From Analogue Totalitarianism to Digital Authoritarianism:
The "Perfect Prompt" as a Heuristic Tool for Analysing Algorithmic Control

Nuno Santos

CEPCEP – Research Centre for Peoples and Cultures of Portuguese Expression
Universidade Católica Portuguesa, Lisbon, Portugal
nunosantos@ucp.pt

aipropagandanazi.com

Abstract

This article proposes the "Perfect Prompt" as a heuristic for analyzing algorithmic governance in the digital age. While existing frameworks — surveillance capitalism, platform governance, informational authoritarianism — theorize the economic and political logics of algorithmic power, none adequately names the phenomenological register in which algorithmic authority is experienced as simultaneously structureless and totalizing. This gap matters for regulation: frameworks such as the EU AI Act and the Digital Services Act rest on assumptions of transparency and auditability that may be insufficient when harm emerges as an unauthored property of optimization systems. Drawing on documented cases of shadowbanning and electoral disinformation, the article articulates the "Perfect Prompt" through three operational concepts — "Digital Submission Labour", "Digital Social Death", and the "banality of code" — and extends the framework to Large Language Models through the concept of "intentionality of code". The analysis operates at three levels — technical infrastructure, the political economy of platforms, and subjective user experience — avoiding both technological determinism and economic reductionism, and offers implications for outcome-based algorithmic accountability in European digital policy.

Keywords: algorithmic governance; digital authoritarianism; platform accountability; EU AI Act; Digital Services Act; surveillance capitalism

1. Introduction

A structural puzzle haunts contemporary critical theory of digital platforms: the experience of algorithmic authority as both absent and omnipresent. Users encounter content moderation decisions with no identifiable decision-maker, recommendation pathways with no visible designer, and exclusions with no responsible author. Existing frameworks — surveillance capitalism (Zuboff, 2019), platform governance (Gillespie, 2018), informational authoritarianism (Guriev & Treisman, 2022) — have produced foundational accounts of the "what" and "how" of algorithmic power: its economic logic, its regulatory implications, its political instrumentalization. What these frameworks share, however, is a structural limitation: they do not adequately theorize the "where" of algorithmic authority — the specific phenomenological register in which the subject experiences control as simultaneously structureless and totalizing. This article addresses that gap.

The article proposes the "Perfect Prompt" as a heuristic for naming this phenomenological register. The concept does not replace existing frameworks; it operates at a different analytical level, naming the moment when an algorithmic system — without a human author, without a visible command, without an identifiable intent — becomes a functional substitute for the charismatic authority that classical political theory associates with totalitarian governance. Three questions organize the analysis. How do mechanisms of algorithmic control — shadowbanning, the dissemination of "Viral Lies," predetermined choice architectures — reconfigure the logics of historical authoritarianism for the digital era? In what ways do users, through what this article calls "Digital Submission Labour," become co-authors of this system

while retaining margins of agency? And what are the implications for European regulatory frameworks premised on transparency and auditability?

The historical comparison between twentieth-century totalitarianism and contemporary platform governance requires methodological care. The comparison advanced here is strictly functional, not moral or empirical. It does not claim equivalence in scale, intention, or consequence between the totalitarian regimes of the twentieth century and contemporary digital platforms. What it claims is narrower: that certain mechanisms for engineering consent, managing information flows, and rendering dissidents invisible exhibit structural continuities across the analog-digital transition, and that holding the historical case alongside the contemporary one reveals features of digital governance that remain invisible when the latter is analyzed in purely technical or economic terms.

The article's central contribution is the "Perfect Prompt" heuristic itself — a conceptual tool for naming the phenomenological experience of algorithmic authority as simultaneously absent and omnipresent. Three supporting concepts operationalize this heuristic for empirical research: "Digital Submission Labour" (the participatory dimension of algorithmic governance), "Digital Social Death" (the structural invisibilization of users), and the "banality of code" (the non-ideological mechanism of algorithmic harm). A further extension addresses Large Language Models (LLMs) through the concept of "intentionality of code," distinguishing emergent algorithmic harm from deliberately engineered influence operations. Together, these contributions offer researchers and regulators a vocabulary that is both theoretically grounded and empirically operationalizable, with direct implications for European digital policy.

