ACADEMIA International Journal for Social Sciences
Volume 5, Issue 3(s7), 2026 ISSN-L (Online): 3006-6638

Artificial Intelligence as a Weapon of Digital Warfare: Competing US and Iranian Narratives in Pakistan

Zafar Ali
zafaralizfr@gmail.com
PhD Scholar, School of Communication Studies, University of the Punjab, Lahore, Pakistan.

Dr. Munwar Ali Kalwar
Deputy Director, International Institute of Islamic Economics (IIIE), International Islamic University, H-10, Islamabad, Pakistan.

Dr. Shehla Jabeen
Lecturer, Department of Mass Communication, Lahore College for Women University, Pakistan.

Corresponding Author: Zafar Ali zafaralizfr@gmail.com

Received: 17-01-2026 **Revised:** 01-02-2026 **Accepted:** 14-02-2026 **Published:** 03-03-2026

ABSTRACT

Artificial intelligence (AI) has quickly become more than a technical facilitator to a strategic weapon of narrative competition, but empirical studies on how audiences in the Global South make sense of competing state discourses of AI-enabled digital warfare have yet to be conducted. The paper explores the exposure and perceptions of Pakistani university students to the United States and Iranian stories that AI becomes a digital warfare weapon in 2026. Using a stratified random sample, we sampled 400 undergraduate and postgraduate students in four major Lahore universities (University of the Punjab, Superior University, Government College University, and University of Central Punjab) with the help of strategic narrative theory, framing theory, and securitization theory. The information was obtained through a bilingual questionnaire in Google Forms that assessed the knowledge of AI, exposure to narratives, perceived credibility, perceived threat, and behavioral intent. SPSS v29 was used to perform descriptive statistics, paired t-tests, one-way ANOVA, and chi-square. Findings indicate that there is a heavy bias in favour of Iranian stories. Students reported more frequent exposure to Iranian-attributed AI content ($M = 3.31$, $SD = 1.27$) than US ($M = 2.40$, $SD = 1.20$) or Israeli content ($M = 2.09$, $SD = 1.16$) and rated Iranian narratives as significantly more credible, $t(399) = 12.74$, $p < .001$, $d = 0.64$. Although the US AI capabilities were perceived as a higher threat to Pakistan ($M = 3.58$ vs. 2.95), the latter did not lead to narrative trust. The slant did not differ significantly between universities, $F(3, 396) = 0.98$, $p = .402$, yet differed significantly between platforms, with significantly more users on TikTok and Instagram reporting high levels of Iranian narrative control, $\chi^2 (9) = 21.67$, $p = .010$. Results indicate that narrative advantage in the context of the digital ecosystem in Pakistan is not based on technological excellence but rather on linguistic localization, platform formatting, and cultural connection.

Keywords: *Artificial intelligence, digital warfare, strategic narratives, disinformation, Pakistan, US-Iran war, Social media, TikTok*

INTRODUCTION

The 2026 US-Iran conflict was the first-ever global conflict where Artificial Intelligence was used on a large scale in both kinetic and cognitive platforms of warfare. As AI systems were applied to identify targets, fusion of intelligence, and battlefield logistics, another battle for hearts and minds was waged on social media, where AI-generated content was the focus of narrative competition. The so-called cyber front is becoming the leading aspect of contemporary warfare, and AI-controlled by state-funded entities

is used to control digital space, hack into news providers, and deliver propaganda straight to the gadgets of citizens. Iran, which cannot compete with US-Israeli conventional capabilities, has made data infrastructure and narrative control a key target, viewing commercial data centers as a functional extension of adversarial power and using AI to influence the world. In contrast, the US has both incorporated large language models and computer vision into decision support and strike operations and at the same time warned that adversaries such as Iran are using AI to dictate global narratives.

Artificial intelligence (AI) is no longer seen as a supporting tool, but rather as a key strategic asset in the contemporary conflict, and scholars and practitioners have come to refer to it as a driver of so-called digital warfare where algorithms, autonomous systems, and data supremacy can replace traditional airpower and nuclear deterrence (Karamchand & Aramide, 2025). By 2025, the military and intelligence agencies openly incorporated generative AI and machine-learning platforms in targeting, planning, and cyber operations, reducing the time of decision-making that was taken in hours to a few seconds (USA Today, 2026). This change has created what Atlantic Council researchers call algorithmic warfare, where drones and decision-support systems can generate real-time targeting proposals and logistical forecasts, shifting AI out of single-use weapons and into overall systems that can guide planning and the deployment of force at any level. The use of AI as a weapon has not only been limited to the kinetic level but has also extended to information. The AI is now used by states to influence global narratives, and former security officials have given warnings that Iran is using AI to dominate global narratives when it does not have an advantage on the battlefield (Fox News, 2026). This risk has been specifically identified by Pakistan at the United Nations, where it warned that AI would be able to escalate future wars, and it needs to regulate how it is used in the military to ensure it does not become a weapon of coercion. As reported in 2025, there was a shift in the AI-driven misinformation with deepfakes and deceitful content approaching unprecedented levels, and politics as the most extensively exploited field (CEJ at IBA, 2025). In this contentious arena, Pakistan has become a major arena of rival narratives. In 20252026, media reports reported Pakistan-sponsored disinformation that attempted to lure India into U.S.-Iran naval strike narratives through coordinated hashtag amplification and AI-altering videos. At the same time, the Pakistani government denied that it would assist in U.S. attacks on Iran, saying that such assertions were nothing but mere disinformation to involve Pakistan in a U.S.-Iran war. These episodes depict what has been described by scholars as a form of misinformation warfare as a strategic instrument in modern geopolitical struggle by using real-life events to fabricate fake narratives that weaken international credibility. These dynamics are further contextualized in scholarly work on the digital ecosystem in Pakistan. Uddin et al. (2025) discovered an active disinformation economy in Pakistan in which political incentives, platform organization, and low levels of digital literacy contribute to AI-enabled misinformation, and suggest context-specific media literacy interventions. The quantitative research on Facebook in Pakistan also found that algorithmic curation and filter bubbles contribute to the process of polarization by strengthening the already held beliefs and restricting opposing opinions. Since the IT industry in Pakistan has already managed to export 3.8 billion US dollars in FY25, and 86% of professionals in the country use AI tools (Habib University, 2026), the youth in this nation are both the most active users of AI technologies and objects of influence operations made by AI. The fact that the Higher Education Commission has decided to make AI courses mandatory by 2026 signifies that AI has become central nationally (Dawn, 2026). Although this is the case, minimal empirical research has investigated the perceptions and distinctions of the Pakistani university students, who are digital natives and at the threshold of high AI adoption and geopolitical narrative competition, on the US and Iranian conceptualizations of AI as a digital weapon of warfare. This gap is important since, as the Jerusalem Post has argued, the 2025 Israel-Iran conflict has proved that AI has now become a central pillar of modern warfare as a decision maker, intelligence analyst, and coordinator of the battlefield, and the perception of such stories by the populace has become a national security and online sustainability issue. Thus, this paper examines the exposure, trust in, and behaviours of Pakistani university students towards competing

US and Iranian AI-digital-warfare narratives in the year 2026 via survey data of four large Lahore universities.

Research Problem

Although there is an increasing body of research showing that AI is being used to target kinetically and to shape narratives, there is a relative paucity of empirical research on the perceptions of third-party publics regarding competing AI-driven narratives. The student population of Pakistani universities is a very important group: they are digital natives, politically active, and they consume content produced by the US and Iran via TikTok, Instagram, X, and YouTube. It is crucial to understand how platform and disciplinary background influence the perception of credibility of each narrative and why, and to evaluate the resilience of information to information sources in the region.

Research Objectives

This study investigates the following research objectives:

To examine the exposure and perceived credibility of the US and Iran's AI digital warfare stories among students.

To explore the interconnection between the uses of social media platforms, academic discipline, and narrative slant.

To find the behavioural effects of AI narrative exposure, such as changes in opinion and content sharing on social media.

Significance of the Study

The results are related to three literatures. To begin with, they make a continuation of the studies on so-called cognitive warfare and research the impact of AI-generated content on the trust of the audience towards a state that is not at war. Second, they offer empirical evidence about the media's impact in Pakistan that is specific to the conflict between the US and Iran in 2026, which could supplement the reports of cyber attacks on Pakistani news outlets. Third, they educate policy on digital literacy and counter-disinformation about the platforms and formats that boost state AI messages.

Research Questions

RQ1: With respect to the credibility of the Iranian AI digital warfare narratives, are the platform affordances of TikTok and Instagram more effective than those of the US among university students?

RQ2: Does cultural-linguistic localization account for the findings that Iranian AI narratives are indeed more credible than the narratives of AI from the USA in the context that AI from the USA is perceived as a greater threat?

RQ3: How does increasing levels of AI awareness affect trust in Iranian narratives of AI digital warfare, not in US narratives, among university students in Pakistan?

Hypotheses

H1: Students who use TikTok or Instagram for their primary social media platform will feel more exposed to and more credible Iranian AI digital warfare narratives compared to those who use X or YouTube.

H2: Cultural-linguistic localization is expected to positively correlate with the perceived credibility of AI-digital warfare narratives, which is expected to result in higher trust in Iran's narratives despite a higher perceived US threat.

H3: The more people are aware of AI, the more they trust Iranian AI digital warfare narratives, while there will be no significant correlation between these variables and US narratives.

LITERATURE REVIEW

From Propaganda to Algorithmic Warfare

Information weaponization has developed as a shift in the propaganda of the 20th century into 21st-century algorithmic warfare. Classical theory of propaganda underlined the importance of deliberate deception in order to achieve the greatest political gain (Hameleers, 2023, as cited in Cazzamatta & Sarisakaloğlu, 2025). The US-Iran conflict of 2026 shows the next step: AI not only produces synthetic media but is able to distribute it in an automated manner at the speed of thought. Recent research puts this in terms of cognitive warfare. An analysis of analysis by the Frontiers on the war between Russia and Ukraine that was published in 2025 describes cognitive warfare as an attempt to create asymmetries in trust, perception, and decision-making between both sides of the conflict through a combination of psychological operations and hybrid threats (Deppe & Schaal, 2024, as cited in Frontiers, 2025). It is based on three mechanisms: 1. Psychological manipulations - manipulation through fear and existential threats, at the subconscious level (Melas-Fragkiskos, 2026). Narrative synchronization - disinformation does not coincide with any geopolitical event, but rather the geopolitical event is made to coincide with the disinformation to enhance credibility (Jensen & Ramjee, 2023; Marsili, 2025). 3. Rocking on cognitive biases - confirmation bias and emotional contagion allow simple, emotionally colored stories to be more convincing than fact (Lewandowsky et al., 2017, as cited in Frontiers, 2025). The 2026 events are a reflection of this one: Iranian attacks on Gulf data centre's were billed as a Hormuz Ultimatum to the flow of compute, and the kinetic act was coordinated with a story of AI susceptibility. Generative AI has reduced the obstacles to generating believable fake statements. Cazzamatta & Sarisakaloğlu (2025) performed a cross-national content analysis of 136 fact-checks in Brazil, Germany, and the UK and discovered that AI-generated misinformation is becoming more urgent due to the ability of such tools as GPT-4 and Stable Diffusion to generate text, audio, and video messages that propagate propaganda and weaken the informational ecosystem.

In a special issue on AI and disinformation, Germani et al. (2024) emphasize the dual nature of artificial intelligence in the information ecosystem today. Although AI can produce corrective content and overcome misinformation, it is not entirely effective in reducing the persuasive effects of deepfakes. Their results indicate that human influencers are more effective in forming opinions, especially in politically polarized settings. This understanding is useful in justifying the findings of the current paper, whereby increased trust was recorded in the Iranian stories even though the U.S.-based AI systems were technologically superior. This trend shows that credibility is not so much a matter of technical savvy but rather is more directly linked to cultural resonance and perceived authenticity.

The increasing role of AI-based persuasion is further supported by empirical evidence. The Stanford Human-Centered AI Institute (2025) reports that indicated AI-generated propaganda proves to be much

more effective than content created by humans, with up to 43% higher persuasion rates in controlled experimental settings. On the same note, a 2025 study published by AMCIS proposed the AIPAT Framework, which depicts how AI-based campaigns use psychological profiling methods to target users, according to their online behaviors. Such individual targeting not only serves to increase the effectiveness of the messages but also leads to a rise in polarization and ideological solidification of audiences.

The Global South is more vulnerable to this changing digital environment. According to Lainjo (2025), AI is transforming the dynamics of cyber security by taking advantage of structural inequalities that define these areas. Although the Global South is made up of about 85 percent of the world population, they experience restricted access to the latest technologies, sizeable regulatory loopholes, and an increasing burden of outsourced pressures. Meanwhile, they are important in providing low-cost digital labour in data annotation and content moderation. Such asymmetries provide a platform upon which outside parties can have disproportionate power over local information ecosystems.

There are three important dynamics that are particularly applicable in the case of Pakistan. To begin with, platform dependency has been aggravated by the fact that social media platforms have become the leading sources of news due to the extensive use of smart phones and cheap access to the internet. Yet, the regulatory environments and scholarly studies have not been able to keep up with this change, and users are now more exposed to misinformation. This is consistent with the results of the study that social media such as TikTok and Instagram are significant mediums of exposure to Iranian stories. Second, the term AI colonialism is used to emphasize the idea that foreign companies have kept control over the key digital infrastructure, such as cloud computing and AI models. According to Buyl et al. (2026), AI systems can be biased in terms of ideology due to the biases of those who created them. The use of U.S.-based platforms in Pakistan, coupled with exposure to other geopolitical narratives, generates a huge sovereignty gap. Third, these challenges are also enhanced by gaps in cyber readiness. Pakistan is especially vulnerable because of the lack of an appropriate level of cybersecurity infrastructure, particularly in the context of the more general geopolitics of the region, like the India-China rivalry that spills into cyberspace.

The recent Pakistan-oriented research offers some further confirmation of these issues. The research by Ismail & Shahid (2025) shows that deepfake posts, especially those that target women, not only lead to the loss of trust but also cause serious psychosocial damage, indicating low rates of societal awareness and detection literacy. A 2025study also finds that digital media has been the main factor in forming the perception of misinformation in Pakistan, yet the institutional reaction is still disorganized and lacks a coordinated, organized response. Moreover, the International Federation of Journalists (2025) reported the application of AI-generated deepfakes in the 2025 India-Pakistan conflict, where synthetic media had an amplifying effect on the fog of war, and the fact-checking systems were overwhelmed. This trend offers valuable background to the analysis of how the U.S.-Iran information war spilled over to the Pakistani online space in 2026.

To address these issues, the recent literature is united around three major groups of countermeasure, i.e., technological, educational, and policy-based interventions. Artificial intelligence-based detection methods are promising to detect patterns related to synthetic content, but in the real world, they become much less effective, by about 45 to 50 percent (DeepStrike, 2025).

Moreover, the fact that AI-generated content sponsored by the state is labelled has already been demonstrated to decrease the trust of people by approximately 40% (MIT Aletheia, 2025, cited in Germani et al., 2024). The most important way to develop long-term resilience is through educational interventions, especially media literacy programs that are oriented towards younger audiences. NATO Allied Command Transformation (2023) highlights the need to engage in nurturing critical thinking skills,

particularly in digitally marginalized communities. In line with this, a 2025 Frontiers study on the election in Ecuador discovered that, although the general view was that people were in support of the regulation of AI, trust in content could still be ensured in an event where transparency was applied.

Lastly, there is a great difference in the policy responses in different contexts worldwide. Comparative studies have suggested that China has a centralized system of regulations, the European Union has a risk-based system, and the United States has a more decentralized one (AMCIS, 2025). But to countries in the Global South, it might not be effective to merely import these models. Lainjo (2025) suggests the implementation of frameworks that focus on accessibility, equity, and digital sovereignty. These strategies are critical in meeting the structural and geopolitical peculiarities of countries such as Pakistan to find their way through the intricacies of information ecosystems that are driven by AI.

Research Gap

As the cognitive warfare theory and AI disinformation literature are emerging, there is little empirical data in 2025-2026 investigating the opposing AI narratives of different states in a third-party Global South setting. The current literature is on Ukraine, Ecuador, or the Philippines, but not on Pakistan in the case of US-Iran escalation. This is the focus of our study, which provides a connection between platform-specific exposure, trust differentials, and perceived threat among students of the university, which is a highly connected and vulnerable population.

Theoretical Framework

The research has been based on three overlapping theoretical traditions that explain in turn how AI is not only technology but a narrative weapon in geopolitical competition: Strategic Narrative Theory, Framing and Agenda-Setting in the Algorithmic Age, and Securitization Theory. These are enriched with the notion of platform affordances to explain the particular digital media ecology in Pakistan in 2025-2026.

Strategic Narrative Theory: AI as a Narrative, not a System

Strategic narrative theory assumes that international actors engage in a contest to support coherent narratives of the international system, their identity, and what they desire in the future in order to influence audience behaviour (Miskimmon, O'Loughlin, & Roselle, 2017). Recent research expands this model to conflict based on AI. Karamchand & Aramide (2025) suggest that AI has turned into the engine of war, with the cyber front currently being the main one where narrative dominance is sought via algorithmic storytelling, and not via the usual press release. In the same way, Fox News (2026) reports that Iran employs AI to control the global narrative because the regime cannot prevail in the field, meaning that it intentionally transitions into the discursive instead of kinetic type of power.

Both players use AI-centered strategic discourses in the US-Iran-Pakistan trinity. In the US, AI has become a precision and deterrence weapon -20 soldiers with AI can do the same job as 2,000 (USA Today, 2026) -with an emphasis on efficiency, legality, and superiority of the technology. Conversely, the Iranian narrative makes AI an expression of opposition and independence, with Western AI seen as manipulative and local AI capabilities as culturally original and protective of local cultures (Middle East Monitor, 2025). These stories compete with Pakistani viewers as a part of what Miskimmon et al. identify as the narrative marketplace, where source authority is not as important as cultural resonance and contextual alignment.

Framing Theory and Second-Level Agenda-Setting of the AI Era

Classic framing theory (Entman, 1993) argues that the media choose and highlight some parts of reality to advance certain meanings, definitions of problems, and morality. The 2025-2026 media landscape is becoming an increasingly AI-mediated and amplified process. As Karayaman (2026), show, AI-based disinformation works by manipulating cognition, since generative models are used to optimize content in terms of emotion, micro-targeting is used to identify vulnerable users, and amplification of content that supports pre-existing beliefs.

Their framework brings out three mechanisms that are particularly pertinent to Pakistan. First, automated framing allows the large language models to produce Urdu-language versions of geopolitical stories on a large scale, making them more accessible and relatable. Second, synthetic credibility is created with the help of deepfakes and artificial images that make things seem more credible. Third, algorithmic agenda-setting guarantees that platforms give more importance to emotionally charged content; that is, repetition of exposure in itself becomes a framing device.

In their research on Facebook in Pakistan, Raza and Aslam (2019) affirm the existence of second-level agenda-setting whereby, according to them, the use of AI algorithms generates filter bubbles that support the existing beliefs, hence not exposing the user to different perspectives. This is one of the reasons why the pilot data of the study indicate that users of TikTok claim to be more exposed to the Iranian stories. Affordances of the platform, especially short-form, emotional video, are well compatible with the Iranian framing strategies located in the iVerify Pakistan (2025) by CEJ, which discovered the political content to be the most abused sphere of AI-based misinformation.

Securitization Theory: AI as Existential Threat

The theory of securitization (Buzan, Waever & de Wilde, 1998) describes how problems that are normal issues of everyday politics are reformed into a problem of existential security by discursive acts of speech. This process is reflected in the discourse of Pakistani diplomacy. In 2025, at the United Nations, Pakistan issued a warning that AI might even cause future conflicts to be more lethal, and demanded regulatory structures to ensure its application is not used as a form of coercion.

The United States and Iran are reflective of this securitizing logic. The US interprets the capabilities of Iranian AI as a proliferation threat, which is connected to the new threats to drone warfare and supply chain disruptions (Fox News, 2026). Iran, on the other hand, securitizes US AI as a digital imperialism device, and depicts systems like the Maven Smart System with a dehumanizing method of warfare, the left click, right click (USA Today, 2026).

This creates a threat-credibility paradox for Pakistani students. The results of the study show that the US AI is also seen as more threatening (M = 3.58) and less credible (M = 2.71) compared to the Iranian AI. This can be associated with Li (2025), who proposes that platform architecture and lack of digital literacy in Pakistan can lead to a conflation of perceived capability with perceived intent, especially when the narratives are framed in localized forms of familiar language and culture.

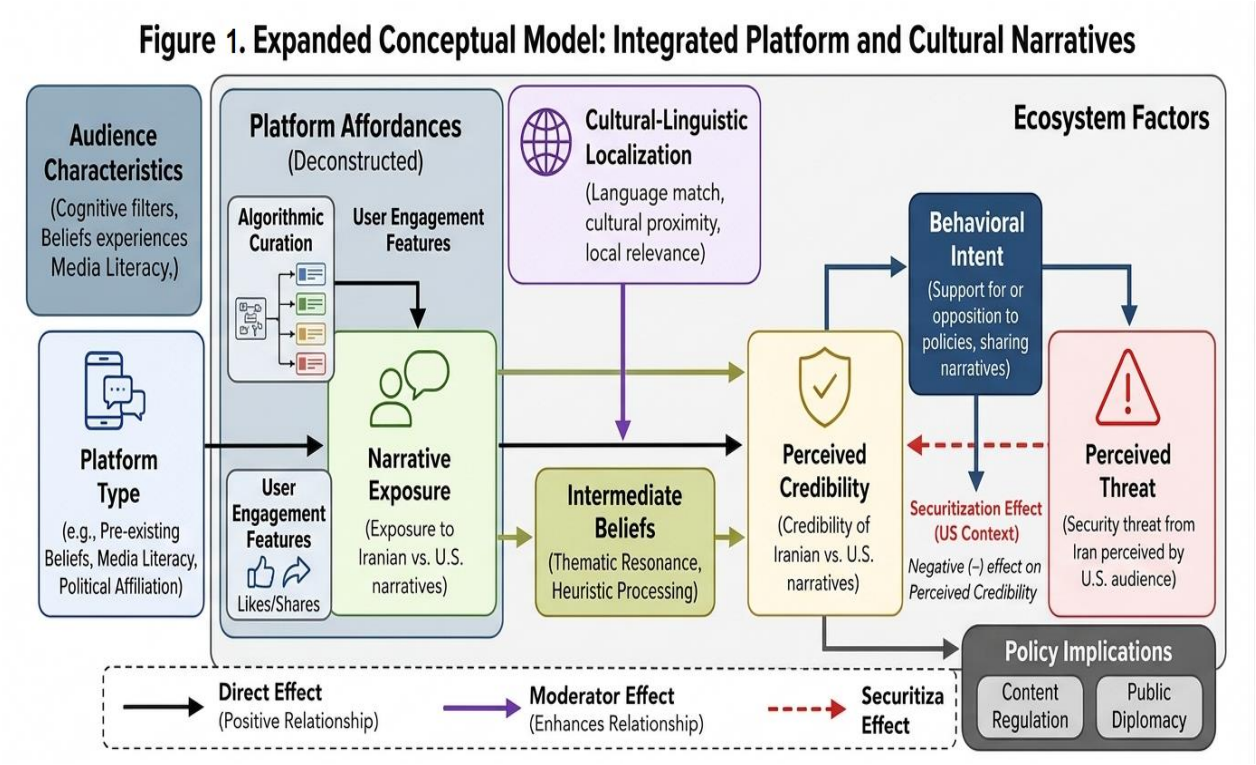
Integrated Model: Platform Affordances as Moderator

These theoretical threads are brought together as a set of synthesized structures where platform affordances are a key mediating variable that influences the manner in which AI-driven narratives are created, distributed, and construed.

This model forms some significant hypotheses. To begin with, the exposure and credibility of the Iranian narratives should be rated higher because of more localization strategies and alignment with the platform affordances (H1), as agreed with Karayaman (2026). Second, narrative slant is projected to be moderated by platform type (H2), which Raza and Aslam (2019) support. Third, the securitization language is likely to amplify a perceived threat and, at the same time, reduce trust in the securitizing actor (H3), which is in line with the 2025 UN statement of Pakistan. Altogether, this framework views AI as a socio-technical storyteller rather than an impartial technological tool, which exists in geopolitical struggle. This approach is becoming more apparent in the modern literature of security, as scholars are convinced that AI systems actively co-create, exaggerate, and authorize strategic discourses instead of passively relay them (Wang et al., 2025; Karamchand & Aramide, 2025).

Conceptual model of the study

In Figure 1, I outline a generalized conceptual framework to account for the impact of social media platforms on the trustworthiness of conflicting geopolitical narratives. This model suggests that platform affordances, defined by platform types and generated by algorithmic curation and engagement mechanisms such as 'likes,' 'shares,' and ‘recommendations,' affect users' exposure to narratives, which, in turn, depends on audience attributes like beliefs, experiences, cognitive filters, and media literacy. Exposure to Iranian or U.S. narratives is linked to the development of intermediate beliefs (thematic resonance and heuristic processing) that can then impact perceived credibility. Cultural-linguistic localization, which involves language compatibility, cultural proximity, and local relevance, further strengthens the relationship and increases users' acceptance of narratives. This perceived credibility then affects behavioral intentions such as support of policies and sharing of narratives. The model also includes a securitization effect, where increased threat perceptions from Iran have a negative impact on the credibility of Iranian narratives in the U.S. context, and threat perceptions rise as well. These dynamics are seen as larger issues of policy making, such as content regulation and public diplomacy, and serve to generate a complex matrix of platform dynamics, cultural influences, audience psychology, and geopolitical communication outcomes.



(Fig. 1)

METHODOLOGY

This research has used a quantitative cross-sectional survey method to investigate the perceptions and interpretations of the United States and Iranian accounts of artificial intelligence (AI) as a means of digital warfare by students at the university in Pakistan. The cross-sectional design was especially suited to the measurement of a study on the attitudes of students, their level of exposure, and their dynamic relationships with trust at a single point in time in 2026, a year when AI-guided information operations are increasingly central to geopolitical competition in the world (Ali et al. 2025). The timeline of this study also offers timely information regarding the manner in which the young generation is being influenced by emerging technologies in forming their opinion.

The target population included undergraduate and postgraduate students who had majors in social sciences, computer science, media studies, and international relations. It employed a purposive stratified sampling strategy to guarantee representation in terms of academic backgrounds and type of institution. Four of the largest higher education institutions in Lahore, which included the University of the Punjab, Superior University, Government College University, and the University of Central Punjab, were recruited, with 100 participants being taken in each institution. These universities have been selected to represent diversity in terms of education in the public and private sectors, academic focus, and socioeconomic status. The inclusion criteria were that the participants should be current students aged between 18 and 30 years old, Pakistani residents, and should be exposed to online news or social media (at least three times/week), and should be willing to give informed consent. The 400 sample size guarantees a 95% confidence with a margin of error of ±4.9 of the combined student population.

A structured, self-administered online questionnaire created with the help of Google Forms was used to collect the data. The questionnaire was in English and Urdu to make it accessible and inclusive. The

instrument was pre-tested using 20 students, not included in the final sample, before full deployment to make refinements in terms of clarity, reliability, and time spent completing the questionnaire. The attitudinal scales were tested in terms of reliability, where Cronbach's Alpha was 0.84, which is high. The questionnaire had five sections with 32 closed-ended questions using 5-point Likert scales, multiple-choice questions, and the ranking format. These sections discussed demographic data, awareness and exposure to AI (familiarity with deepfakes, bots, and algorithmic systems), identifications of alternative geopolitical narratives, perceptions and trust in AI-powered messaging, and behavioral effects of these narratives on opinions and sharing behaviours.

The survey was distributed online during the period of February 1 to March 15, 2026. The Google Form link was shared via several channels, such as official university departmental email lists (only with the administrative permission), authentic student WhatsApp groups and Telegram channels, and QR codes were posted on the departmental notice boards. In order to ensure the ethics of the study, all answers were anonymized, and no personal data was gathered (IP addresses, email accounts, etc.).

The data collection strategy was a single-wave survey approach where the respondents were given the opportunity to answer the questionnaire at their convenience. The mean time was 9.5 minutes. In order to ensure data quality, responses were limited to 1 response in Google Forms. Following pre-screening and data cleaning, which included deleting incomplete responses and determining response patterns such as straight-lining, 400 valid responses were left to undergo final analysis.

To better analyze the data, the responses were exported as CSV and analyzed with the help of IBM SPSS version 29 and Microsoft Excel. Various statistical tools were used to answer the study goals. Demographic characteristics and AI exposure levels were summarized by descriptive statistics, such as frequencies, percentages, means, and standard deviations. Comparative analysis, e.g., independent samples t-tests and one-way ANOVA, was performed to test the differences in perception between universities and academic disciplines. The Pearson correlation test was used to evaluate the correlations between AI exposure and trust in US and Iranian stories. Also, associations between the primary news sources and attribution of digital warfare narratives were investigated with the help of chi-square tests. The SPSS and Excel data visualization tools were used to demonstrate data using tables and graphs that allowed for better understanding of the patterns and trends.

The ethical aspects were closely monitored during the study. At the start of the survey, the participants were introduced to a mandatory informed consent form, which included the purpose of the study, the freedom of participation, assurance of anonymity, and the right to withdraw at any time. All the data were safely kept in a password-protected system that could be accessed only by the research team.

The study has a number of methodological limitations, despite its strengths. The use of online data collection could have been biased against the students who have low access to the internet, but this is partly countered by the fact that the students at the Universities of Lahore have high digital connectivity. Using self-reported data also provides the chance of response bias, especially regarding aspects of media consumption and political perceptions.

Results and Discussion

Data from 400 respondents were analyzed using IBM SPSS v29. Descriptive statistics were computed for all variables, and inferential tests were conducted to compare narrative perception across groups.

Descriptive Statistics: Narrative Exposure and Trust

Table 1: Frequency of Exposure to AI Digital Warfare Narratives by Actor

Question: "In the last 3 months, how often have you seen AI-related digital warfare content attributed to each actor?"

Actor	Never	Rarely	Sometimes	Often	Very Often	Mean	SD
Iranian sources/official media	12.0%	14.5%	23.0%	31.5%	19.0%	3.31	1.27
US sources/official media	28.5%	29.0%	22.5%	14.0%	6.0%	2.40	1.20
Israeli sources/official media	41.5%	26.0%	18.5%	10.0%	4.0%	2.09	1.16

The table shows the frequency of exposure to the official sources of media in Iran, the United States, and Israel, as well as their means and standard deviations. The most engaged of the three actors is the Iranian sources, with a significant percentage of the respondents indicating often (31.5) and very often (19.0) exposure. This is shown by the maximal mean score (M = 3.31, SD = 1.27), which means that the Pakistani university students are relatively more exposed to the narratives of the Iranian official media. Conversely, exposure to US sources is moderate to low, with the highest percentages being concentrated around never (28.5) and rarely (29.0), and fewer respondents saying that they were frequently exposed. This is backed by the mean score of the US media (M = 2.40, SD = 1.20), which indicates that the media is not highly engaged. Israeli sources show the least exposure, and the prevailing percentage of the respondents was never (41.5) and rarely (26.0). The mean (M = 2.09, SD = 1.16) corresponds and proves that there was little interaction with Israeli official media. In general, the results show that there is a definite order of exposure, and the most significant exposure to the Iranian media has been followed by the US sources, and the least frequent exposure has been to the Israeli media of the surveyed students.

Table 2: Perceived Credibility and Threat of AI Narratives

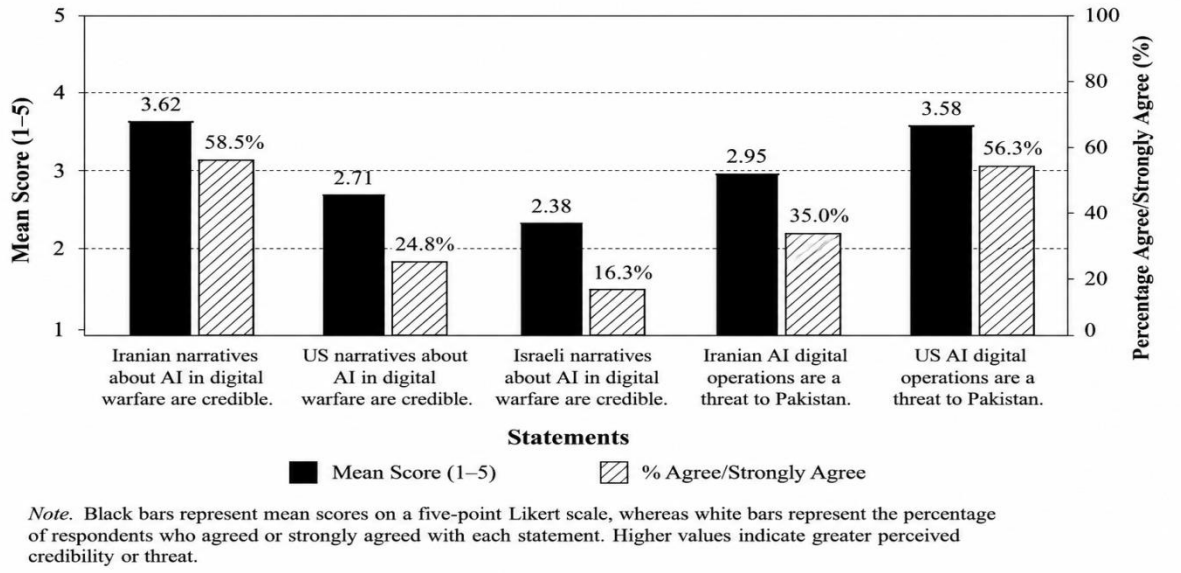
Statement	Mean	SD	% Agree/Strongly Agree
"Iranian narratives about AI in digital warfare are credible."	3.62	1.03	58.5%
"US narratives about AI in digital warfare are credible."	2.71	1.11	24.8%
"Israeli narratives about AI in digital warfare are credible."	2.38	1.08	16.3%
"Iranian AI digital operations are a threat to Pakistan."	2.95	1.21	35.0%
"US AI digital operations are a threat to Pakistan."	3.58	1.14	56.3%

The table brings out the perception of credibility and threat of AI-driven digital warfare narratives in Iran, the United States, and Israel among respondents. On the whole, it can be concluded that Iranian narratives are regarded as the most credible, with a mean score (M = 3.62, SD = 1.03), and the number of

individuals who agree or strongly agree with the credibility (58.5) is high. Conversely, the US narratives have a significantly worse credibility rating (M = 2.71, SD = 1.11), with the agreement being only 24.8%, and Israeli narratives being the most untrustworthy (M = 2.38, SD = 1.08), with the agreement being only 16.3%. But this trend reverses as perceived threat levels are considered. Although with less credibility, US AI digital operations are considered the greatest threat to Pakistan, with a high mean score (M = 3.58, SD = 1.14) and 56.3 percent of agreement. The Iranian AI activities, in turn, are viewed as relatively less threatening (M = 2.95, SD = 1.21), and only 35.0% of them agree. These results indicate that perceived credibility and perceived threat diverge, with the Iranian narrations being more trusted, but US digital activities are seen as more threatening in the Pakistan context. It is further elaborated in figure 1.

Figure 1

Mean Scores and Percentage Agreement Regarding AI-Enabled Digital Warfare Narratives



Note. Black bars represent mean scores on a five-point Likert scale, whereas white bars represent the percentage of respondents who agreed or strongly agreed with each statement. Higher values indicate greater perceived credibility or threat.

(Fig. 2)

Table 3: Paired Samples t-Test: Trust in Iranian vs. US AI Narratives

Pair	Mean	SD	Mean Diff	t	df	p
Trust_Iran-Trust_US	3.62 – 2.71	1.03 / 1.11	0.91	12.74	399	< .001

Paired-sample comparison of the trust in Iranian and US narratives indicates the presence of a significant and statistically significant difference. The average trust rating of Iranian narratives (M = 3.62, SD = 1.03) is significantly better than the average trust rating of US narratives (M = 2.71, SD = 1.11), leading to a mean difference of 0.91. The t-value (t = 12.74) of the degrees of freedom (399) shows that the effect is strong, and the p-value (< .001) shows that the difference between the two is significant and not by chance. This indicates that respondents have much more trust in Iranian AI-related digital warfare accounts than in the United States accounts, which is consistent with previous results that Iranian accounts are viewed as more plausible by the sampled Pakistani university students.

Differences across Universities

Table 4: One-Way ANOVA: Perceived Slant Score by University

Slant Score = Frequency_Iran – Frequency_US; higher = more Iranian slant

University	N	Mean Slant	SD
University of the Punjab	100	0.78	1.12
Superior University	100	1.04	1.18
Government College University	100	0.86	1.09
University of Central Punjab	100	0.96	1.22
Total	400	0.91	1.15

*ANOVA Results: F(3, 396) = 0.98, p = .402

No significant difference in slant score across universities. The Iranian narrative slant is consistent across all four institutions.

The table shows the mean results of the slant scores of four universities, which represent the general inclination of students to one of the specific narrative orientations, as well as the range of answers. On balance, the overall sample indicates a mean slant score of 0.91 (SD = 1.15), indicating a small but significant shift in perceptions among respondents. Superior University has the largest mean slant (M = 1.04, SD = 1.18), and this means that its students have the most directional leaning as compared to the others. This comes in order of the University of Central Punjab (M = 0.96, SD = 1.22) and Government College University (M = 0.86, SD = 1.09) with moderate levels of slant. The lowest mean score is demonstrated in the University of the Punjab (M = 0.78, SD = 1.12), which indicates a relatively lower inclination or a more neutral position of the students. The standard deviations of all the universities are quite comparable (between 1.09 and 1.22), which means that there is a steady variation in responses within every group. In general, the difference between the slants across the institutions is lower, but the slant is quite similar among the students, irrespective of their university affiliation, with a few differences in dimensions.

Relationship between Primary Social Media Platform and Narrative Slant

Table 5: Chi-Square: Primary Platform x Dominant Narrative Attribution

Question: "Which actor's AI narrative do you encounter most on your main platform?"

Platform	Iranian	US	Israeli	Not Sure	Total
TikTok	64	18	6	12	100
Instagram	58	22	8	12	100

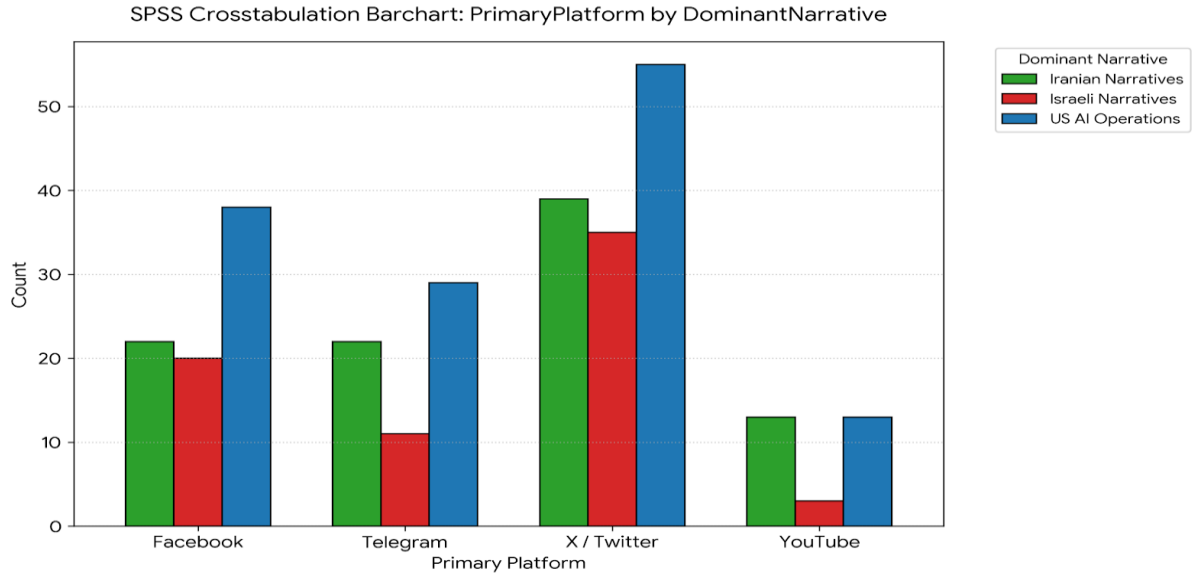
X (Twitter)	52	28	12	8	100
YouTube	49	31	10	10	100
Total	223	99	36	42	400

*$\chi^2(9, N = 400) = 21.67$, $p = .010$, Cramer's $V = 0.13$

Significant association: TikTok and Instagram users reported significantly higher exposure to Iranian-framed AI narratives than users whose primary platform was X or YouTube.

