sciendo

STUDIES IN LOGIC, GRAMMAR AND RHETORIC 70 (83) 2025

DOI: 10.2478/slgr-2025-0056

Zoya Slavina
University of Białystok
Poland
e-mail: z.slavina@uwb.edu.pl, slavina.aiethics@gmail.com
ORCID: 0009-0007-6346-3802

AI AS A POLITICAL ARTEFACT: TECHNOLOGICAL MEDIATION IN CONSTRUCTIVIST INTERNATIONAL RELATIONS

Abstract. This article advances a constructivist account of artificial intelligence (AI) as a political artefact in international relations (IR). Drawing on technopolitics, Actor–Network Theory (ANT), and postphenomenology, it argues that AI systems should be examined not only through the meanings stakeholders ascribe to them, but also through how design choices materialize normative commitments that preconfigure perception, judgment, and action. ANT highlights the agency of non-human actants and the affordances inscribed in technical artefacts, while postphenomenology specifies the mediation modes – embodiment, hermeneutic, alterity, and background – through which increasingly autonomous and data-driven systems reshape IR decision environments. The article is conceptual in nature, supported by illustrative evidence from decision-support systems (DSS) used in NATO and the European Union, including AI-enabled tools for operational command and conflict early warning. These cases demonstrate how DSS redistribute epistemic authority and responsibility across technical and institutional layers, accelerating sense-making under time pressure while introducing new risks. The contribution addresses a research gap by integrating insights from postphenomenology and the sociology of technology into IR, proposing a permissive ontology that treats (semi-)autonomous systems as analytical actants without assuming personhood, and offering conceptual tools for examining mediation in practice and the political implications of disruptive technologies.

Keywords: artificial intelligence (AI), international relations constructivism, postphenomenology, decision-support systems (DSS), technological mediation.

Introduction

The deepening integration of artificial intelligence (AI) into international political practices and decision-making increasingly exposes blind spots in major international relations (IR) theorising that has not yet captured the changes disruptive informational technologies bring. Despite constructivist recognition that ideas and norms have a substantial impact, IR discipline has not yet offered a comprehensive account of technology as

constitutive of international politics. As information systems are delegated evaluative tasks, compress decision time – often operating with relatively opaque logics and diffused accountability – familiar explanatory concepts that assume human decision-makers are expected to keep pace. At the same time, algorithmic systems can materially improve practice by extracting significant patterns from data, improving analytical insights and speed, fusing multiple data sources, or even reducing bias in decision-making; yet those insights are only as reliable as the data on which they rest (e.g. incomplete, biased, or outdated datasets can mislead), and system architectures (e.g., design, ethical accounts). Furthermore, the implementation and the context of deployment have a significance, too. The appearance of competence that comes with large-scale data analysis may invite overconfidence and over-reliance (Passi & Vorvoreanu, 2022); as with any machinery, sober boundaries and capacities must be kept in view if sovereign decision-making is to remain responsible. As examples presented later in this article indicate, in AI-enabled IR settings, these tensions show up real-world applications such as decision-support systems, where model architecture and data quietly steer what counts. In this view, recent research warns that fully automated decision-making in specific task domains risks a loss of practical wisdom in consequential choices if lacking conscious and meaningful implementation; while more broadly, IR faces a crisis of relevance if it cannot explain the contributions of algorithmic systems, ICT infrastructures, and autonomous technologies (e.g. algorithms, big-data infrastructures, robotics, AI etc.) to world politics, a challenge that calls for "flattening the hierarchy among the object/subjects we investigate" so that nonhumans figure in our explanations where appropriate (Kiggins, 2025, p. 313).

Constructivism – understood here in the IR sense, and acknowledging the debate over whether it is a theory or an analytical approach (Reus-Smit, 2022) – remains a promising starting point because of its focus on social constitution. Crucially, constructivism does not assume complete technological neutrality; as it foregrounds the meanings and identities that actants ascribe to technology in contexts. The concept of this article is to open a dialogue with science and technology studies (STS) – more precisely, with technopolitics and the sociology/philosophy of technology – to attend to the inherent political and ideological properties of technological artefacts. In this way, the constructivist meanings ascribed to the artefacts are enriched by the inherent political properties recognised in STS. Here, AI systems are approached as a new type of political artefact with autonomous and smart capacities, whose design and deployment increasingly mediate IR practices. To avoid confusion with IR's conventional use of "actors" (usually states, international organisations, firms, or NGOs with recognised

agency), I use a substitutive term "actant" drawing on Actors Network Theory (ANT) to define any participant – human or nonhuman – whose presence makes a difference in a networked course of action. This allows AI systems, infrastructures, and other influential artefacts to be analysed alongside human decision-makers.

The article is primarily conceptual and lightly illustrative. It refrains from exhaustive excursions into either philosophy of technology or IR theory, and instead stages a balanced meeting point sufficient to make the argument. From STS, the article draws on: (i) Winner and subsequent work arguing that artefacts inherently possess political properties; (ii) a brief engagement with Actor–Network Theory (ANT) and the question of artificial autonomy; and (iii) Don Ihde's postphenomenology for a precise vocabulary of technological mediation (*embodiment*, *hermeneutic*, *alterity*, *background*). A case study of decision-support systems (DSS) illustrates how technology mediates perception, judgement, and action in IR decision-making. Where helpful, the author circles back explicitly to how it is relevant in the context of AI and contemporary international practice.

The aim is to enrich constructivism by integrating insights from technopolitics/STS and postphenomenology; so that constructivist analysis can register both *ascribed meanings* and *the inherent political properties* of increasingly autonomous technologies. The guiding hypothesis is that the example of AI systems as political artefacts both embody values and mediate international practices; to account for this dual role, constructivism requires a mediation-focused vocabulary informed by technopolitics/STS and postphenomenology. Three questions organise the article: (1) How do AI systems, as socio-technical artefacts, embody political and ethical values, and what does this imply for constructivism's emphasis on ascribed meanings? (conceptual clarifications and the politics of artefacts); (2) In what ways can ANT and postphenomenology specify the mediations through which AI shapes international norms and practices?; and (3) What do contemporary cases – like DSS – reveal about the mutual shaping of international order and technology? These questions keep the thread to AI-in-IR explicit, while leaving room for broader lessons about technology in world politics.