2. Methodology and Analytical Approach

This article adopts a conceptual paper approach with empirical illustration, drawing on philosophy of technology, political communication studies, and digital sociology. Empirical cases were selected according to three criteria: (1) evidentiary robustness — only phenomena documented by peer- reviewed research with verifiable datasets were considered, such as YouTube' s recommender system (Covington et al., 2016; Ribeiro et al., 2020) and shadowbanning practices (Delmonaco & Haimson, 2024; Haimson & Delmonaco, 2025); (2) platform heterogeneity — cases span different architectures (video recommendation, text-based microblogging, messaging applications) to avoid inferences confined to a single platform; and (3) analytical pertinence — each case illustrates a distinct operational concept rather than serving as anecdotal reinforcement.

The use of totalitarian precedents to analyze digital governance is itself contested within the critical literature. Scholars such as Chun (2021) and Bratton (2015) have cautioned against historical analogies that may collapse qualitative distinctions between political regimes, while others — notably Fuchs (2018) and Stanley (2018) have defended the productive function of such comparisons for identifying structural continuities in propaganda and control techniques. This article aligns with the latter position while acknowledging its risks: the analogy is deployed as a diagnostic lens, not as an equivalence claim, and its validity depends on its capacity to illuminate features of digital governance that remain obscured in purely technical or economic readings.

The choice of YouTube as the primary deep- dive case (Section 4.3) merits explicit justification. Unlike TikTok, whose recommendation architecture remains substantially undisclosed, or Facebook, whose internal research has only partially entered the public domain through whistleblower disclosures, YouTube offers a rare instance of technical self-documentation. Google engineers published the Covington, Adams, and Sargin (2016) paper in a peer- reviewed venue, making the optimization objective — and its absence of civic

reasoning — a matter of verifiable record rather than inference. It is precisely this transparency-within- opacity that makes YouTube analytically productive.

A second methodological note concerns the epistemological status of the conceptual paper. This approach does not produce statistical generalization. It produces transferability: a set of analytical tools sufficiently developed to be applied, tested, and refined by other researchers in other empirical contexts (Sandvig et al., 2014). The "Perfect Prompt" framework is best understood not as a theory to be falsified but as a heuristic to be mobilized.

3. Theoretical Framework

The study of totalitarianism and propaganda has produced foundational contributions from Arendt (1951, 1963), Ellul (1965), and Foucault (1975). These authors work from distinct, at times incompatible premises: Arendt locates totalitarian control in the fusion of ideology and terror; Foucault shifts the analysis toward capillary, impersonal power operating through self-discipline; Ellul argues that propaganda in technological societies depends not on ideological adherence but on the social integration of the individual into the technical system. This article does not treat these authors as a complementary bloc but mobilizes them as alternative explanatory framings whose productive tension helps identify what is specific about algorithmic control.

In the digital era, the discussion has expanded to encompass new technological infrastructures, with particular emphasis on Zuboff's (2019) "surveillance capitalism", Heidegger's (1977) "Gestell", and Bourdieu's (1986) theories of field and capital. Guriev and Treisman (2022) have proposed "informational authoritarianism", framing political control as the manipulation of information rather than overt coercion. Here too, tensions are instructive: Zuboff emphasizes the extractive logic of behavioral data for commercial profit; Guriev and Treisman focus on

the political instrumentalization of information by state actors; and Heidegger offers an ontological reading of technology that exceeds both economic and political registers.

It is precisely within these tensions that the "Perfect Prompt" acquires its analytical distinctiveness. The concept does not seek to replace "surveillance capitalism" or "algorithmic governance" but to name something these frameworks presuppose yet rarely theorize directly: the moment at which the algorithmic system becomes a functional substitute for the charismatic authority of twentieth-century totalitarianism, producing consent without command. Whereas Zuboff (2019) describes the economic infrastructure of data extraction and the algorithmic governance literature (Gillespie, 2018; O'Neil, 2016) analyzes the rule-making power of code, the "Perfect Prompt" names the phenomenological effect of this infrastructure on the user: the experience of an authority that is simultaneously absent (no visible leader) and omnipresent (shaping every informational horizon). In this sense, the concept functions at a different analytical level — not the level of political economy (Zuboff) or legal-technical regulation (Lessig, 1999), but the level of the subject's phenomenological relationship with algorithmic mediation. This is what allows it to engage productively, rather than redundantly, with adjacent frameworks.

Amoore (2019) and Bucher (2017) have developed adjacent arguments concerning the phenomenological dimensions of algorithmic experience — the former through the figure of "doubt" in machine learning, the latter through the "algorithmic imaginary". The "Perfect Prompt" builds on these contributions while shifting the analytical emphasis toward the political register: specifically, toward the conditions under which algorithmic mediation functions as authority in its own right.

Table 1 presents the operational concepts developed from this theoretical positioning and their relation to existing frameworks.

Table 1. Operational Concepts, Theoretical Basis, and Relation to Existing Frameworks

Operational Concept	Analytical Definition	Theoretical Basis	Relation to Existing Frameworks
Perfect Prompt	Heuristic naming totalising algorithmic governance, substituting human will with machine logic.	Heidegger (1977); Zuboff (2019)	Names the phenomenological register absent in Zuboff (political economy) and Gillespie (platform governance).
Algorithmic Aura	Technical simulation of authenticity in synthetic content, endowing it with a sheen of truth.	Benjamin (1968)	Inverts Benjamin's thesis: AI-generated content does not lack aura but fabricates it.
Digital Social Death	Structural invisibilisation of subjects through algorithmic exclusion.	Arendt (1951); Foucault (1975); Noble (2018)	Complements Noble's algorithmic oppression by naming the phenomenological experience of exclusion.
Digital Submission Labour	Online activity that, by feeding algorithms data, contributes to maintaining the control system.	Bourdieu (1986); Zuboff (2019)	Specifies the participatory dimension underspecified in Zuboff's extractive framing.
Viral Lie	Decentralised, organic disinformation with high adaptability, driven by radicalised users.	Ellul (1965); McIntyre (2018)	Distinguishes organic disinformation from Woolley & Howard-type computational propaganda.

Source: Author's own elaboration.

4. Analysis

4.1 From Broadcast Architecture to Networked Propagation

Twentieth-century totalitarian propaganda operated through what might be called a broadcast architecture of power: a single emitter — ministries of propaganda, state-controlled media — directed messages to populations positioned as passive receivers. This architecture was inherently fragile: it was legible. The source of propaganda was identifiable, its apparatus was visible, and resistance — however costly — could at least orient itself against a known target.

The digital era does not merely accelerate this model; it structurally inverts it. When disinformation spreads through networks of millions of private messaging chats, as

documented in the Brazilian electoral context (Evangelista & Bruno, 2019), there is no ministry to identify, no broadcast tower to destroy, and no single source against which critique can orient itself. The "Viral Lie" is structurally leaderless — which is not to say it is spontaneous. Ellul's (1965) insight is essential: propaganda in technological societies does not require ideological adherence but rather the social integration of the individual into the technical system. Users who share disinformation need not be committed ideologues; they need only be sufficiently embedded in the informational ecosystem for content to arrive pre-validated by the network through which it travels. The lie spreads not because people believe it but because people trust the channel through which it arrives.

This reconfigures the relationship between propaganda and consent. Under analogue authoritarianism, consent was manufactured at the center and distributed to the periphery. Under digital authoritarianism, consent is manufactured at the periphery and aggregated at the center: platform engagement metrics reflect and amplify the emotional investments of millions of individual users, producing feedback loops in which the system learns, at massive scale, which emotional registers — outrage, fear, tribal solidarity — most effectively sustain attention. The "Perfect Prompt" does not tell users what to believe; it discovers, through billions of micro-interactions, what they are already inclined to believe and amplifies it. Large-scale disinformation campaigns documented in the 2018 Brazilian elections, Brexit, and the 2022 invasion of Ukraine (Allcott & Gentzkow, 2017; Evangelista & Bruno, 2019; Geissler et al., 2023) demonstrate this dynamic operationally.

4.2 The Dissolution of Central Authority

The figure of the twentieth-century totalitarian leader was constrained by biology and visibility. The "Perfect Prompt", by contrast, presents itself as ethereal and ubiquitous. This synthesized algorithmic authority requires no podium; it resides in the cloud, operating through

recommender systems that adjust messages, exposure rhythms, and visibility hierarchies according to each user's behavioral profile. Power ceases to present itself as explicit command and instead operates as invisible curation of the informational environment.

The distinction between visible and invisible authority has profound consequences for accountability. When authority is visible — embodied in a leader, an institution, or a decree — it can be named, opposed, and held responsible. When authority is invisible — distributed across billions of recommendation events, each technically defensible as a response to user behavior — accountability dissolves. Who is responsible for a radicalization pathway that no engineer designed, no executive ordered, and no user consciously chose?

The Algorithmic Aura illuminates a second dimension of this invisibility: synthetic authenticity. Benjamin (1968) argued that mechanical reproduction destroys the "aura" of the original work of art. The Algorithmic Aura inverts this logic: AI-generated content does not lack aura; it fabricates it. A deepfake video does not present itself as a copy of an original; it presents itself as an original, complete with the gestural, vocal, and affective signatures of authenticity. The result is an epistemic environment in which authentication becomes structurally impossible for the ordinary user.

This epistemic impossibility is not accidental. Authoritarian governance — whether by state actors or platform architectures — benefits from generalized epistemic uncertainty. When users cannot reliably distinguish authentic from synthetic, they face two dysfunctional responses: either they trust nothing (paralysis) or they trust their prior commitments regardless of evidence (confirmation bias amplification). Both responses are, from the perspective of the "Perfect Prompt", functional: paralysis reduces civic participation; confirmation bias increases engagement. The algorithm does not care which response the user adopts, so long as they remain on the platform.

What makes the “Perfect Prompt” categorically different from the totalitarian propaganda apparatus is therefore not its power — the reach of digital platforms vastly exceeds that of any twentieth-century ministry — but its indifference to content. Goebbels had a message. The algorithm has a metric.

4.3 The Banality of Code: The YouTube Recommender Case

If twentieth-century authoritarianism was driven by visceral hatred, the “Perfect Prompt” is driven by the cold logic of optimization. The regime of algorithmic exclusion does not manifest as acts of explicit violence but as Digital Social Death: the infrastructural invisibilization of voices the system identifies — without consciousness, without malice — as less “engaging”. The banality of code — proposed here in dialogue with Arendt’s (1963) “banality of evil” and Lessig’s (1999) “code is law” — describes this mechanism: it is not ideological hatred that drives the system, but the relentless pursuit of KPIs. Digital evil is terrible precisely because it is indifferent; the algorithm does not “hate”; it merely optimizes.

This is best understood through a deeper empirical case. Covington, Adams, and Sargin (2016), in their technical paper describing YouTube’s deep neural network architecture, disclose the explicit optimization objective: “expected watch time per impression”. The paper does not address content quality, informational accuracy, or civic consequences. The objective function is strictly behavioral. This is the banality of code in concrete form: an engineering choice, internally defensible, containing no ideological malice.

Ribeiro et al. (2020) — auditing over 300,000 YouTube videos and 79 million comments across 349 channels — demonstrate that the same system produces systematic radicalization pathways. Users who begin on moderate content are progressively directed toward increasingly extreme channels, with documented migration from Intellectual Dark Web content to Alt-Lite and Alt-Right ecosystems. The mechanism is not ideological curation; it is a statistical artifact

of watch-time optimization interacting with the attention profiles of viewers predisposed to radical content. No engineer designed radicalization. The system produces it as a by-product of the KPI it was built to maximize. This is what distinguishes the banality of code from Arendt's banality of evil: Eichmann at least had to fill out forms; the algorithm does not even need that minimal bureaucratic gesture. Consequences emerge unbidden from the optimization objective.

Analogously, Facebook's moderation systems, by prioritizing removal of content that violates "community standards", systematically silence marginalized voices — not out of malice, but due to biases embedded in model training (Gillespie, 2018). Designed to scale moderation globally, these systems reproduce cultural and linguistic biases, resulting in algorithmic exclusion that, although unintentional, has real political effects (Gorwa et al., 2024). Empirical cases of shadowbanning on Twitter/X and TikTok (Delmonaco & Haimson, 2024) document psychological impacts, including depression and anxiety (Haimson & Delmonaco, 2025). The common thread is the absence of a deciding subject: the exclusion is real, its effects are measurable, but no individual intended it. This is the specific form of violence the "Perfect Prompt" enacts — violence without an author.

4.4 Digital Submission Labour and Participatory Subjugation

One of the most insidious features of digital authoritarianism is how users become co-authors of their own subjugation. Every time users train an algorithm with their preferences, they actively help build the system that categorizes them and may ultimately silence them. Freedom of expression on social networks is, in practice, the freedom to feed the control system. This is Digital Submission Labour (Bourdieu, 1986; Zuboff, 2019), observable in practices such as continuous scrolling, reactive interaction, and implicit feedback to recommender systems,

through which apparently free online activity is converted into capital for the algorithmic system.

In Bourdieusian terms (Bourdieu, 1986), the digital platform functions as a field in which users accumulate and expend symbolic capital — followers, likes, reach, virality — according to rules they did not set and cannot fully perceive. The habitus of the digital user — the embodied disposition to seek validation through engagement metrics, to calibrate expression to what the algorithm rewards — is formed through repeated participation in the field, internalizing its logic until it feels like second nature. The user does not experience this as submission; they experience it as expression.

This foregrounds a tension the argument must acknowledge rather than resolve: agency within structural constraint. Digital Submission Labour is not slavery. Users can and do resist, subvert, and exit — albeit at different costs depending on their position within the platform's economy of attention. A researcher who deactivates a platform loses a professional network; a marginalised activist who is shadowbanned loses a lifeline. The structural constraint is real but not uniform. What the "Perfect Prompt" framework contributes here is a vocabulary for naming the specific form of power at work: not direct coercion (the totalitarian model) nor pure market choice (the liberal model), but participatory entrapment — a system in which the conditions of participation are simultaneously the conditions of one's own subjugation, and in which the cost of exit is calibrated, whether by design or by structural effect, to be prohibitive for those who most need the platform's affordances.

4.5 The Perfect Prompt Literalised: Large Language Models

The preceding sections have developed the "Perfect Prompt" as a heuristic concept. There is, however, a literalization of the metaphor that the analysis must confront: the emergence of Large Language Models (LLMs) as infrastructure for a new phase of digital authoritarianism.

If the recommender algorithm was the first generation of the “Perfect Prompt” — shaping what users see without shaping what is said — the LLM constitutes a second, more radical generation: a system that not only curates discourse but also generates it.

The irony is not incidental. The term “Perfect Prompt” refers, in generative AI vocabulary, to the ideal instruction that elicits a language model’s most compliant, most calibrated output. Prompt optimization is the technical practice of finding the precise formulation that causes the system to produce the desired response — reliably, at scale, without ideological resistance. Transposed to the political register, this is the dream of every propaganda ministry: a communication apparatus that produces the right message, in the right register, for the right audience, without the unreliability of human intermediaries.

Documented influence operations have already demonstrated the operational use of LLMs for the mass production of synthetic political content. Research has shown that GPT-class models can generate persuasive political messaging indistinguishable from human-written content, at a fraction of the cost and at a speed that outpaces traditional fact-checking infrastructures (Goldstein et al., 2024). The Stanford Internet Observatory has documented networks of AI-generated social media profiles (Goldstein et al., 2023) — synthetic personas complete with profile photographs, posting histories, and engagement patterns — deployed in coordinated inauthentic behavior campaigns. What the recommender algorithm did to the distribution of information, the LLM does to its production: it removes the human bottleneck.

This deepens the “banality of code” argument in a direction Arendt’s framework did not anticipate. In the YouTube case, algorithmic harm was a by-product of an objective function — radicalization emerged unbidden from watch-time optimization. In the LLM case, harm can be deliberately engineered through the objective function itself: a model fine-tuned on propaganda corpora, optimized for persuasive political content, and deployed through synthetic persona networks is not producing harm as an emergent by-product but as a designed output.

The “banality of code” gives way to what might be called the intentionality of code — a system in which the gap of complexity no longer separates engineering choice and political consequence but is directly aligned. This represents a qualitative escalation in the political danger posed by algorithmic systems, and one the “Perfect Prompt” framework is positioned to name: the movement from a system that governs discourse by shaping distribution to one that governs it by automating production.

5. Discussion: Implications for European Digital Policy

5.1 Responding to the Principal Objection

The most foreseeable objection to this argument is that comparing twentieth-century totalitarianism with digital platform governance trivializes historical atrocity. This objection warrants a direct response.

The comparison advanced here is strictly functional, not moral. The article does not claim that platform companies are equivalent to totalitarian regimes or that algorithmic exclusion matches the systematic violence of historical totalitarianism. It makes a narrower claim: that certain mechanisms for engineering consent, managing information flows, and rendering dissidents invisible exhibit structural continuities across the analog-digital transition. The “banality of code” is not an accusation; it is a diagnostic tool. Just as Arendt’s (1963) “banality of evil” was not intended to minimize Eichmann’s culpability but to illuminate the bureaucratic conditions that made mass participation in atrocity possible, the “banality of code” illuminates the engineering conditions that make mass participation in algorithmic control structurally unavoidable. The comparison is a lens, not an equation.

5.2 The Location of Agency

A second, more technically precise objection concerns agency. If humans design algorithms, the argument goes, the "dissolution of central authority" is overstated: there are identifiable authors behind every design decision. This is correct but misidentifies the scale. The argument is not that human authorship disappears; it is that human authorship becomes non-attributable at the level of systemic consequence. The engineer who specifies a watch-time metric does not author the radicalization pathway that emerges from the interaction of that metric with the attention profiles of millions of users. The gap between an engineering decision and its social consequence is not a gap of intention but a gap of complexity. This is what makes the "Perfect Prompt" politically more dangerous than classical authoritarianism, not less. A minister of propaganda has a name, a phone number, and a trial waiting. An engagement metric has none of these.

5.3 Implications for the AI Act and the Digital Services Act

The "Perfect Prompt" framework does not merely diagnose a problem; it reframes how solutions are discussed. Existing regulatory frameworks — most notably the European AI Act and the Digital Services Act — rest on assumptions about transparency and auditability that they themselves call into question. Transparency requirements assume that once an algorithm's workings are disclosed, users and regulators can evaluate its effects. Auditability requirements assume that systemic consequences can be attributed to identifiable design choices. The "banality of code" argument suggests both assumptions may be insufficient: a fully transparent watch-time optimization function reveals nothing objectionable in itself; the harm is not in the function but in its emergent interaction with human attention at scale.

This suggests European regulatory frameworks need to move beyond process transparency (disclosing how the algorithm works) toward outcome accountability (establishing responsibility for the algorithm's outputs). The distinction between intent and effect that

currently shields platform companies from liability — "we did not design radicalization; it is a by-product of user behavior" — is precisely the structure the "banality of code" concept names and challenges. Three concrete regulatory implications follow.

First, DSA implementation should incorporate mandatory outcome-based algorithmic auditing (Sandvig et al., 2014) that evaluates not whether platforms comply with disclosure requirements but whether their engagement-optimization systems systematically produce outcomes — radicalization, epistemic polarization, political demobilization — harmful to democratic pluralism regardless of intent.

Second, the AI Act's risk categorization should explicitly include phenomenological harms: the psychological effects of shadowbanning (Haimson & Delmonaco, 2025), the epistemic effects of synthetic authenticity, and the systemic effects of engagement optimization on democratic deliberation. These harms are currently difficult to classify under the Act's risk categories precisely because they lack identifiable authors.

Third, the emergence of intentionality in code in fine-tuned LLMs deployed in influence operations requires regulatory responses that cannot rely on distinguishing human-authored from machine-authored content at scale. The Code of Practice on Disinformation and the AI Act's provisions on synthetic content need mechanisms for detecting and attributing systems rather than individual pieces of content.

5.4 Limitations and Alternatives

Resistance strategies exist. Migration to decentralized platforms such as Mastodon or PeerTube (La Cava et al., 2021), literacy projects such as AlgorithmWatch, and regulatory frameworks such as the AI Act and DSA represent institutional efforts. However, decentralized platforms still struggle to reach critical mass (La Cava et al., 2021); literacy projects encounter algorithmic opacity and power asymmetries between users and platforms (UNESCO, 2021);

even advanced regulatory frameworks depend on political will and are subject to pressure from major technology firms (Gorwa et al., 2024). Long-term effectiveness depends on articulating local resistance with systemic changes in platform governance.

The article's limitations should be made explicit. It does not demonstrate causality between algorithmic mechanisms and specific political outcomes. Empirical cases are illustrative rather than exhaustive. The concept of "digital authoritarianism" does not replace analyses specific to particular platforms, regimes, or cultural contexts. These delimitations do not weaken the heuristic proposal; they define its scope of application.

6. Conclusion

This article has contributed to the literature on digital authoritarianism by conceptualizing the "Perfect Prompt" as a heuristic for analyzing algorithmic governance. By articulating this concept alongside "Digital Submission Labour", "Digital Social Death", and the "banality of code", and by extending it to LLMs through the "intentionality of code", the analysis shows how logics of control are reconfigured in the digital era, shifting from centralized authority to a regime of distributed, participatory exclusion. The main theoretical contribution is naming the phenomenological register that existing frameworks presuppose but do not adequately theorize, with direct implications for European digital policy.

A multifaceted research agenda follows. Future studies could (1) empirically map the prevalence of "Digital Social Death" across platforms using algorithmic auditing methods and longitudinal studies of psychosocial impact (Haimson & Delmonaco, 2025; Sandvig et al., 2014); (2) compare the governance architectures of centralized and decentralized platforms (La Cava et al., 2021); (3) normatively explore "ethics-by-default" AI, integrating algorithmic transparency and democratic participation in design (Couldry & Mejias, 2019; UNESCO, 2021); (4) assess whether DSA and AI Act implementation can move from process

transparency to outcome accountability. The question is not whether technology controls us, but how we, as a society, can learn to control technology — and ourselves — as agents of a digital era still to be written.

Declaration on the Use of Generative AI

During the preparation of this manuscript, the author used Claude (Anthropic) for editorial assistance, including linguistic refinement and structural revision. The author reviewed and edited all content and assumes full responsibility for the conceptualization, argumentation, and conclusions presented.

References

Allcott, H., & Gentzkow, M. (2017). Social media and fake news in the 2016 election. Journal of Economic Perspectives, 31(2), 211–236. https://doi.org/10.1257/jep.31.2.211

Amoore, L. (2019). Doubt and the algorithm: On the partial accounts of machine learning. Theory, Culture & Society, 36(6), 147–169. https://doi.org/10.1177/0263276419851846

Arendt, H. (1951). The origins of totalitarianism. Harcourt Brace Jovanovich.

Arendt, H. (1963). Eichmann in Jerusalem: A report on the banality of evil. Viking Press.

Benjamin, W. (1968). The work of art in the age of mechanical reproduction. In H. Arendt (Ed.), Illuminations (pp. 217–251). Schocken Books.

Bourdieu, P. (1986). Distinction: A social critique of the judgement of taste. Harvard University Press.

Bratton, B. H. (2015). The stack: On software and sovereignty. MIT Press.

Bucher, T. (2017). The algorithmic imaginary: Exploring the ordinary affects of Facebook algorithms. Information, Communication & Society, 20(1), 30–44. https://doi.org/10.1080/1369118X.2016.1154086

Chun, W. H. K. (2021). Discriminating data: Correlation, neighborhoods, and the new politics of recognition. MIT Press.

Couldry, N., & Mejias, U. A. (2019). The costs of connection: How data is colonizing human life and appropriating it for capitalism. Stanford University Press.

Covington, P., Adams, J., & Sargin, E. (2016). Deep neural networks for YouTube recommendations. Proceedings of the 10th ACM Conference on Recommender Systems, 191–198. https://doi.org/10.1145/2959100.2959190

Delmonaco, D., & Haimson, O. L. (2024). "What are you doing, TikTok?": How marginalized social media users make sense of shadowbanning. Proceedings of the ACM on Human-Computer Interaction, 8(CSCW1), 1–26. https://doi.org/10.1145/3637431

Ellul, J. (1965). Propaganda: The formation of men's attitudes. