli Stati Uniti, come pure campagne di phishing progettate per rubare informazioni e fare propaganda antisemita. Le notizie giornalistiche al riguardo vengono confermate da un rapporto di Microsoft riportato nel blog aziendale a febbraio 2024, "Da quando Hamas ha attaccato Israele nell'ottobre 2023, attori allineati al governo iraniano hanno lanciato una serie di attacchi informatici e operazioni di influenza (IO, Influence Operations, *nda*) mirate a sostenere la causa di Hamas e indebolire Israele, i suoi alleati politici e partner commerciali". E aggiunge: "Molte delle operazioni immediate dell'Iran dopo il 7 ottobre sono state frettolose e caotiche, a indicare uno scarso o nessun coordinamento con Hamas, ma hanno comunque ottenuto un successo crescente" (Microsoft Threat Intelligence, 2024). Tuttavia, nonostante le affermazioni degli attori coinvolti, molti "attacchi" nei primi giorni della guerra riguardano la pubblicazione di dati inattuali e falsi, e la stessa divulgazione di credenziali appare frutto di un accesso preesistente alle reti israeliane. La stessa notizia della loro divulgazione costituisce però un'operazione di influenza che sollecita l'attivazione progressiva di un network di attori, hacktivisti e nation state hacker, capaci di intervenire, in maniera opportunistica, e via via più coordinata, a favore di altri gruppi ideologicamente allineati.

Un altro rapporto, questa volta di Google, sostiene che l'Iran ha originato quasi l'80% di tutte le attività di phishing contro Israele nei sei mesi precedenti gli attacchi del 7 ottobre "Le operazioni di hack-and-leak rimangono una componente chiave negli sforzi degli attori delle minacce" (Joyce & Huntley, 2024). È lo stesso modus operandi degli attori in gioco nel conflitto russo-ucraino ma, a differenza di quest'ultimo, nel conflitto cibernetico tra Iran e Israele "le operazioni informatiche sembrano essere eseguite indipendentemente dalle

[84] Lo shadowbanning è una forma di censura che comporta il blocco, totale o parziale, dei contenuti di un utente da alcune aree geografica o all'interno di una comunità online. Mentre il bando diretto degli utenti è uno strumento visibile al pubblico, lo shadowbanning non è visibile. Gli utenti non sanno che la portata del loro account è stata limitata o del perché ciò sia avvenuto.

azioni cinetiche e di battaglia, in quanto tali capacità informatiche possono essere rapidamente implementate a un costo inferiore contro i rivali regionali senza un confronto militare diretto". Tra le varie tecniche usate, almeno due gruppi di hacktivisti di nome *Karma* e *Handala hack* sfruttano i ceppi di malware noti come *Bibi-Windows Wiper*, *Bibi-Linux Wiper* (Blackberry, 2023), che rimandano al soprannome del premier israeliano Netanyahu, Bibi, per operare attacchi distruttivi volti a eliminare rispettivamente i file di Windows e Linux nei sistemi informativi e nei servizi informatici di Israele[85]. In base alle analisi di Swascan su Bibi Windows: "L'artefatto, similarmente a quanto è avvenuto durante la guerra Russo-Ucraina, è stato utilizzato come strumento di guerra ibrida atto a effettuare azioni distruttive nei confronti delle infrastrutture critiche di Israele, contribuendo di fatto all'offensiva militare e strategica di Hamas. La minaccia, eseguendo una sovrascrittura e una fase di "locking" dei files (senza però chiederne alcun riscatto) pone BiBi Wiper in una condizione diversa da una minaccia Ransomware. L'unico obiettivo del Wiper è quello di rendere i dati dei target system non accessibili e non utilizzabili" (Swascan, 2023).

10.3 L'uso della propaganda nel conflitto palestinese

Le evidenze di un coinvolgimento di civili e hacktivisti nella lunga e sanguinosa guerra iniziata nell'ottobre 2023, appariranno con sempre maggiore chiarezza quando, dopo un anno di guerra combattuta sul terreno dalle Israeli Defence Forces (IDF) contro i miliziani palestinesi, il confronto si sposterà decisamente sui media e i social network, facendo larghissimo uso di fake news e propaganda.

In seguito all'invasione di terra nella striscia di Gaza, dopo aver distrutto la gran parte delle abitazioni e polverizzato la resistenza armata di Hamas e degli altri gruppi armati palestinesi coalizzati nell'attacco del 7 ottobre, Israele farà un largo uso di immagini, video e audio artefatti, anche ricorrendo all'Intelligenza Artificiale, per creare incertezza e dubbio nell'opinione pubblica e nel fronte politico che gli chiede di fermarsi. È la risposta alla diffusione dei video degli ostaggi israeliani in mano alle milizie palestinesi usati per fare pressione su Israele e contrastare le oceaniche manifestazioni dei cittadini di Tel Aviv che chiedono a gran voce il cessate il fuoco, il rilascio dei connazionali ostaggio di Hamas e le dimissioni del premier Netanyahu che considerano un criminale di guerra. Gli esempi sono diversi. Abbiamo già citato il video generato con l'Intelligenza Artificiale in cui appare la Striscia di Gaza trasformata in un resort di lusso con al centro una statua d'oro di Donald Trump. Ma ne sono stati realizzati altri.

[85] Si tratta di un wiper che distrugge tutti i file tranne quelli con estensione .exe, .dll e .sys, poiché questi tipi di file sono essenziali per il funzionamento del computer. Queste estensioni sono integrate nel codice per essere ignorate, in modo che il malware possa completare il suo percorso distruttivo.

In seguito al bando dell'Unrwa[86] a operare nella striscia di Gaza, in quanto ritenuta collusa con le azioni terroristiche di Hamas, il governo israeliano decide di affidare la distribuzione degli aiuti per la popolazione civile palestinese alla fondazione israelo-americana Gaza Humanitarian Foundation (GHF). Amnesty International ritiene che la decisione sia parte di una strategia per usare la fame come arma di guerra, la così detta "guerra delle calorie". La distribuzione di cibo alla popolazione affamata da parte della GHF si caratterizza da subito per una serie di attacchi armati verso i civili in attesa di ricevere gli aiuti alimentari (Amnesty International, 2025). Un fatto che desterà la riprovazione di intellettuali, governi, associazioni umanitarie e dell'opinione pubblica mondiale. Nel tentativo di contrastare questo fallimento diplomatico, sul canale Youtube del Ministero degli Esteri israeliano sarà pubblicato un video (Israel Foreign Affairs Ministry, 2025), generato dall'Intelligenza Artificiale, per ribaltare l'evidenza delle uccisioni di civili in cerca di cibo e creare una narrativa capace di dimostrare la bontà dell'azione di supporto alle vittime della guerra, usando tutti i canali commerciali possibili, compreso l'acquisto di pubblicità su Google, e invadere la Rete con contenuti allineati a questa narrativa (Fanpage, 2025).

A dimostrazione che l'uso dell'IA generativa non è affatto episodico nella guerra di propaganda portata da Israele, Gila Gamliel, ministra della Scienza e della Tecnologia nel governo Netanyahu, a fine luglio 2025 posta sulla piattaforma X un altro video generato con l'IA dove le macerie di Gaza diventano grattacieli, uno con la scritta Trump, in uno scenario pieno di giovani israeliani che si divertono nei locali sulla spiaggia di Gaza. Nel video compare anche il premier Netanyahu che passeggia con la consorte sul lungomare, ringiovanito di trent'anni, al pari di Trump, che pure vi compare con la first lady Melania. Anche qui musica, festa, e fuochi di artificio, stavolta senza i Palestinesi (Rainews24, 2025).

L'uso dell'IA a fini propagandistici è stato segnalato anche in Iran. In questo caso è Microsoft a raccontare in un suo rapporto come, all'inizio di dicembre 2023, l'Iran abbia interrotto la normale programmazione televisiva sostituendola con servizi basati su notizie false usando un presentatore generato dall'Intelligenza Artificiale. È la prima operazione di influenza iraniana rilevata da Microsoft in cui l'Intelligenza Artificiale ha svolto un ruolo chiave nella comunicazione della Repubblica Islamica, rappresentando un esempio della rapida e significativa espansione della portata delle operazioni iraniane dall'inizio del conflitto tra Israele e Hamas. L'interruzione ha raggiunto il pubblico negli Emirati Arabi Uniti, nel Regno Unito e in Canada (Watts, 2024).

[86] L'UNRWA, acronimo di Agenzia delle Nazioni Unite per il Soccorso e l'Occupazione dei Rifugiati Palestinesi nel Vicino Oriente, è un'agenzia sussidiaria dell'Assemblea Generale delle Nazioni Unite che fornisce assistenza e protezione ai rifugiati palestinesi.

La guerra di propaganda tra Israele e Hamas, basata sull'uso dei social media, ha tuttavia una lunga storia. Israele comincia a usare i social media nel 2008 durante l'ennesima operazione militare contro Hamas, denominata *Operazione Piombo fuso*. La decisione di usare i social da parte dello Stato ebraico nasce quando i suoi governanti si accorgono dell'errore di non far entrare i civili a Gaza, lasciando ai soli Palestinesi la testimonianza univoca degli orrori perpetrati in quell'occasione. Per questo, con l'apertura di un account Twitter e la pubblicazione di alcuni video su Youtube, Israele passa all'azione e recluta la *Jewish Internet Defence Force* per organizzare campagne di comunicazione contro la propaganda antisemita islamica. L'esercito israeliano organizzerà delle "sale operative Hasbara", veri e propri centri di propaganda dentro le università israeliane, per modificare perfino i contenuti delle pagine di Wikipedia. Il direttore dell'informazione israeliana in quell'occasione riconoscerà che i fronti della battaglia sono diventati tre: il campo di battaglia, il fronte cibernetico, e il mondo dei social media (Colon, 2024).

Un altro esempio di questa lunga "competizione social" è quello che è accaduto in Israele il 14 novembre 2012, quando il redattore responsabile delle piattaforme di social media in lingua inglese, il giovane riservista Daniel Rubenstein, pubblica un tweet sulla "vasta campagna contro obiettivi terroristici" lanciata dall'IDF. Due minuti dopo il caporale posta il video di un attacco aereo e dell'auto in fiamme di Al-Jabari, leader palestinese, totalizzando quasi cinque milioni di visualizzazioni. Infine, pubblica in un terzo tweet, in cui la grafica computerizzata spiega al mondo le ragioni dell'uccisione di Al-Jabari. Questo live tweeting dell'operazione viene considerato, in Israele e non solo, come un "punto di svolta nella comunicazione militare". Durante gli otto giorni dell'operazione *Pilastro di Difesa*, il nome dato all'intervento militare israeliano, i contendenti ingaggiano una guerra di hashtag su Twitter, che si risolve rapidamente a vantaggio di Hamas, il cui hashtag #gazaunderattack ("Gaza sotto attacco") è retwittato circa 2,6 milioni di volte, contro le 283.050 di #israelunderfire ("Israele sotto tiro"). L'IDF però analizza i social media in tempo reale per adattare le sue azioni sul campo. In uno studio condotto su centinaia di migliaia di tweet che confrontava le reazioni degli utenti con l'andamento del conflitto, un ricercatore americano ha scoperto che le azioni di Israele erano strettamente correlate all'umore prevalente sui social media. Un picco di simpatia online per Hamas, ad esempio, veniva immediatamente seguito da una riduzione del numero di attacchi aerei e da un aumento dei messaggi della propaganda israeliana.

I social media sono diventati così parte integrante delle operazioni israeliane: nel 20I3 la sezione media esteri e social dell'IDF contava una cinquantina di soldati multilingue, intorno ai 21 anni, la metà dei quali impiegati esclusivamente a postare testi sul web in inglese, francese, spagnolo e arabo, come nella Fabbrica di troll di San Pietroburgo. Il team,

composto soprattutto da donne, comprende grafici, videomaker e redattori. Il motivo è chiaro nelle parole del tenente colonnello David Lerner “Se non sei presente nello spazio dei social media stai cedendo quello spazio al nemico” (Colon, 2024).

Il confronto tra Israele e i suoi nemici continuerà a lungo su tutti i social, da Facebook a Twitter, dai notiziari online ai risultati dei motori di ricerca.

La disinformazione, parte essenziale di questo scontro è, da entrambe le parti, facilitata, anche nel 2023, dal divieto imposto da Israele alla stampa internazionale di entrare nella Striscia di Gaza, dove l’IDF continua senza sosta nella sua rappresaglia, uccidendo, secondo Reporters sans frontières, oltre 200 giornalisti (Reporters sans frontières, 2025). Un clima che faciliterà la moltiplicazione di storie ritenute da molti analisti senza fondamento, come quella degli stupri di guerra di Hamas e della decapitazione di neonati israeliani (Iannuzzi, 2024, pp 59-67). Le maglie della censura israeliana, tuttavia, continueranno a far trapelare video di soldati israeliani che invadono le case di Gaza e che, a favore di telecamera, documentano per i social la distruzione di arredi e oggetti personali dei palestinesi in fuga, spesso con montaggi rudimentali e colonne sonore in stile Apocalipse Now. I responsabili, individuati dagli utenti dei social network più diffusi con un’operazione di *crowdsourcing intelligence*, diventeranno il bersaglio di feroci critiche su piattaforme come TikTok che nel frattempo diffonde una serie di video, accompagnati da false dichiarazioni, per rivendicare interventi umanitari cinesi mai avvenuti, come la distribuzione di cibo e medicine da parte dell’aero flotta di Pechino (Langford, 2025).

Un altro esempio della capacità israeliana di usare i media a proprio favore aggirando i tradizionali meccanismi di verifica giornalistica, è costituita infine dalla produzione di prove artefatte dell’esistenza di una roccaforte di Hamas sotto l’ospedale al Shifa. Parliamo dell'analisi di decine di animazioni dell'esercito israeliano, utilizzate per giustificare gli attacchi a Gaza e amplificate da organi di stampa internazionali, provenienti non da servizi segreti come affermato da Israele, ma da servizi digitali, librerie commerciali e creatori di contenuti, perfino dalle animazioni in 3D di un museo marittimo scozzese. Secondo l’inchiesta che il magazine +972 ha svolto con fact checker, giornalisti ed esperti video tra Israele, Svizzera e Scozia, il video animato, pubblicato dall'esercito israeliano il 27 ottobre 2023 e messo online successivamente, per mostrare cosa si nascondesse sotto l'ospedale Al-Shifa, il più grande complesso medico di Gaza, è un falso, che era stato così presentato: “The Shifa Hospital is not only the largest hospital in Gaza but it also acts as the main headquarters for Hamas’ terrorist activity” (Israel Defence Forces, 2025). L’animazione mostra tunnel sotterranei, bunker e una sala di comando di Hamas, il tutto rappresentato

attraverso una grafica 3D accattivante. Tutti elementi presentati come prove schiaccianti: "Si basano sull'intelligence israeliana", aveva detto Mark Regev, allora consigliere senior del Primo Ministro Benjamin Netanyahu, durante un'intervista rilasciata lo stesso giorno alla CNN aggiungendo: "Queste informazioni sono inattaccabili". Non era così, come avrebbero dimostrato il magazine +972 Magazine e Local Call, sito israeliano di informazione indipendente, insieme al collettivo di ricerca Viewfinder, alla rete svizzera SRF e alla testata scozzese The Ferret, che hanno analizzato 43 animazioni prodotte dall'esercito israeliano dal 7 ottobre 2023. Il primo raid israeliano contro l'ospedale non sarebbe comunque avvenuto prima di metà novembre. Ma la narrazione era già stata definita. Il video è stato diffuso simultaneamente sugli account Telegram, Facebook, YouTube, X e Instagram dell'esercito. Sul profilo X di Netanyahu, ha finora ottenuto circa 50 milioni di visualizzazioni. Nelle settimane successive al suo rilascio, decine di testate internazionali lo avrebbero ritrasmesso per il proprio pubblico, immancabilmente accompagnato dall'affermazione di Israele che l'ospedale fungeva da "principale base operativa" di Hamas a Gaza. "Ma nessuna base del genere è mai stata scoperta. Inoltre, la sala di comando mostrata nel video non era un caso isolato; era già apparsa più di un anno prima in un'altra animazione pubblicata dall'esercito israeliano, che illustrava quello che si diceva fosse un tunnel sotto una scuola dell'Agenzia delle Nazioni Unite per il Soccorso e l'Occupazione (UNRWA) a Gaza". Le strade circostanti nel video di "Al-Shifa" erano popolate dalle vetrine di un pacchetto commerciale di risorse 3D, piene di locali fittizi come "Fabio's Pizzeria", "Andre's Bakery" e "Revolution Bike Shop".
L'animazione di Al-Shifa sarebbe diventata uno degli esempi più noti della nuova strategia comunicativa israeliana in tempo di guerra. "Ma ha anche segnato l'inizio di una fase accelerata di produzione all'interno dell'Unità portavoce dell'IDF: dopo aver pubblicato solo una manciata di visualizzazioni 3D prima del 7 ottobre, l'unità ha da allora diffuso decine di video simili che raffigurano presunti siti terroristici a Gaza, in Libano, in Siria e in Iran" (Ziv, 2025).

La campagna di disinformazione e propaganda di Israele, chiamata *Hasbara* (la spiegazione), si svolge parallelamente alla strategia degli omicidi mirati dei giornalisti nel tentativo di contenere i danni all'immagine di stato democratico con cui si presenta al mondo. Il Magazine israeliano +972 in un articolo a firma di Yuval Abraham porterà infatti alla luce l'esistenza di un'unità speciale, chiamata *Legitimation cell*, con il compito di scoprire, ed eventualmente creare le prove dei legami tra i giornalisti ancora presenti sul terreno e Hamas (Abraham, 2025). É il caso di Anas al Sharif, noto giornalista della tv del Qatar, Al Jazeera, ucciso con un "missile intelligente" alla vigilia di Ferragosto 2025 insieme ad altri quattro colleghi. Il governo israeliano rivendicherà l'omicidio sostenendo che il giornalista

era a capo di un'unità terroristica di Hamas dedita al lancio di missili. La tv per cui lavorava smentirà categoricamente. Il 24 agosto nei network internazionali circola la notizia dell'ennesimo omicidio mirato di un operatore dell'informazione nel conflitto in Palestina: il sindacato dei giornalisti palestinesi lo identifica come Khaled al-Madhoun e per la Ong Sheeren.ps è il numero 270 dei giornalisti eliminati dall'IDF. Qualche giorno dopo, Israele eliminerà altri cinque giornalisti in un doppio attacco, causando numerose reazioni diplomatiche e di condanna che non porteranno però alla fine delle uccisioni.

10.4 L'apporto di Big Tech alla guerra in Medioriente

"Corporate actors are deeply entwined in the system of occupation, apartheid and genocide in the occupied Palestinian territory," the Special Rapporteur said. "For decades, Israel's repression of Palestinian people has been scaffolded by corporations, fully aware of and yet indifferent to, decades of human rights violations and international crimes."
(United Nations, 2025)

Ma l'Intelligenza Artificiale non viene usata solo per video propagandistici: Israele ha già riversato nel cloud (Abraham, 2024a) enormi database di volti palestinesi che successivamente saranno usati per individuare presunti terroristi (Robinson-Erly, 2024). La tecnica, nota dai manuali dell'intelligence israeliana, è stata ancora una volta denunciata dal magazine di Tel Aviv +972 e da Local Call che, in un'inchiesta congiunta, riescono a dimostrare come l'esercito israeliano abbia segnalato decine di migliaia di cittadini di Gaza in quanto sospettati di omicidio, utilizzando un sistema di puntamento d'arma basato sull'Intelligenza Artificiale, con scarsa supervisione umana e una politica molto permissiva nell'individuare i futuri bersagli (Abraham, 2024b), un'accusa che troverà nuove conferme nel rapporto Onu (United Nations, Human Rights Council, 2025) della relatrice speciale per i territori palestinesi delle Nazioni Unite, Francesca Albanese, presentato il 30 giugno al Consiglio dei diritti umani delle Nazioni Unite (Piccolo, 2025).

Secondo il rapporto, nel 2021 Amazon e Alphabet hanno siglato un contratto da 1,2 miliardi di dollari con Tel Aviv per fornire spazi cloud e sistemi di elaborazione dati. Nel 2023 Microsoft ha invece 'soccorso' l'esercito israeliano che stava esaurendo i suoi spazi di archiviazione dati mettendo a disposizione le sue strutture. Anche la statunitense Palantir, che nel gennaio 2024 ha tenuto il suo consiglio di amministrazione a Tel Aviv "in segno di solidarietà", contribuisce a sviluppare per Israele i sistemi di Intelligenza Artificiale utilizzati per indirizzare le operazioni belliche su Gaza. Secondo diversi autori, "Gaza è un gigantesco laboratorio per testare sul campo sistemi carcerari e di sorveglianza avanzati. Sistemi di sorveglianza biometrica, sorveglianza attraverso droni, utilizzo di Intelligenza Artificiale e analisi dei dati per supportare le forze militari, e reti avanzate di check point. Per queste

tecnologie Tel Aviv si avvale di numerose società estere, a cominciare dalle statunitensi Ibm ed Hewlett Packard. Microsoft, in Israele da quasi 40 anni, fornisce sistemi e tecnologie per la sorveglianza nelle colonie occupate" (Remocontro, 2025).

Microsoft nega, ma una successiva inchiesta di The Guardian, +972 e Local call ha dimostrato che, grazie alla capacità di archiviazione quasi illimitata di Azure, il cloud di Microsoft, l'Unità d'élite cyber israeliana 8200 ha creato un nuovo potente strumento di sorveglianza di massa: un sistema intrusivo che raccoglie e archivia le registrazioni di milioni di chiamate telefoniche effettuate ogni giorno dai palestinesi a Gaza e in Cisgiordania. Microsoft ha assicurato che il Ceo Satya Nadella non era a conoscenza del tipo di dati che Unit 8200 prevedeva di archiviare nel cloud. Ma una serie di documenti Microsoft precedentemente trapelati, e diverse interviste con numerose fonti sia interne all'azienda che nell'intelligence militare israeliana, hanno rivelato la vera natura del progetto. Secondo i documenti raccolti dal Guardian, a luglio 2025 già 11.500 terabyte di dati militari israeliani - equivalenti a circa 200 milioni di ore di audio - erano conservati nei server Azure di Microsoft nei Paesi Bassi e in Irlanda (Davies & Abraham, 2025). Il 25 settembre 2025, Microsoft annuncia la decisione di negare l'accesso a questi database telefonici all'unità 8200 dell'IDF dimostrando la fondatezza delle denunce di attivisti e giornalisti.