The article's scope is intentionally limited, concentrating on the connections between constructivism, technopolitics, and postphenomenology. Other critical frameworks in International Relations and the philosophy of technology – such as practice theory, relational materialism, critical theory, and posthumanist viewpoints – provide further insights into materiality and technology, enhancing the argument considerably. These directions are not analysed here directly due to previously stated limits; nonetheless, they signify an important direction for the continued advancement of the framework.

Conceptual Clarifications: Constructivism, Technopolitics, and IR

One of the pressing challenges in contemporary IR is that traditional theoretical frameworks yet to account for how autonomous/semi-autonomous information systems shape political decisions and operations. Despite the fact that other scholars have observed this phenomenon and the need to introduce comprehensive frameworks (Kiggins, 2025; Estevens, 2024; Eriksson & Newlove-Eriksson, 2021), no such account has yet been identified, particularly those that are connected to postphenomenology to better understand technological mediation. As Kiggins (2025) argues, with the rise of big data and AI, policymakers increasingly rely on computational systems to generate insights and sometimes to *steer choices*. This impacts human-centric decision models by introducing nonhuman actants – algorithms, AI/ML models, and other increasingly smart, self-improving, and autonomous technologies – presence of which makes a difference in decision outcomes. In AI-enabled IR contexts, the consequences are substantially concrete: delegation of evaluative tasks and reduction of decision time coupled with prevailing opacity and accountability drift can all reconfigure the practice of judgement, even when the formal authority remains human. Important to remember that these very capacities pertain positive characteristics that are relevant for the contemporary IR practiced (more efficient processing of information, deep analytics and patterns discovery, and automation of mundane tasks). However, the goal is to introduce the new capabilities of technologies and account for their impact. This is particularly notable for high-stakes and fast-paced environments, such as military context or rapid diplomatic responses.

From a (neo)realist perspective, technology is primarily regarded as a force multiplier and a component of material capabilities. It has the potential to significantly influence military effectiveness, surveillance, and command-and-control (often discussed under "revolutions in military affairs"), and specific technologies – most notably an example of nuclear weapons – reshape deterrence dynamics (Eriksson & Newlove-Eriksson, 2021). However, none of this alters the anarchic, state-centric nature of international politics or the primacy of power balancing. On the other hand, institutional/liberal perspectives prioritize plural actors, interdependence, and the role of institutions in facilitating collaboration. In this context, technology has been interpreted as a catalyst for connectivity and (occasionally) democratization (e.g. early Internet-era arguments), but a subsequent pessimistic shift has emphasized surveillance, vulnera-

bility, and technologically induced risk within networked societies (Eriksson & Newlove-Eriksson, 2021). However, in both paradigms, technology is typically viewed as an instrumental factor – either as a facilitator/constraint on cooperation or as a capability – rather than as constitutive of perception and practice. Nevertheless, another major theory in IR is constructivism, which offers the potential to more effectively analyse the social construction of political phenomena. Despite not assuming complete technological neutrality and prioritising the social construction of identity and meaning, constructivism still lacks a comprehensive vocabulary for identifying and analysing technological mediation and the inherent political properties of technological artefacts.

Below is a simplified table that compares the traditional IR decision-making models to the ones that are currently in effect:

Table 1
Comparison of Traditional IR Decision-Making and modern conditions

Aspect	Traditional IR Conditions	Current IR Conditions
Decision-making Agent	Human policymakers	Hybrid human–machine networks
Analytical Process	Manual analysis, expert judgment	Real-time data processing, algorithmic learning
Adaptability	Slow, often bureaucratic	Dynamic and adaptive (or at least required to be such)
Transparency	Memorable through public debate	Often opaque due to algorithmic complexity
Policy Outcomes	Based on negotiated agreements	Influenced by data-based/algorithmic recommendations

While widely recognised as a mainstream approach since the 1990s, constructivism has long been a subject of debate within IR – as a coherent theory for systemic analysis (e.g., Wendt) *and/or* as an analytical orientation that foregrounds social construction (Reus-Smit, 2022). Either way, constructivism is well-placed to address technology as a political artefact because of its sensitivity to how norms, identities, and practices are constituted. Importantly, constructivism does not presume technological neutrality; rather, it typically focuses on the meanings and identities that actors ascribe to technologies in context. For AI in IR, this means analysing not only the meaning that diplomats, officials, or analysts *ascribe* to AI systems, but also how these systems themselves (thought their dataset design, thresholds, defaults, and interfaces) pre-figure what becomes visible, legible, and actionable in practice.

Against this backdrop, *technopolitics* has been proposed among IR paradigms (rooted in STS) to stress the tight entanglement of technology and

power (Eriksson & Newlove-Eriksson 2021; Schaupp, 2021). Related work (e.g., Kahn & Kellner, 2007; Kellner 2021; Estevens, 2024) emphasises that technologies are woven into economic and political power relations rather than sitting outside them. This is indeed deeply relevant for IR discipline, yet much of the technopolitics literature is still oriented around technology-society power configurations and the politics of innovation, with less sustained attention to the structure of international order *per se*.

