*Corresponding Author's Email: krbilgin@baskent.edu.tr
Proceedings of the World Conference on Social Sciences and Humanities
Vol. 3, Issue. 1, 2026, pp. 18-25
DOI: https://doi.org/10.33422/shconf.v3i1.2047

ISSN: 3030-0819 online

Diamond Scientific Publishing

Legitimacy Criteria for Strategic Communication in the Age of Artificial Intelligence

Kıvılcım Romya Bilgin

Başkent University, Faculty of Communication, Ankara, Türkiye

Abstract

Traditionally, strategic communication has focused on meaning-making, agenda-setting, persuasion, and reputation management. The rise of artificial intelligence (AI), especially machine-learning based profiling and prediction systems, has transformed this field radically. Communication is no longer limited to the design of a message, but increasingly takes the form of an algorithmically structured architecture of behavioral influence based on a data-driven cycle of collection, profiling, personalization, and optimization. These processes increase precision and effectiveness, but they also raise serious problems of legitimacy. The undefined and unregulated scope and limits of AI-enabled influence may lead strategic communication to move from persuasion to opaque manipulation. Here, legitimacy is used as the central analytical framework to argue that AI-powered strategic communication requires a clearly articulated normative boundary. Legitimacy is viewed as a multi-dimensional construct, encompassing legal conformity, transparency, accountability, proportionality, and the protection of fundamental rights. Drawing on three major European regulatory frameworks (the EU General Data Protection Regulation (GDPR), the Artificial Intelligence Act, and the Digital Services Act), this paper develops a set of evaluative criteria to assess the legitimacy of AI-based strategic communication processes. The study synthesizes regulatory principles and communication theory to distinguish between legitimate strategic influence and manipulative algorithmic intervention in democratic governance structures.

Keywords: algorithmic governance, democratic accountability, fundamental rights, strategic communication, legitimacy

1. Introduction

The digitization of public life has created a new type of communicative actor. The algorithmic system generates, filters, targets, and optimizes messages at a scale and speed no human communicator can rival. As a result, strategic communication, or the purposeful use of communication by an organization to fulfill its mission (Hallahan et al., 2007), has undergone a structural transformation. Where a strategist once wrote a message for a general audience,

today's AI-enabled practitioner uses thousands of micro-targeted variants calibrated to individuals' inferred psychological profiles and behavioral histories.

This shift carries profound normative implications. When communication is designed not merely to inform or persuade but to preemptively shape the informational environment in which decisions are made, the boundary between influence and manipulation becomes contested. Democracies presuppose informed citizens capable of autonomous deliberation, and algorithmic architectures of behavioral influence, deployed without transparency or accountability, may systematically undermine this precondition (Zuboff, 2019). This paper responds by developing a legitimacy-centered analytical framework for AI-driven strategic communication. Legitimacy is understood not merely as legal compliance but as a broader normative condition encompassing democratic acceptability and ethical justifiability (Suchman, 1995; Tyler, 2006) and it provides a stable foundation for assessing whether a given use of AI in strategic communication is appropriate within a democratic constitutional order. The study establishes legitimacy as the appropriate normative standard for this domain, operationalizes it as a five-dimensional construct and it translates emerging European regulatory principles into actionable evaluative criteria. The central research question guiding this study is: under what conditions can AI-driven strategic communication be considered legitimate within democratic governance structures, and how can legitimate influence be distinguished from manipulative algorithmic intervention? In scope, this is a conceptual framework paper. It proceeds through normative theory-building and doctrinal analysis of the three regulatory instruments and does not present new empirical data.

2. Literature Review

In the early 2000s, strategic communication emerged as a separate field, merging public relations, political communication, organizational communication, and military traditions (Hallahan et al., 2007). The integration of AI demands a more fundamental re-conceptualization: algorithms are taking up not only the distribution, but more and more the production of strategic messages. What does agenda-setting mean when the agenda is set by a recommendation algorithm? Studies of computational propaganda have shown how automation and algorithmic amplification can be weaponized against democratic discourse (Woolley & Howard, 2018). Recent empirical work has begun to test the assumptions underlying these concerns: large-scale survey experiments indicate that microtargeting can confer a measurable persuasive advantage over non-targeted messaging, although the advantage is context-dependent (Tappin et al., 2023; Zarouali et al., 2022). Beyond the European Union, scholarship on platform governance has traced how states with divergent political systems are converging on the regulation of platform power, underscoring that the legitimacy questions raised here are global rather than exclusively European (Flew, 2021).

Missing normative vocabulary is provided by the legitimacy theory. In organizational sociology, legitimacy refers to the degree to which an organization's actions are seen as appropriate, proper, and desirable by relevant audiences; Suchman (1995) identifies pragmatic, moral, and cognitive varieties. This needs to be complemented by political theory for AI-driven strategic communication. Democratic legitimacy in the Habermasian tradition demands accountability of practices to open, inclusive public deliberation (Habermas, 1996). Tyler's (2006) social-psychological account demonstrates that citizens accept institutional authority when procedures are fair, transparent, and respectful of their dignity. Legitimacy is thus a property of processes as well as outcomes: was the individual informed, was their data used with knowledge and consent? While recent scholarship has applied concepts of legitimacy to

AI systems in general (Diakopoulos, 2016; Floridi et al., 2018), no comprehensive framework has been tailored to strategic communication in AI-mediated environments.

3. Theoretical Framework: Legitimacy as a Multidimensional Construct

In this field, legitimacy is not a binary property, but a multidimensional and graduated condition. A practice can be legally acceptable but ethically unjustifiable, or technically transparent but democratically unacceptable. This paper suggests five conceptually distinct but practically connected dimensions. The threshold condition is legal conformity: a practice that is illegal under the relevant law cannot be legitimated. But law is a lagging indicator of normative development. Before any legal violation was found, the Cambridge Analytica affair was already condemned as a threat to democratic integrity (Cadwalladr & Graham-Harrison, 2018). Legal conformity is thus a necessary but not a sufficient condition. Transparency means that people need access to accurate and intelligible information about the systems that impact them – such as disclosure of the use of AI, explanation of the logic of profiling and targeting, and auditability by independent regulators and researchers. The framework is based on a "thick" notion of transparency where the simple publication of information may meet formal requirements but may not contribute to meaningful accountability (Fung et al., 2007). Recent work on "transparency by design" likewise argues that such requirements must be built into AI systems from the outset rather than retrofitted after deployment (Felzmann et al., 2020). Accountability means that responsible actors are identifiable and answerable. In complex AI systems, responsibility is shared among data providers, model developers, platform operators and end-users – the "accountability gap" of AI governance (Mittelstadt et al., 2016). Legitimacy requires clear assignment of responsibility, meaningful human oversight and effective redress. Proportionality, a core principle of European fundamental rights law, places a structural limit on "influence inflation": the extent to which conduct may be influenced must be appropriate to the communicative aim and no more than is necessary. Finally, protection of fundamental rights is a substantive constraint on the other four dimensions, namely privacy and data protection, freedom of expression and thought, non-discrimination and cognitive liberty, that is, the right to form beliefs and make decisions free from external manipulation (Bublitz, 2013; Farahany, 2023).

A central theoretical challenge is to delineate legitimate influence from manipulation. Susser et al. (2019) define manipulation as influence that bypasses or subverts the rational agency of its targets, rather than engaging it. It does not work by giving reasons but by exploiting psychological vulnerabilities, manufacturing false beliefs, or creating informational asymmetries that rule out autonomous choice. Concrete manipulative modalities include dark patterns that extract non-voluntary "consent," micro-targeting exploiting emotional vulnerabilities, synthetic media, and behavioral prediction models timed to moments of maximum susceptibility. Experimental research confirms that such interface designs measurably steer users into choices they would not otherwise make (Luguri & Strahilevitz, 2021). By contrast, legitimate influence is a matter of rational agency. It provides accurate information, real reasons and respects the audience's right to reach different conclusions. Importantly, the legitimacy of a communicative practice does not depend on the empirical success of the attempted influence. The behavioral effectiveness of micro-targeting is empirically disputed (Tappin et al., 2023), but a manipulative design is illegitimate even if it fails to have the intended effect.

4. Regulatory Analysis: The European Framework

The European Union has created the world's most comprehensive regulatory framework for AI-mediated communication – a layered architecture where the GDPR, the AI Act and the DSA operationalise the value commitment of Article 2 TEU.

4.1. The General Data Protection Regulation (GDPR)

The processing of personal data under Regulation 2016/679 (GDPR; European Parliament and Council, 2016) shall be lawful, fair and transparent and subject to the principles of purpose limitation and data minimisation. The most important provisions for AI-driven strategic communication are those concerning profiling and automated decision making. Under Article 22, data subjects have the right not to be subject to a decision based solely on automated processing that produces legal effects or similarly significant effects. Profiling – the core mechanism of algorithmic targeting – is subject to a lawful basis under Article 6, but consent obtained through dark patterns or denial of service is not real consent; and where processing relates to political opinions (Article 9), explicit consent is required, a provision of particular importance for political micro-targeting.

But in practice the exercise of such rights is systematically underperformed. So formal rights are not sufficient to secure legitimacy. But two elements of algorithmic profiling test the protective scope of Article 9. First, special category characteristics such as political opinion or health status are increasingly inferred from ordinary behavioral data rather than collected directly. Second, the exception for data "manifestly made public" by the data subject (Article 9(2)(e)) is invoked to justify large-scale scraping of public social media content, even if making a post public is not the same as consenting to its use for psychographic profiling.

4.2. The Artificial Intelligence Act (AI Act)

The AI Act (Regulation 2024/1689; European Parliament and Council, 2024) sets out the first ever comprehensive legal framework for AI systems, based on a risk-based approach, from prohibited practices to minimal-risk systems. Article 5 prohibits AI systems that use subliminal techniques to materially distort a person's behavior in a manner that is likely to cause harm or systems that exploit vulnerabilities related to age, disability or social and economic situation. This ban provides legal expression to the influence–manipulation divide: influence mediated by AI that is below the level of conscious awareness and causes material behavioral change crosses the line into forbidden manipulation. With the transparency obligations under the Act, the transparency dimension has come into positive law.

Chapter V of the Act extends these obligations up the value chain, to providers of general-purpose AI models that increasingly generate the content of strategic communication itself. Providers shall keep technical documentation and provide downstream deployers with the information necessary to understand a model's capabilities and limitations (Article 53), and models identified as posing systemic risk shall be subject to further obligations regarding risk assessment, mitigation and incident reporting (Article 55). This value-chain allocation of responsibility speaks directly to the accountability gap identified in Section 3: legitimacy obligations attach not only to the communicator who deploys an AI-generated message but also to the actors who build and supply the underlying models.

4.3. The Digital Services Act (DSA)

The DSA (Regulation 2022/2065; European Parliament and Council, 2022) regulates online platforms as intermediaries in the distribution of information and influence. Article 27 requires platforms to disclose the parameters of their recommendation systems, Article 38 requires very

large online platforms (VLOPs) to provide recommendations that are not based on profiling and Article 40 provides for access to platform data for vetted researchers. Most importantly, Article 34 obligates VLOPs to carry out annual assessments of systemic risks, including negative effects on fundamental rights, civic discourse, and electoral processes – a formal requirement to consider the democratic implications of algorithmic influence architectures that is not clearly mirrored in regulatory precedent. The DSA also operationalises transparency and redress at the level of individual users.

Articles 26 and 39 require that advertisements be labelled as such, that users are informed who paid for them and how they were targeted – with targeting based on special categories of data prohibited outright – and that VLOPs maintain publicly searchable advertisement repositories, turning persuasion itself into an auditable activity. This is complemented by the statement-of-reasons requirement (Article 17), internal complaint-handling (Article 20), out-of-court dispute settlement (Article 21) and the right to lodge complaints with Digital Services Coordinators (Article 53), which offer the tangible redress chain required by the accountability criterion.

5. Discussion: Evaluative Criteria and Implications

Synthesizing the theoretical framework and the regulatory analysis yields five evaluative criteria, summarized in Table 1. Each must be satisfied substantively rather than formally, and the fifth operates as an overriding constraint on the other four.

Table 1. Legitimacy criteria and example audit indicators for AI-driven strategic communication

Criterion	Core requirement	Primary regulatory anchor	Example audit indicators
Legal conformity	Substantive compliance with all applicable law	GDPR Arts. 5–9; AI Act; DSA	Documented lawful basis; records of processing; data protection impact assessment
Transparency	Disclosure, explanation, and auditability ("thick" standard)	GDPR Arts. 13–15; AI Act Art. 50; DSA Arts. 26–27, 39–40	AI-use disclosure present; advertisement labels and repository entries; researcher data access
Accountability	Identifiable responsible actors, human oversight, redress	GDPR Art. 22; AI Act Arts. 53, 55; DSA Arts. 20–21	Named responsible owner; human-oversight procedure; functioning complaint and redress channels
Proportionality	Influence commensurate with objective; least intrusive means	EU Charter; AI Act Art. 5	Documented necessity assessment; least-intrusive-means justification
Fundamental rights	No threat to privacy, expression, thought, non-discrimination, cognitive liberty	EU Charter Arts. 7, 8, 10, 11, 21	Fundamental rights impact assessment; no special-category targeting

Source: Author's elaboration

Two clarifications concern the framework's operationalization. First, proportionality – originally a doctrine of constitutional adjudication – translates to communication contexts as a three-step test: is the influence technique suitable for a legitimate communicative objective; is it the least intrusive means available (for instance, contextual rather than psychographic targeting); and are its expected benefits proportionate to the intrusion into the audience's decisional environment? Second, each criterion can be assessed through observable audit indicators, examples of which are listed in Table 1: the presence and quality of disclosures, the

existence of a designated accountable owner and functioning complaint channels, documented necessity assessments, and completed fundamental rights impact assessments. These indicators are indicative rather than exhaustive; their purpose is to translate the criteria into a practicable checklist for internal governance and external audit.

The proposed framework is meant to be not only a normative reference point for scholars, but also a practical assessment tool for organizations implementing AI in communication processes. This operationalization of legitimacy into observable dimensions allows communication scholars and policymakers to evaluate whether AI-supported communication practices align with democratic values, institutional accountability, and fundamental rights throughout the communication process.

A brief vignette illustrates how the framework operates in practice. Consider a political campaign that deploys a generative AI chatbot to hold personalized persuasion conversations with voters identified through behavioral profiling. Legal conformity requires a lawful basis for the profiling and, where political opinions are inferred, explicit consent (GDPR Articles 6 and 9). Transparency requires that voters be informed they are conversing with an AI system (AI Act Article 50) and, where the conversations are delivered as platform advertising, that the sponsor and the targeting logic be disclosed (DSA Article 26). Accountability requires identifiable responsible actors along the whole value chain – campaign, platform, and model provider alike (AI Act Articles 53 and 55). Proportionality asks whether persuasion calibrated to individual psychology is necessary for the campaign's legitimate objective, or whether less intrusive messaging would suffice. Finally, if the chatbot times its appeals to moments of inferred emotional vulnerability, it crosses from persuasion into manipulation prohibited by Article 5 of the AI Act. The vignette shows that the criteria operate jointly: a deployment may satisfy some and fail others, and failure on fundamental rights is disqualifying.

Applying the framework involves a trade-off: transparency may be in conflict with commercial confidentiality, proportionality with effectiveness. In fact, these tensions are conflicts of interest between communicators, platforms, regulators, and audiences; the framework cannot resolve these conflicts, but it does provide a shared normative language for managing them. Three implications for practitioners are to build internal AI governance structures to assess legitimacy before deployment, to invest in comprehensible transparency infrastructure for genuine informed consent, and to engage the European regulatory framework as an opportunity, because legitimate AI communication practices build the trust essential for long-term effectiveness.

6. Conclusion

In this paper we have developed a legitimacy-focused framework for AI-powered strategic communication and applied it to the main European regulatory instruments. The analysis shows that the EU's integrated framework provides a solid legal basis for each of the five dimensions. The GDPR anchors legal compliance and rights protection. The AI Act anchors proportionality and accountability. The DSA anchors transparency and democratic accountability in the platform architecture. The influence–manipulation distinction developed here as a theoretical construct has thus taken on binding legal form in Article 5 of the AI Act. The evaluative criteria summarized in Table 1 translate this foundation into an applicable assessment tool for researchers, regulators and practitioners.

A principal limitation of the framework is its EU-centrism: the criteria are anchored in a specific regulatory architecture and in European fundamental rights doctrine. The underlying dimensions nonetheless travel. Legal conformity, transparency, accountability, proportionality

and rights protection can be mapped onto other regimes – for instance, consumer protection enforcement against manipulative interface design in the United States, emerging data protection statutes in other regions, or algorithm-related platform rules in non-European jurisdictions – although the strength of each regulatory anchor will vary with local law (Flew, 2021). Testing this portability is a task for comparative research.

Future research could apply the proposed framework to empirical case studies of governments, international organizations, corporations or digital platforms to assess how legitimacy challenges arise in different communication contexts. Comparative analyses of regulatory environments outside the European Union may also contribute to refine and validate the proposed legitimacy criteria. Electoral, public health and crisis communication studies in specific domains, as well as empirical evaluations of whether the regulatory framework delivers legitimate outcomes in practice would further enrich the framework's utility.

More broadly, the stakes are not only academic but political: the conditions of legitimate AI-mediated strategic communication are conditions for the survival of democratic public discourse in an age of algorithmic mediation. This study argues that legitimacy should not be considered a secondary ethical consideration, but a core strategic requirement of AI-enabled communication in democratic societies. These conditions are articulable, operationalizable, and enforceable by sustained institutional commitment.

References

Bublitz, J. C. (2013). My mind is mine!? Cognitive liberty as a legal concept. In E. Hildt & A. G. Franke (Eds.), *Cognitive enhancement: An interdisciplinary perspective* (pp. 233–264). Springer. https://doi.org/10.1007/978-94-007-6253-4_19

Cadwalladr, C., & Graham-Harrison, E. (2018, March 17). Revealed: 50 million Facebook profiles harvested for Cambridge Analytica in major data breach. *The Guardian*.

Diakopoulos, N. (2016). Accountability in algorithmic decision making. *Communications of the ACM, 59*(2), 56–62. https://doi.org/10.1145/2844110

European Parliament and Council. (2016). Regulation (EU) 2016/679 on the protection of natural persons with regard to the processing of personal data (General Data Protection Regulation). *Official Journal of the European Union*, L 119, 1–88.

European Parliament and Council. (2022). Regulation (EU) 2022/2065 on a Single Market for Digital Services (Digital Services Act). *Official Journal of the European Union*, L 277, 1–102.

European Parliament and Council. (2024). Regulation (EU) 2024/1689 laying down harmonised rules on artificial intelligence (Artificial Intelligence Act). *Official Journal of the European Union*, L 2024/1689.

Farahany, N. A. (2023). *The battle for your brain: Defending the right to think freely in the age of neurotechnology*. St. Martin's Press.

Felzmann, H., Fosch-Villaronga, E., Lutz, C., & Tamò-Larrieux, A. (2020). Towards transparency by design for artificial intelligence. *Science and Engineering Ethics, 26*(6), 3333–3361. https://doi.org/10.1007/s11948-020-00276-4

Flew, T. (2021). *Regulating platforms*. Polity Press.

Floridi, L., Cowls, J., Beltrametti, M., Chatila, R., Chazerand, P., Dignum, V., Luetge, C., Madelin, R., Pagallo, U., Rossi, F., Schafer, B., Valcke, P., & Vayena, E. (2018).

AI4People – An ethical framework for a good AI society: Opportunities, risks, principles, and recommendations. *Minds and Machines, 28*(4), 689–707. https://doi.org/10.1007/s11023-018-9482-5

Fung, A., Graham, M., & Weil, D. (2007). *Full disclosure: The perils and promise of transparency*. Cambridge University Press. https://doi.org/10.1017/9780521699617

Habermas, J. (1996). *Between facts and norms: Contributions to a discourse theory of law and democracy* (W. Rehg, Trans.). MIT Press. https://doi.org/10.7551/mitpress/1564.001.0001

Hallahan, K., Holtzhausen, D., van Ruler, B., Verčič, D., & Sriramesh, K. (2007). Defining strategic communication. *International Journal of Strategic Communication, 1*(1), 3–35. https://doi.org/10.1080/15531180701285244

Luguri, J., & Strahilevitz, L. J. (2021). Shining a light on dark patterns. *Journal of Legal Analysis, 13*(1), 43–109. https://doi.org/10.1093/jla/laaa006

Mittelstadt, B. D., Allo, P., Taddeo, M., Wachter, S., & Floridi, L. (2016). The ethics of algorithms: Mapping the debate. *Big Data & Society, 3*(2), 1–21. https://doi.org/10.1177/2053951716679679

Suchman, M. C. (1995). Managing legitimacy: Strategic and institutional approaches. *Academy of Management Review, 20*(3), 571–610. https://doi.org/10.5465/amr.1995.9508080331

Susser, D., Roessler, B., & Nissenbaum, H. (2019). Technology, autonomy, and manipulation. *Internet Policy Review, 8*(2), 1–22. https://doi.org/10.14763/2019.2.1410

Tappin, B. M., Wittenberg, C., Hewitt, L. B., Berinsky, A. J., & Rand, D. G. (2023). Quantifying the potential persuasive returns to political microtargeting. *Proceedings of the National Academy of Sciences, 120*(25), Article e2216261120. https://doi.org/10.1073/pnas.2216261120

Tyler, T. R. (2006). *Why people obey the law*. Princeton University Press. https://doi.org/10.1515/9781400828609

Woolley, S. C., & Howard, P. N. (Eds.). (2018). *Computational propaganda: Political parties, politicians, and political manipulation on social media*. Oxford University Press. https://doi.org/10.1093/oso/9780190931407.001.0001

Zarouali, B., Dobber, T., De Pauw, G., & de Vreese, C. (2022). Using a personality-profiling algorithm to investigate political microtargeting: Assessing the persuasion effects of personality-tailored ads on social media. *Communication Research, 49*(8), 1066–1091. https://doi.org/10.1177/0093650220961965

Zuboff, S. (2019). *The age of surveillance capitalism: The fight for a human future at the new frontier of power*. PublicAffairs.