Prima di questi accadimenti, Israele sarà protagonista di un complesso attacco alla filiera di produzione e approvvigionamento di prodotti tecnologici da parte di Hezbollah, sfruttando la manomissione dei cercapersone, i pager, usati dai suoi miliziani in Libano per coordinare il lancio di missili che hanno ripetutamente colpito il nord di Israele. Tra il 17 e il 18 settembre 2024 i pager saranno fatti esplodere per mettere fuori gioco i miliziani capaci di manovrarli colpendo anche tutti coloro che gli erano vicini, ferendoli o uccidendoli (CNN, 2024).

Come ci sono riusciti? Diversi analisti ritengono che tutto sia basato sull'IA e che un ruolo centrale l'abbia giocato Palantir, azienda Usa di IA e data processing. Il giornalista Michael Steinberger nel libro "The Philosopher in the Valley", una dettagliata biografia di Alex Karp, attuale Ceo di Palantir Technologies, sostiene che i sistemi di IA prodotti dalla sua azienda siano stati utilizzati da Israele durante le operazioni del 2024 in Siria e in Libano, inclusa quella che il libro chiama *Operazione Grim Beeper*, in cui i pager furono innescati con trappole esplosive (Albanese & Giangiulio, 2025). Palantir sviluppa software per l'integrazione dei dati: la sua tecnologia assimila enormi quantità di informazioni e identifica rapidamente modelli, tendenze e connessioni di una complessità ingestibile per gli analisti umani. Fondata nel 2003 per supportare il governo degli Stati Uniti nella guerra al terrorismo, inizialmente finanziata dalla CIA, Palantir è oggi un colosso globale da 400 miliardi di dollari, e il suo software è utilizzato dai principali servizi segreti (canadesi,

francesi, israeliani, tra cui il Mossad), dall'esercito statunitense, dal FBI, dalla NSA e dall'ICE, oltre che da note aziende belliche, la Lockheed Martin, ad esempio, ma anche da imprese italiane, come Stellantis e Ferrari e finanche da una clinica universitaria come il Policlinico Gemelli di Roma.

Già nel libro *Cybermania* di Eviatar Matania e Amir Rapaport (2022) si parla di come le operazioni di intelligence mirate a individuare specifici bersagli ormai usino i dati raccolti da satelliti, droni e sistemi di sorveglianza filtrati da algoritmi. Il sistema "Habsora" (in ebraico "the Gospel") sfrutta centinaia di algoritmi per individuare potenziali obiettivi tra i dati accumulati in un enorme bacino digitale chiamato "the pool" (Pisa, 2024). Gli algoritmi setacciano intercettazioni, foto satellitari e post sui social network per segnalare coordinate di presunte strutture sotterranee, tunnel o depositi di armi. "Usando il riconoscimento delle immagini del software, i soldati possono scovare minuscoli cambiamenti in anni di riprese satellitari di Gaza che suggeriscono come Hamas abbia piazzato un lanciarazzi o scavato un nuovo tunnel su terreni agricoli" scrive il *Washington Post* sulla base delle rivelazioni di un ex dirigente militare che ha lavorato a questi sistemi di Intelligenza Artificiale (The Washington Post, 2024). Altri programmi, come "Lavender", utilizzano punteggi in percentuale per stimare la probabilità che una persona appartenga a gruppi armati. Elementi come la presenza in determinate chat o l'uso frequente di più linee telefoniche possono alzare il livello di sospetto. Applicazioni come "Hunter" e "Flow", invece, consentono ai soldati israeliani sul campo di battaglia di accedere a dati in tempo reale, inclusi video real-time delle zone a cui si avvicinano e perfino le stime su possibili vittime civili. Questi sistemi si interfacciano con "Gospel", potenziando l'intero processo di acquisizione degli obiettivi.
Come ha ricostruito Tech Policy Press, pochi mesi dopo che la GHF è stata istituita, nel febbraio 2025, per gestire la distribuzione degli aiuti a Gaza "sono emersi piani per rendere obbligatorio l'uso di strumenti di riconoscimento facciale nei siti di distribuzione, portando i gruppi per i diritti umani a condannare questo ipotetico meccanismo di controllo e sorveglianza basato sulla "biometria in cambio di cibo". Poco tempo dopo, a luglio, un appaltatore della GHF ha confermato l'uso di telecamere per il riconoscimento facciale nei siti di distribuzione. Questa militarizzazione della biometria e del riconoscimento facciale segna un'ulteriore espansione dell'uso dell'intelligenza artificiale (IA) nell'occupazione della Palestina" (Fathallah, 2025).

L'intelligenza artificiale è stata parte integrante della strategia di Israele come potenza militare occupante. Per anni, l'intelligenza artificiale è stata alla base di molti sistemi di sorveglianza e guerra, tra cui fucili d'assalto che trasferiscono coordinate, sistemi sensore-

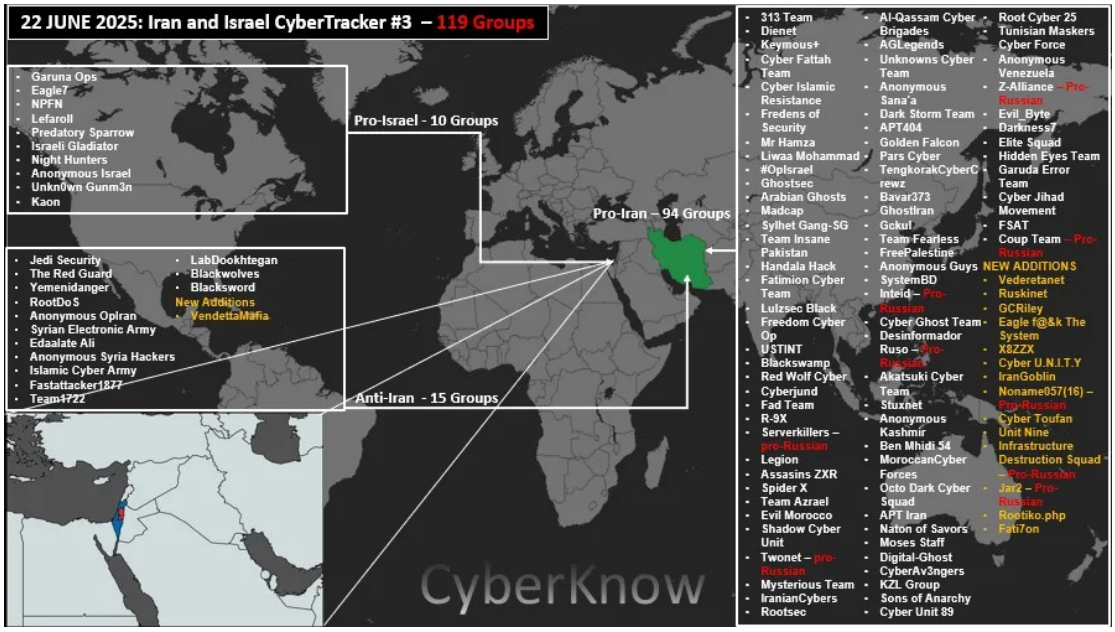
tiratore che individuano i bersagli, veicoli blindati progettati per manovrare nel modo più ottimale e droni che colpiscono obbiettivi umani. Due anni prima dell'assalto del 2023, l'esercito israeliano aveva già definito la sua operazione militare a Gaza del 2021 come "la prima guerra di intelligenza artificiale", annunciando l'utilizzo di Gospel che successivamente è stato usato anche per generare elenchi di edifici da colpire con attacchi aerei.

C'è tuttavia un altro modo in cui Big Tech supporta il genocidio dei Gazawi: limitando la sovranità digitale delle entità internazionali che usano i suoi software. Nel caso di Israele e della Corte penale internazionale rendendo inaccessibile la posta elettronica dei procuratori della Corte come conseguenza di un executive order del *Commander in chief*, Donald Trump[87]. Il 6 febbraio 2025 il presidente USA Donald Trump ha infatti imposto una serie di sanzioni alla Corte penale internazionale per le indagini su personale statunitense e su alcuni alleati, incluso Israele. In applicazione di tale decisione Microsoft ha disabilitato l'accesso all'account di posta elettronica del procuratore capo della Corte penale internazionale che il 21 novembre 2024 aveva spiccato un mandato d'arresto internazionale contro Bibi Netanyahu e il suo ex ministro della Difesa Yoav Gallant.

10.5 La Cyberguerra Israele-Iran

Il 12 giugno 2025 Israele lancia un attacco aereo verso l'Iran che, secondo l'Agenzia internazionale per l'energia atomica, Aiea, sta velocemente giungendo a una piena capacità nucleare. L'Operazione, denominata Rising Lion, è accompagnata da una serie di attacchi informatici per tutta la prima fase del conflitto. Attacchi DDoS, defacciamenti, esfiltrazione di dati sensibili e di informazioni riservate. Si creano due fronti di hacktivisti: 89 a favore dell'Iran, 10 a favore di Israele e 15 anti-iraniani, per un totale di 119 gruppi, secondo le stime di CyberKnow (2025), l'azienda australiana specializzata in Osint e cyber threat intelligence.

[87] Monti, A. (2025). Così gli USA controllano la sovranità digitale delle istituzioni internazionali e della UE. La Repubblica. Disponibile in: https://www.repubblica.it/tecnologia/blog/lettere/2025/07/13/news/cosi_gli_usa_controllano_la_sovranita_digitale_delle_istituzioni_internazionali_e_della_ue-424728328/ [13 luglio 2025]



Cyberknow: cybertracker 22 giugno 2025

Anche in occasione di questo conflitto, condotto con armi convenzionali, e facendo uso di omicidi mirati che uccidono alti funzionari iraniani, compreso il capo dell'intelligence, comandante della forza Quds, Saeed Izadi, lo schema si ripete: gli attacchi cinetici, e cioè prima i bombardamenti israeliani su siti sensibili iraniani e poi, di rimando, i lanci di missili balistici dall'Iran contro Israele, sono accompagnati dagli attacchi cibernetici e dalla disinformazione. Quello che pare un dominio separato, il mondo cyber, anche in questo caso ha una prosecuzione nel mondo fisico.

La propaganda viene subito attivata sui due fronti del conflitto. I media riportano un attacco alla banca iraniana Sepah, ma l'agenzia di stampa statale Irna sosterrà che le transazioni nella repubblica islamica non ne hanno sofferto. Ugualmente sarà diffusa la notizia di un attacco informatico al Cryptoexchange iraniano *Nobitex* con l'effetto di sottrargli circa 90 milioni di dollari di cryptovalute. L'Exchange dirama una comunicazione in cui informa la clientela che ogni asset sarà rimborsato.

Le autorità della Repubblica Islamica dell'Iran sin dal 13 giugno decidono massicce restrizioni dell'accesso a Internet, il traffico si riduce all'80%. Ai funzionari del regime viene raccomandato di interrompere l'utilizzo di qualsiasi dispositivo connesso, anche di WhatsApp, per evitare di essere geolocalizzati e diventare un target degli israeliani, tecnica usata dagli stessi cybersoldati iraniani per identificare e colpire una base militare israeliana.

Dall'inizio del conflitto si registrano diversi tentativi di interruzione del funzionamento di infrastrutture critiche nei settori energetico e delle telecomunicazioni, con attacchi a centrali

elettriche, raffinerie e impianti petrolchimici. I gruppi statuali e degli attivisti cercano infatti di infiltrarsi in dighe, aeroporti e centrali energetiche sfruttando vulnerabilità nei sistemi di controllo industriale (Industrial Control Systems, ICS, e SCADA, Supervisory Control & Data Acquisition) per causare blackout o disservizi, intercettare dati e compromettere la sicurezza delle comunicazioni militari e civili.

Gli iraniani, per i quali la cyberwarfare è parte della dottrina militare della *soft war*, vedono schierati al proprio fianco alcune cybergang come *Handala*, da sempre impegnata in attacchi ransomware, che stavolta però non appare interessata all'ottenimento di un riscatto ma a creare caos e incertezza nel cyberspace israeliano assumendo un profilo hacktivista.

Durante gli attacchi missilistici contro le città di Tel Aviv, Haifa, Ber 'Sheva, vengono inoltre condotti attacchi DDoS mirati ai siti web delle stazioni radio israeliane per creare confusione e ostacolare la diffusione degli alert di allarme. In seguito, vengono divulgate notizie di attacchi contro centri di ricerca nucleare e militare, con diffusione di malware per il furto di informazioni. A farne le spese, il centro di ricerca Weizmann, come parte di una campagna di phishing avente come obbiettivo istituzioni accademiche e del settore israeliano della difesa, azione motivata dal coinvolgimento delle università israeliane nel sistema militare e di sicurezza del Paese. Ed è infatti proprio il gruppo pro-pal *Handala Hack* che il 18 giugno 2025 annuncia una fuga di dati di 425 GB dall'azienda israeliana Mor Logistics e l'ottenimento dell'accesso a 4 TB di documenti classificati del Weizmann Institute of Science, colpito da un attacco missilistico iraniano il giorno prima (Daily DarkWeb, 2025).

Nel canale Telegram *APTIran* i gestori rivendicano gli attacchi contro Israele come ritorsione per i bombardamenti subiti. In aggiunta, il gruppo diffonde sui suoi canali Telegram una serie di informazioni relative agli attacchi a servizi e infrastrutture iraniane e, in un post significativo, fornisce consigli ai potenziali target iraniani ricordando che, in un "teatro di guerra digitale", "l'utilizzo di tecnologie non prodotte da vendor affidabili rappresenta un rischio diretto per le infrastrutture critiche del Paese [...] in quanto ogni componente importata o sviluppata da soggetti esterni può diventare uno strumento di intrusione, controllo o sabotaggio da parte del nemico" (Red Hot Cyber, 2025). Il gruppo mette in guardia i connazionali dalla possibile presenza di backdoor nella tecnologia in uso nel Paese illustrando bene uno dei rischi centrali alla sovranità digitale.

Terminata la "guerra dei dodici giorni" gli attacchi cibernetici e le operazioni di influenza non smettono. L'azienda israeliana Check Point Software individua una campagna di phishing potenziata con l'Intelligenza Artificiale da parte di attori iraniani, così detti APT35

(Lakshmanan, 2025), che, a partire da metà giugno 2025, ha preso di mira cittadini israeliani utilizzando false e-mail e messaggi WhatsApp personalizzati, redatti con strumenti di Intelligenza Artificiale, come suggeriscono, secondo i ricercatori, il layout strutturato e l'assenza di errori grammaticali. Al target della campagna, esperti israeliani di Intelligenza Artificiale, veniva paradossalmente chiesto supporto per un sistema di rilevamento delle minacce basato sull'IA, proprio per contrastare l'ondata di attacchi informatici che aveva preso di mira il loro paese, Israele, a partire dal 12 giugno.

Successivamente, alla fine della guerra, nei primi di agosto 2025, la società Security Scorecard decide di rilasciare pubblicamente un rapporto in cui viene chiarito come hanno operato gli attori filoiraniani delle minacce, noti e meno noti, durante il conflitto dei dodici giorni. STRIKE, il gruppo di threat intelligence di Security Scorecard ha analizzato infatti oltre 250.000 messaggi provenienti da 178 gruppi attivi, e ha potuto in tal modo rilevare una campagna digitale altamente coordinata che rispecchiava le azioni militari sul campo. L'analisi condotta ha individuato tre principali categorie di attori:

1) hacktivisti vagamente affiliati che operano senza supervisione diretta ma allineati con le priorità del Corpo delle Guardie della Rivoluzione Islamica (IRGC);

2) cluster strutturati allineati all'IRGC che eseguono campagne mirate;

3) gruppi interamente sponsorizzati dallo Stato come Imperial Kitten (noto anche come Tortoiseshell, Cuboid Sandstorm o Yellow Liderc).

Queste entità, concentrate su settori ad alto valore, tra cui istituzioni finanziarie, agenzie governative e organi di informazione, utilizzano attacchi di SQL injection, DDoS e metodi di esfiltrazione dei dati per la raccolta di informazioni e l'interruzione delle comunicazioni per danneggiare gli avversari. Operazioni che prevedono anche tecniche di ricognizione, analisi delle vulnerabilità per exploit zero-day e distribuzione di script malware personalizzati, il tutto programmato per coincidere con attacchi aerei e incursioni al confine (Security Scorecard, 2025).

In base alle analisi della società SOCRadar (2025) invece, il conflitto Iran-Israele del 2025 ha portato a un'impennata dell'attività informatica, con oltre 600 segnalazioni di attacchi informatici su oltre 100 canali Telegram tra il 12 e il 27 giugno 2025. Gli hacker filoiraniani non avrebbero colpito solo Israele, il paese maggiormente preso di mira con 441 segnalazioni di attacchi, ma anche Stati Uniti (69), India (34) e nazioni mediorientali come Giordania (33) e Arabia Saudita (13).

I principali gruppi di hacktivisti in quei giorni sarebbero stati Mr. Hamza, Keymous, Mysterious Team, Team Fearless, GARUDA_ERROR_SYSTEM, Dark Storm Team, Arabian

Ghosts, Cyber Fattah, CYBER U.N.I.T.Y e NoName057 (16). Anche in questo caso le alleanze tradizionali, ad esempio con attori russi, si consolidano nel cybersapce. "Dall'inizio della guerra, hacker sponsorizzati dallo Stato, hacktivisti di entrambi i Paesi e cybercriminali di nazioni non partecipanti, dall'Asia meridionale alla Russia fino al Medio Oriente, sono diventati attivi", ha affermato la società di threat intelligence. "Israele è stato il principale obiettivo degli attacchi DDoS, con 357 segnalazioni, che rappresentano il 74% di tutta l'attività DDoS".

Evidenziando l'impennata dell'attività degli hacktivisti durante il conflitto, Lidia López Sanz, ricercatrice di Outpost24 KrakenLabs, ha affermato che oltre 80 diversi gruppi di hacktivisti stanno "conducendo o supportando attivamente" operazioni informatiche offensive contro Israele e i suoi alleati, aggiungendo che presunte entità “faketiviste” come Cyber Av3ngers, Handala e Predatory Sparrow, gruppo filo-israeliano, probabilmente operano con il supporto dello Stato o direttamente sotto la sua direzione (López Sanz, 2025). Tra i collettivi di hacktivisti che hanno espresso solidarietà all'Iran ci sono DieNet, Mysterious Team Bangladesh, Team Insane Pakistan, Z-Alliance, Server Killers, Akatsuki Cyber Team, GhostSec, Keymous+, Inteid, Anonymous Kashmir e Mr Hamza Cyber Force.

"Il drammatico aumento delle operazioni informatiche degli hacktivisti a seguito delle recenti escalation geopolitiche tra Israele e Iran sottolinea il ruolo sempre più centrale che il conflitto informatico gioca nella guerra moderna", ha detto. "Gli hacktivisti motivati da ideologie, insieme ai possibili *faketivisti* degli stati nazionali, hanno chiaramente dimostrato la loro disponibilità a sfruttare le tensioni geopolitiche per perseguire diversi obiettivi strategici". La ricercatrice, senior threat intelligence analyst di Outpost, li chiama “faketivisti” proprio per chiarire che si tratta di attori statali che vestono a piacimento i panni degli hacktivisti per consentire ai governi che fanno la guerra cibernetica usando dei proxy di non subire immediate ritorsioni negando l’attribuzione statale degli attacchi per mettersi al riparo da risposte distruttive coordinate nel cyberspace (The Hacker News, 2025). Nel caso della guerra dei dodici giorni non è bastato.

11. Conclusioni

Il 18 novembre 2025 il Ministro italiano della Difesa, Guido Crosetto, diffonde un paper sui rischi della minaccia ibrida: cyberattacchi, disinformazione, strozzature nella supply chain di forniture critiche. Il paper afferma senza mezzi termini che la sovranità digitale dell’Italia è a rischio.

Il 19 novembre 2025 l'Italia firma a Berlino, insieme agli altri Stati membri dell'Unione europea, la *Declaration for European Digital Sovereignty*. L'obiettivo è consentire all'UE di controllare infrastrutture digitali, dati e tecnologie secondo le proprie norme e valori, evitando dipendenze critiche da attori esterni e garantendo allo stesso tempo la cooperazione con partner affidabili (European Council, 2025).

Il 7 dicembre 2025 Elon Musk chiude d'imperio l'account X della Commissione UE come rappresaglia a una multa comminata dagli organismi europei alla sua azienda.

Il 12 dicembre 2025 Il governo tedesco convoca l'ambasciatore russo al ministero degli Esteri accusando la Federazione russa di un grave attacco informatico al controllo del traffico aereo e di una costante campagna di disinformazione durante la campagna elettorale federale, oltre che di interferenze nelle elezioni.

Il 18 dicembre la notizia di un falso colpo di Stato in Francia, annunciato da un video generato con l'Intelligenza Artificiale (IA) e diffuso su Facebook, suscita scalpore a livello globale raggiungendo circa 13 milioni di visualizzazioni in poche ore. La richiesta formale di rimozione avanzata dal governo francese viene respinta dalla piattaforma perché il contenuto, pur palesemente falso, non violerebbe gli standard interni.

La notte di capodanno 2025, La Poste, il principale servizio postale francese subisce un potente attacco informatico, successivo a quello rivendicato da NoName(057)16 in precedenza il 22 dicembre, bloccando il tracciamento delle spedizioni, i suoi siti web e la stessa banca postale.

Al primo gennaio 2026 circa 130 mila personalità hanno firmato una dichiarazione pubblica in cui si oppongono alla creazione di una IA super-intelligente prima della definizione di un robusto quadro regolatorio, e auspicando una moratoria del suo sviluppo in assenza di prove concrete che dimostrino di poterla controllare. La dichiarazione è stata sottoscritta tra gli altri, dagli inventori del deep machine learning, i "creatori" cioè della moderna IA, Joshua Bengio e Geoffrey Hinton, e poi Yuval Harari, Paolo Benanti, Susan Rice, Max Tegmark, Audrey Tang, Stuart Russel, Steve Bannon, Lawrence Lessig e molti altri.

Questi avvenimenti rappresentano il riscontro più recente alle ipotesi che hanno guidato la nostra tesi. E cioè che la sovranità digitale dei singoli paesi e quindi l'autonomia, la stabilità e l'indipendenza dei singoli Stati è messa a rischio da cyberattacchi, campagne di disinformazione, tensioni geopolitiche e strapotere delle Big Tech in un contesto di guerra ibrida in cui l'IA rischia di sfuggire al controllo dei suoi stessi produttori. Ma questi eventi mostrano anche la mutazione dell'Internet delle origini da strumento di connessione tra gli individui a strumento che può potenziare dinamiche di guerra e di ricatto.

Nel corso della nostra ricerca abbiamo provato a spiegare come l'Internet aperta e collaborativa delle origini si sia successivamente strutturata attorno a concentrazioni di potere che ne contraddicono profondamente la natura distribuita e neutrale. Un numero ristretto di attori, prevalentemente privati e geograficamente localizzati, controlla porzioni decisive dell'infrastruttura, dei servizi, dei dati e perfino degli standard. Tuttavia, come spiega Luca Sambucci nella nostra intervista "le spinte verso una sovranità digitale nazionale, se ben orientate e coordinate fra Paesi, possono fungere da contrappeso a questi cluster di potere", promuovendo una distribuzione più democratica delle risorse digitali.

Significherebbe impostare una sovranità digitale che incentivi lo sviluppo di infrastrutture aperte e trasparenti, che sostenga l'interoperabilità dei sistemi e rilanci la decentralizzazione come principio architetturale e politico. É questo il punto di vista espresso da altri intervistati come Barbara Carfagna e Michele Mezza. Secondo la giornalista il modello introdotto da alcuni paesi arabi e caucasici, basato su una forte regolamentazione, localizzazione del dato e diversificazione dei fornitori è la direzione da seguire; secondo il professore Michele Mezza invece il modello vincente è la contrattazione sulla conoscibilità del dato e la trasparenza del suo utilizzo unite all'alfabetizzazione degli utenti.

Di fronte agli interessi di Big Tech però, come sostiene il professore Colajanni, tutto questo è assai difficile e forse siamo fuori tempo. Possibilista è invece il diplomatico Massimo Marotti che ritiene sempre possibile una negoziazione, necessaria per non soccombere al potere delle piattaforme e garantire l'autonomia decisionale delle singole nazioni.

Mentre, come abbiamo spiegato, l'Europa ha tracciato il suo *Nomos der Erde* facendo della regolamentazione la linea rossa della sovranità digitale, grazie a leggi orientate a tutelare dati (GDPR, DSA, DMA, AI ACT, Data Act) e infrastrutture (NIS, CRA) da ingerenze non controllabili, a impedire il percorso verso una sovranità digitale dove siano ben bilanciati gli interessi pubblici e privati, è intervenuta una guerra dell'informazione che quando non manipola e inquina i dati, distrugge le infrastrutture necessarie al loro legittimo utilizzo. É la cyberwar, elemento centrale della guerra ibrida che oggi è combattuta su diversi fronti, in particolare nei due conflitti presi ad oggetto del nostro esame: la guerra in Medioriente scoppiata nel 2023 e la guerra nel Donbass iniziata nel 2022.

Dalle nostre analisi emerge che, come sanno gli strateghi militari, ogni guerra può essere considerata ibrida. Gli attori dei conflitti usano da sempre tutti gli strumenti a loro disposizione per prevalere sull'avversario, dalle armi convenzionali al terrorismo, dalle interferenze economiche all'hacker warfare fino alla guerra cognitiva, cioè tecniche di manipolazione delle percezioni, basate su propaganda e disinformazione.

In seguito alla progressiva dipendenza da software e algoritmi, con lo sviluppo esponenziale dell'intelligenza artificiale e la corsa agli armamenti digitali, siamo entrati in una nuova era dei conflitti, siamo nell'era della guerra algoritmica. Una guerra che mette a rischio la sovranità digitale e quindi il benessere e l'incolumità stessa dei cittadini. La Rete è diventata spazio e strumento di conflitto aperto a cui partecipano anche i civili.

Con l'avvento di Internet e della progressiva digitalizzazione delle attività umane infatti, tutti i soggetti, civili e militari, singoli e organizzati, possono partecipare in maniera aperta, consapevole e mercenaria ai conflitti usando sia mezzi non militari sia tecniche e strumenti di tipo militare. Si pensi all'importanza delle telecomunicazioni spaziali per la vita quotidiana, dai commerci all'informazione alla difesa: ebbene, finora, nel conflitto russo-ucraino ci sono stati circa 170 attacchi satellitari, in quello tra Israele-Hamas e l'Iran gli attacchi sono stati 117.

La guerra non è più soltanto affare di eserciti e i moderni conflitti in cui sia i cittadini che le aziende di telecomunicazioni e le piattaforme digitali prendono parte, lo rende palese.

Nella nostra analisi abbiamo osservato che agli attacchi cinetici si accompagnano regolarmente gli attacchi cibernetici e che entrambi sono preceduti e seguiti da campagne di disinformazione.

Un elemento di assoluto rilievo in questo contesto è rappresentato proprio dall'uso di algoritmi e IA generative, elementi centrali della disinformazione e dell'automazione di tecniche e tattiche di guerra. Il coinvolgimento di attori non statali, il loro utilizzo strategico da parte dei governi, gli attacchi alle infrastrutture civili, condotti anche attraverso ransomware gangs, per colpire la *supply chain* delle aziende dei paesi belligeranti, hanno l'obbiettivo sia di interferire con la produzione manifatturiera sia con l'erogazione di servizi essenziali, mettendo a rischio la sovranità digitale e trasformando Internet in una trincea di guerra. Amaro preludio della fine dell'utopia di un mondo pacifico perché iperconnesso e interdipendente grazie alla Rete.

Un altro elemento che abbiamo osservato durante il nostro lavoro di ricerca riguarda la partecipazione dei civili alle operazioni cibernetiche assumendo il ruolo di operatori e hub di informazione per le parti in guerra, mentre gli attivisti digitali vengono reclutati all'interno di strutture organizzate di tipo paramilitare. La visibilità delle loro azioni e le ricompense economiche ottenute si aggiungono alle motivazioni politico ideologiche che li spingono a partecipare ai conflitti.

Uno dei fenomeni più significativi dei vari conflitti che si sono verificati negli ultimi anni può infatti essere identificato nella militarizzazione dell'attivismo politico all'interno del cyberspazio che unisce tecniche di hacking e disinformazione. Per circa trent'anni,

l'hacktivism, con le sue incursioni nel mondo della comunicazione digitale ha rappresentato un modo per rivendicare il proprio antagonismo, individuale e collettivo, senza porre rischi significativi alle organizzazioni colpite pur provocando danni variamente quantificabili. Oggi, alla metà degli anni '20 del nuovo secolo, essendo diventato più organizzato, strutturato e sofisticato, l'hacktivism ha inaugurato un nuova fase evolutiva entrando nelle guerre guerreggiate. Solo nel conflitto Russo-Ucraino si contano oltre cento gruppi di hacktivisti che parteggiano per l'una o per l'altra parte in conflitto. Lo stesso accade negli scenari di guerra in Medio Oriente e, parzialmente, in Europa e Stati Uniti. I numeri di questa presenza aumentano costantemente grazie alle caratteristiche proprie dell'uso del cyberspace come teatro di scontro e strumento di conflitto: accessibile a chiunque, a-territoriale, impalpabile. Poiché molti gruppi di hacktivisti hanno un'agenda politica legata agli Stati, questi ultimi sono interessati a supportarli in maniera sempre più rilevante sia in tempo di pace che in tempo di guerra.

É in questo scenario che si afferma la guerra cognitiva.

Ogni società ricorre a delle narrazioni per fare progredire i gruppi sociali che la compongono verso mete utili alla collettività. Il senso e la direzione di queste narrazioni, politiche, sociali e religiose, cambia nei secoli, ed è in genere indifferente alla nozione di verità, concetto mobile e sfuggente per definizione. La creazione del consenso intorno a queste narrazioni si basa su storie condivise e la loro forza dipende dall'innesco di meccanismi psicologici. Questi principi sono manipolati costantemente da specifici attori. La propaganda è una forma di narrazione e condivisione di storie collettive, mentre la disinformazione si basa su distorsioni narrative, bias psicologici e tecnologie persuasive. Con l'avvento del digitale e dei social network è più facile propagandare narrazioni vere, false, o inventate. Possono essere automatizzate, elicitano risposte rapide, non sono facilmente verificabili. Gli attori della disinformazione lo sanno, e mescolano sapientemente il vero con il falso per far progredire la propria agenda politica ed economica.

È l'apoteosi dei servizi di intelligence che operano secondo la logica delle misure attive, l'insieme di strumenti volti a manipolare la percezione di un target, per portare il loro attacco, l'attacco alla mente.

Lo scenario descritto nel lavoro di ricerca lascia poco spazio all'ottimismo, eppure non è possibile rassegnarsi. Mentre è necessario approfondire le tematiche trattate e metterle costantemente alla prova dei fatti, nuove direzioni di ricerca possono essere sviluppate.

Il primo dicembre 2022 la Svezia ha creato la prima agenzia governativa al mondo per la difesa psicologica con l'obbiettivo di contrastare le influenze esterne al processo di formazione delle decisioni dei propri cittadini e salvaguardare così il processo politico elettorale su cui si basa la democrazia liberale rappresentativa. Favorire la conoscenza dei meccanismi di manipolazione delle percezioni è la strategia da seguire. Irrobustire i meccanismi di contrasto alla disinformazione con iniziative dedicate a cui partecipino i conglomerati mediali è un altro possibile fattore di intervento. Obbiettivo sarà quello di contrastare la disinformazione sul nascere in una logica di pre-bunking, e non solo di fact-checking, senza rinunciare a investire nell'informazione di qualità.

Il governo italiano dal canto suo sta discutendo in Parlamento una serie di misure di contrasto alla minaccia ibrida e il Ministro Crosetto propone finanche la creazione di un "Centro di contrasto alla minaccia ibrida". In questo caso vanno indagate le possibilità di realizzare una difesa proattiva con la partecipazione di enti e istituzioni in una prospettiva di "protezione civile cibernetica" che si affianchi alle realtà già esistenti che si occupano di resilienza informatica come l'Agenzia per la Cybersicurezza Nazionale. Investire in formazione e consapevolezza, ad ogni livello, dalla scuola all'università, con la partecipazione congiunta del sistema dei media e il supporto delle istituzioni, è una delle opzioni da esplorare.

E questo perché, come dice Marco Ramilli nella sua testimonianza, "Non possiamo pensare di esercitare sovranità su ciò che non comprendiamo. Serve educare cittadini, lavoratori, imprese e decisori pubblici a riconoscere le potenzialità ma anche i rischi del digitale – dai bias algoritmici agli attacchi cibernetici – rendendo la società più consapevole, resiliente e capace di scegliere con cognizione di causa".

L'Unione Europea è avvertita dei pericoli descritti e, mentre non si lascia intimidire dalle rappresaglie delle Big Tech, continua a sostenere lo sviluppo di tecnologie quantistiche, dei supercomputer, delle startup tecnologiche e dell'Intelligenza Artificiale come volano di crescita e benessere. Ma l'Europa potrebbe anche far contare maggiormente il suo peso economico e negoziare alleanze strategiche per l'approvvigionamento di materie prime, la cooperazione scientifica e tecnologica, lo sviluppo di nuovi mercati, anche con paesi finora esclusi dalla cooperazione civile, militare e industriale nel quadro delle attuali alleanze.

Infine, se la logica dell'impiego della forza potrà difficilmente essere contrastata con i mezzi della diplomazia a causa della rinnovata aggressività politica e militare di Usa, Cina, Iran, Russia e Israele, ci sono ancora diverse strade da percorrere sia per la ricerca che per la società civile in modo da informare scelte politiche all'altezza dei tempi che viviamo. Intanto chiedere sempre il rispetto del Diritto internazionale, secondo cui i civili, i feriti e i più fragili

non possono essere obiettivo di guerra, neanche di quella informatica. Secondo, giungere a una regolamentazione efficace dell'impiego dell'Intelligenza Artificiale, attraverso lo sforzo collettivo di tutte le nazioni, per scongiurare la definitiva perdita di controllo su una tecnologia che, se male utilizzata, è in grado di cancellare la vita umana così come la conosciamo. Infine, sostenere gli sforzi del dialogo così come viene tentato presso i consessi internazionali, ad esempio presso l'Internet governance forum delle Nazioni Unite a cui partecipano tutti gli stakeholder della Rete, cittadini, associazioni, aziende e governi, e che potrebbe trasformarsi in un consesso decisionale e non solo consultivo. La ricerca in questo caso dovrà riguardare l'esplorazione di soluzioni condivise per tenere insieme le esigenze di autogoverno digitale e l'apertura e la collaborazione fra gli Stati per sviluppare una Rete sicura, aperta e resiliente, che sia strumento di democrazia e di sviluppo.

Ultimo ma non ultimo, sostenere le iniziative di collaborazione internazionale fra tutti i paesi già impegnati nella protezione delle infrastrutture critiche, sanità, trasporti, energia, telecomunicazioni, finanza, sistemi elettorali, all'interno di un mondo che è senza confini, quello della Rete, dove gli effetti di un attacco informatico possono dilagare verso luoghi imprevisti; facilitare il coordinamento normativo e repressivo contro le estorsioni informatiche (il ransomware), allargando il numero dei paesi partner della Counter Ransomware initiative, che oggi sono 61; lavorare concordemente allo sviluppo sicuro dell'intelligenza artificiale secondo i dettami della Bletchley Park Declaration a cui già partecipano tutti i paesi europei e del Commonwealth. Queste tre direttrici di intervento, già perseguite all'interno del G7 Cybersecurity Working Group, voluto dall'Italia, potrebbero essere adottate da altri paesi progressivamente associabili agli sforzi in atto, superando anche le tradizionali appartenenze e le alleanze geopolitiche del secolo scorso.

Il compito della ricerca potrebbe essere proprio quello di dimostrare che esistono alternative all'impiego dell'hard power, storicizzandone gli esiti catastrofici, e anticipando gli esiti di scelte potenzialmente sbagliate. Per garantire la pace e la democrazia in un mondo complesso e in costante evoluzione, la logica della forza e il dialogo devono necessariamente trovare un equilibrio.

12. Le interviste

Per mettere alla prova le ipotesi di questo lavoro abbiamo intervistato alcuni tra i maggiori esperti italiani di innovazione digitale, cybersecurity e Intelligenza Artificiale. Abbiamo pertanto posto sei domande aperte a docenti universitari (Michele Colajanni), saggisti (Michele Mezza), giornalisti (Barbara Carfagna), imprenditori (Luca Sambucci e Marco Ramilli), e figure istituzionali (Massimo Marotti), chiedendogli cosa pensano delle minacce

alla sovranità digitale, del ruolo della disinformazione nei processi democratici, della possibilità di bilanciare la sovranità digitale con la sicurezza, l'apertura e la neutralità delle Rete.

La selezione degli interlocutori intervistati è stata determinata dalla scelta di alcuni criteri: la conoscenza approfondita dei temi affrontati nel percorso di tesi; la numerosità delle pubblicazioni realizzate, sia di quelle scientifiche che di quelle divulgative; l'impegno didattico e pedagogico, nei settori dell'educazione, dell'istruzione e della divulgazione mediale. I testimoni scelti sono a vario titolo coinvolti in percorsi accademici e di ricerca e sono ben noti all'interno del dibattito pubblico, i loro contributi compaiono frequentemente sulla carta stampata, alla radio e in televisione, e sono considerati degli influencer del settore tecnologico in virtù dei riscontri ottenuti dalla loro presenza sui social media come LinkedIn, Facebook, X, e YouTube.

I loro interventi e le loro riflessioni hanno generato negli anni consapevolezza e conoscenza sia presso i decisori pubblici, essendo sentiti presso le commissioni parlamentari competenti, e presenti nelle iniziative istituzionali di Camera e Senato, sia presso il pubblico generico attraverso l'utilizzo di media online e social media.

Pur nella diversità delle opinioni, nelle loro interviste si individua un pensiero affine, che seppure non consenta una sintesi immediata, si fonda su alcuni punti di vista prevalenti. Tutti ritengono la sovranità digitale necessaria, uno scopo da perseguire per il benessere del nostro paese, l'Italia e per l'Europa; tutti, tranne uno, il professore Michele Colajanni, ritengono che sia possibile perseguire delle politiche di sovranità digitale basate sulla regolamentazione nazionale ed europea; nella maggioranza ritengono che sovranità digitale e apertura della Rete possano convivere; tutti gli intervistati ritengono che lo sviluppo della cybersecurity rappresenti il fondamento di altre forme di sicurezza, economica, sociale e politica. Infine, ritengono, unanimemente, che la guerra algoritmica sia una realtà del nostro tempo da indagare e con cui è necessario confrontarsi in ambito sociale e istituzionale.

Michele Mezza

Intervista a Michele Mezza, giornalista, per lungo tempo impegnato in Rai prima come inviato e poi come responsabile di progetti di digitalizzazione delle news. Docente alla Federico II di Napoli e autore di diverse pubblicazioni sul tema della negoziazione sociale degli algoritmi e della cybersecurity come categoria dell'informazione.

Professore Mezza, che cos'è secondo lei la sovranità digitale?

È un concetto non meno fluido di quanto non stiano diventando tutte le tecnicalità digitali, che ormai stanno evolvendo ad un ritmo sempre più frenetico. In questa trasformazione si intravvede come tendenza principale la spinta al decentramento degli accessi e delle personalizzazioni dei dispositivi, in particolare nel campo delle nuove intelligenze artificiali. Per tanto direi che per uno stato la sovranità digitale è la capacità di permettere ai propri cittadini di essere autonomi e indipendenti proprio nelle fasi di selezione e personalizzazione dei meccanismi digitali, riducendo la subalternità rispetto alla proprietà degli stessi meccanismi alla pura fase di impostazione del sistema. Come recitava l'ultimo aforisma di Carl Schmitt, il padre del concetto di autorità statale, uno stato è sovrano se controlla insieme al potere di dichiarare lo stato di emergenza anche le dinamiche delle onde elettromagnetiche, che nel 1985, anno in cui Schmitt, prima di morire, concepì l'aforisma, erano il sinonimo delle nove forme di relazioni tecnologiche.

La sovranità digitale è un rischio oppure un'opportunità per il buon funzionamento di Internet?

La sovranità è la matrice del pluralismo computazionale, e come tale, è il motore di una differenziazione dei linguaggi e delle applicazioni che la storia di Internet ci dice essere la vera origine della sua straordinaria esperienza. Io penso che oggi il problema della rete, in particolare dell'ultimo miglio di questa innovazione, che è appunto il fenomeno delle intelligenze artificiali, sia un'ancora ridotto margine di negoziazione sociale, cioè di capacità da parte di soggetti quali ad esempio le comunità territoriali o le categorie professionali, di contrattare con i titolari dei sistema procedure e modelli di adattamento e personalizzazione, introducendo forme di evoluzione e sviluppo di ulteriore creatività. In questa logica rivendicare sovranità maggiori significa animare questa dialettica propositiva.

Quali sono secondo lei i rischi per i singoli Stati determinati dall'assenza di sovranità digitale?

Potremmo parlare di una sorta di neocolonialismo digitale, che da una parte creerebbe pericolose servitù tecnologiche, culturali e civili, con un dominio assoluto da parte dei centri proprietari, ma dall'altro, come stiamo per altro verificando in questi anni, comporterebbe un rallentamento dei processi di innovazione che sarebbero frenati dai monopoli di fatto che si stanno realizzando. Se per motore di ricerca si intende solo Google, o per intelligenza artificiale prevalentemente ChatGPT è ovvio che siamo sull'orlo di un feudalesimo tecnologico. Mi pare utilissimo il suggerimento venuto nel giugno del 2025 direttamente da Jensen Huang, CEO di Nvidia, la principale azienda produttrice di chip GNU ad alto potenziale per l'intelligenza artificiale, che ha sollecitato ogni soggetto, pubblico o privato, a cimentarsi

nella progettazione di proprie intelligenze, dando una spinta alle forme di innovazione e rendendosi sempre più autonomi.

La guerra algoritmica portata da alcuni attori nel cyberspace minaccia le funzioni basilari degli stati e come?

Dobbiamo constatare che ormai quella che si definisce guerra algoritmica è una condizione relazionale sia a livello geopolitico che a quello interno di ogni singolo Stato, che tende ormai a sostituire strutturalmente lo stato di pace. La cosiddetta guerra ibrida, che prevede forme di interferenza sia materiale, con il tentativo di mano mettere documenti e di bloccare sistemi informatici, sia immateriale, con la manomissione del senso comune di un paese o di una comunità, nel processo di formazione di una propria opinione pubblica e reputazionale, è una condizione che tende a diventare endemica. In questa guerra ibrida, lo spiegava tempo fa il Capo di Stato Maggiore della federazione russa generale Gerasimov, l'obbiettivo è rendere lo stato avversario meno reattivo e capace di assicurare ai propri cittadini l'autonomia cognitiva nella lettura e decifrazione degli eventi, rendendo lo stesso stato vulnerabile nella contrapposizione informazionale.

É possibile garantire libertà e apertura della Rete senza rinunciare alla sovranità digitale?

Io credo che la sovranità aumenti e non limiti la libertà in rete. Certo il concetto di libertà non deve risolversi in un arbitrio da parte dei più forti. Un esempio ci può aiutare: imporre la trasparenza per gli autori di messaggi, ossia di rendere subito identificabile se un post è stato elaborato da un essere umano o da una batteria di bot, questo alza la soglia di libertà, perché permette ad ognuno di interpretare correttamente i messaggi che gli arrivano. Si tratta di ridurre i margini di totale discrezionalità che i proprietari di tecniche e di competenze avanzate hanno accumulato in rete a danno dei comuni cittadini. Come diceva un pedagogo dell'80, Jean batiste Lacordaire, fra un forte e un debole la legge libera e la libertà opprime.

Quale può essere la risposta al rischio di attacchi cibernetici, alla disinformazione, all'abuso dell'intelligenza artificiale?

Secondo la regola che un algoritmo può essere contrastato solo da un altro algoritmo dovremmo elaborare piattaforme e soluzioni in grado di rilevare la manipolazione. Io non penso che il vero problema sia la palese falsificazione di informazioni, che in rete durano poco, quanto la proliferazione di canali individuali che sulla base di profilazioni arbitrarie degli utenti sono in grado di suggestionare e alterare la percezione critica di ogni individuo. Come ha certificato un centro di ricerca di Zurigo, un sistema di profilazione che può

esprimersi con lo stesso linguaggio del suo interlocutore ha una potenza di convincimento sette volte superiore a quella di un essere umano. Questo gioco va reso evidente e non ordinario, mediante infrastrutture pubbliche.

Michele Colajanni

INTERVISTA a **Michele Colajanni**, Professore ordinario presso il "Dipartimento di Informatica: Scienza e Ingegneria" dell'Università di Bologna e titolare dei corsi di Cybersecurity e di Scalable and reliable services delle lauree magistrali di Ingegneria Informatica e di Informatica.

Che cos'è secondo lei la sovranità digitale.