> The strength of [the theory of technological politics] point of view is that it takes technical artifacts seriously. Rather than insist that we immediately reduce everything to the interplay of social forces, it suggests that we pay attention to the characteristics of technical objects and the meaning of those characteristics. A necessary complement to, rather than a replacement for, theories of the social determination of technology, this perspective identifies certain technologies as political phenomena in their own right. It points us back, to borrow Edmund Husserl's philosophical injunction, *to the things themselves*" (Winner 1980, p. 123; emphasis original)

For this reason, it is possible to view technopolitics as complementary to IR constructivism (or other major theories) rather than as a competing paradigm; its focus on artefacts' political qualities enhances constructivist analysis of norms and practices. In AI terms, technopolitics helps locate how value-laden architectures (e.g., prioritising speed over contestability, or precision over recall) shape decision environments; constructivism then traces how those environments stabilise or reconfigure in international practices.

The Politics of Artefacts

Among the foundational arguments in the sociology and philosophy of technology is Langdon Winner's statement that artefacts are not neutral instruments but *inherently political*. The argument he is making is not that technical objects independently determine politics; rather, it is that design decisions and system configurations are intrinsically linked to power, inclusion/exclusion, and authority. In Winner's formulation, our habitual way of treating technologies as value-free tools obscures their ability to produce consequences before their deployment:

> To our accustomed way of thinking, technologies are seen as neutral tools that can be used well or poorly [...] But we usually do not stop to inquire whether a given device might have been designed and built in such a way that it produces a set of consequences logically and temporally *prior* to any of its professed uses (Winner, 1980, p. 125; emphasis original).

He further argues that political effects need not follow from conspiracy or malign intent: neglect, routinised assumptions, and institutional convenience are sufficient to instil values into everyday artefacts. In AI-enabled decision-support systems (DSS), for instance, the choice of training data, the calibration of acceptable error rates (including false positives and false negatives), and the prioritisation of certain variables often embed substantive assumptions about the nature of the world (and its inhabitants) and the cases to be evaluated, prior to its deployment. To illustrate on a commonly known example using the mundane ATM, Lucas Introna shows how design presupposes a *type* of user and thereby excludes or marginalises others:

> The ATM bank machine assumes a particular person in front of it... able to see the screen, read it, remember and enter a [PIN] code [...] In this way the ATM embodies a particular understanding of the world (humans) in front of it (Introna, 2018, p. 12).

This example also well illustrates that artefacts carry "scripts" (in the ANT terms) – expectations and action programs – that configure the roles available to humans who meet them. With increasingly complex information systems, the problem deepens since many assumptions are inaccessible to inspection, in some cases even by experts, where:

> Everyday technologies such as utensils are mostly open to scrutiny by the reflexive user. Information technology, on the other hand, is mostly not open to such scrutiny. Important assumptions and biases in information technology are mostly obscured and subsumed in 'black-boxes'[1] in ways that make it difficult even for the experts to scrutinize them. It is not possible for the ordinary computer user to scrutinize the assumptions and biases embedded in the code of the Microsoft Windows operating system (Introna, 2018, p. 13).

For IR contexts, it highly matters due to *the high-stakes nature of decisions*, as more explainable and interpretable qualities are required to ensure meaningfully informed decision-making[2]. Philip Brey has long argued that technological design makes ethics immanent to artefacts rather than an afterthought (Brey, 2000). Existing approaches aim to turn those ethical choices into intentional decisions, by involving frameworks such as ethics-by-design or human-centric approaches[3], among others. This implies identifying that choices about data provenance, labelling (e.g. defining and attributing definitions), model architecture, explanation facilities, and override mechanisms constitute the normative decisions with distributional consequences. In other words, what looks like a purely technical decision may well be a political-ethical settlement about who or what is to be trusted (and who is not), how uncertainty is to be handled, and which errors are tolerable.

Framework: ANT, Artificial Agency, and Postphenomenology

In Actor–Network Theory (ANT), technologies can be identified as actors or *actants*: participants whose presence makes a difference in a course of action; analytically on par with generally perceived humans' autonomy and impact (Latour, 2005). ANT traces associations to show how practices stabilise when heterogeneous elements – people, institutions, artefacts, documents, standards – hold together long enough to act (Latour, 2005). Among central notions in this discussion is *inscription*[4]: values, assumptions, and "programs of action" embedded in design features, procedures, interfaces, and data schemas that prefigure what becomes visible, legible, and actionable (Introna, 2024). While ANT theory is quite extensive, for the present purposes the triad *actants–inscriptions–networks* is sufficient; allowing for examining how increasingly autonomous, self-adjusting and smart technologies – and subsequent Information and Communication Technology (ICT) assemblies – cohere with the practical effects. This provides a foundation for us to treat new types of disruptive and (semi-)autonomous technologies as *artificial agents*[5] – in ANT's idiom, *artificial actants*. Furthermore, Floridi notes that AI represents a "divorce" between intelligence from autonomy and not its union: tasks can be taught and automated that are low in difficulty (*easy*) yet sufficiently complex to provide the "rules" that constitute the process (*complex*), thus addressing the challenge of computing and navigating real-world complexity without granting human-like intelligence status, and enabling delegation of "smart" tasks (Floridi, 2023).

However, despite shared interest in how realities are produced through practice, constructivism and ANT diverge ontologically and epistemologically on aspects like agency, reflection, and the role of values (Jakobsen, 2017). For instance, constructivism treats actors as reflective agents who interpret situations, evaluate alternative possibilities, and act based on value commitments that shape what they consider meaningful and achievable. Here, reality is understood as the outcome of this evaluative relationship between actors and their environment. By contrast, ANT does not primarily account for reflection and values, describing action as an effect of associations in heterogeneous networks of human and nonhuman actants; "possibilities" are not chosen among but generated through network configurations. As a result, ANT offers a powerful account of how artefacts participate in world-making, but provides limited resources for explaining why particular configurations are adopted or contested. For example, two or-

ganizations may have access to the same decision-support system yet enact it differently: a military command may stress speed and automation, while a diplomatic agency may emphasize contestability and accountability. Constructivism explains these differences in terms of value commitments – such as preferences for delegation, transparency, or risk comfort – whereas ANT can describe how each network stabilizes around different elements, but has fewer tools for explaining why these value priorities structure the network in the first place. Both theories have considerable potential to deepen analysis of technological impact in international relations, yet their differences highlight the need for careful conceptual work when integrating them into a unified framework.

