

NEW FACE OF IR: **SOCIAL MEDIA** AND DIGITALISATION

Authors 76
ARDA ERCAN 6



BIDGE Publications

NEW FACE OF IR: SOCIAL MEDIA AND DIGITALISATION

Authors: ARDA ERCAN

ISBN: 978-625-372-875-5

1st Edition

Page Layout By: Gozde YUCEL

Publication Date: 20.12.2025

BIDGE Publications

All rights reserved. No part of this work may be reproduced in any form or by any means, except for brief quotations for promotional purposes with proper source attribution, without the written permission of the publisher and the editor.

Certificate No: 71374

All rights reserved © BIDGE Publications

www.bidgeyayinlari.com.tr - bidgeyayinlari@gmail.com

Krc Bilişim Ticaret ve Organizasyon Ltd. Şti.

Güzeltepe Mahallesi Abidin Daver Sokak Sefer Apartmanı No: 7/9 Çankaya / Ankara



İçindekiler

Chapter One : The Digital Turn in Global Affairs	7
Introduction	7
The Evolution of Social Media in International Relations	7
The Rise of Twiplomacy	8
Artificial Intelligence: The New Frontier of IR	8
The Convergence: AI-Driven Digital Diplomacy	9
Digital Geneva Convention.....	9
The Tripartite Structure	10
Pillar I: State Commitment (The Treaty)	10
Pillar II: The Tech Sector (Digital Switzerland)	10
Pillar III: An Independent Attribution Organization	11
Theoretical and Legal Foundations.....	11
Challenges and Current Status (2025)	11
Case Study: "Digital Switzerland" in Action (2022–2025)	12
The 2025-2026 UN Cyber-Treaty Negotiations.....	12
The "Attribution Challenge" and AI	13
Conclusion	14
Chapter Two: The Constitutive Web: Social Media, Identity, and the Transformation of International Relations	15
The Digital Turn in International Relations Theory.....	15
Thesis Statement	16

The Theoretical and Empirical Scope of Inquiry	16
The Theoretical Nexus: Constructivism, Identity, and Digital Affordances	17
Identity and Intersubjective Structure in IR	17
Social Media Affordances and Identity Processes	17
Conceptualizing Collective Political Identity in the Digital Age	18
Identity Performance in Statecraft: Digital Diplomacy and Nation Branding : Digital Public Diplomacy (PD) and Strategic Projection	21
Visual Diplomacy and the Reinforcement of State Identity	21
Case Study: Targeted Identity Projection in Conflict.....	22
Ethical and Security Dilemmas in Digital Diplomacy	23
Transnationalism and Identity Mobilization: Non-State Actors as Constitutive Agents: The New Geopolitics of Non-State Actors (NSAs)	24
Digital Diasporas and Homeland Conflict Identity.....	24
Social Media as a Catalyst for Collective Identity Formation and Protest.....	25
The Dark Side of Digital Identity: Polarization, Manipulation, and Hybrid Warfare Weaponizing Identity and Hybrid Threats	26
Algorithmic Acceleration of Affective Polarization	27
Identity and Conflict: The Case of Ukraine	28
The Limits of Disinformation’s Strategic Effectiveness	29

Governance, Ethics, and the Future of the Digital Public Sphere: Digital Authoritarianism and Identity Control	30
The Ethical Imperative of Platform Design	31
Sovereignty, Content Moderation, and Trust	32
The Need for Digital Ethics in International Relations.....	33
Conclusion	33
Synthesis of Findings: Social Media as the New Arena for Identity Struggle	33
Theoretical Contributions and Policy Relevance	34
Future Research Agenda	34
Kaynakça.....	47
Chapter Three: The Digital Self On The World Stage: Social Media, Identity Construction, And The Transformation Of International Relations.....	37
Introduction.....	37
Theoretical Framework: Constructivism and the Digital Turn ..	37
Constructivism in the Digital Age.....	38
Soft Power vs. Sharp Power.....	38
The Architecture of Digital Diplomacy	38
Phase 1: Information Dissemination (Web 1.0)	39
Phase 2: Engagement and Twiplomacy (Web 2.0).....	39
Phase 3: Algorithmic Diplomacy and AI (Web 3.0)	39
Identity Construction in Conflict: The Narrative Wars	39

The Ukraine-Russia War: The First "TikTok War"	39
China's "Wolf Warrior" Diplomacy	40
Non-State Actors and the "Digital Nation"	40
Digital Diasporas.....	40
Terrorist Identity Construction.....	41
Challenges and Ethical Implications.....	41
Algorithmic Bias and Digital Orientalism	41
The Erasure of Truth	41
Conclusion	41
Case Study.....	43
The First "TikTok War" – Digital Identity and Narrative Control in the Russia-Ukraine Conflict.....	43
Introduction.....	43
Ukraine's Digital Strategy: The Connective Strategic Narrative	43
Zelenskyy and "Selfie Diplomacy"	43
Horizontal Mobilization: The IT Army	44
Russia's Digital Strategy: The Firehose of Falsehood	44
Fragmented Identities.....	44
Cyber-Enabled Disinformation: "Ghostwriter"	45
Sub-Case: Meme Warfare and "Saint Javelin"	45
Conclusion: The Asymmetry of Authenticity	45

Chapter One : The Digital Turn in Global Affairs: Analyzing the Intersection of Social Media, Twiplomacy, and Artificial Intelligence in International Relations

Introduction

The traditional Westphalian system of International Relations (IR), characterized by formal state-to-state interactions and closed-door negotiations, is undergoing a radical paradigm shift. The advent of the Information Age has eroded the state's monopoly on information, giving rise to a new diplomatic ecosystem where power is increasingly defined by connectivity and narrative control rather than coercion alone. This essay analyzes the concept of "Social Media IR Relations," a term denoting the integration of digital communication platforms into the fabric of foreign policy. By synthesizing the phenomena of "Twiplomacy" and the emergent capabilities of Artificial Intelligence (AI), this study elucidates how digital tools are redefining diplomatic norms, altering the speed of political engagement, and creating new security challenges.

The Evolution of Social Media in International Relations

Social media has transitioned from a peripheral communication tool to a central arena for international politics. In the theoretical framework of Constructivism, IR is shaped by shared ideas, identities, and norms (Wendt, 1999). Social media accelerates the construction and deconstruction of these norms by allowing immediate, unmediated interaction between global leaders and foreign publics. This phenomenon extends Joseph Nye's concept of "Soft Power"—the ability to attract and persuade—into the digital

realm, often termed "Cyber Soft Power" or "Digital Diplomacy" (Nye, 2004).

Unlike the "CNN Effect" of the 1990s, where 24-hour news cycles pressured politicians to react, the "Social Media Effect" facilitates real-time interaction and mobilization. The Arab Spring of 2011 serves as a seminal case study where social media platforms facilitated the rapid organization of non-state actors, challenging state sovereignty and demonstrating the disruptive potential of digital networks (Khondker, 2011).

The Rise of Twiplomacy

A specific subset of digital diplomacy, "Twiplomacy" (Twitter Diplomacy), refers to the use of Twitter (now X) and other micro-blogging sites by world leaders and international organizations to conduct diplomatic outreach. Twiplomacy represents a flattening of diplomatic hierarchy, allowing heads of state to bypass traditional foreign ministries and speak directly to global audiences (Manor, 2019).

This direct engagement offers significant advantages, including the ability to control state narratives, gauge public sentiment instantly, and demonstrate transparency. However, it also introduces "diplomacy by hashtag," where complex geopolitical nuances are reduced to 280-character soundbites. This brevity can lead to misunderstandings and the rapid escalation of tensions, as observed in various diplomatic spats where impulsive tweets replaced measured diplomatic *démarches*. The informality of Twiplomacy often blurs the line between personal opinion and official state policy, creating ambiguity in international law and protocol.

Artificial Intelligence: The New Frontier of IR

The integration of Artificial Intelligence (AI) into IR signifies a shift from human-centric to algorithmically-augmented diplomacy. AI influences IR primarily through three mechanisms: predictive analytics, autonomous systems, and information warfare.

First, AI enhances decision-making by processing vast datasets to predict political instability, migration patterns, or economic shifts, allowing states to engage in proactive rather than reactive diplomacy. Second, in the security domain, AI-driven autonomous weapon systems (LAWS) challenge the ethical and legal frameworks of war, prompting debates on accountability and the dehumanization of conflict.

Third, and most critical to social media relations, is the use of AI in narrative control. AI algorithms now curate the information consumed by global audiences, creating "echo chambers" that can reinforce polarization. Furthermore, "Deepfakes"—hyper-realistic AI-generated audio or video—pose an existential threat to diplomatic trust. A fabricated video of a world leader declaring war or making offensive remarks could trigger irreversible diplomatic fallout before verification is possible (Chesney & Citron, 2019).