Ci sono molteplici livelli da considerare e molteplici misure adottabili o meno: tecnologico (reti e server), dati, servizi, persone. Dal punto di vista tecnologico, la sovranità digitale è un miraggio politico irrealizzabile da dare in pasto a sovranisti ignoranti. La tecnologia digitale, e non solo, nasce come aggregazione di centinaia di componenti. Nessun singolo Paese, neanche la Cina né gli Stati Uniti e tantomeno l'Europa, è in grado di essere sovrano e autonomo su tutte le tecnologie che compongono un moderno servizio digitale. Diverso è il discorso dei servizi digitali che sono completamente nelle mani di aziende americane e cinesi. Diverso ancora è il dominio dei dati. Come si possa garantire il sovranismo europeo sui dati nel momento in cui i servizi che elaborano tali dati non sono nelle mani di aziende europee è una contraddizione che si è voluta "risolvere" con la localizzazione dei server in territorio europeo su cui preferirei non esprimere commenti. Le persone sono l'obiettivo fondamentale di tutti gli attacchi delle campagne disinformative esterne e interne che si alimenta con l'ignoranza. La protezione nascerebbe da un livello di cultura maggiore, non dalla tecnologia.

La sovranità digitale è un rischio oppure un'opportunità per il buon funzionamento di Internet?

Dopo cinquant'anni di sviluppo di Internet con spirito aperto, interoperabile e collaborativo, la completa sovranità digitale è irrealizzabile. Tuttavia, sono già in corso molteplici tentativi di balcanizzazione di Internet (dal firewall cinese alla Runet sovrana russa fino al progetto DNS4EU e all'Ambizioso progetto Digital India). Sono iniziative politiche che riflettono un clima di sfiducia tra gli attori principali ma che, come tutte le balcanizzazioni, non porteranno

a niente di buono. I Balcani sono prossimi e spero che non abbiamo dimenticato gli eventi degli anni '90. La situazione potrebbe sempre virare verso un'auspicabile razionalità e collaborazione, ma per ora la direzione appare chiara e i suoni di scudi e tamburi risuonano in lontananza.

Quali sono secondo lei i rischi per i singoli Stati determinati dall'assenza di sovranità digitale?

Il digitale è un dominio fondamentale a tutti i quattro livelli di cui sopra. I Paesi che non lo controllano sono destinati a soccombere. Se ci limitiamo ai Paesi europei, non controlliamo la tecnologia digitale, i servizi né i dati. Potremmo migliorare la sovranità dei nostri pensieri promuovendo la cultura in senso lato, non la cultura digitale. Le sembra che vi sia un'iniziativa politica e progetti seri in tale direzione? La mia impressione è che la direzione sia opposta e non solo a causa dei politici recenti. Anche Trump è il prodotto, non la causa.

La guerra algoritmica portata da alcuni attori nel cyberspace minaccia le funzioni basilari degli stati e come?

Questo è evidente. Gli attori si muovono essenzialmente con due obiettivi, talvolta sovrapposti: politici ed economici. Entrambe le categorie utilizzano gli stessi mezzi: profilazione sempre più pervasiva degli utenti, soddisfazione dei desideri (indotti, non reali), assuefazione e "bombardamento informativo". Il sistema sta funzionando benissimo anche grazie all'AI che rappresenta uno strumento potente per tali scopi. Ci vuole molta competenza, cultura, tempo da dedicare per capire e difendersi; purtroppo, queste caratteristiche non rappresentano la maggioranza. E le minoranze, da sempre, sono destinate a soccombere.

É possibile garantire libertà e apertura della Rete senza rinunciare alla sovranità digitale?

È la domanda da un milione di dollari a cui rispondo con un'altra domanda. Ma lei è proprio sicuro che con una Rete chiusa e sovrana, circoscritta a ciascun Paese, la situazione potrebbe migliorare per i cittadini oppure il controllo e la disinformazione potrebbe prosperare senza limiti? L'esempio più adeguato in tal senso è rappresentato dalla Cina che non userei come faro di libertà e democrazia. Partiamo dagli obiettivi: se privilegiamo libertà e democrazia, massima apertura e collaborazione.

Quale può essere la risposta al rischio di attacchi cibernetici, alla disinformazione, all'abuso dell'intelligenza artificiale?

La risposta a questa domanda vale 100 milioni di dollari! Tanti obiettivi, molto diversi tra loro.

I rischi di attacchi cibernetici non si riducono con il sovranismo. Si possono limitare gli effetti concentrando le difese sulle priorità di un Paese. Se si passa da qualche centinaio a decine di migliaia di obiettivi, disperdiamo le nostre forze e le possibilità di difenderci (ogni riferimento alla NIS2 non è causale...).

La disinformazione si combatte solo con la cultura e con lo stimolo al pensiero critico.

L'abuso dell'AI, in realtà, parte dall'abuso che subiamo sui dati digitali. I modelli sono disponibili e utilizzabili, talvolta open, ma ci mancano i dati. Non è facile, forse è impossibile, recuperare il gap costituito da cinquant'anni di dati acquisiti dai nostri PC, vent'anni di dati acquisiti dai nostri smartphone e da qualche anno di dati acquisiti dagli oggetti smart che installiamo in casa o indossiamo. Quale potrebbe essere una vera contromisura sovranista? Creare uno smartphone interamente europeo dotato di elettronica, software e app EU, e obbligare tutti i cittadini all'uso? Se le sembra un'ipotesi irrealizzabile, non capisco perché si continui a parlare seriamente di sovranismo digitale europeo. Chiacchiere, chiacchiere vane, senza una concreta ipotesi di messa a terra di un progetto. Se si passasse dalle chiacchiere a un'ipotesi progettuale, se ne vedrebbe l'assurdità, come la mia idea provocatoria di uno "smartphone total EU". Questa situazione non cambierà perché le chiacchiere colpiscono meglio l'immaginario e non hanno bisogno di essere veramente realizzate.

Massimo Marotti

INTERVISTA a **Massimo Marotti**, già ambasciatore italiano in Iraq e in Libano, è Direttore del Servizio Strategia e Cooperazione dell'Agenzia per la cybersicurezza nazionale, ACN.

Che cos'è secondo lei la sovranità digitale.

Per uno Stato che fonda il proprio potere su istituzioni democratiche, la sovranità digitale è la capacità di contenere e compensare influenze e controlli estranei alla comunità nazionale di cui esso è espressione sulle tecnologie trasformative che investono l'uso, la custodia ed il flusso dei dati.

La sovranità digitale è un rischio oppure un'opportunità per il buon funzionamento di Internet?

È una forma di assicurazione contro i rischi estremi dell'impiego di Internet.

Quali sono secondo lei i rischi per i singoli Stati determinati dall'assenza di sovranità digitale?

Dipende dal contesto e dagli scenari su scala globale. In un mondo ad economia integrata l'assenza di sovranità digitale per uno stato medio introduce una ulteriore forma di dipendenza nell'esercizio delle proprie funzioni ed in quello delle attività umane della propria comunità.

La guerra algoritmica portata da alcuni attori nel cyberspace minaccia le funzioni basilari degli stati e come?

Compromette la sicurezza economica e apre la porta a nuove forme di dipendenza da Stati ostili e no.

É possibile garantire libertà e apertura della Rete senza rinunciare alla sovranità digitale

Dove l'acqua dolce incontra l'acqua salata, ha detto Floridi, crescono solo le mangrovie. Libertà e apertura della rete spero siano le nostre mangrovie.

Quale può essere la risposta al rischio di attacchi cibernetici, alla disinformazione, all'abuso dell'intelligenza artificiale?

La conoscenza e lo sviluppo delle capacità cognitive. E molti soldi.

Barbara Carfagna

INTERVISTA a **Barbara Carfagna**, è giornalista, conduttrice del TG1, autrice di "Codice" su Raiuno e del Podcast "Codice Beta" per Radiouno. È autrice di reportage internazionali e documentari sull'impatto del digitale, la cybersicurezza, la politica e l'innovazione. Editorialista de IlSole24Ore. Ha lavorato in Africa, Asia, Europa e Medio Oriente, collaborando anche con importanti testate italiane. È membro dell'Aspen Institute, moderatrice di conferenze internazionali e docente in corsi di formazione per manager, forze armate e pubblica amministrazione.

Che cos'è secondo lei la sovranità digitale.

La sovranità digitale è la capacità di uno Stato di imporre la propria architettura su dati, infrastrutture, tecnologie critiche e piattaforme. Significa poter decidere *come* e *da chi* si raccolgono, si processano e governano le informazioni: una condizione necessaria per tutelare diritti come privacy, libertà di espressione e sicurezza, oggi ridefiniti nelle interfacce delle grandi piattaforme. Quindi fuori dalla possibilità della maggior parte degli Stati.

In assenza della possibilità di ottenere una sovranità digitale completa (senza tecnologie proprietarie) ho trovato interessante, come giornalista, esplorare due Paesi che utilizzano tecnologie sia americane che cinesi. Entrambi si sono concentrati sulla gestione del dato. Arabia Saudita[88] e Azerbaijan[89].

Pur utilizzando infrastrutture tecnologiche esterne (USA, Cina, Europa), questi paesi cercano di mantenere il controllo sulla regolamentazione interna:

a. Hosting locale obbligatorio dei dati: i dati dei cittadini e delle PA devono risiedere fisicamente sul suolo nazionale (es. *Saudi Cloud First Policy*),

b. Restrizioni sull'accesso esterno ai dati e licenze locali obbligatorie per i provider globali.

L'Azerbaigian, ad esempio, ha avviato centri nazionali di cybersecurity e sistemi cloud pubblici gestiti da agenzie statali o semi-statali, anche se basati su software esteri.

In assenza di tecnologia proprietaria, questi paesi diversificano i fornitori per evitare una dipendenza eccessiva da un solo blocco geopolitico:

- L'Arabia Saudita lavora sia con aziende americane (Google Cloud, Oracle, IBM), sia con Huawei (Cina), sia con startup israeliane in ambito cybersecurity (non ufficialmente).

- L'Azerbaigian collabora con fornitori europei, turchi, israeliani e russi, mantenendo un equilibrio tra interessi geopolitici ad hoc

L'Arabia Saudita ha fondato entità come SDAIA (Saudi Data and Artificial Intelligence Authority), che agiscono come filtri strategici tra tecnologia straniera e governance nazionale.

[88] https://www.youtube.com/watch?v=fsSEfY20Fnk&t=27s

[89] https://www.youtube.com/watch?v=qVKBqyahQ6A&t=238s

L'Azerbaigian ha istituito strutture come l'IDDA (Agenzia per lo Sviluppo Digitale) o AzInTelecom, per garantire che anche tecnologie estere siano implementate secondo priorità nazionali, spesso attraverso adattamenti locali, auditing e localizzazione linguistica.

Anche senza tecnologie proprietarie, questi paesi controllano l'accesso ai servizi digitali pubblici attraverso:

- Sistemi di identità digitale proprietari (es. Nafath e Absher in Arabia Saudita, sistemi integrati con ASAN in Azerbaigian).
- E-government e portali unificati: tutti i servizi pubblici passano attraverso strutture governative nazionali, anche se ospitati su cloud stranieri.

Entrambi i paesi stanno cercando di:

- Formare talenti locali per ridurre la dipendenza esterna nel medio-lungo periodo.
- Attrarre centri R&D stranieri in loco con incentivi fiscali, come fa la Saudi Digital Academy o l'Azerbaijan Cybersecurity Academy in collaborazione con Technion (Israele).

La sovranità digitale è un rischio oppure un'opportunità per il buon funzionamento di Internet?

Un'opportunità, tranne quando serve a giustificare atteggiamenti autoritari. In particolare, oggi l'opportunità è data dalle gestione autonoma dei dati strategici e della sicurezza delle infrastrutture e dalla costruzione di capacità tecnologiche autonome soprattutto in ambito militare o civile nel campo del controllo delle filiere produttive per evitare episodi come quello dei pager in Libano.

Quali sono secondo lei i rischi per i singoli Stati determinati dall'assenza di sovranità digitale?

La perdita del controllo sui dati (e quindi la cessione delle capacità di influenza) dei cittadini e delle istituzioni apre alle manipolazioni e quindi alla perdita di sicurezza nazionale, oltre che di competitività economica.

La guerra algoritmica portata da alcuni attori nel cyberspace minaccia le funzioni basilari degli stati e come?

Certamente. In tutto: possibilità di attacchi cyber a infrastrutture; disinformazione, attacchi a singoli individui competenti in materie sensibili o con capacità utili in guerra. Tutto potenziato (o attutito) dalla mancata o implementata formazione della cittadinanza in alcuni Paesi.

É possibile garantire libertà e apertura della Rete senza rinunciare alla sovranità digitale?

È la sfida europea. Non siamo ancora in grado di dire se stia funzionando, anche perché non conosciamo l'evoluzione degli agenti AI, come opereranno e quanto sarà realmente possibile interferire

Quale può essere la risposta al rischio di attacchi cibernetici, alla disinformazione, all'abuso dell'intelligenza artificiale?

Senz'altro la formazione, la consapevolezza. Poi sanzioni efficaci, controllo delle filiere produttive, fiducia, se ben riposta, nella cooperazione internazionale, controllo sui dati da parte di chi non possiede tecnologie proprietarie.

Marco Ramilli

INTERVISTA a **Marco Ramilli**, imprenditore nel campo della cybersecurity e dell'Intelligenza Artificiale. Ramilli è anche ricercatore, dottorato all'Università di San Diego negli Usa e collabora a ricerche e progetti con l'Università di Bologna. Ha creato IdentifAI, una startup innovativa che secondo il NIST è la prima di 100 aziende a livello mondiale nel settore del contrasto ai Deepfake.

Che cos'è secondo lei la sovranità digitale.

Per me, la sovranità digitale significa la capacità di uno Stato, e in particolare dell'Italia, di orientare consapevolmente le proprie risorse economiche, umane e normative per promuovere lo sviluppo di un ecosistema digitale nazionale. Non si tratta di illudersi di poter sostituire infrastrutture tecnologiche già esistenti, costruite prevalentemente da grandi attori internazionali, ma di comprendere che il digitale è un settore in continua evoluzione. Ogni giorno nascono nuovi paradigmi, tecnologie, piattaforme. La vera sfida, e l'opportunità, sta nel costruire le condizioni per cui la prossima infrastruttura fondamentale – che sia un'intelligenza artificiale, una rete di servizi digitali o una nuova architettura di connettività – possa nascere in Italia. Per farlo, serve investire nei talenti, sostenere la nascita e la crescita

di nuove organizzazioni, pubbliche e private, e creare un ambiente normativo e culturale che favorisca l'innovazione. Solo così potremo esercitare una vera sovranità nel mondo digitale: non chiudendoci, ma guidando con visione e coraggio il digitale che verrà.

La sovranità digitale è un rischio oppure un'opportunità per il buon funzionamento di Internet?

La sovranità digitale rappresenta una straordinaria opportunità per proiettare l'Italia verso un futuro in cui innovazione e identità nazionale crescano insieme. Investire in infrastrutture e imprese locali nel settore digitale non significa solo costruire tecnologia, ma dare forma a un ecosistema in grado di valorizzare le nostre eccellenze più autentiche. L'Italia ha una combinazione unica di elementi: un patrimonio umano fatto di talenti creativi e tecnici di altissimo livello, un territorio ricco di paesaggi straordinari, una cultura enogastronomica riconosciuta a livello mondiale e un'eccellenza storica nella meccanica e nel manifatturiero. Il digitale può diventare il "collante" che unisce tutto questo, mettendo in rete competenze, risorse e territori, e offrendo un vantaggio competitivo che pochi altri Paesi possono anche solo immaginare. Non si tratta solo di usare la tecnologia, ma di plasmarla a partire da ciò che rende l'Italia unica: se coltiviamo questa visione, possiamo davvero costruire un modello italiano di sviluppo digitale, innovativo e profondamente radicato nella nostra identità culturale e produttiva.

Quali sono secondo lei i rischi per i singoli Stati determinati dall'assenza di sovranità digitale?

L'assenza di una strategia di sovranità digitale comporta rischi significativi, che si possono raggruppare in tre grandi ambiti: la dipendenza economica, la vulnerabilità della sicurezza nazionale e il mancato sviluppo economico:

- Dipendenza economica da stati terzi

Se non investiamo nella costruzione di un ecosistema digitale autonomo, rischiamo di lasciare una quota crescente di valore economico – in termini di ricchezza, posti di lavoro, capacità imprenditoriale e know-how – nelle mani di attori stranieri. Questo significa non solo perdere controllo su un settore strategico in continua espansione, ma anche rinunciare alla possibilità di orientare lo sviluppo economico del nostro Paese secondo le nostre priorità e il nostro modello sociale.

- Vulnerabilità della sicurezza nazionale:

Tutte le infrastrutture critiche del Paese – dalla sanità alla pubblica amministrazione, dalle reti energetiche ai sistemi militari – si basano oggi su componenti digitali. Affidare lo sviluppo, la gestione e l'innovazione di questi sistemi a fornitori esterni, spesso legati ad altri interessi geopolitici, espone l'Italia a un rischio strutturale: quello di non poter garantire pienamente la propria autonomia operativa, né la riservatezza e la resilienza dei propri dati e delle proprie attività strategiche.

- Mancato sviluppo economico interno:

Infine, senza un'adeguata sovranità digitale, perdiamo l'opportunità di far nascere e crescere imprese locali capaci di competere a livello globale. Questo rallenta la nascita di nuovi poli tecnologici, impoverisce l'ecosistema dell'innovazione e rende il Paese meno attrattivo per i giovani talenti e per gli investimenti di lungo periodo.

In sintesi, non coltivare una vera sovranità digitale non è una scelta neutra: significa cedere potere, influenza e opportunità. Al contrario, adottare una visione strategica in questo ambito è oggi una delle leve più importanti per assicurare indipendenza, prosperità e sicurezza al nostro Paese.

La guerra algoritmica portata da alcuni attori nel cyberspace minaccia le funzioni basilari degli stati e come?

La mancanza di sovranità digitale rappresenta una minaccia diretta alla stabilità, alla continuità e all'affidabilità dei servizi essenziali su cui si basa il funzionamento del Paese. Non si tratta soltanto di protezione dei dati o di difesa da attacchi informatici, ma di garantire che settori vitali – come la sanità, le infrastrutture logistiche, le reti energetiche o i sistemi di decisione pubblica – possano operare in modo autonomo, sicuro e senza interruzioni. In un contesto in cui gran parte di questi servizi si appoggia a piattaforme digitali complesse, spesso sviluppate o gestite da fornitori esterni, il rischio non è solo quello di una violazione, ma anche di un'interruzione o di una degradazione del servizio. L'indisponibilità, anche temporanea, di un sistema sanitario digitale, di un'infrastruttura di trasporto intelligente o di una piattaforma di supporto alle decisioni pubbliche può avere conseguenze gravi e immediate per la vita dei cittadini e per la capacità dello Stato di operare efficacemente.

É possibile garantire libertà e apertura della Rete senza rinunciare alla sovranità digitale?

Assolutamente sì: la sovranità digitale è pienamente compatibile con una rete libera e aperta. Anzi, se ben gestita, può rafforzarla. È importante chiarire che la "rete" – intesa

come infrastruttura di connessione e comunicazione – è solo una componente del mondo digitale. Il digitale, oggi, include molto di più: esperienze digitali complesse, sistemi autonomi, software evoluti, intelligenza artificiale, robotica, applicazioni innovative, e soprattutto nuovi modelli di impresa, nuove economie e nuove forme di lavoro. Sovranità digitale non significa chiudere, limitare o controllare la rete, ma garantire che ciò che si costruisce sopra di essa – dai dati alle applicazioni, dalle piattaforme ai modelli d'uso – sia in linea con i valori, gli interessi e la visione strategica del Paese. Possiamo (e dobbiamo) mantenere aperto l'accesso alla rete, assicurare la neutralità e la trasparenza del traffico digitale, mentre allo stesso tempo sviluppiamo e governiamo in autonomia ciò che riguarda la gestione dei dati, la proprietà delle tecnologie, l'etica delle applicazioni e l'equità dei modelli economici. In altre parole, la vera sfida è costruire sovranità non contro l'apertura, ma dentro l'apertura: avere le competenze, le imprese, le infrastrutture e le visioni capaci di partecipare alla rete globale da protagonisti, e non da semplici utilizzatori passivi.

Quale può essere la risposta al rischio di attacchi cibernetici, alla disinformazione, all'abuso dell'intelligenza artificiale?

Costruire una vera sovranità digitale richiede un'azione articolata su più fronti, nessuno dei quali può essere trascurato. È una sfida sistemica che deve tenere insieme cultura, diritti e impresa:

1. Cultura digitale e consapevolezza collettiva

   Il primo passo è creare una cultura diffusa del digitale e dell'intelligenza artificiale. Non possiamo pensare di esercitare sovranità su ciò che non comprendiamo. Serve educare cittadini, lavoratori, imprese e decisori pubblici a riconoscere le potenzialità ma anche i rischi del digitale – dai bias algoritmici agli attacchi cibernetici – rendendo a società più consapevole, resiliente e capace di scegliere con cognizione di causa.

2. Una "costituzione digitale" per i diritti fondamentali

   Accanto alla cultura serve una solida cornice normativa. Occorre definire chiaramente i diritti fondamentali del cittadino nello spazio digitale: dalla proprietà dei dati alla trasparenza degli algoritmi, dal diritto alla disconnessione fino all'accessibilità e all'inclusione. Una sorta di carta costituzionale digitale che protegga le libertà individuali anche nel contesto virtuale e guidi lo sviluppo tecnologico nel rispetto della dignità umana.

3. Un tessuto produttivo capace di innovare.

4. Infine, la sovranità digitale non può esistere senza un ecosistema economico in grado di progettare, costruire e gestire tecnologie avanzate. Serve investire nella crescita di imprese locali capaci di innovare, creare valore e competere in ambito internazionale. Non basta usare le tecnologie degli altri: dobbiamo essere in grado di svilupparle, adattarle e, quando serve, guidarne l'evoluzione.

In sintesi, la sovranità digitale non è un obiettivo singolo, ma un equilibrio dinamico tra conoscenza, diritti e capacità produttiva. Solo con un'azione coerente su questi tre assi potremo davvero affermare una presenza italiana autorevole e indipendente nel futuro digitale.

Luca Sambucci

INTERVISTA a Luca Sambucci, imprenditore nel campo dell'IA, founder di Noctive Security, una startup di AI Security. Dal 2019 pubblica Notizie.AI, primo sito italiano sull'AI.

Che cos'è secondo lei la sovranità digitale.

Si tratta del nuovo terreno di gioco di una partita che si disputa da millenni. Controllare aspetti strategici del proprio ambiente al fine di garantirsi autonomia decisionale, cercando allo stesso tempo di non trasformare questo controllo in autarchia. L'equilibrio, normalmente fra controllo e partecipazione decisionale, è sempre una questione di punti di vista e, come qualsiasi equilibrio, solo temporaneo.