When it comes to inclusion of actants (e.g. artificial agents) in IR theorising, Kiggins (without appealing to ANT theory) argues that a compatible heuristic for scholars treating all entities as object-subjects...

> [...] may be that of agent. A sufficiently ambiguous term broad enough to account for state, non-state, and inanimate material objects while simultaneously communicating sufficient specificity regarding the capacity for affecting, effecting, and participating in decision-making under certain conditions. Precisely what International Relations scholars seek to understand, describe, and explain (Kiggins, 2025, pp. 315–316).

Read alongside Floridi, this supports treating AI systems that are delegated decision roles as agents in the analytical sense. As Kiggins (2025) also suggests, "[r]obots, artificial intelligence, and other information technologies to which have been delegated decision-making capability in the policy-making process, must be treated with the same ontological and epistemological status as individuals, states, and other actors" (pp. 313–314). At the same time, the objective is not to elevate autonomous systems to the status of individuals. It is to show that treating systems as actants within a specific analytical framework does not imply attributing political intent or responsibility to them. This allows the analysis to capture how technical systems participate in shaping epistemic processes without diminishing the accountability of human decision-makers.

While ANT helps identify *who and what* participates – along with the *in what ways* networks stabilise around inscriptions – postphenomenology in technology supplies a vocabulary for *how* technologies, once in use, mediate perception, judgement, and action. To be put briefly, phenomenology studies how the world is given in experience; while postphenomenology, from the perspective of philosophy of technology, develops this line by analysing concrete human–technology relations. Peter-Paul Verbeek's (2001) clarifi-

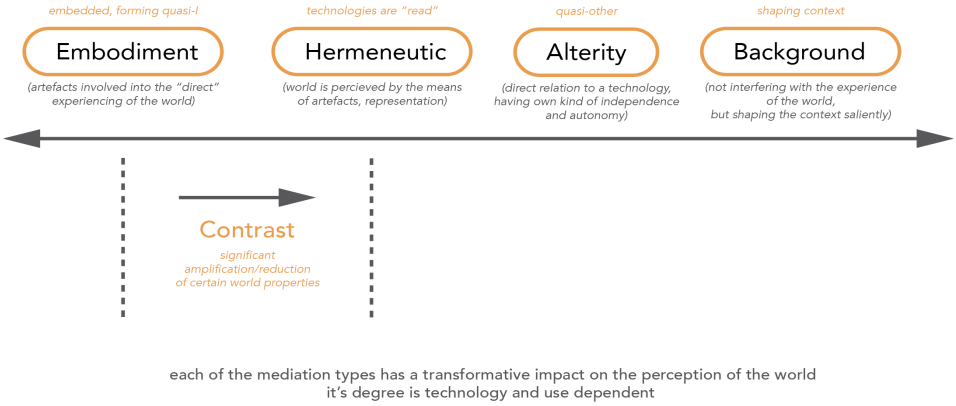
cation helps situate these relations ontologically: mediation should not be seen as something that happens *between* a predefined subject and object; rather, mediating artefacts co-shape both "sides" in the very process of perception and action. On this view, technologies help constitute "who we are" in practice (e.g., how a negotiating team is configured through its tools) and "what the world is" for us (e.g., which events register as salient evidence) (Don Ihde, 1990; Verbeek 2001; Introna 2024). Due to the fact that Heidegger's work dedicated a substantial attention to technology through the lens of phenomenology, this view has occupied an important place in the philosophy of technology (Verbeek, 2001).

Among key figures in philosophy of technology that has developed mediation concepts through postphenomenological understanding is Don Ihde (1990). He accomplished this by foregrounding what people do with particular artefacts, and how these practices co-shape subjectivities and objectivities. Mediation is not simply "in-between", but rather constitutive. Ihde (1990) presents the foundational scheme of mediation *Human–Technology–World* on which other types of mediation introduced later are built upon. To adopt this for IR theorising, the "World" is best specified as the *International*. Here, the *International* refers to the environment in which global political relations unfold, including the structures, practices, and material conditions that shape and are shaped by states, organisations, and other IR participants. Assuming this view ensures adaptability for analysing mediation in diplomacy, security, trade, and compliance or other IR domains. This also provides flexibility in defining what constitutes the International, as different theories and perspectives may include or exclude certain actors, structures, events etc. Thus, I argue that for analysis in IR we can talk of *Human–Technology–International*, bridging the postphenomenological approach to technologies and introducing it into the IR analysis. This would prompt the questions what sort of subjects (whether in STS or IR sense) do we become (e.g. how this changes the ways in which we present ourselves to the world) and what does the IR environment become (e.g. how the context is shaped) with the presence of these mediating technologies?

Don Ihde's postphenomenology identifies four modes of technological mediation (Ihde, 1990, pp. 72–110), each transforming how humans perceive and act in the world (see Figure 1).

Embodiment relations occur when technologies are taken into the body schema, becoming transparent instruments through which the world is perceived (e.g., diplomats using translation earpieces; analysts "seeing through"

Figure 1. The diagram representing the continuum of technological mediation sets, along with an account of contrast based on Don Ihde(1990) and Verbeek's (2001) interpretation of his work



Author's own work.

satellite image enhancement). Hermeneutic relations involve artefacts that provide readable representations of the world, requiring interpretation (e.g., risk dashboards in sanctions compliance; AI-generated conflict maps). As the diagram indicates, embodiment and hermeneutic relations form another continuum structured by contrast: embodiment mediations deviate only slightly from naked perception (low contrast), whereas hermeneutic mediations transform perception more radically by amplifying certain features while reducing others (high contrast). Alterity relations position technologies as quasi-others that respond to or interact with users (e.g., automated visa assessment systems; AI-powered negotiation bots), while background relations describe salient infrastructures that recede from direct attention yet condition our experience (e.g., algorithmic audit trails in compliance regimes; surveillance infrastructures shaping border governance). Important to note, these relations illustrate that mediation always involves degrees of amplification and reduction of certain world properties, while the form of transformation varies: from transparent embodiment to representational hermeneutics, from interactive alterity to contextual background. However, the categorisation of mediation types are simplifications, and technologies often present a mixture of mediation types. This is especially true for more complex ones, like those involving embedded AI into them; for instance, decision-support systems with AI analytics or Augmented Reality (AR) technology in military training contexts. Table 2 below summarises presented mediation types with examples.

Table 2
Summary of the technological mediation, with notes on transformation of perception and IR-relevant examples

Relation	Definition	Contrast (deviation from naked perception)	IR-relevant examples
Embodiment	Technology is incorporated into the body schema, enabling perception *through* it.	**Low contrast:** world looks the same but clarified.	Translation earpieces in diplomacy; enhanced satellite imagery for analysts.
Hermeneutic	Technology provides *representations* of the world that must be interpreted.	**High contrast:** perception is transformed into coded/abstract form.	AI-generated conflict maps; sanctions risk dashboards.
Alterity	Technology encountered as a *quasi-other* that interacts with users.	Not defined by contrast but by interaction stance.	Automated visa assessment systems; consultative systems or chatbots in diplomacy.
Background	Technology recedes from awareness, shaping the *context* of experience.	Not defined by contrast; operates as ambient condition.	Algorithmic audit trails in financial compliance; surveillance infrastructures at borders.

Author's own making.

Contemporary ICT and AI technologies add another layer of complexity. As Floridi notes, much mediation now also occurs in technology-to-technology control and communication, where devices and systems are coordinated autonomously:

> It is not just a technology between us and nature, like an axe (first order), or a technology between us and another technology, like an engine (second order). It is rather a technology between one technology and another technology, like a computerized system controlling a robot painting a car (third order). Due to the autonomous processing power of the digital, we may not even be on – let alone in – the loop (Floridi, 2023, p. 9).

Understanding this "third-order" setting is particularly important in AI control systems within international relations context, as this provides insights into stability and resilience of such networks.

Decision Support Systems: Case Studies

Integration of informational systems into political decision-making is not new. A historical observation illustrated made by Kiggins (2025) analyses the USSR's VRYAN intelligence network that relied on a computing machine to assess the "correlation" of Soviet security relative to the United States, compressing diverse indicators into a single value that shaped high-level perceptions of vulnerability:

> The machine relied on thousands of security and economic data inputs to determine the correlation or, relative security position, of the USSR to the US by assigning a value of 100 to the US. So long as the Soviet Union was at sixty or seventy percent of the US, Soviet leaders were confident that nuclear

> war was avoidable. [...] the VRYAN machine calculated the correlation with the US at forty-five percent sending Soviet leaders into a heightened state of alarm fearing that the USSR had entered a position of vulnerability due to a decrease in its nuclear deterrence capability (pp. 308–309).

In the contemporary settings, advanced ICT are ever more deeply integrated into decision processes and infrastructures. The ways in which data is processed and insights derived with subsequent decision-making and control execution have radically changed, providing an upper hand to those who have the ability to access the advanced technology (and hardware on which it rests upon). All this redefines the IR doctrine, including its conducts, notions of efficiency and human-computer interaction in political contexts. In the following, cases of such systems, called decision support systems (DSS) will be presented and analysed to better understand the impact in practice. The cases presented have an illustrative character to highlight how such systems operate in practice, rather than to provide a full empirical investigation. A deeper analysis of design choices, inscriptions, and implementation controversies would require a substantially larger format and access to primary material. The examples therefore serve to situate the theoretical argument in concrete contexts, while the detailed examination remains a task for future works.

Decision Support Systems in Political and Military Decision-Making: Mediation and Distributed Agency

The integration of DSS into political and military decision-making represents a paradigmatic case of how technologies mediate and transform agency with an impact on international relations, illustrating the shift toward hybrid human–machine decision environments. A decision-support system (DSS) can be broadly understood as a socio-technical arrangement that orchestrates different data sources, processing, analytical models, and human judgment to assist in evaluating options and forming decisions. In international and security contexts, DSS range from risk-assessment tools and intelligence-fusion platforms to strategic-planning interfaces. This type of system are not generally executing decisions itself but rather act as *mediating infrastructures* – artefacts that filter, prioritise, and present information in ways that shape perception, judgement, and action. In this sense, a DSS can be thought of as both an informational instrument and an organisational actor whose inscriptions (e.g. data structures, thresholds, interfaces) contribute to configuring what becomes noticeable and actionable in practice. These systems embed algorithmic pro-

cesses into existing decision structures, providing decision-makers with vast analytical capacity but often at the cost of contextual understanding; thus introducing subtle reconfigurations of responsibility and authority (Csernatoni et al., 2024).

Recently, these systems begun being implemented in both military and civilian contexts, including in initiative and operations of international organisations such as in NATO and the European Union (EU). A NATO special report on AI indicates that it has "the potential to revolutionise warfare" and "promise advancements on decision-making, autonomous systems, and the role of soldiers"; yet, at the same time, "the actual process of incorporating AI into armed forces remains challenging" as the "[p]rocurement processes require adjustments, interoperability across armies cannot be taken for granted, and the prospect of AI-powered weapons raises significant ethical and legal questions" (Clement, 2024, p. ii).