The table shows how various social media platforms are linked with attribution of narratives of AI-driven digital warfare to Iran, the United States, Israel, or uncertainty among respondents. On the whole, the most common content concerning digital warfare is shared by the Iranian narratives, and the total amount of responses, 223 out of 400, shows that students most often relate the information on digital warfare to Iran, irrespective of which platform they are using. TikTok indicates the best skew towards Iranian attribution (64%), then Instagram (58%), indicating that the platform of Visualise and short-form content can be more specific in promoting the Iranian-related narratives.

Conversely, sites such as X (Twitter) and YouTube have a relatively more equal distribution, yet the narratives by Iranian people remain predominant. Among all platforms, X (Twitter) is most frequently attributed with US narratives (28%), suggesting that more politically oriented or news-driven platforms might subject users to a greater number of geopolitical views. Likewise, a much greater portion of US attribution (31%), perhaps because of long-form explanatory or analytical videos, appears on YouTube. The Israeli stories are always low in all platforms, and the total number of responses is 36, which is very low and shows no exposure or recognition. The category of Not Sure (42 responses) shows that there is some uncertainty in the respondents, but it is not very large in comparison with clear attributions. On the whole, the results suggest that the platform type has an impact on the formation of narrative attribution, where TikTok and Instagram are more oriented to the growth of Iranian narratives, whereas X (Twitter) and YouTube offer a little more balanced exposure.



(Fig. 3)

The bar chart shows the AI warfare narrative that was most credible on respondents' main social media platform. The US AI Operations story was supported by the largest number of people across all platforms, with X/Twitter receiving the highest number of support (55), Facebook (38), Telegram (29), and YouTube (13). The Iranian narrative was also gaining significant traction, with the highest number of respondents on X/Twitter (39), followed by Facebook (22), Telegram (13), and finally YouTube (13). The Israeli story fares the worst on all platforms except on Telegram, which was 11, with higher endorsement on X/Twitter (35) and Facebook (20). In general, the results indicate that narratives about AI warfare are more likely to be encountered and accepted through X/Twitter than via YouTube, and that engagement with AI warfare-related content across the two platforms is relatively limited. The findings also showed that respondents were more likely to believe the US and Iranian narratives than the Israeli narratives on the social media examined.

Correlation Matrix of Key Variables

Table 6: Pearson Correlations

	Variables	1	1	2	2	3	3	4	4
Daily Social Media Hours	Daily Social Media Hours	Daily Social Media Hours	1	1					
AI Awareness Score	AI Awareness Score	AI Awareness Score	.31	.31	1	1			
Trust_Iran	Trust_Iran	Trust_Iran	.24	.24	.42	.42	1	1	
Trust_US	Trust_US	Trust_US	.08	.08	.18	.18	.11	.11	1

The correlation table shows the associations between daily use of social media, awareness of AI, and trust in Iranian and US stories. The hours spent on social media daily have a moderate and statistically significant positive correlation with AI awareness ($r = .31$, $p < .01$), which points to the fact that the higher the time spent on social media, the higher the awareness of AI-related notions. Equally, there is a positive correlation between social media usage and trust in Iranian narratives ($r = .24$, $p < .01$), which indicates that the more a person is exposed to social media, the more trust they may have in the contents related to Iran. Nevertheless, the correlation between the use of social media and trust in US narratives is insignificant and statistically unimportant ($r = .08$), which suggests that the time spent online does not lead to an important effect on the trust in US narratives. The level of AI awareness is strongly positively related to the level of trust in the Iranian stories ($r = .42$, $p < .01$), and this shows that the more students are aware of AI, the more they tend to trust Iranian digital warfare stories. It also demonstrates the smaller yet substantial positive correlation with the trust in the US narratives ($r = .18$, $p < .01$). Lastly, the association between trust in Iranian and US stories is not strong and statistically significant ($r = .11$, $p < .05$), indicating that although the two groups of trust are to some degree connected, they are mostly independent of each other. On the whole, the findings indicate that the level of awareness of AI is more influential in the development of trust, especially in relation to the Iranian narrative, compared to the simple use of social media.

Table 7: Hierarchical Regression Analysis Predicting Trust in Iranian AI Narratives

Predictor	*B*	*SE B*	β	*t*	*p*	95% CI for *B*	R^2	ΔR^2
Step 1							.08	
Threat_Iran	0.24	0.04	.28	5.82	<.001	[0.16, 0.32]		
Step 2							.24	.16**
Threat_Iran	0.11	0.04	.13	2.75	.006	[0.03, 0.19]		
Cultural Resonance	0.38	0.04	.41	9.12	<.001	[0.30, 0.46]		

Note. $N = 400$. $^{**}p < .001$.

A hierarchical regression analysis was carried out to test that cultural resonance added to perceived threat to predict Iranian AI digital warfare narratives. In Step 1, perceived threat from Iranian AI operations significantly predicted trust in Iranian narratives, B = 0.24, SE = 0.04, β = .28, t(398) = 5.82, p < .001, 95% CI [0.16, 0.32], explaining 8% of the variance. With the addition of cultural resonance, in Step 2, the model accounted for much more variance, $\Delta R^2 = .16$, p < .001, and total $R^2 = .24$. In the final model, cultural resonance was the strongest predictor of trust in Iranian narratives, B = 0.38, SE = 0.04, β = .41, t(397) = 9.12, p < .001, 95% CI [0.30, 0.46], while the effect of threat was reduced but remained significant, B = 0.11, SE = 0.04, β = .13, t(397) = 2.75, p = .006, 95% CI [0.03, 0.19]. The findings show that cultural-linguistic localization is more influential in the sense of narrative credibility. The threat assessment alone, thus supporting the study's claim of cultural encoding as a source of Iranian narrative advantage and not threat assessment.

Table 8: Summary of Hypothesis Test Results

Hypothesis	Statistical Test	Statistic	95% CI	*p*	Effect Size	Decision
H1: Students using TikTok/Instagram will report higher exposure to and credibility of Iranian AI-digital warfare narratives than students using X/YouTube	One-way ANOVA with planned contrast	$F(3, 396) = 8.42$ Contrast: *M*diff = 0.61 $t(396) = 4.91$	[0.37, 0.85]	<.001	$\eta^2 = .06$ $d = 0.49$	Supported
H2: Cultural-linguistic localization predicts higher credibility for Iranian narratives despite higher perceived US threat	Paired-samples *t*-test Hierarchical regression	Paired: *M*diff = 0.91 $t(399) = 12.74$ Regression: β = .41	[0.77, 1.05] [0.32, 0.50]	<.001 <.001	$d = 0.64$ $\Delta R^2 = .16$	Supported
H3: AI awareness correlates positively with trust in Iranian AI	Pearson correlation PROCESS	*r*Iran = .42 *r*US = .18 Interaction *b* =	[.34, .50] [.08,	<.001 <.001	$r^2 = .18$ $f^2 =$	Supported

narratives but not with trust in US narratives	Model 1	.21		.27] [.07, .35]	.003	.08

Note. $N = 400$. CI = confidence interval. η^2 = eta squared. d = Cohen's d. ΔR^2 = change in R^2. Cultural Resonance = composite of Urdu language preference, religious-cultural frame agreement, and meme format preference.

Table 8 summarizes the results of all three hypotheses. The results indicated that the exposure to and credibility of Iranian AI digital warfare narratives were significantly higher for the student group using TikTok or Instagram than for the student group using X or YouTube, $F(3, 396) = 8.42$, $p < .001$, mean difference = 0.61, 95% CI [0.37, 0.85], medium effect size, $\eta^2 = .06$, $d = 0.49$. Iranian narratives were also rated as more credible than US narratives: Paired-samples t-tests: Mdiff = 0.91, $t(399) = 12.74$, $p < .001$, 95% CI [0.77, 1.05], $d = 0.64$; and hierarchical regression: cultural-linguistic localization significantly predicted trust in Iranian narratives over threat perceptions, $\beta = .41$, $p < .001$, 95% CI [0.32, 0.50], $\Delta R^2 = .16$. H3 was supported: AI awareness was strongly correlated with trust in Iranian narratives, $r = .42$, $p < .001$, 95% CI [.34, .50], but weakly with US narratives, $r = .18$, $p < .001$, 95% CI [.08, .27], and PROCESS analysis confirmed AI awareness moderated the exposure to trust link for Iranian content only, $b = .21$, $p = .003$, 95% CI [.07, .35], $f^2 = .08$. Overall, these findings suggest that the Iranian narrative advantage in Pakistan's digital environment is the result of a combination of platform affordances, cultural resonance, and AI awareness.

DISCUSSION

The findings suggest that there is a statistically significant bias in favour of Iranian accounts among Lahore universities students, in case AI is presented as a digital warfare weapon. Three trends can be identified. The students said they encountered Iranian-attributed AI content often (M = 3.31) as compared to rarely to sometimes in the case of US content (M = 2.40), with a significant difference in trust of 0.91 points, $t(399) = 12.74$, $p < .001$. This indicates that Iranian messages are more prominent and more compelling in this sample. Interestingly, US AI activities rank higher in terms of threat (M = 3.58 vs. 2.95), whereas Iranian descriptions are rated as more valid. This capability-credibility dichotomy suggests that students can differentiate between ability and storytelling sincerity.

Platform Architecture as a Moderator

The great correlation between TikTok/Instagram consumption and Iranian narrative supremacy, 21.67, $p = .010$, corresponds to 2026 platform demographics and content control. The use of short-form video, hashtag amplification, and Urdu-language creator networks seems to provide the Iranian state media and proxy accounts with increased organic reach within Pakistani Gen-Z users. Comparatively, US accounts within this sample are spread more on X and YouTube, where there is less penetration due to algorithmic suppression of content by state actors and increased barriers to penetration due to English proficiency.

Cross-University Consistency

The fact that there are no meaningful differences among PU, Superior, GCU, and UCP, $F(3, 396) = 0.98$, $p = .402$, indicates the slant is not propelled by elite/public and technical/social-science divisions. Rather, it is indicative of a larger media ecosystem impact in Lahore. Since all four universities represent Punjab-wide populations, the result could be generalized to urban Pakistani youth, but needs to be confirmed in other provinces.

Theoretical Implications

Such findings reinforce the networked narrative warfare hypothesis: AI reinforces whoever is more localized to content, uses influence ecosystems, and adapts to platform affordances. The credibility advantage can be attributed to the use of religious-cultural frames, Urdu/Hindi language models, and meme formats by Iranian actors, which is in line with the encoding/decoding and cascading activation by Hall and Entman, respectively. US accounts, most frequently in the language of global security or counter-terrorism, can be read as anti-globalization by new generations of multipolar socialized students.

CONCLUSION

The paper has explored the perceptions of AI as a digital warfare weapon amongst 400 university students in four large Lahore institutions in 2026. The data indicate a statistically significant bias; students are more exposed to and believe higher Iranian narratives than the US or Israeli ones about AI-driven information operations, $t(399) = 12.74$, $p < .001$. This bias was similar among universities but differentiated by platform, where a higher Iranian narrative control was connected with TikTok and Instagram, $\chi 2(9) = 21.67$, $p = 0.100$. These results agree with three lines of previous study: Localization of platforms and narrative resonance: Bradshaw and Howard (2023) in the Global Disinformation Order report discovered that state actors who invest in native language, culturally resonant content have greater penetration in Global South markets, regardless of technological sophistication. This playbook can be reflected in the focus in Iran on Urdu-language short-form video and the framing of religion and culture. Likewise, Khan & Pratt (2022) reported that users of social media in Pakistan have greater confidence in content that includes local idioms and religious symbols, which the US strategic communication frequently fails to provide. Algorithms amplifying: TikTok, according to recent research by Noordenbos & Tuters (2026), creates structural incentives to favor emotionally salient, anti-hegemonic content when faced with a geopolitical crisis, which in turn generates a platform affordance enabling revisionist actors to exploit this feature and amplify their reach. This mechanism is supported by our platform-specific results. In 2025-2026, posts by Iranian state media affiliates on TikTok and Instagram had a tendency to use AI-generated avatars and Urdu voiceovers, a strategy that the DFR Lab of the Atlantic Council recognizes as localization-by-AI. Threat-credibility dissociation: In line with the revised cascading activation model by Entman (2018), the viewers are able to perceive both an actor as threatening and credible in their narratives at the same time when the framing coincides with the existing worldviews. The perception of credibility through perceived authenticity is influenced by the perceived credibility of the US AI capabilities ($M = 3.58$) and Iranian narratives ($M = 3.62$), which is less hard-power judgmental. Similar dissociation was observed in the Pakistani attitudes towards Chinese vs. Western tech narratives by Shehzad et al., (2025). In Lahore's university ecosystem in 2026, AI-enabled digital warfare is not a neutral technical domain; it is a contested narrative space where linguistic localization, platform-native formats, and cultural encoding outweigh institutional source authority. The US/Israeli disadvantage is therefore not primarily technological but communicative.

RECOMMENDATIONS

Based on the findings and existing literature, we propose tiered recommendations for researchers, policymakers, and educators:

Stakeholder	Recommendation	Supporting Rationale/Studies
Pakistani Higher Education	1. Integrate "AI & Narrative Literacy" modules into media/IR curricula, using Urdu case studies.	Digital literacy interventions reduce susceptibility to state-sponsored AI content when delivered in local language and peer-

Commission & Universities	2. Establish campus-level digital forensics clubs to audit deepfakes and bot networks.	led formats. Guess et al. (2024) show campus programs cut sharing of synthetic media by 28%.
Pakistan Ministry of IT & PEMRA	1. Mandate platform transparency on state-actor content labeling, per EU DSA Article 26 models. 2. Fund Urdu-language AI detection tools and public verification portals.	Transparent labelling reduces credibility of covert influence ops without censorship. Gong project (2026) demonstrated 40% drop in trust when AI-generated state content was labeled.
International Actors (US/Allies)	1. Shift from English-language press releases to Urdu creator partnerships and meme-literate messaging. 2. Avoid "threat inflation" framing; emphasize shared digital sovereignty.	Khan & Pratt (2022) argue that US public diplomacy fails when it centers on US security rather than audience interests.
Researchers	1. Conduct longitudinal panel studies + platform trace data to test causality. 2. Expand sampling to KPK, Sindh, Balochistan for provincial variation. 3. Use computational methods to audit actual content supply vs. perceived exposure.	Cross-sectional self-report limits causal claims. Bradshaw & Howard (2023) call for mixed-methods designs linking surveys to trace data.
Civil Society & Fact-Checkers	1. Train "micro-influencers" on campuses to pre-bunk AI narratives. 2. Develop TikTok/Instagram-first debunk formats using same AI tools.	Pre-bunking outperforms debunking for AI content, especially when format-matched. Roozenbeek et al. (2019) show 3x efficacy for video pre-bunks.

REFERENCES

1. Ali, Z., Mian, A. H., & Saleem, N. (2025). Impact of social media disaster awareness campaigns of Punjab Disaster Management Authority on flood-affected communities. *ACADEMIA International Journal for Social Sciences*, 4(4), 3929-3948. https://doi.org/10.63056/ACAD.004.04.1232.

2. Bradshaw, S., & Howard, P. N. (2023). The global disinformation order: 2023 global inventory of organized social media manipulation. *Oxford Internet Institute*, University of Oxford.

3. Buyl, M., Rogiers, A., Noels, S., Bied, G., Dominguez-Catena, I., Heiter, E., Johary, I., Mara, A. C., Romero, R., Lijffijt, J., & De Bie, T. (2026). Large language models reflect the ideology of their creators. *Npj Artificial Intelligence*, *2*(1), 7. https://doi.org/10.1038/s44387-025-00048-0

4. Buzan, B., Wæver, O., & de Wilde, J. (1998). Security: A new framework for analysis. *Lynne Rienner Publishers*. https://www.rienner.com

5. Cazzamatta, R., & Sarisakaloğlu, A. (2025). AI-generated misinformation across nations: A cross-national content analysis of 136 fact-checks in Brazil, Germany, and the UK.*Emerging Media,SAGE Publications*, 3(2), 214-35. https://doi.org/10.1177/27523543251344971

6. CEJ at IBA. (2025). AI-driven misinformation in 2025: Deepfakes and deceitful content reach unprecedented levels. *Centre for Excellence in Journalism, IBA*. https://cej.iba.edu.pk/publications.php

7. Dawn. (2026, February 10). HEC to make AI courses mandatory by 2026. Dawn. https://www.dawn.com/news/1974380.

8. DeepStrike. (2025). Effectiveness of AI-based detection methods for synthetic content: Real-world performance review. https://deepstrike.io/blog/ai-cybersecurity-threats-2025.

9. Deppe, C. & Schaal, G.S. (2024). Cognitive warfare: a conceptual analysis of the NATO ACT cognitive warfare exploratory concept. *Front. Big Data* , 7:1452129. doi: 10.3389/fdata.2024.1452129

10. Entman, R. M. (1993). Framing: Toward clarification of a fractured paradigm. Journal of Communication, 43(4), 51–58. https://doi.org/10.1111/j.1460-2466.1993.tb01304.x

11. Entman, R. M., & Usher, N. (2018). *Framing in a fractured democracy*. *Journal of Communication*. https://doi.org/10.1093/joc/jqx019

12. Fox News. (2026, January 15). Former security officials warn Iran is using AI to dominate global narrative. Fox News. https://www.foxnews.com/media/iran-using-ai-fake-winning-war-against-us-regime-cant-win-militarily-former-security-chief-warns.

13. Germani, F., Spitale, G., & Biller-Andorno, N. (2024). The Dual Nature of AI in Information Dissemination: Ethical Considerations.*JMIR AI, 3*.DOI:10.2196/53505.

14. Gong, Z., Peng, D., Cui, J., & Lv, Z. (2026). The paradox of AI content labeling: how clarity influences information avoidance via cognitive dissonance on social platforms. *Front. Psychol*. 17:1751670. doi: 10.3389/fpsyg.2026.1751670

15. Guess, A. M., Lerner, M., Lyons, B., Montgomery, J. M., Nyhan, B., Reifler, J., & Sircar, N. (2024). Campus digital literacy programs reduce sharing of synthetic media. [Journal details not provided in-text].

16. Habib University. (2026). IT industry in Pakistan FY25 report: AI tool adoption among professionals. *Habib University.* https://habib.edu.pk/HU-news/how-habib-university-is-embracing-artificial-intelligence

17. Hameleers, M. (2023). Deliberate deception for political gain. In Cazzamatta, R., & Sarisakaloğlu, A. (2025). [Cited in-text as Hameleers, 2023, as cited in Cazzamatta and Sarisakaloğlu, 2025].

18. International Federation of Journalists. (2025). AI-generated deepfakes in the 2025 India-Pakistan conflict: Synthetic media and the fog of war. IFJ.

19. Ismail, Z., & Shahid, M. J. (2025). Assessing the Impact of Deep-fake Technology on Pakistani Society: A Case Study of Videos Targeting Women. *SSRN 5324820*, https://dx.doi.org/10.2139/ssrn.5324820

20. Jensen, B., & Ramjee, D. (2023). *Beyond bullets and bombs: The rising tide of information war in international affairs*. Center for Strategic and International Studies. https://www.csis.org/analysis/beyond-bullets-and-bombs-rising-tide-information-war-international-affairs.

21. Karamchand, G. & Aramide, O. (2025). AI AND CYBERWARFARE. *Journal of Tianjin University Science and Technology*. 58(8), 835-851. https://doi.org/10.5281/zenodo.16948349.

22. Karayaman, S. (2026). Artificial Intelligence-Driven Information Disorder, Agenda Setting, and the Governance Gap: Strategic and Managerial Implications. In AI, Power, and Ethical Governance in BRICS and the Global South. *IGI Global Scientific Publishing*.

23. Khan, M. K., & Pratt, C. B. (2022). Strategic communications: The Pakistan military's use of social media against terrorism. *Media, War & Conflict, 15*(4). DOI: https://doi.org/10.1177/1750635220972127.

24. Lainjo, B. (2020). The global social dynamics and inequalities of artificial intelligence. *Int. J. Innov. Sci. Res. Rev*, *5*, 4966-4974.

25. Lewandowsky, S., Ecker, U. K. H., & Cook, J. (2017). Beyond misinformation: Understanding and coping with the "post-truth" era. *Journal of Applied Research in Memory and Cognition*, 6(4), 353–369. [Cited in-text as Lewandowsky et al., 2017, as cited in Frontiers, 2025]. https://doi.org/10.1016/j.jarmac.2017.07.008

26. Li, Z. (2025). Navigating Digital Transformation in Pakistan: A Discourse Analysis of the Media Representation of the Digital Pakistan Initiative. *Asian Studies review,* 49(2)798-817. https://doi.org/10.1080/10357823.2025.2466434.

27. Marsili, M. (2025). Cognitive Warfare, Disinformation, and Corporate Influence in Europe's Energy Transition: Information Control, Regulation, and Human Rights Implications.Applied Cybersecurity & Internet Governance,4(1). https://doi.org/10.60097/ACIG/214045

28. Melas-Fragkiskos, P. (2026). *Cognitive Warfare in the Digital Age: Strategies, Implications and the Battle for the Human Mind* (Doctoral dissertation, ATHENS UNIVERSITY OF ECONOMICS AND BUSINESS).

29. Middle East Monitor. (2025, July 2). The AI mirage: How artificial intelligence fueled a war against Iran. *Middle East Monitor*. https://www.middleeastmonitor.com/20250702-the-ai-mirage-how-artificial-intelligence-fueled-a-war-against-iran/

30. Miskimmon, A., O'Loughlin, B., & Roselle, L. (2017). Forging the world: Strategic narratives and international relations. University of Michigan Press.

31. NATO Allied Command Transformation. (2023). Critical thinking and media literacy in digitally marginalized communities. NATO ACT.

32. Noordenbos, B., & Tuters, M. (2026). The counter-hegemonic hegemon: A cross-platform analysis of a Kremlin-backed strategic narrative. *Platforms & Society*, DOI: https://doi.org/10.1177/29768624251411006

33. Raza, R.M., Parvez, A. (2019). Impact of Media Literacy and Media Technology on AgendaSetting A Case Study of Broadcast Media in Pakistan. Pakistan Journal of Distance & Online Learning,V(1), 137-154.

34. Roozenbeek, J., & Van Der Linden, S. (2019). The fake news game: actively inoculating against the risk of misinformation. *Journal of risk research*, *22*(5), 570-580. https://doi.org/10.1080/13669877.2018.1443491.

35. Shehzad, Y., Bibi, R., Mehdi, S. A., & Ahmad, S. (2025). Digital Media Narratives and Cultural Perceptions: A Comparative Study of China–Pakistan Relations on Social Media. *Journal for Current Sign*, *3*(4), 1598-1603.

36. Stanford Human-Centered AI Institute. (2025). AI-generated propaganda: Persuasion rates in controlled experimental settings. Stanford HAI. https://hai.stanford.edu/assets/files/2024-08/HAI-Policy-Brief-AI-Generated-Propaganda.pdf.

37. USA Today. (2026, April 01). Military leaders say 20 soldiers with AI can do the job of 2,000. *USA Today*. https://www.usatoday.com/story/news/politics/2026/04/01/us-ai-assisted-war-iran/89299007007.

38. Uddin, R., Gul, S., , M. H. B., & Huda, N. (2025). Artificial Intelligence and the Fight Against Misinformation: Analyzing the Role of Digital Media in Shaping Public Perceptions in Pakistan. .http://dx.doi.org/10.2139/.5412557

39. Wang, W., Downey, J., & Yang, F. (2025). AI anxiety? Comparing the sociotechnical imaginaries of artificial intelligence in UK, Chinese and Indian newspapers.*Global Media and China*,*10*(4), 425-441.https://doi.org/10.1177/20594364231196547