The Convergence: AI-Driven Digital Diplomacy

The intersection of social media and AI creates a potent feedback loop. AI bots and "troll farms" can artificially amplify specific narratives, creating a false impression of public consensus—a tactic known as "astroturfing." This manipulation undermines the democratic potential of Twiplomacy, turning social media platforms into battlegrounds for cognitive warfare.

In this environment, the "Audience" in IR becomes a contested strategic asset. States utilize AI to micro-target foreign populations with tailored propaganda, effectively weaponizing the open nature of the internet. This necessitates a re-evaluation of sovereignty, as digital borders are porous and easily penetrated by algorithmic influence operations.

Digital Geneva Convention

The concept of a "**Digital Geneva Convention**" (DGC) was first formally proposed by Microsoft President Brad Smith at the 2017 RSA Conference. It aims to create a legal framework that mirrors the original 1949 Geneva Convention, which established protections for civilians during traditional warfare. In the context of

International Relations (IR), the DGC addresses the "Wild West" nature of cyberspace, where nation-state attacks on civilian infrastructure often go unpunished.

The following is a detailed breakdown of the proposal's core pillars, its implementation mechanisms, and its status as of 2025.

The Tripartite Structure

The DGC is not envisioned as a single treaty but as a three-layered ecosystem involving governments, the technology sector, and a neutral oversight body.

Pillar I: State Commitment (The Treaty)

Governments would sign a multilateral treaty pledging to exercise restraint in cyberspace. The key commitments include:

No Targeting of Civilians: States must refrain from attacking critical infrastructure (power grids, hospitals, water systems) and the global financial system.

Non-Proliferation: A commitment to not distribute or sell cyberweapons and to restrict their development to precise, non-replicable tools.

Reporting Vulnerabilities: Rather than stockpiling "Zero-Day" vulnerabilities for offensive use, states would agree to report them to vendors to improve global security.

Pillar II: The Tech Sector (Digital Switzerland)

This pillar calls for the global technology industry to act as a "**Digital Switzerland**"—a neutral party that protects and assists all users regardless of their nationality or the identity of the attacker.

Neutrality: Tech companies would refuse to assist governments in offensive cyber operations.

First Responders: Similar to the Red Cross in physical war, tech firms would act as "digital first responders," patching systems and defending users during state-sponsored attacks.

Pillar III: An Independent Attribution Organization

The most innovative (and controversial) aspect is the creation of a global, independent body of experts from the private sector, academia, and civil society.

Technical Attribution: This body would investigate cyberattacks to provide objective, public evidence of which state was responsible.

The "Shaming" Mechanism: By providing credible attribution, this body would create political costs for rogue states, similar to how the IAEA monitors nuclear activity.

Theoretical and Legal Foundations

The DGC builds upon several existing frameworks within International Relations:

The UN GGE & OEWG: It seeks to codify the "11 voluntary norms of responsible state behavior" established by the UN Group of Governmental Experts into binding international law.

The Tallinn Manual: While the Tallinn Manual interprets *how* existing international law applies to cyberwarfare, the DGC seeks to create *new* specific protections for the digital age.

Paris Call for Trust and Security: Many principles of the DGC were later integrated into the **Paris Call (2018)**, which now has over 1,200 signatories including states, companies, and NGOs.

Challenges and Current Status (2025)

As of late 2025, the DGC remains a "soft law" framework rather than a universally ratified treaty. Several factors hinder its full adoption:

The Attribution Problem: Despite AI advancements, definitive "technical" attribution remains difficult to separate from "political" attribution, as states can use proxies or false-flag operations.

Sovereignty vs. Censorship: Authoritarian states often view digital norms as a way for Western powers to impose "internet freedom" standards that threaten their domestic control.

The AI Acceleration: The 2025 landscape is dominated by **AI-enabled offensive cyber tools**. This has led to the **Geneva Cyber Week 2025**, where discussions shifted toward the "Dystopia Scenario"—the risk of AI systems autonomously launching cascading attacks on essential services.

Case Study: "Digital Switzerland" in Action (2022–2025)

The concept of technology companies acting as a "Digital Switzerland"—neutral, non-combatant defenders—transitioned from theory to necessity during the conflict in Ukraine and subsequent regional tensions in 2024.

Evacuation of Sovereign Data: In a landmark application of DGC principles, Microsoft and Amazon assisted the Ukrainian government in migrating critical state registries and ministerial data to "sovereign cloud" centers across Europe. This effectively created a **"Digital Government-in-Exile,"** ensuring that even if physical ministries were destroyed, the state's digital identity remained intact.

The Neutrality Paradox: While the DGC calls for tech companies to remain neutral, 2025 has seen a "Strategic Alignment" shift. Companies now often prioritize **Digital Resilience** for civilians over strict political neutrality. For instance, by providing real-time threat intelligence to civilian infrastructure (hospitals, power grids), tech firms act as "digital red crosses," even if those actions indirectly disadvantage an aggressor state.

The 2025-2026 UN Cyber-Treaty Negotiations

The diplomatic momentum has shifted from "voluntary norms" to a binding **UN Convention against Cybercrime**, which opened for signature in Hanoi in October 2025. This treaty serves as the legislative backbone that the DGC lacks.

AI-Specific Clauses: A critical addition to the 2026 framework is the inclusion of **"AI-Enabled Misuse"** clauses. These address:

Autonomous Malicious Actors: Legal frameworks that hold states accountable if their "autonomous" AI agents trigger systemic

financial or infrastructure failures, even without a direct "human-in-the-loop" command.

Deepfake Prohibitions: Proposed mandates for "digital watermarking" of AI-generated content used by diplomatic entities to prevent the erosion of trust in Twiplomacy.

Technological Neutrality: Negotiators have adopted "technologically neutral" language. Instead of banning specific AI models (which would be obsolete by the time the treaty is ratified), the focus is on **Criminal Intent and Effect**—criminalizing the *result* of a digital attack on civilians regardless of whether it was executed via a botnet or a generative AI agent.

Comparison of Frameworks

Feature	Original Geneva Convention (1949)	Digital Geneva Convention (Proposal)	UN Cybercrime Treaty (2025/26)
Primary Actor	Nation-States	States + Tech Giants	Nation-States (UN Members)
Core Value	Physical Neutrality	Digital Resilience/Neutrality	Law Enforcement Cooperation
Enforcement	International Court of Justice	Market/Social Reputation	Mutual Legal Assistance (MLA)
AI Focus	None	Algorithmic Accountability	AI-Enabled Evidence Sharing

The "Attribution Challenge" and AI

One of the DGC's pillars is the **Independent Attribution Organization**. In 2025, AI has made this both easier and more difficult:

AI for Defense: AI can now trace the "coding fingerprint" of an attack back to specific state-sponsored groups in milliseconds.

AI for Deception: Conversely, AI allows states to use "Large Language Model (LLM) Wrappers" to mimic the attacking style of *other* countries, creating a crisis of "False-Flag Diplomacy."

Conclusion

Social Media IR relations represent more than a technological upgrade; they constitute a fundamental restructuring of how states interact. Twiplomacy has dismantled the walls of the embassy, inviting the public into the diplomatic process, while AI has provided the tools to monitor, influence, and manipulate that public on an unprecedented scale. As these technologies mature, the discipline of International Relations must adapt to a "Post-Truth" era where the speed of information often outpaces the capacity for verification. Future diplomatic success will depend not only on traditional statecraft but also on "digital literacy"—the ability to navigate the complex, algorithmically driven information environment while maintaining the integrity of international norms.

The Digital Geneva Convention represents a shift toward **"Multi-stakeholder Diplomacy."** It acknowledges that in the digital era, private tech giants hold more "territory" (servers and data) than some nations. Therefore, a secure global order requires a partnership where the private sector is not just a participant, but a protected and neutral arbiter of digital peace.

Also the Digital Geneva Convention is evolving from a idealistic proposal by a single corporation into a **multi-stakeholder governance model**. In the IR landscape of 2025, "sovereignty" is no longer just about territory; it is about **Digital Sovereignty**—the ability of a state to protect its data, its algorithms, and its citizens' reality from foreign digital interference.

Chapter Two: The Constitutive Web: Social Media, Identity, and the Transformation of International Relations

The Digital Turn in International Relations Theory

The study of International Relations (IR) has historically centered on material power, rational choice, and institutional structures. The advent and global pervasiveness of social media platforms, however, represent a structural shift that demands a fundamental reassessment of these traditional frameworks. Social media has transcended its initial role as a mere communication tool, evolving into a pervasive technological and social environment that actively constitutes political and social reality (Clapton, 2025). This digital turn necessitates scholarly engagement, particularly in addressing the "cultural basis of political action and meaning-making" that is now deeply interwoven with digital platforms (Clapton, 2025).