In 2023, NATO's Science and Technology Organization tested a new AI system called ANTICIPE (Augmented Near Real-Time Instrument for Critical Information Processing and Evaluation) to aid decision-making in operational settings (Nadibaidze, et al., 2024). Thales, a French company that is frequently referred to as a leader in advanced technologies and a long-standing French military partner, collaborated on the development of ANTICIPE. This system integrates a war-gaming simulator that ingests inputs from multiple sources – supported by an orchestration layer coordinating sensors, reports, and other data streams – and incorporates machine learning models for pattern recognition and scenario-relevant inference. Its innovative approach breaks down commanders' critical information requirements (CCIRs) into "weak signal" cues and triggers, using automation to detect patterns that human analysts might miss (Petrović et al., 2024). By prioritizing these signals and visualizing threats, the system can alert commanders in near-real-time with recommended responses coded by severity (e.g. preventive action at yellow, mitigation at orange, major decision at red). The ANTICIPE prototype was trialled during NATO's Steadfast Jupiter exercise to examine how an AI tool can accelerate the OODA loop (Observe–Orient–Decide–Act) in command and control (Petrović et al., 2024). This instance serves as an illustration of NATO's efforts to "de-complexify" information overload and aid human operators in making more informed decisions during multi-domain operations.

While the European Union is not a military alliance, it plays a critical role in the coordination of security policy and international diplomacy. The EU has evolved its own Conflict Early Warning System (EWS) over the past decade to facilitate diplomatic decision-making and preven-

tative action. The EU EWS is a real-world DSS risk-scoring tool that was established in 2012 and has been updated over the years. It generates periodic assessments that are reviewed by analysts and inform the European External Action Service (EEAS) in prioritizing conflict prevention efforts with expert qualitative analysis (Bressan, 2024). The system combines open-source indicators (e.g. conflict history, socio-economic data, social media trends) with inputs from member states' intelligence and on-the-ground diplomatic reports to flag countries at risk of violence or instability (De Agostini & Giovanardi, 2025). Experts in the past have referenced the EU's early warning system as a model for integrating AI-driven forecasting with human judgment (Deen et al., 2021). This system employs algorithmic models, such as machine-learning-based risk indexes, to identify escalation signals, but it then augments these signals with political analysis and field expertise before initiating any policy response. The EWS has contributed to EU decision-making in foreign policy by identifying emerging crises (for example, signs of destabilization in countries like Mali or Sudan) early enough to plan mediation or humanitarian interventions (Bressan, 2024).

At the same time, research indicates that the willingness to use Decision Support Systems (DSS) depends not only on their technical performance but also on how they are perceived by human operators. As illustrated in the model presented by the Göhler Bornsch et al. (2024) both deductive and inductive factors influence user attitudes through the mediating variables of trust and acceptance. Deductive factors such as resilience, speed, explainability, and transparency reflect system design qualities theoretically associated with reliability and ethical accountability. Inductive factors – traceability, controllability, experience with the system, and usefulness – derive from empirical observations of user interaction and technology acceptance. These variables influence users' assurance that a DSS is reliable, interpretable, and responsive to human oversight. Trust and acceptance thus act as essential components: even a highly capable system will not be adopted if it lacks explainability or control transparency while the user requires it. This framework is consistent with the recognised models of human–AI interaction and technology acceptance, indicating that users' tolerance for DSS is contingent upon a balance between their perceived agency, transparency / understanding and the technical performance of the system. The model suggests that in order for AI-enabled DSS to be adopted in military decision-making, developers and commanders must prioritize the design of features that promote interpretability, user trust, and perceived usefulness, in addition to accuracy and speed.

From the perspective of international relations theory, Decision Support Systems (DSS) exemplify how technological mediation reorganizes both power and knowledge within global governance structures. As Raluca Csernatoni et al. (2024) argue, the rise of AI-enabled infrastructures has produced a diffuse assemblage of state, military, and corporate actors who collectively exercise forms of algorithmic governance that transcend traditional borders and hierarchies. In this sense, AI-driven DSS embody a reconfiguration of authority in which decision-making is distributed across technical systems and institutional layers, rather than remaining exclusively within sovereign political domains. This development complicates classical IR conceptions of sovereignty, control, and rational agency where decisions once grounded in deliberative political human negotiations are now partially automated, data-driven, and mediated through digital architectures maintained by a hybrid network of governmental, industrial, and research entities (Csernatoni et al., 2024; De Agostini & Giovanardi, 2025).

Moreover, scholars increasingly warn that the spatial and cognitive distancing produced by such technologies risks normalizing forms of remote governance and depersonalized control. As Johan Eriksson and Lindy M. Newlove-Eriksson (2021) and Fabio Cristiano et al. (2023) note, the transition from battlefield to control room replaces direct human engagement with the interface of automation, potentially eroding moral responsibility and situational awareness. The broader implication is that technological mediation – while enhancing resilience, speed, and transparency – introduces a profound *reordering of decision authority* that demands renewed theoretical and normative scrutiny within international relations.

Implications for the Constructivist Theory of International Relations

The agent–structure concept in International Relations requires revisiting to provide analytic space for increasingly ubiquitous incorporations of (semi-)autonomous, self-learning and advanced informational systems, like artificial intelligence, that constitute to decision-making (whether coupled with human or autonomously), information mediation, and subsequent outcomes. In this context, Actor–Network Theory (ANT) provides a vocabulary that allows to capture AI systems – including their data orchestration, interfaces, along with organizational procedures – as actants whose inscriptions assist in the stabilization of certain practices by determining what is recognised, visible and actionable. The treatment of these artifacts as par-

ticipants does not anthropomorphize them; rather, the goal is to emphasize the fact that design decisions precede any formal policy decision by establishing "programs of action" provided by the technological affordances with distributive repercussions. In short, artefacts are "active" not only in use and the meanings ascribed to them by the stakeholders, but also as actants that steer certain choices and decisions.

As previously mentioned, constructivism can integrate the insights from technopolitics by analysing the embedding of norms and power in the design and technical structure of artifacts; having an impact on how certain structures and powers form with social consequences. Winner's (1980) appeal to "pay attention to the characteristics of technical objects and the meaning of those characteristics" resonates with this view on how norms and identities are constituted (p. 123). Technopolitics, while cannot be recognised as a substantial IR framework, provide tools and vocabulary to analyse how value-laden technological design architectures influence experiences and environments, while constructivism explains how these environments subsequently stabilize and gain legitimacy in international political practice.