La sovranità digitale implica il controllo dei diversi strati che intrecciano il ciclo di vita di un dato. L'infrastruttura, con cavi e data center, l'hardware, con chip e sensori, il software, con le piattaforme di gestione e gli algoritmi che elaborano il dato trasformandolo in informazione e quindi in decisione.

Eppure, la sovranità digitale non si esaurisce nella mera capacità tecnica di possedere o localizzare questi elementi entro i confini nazionali. Si estende alla possibilità di esercitare un controllo normativo, politico ed economico su di essi. Significa, in altre parole, poter decidere chi può accedere ai dati e in che modo possono essere trattati.

La sovranità digitale è un rischio oppure un'opportunità per il buon funzionamento di Internet?

In pochi oggi si preoccupano del buon funzionamento di Internet. Oppure, nella migliore delle ipotesi, danno a esso una definizione allineata ai propri interessi. A voler essere puristi e richiamare i princìpi degli anni Novanta, dove Internet era immaginato come uno spazio aperto, distribuito, neutrale e interoperabile, l'applicazione di una sovranità digitale piegata agli interessi di pochi attori rappresenterebbe un rischio. La frammentazione introdotta da politiche nazionali troppo rigide, ma anche da standard tecnici assoggettati alle grandi aziende tecnologiche, può trasformare la rete globale in una somma di intranet regionali, ostacolando la libera circolazione delle informazioni e la cooperazione internazionale.

Eppure, la sovranità digitale potrebbe anche rappresentare un'opportunità inattesa. La realtà è che, al di là delle nobili intenzioni originarie, Internet si è di fatto strutturata attorno a concentrazioni che ne contraddicono profondamente la natura distribuita e neutrale. Un numero ristretto di attori, prevalentemente privati e geograficamente localizzati, controlla porzioni decisive dell'infrastruttura, dei servizi, dei dati e perfino degli standard. A questo punto le spinte verso una sovranità digitale nazionale, se ben orientate e coordinate fra Paesi, possono fungere da contrappeso a questi "cluster di potere", promuovendo una distribuzione più democratica delle risorse digitali. Significherebbe impostare una sovranità digitale che incentivi lo sviluppo di infrastrutture aperte e trasparenti, sostenga l'interoperabilità dei sistemi e rilanci la decentralizzazione come principio architetturale e politico.

Quali sono secondo lei i rischi per i singoli Stati determinati dall'assenza di sovranità digitale?

Per comprendere davvero i rischi legati all'assenza di sovranità digitale bisogna partire dal percorso che va dal dato alla decisione. Il dato, da solo, è inerte. Diventa utile solo quando viene raccolto, trattato, interpretato e trasformato in informazione utile per prendere decisioni. È in questo processo che si gioca la capacità di uno Stato di agire in modo autonomo ed efficace. Ogni fase, dalla raccolta alla registrazione, dall'elaborazione all'interpretazione, coinvolge tecnologie, infrastrutture e competenze che possono trovarsi dentro o fuori dal perimetro nazionale. Quando sono fuori, perdiamo pezzi di sovranità.

Il rischio non è solo che i dati vengano trafugati (cosa più che possibile grazie a leggi come il Cloud Act statunitense) o che qualcuno ne faccia un uso improprio. È che il valore delle decisioni che prendiamo sia condizionato da strumenti che non controlliamo, e che potrebbero restituirci un quadro distorto della realtà. Modelli di intelligenza artificiale opachi, infrastrutture cloud soggette a legislazioni straniere, software critici sviluppati senza

trasparenza sono solo alcuni esempi di come la mancanza di sovranità digitale si traduce in una perdita di capacità decisionale. E questa perdita, quando riguarda ambiti sensibili come la sicurezza o i diritti fondamentali, diventa un problema politico oltre che tecnico.

Per non parlare dei rischi derivanti dal "lock-in" tecnologico. Sistemi perfettamente integrati fra loro, ma realizzati da altri, diventano talmente comodi e diffusi che smettere di usarli comporta costi enormi, spesso insostenibili. Più si raffina l'interdipendenza tra questi strumenti, più diventa difficile sostituirne anche solo uno senza compromettere l'intero flusso operativo. Un vincolo silenzioso, che si rafforza a ogni iterazione, a ogni aggiornamento, a ogni aumento di complessità del workflow. Nel frattempo, si perde la capacità di scegliere, di adattare, di governare l'evoluzione tecnologica in funzione dei propri interessi, una vulnus che porta a una vera e propria perdita di sovranità strategica.

La guerra algoritmica portata da alcuni attori nel cyberspace minaccia le funzioni basilari degli stati e come?

Per capire cosa rappresenta oggi una guerra algoritmica, occorre fare uno sforzo di astrazione. Non stiamo parlando solo di un attacco diretto a un'infrastruttura, tipo un malware che paralizza un ospedale. Parliamo di qualcosa di più sottile, che concerne la progressiva erosione della capacità di uno Stato di operare in modo affidabile, continuo e autonomo nei suoi ambiti essenziali, come difesa, economia, energia, informazione. Nel momento in cui queste funzioni passano sempre più attraverso strati digitali, diventano inevitabilmente vulnerabili a un diverso tipo di attacco, quello che distorce, degrada, interrompe. Non servono le bombe, bastano gli algoritmi. E a differenza degli attacchi tradizionali, qui l'identificazione del colpevole è incerta, la risposta è complicata, il tempo di reazione si misura in ore anziché in settimane.

Gli esempi servono solo a rendere visibile questa logica. Quando un ransomware costringe il governo americano a dichiarare lo stato d'emergenza per garantire il rifornimento di carburante (pensiamo a Colonial Pipeline nel 2021) non è solo un problema di sicurezza informatica, ma una dimostrazione di vulnerabilità sistemica. Quando un'operazione come NotPetya paralizza logistica, trasporti e finanza a livello globale, la lezione oltre che tecnica è geopolitica. Quando le piattaforme digitali vengono usate per interferire nei meccanismi di selezione del potere come le elezioni, non ci troviamo più nel campo della propaganda, siamo in quello della stabilità democratica.

Il punto è che questa forma di conflitto complicata da capire, da gestire e da comunicare ai cittadini, per non parlare dell'ibridazione che la distribuisce su molteplici piattaforme, colpisce ciò che rende uno Stato tale. La sua capacità di decidere, di garantire continuità, di ispirare fiducia. E lo fa sfruttando le interdipendenze che lo Stato stesso ha costruito, in nome dell'efficienza e dell'innovazione. Ma se tutto questo si fonda su tecnologie che non controlliamo, e se le decisioni dipendono da informazioni che possono essere manipolate, allora il problema è capire dove finisce il potere pubblico e dove inizia la vulnerabilità digitale sistemica.

É possibile garantire libertà e apertura della Rete senza rinunciare alla sovranità digitale?

La tensione tra libertà della Rete e sovranità digitale è reale, ma non è insormontabile. Il problema nasce nel momento in cui si pensa ai due concetti come opposti. Da un lato l'ideale di Internet come spazio aperto, fluido, senza confini. Dall'altro la necessità per gli Stati di esercitare un controllo su ciò che accade nei propri spazi digitali. Ma si tratta, a mio modesto avviso, di una dicotomia fuorviante. Perché libertà e sovranità non sono necessariamente in contraddizione. Dipende da che tipo di sovranità intendiamo e da come la esercitiamo.

Se per sovranità digitale intendiamo un sistema chiuso, isolato, con barriere arbitrarie all'accesso e alla circolazione dei contenuti, allora sì, l'apertura della Rete ne risulterebbe gravemente compromessa. Ma se la intendiamo come capacità di uno Stato (o meglio, di una comunità politica) di definire regole chiare, trasparenti e democraticamente negoziate per proteggere i propri interessi strategici ed evitare "cluster di potere" non necessariamente inclini alle regole democratiche, allora non solo la libertà non viene sacrificata, ma può essere rafforzata.

La libertà della Rete non è assenza di regole. È assenza di arbitrio. Una anarchia totale non garantirebbe affatto la libertà, poiché favorirebbe chi dispone di maggiori risorse, strumenti o visibilità per imporre il proprio volere. Infatti, oggi, troppo spesso, l'arbitrio non viene dagli Stati, ma da piattaforme private che stabiliscono unilateralmente cosa si può dire, vedere, monetizzare o rimuovere. La vera libertà online nasce invece da un insieme di norme chiare, trasparenti e stabili, definite in modo democratico, che valgano per tutti e siano applicate con procedure verificabili. In questo senso, una sovranità digitale ben costruita, che promuove trasparenza, tutela i diritti, impone standard di interoperabilità e protegge lo spazio informativo da manipolazioni esterne, non limita necessariamente l'apertura della Rete, ma può renderla più solida, più sostenibile, più democratica.

Dovremmo dunque progettare una sovranità che non sia difensiva, ma generativa. Che non costruisca muri, ma architetture pubbliche. Che non sostituisca l'apertura con il controllo, ma il controllo opaco con la responsabilità condivisa. Non è la sovranità il problema. È come decidiamo di esercitarla.

Quale può essere la risposta al rischio di attacchi cibernetici, alla disinformazione, all'abuso dell'intelligenza artificiale?

Uno degli effetti più profondi, e meno discussi, dell'intelligenza artificiale è la sua capacità di democratizzare la pericolosità. In passato, per condurre un attacco informatico sofisticato, coordinare una campagna di disinformazione o costruire un ordigno esplosivo servivano competenze elevate, accesso a risorse specifiche, conoscenze particolari, spesso una rete strutturata di supporto. Oggi, grazie ai modelli di AI generativa opportunamente modificati, questo scenario sta cambiando radicalmente. E lo sta facendo in fretta.

Non serve molta immaginazione per proiettare nel futuro prossimo un quindicenne, o chiunque disponga di una minima motivazione, con accesso a LLM open-source, facilmente modificabili, installabili in locale, non soggetti ad alcun tipo di filtro o tracciamento. A quel punto potrà ottenere, in pochi secondi, istruzioni dettagliate su come avvelenare una falda acquifera, costruire un malware sofisticato, manipolare le opinioni pubbliche con campagne automatizzate, creare deepfake per screditare chiunque, o analizzare automaticamente le difese di un'azienda per poi penetrare - sempre automaticamente - il suo sistema informatico. Si tratta di un cambiamento strutturale nel rapporto tra potenza tecnologica e limitazione pratica del danno individuale. L'intelligenza artificiale abbatterà completamente la soglia tecnica di ingresso nell'ecosistema delle minacce, rendendo chiunque, anche chi finora non veniva preso in considerazione per mancanza di competenze o di accesso alle informazioni, un potenziale pericolo pubblico.

L'asimmetria storica tra attacco e difesa si sta spostando in modo netto a favore degli attaccanti. E questa è una condizione che nessun sistema sociale, politico o tecnologico può reggere a lungo. Non sarà sufficiente rafforzare le difese tecniche, servirà un nuovo approccio culturale e strategico, che riconosca la natura accessibile, distribuita e adattiva delle minacce future.

La risposta non potrà più essere cercata nell'idea di una difesa impenetrabile. Quel paradigma ormai è superato. Quando la superficie d'attacco è ovunque e chiunque può

potenzialmente colpire con azioni sofisticate, l'obiettivo non è più evitare la caduta, ma saper cadere meglio degli altri. La parola chiave è "resilienza", un mantra in questi anni spesso recitato a sproposito. Resilienza intesa non come semplice resistenza passiva, ma come capacità strutturale di assorbire un impatto, contenerlo, comprenderlo rapidamente e rialzarsi. Con la minima perdita possibile di funzione. Resilienza significa accettare che l'incidente non sia un'eccezione ma una condizione di sistema, e organizzare la risposta di conseguenza.

Resilienza, in questo contesto, vuol dire anche preparazione culturale. Avere cittadini, amministratori, imprese che sanno riconoscere un attacco, distinguere un'informazione manipolata, reagire senza panico, comunicare con trasparenza. La vulnerabilità oggi non sta solo nei sistemi, ma nella nostra capacità collettiva di interpretare e governare ciò che ci colpisce. Prepararsi al peggio non è segno di sfiducia nel progresso tecnologico o nelle capacità delle Istituzioni, ma condizione per poter continuare a beneficiare della tecnologia.

13. Sitografia

- https://t.me/itarmyofukraine2022
- https://t.me/itarmyofukraine2022/1637
- https://t.me/c/1228309110/34219
- https://www.ccc.de/en/
- https://web.archive.org/web/20010201081900/http://www.netstrike.it/
- http://www.billboardliberation.com/
- https://it.wikipedia.org/wiki/Partito_Pirata
- https://web.archive.org/web/20120208220331/http://www.rtmark.com/gatt.html
- https://web.archive.org/web/*/www.thehacktivist.com

14. Bibliografia

1. Abraham, Y. (2025). ‘Legitimization Cell’: Israeli unit tasked with linking Gaza journalists to Hamas. +972. Disponibile in: https://www.972mag.com/israel-gaza-journalists-hamas-hasbara/ [14 agosto 2025]

2. Abraham, Y., (2024). ‘Order from Amazon’: How tech giants are storing mass data for Israel’s war. +972. Disponibile in: https://www.972mag.com/cloud-israeli-army-gaza-amazon-google-microsoft/ [4 agosto 2024]

3. Abraham, Y., (2024). ‘Lavender’: The AI machine directing Israel’s bombing spree in Gaza, +972. Disponibile in: https://www.972mag.com/lavender-ai-israeli-army-gaza/ [3 aprile 2024].

4. Access Now. (2023). Palestine unplugged: how Israel disrupts Gaza's internet, Access now. Disponibile online in: https://www.accessnow.org/publication/palestine-unplugged/ [10 novembre 2023].

5. Agenzia per la cybersicurezza Nazionale (2024). Problematiche relative ad aggiornamento CrowdStrike (BL01/240719/CSIRT-ITA) -Aggiornamento. ACN. Disponibile in: https://www.acn.gov.it/portale/w/problematiche-relative-ad-aggiornamento-crowdstrike-bl01/240719/csirt-ita-aggiornamento [19 luglio 2024].

6. AI Forensics (2025). AI Generated Algorithmic Virality. Ai Forensics. Disponibile in: https://aiforensics.org/work/gen-ai-slop [31 luglio 2025]

7. American Sunlight Project (2024). NEW REPORT: Russian propaganda may be flooding AI models. ASP. Disponibile in: https://www.americansunlight.org/updates/new-report-russian-propaganda-may-be-flooding-ai-models.

8. Amnesty International. (2025). Gaza, fame come arma di genocidio. Amnesty International. Disponibile in: https://www.amnesty.it/gaza-fame-come-arma-di-genocidio/ [3 luglio 2025]

9. ANSA. (2022). Musk compra Twitter e licenzia subito quattro top manager. ANSA. Disponibile in: https://www.ansa.it/sito/notizie/tecnologia/internet_social/2022/10/28/musk-compra-twitter-e-licenzia-subito-quattro-top-manager_4974ef3e-a6d1-4fb2-a103-1ea6a29cad8b.html [28 ottobre 2022].

10. Anthropic (2024). Clio: A system for privacy-preserving insights into real-world AI use. Anthropic. Disponibile in: https://www.anthropic.com/research/clio.

11. Anthropic (2025). Detecting and countering misuse of AI: August 2025. Anthropic. Disponibile in: https://www.anthropic.com/news/detecting-countering-misuse-aug-2025 [27 agosto 2025]

12. Anthropic (2025). Disrupting the first reported AI-orchestrated cyber espionage campaign. Disponibile in: https://www.anthropic.com/news/disrupting-AI-espionage?_bhlid=5ccdd567e96de7457f6c4c8c18651b8d1e6afe22 [13 novembre 2025]

13. Agrafiotis. I., Nurse, J. R. C.. Goldsmith, M., Creese, S. & Upton, D., (2018). A taxonomy of cyber-harms: Defining the impacts of cyber-attacks and understanding, how they propagate, *Journal of Cybersecurity*, Volume 4, Issue 1, 2018, tyy006, https://doi.org/10.1093/cybsec/tyy006.

14. Albanese, A., Giangiulio, G. (2025). #ISRAELHAMASWAR. Israel, Palantir, and the 2024 pager explosions in Lebanon. Disponibile in: https://www.agcnews.eu/israelhamaswar-israel-palantir-and-the-2024-pager-explosions-in-lebanon/ [21 dicembre 2025]

15. Arditti, R. (2025). Hard Power. Perchè la Guerra cambia la storia. Roma/Cesena: Giubilei Regnani.

16. Asaf, D. (2025). Microsoft cloud services disrupted by Red Sea cable cuts. Disponibile in: https://www.bbc.com/news/articles/c3rvx470yg8o [7 settembre 2025].

17. Ashraf, C., (2021). Defining cyberwar: towards a definitional framework, Defense & Security Analysis, 37:3, 274-294, DOI: 10.1080/14751798.2021.1959141.

18. Assange, J. (2012). Internet è il nemico. Conversazione con Jacob Appelbaum, Andy Muller-Mauguhn e Jeremie Zimmermann. Milano: Giangiacomo Feltrinelli Editore.

19. Atlantic Council. (2024). Romania annulled its presidential election results amid alleged Russian interference. What happens next? Atlantic Council. Disponibile in: https://www.atlanticcouncil.org/blogs/new-atlanticist/romania-annulled-its-presidential-election-results-amid-alleged-russian-interference-what-happens-next/ [6 dicembre 2024].

20. Autonome A.f.r.i.k.a. gruppe, Luther Blisset, Sonja Brunzels (2001), Comunicazione-guerriglia. Tattiche di agitazione gioiosa e resistenza ludica all'oppressione, Roma, DeriveApprodi.

21. Bachini, V., Tesconi, M. (2020). Fake People, storie di social bot e bugiardi digitali. Torino, Codice Edizioni.

22. BadSha, N., Reuters. (2024). Telegram app founder Pavel Durov reportedly arrested at French airport, The Guardian. Disponibile in https://www.theguardian.com/media/article/2024/aug/24/telegram-app-founder-pavel-durov-arrested-at-french-airport [25 agosto 2024].

23. Roberto Baldoni, Managing the cyber risk in a multipolar world, International Journal of Critical Infrastructure Protection, Volume 39, 2022, 100578, ISSN 1874-5482, https://doi.org/10.1016/S1874-5482(22)00062-2.

24. Baldoni, R. (2024). Charting digital sovereignity. A survival playbook. Amazon.

25. Baldoni, R. (2025). Sovranità digitale. Cos'è e quali sono le principali minacce al cyberspazio nazionale. Bologna: Il Mulino.

26. BBC News. (2022). Bucha killings: Satellite image of bodies site contradicts Russian claims. BBC. Disponibile in: https://www.bbc.com/news/60981238 [11 aprile 2022].

27. Barbano, A. (2012). Manuale di giornalismo. In collaborazione con Vincenzo Sassu. Pp 262-279. Roma-Bari: Gius. Laterza & Figli S.p.A.

28. Barlow, J.P. (1996). Dichiarazione d'indipendenza del cyberspazio. Davos, Svizzera, 8 febbraio 1996. Disponibile in: https://www.eff.org/cyberspace-independence

29. Bartolomei, R. (2025). Cavi sottomarini bersaglio. Il Giorno, [28 settembre 2025].

30. Rocco Bellanova, Helena Carrapico & Denis Duez (2022). Digital/sovereignity and European security integration: an introduction, European Security, 31:3, 337-355, DOI: 10.1080/09662839.2022.2101887

31. Benanti, P. (2025). Il crollo di Babele. In Mario Caligiuri, (a cura di) Il fuoco di Prometeo. Intelligence e intelligenza artificiale. Soveria Mannelli: Rubbettino Editore.

32. E. Bender, T. Gebru, A. McMillan-Major, & S. Shmitchell. (2021). *On the Dangers of Stochastic Parrots: Can Language Models Be Too Big? Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency, New York, Association for Computer Machinery – ACM,* (*March 2021*) DOI: 10.1145/3442188.3445922

33. Bentivegna, S., Boccia Artieri, G. (2021). Voci della democrazia. Il futuro del dibattito pubblico. Pp 106-109. Bologna: Il Mulino.

34. Sergio López Bernal, Alberto Huertas Celdrán, Gregorio Martínez Pérez, Neuronal Jamming cyberattack over invasive BCIs affecting the resolution of tasks requiring visual capabilities, Computers & Security, Volume 112, 2022, 102534, ISSN 0167-4048, https://doi.org/10.1016/j.cose.2021.102534.

35. Bernays, E. L. (2020). Propaganda. L'arte di manipolare l'opinione pubblica, a cura di Raffaele Scelsi, Milano, Shake Edizioni 2020; ed. or. *Propaganda, 1928,* Horace Liveright, New York.

36. Bey, H. (2007). T.A.Z. Zone temporaneamente autonome, ShaKe; tit. or. T.A.Z.: The Temporary Autonomous Zone, Ontological Anarchy, Poetic Terrorism, Autonomedia, New York, 1991.

37. Bjola, C. (2024). Algorithmic invasions: How information warfare threatens NATO's eastern flank. Nato Review. Disponibile in: https://www.nato.int/docu/review/articles/2025/02/07/algorithmic-invasions-how-information-warfare-threatens-nato-s-eastern-flank/index.html [07 febbraio 2025].

38. Berardi, F. Bifo (2011), La sollevazione. Collasso europeo e prospettive del movimento. Lecce: Manni Editori.

39. Bernal, A. Carter, C., Singh, I., Cao, K., & Madreperla, O. (2020). Cognitive warfare. An attack on truth and thought. Nato, Johns Hopkins University. Disponibile online: https://innovationhub-act.org/wp-content/uploads/2023/12/Cognitive-Warfare.pdf

40. Bianchi, L. (2021). Complotti. Da QAnon alla pandemia, cronache dal mondo capovolto. Roma: Minimum Fax.

41. Bigazzi, F., Fertilio, D., Germani, S. (2022). Bugie di guerra. La disinformazione russa dall'Unione sovietica all'Ucraina. Roma: Paesi Edizioni

42. Blackberry. (2023). BiBi Wiper Used in the Israel-Hamas War Now Runs on Windows. Blackberry. Disponibile in: https://blogs.blackberry.com/en/2023/11/bibi-wiper-used-in-the-israel-hamas-war-now-runs-on-windows [11 ottobre 2023].

43. Bolter, J. D., Grusin, R. (2000), Remediation: Understanding New Media. Mit Press.

44. Borghi, E. (2025). Sotto Attacco. Cosa sta accadendo, cosa potrebbe accadere, come ne usciremo, pp 221-237. Soveria Mannelli, Rubbettino.

45. Borgia, F. (2022). Diritto e conflitti: il ruolo dei nuovi attori e della tecnologia. In M. Bressan, G. Cuzzelli, (a cura di), Da Clausewitz a Putin: la guerra nel XXI secolo. Riflessioni sui conflitti nel mondo contemporaneo (pp. 87-102), Milano: Ledizioni.

46. Brady, William J., Ana P. Gantman, and Jay J. Van Bavel. "Attentional capture helps explain why moral and emotional content go viral." *Journal of Experimental Psychology: General* 149.4 (2020): 746.

47. Brera, P. (2022). Vivere a Kiev sotto attacco hacker. "É tutto bloccato, ci tengono in pugno". La Repubblica, 17 febbraio 2022.

48. Brooks, R. R., Oxcelik, I., Oakley, J. & Tusing, N. (2021). Distributed Denial of Service (DDoS): A History, IEEE.

49. Calise, M., Musella, F. (2019). Il Principe digitale. Roma: Laterza.

50. Canducci, M. (2025). Empatia Artificiale. Come ci innamoreremo della macchine e perché non saremo ricambiati. Milano: Egea.

51. Cappella, N., Hall Jamieson, K., (2008). Echo Chamber: Rush Limbaugh and the Conservative Media Establishment. Oxford: Oxford University Press.

52. Carboni, K. (2024). Un deepfake del presidente Biden ha detto agli elettori di non votare. Wired Italia. Disponibile in: https://www.wired.it/article/deepfake-joe-biden-elezioni-stati-uniti/ [23 gennaio 2024]

53. Carfagna, B. (2024). La Guerra è diversa, rivoluzionata dalle nuove frontiere della tecnologia. Il Sole 24 Ore, 11 dicembre 2024

54. Carlin, J., P. (2018) Dawn of the Code War. America's battle Against Russia, China, and the Rising Global Cyber Threat. With Garret >M. Graff. New York, Public Affairs

55. Cerf, V. (2022). Sulla governance di Internet. In Abba, L., Lazaroni, A., Pietrangelo, M. La Internet governance e le sfide della trasformazione digitale, pp 17-22. Napoli, Editoriale Scientifica.

56. Cernicchiaro, G. (2022). La cyber guerra e l'art. 5 del trattato NATO. Altalex. Disponibile in: https://www.altalex.com/documents/news/2022/06/01/cyber-guerra-art-5-trattato-nato [1° giugno 2022].

57. CheckFirst (2025). Content Analysis: a.network.news-pravda.com. Check First. Disponibile in: https://portal-kombat.com/

58. Chomsky, N., Sherman, E.S. (1998). La fabbrica del consenso. Ovvero la politica dei mass media. Milano: Marco Tropea Editore.

59. Cialdini, R. (1993). Influence. Science and practice. New York. HaperCollins College Publishers

60. Cicerone. (2007). L'arte di comunicare. A cura di Paolo Marsich. PP. 21. Milano: Arnoldo Mondadori Editore S.p.A.

61. CNN. (2024). How did pagers explode in Lebanon and why was Hezbollah using them? Here's what we know. Disponibile in: https://edition.cnn.com/2024/09/17/middleeast/lebanon-pager-attack-explosions-hezbollah-explainer-intl-latam [17 settembre 2024].

62. Crescenzi, C. (2023). L'Italia nel mirino degli hacktivisti pro-Palestina. Wired. Disponibile in: https://www.wired.it/article/italia-hacktivisti-pro-palestina-mysterious-team/ [25 ottobre 2023]

63. Cillario. L., Finelli, R. (a cura di). (1998). Capitalismo e conoscenza. L'astrazione del lavoro nell'era telematica. pp. 11-40. Roma: Manifestolibri s.r.l.

64. Cillario, L. (1990). Il capitalismo cognitivo. Sapere, sfruttamento e accumulazione dopo la rivoluzione informatica. In AA. VV. Trasformazione e persistenza. pp. 69-81. Roma: Franco Angeli. Cit. in Alcune tesi su capitalismo, marxismo e postmodernità. In Cillario. L., Finelli, R. (a cura di). (1998). Capitalismo e conoscenza. L'astrazione del lavoro nell'era telematica. Roma: Manifestolibri s.r.l.

65. M. Cinelli, G. De Francisci Morales, A. Galeazzi, W. Quattrociocchi, & M. Starnini, The echo chamber effect on social media, Proc. Natl. Acad. Sci. U.S.A. 118 (9) e2023301118, https://doi.org/10.1073/pnas.2023301118 (2021).

66. Cingari, P. (2025). I magnifici sette superano il Pil dell'Ue: è bolla tecnologica? Euronews. Disponibile in: https://it.euronews.com/business/2025/10/04/i-magnifici-sette-superano-il-pil-dellue-e-bolla-tecnologica [4 ottobre 2025].

67. Cingolani, R. (2019). L'altra specie. Otto domande su noi e loro. Bologna: Il Mulino

68. Cisa, (2025). CISA and Partners Release Joint Advisory on Countering Chinese State-Sponsored Actors Compromise of Networks Worldwide to Feed Global Espionage Systems. Cisa. Disponibile in: https://www.cisa.gov/news-events/news/cisa-and-partners-release-joint-advisory-countering-chinese-state-sponsored-actors-compromise [27 agosto 2025].

69. CyberExpress. (2023). RedAlert Cyberattack by AnonGhost Sends Alarming Nuclear Bomb Fake Alerts, disponibile in: https://thecyberexpress.com/redalert-cyberattack-anonghost-nuclear-alerts/ [29 ottobre 2023].

70. Cyber Infrastructure & Security Agency, (2021). Cyber-Attack Against Ukrainian Critical Infrastructure, Disponibile in https://www.cisa.gov/news-events/ics-alerts/ir-alert-h-16-056-01 [20 luglio 2021], Cyber Infrastructure & Security Agency.

71. Clusit, (2025). Rapporto Clusit sulla cybersecurity in Italia e nel mondo. Clusit. Disponibile in: https://clusit.it/wp-content/uploads/download/Rapporto_Clusit_03-2025_web.pdf [11 marzo 2025]

72. Colajanni, M. (2018). La singolare parabola dell'informatica: dalla tecnoutopia di Internet all'ingovernabilità del Web alla sorveglianza di massa nell'era dei Big Data. In Phronesis. Ventennale di intelligence, AA.VV. Pp 109-120. Roma: Eurilink University Press.

73. (M. Colajanni, comunicazione personale, 10 agosto 2025)

74. Coleman, G. (2016). I Mille volti di Anonymous. La vera storia del gruppo hacker più provocatorio al mondo, Roma, Stampa Alternativa; tit. or. Hacker, Oaxes, Whistleblower, Spy: The Many Faces of Anonymous, Verso, London, 2014.

75. Colon, D. (2024). La guerra dell'informazione. Gli stati alla conquista delle nostre menti. Torino: Einaudi.

76. Cristadoro, N. (2018). La dottrina Gerasimov: la filosofia della guerra non convenzionale nella strategia russa contemporanea. Roma: Libellula.

77. Cristianini, N. (2023). La scorciatoia. Come le macchine sono diventate intelligenti senza pensare in modo umano. Il Mulino: Bologna.

78. Cristianini, N. (2024). Machina sapiens L'algoritmo che ci ha rubato il segreto della conoscenza. Il Mulino: Bologna.

79. Critical Art Ensemble, (1995). Sabotaggio elettronico. Il primo gruppo americano di critica e attacco ai mass media. Roma: Castelvecchi.

80. Critical Art Ensemble, (1998). Disobbedienza Civile Elettronica e altre idee impopolari: come sopravvivere e resistere nella società del controllo, Roma, Castelvecchi; tit or. Critical Arts Ensemble, Civil Disobedience, Autonomedia, New York, 1996.

81. Csirt Italia. (2022). Attacchi DDOS ai danni di soggetti nazionali e internazionali avvenuti a partire dall'11 maggio 2022: Analisi e mitigazione. ACN. Disponibile in https://www.csirt.gov.it/contenuti/attacchi-ddos-ai-danni-di-soggetti-nazionali-ed-internazionali-avvenuti-a-partire-dall11-maggio-2022-analisi-e-mitigazione-bl01-220513-csirt-ita. [10 settembre 2023] CSIRT Italia

82. Curioni, A, Giannuli, A., (2019). Cyberwar. La guerra prossima ventura, Milano – Udine: Mimesis edizioni.

83. Cyberknow, (2025), Iran-Israel War Cyber Tracker, Disponibile online [https://cyberknow.substack.com/p/iran-israel-war-cyber-tracker?img=https%3A%2F%2Fsubstack-post-media.s3.amazonaws.com%2Fpublic%2Fimages%2F85c530a6-1747-4498-a458-09fa98f03956_895x503.png&open=false] Cyberknow newsletter

84. Da Empoli, S. (2023). L'economia di ChatGPT. Tra false paure e veri rischi. Milano: Egea S.p.A.

85. Daily DarkWeb (2025). Israeli Organizations Weizmann Institute, Mor-logistics, and Agura B.C. LTD Allegedly Breached by Handala Hacking Group, DD, disponibile on line in [https://dailydarkweb.net/israeli-organizations-weizmann-institute-mor-logistics-and-agura-b-c-ltd-allegedly-breached-by-handala-hacking-group/], 18 giugno 2025

86. DARPA (2025). AIxCC: AI Cyber Challenge. Darpa. Disponibile in: https://www.darpa.mil/research/programs/ai-cyber

87. Davies, H., Abraham, Y., (2025). 'A million calls an hour': Israel relying on Microsoft cloud for expansive surveillance of Palestinians. The Guardian. Disponibile in: https://www.theguardian.com/world/2025/aug/06/microsoft-israeli-military-palestinian-phone-calls-cloud [6 agosto 2025]

88. Debenedetti, E., Shumailov, I., Fan, T., Hayes, J., Carlini, N., Fabian, D., ... & Tramèr, F. (2025). Defeating prompt injections by design. *arXiv preprint arXiv:2503.18813*.

89. Debord, G. (2001). La società dello spettacolo. Roma: Baldini & Castoldi; Ed. or. La Société du spectacle. Paris: Edition Gallimard. 1992.

90. Delmastro, M., Nicita, M. (2019). Big data. Come stanno cambiando il nostro mondo. Bologna: Il Mulino.

91. Denning, D. E. (1999). "Activism, Hacktivism, and Cyberterrorism: The Internet as a Tool for Influencing Foreign Policy", Global Problem Solving Information Technology and Tools, December 10, 1999, Disponibile in https://nautilus.org/global-problem-solving/activism-hacktivism-and-cyberterrorism-the-internet-as-a-tool-for-influencing-foreign-policy-2/

92. De Seriis, M., Marano, G. (2008). Net.Art. L'arte della connessione. Milano. Shake.

93. De Seriis, M. (2017). Hacktivism, On The use of botnet in Cyberattacks, Theory, Culture and Society, 2017, Vol 34(4) 131-152.

94. Di Corinto, A. (2002). Don't Hate the media, become the media. In La sfida al G8, AA. VV., pag. 157-177. Roma: Manifestolibri.

95. Di Corinto, A., Tozzi, T. (2002). Hacktivism. La libertà nelle maglie della rete. Roma: Manifestolibri.

96. Di Corinto, A. (2006). Revolution OS II. Software libero, proprietà intellettuale, cultura e politica. Milano: Apogeo Editore.

97. Di Corinto, A. (2009). Tecnologie Persuasive. Il Sole 24 Ore. [24 settembre 2009]

98. Di Corinto, A. (2010). Con Wiki, senza amare Julian. Hacker italiani a favore della trasparenza ma non dell'australiano. Il Sole 24 Ore. [14 dicembre 2010]

99. Di Corinto, A. (2014). Un dizionario hacker. Lecce: Manni Editori.

100. Di Corinto, A. (2014). Anonymous ruba gli account del Ku Klux Klan: operazione "Giù il cappuccio", rivelati esponenti, La Repubblica. Disponibile in: https://www.repubblica.it/tecnologia/2014/11/18/news/anonymous_ruba_gli_account_del_ku_klux_klan-100806973/. [17 novembre 2014]

101. Di Corinto, A. (2015). Anonymous minaccia l'Arabia saudita: "Non uccidete Alì". La Repubblica. Disponibile in: https://www.repubblica.it/tecnologia/sicurezza/2015/09/30/news/anonymous_arabia_saudita-124004119/. [30 settembre 2015]

102. Di Corinto, A. (2015). Anonymous: "Abbiamo violato la rete jihadista", La Repubblica. Disponibile in: https://www.repubblica.it/tecnologia/2015/02/08/news/anonymous_abbiamo_violato_la_rete_jihadista-106820821/ [08 febbraio 2015]

103. Di Corinto, A. (2017). Internet non è nata come progetto militare, mettetevelo in testa. In: (a cura di): Abba L, Di Corinto, A., Il futuro trent'anni fa Quando Internet è arrivata in Italia. p. 18-22, San Cesario di Lecce: Manni Editori.

104. Di Corinto, A. (2017). Cybersicurezza, l'allarme degli esperti: ''Borse mondiali nel mirino degli hacker''. La Repubblica. Disponibile in: https://www.repubblica.it/tecnologia/sicurezza/2017/04/06/news/security_analist_summit_2017-162331025/ [6 aprile 2017]

105. Di Corinto, A. (2017). Peggio del Datagate: i segreti della cripta (Vault 7), svelati da Wikileaks. Startupitalia. Disponibile in: https://startupitalia.eu/tech/cybersecurity/peggio-del-datagate-i-segreti-della-cripta-vault-7-svelati-da-wikileaks/ [8 marzo 2017]

106. Di Corinto, A. (2018). Le guerre del futuro si combatteranno nei nostri cuori. Il Manifesto. Disponibile in: https://ilmanifesto.it/le-guerre-del-futuro-si-combatteranno-nei-nostri-cuori [21 gennaio 2018].

107. Di Corinto, A. (2018). Wannacry e non solo, le cyber-armi della Nsa sono ancora in giro. Il Manifesto. Disponibile in: https://ilmanifesto.it/wannacry-e-non-solo-le-cyber-armi-della-nsa-sono-ancora-in-giro [18 maggio 2018].

108. Di Corinto, A. (2018). La guerra degli hashtag e il mostro del web. Il Manifesto. Disponibile in https://ilmanifesto.it/la-guerra-degli-hashtag-e-il-mostro-del-web-2/ [31 maggio 2018].

109. Di Corinto, A. (2019). Perché gli umani attaccano i sistemi basati sull'Intelligenza Artificiale. La Repubblica. Disponibile in: https://www.repubblica.it/tecnologia/sicurezza/2019/07/18/news/perche_gli_umani_attaccano_i_sistemi_basati_sull_intelligenza_artificiale-231461578/ [18 luglio 2019].

110. Di Corinto, A. (2020). Dark Basin e i suoi fratelli, i segreti degli hacker mercenari. La Repubblica. Disponibile in: https://www.repubblica.it/tecnologia/sicurezza/2020/06/12/news/dark_basin_e_i_suoi_fratelli_i_segreti_degli_hacker_mercenari-259011620/ [12 giugno 2020].

111. Di Corinto, A. (2020), Il Dragone attacca con le fake news, sicuri di riconoscerle? Il Manifesto. Disponibile in: https://ilmanifesto.it/il-dragone-attacca-con-le-fake-news-sicuri-di-riconoscerle [20 luglio 2020].

112. Di Corinto, A. (2020). Iran vs Usa, la cyberguerra è solo agli inizi. La Repubblica. Disponibile in: https://www.repubblica.it/tecnologia/sicurezza/2020/01/09/news/iran_vs_usa_la_cyberguerra_e_solo_agli_inizi-245335228/ [9 gennaio 2020].

113. Di Corinto, A. (2021), Corea del Nord: cybercrime di Stato per finanziare il programma nucleare, La Repubblica. Disponibile in: https://www.repubblica.it/esteri/2021/02/11/news/nord_corea_cybercrime_di_stato_per_finanziare_il_programma_nucleare-287136346/ [11 febbraio 2021].

114. Di Corinto, A. (2021). Perché è il momento di parlare di neuro-privacy. Wired. Disponibile in: https://www.wired.it/attualita/tech/2021/01/29/privacy-neuro-privacy-garante/?refresh_ce=[29 gennaio 2021]

115. Di Corinto, A. (2021). Splinternet, la frammentazione della Rete è servita. E dobbiamo preoccuparci. La Repubblica. Disponibile in: https://www.repubblica.it/tecnologia/2021/06/11/news/il_rischio_splinternet_ovvero_la_balcanizzazione_della_rete-304644266/ [11 giugno 2021].

116. Di Corinto, A. (2021). "Il maxi-attacco ransomware a Kaseya riguarda anche l'Italia. E ci fa tremare", La Repubblica, Disponibile in: https://www.repubblica.it/tecnologia/2021/07/05/news/ransomware_attacco_kaseya_italia_danni-308945610/ [5 luglio 2021].

117. Di Corinto, A. (2021). "Perché l'attacco a SolarWinds è stato così devastante? La Repubblica, Disponibile in: https://www.italian.tech/2021/09/18/news/perche_l_attacco_a_solarwinds_e_stato_cosi_devastante_-318027621/ [18 settembre 2021].

118. Di Corinto, A. (2022). Ricatti informatici, il 2021 è un anno da dimenticare. La Repubblica. Disponibile in: https://www.repubblica.it/tecnologia/2022/01/04/news/ricatti_informatici-332611336/ [4 gennaio 2022].

119. Di Corinto, A. (2022). Attacco informatico a Thales Group: LockBit 2.0 ha pubblicato i dati rubati. La Repubblica. Disponibile in: https://www.repubblica.it/tecnologia/2022/01/18/news/attacco_informatico_a_thalesgroup_lockbit_2_0_avrebbe_pubblicato_i_dati_rubati-334299440/ [18 gennaio 2022]

120. Di Corinto, A. (2022). #OpRussia: Anonymous dichiara guerra a Putin nel cyberspazio, La Repubblica. Disponibile in: https://www.repubblica.it/tecnologia/2022/02/25/news/oprussia_anonymous_dich

iara_guerra_a_putin_nel_cyberspace_insieme_a_ghostsec-339254926/ [4 marzo 2022]

121. Di Corinto, A. (2022). La guerra in Ucraina è anche sul web: allarme per un virus che cancella la memoria dei computer. La Repubblica. Disponibile in: https://www.repubblica.it/tecnologia/2022/02/24/news/ucraina_sotto_attacco_cibernetico_allarme_per_un_virus_che_cancella_la_memoria_dei_computer_killdisk-339103442/ [24 febbraio 2022].

122. Di Corinto, A. (2022). La guerra in Ucraina è anche sul web. In: DEMOCRAZIA FUTURA. ISSN 2785-0811. VI-VII:(2022), pp. 505-515.

123. Di Corinto, A., (2022). Perché la gang ransomware Conti si è schierata con la Russia. La Repubblica. Disponibile in:https://www.repubblica.it/tecnologia/2022/03/01/news/perche_la_gang_ransomware_conti_si_e_schierata_con_la_russia-339788516/ [1° marzo 2022]

124. Di Corinto, A., Rociola, A., (2022). Attacco hacker all'Italia. Cos'è Killnet, il gruppo russo che lo ha rivendicato, La Repubblica. Disponibile in: https://www.repubblica.it/tecnologia/2022/05/11/news/attacco_hacker_italia_russia_killnet-349111881/ [11 maggio 2022].

125. Di Corinto, A. (2022). Data commons: privacy e cybersecurity sono diritti umani fondamentali. In Abba, L., Lazaroni, A., Pietrangelo, M., (2022). La Internet governance e le sfide della trasformazione digitale, pp 45-51. Napoli, Editoriale Scientifica.

126. Di Corinto, A. (2022). Hacking e disinformazione, la scuola russa, Il Manifesto. Disponibile in: https://ilmanifesto.it/hacking-e-disinformazione-la-scuola-russa [29 settembre 2022].

127. Di Corinto, A. (2023). La propaganda computazionale e le interferenze hacker. Mondo Digitale, 103 (2023), ISSN: 1720-898X. Disponibile in: https://mondodigitale.aicanet.it/la-propaganda-computazionale-e-le-interferenze-hacker/

128. Di Corinto, A (2024). The Role Of Disinformation, Propaganda And Active Measures In Cyber Warfare. Noname(057)16 Travels To Italy. In G. D'Angelo, F. Luccio, F. Palmeri, Proceedings of the 8th Italian Conference on Cyber Security, ITASEC 2024, CEUR-WS.org, ISSN 1613-0073.

129. Digital Forensic Research Lab, (2021). Weaponized: How rumors about COVID-19's origins led to a narrative arms race. DFRLab. Disponibile in:https://www.atlanticcouncil.org/in-depth-research-reports/report/weaponized-covid-19 [14 febbraio 2021]

130. Digital Forensic Research Lab, (2022). How ten false flag narratives were promoted by pro-Kremlin media. Medium. Disponibile in: https://medium.com/dfrlab/how-ten-false-flag-narratives-were-promoted-by-pro-kremlin-media-c67e786c6085 [18 febbraio 2022].

131. Dominguez, R. (2003). Illegal Knowledge? Strategies for new media activism. Electronic book review, Disponibile in https://electronicbookreview.com/essay/illegal-knowledge-strategies-for-new-media-activism/ [10 agosto 2023].

132. Donini, A. (1991). Breve storia delle religioni. Roma, Newton Compton Editori.

133. Doroshenko, L., Lukito J. (2021). „Trollfare: Russia's Disinformation Campaign During Military Conflict in Ukraine", International Journal of Communication 15(2021), 4662–4689, http://ijoc.org. USA, Austin, p. 464.

134. Downing, J. D.H. et T. Villareal Ford, G. Gil, L. Stein, (2001). Radical Media. Rebellious communication and Social Movements, Sage Publications Inc.

135. ENISA. (2024). ENISA Threat Landscape 2024. Enisa. Disponibile in: https://www.enisa.europa.eu/publications/enisa-threat-landscape-2024 [19 settembre 2024].

136. European Council (2025). Declaration for European Digital Sovereignty. European Council. Disponibile in: https://www.consilium.europa.eu/en/policies/european-declaration-on-digital-rights/ [19 novembre 2025].

137. Europarlamento, (2021), The impact of disinformation on democratic processes and human rights in the world, https://www.europarl.europa.eu/RegData/etudes/STUD/2021/653635/EXPO_STU(2021)653635_EN.pdf [aprile 2021]

138. European Union, (2022). DIRECTIVE (EU) 2022/2557 OF THE EUROPEAN PARLIAMENT AND OF THE COUNCIL. on the resilience of critical entities and repealing Council Directive 2008/114/EC. Eur-lex. Disponibile in: https://eur-lex.europa.eu/eli/dir/2022/2557/oj.

139. European Union (2022). REGULATION (EU) 2022/2554 OF THE EUROPEAN PARLIAMENT AND OF THE COUNCIL. on digital operational resilience for the financial sector and amending Regulations (EC) No 1060/2009, (EU) No 648/2012, (EU) No 600/2014, (EU) No 909/2014 and (EU) 2016/1011 Eur-lex. Disponibile in: https://eur-lex.europa.eu/eli/reg/2022/2554/oj/eng [27 dicembre 2022].

140. European Union, (2025). EU adopts blueprint to better manage European cyber crises and incidents. European Union. Disponibile in: https://www.consilium.europa.eu/en/press/press-releases/2025/06/06/eu-adopts-blueprint-to-better-manage-european-cyber-crises-and-incidents/ [6 giugno 2025].

141. Europol, (2025). Global operation targets NoName057(16) pro-Russian cybercrime network. Europol. Disponibile in: https://www.europol.europa.eu/media-press/newsroom/news/global-operation-targets-noname05716-pro-russian-cybercrime-network [16 luglio 2025].

142. Facta News, (2025). Il nuovo piano della disinformazione russa per "infettare" i chatbot AI utilizzati in Occidente. Newsguard. Disponibile in: https://www.facta.news/articoli/pravda-disinformazione-russa-chatbot-ia [11 luglio 2025].

143. Fadda, D. (2025). Anonymous rivendica il defacement di 100 siti russi sotto #OpRussia: analisi tecnica e impatti. Insicurezza digitale. Disponibile in: https://insicurezzadigitale.com/anonymous-rivendica-il-defacement-di-100-siti-russi-sotto-oprussia-analisi-tecnica-e-impatti/ [6 agosto 2025]

144. Fang, R., Bindu, R., Gupta, A., Zhan, Q., & Kang, D. (2024). Llm agents can autonomously hack websites. *arXiv preprint arXiv:2402.06664*.

145. Fanpage, (2025). Chi paga per farci vedere i video di Israele che aiuta Gaza su YouTube: la nostra indagine su Google Ads disponibile online: https://www.fanpage.it/https://www.fanpage.it/innovazione/tecnologia/chi-paga-per-farci-vedere-i-video-di-israele-che-aiuta-gaza-su-youtube-la-nostra-indagine-su-google-ads/ [17 giugno 2025]

146. Fathallah, S. (2025). Artificial Intelligence and the Orchestration of Palestinian Life and Death. Tech Policy Press. Disponibile in: https://www.techpolicy.press/artificial-intelligence-and-the-orchestration-of-palestinian-life-and-death/

147. Feltri, S. (2022). Il partito degli influencer. Perché il potere dei social network è una sfida alla democrazia. Torino: Einaudi.

148. Ferraris, M. (2017). Postverità e altri enigmi. Bologna: Il Mulino.

149. Fiammeri, B. (2023). Cybersecurity, cosa succede dopo le dimissioni di Baldoni? Oggi Mantovano al Copasir, 2023. Il Sole 24 Ore. Disponibile in: https://www.ilsole24ore.com/art/cybersecurity-cosasuccede-le-dimissioni-baldoni-pole-prefetto-frattasi-AEFIggzC [8 marzo 2023]

150. Filippi, F. (2022). Guida semiseria per aspiranti storici social. Torino: Bollati Boringhieri Editore.

151. Filippi, F. (2024). Cinquecento anni di rabbia. Rivolte e mezzi di comunicazione da Gutenberg a Capitol Hill. pp- 171-179. Torino: Bollati Boringhieri Editore.

152. Fogg, B.J. (2005). Tecnologia della persuasione. Un'introduzione alla captologia, la disciplina che studia l'uso dei computer per influenzare idee e comportamenti. Milano: Apogeo Editori-Maggioli Editore

153. Floridi. L., (2017). La quarta rivoluzione. Come l'infosfera sta cambiando il mondo. Milano: Raffaello Cortina Editore.

154. Floridi, Luciano, The Fight for Digital Sovereignty: What It Is, and Why It Matters, Especially for the EU (April 15, 2021). Available at SSRN: https://ssrn.com/abstract=3827089 or http://dx.doi.org/10.2139/ssrn.3827089

155. Floridi, L. (2025). La differenza fondamentale. Artificial Agency: una nuova filosofia dell'intelligenza artificiale. Milano: Mondadori.

156. France24, AFP. (2024). S. Korea administrative robot defunct after apparent suicide. France24. Disponibile in https://www.france24.com/en/live-news/20240626-s-korea-administrative-robot-defunct-after-apparent-suicide [26 giugno 2024].

157. Frumento, E. (2025). Modelli di Intelligenza Artificiale "abliterated": la nuova frontiera del cybercrime accessibile a tutti. Cefriel. Disponibile in: https://www.cefriel.com/whitepaper/modelli-di-intelligenza-artificiale-abliterated-la-nuova-frontiera-del-cybercrime-accessibile-a-tutti/#gf_135 [14 ottobre 2025]

158. Gabanelli, M., Santucci, G. (2023). L'Escalation in corso della guerra in rete. in Il Corriere della Sera, 23 ottobre 2023.

159. Gaggi, M. (2022). Big Tech. Con la guerra finisce l'era dell'irresponsabilità. IL Corriere della Sera. Disponibile in: https://www.corriere.it/economia/aziende/22_marzo_09/02-economia-unocorriere-web-sezioni-da149d3c-9ff4-11ec-82d5-6f137f6a69fd.shtml [10 marzo 2022]

160. Galeotti, M. (2017). Controlling Chaos: How Russia manages its political war in Europe, European Council on Foreign Relations, https://ecfr.eu/publication/controlling_chaos_how_russia_manages_its_political_war_in_europe/

161. Gerasimov, V. (2013). Cennost' naukiv predvidenii, Voyenno-Promy-shlennyj. Army Press. Disponibile in: https://www.armyupress.army.mil/Journals/Military-Review/Online-Exclusive/2019-OLE/November/Orenstein-Gerasimov

162. Germani, L.S. (2017). La minaccia della disinformazione. Panoramica Introduttiva. in Germani, L.S. (a cura di): Disinformazione e manipolazione delle percezioni. Una nuova minaccia al sistema-paese. Roma: Eurilink University Press.

163. Giannuli, A. (2012). Come i servizi segreti usano i media. Con quali tecniche l'intelligence influenza e interpreta l'informazione? E come possiamo utilizzarle anche noi per leggere le notizie tra le righe? Milano: Ponte alle Grazie/Adriano Salani S.p.A.

164. Giaquinta, F. (2025). Deepfake elettorali su Facebook: la frode politica è modello di business. Agenda Digitale. Disponibile in: https://www.agendadigitale.eu/cultura-digitale/deepfake-elettorali-su-facebook-la-frode-politica-e-modello-di-business/ [23 ottobre 2025].

165. Goode, L. (2015). Anonymous and the Political Ethos of Hacktivism, Popular Communication, 13:1, 74-86, DOI: 10.1080/15405702.2014.978000.

166. GPDP (2025). Potere e Responsabilità. La cultura della protezione dei dati. Relazione del Presidente Pasquale Stanzione 2024, GPDP. Disponibile in: https://www.garanteprivacy.it/web/guest/home/docweb/-/docweb-display/docweb/10150238 [15 luglio 2025].

167. Graphika, (2020), Spamouflage Goes to America, disponibile online https://graphika.com/reports/spamouflage-dragon-goes-to-america [1 agosto 2020]

168. Graphika, (2025). Keeping Up With The Hacktivists. Examining How International Hacktivist Groups Pursue Attention, Select Targets, and Interact in an Evolving Online Landscape. Graphika. Disponibile in: https://www.graphika.com/reports/keeping-up-with-the-hacktivists

169. Greenberg, A. (2019). Sandworm. A new era of cyberwar and the hunt for Kremlin's most dangerous hackers, New York, DoubleDay.

170. Grynbaum, M. M., Mac, R. (2023). The Times Sues OpenAI and Microsoft Over A.I. Use of Copyrighted Work NYT. Disponibile in: https://www.nytimes.com/2023/12/27/business/media/new-york-times-open-ai-microsoft-lawsuit.html [27 dicembre 2023].

171. Han, Byung-Chul. (2023). Infocrazia. Le nostre vite dominate dalla rete. Torino: Giulio Einaudi editore S.p.A.; ed.or. Infokratie. Berlin: Msb Matthes & Seitz, 2021.

172. Harari, Y. N. (2018). 21 Lezioni per il XXI secolo. Milano: Bompiani.

173. Harari, Y., N. (2024). Nexus. Breve storia delle reti di informazione dall'età della pietra all'IA. Milano: Bompiani.

174. Harding, H., (2017). Media Lies and Brexit: A Double Hammer-Blow to Europe and Ethical Journalism. Ethical journalism network. Disponibile in: https://ethicaljournalismnetwork.org/ethics-in-the-news-media-lies-and-brexit [2017]

175. Hartford, E. (2023). Uncensored Models. Eric Hartford. Disponibile in: https://erichartford.com/uncensored-models

176. Hill, K. (2022). Facial Recognition Goes to War. The New York Times. Disponibile in https://www.nytimes.com/2022/04/07/technology/facial-recognition-ukraine-clearview.html [7 aprile 2022]

177. Hunger, F. (2019). How to hack artificial intelligence. In: Colakides, Y, Garret, M., Gloerich, I. State Machines. (a cura di), Reflections and actions at the edge of digital citinzeship, finance and art (pp 129-144). Amsterdam, Institute of Network Cultures.

178. Iannuzzi, R. (2024). Il 7 ottobre tra verità e propaganda. L'attacco di Hamas e i punti oscuri della narrazione israeliana, pp. 59-67. Fazi Editore, Roma

179. ICT Security Magazine (2025). La guerra algoritmica del 2025: LLM malevoli vs difensivi. ICT Security Magazine. Disponibile in: https://www.ictsecuritymagazine.com/notizie/llm-malevoli/ [9 agosto 2025].

180. Iezzi, P. (2023). Cyber e Potere, L'escalation delle ostilità digitali e i nuovi rischi per le infrastrutture strategiche. Milano: Mondadori Libri S.p.A.

181. Iezzi, P., Razzante, R. (2024). Algoritmo Criminale. Come Mafia, cyber e AI riscrivono le regole del gioco. Milano: Il Sole 24 Ore S.p.A.

182. Industrial Cyber, (2025). Espionage, ransomware, hacktivism unite as nation-states use criminal proxies, cyber tooling to advance geopolitical agendas. Industrial Cyber. Disponibile in: https://industrialcyber.co/medical/espionage-ransomware-hacktivism-unite-as-nation-states-use-criminal-proxies-cyber-tooling-to-advance-geopolitical-agendas/ [8 settembre 2025]

183. Ishiguro, H. (2021). Come costruire un essere umano. Arezzo/Milano: Wudz.

184. ISPI, (2020). On the Geopolitics of Russia's Sovereign Internet Law. ISPI. Disponibile in: https://www.ispionline.it/en/publication/geopolitics-russias-sovereign-internet-law-25428 [18 marzo 2020].

185. Israel Foreign Affairs Ministry, (2025). Livelli record di aiuti umanitari a Gaza. Israel Foreign Affairs Ministry. Disponibile in: https://www.youtube.com/watch?v=Cs7cpkDKTh0 [giugno 2025]

186. Israel Defence Forces, (2025). Home to Hamas' Headquarters, This is an IDF 3D Diagram of the Shifa Hospital. IDF. Disponibile in: https://www.youtube.com/watch?v=6pTYHBZVgVQ

187. Istituto della Enciclopedia Italiana, (1987). Guerra, in Vocabolario della Lingua Italiana, pp. 710-711. Roma: Istituto della Enciclopedia Italiana.

188. ITU, (2024). Global Cybersecurity Index 2024. ITU. Disponibile in: https://www.itu.int/epublications/publication/global-cybersecurity-index-2024.

189. Joyce, S., Huntley, S. (2024), Tool of First Resort: Israel-Hamas War in Cyber, in Google, disponibile online https://blog.google/technology/safety-security/tool-of-first-resort-israel-hamas-war-in-cyber/ [14 febbraio 2024].

190. Kahneman, D. (2012). Thinking, Fast and Slow. London: Penguin Books.

191. Kaye, D. (2021). Libertà Vigilata. La lotta per il controllo di Internet. Roma: Treccani; ed. or. 2019. Speech Police. The global struggle to govern the Internet. Columbia global report.

192. Kapuściński, R. (2006). Il cinico non è adatto a questo mestiere. Conversazioni sul buon giornalismo. A cura di Maria Nadotti. Roma: e/o.

193. Kharpal, A. (2025). Germany tells Apple, Google to block DeepSeek as the Chinese AI app faces rising pressure in Europe, CNBC. Disponibile in:

https://www.cnbc.com/2025/06/27/germany-tells-apple-google-to-block-deepseek-ai-app.html [27 giugno 2025].

194. Klein, N. (2000) No Logo. Taking Aim at the Brand Bullies, Usa, Picador.

195. Klincewicz, M., Alfano, M., Fard, A. E. (2025). Slopaganda: The interaction between propaganda and generative AI. ArXiv preprint https://doi.org/10.48550/arXiv.2503.01560.

196. Körömi, C., Haeck, P., & Cheslow, D. (2025). Zuck goes full Musk, dumps Facebook fact-checking program, Politico.eu. Disponibile in: https://www.politico.eu/article/mark-zuckerberg-full-elon-musk-dump-facebook-fact-checker/ [17 gennaio 2025].

197. Labatut, B. (2023). Maniac. Milano: Adelphi Edizione S.P.A.

198. Lakshmanan, R. (2025). Iranian APT35 Hackers Targeting Israeli Tech Experts with AI-Powered Phishing Attacks, The Hacker News. Disponibile in: https://thehackernews.com/2025/06/iranian-apt35-hackers-targeting-israeli.html [26 giugno 2025]

199. Lamperti, F. (2025). La Cina lancia un'identità unica digitale. E crea un nuovo modello di controllo di Internet, La Repubblica. Disponibile in: https://www.repubblica.it/tecnologia/2025/07/04/news/cina_identntita_digitale_internet_id_come_funziona-424710481 [4 luglio 2025].

200. Lange, L. (2024). Decoding China's AI powered 'Algorithmic Cognitive Warfare". Special Competitive Studies Project, CSP. Disponibile in: https://www.scsp.ai/resource/decoding-chinas-ai-powered-algorithmic-cognitive-warfare/ [21 novembre 2024]

201. Langford, C. (2025). False claims of Chinese airdrops to Gaza. UK Defence Journal. Disponibile in: https://ukdefencejournal.org.uk/false-claims-of-chinese-airdrops-to-gaza/ [18 maggio 2025].

202. La Repubblica, (2012). Anonymous attacca la Costituzione "Il popolo deve difendersi dai tiranni". La Repubblica. Disponibile in: https://www.repubblica.it/tecnologia/2012/03/05/news/anonymous_attacca_la_costituzione_il_popolo_deve_difendersi_dai_tiranni-30978767/ [5 marzo 2012]

203. La Repubblica, (2022). Il video deepfake di Zelensky che ordina agli ucraini di arrendersi. La Repubblica. Disponibile in: https://www.repubblica.it/tecnologia/dossier/tech/2022/03/17/video/il_video_deepfake_di_zelensky_che_ordina_agli_ucraini_di_arrendersi-423327466/ [17 marzo 2022]

204. La Repubblica (2025). "Colpo di stato in Francia": il video fake che ha fatto infuriare Macron. La Repubblica. Disponibile in: https://www.repubblica.it/video/socialnews/2025/12/18/video/colpo_di_stato_in_francia_il_video_fake_che_ha_fatto_infuriare_macron-425047518/ [18 dicembre 2025]

205. Mark A. Lemley, The Splinternet, 70 *Duke Law Journal* 1397-1427 (2021) Disponibile in: https://scholarship.law.duke.edu/dlj/vol70/iss6/3.

206. Levy, S. (1996). Hackers, gli eroi della rivoluzione informatica, Milano, ShaKe Edizioni Underground; tit. originale Hackers, Heroes of the informatic revolution, Ed. or. 1984

207. Lewandowsky, S., Van Der Linden, S. (2021). Countering misinformation and fake news through inoculation and prebunking. European Review of Social Psychology, 32(2), 348-384.

208. Loewenstein, A. (2024), Laboratorio Palestina: Come Israele esporta la tecnologia dell'occupazione in tutto il mondo. Roma: Fazi Editore.

209. Longo, A., Scorza, G. (2021). Intelligenza Artificiale. L'impatto sulle nostre vite, diritti e libertà, pp. 17. Firenze: Mondadori Education S.p.A.

210. López Sans, L. (2025). How hacktivist cyber operations surged amid Israeli-Iranian conflict. Outpost24. Disponibile in: https://outpost24.com/blog/hacktivist-cyber-operations-iran-israel/ [1° luglio 2025].

211. Malik, A. (2025). TikTok will let you choose how much AI-generated content you want to see. Disponibile in: https://techcrunch.com/2025/11/18/tiktok-now-lets-you-choose-how-much-ai-generated-content-you-want-to-see/ [18 novembre 2025].

212. Mandiant Intelligence, (2022). Hacktivists Collaborate with GRU-sponsored APT28. Mandiant Intelligence. Disponibile in: Hacktivists Collaborate with GRU-sponsored APT28, 2022 https://www.mandiant.com/resources/blog/gru-rise-telegram-minions [10 settembre 2023].

213. Mandić, J. & Klarić, D. (2023). Case study of the russian disinformation campaign during the war in Ukraine – propaganda narratives, goals, and impacts. *National Security and the Future, 24* (2), 97-140. https://doi.org/10.37458/nstf.24.2.5

214. Mari, G. (2025). L'orchestra di Goebbels. Ordini e veline alla stampa per manipolare le masse. Torino: Lindau s.r.l.

215. Martin, A. (2022). US military hackers conducting offensive operations in support of Ukraine, says head of Cyber Command. Sky News. Disponibile in: https://news.sky.com/story/us-military-hackers-conducting-offensive-operations-in-support-of-ukraine-says-head-of-cyber-command-12625139 [1 giugno 2022].

216. Mascitti, M. (2025). All'Intelligenza Artificiale si può far dire qualsiasi cosa, con un po' di allenamento. Facta news. Disponibile in: https://www.facta.news/articoli/intelligenza-artificiale-chatbot-disinformazione [23 aprile 2025].

217. Massa, Alessandra & Anzera, Giuseppe. (2023). States vs. Tech Giants. In The Routledge Handbook of Soft Power (pp.264-274) 10.4324/9781003189756-20.

218. Massarelli, F. (2012). La collera della Casbah. Voci di rivoluzione da Tunisi, Milano: Agenzia X.

219. Matania, E., Rapaport A., (2022) Cybermania. How Israel became a global powerhouse in the domain that is revolutionizing the future of humanity, pp. 213-217. Israel, Cybertech-Arrowmedia.

220. Mazzuccato, M. (2014). Lo stato innovatore. Bari: Laterza.

221. McQuail, D. (2025). Propaganda. Treccani; Denis McQuail, Propaganda, in Enciclopedia della Scienze Sociali. Istituto della Enciclopedia Italiana, Roma, 1997.

222. Meikle, G. (2004). Disobbedienza civile elettronica. Mediattivismo, come costruire una nuova sfera pubblica, Milano: Apogeo.

223. Messina, F. (2025). Attacco hacker colpisce il vertice NATO con attacchi DDoS: rivendicato da un gruppo filorusso, HTML. II. Disponibile in: https://www.html.it/magazine/attacco-hacker-colpisce-il-vertice-nato-con-attacchi-ddos-rivendicato-da-un-gruppo-filorusso/ [24 giugno 2025]

224. Mezza, M. (2022). Net-War, Come il giornalismo sta cambiando la guerra. Roma: Donzelli Editore.

225. Mezza, M. (2024), Connessi a morte. Guerra. media e democrazia nella società della cybersecurity. Milano: Donzelli Editore.

226. Mhalla, A. (2025). Tecnopolitica. Come la tecnologia ci rende soldati. Torino: add Editore.

227. Microsoft Digital Security Unit, (2022). An overview of Russia's cyberattack activity in Ukraine, Disponibile in: https://query.prod.cms.rt.microsoft.com/cms/api/am/binary/RE4Vwwd. [10 settembre 2023] Microsoft Digital Security Unit

228. Microsoft Threat Intelligence, (2024). Iran surges cyber-enabled influence operations in support of Hamas, Microsoft, disponibile in: https://www.microsoft.com/en-us/security/security-insider/intelligence-reports/iran-surges-cyber-enabled-influence-operations-in-support-of-hamas [26 febbraio 2024]

229. Miller, C. (2024). Chip war. La sfida tra Cina e USA per il controllo della tecnologia che deciderà il nostro futuro. Garzanti, Milano.

230. Molino, W., Porro, S. (2003). Disinformation Technology. Come si manipola l'informazione per divertirsi, fare soldi e magari provocare una guerra. Milano: Apogeo Editore.

231. Monti, A. (2025). Così gli USA controllano la sovranità digitale delle istituzioni internazionali e della UE. La Repubblica. Disponibile in: https://www.repubblica.it/tecnologia/blog/lettere/2025/07/13/news/cosi_gli_usa_controllano_la_sovranita_digitale_delle_istituzioni_internazionali_e_della_ue-424728328/ [13 luglio 2025].

232. Moulier-Boutang, Y. (2002). (a cura di) L'età del capitalismo cognitivo. Innovazione, proprietà e cooperazione delle moltitudini. Verona: Ombre corte

233. Susan Morgan (2018), Fake news, disinformation, manipulation and online tactics to undermine democracy, Journal of Cyber Policy, 3:1, 39-43, DOI: 10.1080/23738871.2018.1462395

234. Mosca, G. (2025). Che cos'è il vibe-hacking e perché l'intelligenza artificiale può renderlo più diffuso. La Repubblica. Disponibile in: https://www.repubblica.it/tecnologia/2025/08/28/news/che_cose_e_il_vibe-hacking_e_perche_l_intelligen-za_artificiale_puo_renderlo_piu_diffuso-424812797/ [28 agosto 2025]

235. Mott, F. L. (1941). American Journalism: A History of Newspapers in the United States Through 250 Years, 1690-1940, Routledge/Thoemmes Press, 2000 (reprint)

236. Musiani, F. (2022). *Infrastructuring* digital sovereignty: a research agenda for an infrastructure-based sociology of digital self-determination practices, Information, Communication & Society, 25:6, 785-800, DOI: 10.1080/1369118X.2022.2049850

237. Mussi, C. (2024). Innamorato di un chatbot, un adolescente americano si è suicidato a 14 anni. La madre fa causa all'app: «Pericolosa e non testata». Il Corriere della sera. Disponibile in https://www.corriere.it/tecnologia/24_ottobre_23/innamorato-di-un-chatbot-un-adolescente-americano-si-e-suicidato-a-14-anni-la-madre-fa-causa-all-app-pericolosa-e-non-testata-45697ab8-7316-46d3-bc01-60de5bb74xlk.shtml [23 Ottobre 2024]

238. Nato Cooperative Cyber Defence Centre of Excellence, (2017). Tallinn Manual 2.0 on the International Law Applicable to Cyber Operations, Cambridge University Press.

239. Nato, (2025). Statement of condemnation by the North Atlantic Council concerning Russian malicious cyber activities. Disponibile in: https://www.nato.int/cps/en/natohq/official_texts_237067.htm [18 luglio 2025].

240. Newsguard, (2025). Centro di monitoraggio sull'IA: oltre 1.200 siti inaffidabili di 'notizie generate dall'Intelligenza Artificiale' (in continua crescita) e le principali narrazioni false prodotte da strumenti basati sull'IA. Newsguard. Disponibile in: https://www.newsguardtech.com/it/special-reports/centro-monitoraggio-ia/ [10 febbraio 2025]

241. Nilus, S. A. (2023). Protocolli dei savi di Sion. Milano, Edizioni Clandestine. Il testo originale comparve come ultimo capitolo del libro *Velikoe v malom i antichrist,*

kak blizkaja političeskaja vozmožnost'. Zapiski pravoslavnogo (*Il Grande all'interno del Piccolo e Anticristo, un'Imminente Possibilità Politica. Note di un Credente Ortodosso*), sull'avvento dell'Anticristo pubblicato nel 1903.

242. Ninotti, L., Colatin De Tomas, S. (2022). Analysis of the Russian-Speaking Threat Actor NoName 057(16), Yarix. Disponibile in: https://labs.yarix.com/2022/10/analysis-of-the-russian-speaking-threat-actor-noname-05716/ [13 ottobre 2022]

243. Olson, P. (2012). We Are Anonymous: Inside the Hacker World of LulzSec. New York: Little, Brown and Company.

244. Open AI. (2025). Disrupting malicious use of AI, Disponibile in: https://openai.com/global-affairs/disrupting-malicious-uses-of-ai-june-2025/ [5 giugno 2025].

245. Ottaviani, M. F. (2022). Brigate Russe. La guerra occulta del Cremlino tra troll e hacker. Milano: Ledizioni LediPublishing.

246. Paquinelli, M. a cura di, (2002). Media Activism. Strategie e pratiche della comunicazione indipendente, Roma: DeriveApprodi.

247. Perry, B., (2015). Non-Linear Warfare in Ukraine: The Critical Role of Information Operations and Special Operations, Small Wars Journal,11(8) (2015), p. 4–5.

248. Piccolo, R. (2024). Cosa sappiamo sui due cavi sottomarini tranciati nel Mar Baltico. Wired. Disponibile in: https://www.wired.it/article/cavi-sottomarini-baltico-russia/ [20 novembre 2024].

249. Piccolo, R. (2025), Il rapporto di Francesca Albanese, relatrice delle Nazioni Unite, accusa le big tech di sostenere l'occupazione di Israele. Wired Italia. Disponibile in https://www.wired.it/article/rapporto-francesca-albanese-aziende-big-tech-profitto-israele/ [3 luglio 2025].

250. Piro, N. (1998). Cyberterrorismo. Come si organizza un rapimento virtuale. Roma: Castelvecchi.

251. Pisa, P. (2022). L'Ucraina pensa di fermare la guerra inviando foto dei soldati morti alle loro madri in Russia. La Repubblica. Disponibile in: https://www.repubblica.it/tecnologia/2022/04/22/news/ucraina_clarview_riconoscimento_facciale_soldati_russi_morti-346452874/ [22 aprile 2022].

252. Pisa, P. (2024). In questo momento l'arma più pericolosa dell'esercito israeliano è l'IA. La Repubblica. Disponibile in: https://www.repubblica.it/tecnologia/2024/12/30/news/israele_gaza_intelligenza_artificiale_gospel_errori_rischi-423913614/?utm_source=firefox-newtab-it-it [30 dicembre 2024]

253. Preziosa, P. (2025). Il silente avvelenamento dei pozzi digitali. Il nuovo sabotaggio cognitivo secondo Preziosa. Formiche. Disponibile in:

https://formiche.net/2025/07/il-silente-avvelenamento-dei-pozzi-digitali-il-nuovo-sabotaggio-cognitivo-secondo-preziosa/#content [14 luglio 2025]

254. Rainews24. (2023). Papa Francesco col piumino bianco da trapper fa milioni di clic: la foto è un fake, Radio Televisione Italiana, disponibile in: https://www.rainews.it/articoli/2023/03/papa-francesco-col-piumino-bianco-da-trapper-fa-milioni-di-clic-la-foto--fake--d054802c-30ae-4632-9d5a-193567a29726.html [27 marzo 2023].

255. Rainews24, (2025). "Ecco la Gaza del futuro", il video della ministra israeliana trasforma la Striscia in Las Vegas. Rainews24. Disponibile in: https://www.rainews.it/video/2025/07/ecco-la-gaza-del-futuro-il-video-della-ministra-israeliana-trasforma-la-striscia-in-las-vegas--3f6370bb-b939-45f4-aedd-b8c73d8e6bce.html

256. Ramkunar, A., Bade, G. (2025). Trump Goes to Bat for Big Tech in Global Trade Talks. The Wall Street Journal. Disponibile in: https://www.wsj.com/politics/policy/trump-trade-talks-tech-industry-e1061e53?mod=djemCybersecruityPro&tpl=cs [22 luglio 2025].

257.

258. Rampini, F. (2015). Rete padrona. Amazon, Apple, Google & co. Il volto oscuro della rivoluzione digitale. Milano: Feltrinelli.

259. Rapetto, U., Di Nunzio, R. (2001). Le Nuove Guerre. Dalla Cyberwar ai Black Bloc, dal sabotaggio mediatico a Bin Laden. Milano: RCS Libri.

260. Red Hot Cyber, redazione (2025). È cyberwar tra Israele ed Iran! APTIran colpisce Israele e avverte sul pericolo delle backdoor, Red Hot Cyber, disponibile in: https://www.redhotcyber.com/post/e-cyberwar-tra-israele-ed-iran-aptiran-colpisce-israele-e-avverte-sul-pericolo-delle-backdoor/ [17 giugno 2025].

261. Reeves, B., Nass, C. (1996), The Media Equation: How People Treat Computers, Television, and New Media like Real People and Places, Cambridge University Press: Cambridge.

262. Reporters' sans frontiers (2025), Gaza: RSF, CPJ and over 200 media outlets call for opening the Strip to foreign journalists and protecting Palestinian reporters. Reporters' sans frontiers. Disponibile in: https://rsf.org/en/gaza-rsf-cpj-and-over-130-media-outlets-call-opening-strip-foreign-journalists-and-protecting [update 12 giugno 2025].

263. Remocontro (2025). Le bombe Israeliane su Gaza e la diplomazia Usa contro l'analista Onu. Libera Informazione. Disponibile in: https://www.liberainformazione.org/2025/07/16/le-bombe-israeliane-su-gaza-e-la-diplomazia-usa-contro-lanalista-onu/

264. Rid, T. (2022). Misure Attive. Storia segreta della disinformazione, Roma, Luiss University Press; tit. or. Active Measures: The Secret History of Disinformation and Political Warfare, Ferrar Straus & Giroux 2021.

265. Robinson-Erly, N., (2024), How Israel uses facial-recognition systems in Gaza and beyond, The Guardian, disponibile online https://www.theguardian.com/technology/2024/apr/19/idf-facial-recognition-surveillance-palestinians [19 aprile 2024]

266. Rodrigues Fowler, Y., Goodman, C., (2017). How Tinder Could Take Back the White House, The New York Times, Disponibile in: https://www.nytimes.com/2017/06/22/opinion/how-tinder-could-take-back-the-white-house.html [22 giugno 2017]

267. Rodriguez, S. (2021). "Twitter locks Trump's account following video addressing Washington rioters". CNBC, disponibile in: https://www.cnbc.com/2021/01/06/twitter-pledges-action-on-any-calls-for-violence-in-capitol-riot.html [6 gennaio 2021]

268. Michael Robinson, Kevin Jones, Helge Janicke, Cyber warfare: Issues and challenges, Computers & Security, Volume 49, 2015, Pages 70-94, ISSN 0167-4048, https://doi.org/10.1016/j.cose.2014.11.007.

269. Rumiati, R. (1990). Giudizio e Decisione. Teorie e applicazioni della psicologia della decisione. Bologna, Il Mulino.

270. Santaniello, M. & Palladino, N. (2022). Discourse Coalitions in Internet Governance: Shaping Global Policy by Narratives and Definitions. In Internet Diplomacy. Shaping the Global Politics of Cyberspace Pag.61-84 Lanham, Maryland, United States Rowman & Littlefield Publishers. ISBN:978-1-5381-6117-3

271. Santaniello, M. (2021). Lower the top. La sfida alle piattaforme digitali, tra sovranità statuale e saperi sociali. In Comunicazione Punto doc. (2021). Lower the top. La sfida alle piattaforme digitali, tra sovranità statuale e saperi sociali. N. 25 ago-dic. 2021. Bologna Fausto Lupetti editore.

272. Savelli, F. (2022). Così Mosca si prepara a lasciare Internet (e cos'è Runet: l'intranet per soli russi). Il Corriere della Sera, disponibile in: https://www.corriere.it/economia/consumi/22_marzo_10/cosi-mosca-si-prepara-lasciare-internet-cos-runet-l-intranet-soli-russi-41b0e326-9fc6-11ec-82d5-6f137f6a69fd.shtml [10 marzo 2022].

273. Sbaraglia, G. (2022). Cyber Security. Kit di sopravvivenza. Seconda edizione, pp.170-173. Firenze, goWare.

274. Security Scorecard, (2025). From the Depths of the Shadows: IRGC and Hacker Collectives Of The 12-Day War. Security Scorecard. Disponibile in: https://securityscorecard.com/blog/from-the-depths-of-the-shadows-irgc-and-hacker-collectives-of-the-12-day-war/ [5 agosto 2025]

275. Servizio Operazioni e gestione delle crisi cibernetiche, (2024). La tassonomia cyber dell'ACN. Agenzia per la cybersicurezza nazionale. Disponibile in: https://www.acn.gov.it/portale/w/la-tassonomia-cyber-dellacn [31 luglio 2024]

276. Servizio Operazioni e gestione delle crisi cibernetiche, (2025). Operational summary primo semestre 2025. Dati e indicatori della minaccia cyber in Italia. Agenzia per la cybersicurezza nazionale. Disponibile in: https://www.acn.gov.it/portale/w/operational-summary-1-semestre-2025 [4 agosto 2025]

277. Robert Schuman Centre. (2025). Global Risk to the EU. European university institute (EUI) in Florence. Disponibile in: https://europeangovernanceandpolitics.eui.eu/global-risks-eu/

278. Schulz, J. (2025). Cybercriminal abuse of large language models. Cisco Talos. Disponibile in: https://blog.talosintelligence.com/cybercriminal-abuse-of-large-language-models/ [25 giugno 2025]

279. Sessa, M.G. (2025). Slopaganda: la costruzione del consenso sintetico ai tempi dell'IA. Gli Stati Generali. Disponibile in: https://www.glistatigenerali.com/tecnologia-e-media/a-i/slopaganda/ [10 ottobre 2025].

280. Simone, R. (2008). Il mostro mite. Perché l'Occidente non va a sinistra. Milano: Garzanti.

281. Smith, B., Brown, C. A. (2019). Tools and Weapons: The Promise and The Peril of the Digital Age. Londra: Hodder & Stoughton/Hachette

282. SOCRadar (2025). Iran-Israel conflict, threat landscape report (2025). Socradar. Disponibile in: https://socradar.io/resources/report/iran-israel-conflict-threat-landscape-report/ [30 giugno 2025].

283. SkyTg24 (2025). Gaza secondo Trump, il video realizzato con l'AI con statue d'oro e resort di lusso, disponibile in: https://tg24.sky.it/mondo/2025/02/26/gaza-trump-video-intelligenza-artificiale [26 febbraio 2025].

284. Stato Maggiore della Difesa (2023). Cognitive Warfare. La competizione nella dimensione cognitiva. SMD. Disponibile in: https://www.difesa.it/assets/allegati/31787/4.cognitive_warfare_-_la_competizione_nella_dimensione_cognitiva._ed.2023.pdf [febbraio 2023].

285. Strano Network, (1996). Net strike, no copyright, Et (:- Pratiche antagoniste nell'era telematica, Firenze, AAA Edizioni.

286. Sunstein, C. (2003). Republic.com. Cittadini informati o consumatori di informazioni? Bologna: IL Mulino; ed.or. 2001, Republic.com. Princeton-Oxford: Oxford University Press.

287. Sun Tzu, (V sec. A.C.). The Art of War.

288. Swascan, (2023). BiBi Wiper: analisi malware. Swascan. Disponibile in: https://www.tinextacyber.com/bibi-wiper-analisi-malware/ [18 dicembre 2023].

289. Tass, (2025). Russia moves to create national messaging app as part of digital services platform. Disponibile in: https://tass.com/economy/1980327 [24 giugno 2025].

290. Teti, A. (2024). China Intelligence. Tecniche, strumenti e metodologie di spionaggio e controspionaggio della Repubblica Popolare Cinese. Soveria Mannelli: Rubbettino.

291. Tegmark, M., (2018). Vita 3.0. Essere umani nell'era dell'intelligenza artificiale. Milano: Raffaello Cortina Editore.

292. Thales, (2022). Thales presents the 2022 Thales Cyberthreat Handbook. Thales. Disponibile in: https://www.thalesgroup.com/en/worldwide/security/press_release/thales-presents-2022-thales-cyberthreat-handbook [31 maggio 2022]

293. The Hacker News, (2025). U.S. Agencies Warn of Rising Iranian Cyber Attacks on Defense, OT Networks, and Critical Infrastructure. THN. Disponibile in: https://thehackernews.com/2025/06/us-agencies-warn-of-rising-iranian.html [30 giugno 2025]

294. The Washington Post (2024) Israel built an 'AI factory' for war. It unleashed it in Gaza. The Washington Post. Disponibile in: https://www.washingtonpost.com/technology/2024/12/29/ai-israel-war-gaza-idf/)

295. Tiddi, A. (2002). Precari. Percorsi di vita tra lavoro e non lavoro, Roma, DeriveApprodi.

296. Tiku, N., (2022). The Google engineer who thinks the company's AI has come to life. The Washington Post. Disponibile in: https://www.washingtonpost.com/technology/2022/06/11/google-ai-lamda-blake-lemoine/ [11 giugno 2022].

297. Timberg, J. (2017). As a conservative Twitter user sleeps, his account is hard at work, The Washington Post, Disponibile in: https://www.washingtonpost.com/business/economy/as-a-conservative-twitter-user-sleeps-his-account-is-hard-at-work/2017/02/05/18d5a532-df31-11e6-918c-99ede3c8cafa_story.html [5 febbraio 2017].

298. Tozzi, T. (2019). Le radici dell'hacktivism in Italia, 1969-1989. Dallo sbarco sulla Luna alla caduta del muro di Berlino, Firenze: Accademia di Belle Arti di Firenze.

299. Trend Micro. (2025). The AI-fication of Cyberthreats: Trend Micro Security Predictions for 2026. Trend Micro. Disponibile in: https://www.trendmicro.com/vinfo/us/security/research-and-analysis/predictions/the-ai-fication-of-cyberthreats-trend-micro-security-predictions-for-2026 [25 novembre 2025]

300. Treyger, E., Cheravitch, J. & Cohen, R. S., (2022) Russian disinformation Effort on social media, Rand Corporation.

301. Turing, A. M. (1950) Computing Machinery and Intelligence. *Mind 49:* 433-460.

302. United Nations (2025). Bulletin on actions by the United Nations and other intergovernmental organizations relevant to the questions of Palestine. Disponibile in: https://www.un.org/unispal/document/action-by-un-system-and-intergovernmental-organizations-relevant-to-the-question-of-palestine-july-2025-monthly-bulletin/#_Toc205896782. United Nations [July 2025]

303. United Nations, Human Right Council (2025). From economy of occupation to economy of genocide. Disponibile in: https://www.un.org/unispal/document/a-hrc-59-23-from-economy-of-occupation-to-economy-of-genocide-report-special-rapporteur-francesca-albanese-palestine-2025/ **[Giugno-luglio 2025]**

304. Uva, V. (2025). Privacy e digitale, aziende nel labirinto delle regole Ue. Il Sole 24 Ore, 25 agosto 2025

305. Van der Klaauw, C. (2023). Cognitive warfare. The 21st-Century Game Changer. NATO, Disponibile in**:** https://studylib.net/doc/27295928/nato-cognitive-warfare--the-21st-century-game-changer--2023- [2023]

306. Van Dijck, Poell, T., De Waal, M. (2019). Platform society. Valori pubblici e società connessa. Edizione italiana a cura di Giovanni Boccia Artieri e Alberto Marinelli. Roma, Guerini Scientifica.

307. Van Dijck, J 2021, 'Seeing the forest for the trees: visualizing platformization and its governance', *New Media & Society*, vol. 23, no. 9, pp. 2801–2819

308. Vecchi, B. (2017). Il capitalismo delle piattaforme. Roma: Manifestolibri – la talpa s.r.l.

309. Veneziani, M. (2006). Controinformazione. Stampa Alternativa e giornalismo d'inchiesta dagli anni Settanta ad oggi. Roma: Castelvecchi.

310. Waldman. Katy, (2025). Trump Is the Emperor of A.I. Slop. New Yorker. Disponibile in: https://www.newyorker.com/culture/critics-notebook/trump-is-the-emperor-of-ai-slop [26 aprile 2025]

311. Wardle, C., Derakhshan, H. (2017). Information disorder: Toward an interdisciplinary framework for research and policy making. Council of Europe report.

312. Watts, C. (2024). Iran accelerates cyber ops against Israel from chaotic start. Microsoft. Disponibile in: https://blogs.microsoft.com/on-the-issues/2024/02/06/iran-accelerates-cyber-ops-against-israel/ [6 febbraio 2024]

313. Jason Wei, Xuezhi Wang, Dale Schuurmans, Maarten Bosma, Brian Ichter, Fei Xia, Ed Chi, Quoc Le, & Denny Zhou (2023). Chain-of-Thought Prompting Elicits Reasoning in Large Language Models, arXiv, https://arxiv.org/abs/2201.11903

314. White, G. (2022). Crime dot com. Il potere globale dell'hacking dai virus ai brogli elettorali, Città di Castello, Odoya; (ed or. Crime Dot Com: From Viruses to Vote Rigging, How Hacking Went Global, 2020, Reaktion publishing).

315. Wilye, C. (2019). Il Mercato del consenso: come ho creato e poi distrutto Cambridge Analytica, Milano, Longanesi; ed. or. Mindf*ck: Inside Cambridge Analytica's Plot to Break the World, London, Profile Books, 2019.

316. Samuel C. Woolley & Philip N. Howard, "Computational Propaganda Worldwide: Executive Summary." Samuel Woolley and Philip N. Howard, Eds. Working Paper 2017.11. Oxford, UK: Project on Computational Propaganda. comprop.oii.ox.ac.uk. 14 pp.

317. World Economic Forum. (2025). Global Risks Report 2025. WEF. Disponibile in: https://www.weforum.org/publications/global-risks-report-2025/ [15 gennaio 2025]

318. Yi Liu, Gelei Deng, Zhengzi Xu, Yuekang Li, Yaowen Zheng, Ying Zhang, Lida Zhao, Tianwei Zhang, Kailong Wang, Yang Liu (2023). Jailbreaking ChatGPT via Prompt Engineering: An Empirical Study. ArXiv. Disponibile in: https://arxiv.org/abs/2305.13860

319. Zetter, K. (2015). Countdown to zero Day: Stuxnet and the Launch of the World's First Digital Weapon, New York, Crown.

320. Zorloni, L. (2023). Cosa sappiamo sulle dimissioni del direttore dell'Agenzia per la cybersicurezza nazionale. Wired. Disponibile in: https://www.wired.it/article/baldoni-agenzia-cybersicurezza-dimissioni-soldi-pnrr/ [7 marzo 2023]

321. Zuboff, S. (2019). Il capitalismo della sorveglianza. Il futuro dell'umanità nell'era dei nuovi poteri. Roma: Luiss University Press; ed or. The Age of Surveillance Capitalism. The Fight for the Future at the New Frontier of Power. USA. Public Affairs (2019).

322. Ziv, O. (2025). How a Scottish maritime museum ended up in Israel's 3D propaganda videos. +972. Disponibile in https://www.972mag.com/israeli-army-3d-propaganda-animations/ [8 ottobre 2025]

323. .Zip, (1997). Hot Web. Guida ai siti alternativi e radicali su Internet. Roma: Castelvecchi.

324. Zitelli, A. (2025). Il nuovo piano della disinformazione russa per "infettare" i chatbot AI utilizzati in Occidente. Facta News. Disponibile in: https://www.facta.news/articoli/pravda-disinformazione-russa-chatbot-ia [11 luglio 2025]

15. Ringraziamenti

Questo lavoro di ricerca non sarebbe stato possibile senza lo sprone del professor Alberto Marinelli a cui vanno i miei più sentiti ringraziamenti. Subito dopo ringrazio il professor Paolo Prinetto per avermi aiutato a strutturare la tesi, il professore Mauro Santaniello che l'ha letta dandomi delle utili indicazioni, il professore Christian Ruggiero che non mi ha mai fatto mancare il suo supporto e poi i referee della tesi, il professore Giovanni Boccia Artieri e il professore Michele Mezza.
Un grazie speciale lo rivolgo ai testimoni intervistati e cioè a Barbara Carfagna, Luca Sambucci, Marco Ramilli, Massimo Marotti e a Michele Mezza.

Voglio anche ringraziare sentitamente il professor Roberto Baldoni che mi ha insegnato buona parte delle cose di cui parlo nel libro. Un ringraziamento affettuoso va anche al Direttore Generale di ACN, Bruno Frattasi: lavorare con lui mi ha arricchito professionalmente e culturalmente.

Desidero ringraziare anche l'ispettore Domenico Runieri di ACN che mi ha sostenuto con le sue attenzioni in momenti difficili.

L'ultimo necessario ringraziamento va a mia moglie, Annamaria Lena, il cui incitamento è stato sempre un toccasana. E, naturalmente, ai miei genitori che mi hanno permesso di studiare all'Università, il primo della famiglia.