For scholars rooted in Constructivism, social media offers a critical arena for analyzing how intersubjective beliefs and shared understandings are formed, maintained, and contested on a global scale. This research adopts a constructivist lens, positioning identity not merely as a characteristic of states but as an intersubjectively constructed group marker subject to constant flux and internal contestation (Adler-Nissen & Guzzini, 2025). Analyzing digital politics provides a crucial opportunity for constructivists to revitalize their research program, which has sometimes faced fragmentation under disciplinary pressures within IR (Adler-Nissen & Guzzini, 2025).

Thesis Statement

Social media platforms fundamentally alter the constitution, contestation, and strategic performance of political identity across state and non-state actors in the international system, challenging the stability of intersubjective norms and accelerating global political fragmentation and polarization.

The Theoretical and Empirical Scope of Inquiry

The prominence of social media has generated new dynamics in the international system that relate directly to competition for status and self-esteem. National identifications are characterized as a principal means by which leaders, peoples, and foreign actors seek status and build self-esteem, often stressing the alleged uniqueness of a nation (Mercer, 1995). Social media, through its highly visible and quantifiable metrics—such as retweets, quotes, and likes, which serve as measures of social visibility and approval (Khosro, 2025)—provides an immediate, transparent arena for this competition.

This environmental change shifts the diplomatic focus from achieving material policy outcomes to achieving social visibility and approval, a form of identity inflation where perceived legitimacy in the digital sphere may sometimes overshadow tangible material influence. Furthermore, a critical structural parallel exists between micro-level identity processes and macro-level political phenomena.

Social media profoundly impacts identity development in adolescents, facilitating introspection, storytelling, and dialogue—processes essential for forming commitments and constructing a coherent life story (Handbook of Adolescent Digital Media Use and Mental Health, 2025). However, this digital engagement also introduces identity distress, particularly due to the pressure of idealized self-presentation and engagement in social comparisons (Rhoda, 2025). This analysis suggests that the psychological mechanisms driving micro-identity crises—such as the struggle to reconcile idealized aspirations with reality (Rhoda, 2025)—are mirrored at the macro-level in state foreign policy and nation branding. States engage in national storytelling and diplomatic self-presentation (Duncombe, 2025), and when those idealized digital

narratives conflict with real-world circumstances, it can generate a collective form of identity distress, thereby justifying the integration of sociological theories of identity formation into the core of IR analysis.

The Theoretical Nexus: Constructivism, Identity, and Digital Affordances

Identity and Intersubjective Structure in IR

In Constructivist IR scholarship, identity is often considered the "master variable" driving state behavior (Mercer, 1995). Identities are not static, singular attributes of states, but rather multiple, fluid identifications that leaders and publics seek to impose consistent with their political projects and psychological needs (Mercer, 1995). Crucially, identities are not causes of policy so much as they are targets of constant domestic political contestation, often invoked as rationalizations for policies already pursued (Mercer, 1995).

The framework further posits that social norms constitute a significant aspect of social structure. These norms emerge from the actions and beliefs of actors within specific communities and, in turn, shape those actors' identities and interests (Oxford Research Encyclopedia, 2025). Early constructivist work successfully argued that shared ideas about appropriate state behavior have a profound impact on the nature and functioning of world politics (Oxford Research Encyclopedia, 2025). As the constructivist agenda has matured, scholars have shifted attention to compliance with and contestation over norms, underscoring the necessity of a prior understanding of who constitutes the community and the relationship between subjective and intersubjective reality (Oxford Research Encyclopedia, 2025). Social media directly intervenes in this relationship, becoming the primary site where intersubjective reality is negotiated.

Social Media Affordances and Identity Processes

The unique structural characteristics, or affordances, of social media platforms directly interact with the psychological and

social processes of identity formation. Research shows that digital media affects the dual aspect of identity development: the progression towards self-evaluated commitments and values, and the construction of a coherent life story, or narrative identity (Handbook of Adolescent Digital Media Use and Mental Health, 2025). Social media environments facilitate introspection, storytelling, and dialogue—the mechanisms through which these identity processes occur (Handbook of Adolescent Digital Media Use and Mental Health, 2025).

However, the efficacy of this engagement depends on the nature of use, rather than mere quantity. Active participation in social media, which involves content creation and meaningful engagement, is associated with greater identity exploration, whereas simply spending time on platforms is not (A Systematic Review of Social Media Use and Adolescent Identity Development, 2025). Furthermore, authenticity correlates positively with higher self-concept clarity, suggesting that idealized self-presentation can be counterproductive to coherent identity formation. Conversely, engaging in comparisons on social media correlates with increased identity exploration but also heightened identity distress (A Systematic Review of Social Media Use and Adolescent Identity Development, 2025). The digital environment, therefore, is not uniformly beneficial; it is a complex space where self-reinforcement coexists with the potential for psychological fragmentation.

Conceptualizing Collective Political Identity in the Digital Age

At the collective level, social media shapes the boundaries of political identity. Online communities centered around shared interests or political causes provide users with a strong sense of connection and identity, reinforcing their existing beliefs and fostering a sense of belonging (Rhoda, 2025). This group consolidation, however, often occurs at the cost of diversity and openness, contributing to the formation of polarized echo chambers (Rhoda, 2025). The media ecology created by these platforms serves to accelerate or obstruct serious political discussion and debate, tremendously expanding the parameters within which groups can

voice concerns and organize contentious political activities (Chapman University, 2025).

This dynamic demands that IR scholarship broaden its focus beyond traditional state-centric analyses. The "cultural basis of political action and meaning-making" now requires understanding the everyday politics of how audiences perceive and contest popular cultural artifacts, positioning popular culture not merely as intersecting with politics but as a direct object of political contestation (Clapton, 2025).

The ability of social media to rapidly restructure intersubjective reality represents a profound challenge to traditional IR temporalities. Constructivist change, which relies on the diffusion and institutionalization of new norms and identities (Oxford Research Encyclopedia, 2025), is typically gradual. However, recent experimental data on affective polarization demonstrates that changes in algorithmic curation can induce attitude shifts in users' feelings toward an out-party in a single week that are comparable in size to shifts normally observed over a three-year period (Northeastern University, 2025). This finding demonstrates that social media is not just a mechanism for the constitution of identity; it drastically accelerates the rate of constitutive change, compressing political time and thereby increasing the volatility and unpredictability of intersubjective structures in the international system.

This acceleration is particularly relevant when considering adversarial actions. Identity formation, at the national level, is a process of constructing a "coherent life story" (Handbook of Adolescent Digital Media Use and Mental Health, 2025). Adversarial state and non-state actors, leveraging social media, consciously seek to undermine this narrative coherence. These campaigns function as a form of "narrative counter-formation," strategically targeting the mechanisms of collective identity (storytelling, shared information) through the dissemination of false information designed to fracture, contest, and assault the national narrative reality (Public Diplomacy in Post-Truth Era, 2025). The effectiveness of disinformation lies in its characteristics—volume,

diversity of sources, speed, and repetition (Public Diplomacy in Post-Truth Era, 2025)—which induce widespread feelings of uncertainty and, by extension, collective identity distress, mirroring the negative self-perceptions observed in individuals subjected to online harassment (Rhoda, 2025).

Table 1 summarizes the constitutive relationship between core IR concepts and the digital environment.

Table 1: The Constitutive Link: Integrating Constructivism with Digital Affordances

Constructivist Concept	Definition (IR Theory)	Social Mechanism/Affordance	Media	Constituting Outcome in IR
Identity (National/Collective)	Multiple, fluid identifications invoked for political projects and status seeking (Mercer, 1995)	Storytelling, Presentation, Personalized Communication (Khoso, 2025)	Self-	Nation branding, diaspora mobilization, conflict visualization (New Lines Institute, 2025)
Norms/Social Structure	Shared ideas and beliefs constituting actors' identities and interests (Oxford Research Encyclopedia, 2025)	Content moderation policies, algorithmic prioritization, potent social norms (Handbook of Adolescent Digital Media Use and Mental Health, 2025)		Shaping acceptable diplomatic conduct, accelerating polarization (Northeastern University, 2025)
Intersubjective Reality	Shared understandings determining meaning, authority, and power (Adler-Nissen & Guzzini, 2025)	Social visibility (retweets, likes), viral emotional content (Khoso, 2025)		Validation of performative identity (e.g., digital diplomacy ritual) (Duncombe, 2025), fracturing of objective reality (Public Diplomacy in Post-Truth Era, 2025)

Identity Performance in Statecraft: Digital Diplomacy and Nation Branding : Digital Public Diplomacy (PD) and Strategic Projection

States recognize that soft power—the ability to influence others through attraction to shared values—is increasingly necessary in the global information age, particularly for addressing contemporary security threats (Diva Portal, 2011). Social media platforms have become an integral tool of this public diplomacy (PD), transforming how state actors communicate with international publics (Global Media Journal, 2025). This strategic approach, often termed "Diplomacy 2.0," has gained traction in foreign policy initiatives globally (Bjola, 2025).

States leverage digital platforms for "nation branding," promoting national interests and shaping global perceptions (Nation Branding, 2025). For example, the Chinese government utilizes its primary news agency, Xinhua, and English-language media outlets like China Daily and Global Times to maintain control over media coverage and promote national interests abroad (Nation Branding, 2025). This state engagement in soft power, often combined with hard power (such as economic sanctions or military threats), represents Joseph Nye's concept of smart power, recognizing that a fitted mixture of both elements is often the most effective way to exercise influence (Diva Portal, 2011).

Visual Diplomacy and the Reinforcement of State Identity

Digital diplomacy relies heavily on the performative nature of social media to project specific national identities. Foreign ministries utilize profile pictures and cover photos to communicate their mission and global outlook (University of Southern California, 2025). For instance, a ministry may project a contemporary 21st-century global perspective, while another may deliberately project a different, more nationally focused mission (University of Southern California, 2025).

Scholarly analysis of digital diplomacy increasingly views the myriad images uploaded daily by diplomats not merely as strategic communications but as "ritual, performative, and symbolic

markers" (Duncombe, 2025). The sharing of these visual elements performs a social aesthetic, designed to express and reproduce myths that convey authority, power, and privilege (Duncombe, 2025). In this sense, digital engagement often serves not to revolutionize or disrupt traditional diplomacy, but rather to reinforce its established ideals of international mediation and state authority (Duncombe, 2025).

A tension exists, however, between maintaining diplomatic gravitas and maximizing digital engagement. The use of humor and memes ("Selfie Diplomacy") can be effective for audience engagement but risks working against public expectations of diplomatic conduct, which traditionally revolve around requirements of decorum and sobriety (Manor, 2025). Such memetic engagement, if conducted without clear objectives regarding policy visibility or undermining an opponent's position, risks undermining the authority being ritually constructed (Manor, 2025).

The shift in the digital environment implies that the proximate measure of diplomatic success has been altered. While traditional diplomacy measured success by negotiated outcomes, digital diplomacy prioritizes social visibility and approval—metrics such as retweets and favorites (Khoso, 2025). This means that a Foreign Ministry's perceived success may shift toward achieving identity acceptance and status affirmation in the digital sphere. The objective, therefore, becomes influencing online narrative and image, potentially regardless of direct policy effects, because digital visibility enhances the viewability and legitimacy of content (Khoso, 2025).

Case Study: Targeted Identity Projection in Conflict

The strategic use of social media by state and quasi-state actors highlights the immediate link between identity performance and digital metrics. An analysis of the Taliban's Ministry of Foreign Affairs (MoFA) on Twitter demonstrated that employing cultural diplomacy significantly enhanced their social visibility and approval (Khoso, 2025). In contrast, their exchange diplomacy had a negative impact. Furthermore, social media communication strategies

utilizing personalized communication and even the expression of negative emotions were positively associated with increased visibility (retweets and quotes) and approval (likes) (Khosro, 2025). This shows that identity projection in conflict requires specific, emotionally resonant digital strategies tailored to the platform's affordances, rather than relying on traditional information dissemination.

The need for constant, personalized, and attention-grabbing content—whether through negative emotions or memetic engagement—risks the digital dissipation of soft power through overuse. Soft power is rooted in attraction (Diva Portal, 2011). If digital diplomacy prioritizes entertainment or visibility metrics over substantive content, it risks cheapening the symbolic performance of state authority and reducing the credibility of the actor in serious political contexts. This overuse or misuse of soft power can lead to diminished overall attractiveness, mirroring the way that overuse of hard power can reduce soft power (Diva Portal, 2011).

Ethical and Security Dilemmas in Digital Diplomacy

Digital engagement in international affairs introduces profound ethical challenges (Ethical Dilemmas in Digital Diplomacy, 2025). The increased visibility of diplomatic actions inherently challenges established norms of confidentiality (Ethical Dilemmas in Digital Diplomacy, 2025). Furthermore, digital tools facilitate the widespread dissemination of misinformation and enable covert influence operations, undermining truth-seeking in diplomacy (Ethical Dilemmas in Digital Diplomacy, 2025).

High-profile incidents, such as the 2010 WikiLeaks cable dump, the 2016 Russian DNC hack, and the 2021 Pegasus spyware scandal, demonstrate the complex tensions between transparency and security (Ethical Dilemmas in Digital Diplomacy, 2025). Hacking campaigns and cyber espionage, often perpetrated by state actors, compromise diplomatic cybersecurity, leading to leaks that endanger international stability (Ethical Dilemmas in Digital Diplomacy, 2025). To mitigate these risks, the ethical imperative for diplomats and international bodies involves cultivating cultures of

digital ethics and literacy, establishing standardized cybersecurity protocols, and developing clear social media policies that carefully balance security, transparency, and free speech (Ethical Dilemmas in Digital Diplomacy, 2025).

Transnationalism and Identity Mobilization: Non-State Actors as Constitutive Agents: The New Geopolitics of Non-State Actors (NSAs)

Since the end of the Cold War, non-state actors (NSAs) have assumed an increasingly consequential role in global dynamics, a trend facilitated by the growing capability and resourcing available outside of traditional state structures (Office of the Director of National Intelligence, 2024). NSAs encompass a broad spectrum of entities, including civil society movements, militant groups, commercial firms, and "super-empowered individuals" (Office of the Director of National Intelligence, 2024). They challenge state institutions by providing societal services, security, governance, and importantly, by shaping norms, values, and standards that give them prominence apart from states (Office of the Director of National Intelligence, 2024).

In the cyber domain, non-state actors launch the majority of cyber-attacks (Clingendael Institute, 2020). This complexity is exacerbated by the trend of state actors hiring or utilizing non-state actors as proxies to launch severe cyber-attacks, allowing states to operate covertly and distance themselves from damaging actions (Clingendael Institute, 2020). Dealing with non-state cyber-attackers becomes inherently difficult when states hide behind these entities (Clingendael Institute, 2020).

Digital Diasporas and Homeland Conflict Identity

The digital age has fundamentally transformed the political role of diaspora communities, moving them beyond traditional roles of lobbying or providing remittances. Digital diasporas now leverage social media, encrypted networks, and crowdfunding platforms to exert direct influence on conflict dynamics and international peacebuilding efforts (New Lines Institute, 2025). Examples include

the Ukrainian “IT Army” conducting coordinated cyberattacks, or Palestinian networks reshaping global discourse and mobilizing protests internationally (New Lines Institute, 2025).

Diasporas are defined by a sense of belonging, whether real or symbolic, to their country or origin (International Organization for Migration, 2025). Governments are increasingly seeking to engage these transnational communities to mobilize financial, human, and social resources (International Organization for Migration, 2025). However, diaspora communities are often fragmented and politicized, and their digital mobilization carries significant risks, including the rapid spread of misinformation, online harassment, and transnational repression (New Lines Institute, 2025).

This empowerment of non-state actors through digital platforms leads to the dissolution of the state's monopoly on transnational political identity. Historically, the sovereign state held the exclusive authority to define national identity and control its political projection abroad. Digital platforms bypass this control, enabling digital diasporas—united by a shared, symbolic identity (International Organization for Migration, 2025)—to generate independent narratives, engage in narrative shaping (a form of public diplomacy), and even execute coercive measures (cyberattacks) (New Lines Institute, 2025). This grants non-state actors a significant measure of narrative sovereignty over the collective identity related to a territory or conflict, posing a fundamental challenge to Westphalian notions of political control.

Social Media as a Catalyst for Collective Identity Formation and Protest

Social movement studies acknowledge the pivotal role of communication and mediation in the construction of collective identities, the sustainability of protest movements, and the development and spread of contentious politics (Cammaerts, 2025). The Arab Spring uprisings demonstrated that new media platforms played a crucial role in the planning and mobilization efforts that

successfully challenged existing political systems (Chapman University, 2025).

The outcomes of these revolutions varied, sometimes leading to the rise of extremist groups or the transfer of power to new authoritarian governments (Chapman University, 2025). However, the central takeaway is that social media served as a powerful catalyst for mobilization (Wikipedia, 2025). Online posting and sharing increased dramatically before major protests, creating a rapid feedback loop between citizen footage, social media circulation, and broadcast coverage, which maintained momentum and connected local actions to a wider regional movement (Wikipedia, 2025). This shift in media ecology expands the ability of marginalized individuals and groups—such as Black and Latino(a) youth—to voice concerns, organize protest activities, and effect change in an effort to minimize discrimination (Chapman University, 2025).

The increasing use of hybrid tactics in international conflict complicates traditional attributions of power. Soft power relies on attraction and clear attribution of the actor (Diva Portal, 2011). When states engage in proxy cyber operations or utilize non-state cyber mercenaries (Clingendael Institute, 2020), they intentionally obfuscate attribution. This hybrid approach allows states to wield compulsory power (coercion through cyberattacks) while insulating their national identity from the reputational damage associated with overt hard power or aggressive state actions. This tactic renders accountability complex under existing international law and challenges the foundational norms of responsibility in the digital age.

The Dark Side of Digital Identity: Polarization, Manipulation, and Hybrid Warfare Weaponizing Identity and Hybrid Threats

The structural affordances of social media platforms make democratic societies vulnerable to foreign influence operations (Federal Ministry of the Interior, 2025). Organized actors with political motives actively manipulate information or present it out of context to influence public debate (Federal Ministry of the Interior, 2025). This tactic is strategically defined as "weaponizing identities

and values" to increase societal polarization from the outside (Hybrid CoE, 2025). Hybrid threat actors manipulate internal vulnerabilities based on ethnic, religious, gender, and socioeconomic identities across democratic societies (Hybrid CoE, 2025).

The success of these disinformation campaigns in fracturing societal consensus relies on the platform dynamics: specifically, the volume, diversity of sources, speed, and repetition of the false information disseminated (Public Diplomacy in Post-Truth Era, 2025). These campaigns are designed to undermine public trust in government and aggravate existing social conflicts by focusing on controversial issues (Federal Ministry of the Interior, 2025). State actors may also utilize cyberattacks, such as hacking social media accounts and stealing data from elected officials, to prepare for or execute disinformation activities (Federal Ministry of the Interior, 2025).

Algorithmic Acceleration of Affective Polarization

A crucial finding in digital politics relates to the structural influence of platform architecture on political identity. Algorithms significantly drive political feelings, leading to affective polarization—partisan animosity—at an unprecedented speed (Northeastern University, 2025). Experimental studies have demonstrated this acceleration: users whose feeds were exposed to increased levels of "antidemocratic attitudes and partisan animosity" (AAPA) content saw their feelings toward the opposing party "cool" by approximately two points on a political science attitude scale in just one week (Northeastern University, 2025). Conversely, users exposed to reduced AAPA content saw their feelings "warm" by a similar magnitude (Northeastern University, 2025).

This four-point swing in affective polarization occurred in only seven days, a change comparable to the shift typically observed over three years in the United States (Northeastern University, 2025). This dramatic acceleration elevates the commercial imperative of platform design—the drive for maximum engagement—into a primary geopolitical threat vector. The platform architecture is not politically neutral; it structurally creates a volatile intersubjective

reality that generates rapid ideological fragmentation, even without explicit foreign state intervention.

This structural volatility implies that the algorithmic imperative itself constitutes a geopolitical threat. The core economic model of social media, which prioritizes engagement through the amplification of emotionally charged or polarizing content, ensures a perpetually volatile intersubjective reality. This environment creates immediate fissures in the social fabric that can be easily exploited by hybrid threat actors seeking to induce societal instability (Social Drivers and Algorithmic Mechanisms on Digital Media, 2025). The platform architecture thus becomes a mechanism for continuous identity destabilization.

Identity and Conflict: The Case of Ukraine

Contemporary conflicts provide empirical evidence of the direct relationship between social media, identity, and warfare. In the context of the Russia-Ukraine war, social media data has proven effective in tracking and quantifying shifts in national identity (American Economic Association, 2024). Analysis of VK users' language proficiency and the language used in tweets showed a significant shift from Russian to Ukrainian during both the early phase (2014–2022) and the later phase (2022–) of the conflict (American Economic Association, 2024). This linguistic change observed on social media correlated positively with shifts in ethnic identity reported in nationally representative surveys (American Economic Association, 2024). This demonstrates that digital behavior provides a measurable indicator of rapid, conflict-induced political and ethnic identity change.

Furthermore, social media has transformed the visualization of conflict. Videos posted by Ukrainian soldiers circulating widely at the onset of the Russian invasion fostered empathy for Ukrainians and brought the war closer to the public in an unprecedented manner (Taylor & Francis, 2025). This diffusion of raw, personalized content challenges traditional models of state-controlled, militarized narratives (Taylor & Francis, 2025). Conflict visualization via social media fragments the idealized state identity (e.g., the stoic, hyper-

masculine national hero) and introduces nuanced emotional complexity that challenges established identity norms within warfare (Taylor & Francis, 2025).

The Limits of Disinformation's Strategic Effectiveness

While the power of social media to fracture domestic identity is clear, the strategic effectiveness of disinformation in changing a target state's grand foreign policy decisions—such as alignment or armaments (i.e., the balance of power)—may be exaggerated (Zeitsoff, 2025). Disinformation campaigns, such as those targeting the Baltic states, face three significant obstacles in IR: the fundamental uncertainty generated by international anarchy over information broadcast by adversaries; the pre-existing prejudices of foreign policy elites and ordinary citizens; and the availability of countermeasures (Zeitsoff, 2025).

The primary success of identity weaponization, therefore, lies not in altering foreign policy but in eroding the domestic prerequisites for political stability. Even if Russia's disinformation campaign, for example, failed to shift a Baltic state's alliance structure, the resulting erosion of public trust and societal polarization constitutes a significant strategic success for the hostile actor (Federal Ministry of the Interior, 2025). The erosion of domestic social cohesion is a geopolitical end in itself, as it diminishes a country's collective resilience.

Table 2: Identity Dynamics in the Digital International System

Actor Type	Identity Focus/Goal	Digital Mechanism Used	IR Impact (Positive/Negative)
State/MFA	National Status & Legitimacy	Visual Diplomacy, Personalized Communication, Cultural Diplomacy (Khoso, 2025)	Enhanced soft power, ritualistic reinforcement of authority, risk of inauthenticity (Manor, 2025)
Non-State/Diaspora	Collective Political Identity	Crowdfunding, Encrypted Networks, Narrative Shaping (New Lines Institute, 2025)	Increased transnational political influence, challenging state sovereignty (Office of the Director of National Intelligence, 2024), conflict mobilization
Adversary/Hybrid Actor	Societal Destabilization	Algorithmic Amplification, Targeted Disinformation, Hacking (Northeastern University, 2025)	Rapid affective polarization, erosion of public trust, fracturing of reality (Public Diplomacy in Post-Truth Era, 2025)
Authoritarian Regimes	Domestic Cohesion & Control	Censorship, Surveillance, Data Exploitation (European Consortium for Political Research, 2025)	Repression of political identity (Mercer, 1995), undermined privacy, consolidation of digital authoritarianism (Freedom House, 2018)

Governance, Ethics, and the Future of the Digital Public Sphere: Digital Authoritarianism and Identity Control

The internet, once heralded as a tool for destabilizing dictatorships, is increasingly utilized by authoritarian regimes to consolidate control (Freedom House, 2018). Digital authoritarianism is defined as the 'use of digital information technology by authoritarian regimes to surveil, repress, and manipulate domestic and foreign populations' (European Consortium for Political Research, 2025). These regimes frequently embrace models of extensive censorship and automated surveillance (Freedom House, 2018).

The unbridled collection and potential exploitation of personal data, exemplified by incidents like the Cambridge Analytica scandal, demonstrate how personal information can be weaponized to influence electoral outcomes globally (Freedom House, 2018). Furthermore, state-sponsored online activities, including the spread of false rumors and hateful propaganda, have incited jarring real-world violence against ethnic and religious minorities in countries like Myanmar, India, and Sri Lanka (Freedom House, 2018). This deliberate sowing of hostility between communities serves the interests of antidemocratic forces and hostile foreign states (Freedom House, 2018).

The Ethical Imperative of Platform Design

The commercial structures underpinning social media, driven by the "attention economy," are inherently designed to maximize user engagement, often through addictive features that raise serious ethical questions and contribute to negative mental health outcomes (Handbook of Adolescent Digital Media Use and Mental Health, 2025). This dynamic has brought increased academic focus on the digital use divide, which pertains not just to unequal access to resources, but to unequal exposure to manipulative design features (Castro & O'Connor, 2025).

The technological features that create this addictive environment could be intentionally redesigned to promote healthier user relationships with social media (Castro & O'Connor, 2025). Furthermore, research suggests that algorithms could be employed for "positive societal impact" by emphasizing the overlap in views between political groups, thereby reducing polarization, or by highlighting constructive and nuanced content (Northeastern University, 2025). Media developers, therefore, carry an ethical responsibility to design social media environments that actively support healthy political identity development, rather than facilitating identity distress and fragmentation (Handbook of Adolescent Digital Media Use and Mental Health, 2025).

Sovereignty, Content Moderation, and Trust

The governance of digital platforms significantly impacts political identity. Content moderation policies determine what narratives and self-presentations are permissible online (Congressional Research Service, 2025). When platforms enforce these rules globally, they effectively assume a quasi-sovereign regulatory function, defining the acceptable boundaries of political identity expression across international borders.

The political context in which social media operates is paramount. Research on political identity in authoritarian contexts reveals that perceptions of corruption significantly moderate the relationship between social media activity (such as information sharing and self-presentation) and the positive formation of political identity (The Influence of Social Media Content, 2025). High corruption perception weakens these positive associations (The Influence of Social Media Content, 2025).

This reveals a critical governance gap in identity resilience. Adversarial actors actively seek to undermine public trust in government through hacking and disinformation (Federal Ministry of the Interior, 2025). If political corruption perceptions are high, the positive potential of social media—its ability to foster political discussion and social cohesion—is neutralized (The Influence of Social Media Content, 2025). Consequently, improving political transparency and tolerance is crucial to unlocking the constructive potential of social media and fostering a broader, more cohesive political identity (The Influence of Social Media Content, 2025).

Countering the weaponization of identity necessitates a “whole-of-society approach” that identifies vulnerable spots and delivers countermeasures through entities the public perceives as trustworthy (Hybrid CoE, 2025). However, when external attacks actively target public trust and transparency is low, this creates a vicious cycle where identity resilience measures cannot be effectively implemented, leading to political paralysis and continued polarization.

The Need for Digital Ethics in International Relations

The integration of digital technologies into statecraft and societal life necessitates an adaptation of ethical frameworks in international relations. The tension between promoting transparency while combating misinformation, and fostering cooperation amidst cyber espionage, presents profound ethical dilemmas (Ethical Dilemmas in Digital Diplomacy, 2025). To maintain the moral foundations of diplomacy in the digital age, policy must mandate enhanced cybersecurity protocols, organizational cultures that value digital ethics and literacy, and clear guidelines for platform engagement (Ethical Dilemmas in Digital Diplomacy, 2025). Ethical adaptation, rooted in transparency and accountability, is critical for preserving the enduring values of dialogue and cooperation in an era of rapid technological change (Ethical Dilemmas in Digital Diplomacy, 2025).

Conclusion

Synthesis of Findings: Social Media as the New Arena for Identity Struggle

Social media has irrevocably altered the landscape of International Relations by establishing itself as a constitutive structural force, rather than a mere conduit of information. Utilizing a constructivist framework, this analysis confirms that political identity—a multiple, fluid, and intersubjectively constructed variable—is highly susceptible to digital affordances.

The core finding is the duality of social media's impact: it empowers non-state actors, such as digital diasporas, granting them a degree of narrative sovereignty that directly challenges the state's monopoly on political identity projection (New Lines Institute, 2025). Simultaneously, the platform architecture, driven by algorithmic engagement metrics, accelerates political fragmentation and affective polarization at a pace previously unseen in IR, generating a volatile environment ripe for exploitation by hybrid threat actors (Northeastern University, 2025).

For states, identity performance has become inextricably linked to achieving social visibility and approval (Khoso, 2025), potentially shifting the metrics of diplomatic success from policy outcomes to digital status affirmation (Duncombe, 2025). However, this necessity to perform publicly exposes statecraft to severe ethical and security risks, including cyber espionage and the dissipation of soft power through overly aggressive or inauthentic digital engagement (Diva Portal, 2011).

Theoretical Contributions and Policy Relevance

The primary theoretical contribution of this study is the documentation and contextualization of the acceleration of constitutive change in the international system. The empirical evidence demonstrating that algorithms can achieve years-worth of identity shifts in a week compels IR scholars to abandon traditional, slower temporal models of normative evolution (Northeastern University, 2025). Future theoretical work must incorporate the concept of algorithmic acceleration when modeling political volatility and conflict escalation.

For policymakers, the focus must shift beyond the exclusive combatting of foreign disinformation to structurally mitigating algorithmic polarization, as the platform design itself is a significant driver of internal fragmentation (Northeastern University, 2025). Furthermore, national security policy must recognize that domestic political factors, particularly transparency and trust, are crucial preconditions for identity resilience. The presence of high corruption perception fundamentally neutralizes the positive potential of social media engagement (The Influence of Social Media Content, 2025), requiring states to prioritize institutional reform alongside digital countermeasure deployment.

Future Research Agenda

The comprehensive study of social media and identity in International Relations suggests several urgent avenues for future scholarly inquiry:

Modeling Volatility and Cooperation: Developing predictive IR models that assess the long-term impact of rapid, algorithmically induced identity change (polarization) on the prospects for international cooperation, crisis escalation, and alliance cohesion.

Comparative Platform Governance: Conducting comparative studies across different regime types (e.g., strong democracy, weak democracy, full authoritarianism) to determine how the interaction between state regulation, platform governance, and citizen political identity formation is modified by local factors such as state corruption and media trust.

Ethical Design and Identity Support: Further research is needed to collaborate with technology developers and psychologists to design digital media environments whose affordances actively support, rather than hinder, healthy political identity formation and dialogue (Handbook of Adolescent Digital Media Use and Mental Health, 2025). This includes developing metrics for success that prioritize collective well-being and nuanced engagement over mere visibility and engagement metrics.

Digital Power and International Law: Analyzing the legal and normative implications of platform content moderation policies, treating them as quasi-sovereign regulatory acts, and developing international norms to ensure that these corporate governance decisions align with human rights and democratic values across borders.

Chapter Three: The Digital Self On The World Stage: Social Media, Identity Construction, And The Transformation Of International Relations

Introduction

The study of International Relations (IR) has long been dominated by Realist conceptions of power, where material capabilities—military might, economic GDP, and territory—dictate the hierarchy of nations. However, the advent of the digital age has introduced a new variable that defies strictly material measurement: the **digital identity**. Today, a tweet from a head of state can trigger a diplomatic crisis, a viral video can undermine a military operation, and a hashtag can mobilize a global diaspora against a regime.

As of the mid-2020s, the digital realm is no longer virtual; it is the primary arena where political narratives are contested.² The central problem this paper addresses is the **ontological shift** in how international actors exist. States no longer just *are*; they must constantly *perform* their identity on social media platforms owned by private corporations.³

This paper seeks to answer the following research question: *How do social media platforms function as constitutive structures of international identity, and in what ways does this "digital identity" alter the practice of diplomacy and conflict?*

Theoretical Framework: Constructivism and the Digital Turn

To understand digital identity in IR, we must look beyond Realism and Liberalism. **Constructivism**, particularly the work of Alexander Wendt, provides the most robust lens. Wendt famously argued that "anarchy is what states make of it" (Wendt, 1992). In the

digital age, we might update this to: **"Identity is what the algorithm makes of it."**

Constructivism in the Digital Age

Constructivism posits that the international system is social, not material.⁴ Identities and interests are constructed through interaction.⁵ Social media accelerates this interaction to real-time.⁶ Where traditional diplomatic signals (e.g., recalling an ambassador) took days or weeks to process, digital signals (e.g., a fiery tweet or a changed profile picture) are instantaneous.⁷

Intersubjectivity: On social media, state identity is co-constituted by the "audience" (global citizens). A state cannot simply declare it is "peace-loving" if the global comment section successfully frames it as an aggressor.

The Performative State: Scholars like Adler-Nissen (2015) argue that diplomacy is a performance. Social media turns IR into a constant, high-stakes theater where "likes" and "shares" serve as proxy metrics for legitimacy.

Soft Power vs. Sharp Power

Joseph Nye's concept of **Soft Power**—the ability to attract rather than coerce—is central to digital identity (Nye, 2004).⁸ However, the digital ecosystem has morphed Soft Power into what some scholars call **Sharp Power** (Walker & Ludwig, 2017).⁹

Soft Power: A nation branding campaign on Instagram promoting tourism.

Sharp Power: Using bot farms to amplify social discord in a rival nation, weaponizing identity politics to weaken the adversary from within.

The Architecture of Digital Diplomacy

Digital diplomacy (or e-diplomacy) has evolved through three distinct phases, each redefining international identity.

Phase 1: Information Dissemination (Web 1.0)

Early digital diplomacy was simply a digitization of press releases. Embassies created websites to broadcast static identities. The flow of information was one-way (State  Public).

Phase 2: Engagement and Twiplomacy (Web 2.0)

The rise of Twitter (now X) and Facebook ushered in "Twiplomacy." Leaders began engaging directly with citizens and each other.¹⁰

Case Study: The Iran Nuclear Deal negotiations saw unprecedented use of Twitter by diplomats to signal red lines and sell the deal to domestic audiences in real-time.

Identity Impact: Diplomats became "personalities." The stiff, formal identity of the state was replaced by the accessible, witty, or aggressive persona of the individual diplomat.

Phase 3: Algorithmic Diplomacy and AI (Web 3.0)

We are now in an era where algorithms determine visibility. State identities are subject to the "black box" of platform moderation.

The Sovereignty Paradox: States claim sovereignty, yet their digital voice can be silenced by a CEO in Silicon Valley (e.g., the de-platforming of political leaders or the labeling of state media accounts).

Identity Construction in Conflict: The Narrative Wars

Digital identity is most visible—and volatile—during conflict. The battleground is no longer just territory, but the **narrative of the self** and the **demonization of the other**.

The Ukraine-Russia War: The First "TikTok War"

The 2022 invasion of Ukraine demonstrated the power of digital identity.

Ukraine's Digital Self: Under President Zelenskyy, Ukraine constructed an identity of the "heroic underdog." High-quality video production, meme warfare (e.g., the "Saint Javelin" image), and

direct appeals to Western pop culture created a global digital solidarity.

Russia's Identity Crisis: Conversely, Russia struggled to project a coherent identity. Its traditional disinformation machine (RT, Sputnik) was largely banned in the West, forcing it to rely on "grey zone" actors and bot networks. The isolation of Russia's digital identity mirrored its geopolitical isolation.

China's "Wolf Warrior" Diplomacy

China's shift to "Wolf Warrior" diplomacy represents a deliberate reconstruction of state identity. Moving away from Deng Xiaoping's maxim "hide your strength, bide your time," Chinese diplomats on Twitter began aggressively countering Western narratives.

Identity Goal: To project strength and reject "victimhood."

Result: While this solidified domestic nationalism (internal identity), it often alienated international audiences (external identity), highlighting the tension between domestic legitimacy and international soft power.

Non-State Actors and the "Digital Nation"

Social media has broken the state's monopoly on international identity.¹¹ Non-state actors can now construct identities that rival nations in influence.

Digital Diasporas

Diasporas use social media to maintain a "long-distance nationalism." They lobby host governments and influence the foreign policy of their home states.

Example: The Armenian diaspora's digital mobilization during the Nagorno-Karabakh conflicts effectively kept the issue on the international agenda, constructing a trans-national identity that transcended borders.

Terrorist Identity Construction

Groups like ISIS were pioneers in digital identity. They produced Hollywood-style propaganda to project an identity of "invincibility" and "statehood" (the Caliphate) to recruit globally. Their digital identity was key to their operational capacity, proving that in the digital age, **brand is power**.

Challenges and Ethical Implications

Algorithmic Bias and Digital Orientalism

The algorithms that curate our feeds are not neutral.¹² They often reinforce existing power hierarchies. Scholars argue this leads to "Digital Orientalism," where the identities of the Global South are filtered through Western-centric algorithms, often reducing complex political struggles to simple, consumable stereotypes for Western audiences.

The Erasure of Truth

When identity is purely performative, truth becomes secondary. "Deepfakes" and AI-generated content threaten the very foundation of diplomatic trust. If a state cannot verify the identity of a counterpart in a video call, the stability of the international system is at risk.

Conclusion

The intersection of digital identity, social media, and International Relations represents a fundamental shift in the nature of global politics. Identity in IR is no longer a static attribute given by geography or history; it is a **fluid, contested, and real-time performance** mediated by private technology platforms.

For diplomats and policymakers, this means that **digital literacy is now a core diplomatic skill**.¹³ Understanding the "code" of social media is as important as understanding the protocol of a state dinner. As we move forward, the states that successfully manage their "digital self"—balancing openness with security, and

narrative appeal with factual credibility—will define the future of the international order.

Case Study

The First "TikTok War" – Digital Identity and Narrative Control in the Russia-Ukraine Conflict

Introduction

The 2022 full-scale invasion of Ukraine by the Russian Federation marks a watershed moment in International Relations (IR), often described as the first "broadband war" or "TikTok War" (Cherevko & Yena, 2023). While information warfare is not new, the scale, speed, and algorithmic nature of the Russia-Ukraine conflict distinguish it from predecessors like the Arab Spring. This case study analyzes how two distinct models of **digital identity construction** collided: Ukraine's "connective strategic narrative" versus Russia's top-down "firehose of falsehood."

Ukraine's Digital Strategy: The Connective Strategic Narrative

Ukraine's resistance was not merely military but ontological. Under President Volodymyr Zelenskyy, the state successfully leveraged what scholars call a "connective strategic narrative"—a bottom-up form of storytelling that invites global participation (Parmelee, 2023).

Zelenskyy and "Selfie Diplomacy"

President Zelenskyy's transformation from a suit-wearing politician to a fatigue-clad wartime leader was a deliberate performance of identity. His handheld video from Kyiv on February 25, 2022, declaring "We are here," served as a pivotal "speech act" in the Constructivist sense (Wendt, 1999).

Immediacy and Authenticity: By using a smartphone camera in low light, Zelenskyy bypassed traditional media filters, creating a sense of intimacy and "raw" truth that resonated with Western digital audiences (Battista, 2025).

The "Affective Public": Zelenskyy's rhetoric transformed passive international observers into an "affective public"—a digitally connected group united by shared sentiment rather than shared geography (Papacharissi, 2015). This emotional mobilization directly influenced Western policy shifts, such as Germany's reversal on arms exports.

Horizontal Mobilization: The IT Army

Unlike traditional state propaganda, Ukraine's digital identity was decentralized. The Ministry of Digital Transformation recruited a global "IT Army" of over 300,000 volunteers. This blurred the Westphalian distinction between **combatant** and **civilian**, as teenage coders in Berlin or Toronto could launch DDoS attacks on Russian banks, effectively "donating" their digital identity to the Ukrainian state (Soesanto, 2022).

Russia's Digital Strategy: The Firehose of Falsehood

In contrast to Ukraine's agile, horizontal strategy, Russia employed a rigid, vertical model of information control, often characterized as the "firehose of falsehood" model (Paul & Matthews, 2016).

Fragmented Identities

Russia faced a dual-audience problem that fractured its digital identity:

Domestic Audience: The narrative was one of a "Special Military Operation" to "denazify" Ukraine, appealing to Soviet nostalgia and imperial identity.

International Audience: Since this narrative was rejected by the West, Russia shifted from *persuasion* to *confusion*. The goal was not to convince the world that Russia was "good," but to flood the

zone with so much noise (e.g., conflicting theories about the Bucha massacre) that the truth became unknowable (Smart et al., 2022).

Cyber-Enabled Disinformation: "Ghostwriter"

Technical reports link Russian intelligence (GRU) to the "Ghostwriter" campaign, which hacked the social media accounts of Ukrainian military officials to post "deepfake" surrender videos (Meta, 2022). This represents the **weaponization of digital identity**: stealing a trusted persona to inject a lethal narrative virus.

Sub-Case: Meme Warfare and "Saint Javelin"

One of the most unique aspects of this conflict was the use of internet memes as instruments of Soft Power.

Saint Javelin: An image of the Virgin Mary holding a FGM-148 Javelin anti-tank missile became a global icon of resistance. Originally a digital art piece, it was monetized to raise over \$1 million for the Ukrainian military.

Theoretical Implication: This phenomenon challenges traditional IR theory. A meme, created by a non-state actor (a Canadian journalist), generated material aid (money/equipment) for a sovereign state. It demonstrates that in the digital age, **cultural capital can be instantly converted into military capital** (Borys, 2022).

Conclusion: The Asymmetry of Authenticity

This case study reveals that in digital IR, authenticity is a strategic asset. Ukraine's digital identity felt "native" to social media—chaotic, emotional, and human. Russia's identity felt "imported" from 20th-century broadcast propaganda—stiff, scripted, and easily debunked by open-source intelligence (OSINT) communities like Bellingcat. The war suggests that future diplomatic success will depend on a state's ability to navigate the algorithmic marketplace of attention, where credibility is currency.

Adler-Nissen, R. (2015). *Diplomatic Agency: The Plurality of International Relations*. Oxford University Press.

Bjola, C., & Holmes, M. (2015). *Digital Diplomacy: Theory and Practice*. Routledge.

Castells, M. (2009). *Communication Power*. Oxford University Press.

Duncombe, C. (2017). Twitter and transformative diplomacy: Social media and Iran–US relations. *International Affairs*, 93(2), 303-314. <https://doi.org/10.1093/ia/iw024>

Hocking, B., & Melissen, J. (2015). *Diplomacy in the Digital Age*. Clingendael Institute.

Manor, I. (2019). *The Digitalization of Public Diplomacy*. Palgrave Macmillan.

Nye, J. S. (2004). *Soft Power: The Means to Success in World Politics*. PublicAffairs.

Sandre, A. (2015). *Twiplomacy: Social Media and the New Diplomacy*. Lynne Rienner Publishers.

Walker, C., & Ludwig, J. (2017). The Meaning of Sharp Power. *Foreign Affairs*.
<https://www.foreignaffairs.com/articles/china/2017-11-16/meaning-sharp-power>

Wendt, A. (1992). Anarchy is what states make of it: The social construction of power politics. *International Organization*, 46(2), 391-425.

Zeitsoff, T. (2017). How Social Media Is Changing Conflict. *Journal of Conflict Resolution*, 61(9), 1970–1991.

Battista, A. (2025). The evolution of digital communication: Zelensky and the use of Instagram in wartime. *Athens Journal of Mass Media and Communications*, 11(1), 1–18.

Borys, C. (2022, May). How a meme became a movement: The story of Saint Javelin. [Lecture transcript]. University of Toronto Munk School of Global Affairs.

Cherevko, T., & Yena, O. (2023). Social media as a battlefield: The phenomenon of the "TikTok War" in Ukraine. *Journal of Strategic Security*, 16(2), 45–62.

Ciuriak, D. (2022). The economic consequences of the Russia-Ukraine war. SSRN Electronic Journal. <https://dx.doi.org/10.2139/ssrn.4063273>

Meta. (2022). Quarterly adversarial threat report. Meta Security. <https://about.fb.com/news/2022/04/state-media-trends-ukraine/>

Papacharissi, Z. (2015). Affective publics: Sentiment, technology, and politics. Oxford University Press.

Parmelee, J. H. (2023). The connective strategic narrative: Social media and the Ukrainian defense. Political Communication, 40(4), 501–522.

Paul, C., & Matthews, M. (2016). The Russian "Firehose of Falsehood" propaganda model. RAND Corporation.

Smart, C., O'Connor, C., & Smith, G. (2022). Weaponized confusion: Russian disinformation tactics in the Ukraine conflict. Institute for Strategic Dialogue.

Soesanto, S. (2022). The IT Army of Ukraine: Structure, tasking, and ecosystem. Center for Security Studies (CSS), ETH Zürich.

Wendt, A. (1999). Social theory of international politics. Cambridge University Press.

Kaynakça

Adler-Nissen, R., & Guzzini, S. (2025). Reclaiming constructivism: Identity and the practice of the study of International Relations. *PS: Political Science & Politics*. <https://www.cambridge.org/core/journals/ps-political-science-and-politics/article/reclaiming-constructivism-identity-and-the-practice-of-the-study-of-international-relations/B4E910023624C5A2EB3AD984747A4B0A> adresinden erişildi.

Bjola, C. (2025). *Diplomacy 2.0: The future of social media in nation branding*. Syracuse University. <https://surface.syr.edu/cgi/viewcontent.cgi?article=1032&context=exchange> adresinden erişildi.

Cammaerts, B. (2025). *Social media and activism*. LSE Research Online. https://eprints.lse.ac.uk/62090/1/Social_media_and.pdf adresinden erişildi.

Castro, A., & O'Connor, K. (2025). Ethics of the attention economy: The problem of social media addiction. *Business Ethics Quarterly*. <https://www.cambridge.org/core/journals/business-ethics-quarterly/article/ethics-of-the-attention-economy-the-problem-of-social-media-addiction/1CC67609A12E9A912BB8A291FDFFE799> adresinden erişildi.

Clapton, W. (2025). *Popular culture, social media, and the politics of identity* (1. baskı). London: Routledge. <https://www.routledge.com/Popular-Culture-Social-Media-and-the-Politics-of-Identity/Clapton/p/book/9781032486413> adresinden erişildi.

Congressional Research Service. (2025). *Social media: Content dissemination and moderation practices*. <https://www.congress.gov/crs-product/R46662> adresinden erişildi.

Duncombe, C. (2025). Social aesthetics of digital diplomacy. *International Political Sociology*, 18(3). <https://academic.oup.com/ips/article/18/3/olae027/7696036> adresinden erişildi.

European Consortium for Political Research. (2025). *Digital authoritarianism*. <https://ecpr.eu/group/digital-authoritarianism> adresinden erişildi.

Federal Ministry of the Interior (BMI). (2025). *Homepage - Disinformation as hybrid threat*. <https://www.bmi.bund.de/SharedDocs/schwerpunkte/EN/disinformation/article-disinformation-hybrid-threat.html> adresinden erişildi.

Freedom House. (2018). *The rise of digital authoritarianism*. <https://freedomhouse.org/report/freedom-net/2018/rise-digital-authoritarianism> adresinden erişildi.

Hybrid CoE. (2025). *Weaponizing social identities: What can we learn from examples of targeted disinformation?* <https://www.hybridcoe.fi/news/weaponizing-social-identities-what->

can-we-learn-from-examples-of-targeted-disinformation/
adresinden erişildi.

International Organization for Migration (IOM). (2025). *Harnessing the contributions of transnational communities and diasporas*.

https://www.iom.int/sites/default/files/our_work/ODG/GCM/IOM-Thematic-Paper-Harnessing-the-contributions-of-transnational-com.pdf adresinden erişildi.

Khoso, A. (2025). Social media diplomacy for social visibility and social approval: Strategic communication of the Taliban de facto government on Twitter. *Journal of Public Diplomacy* (veya ilgili dergi).
[https://www.tandfonline.com/doi/full/10.1080/19331681.2025.2548462?src=](https://www.tandfonline.com/doi/full/10.1080/19331681.2025.2548462?src=adresinden) adresinden erişildi.

Manor, I. (2025). *The "dark side" of digital diplomacy*. USC Center on Public Diplomacy.
<https://uscpublicdiplomacy.org/blog/dark-side-digital-diplomacy>
adresinden erişildi.

Mercer, J. (1995). *National identities and international relations*. Cambridge University Press.
<https://www.cambridge.org/core/books/national-identities-and-international-relations/94021153B07FD201FAF19AB6586F4F4C>
adresinden erişildi.

New Lines Institute. (2025). *How social media has transformed diaspora mobilization*.
<https://newlinesinstitute.org/strategic-technology/digital-diasporas-how-social-media-and-the-internet-are-transforming-diaspora-mobilization-in-homeland-conflict/> adresinden erişildi.

Northeastern University. (2025, 27 Kasım). *How does social media impact political polarization?*
<https://news.northeastern.edu/2025/11/27/social-media-political-polarization-research/> adresinden erişildi.

Office of the Director of National Intelligence (DNI). (2024). *Non-state actors playing greater roles in governance and international affairs*.

https://www.dni.gov/files/ODNI/documents/assessments/NICM-Non-State-Actors_23-01637_05-18-24_.pdf adresinden erişildi.

Oxford Research Encyclopedia. (2025). *Norms and social constructivism in International Relations*. <https://oxfordre.com/internationalstudies> adresinden erişildi.

Rhoda, B. (2025). *The effect of social media on identity formation — How social media influences personal identity, self-perception, and group belonging*. Medium. <https://medium.com/@biodunrhoda/the-effect-of-social-media-on-identity-formation-how-social-media-influences-personal-identity-c550cee715b7> adresinden erişildi.

University of Southern California. (2025). *Selfie diplomacy: An analysis of MFA profile pictures on Twitter*. USC Center on Public Diplomacy. <https://uscpublicdiplomacy.org/blog/selfie-diplomacy-analysis-mfa-profile-pictures-twitter> adresinden erişildi.

Zeitzoff, T. (2025). Disinformation in international politics. *European Journal of International Security*. <https://www.cambridge.org/core/journals/european-journal-of-international-security/article/disinformation-in-international-politics/49121D3134C4079BF51E5012E6BE247A> adresinden erişildi.

NEW FACE OF IR: **SOCIAL MEDIA** AND DIGITALISATION