Postphenomenology specifies *how* mediation proceeds once systems are in use; earlier introduced framework include embodiment (becoming transparent instruments), hermeneutics (producing representations that must be read), alterity (interacting as quasi-others), and background (conditioning action from the periphery). By acknowledging the mediation types that arise in practice, it allows to better trace and understand their impact on our perceptions, interactions, and actions in the world. Furthermore, this allows for the analysis of how the same system can simultaneously amplify and reduce various aspects of the world, thereby redistributing epistemic authority and responsibility. The adoption of a *Human–Technology–International* schema places these relations in the context of IR, diplomatic, security, and governance settings. The concept of International provides enough flexibility to be adapted and interpreted by the various IR theories according to their base ontology and epistemology.

Finally, the field needs a permissive ontology for nonhuman participation. As Kiggins (2025) warns, IR's human-centric theories risk obsolescence as automated and data-driven decision-making involving increasingly autonomous systems expands:

> The agent/structure ideal type from the social constructivist tradition of IR theory may be best suited *to account for the potential agentive capacity to act autonomously inherent in artificial and other information intelligence technologies* [...] as reliance on automated decision-making increases apace with the development of more capable artificial intelligence technology, the risk to IR is

> that some theories on which we rely to explain human decision-making under certain conditions may become obsolete [...] *artificial intelligence technology is capable of agentive capacity and must now be treated as an agent capable of effectuating outcomes in international relations.* (p. 298, emphasis added).

This call also provides a ground for further discussions and analyses regarding how the advancement of major IR theories can happen to ensure the explainability capabilities are continuously in accordance with the changing world.

Conclusion

This article presents ideas of constructivist theory advancement by involving frameworks from technopolitics, ANT and postphenomenology to account for increasingly autonomous, self-learning and data-driven system such as AI in international relations. Conceptually, it brings together ANT's emphasis on artificial actants (e.g. agents) and inscriptions with postphenomenology's account of technological mediation to show how advanced ICT systems such as AI can co-shape perception, judgment, and action in world politics. To account for the inherently political properties of technological artifacts that may preceed their usage (i.e., political in themselves due to the design properties), this perspective is proposed against the background of STS, technopolitics and sociology of technology. These dynamics are analysed by decision-support systems (DSS) case studies in NATO and EU.

The conclusion circles back to the initial research question presented at the beginning of this article: (1) AI systems embody political and ethical values through design decisions that precede policy, which invites constructivists to analyse artefacts as *sites of norm production*, not just objects of meaning; (2) ANT and postphenomenology together specify where and how to study mediation, following the actants and inscriptions concepts, then address these mediation modes – embodiment, hermeneutic, alterity, background – to recognise how political practices are affected and reshaped; (3) Contemporary DSS cases show mutual influence between order and technology: international routines incentivise certain designs while those designs stabilize new expectations about attention, expertise, and responsibility in diplomacy and security. Incorporating these insights equips constructivism and other IR theories to register both ascribed meanings and the inherent political properties of increasingly autonomous systems, without treating technology as either neutral instrument or singular determinant.

Summary

This article develops a mediation-focused constructivist account of artificial intelligence (AI) as a political artefact in international relations. Bringing STS, technopolitics and postphenomenology into dialogue with IR, it argues that AI systems should be analysed not only for the meanings actors ascribe to them but also for how design choices – data provenance, model thresholds, interface defaults – materialize normative commitments that preconfigure perception, judgment, and action. Actor–Network Theory clarifies how heterogeneous participants (humans, organisations, datasets, models, interfaces) hold together through inscriptions, while postphenomenology specifies the modes of mediation (embodiment, hermeneutic, alterity, background) through which autonomous systems like AI reshapes decision environments. The article is primarily conceptual, supported by illustrative evidence from decision-support systems (DSS) in NATO and the European Union, AI-enabled tools for operational command and conflict early warning. These cases reflect how DSS redistribute epistemic authority and responsibility across technical systems and institutional layers, accelerate sense-making under time pressure, while highlighting risks of opacity and accountability drift if not designed and governed with care. The contribution is as follows: a permissive ontology for treating (semi-)autonomous systems as agents in the analytical sense without assuming personhood, and a vocabulary and analytical frameworks for tracing inscriptions and mediation in practice. The framework contributes to the gap in IR by facilitating the development of comprehensive schemes that account for advanced technological properties. In addition, it contributes to the understanding of technological mediation in IR through the lens of postphenomenology.

N O T E S

[1] A black-box system is one whose internal workings cannot be meaningfully inspected or understood by those who use or deploy it, even though its designers or developers may (in principle) have access to that understanding. In high-stakes decision contexts, this opacity matters: without insight into how a system produces outputs, it is difficult to judge when the system performance may fall short or deviate due to its properties, and therefore difficult to assign responsibility, contest decisions, or improve practice.

[2] *Explainable AI* (*XAI*) refers to methods that generate *post-hoc* explanations for a model's outputs – e.g., highlighting which inputs influenced a particular decision, or producing human-readable rationales – without necessarily making the model itself simple or transparent. By contrast, *interpretable AI* denotes models whose internal structure is sufficiently transparent that a knowledgeable observer can follow the steps from input to output (for example, models that are inherently understandable because of their form or constraints).

[3] Ethics-by-design or human-centric approaches start from the premise that ethical and human factors are not external add-ons but integral to technical work. In practice, they require identifying which human values and interests are at stake, how design choices can affect those values (e.g., fairness, transparency, safety, contestability), and how such considerations should be built into requirements, data practices, model choices, interfaces, and governance throughout the system's life cycle.

[4] For a practical example, please refer to pp. 7–8 for the ATM illustration and analysis.

[5] To clarify the status of technological participants, I use a widely cited minimalist account from computer science and multi-agent systems. On this view, a system counts as an agent if it (a) receives and uses data from its environment; (b) takes actions on that basis, without needing step-by-step human control, in pursuit of specified goals; and (c) improves its performance by learning from interactions with the environment (Floridi 2023: 10). This analytical usage does not imply personhood or moral responsibility; it serves to mark systems that act adaptively and autonomously within defined task domains.

BIBLIOGRAPHY

Bressan, S. (2024). *The power and limits of data for peace: Improving the EU Conflict Early Warning System for more effective prevention* [Brief]. Paris: European Union Institute for Security Studies. Retrieved October 15, 2025, from https://www.iss.europa.eu/publications/briefs/power-and-limits-data-peace

Brey, P. (2000). Disclosive computer ethics. *Computers and Society, 30*(4), 10–16.

Clement, S. (2024). *NATO and artificial intelligence: Navigating the challenges and opportunities.* Luxembourg: NATO Parliamentary Assembly, Science and Technology Committee. Retrieved from https://www.nato-pa.int/document/2024-nato-and-ai-report-clement-058-stc

Cristiano, F., et al. (2023). *Artificial intelligence and international conflict in cyberspace.* Oxon & New York: Routledge.

Csernatoni, R., Broeders, D., Andersen, L. H., et al. (2025). Myth, power, and agency: Rethinking artificial intelligence, geopolitics and war. *Minds & Machines, 35*(1), 1–26. https://doi.org/10.1007/s11023-025-09741-0

De Agostini, L., & Giovanardi, M. (2025). *AI and global security: From early warning to AI-assisted diplomacy* (Global Peace and Security Policy Brief). Think7 Canada. Retrieved October 15, 2025, from https://www.think7.org/publications/ai-and-global-security-from-early-warning-to-ai-assisted-diplomacy/

Deen, B., Stoetman, A., & de Bruijne, K. (2021). *From indices to insight: A proposal to enhance the risk assessment of the Dutch early warning/early action process* (Clingendael Report, November 2021). The Hague: Clingendael Institute. Retrieved from https://www.clingendael.org/sites/default/files/2021-11/Report_Early_warning_early_action_Nov_2021.pdf

Estevens, J. (2024). International relations theories and the impact of the Fourth Industrial Revolution on global affairs. In D. Hernández Martínez & J. M. Calvillo Cisneros (Eds.), *International relations and technological revolution 4.0* (pp. 9–24). Cham: Springer. https://doi.org/10.1007/978-3-031-66750-3_2

Göhler Bornsch, G.-F., Bornschegl, O., & Tzschoppe, S. (2024). *Factors affecting the willingness to use decision support systems in a military context* [Presentation]. NATO Science and Technology Organisation, System Analysis and Studies Panel. https://doi.org/10.14339/MP-SAS-OCS-ORA-2023-11P-PDF. Retrieved October 15, 2025, from https://www.sto.nato.int/document/factors-affecting-the-willingness-to-use-decision-support-systems-in-a-military-context-presentation/

Ihde, D. (1990). *Technology and the lifeworld: From garden to Earth.* Bloomington & Indianapolis: Indiana University Press.

Introna, L. (2024). Ethics and information technology: Phenomenology. In *Stanford Encyclopedia of Philosophy.* Retrieved October 10, 2025, from https://plato.stanford.edu/entries/ethics-it-phenomenology/

Jakobsen, M. (2017). A comparison of pragmatic constructivism and actor network theory. *A Philosophy of Management Accounting*, 245–259. https://doi.org/10.4324/9781315680736-12

Kahn, R., & Kellner, D. (2007). Globalization, technopolitics, and radical democracy. In L. Dahlberg & E. Siapera (Eds.), *Radical democracy and the Internet* (pp. 17–36). London: Palgrave Macmillan. https://doi.org/10.1057/9780230592469_2

Kellner, D. (2021). Globalization, technopolitics and revolution. In *Technology and democracy: Toward a critical theory of digital technologies, technopolitics, and technocapitalism* (pp. 153–177). Wiesbaden: Springer VS. https://doi.org/10.1007/978-3-658-31790-4_6

Kiggins, R. D. (2025). Big data, artificial intelligence, and autonomous policy decision-making: A crisis in international relations theory? In R. D. Kiggins (Ed.), *The political economy of robots* (pp. 297–327). Cham: Palgrave Macmillan. https://doi.org/10.1007/978-3-031-95171-8_13

Nadibaidze, A., Bode, I., & Zhang, Q. (2024). *AI in military decision support systems: A review of developments and debates.* Odense: Center for War Studies, University of Southern Denmark.

Passi, S., & Vorvoreanu, M. (2022*). Overreliance on AI: Literature Review.* Microsoft Aether: AI Ethics and Effects in Engineering and Research. Microsoft Corporation. https://www.microsoft.com/en-us/research/wp-content/uploads/2022/06/Aether-Overreliance-on-AI-Review-Final-6.21.22.pdf

Petrović, P., Tavares Martins Dias, M., Grasso, N., Vicente Botella Berenguer, V., & Denieul, M. (2024). *The role of AI decision-making for land-based operations* (Food for Thought 11–2024). Brussels: FINABEL – European Land Force Commanders Organisation.

Reus-Smit, C. (2022). Constructivism. In R. Devetak & J. True (Eds.), *Theories of international relations* (6th ed., pp. 387–426). London: Bloomsbury Publishing.

Schaupp, S. Technopolitics from Below: A Framework for the Analysis of Digital Politics of Production. Nanoethics 15, 71–86 (2021). https://doi.org/10.1007/s11569-021-00386-8

Verbeek, P.-P. (2001). Don Ihde: The technological lifeworld. In H. Achterhuis (Ed.), *American philosophy of technology: The empirical turn* (pp. 119–146). Bloomington: Indiana University Press.

Winner, L. (1980). Do artifacts have politics? *Daedalus, 109*(1), 121–136.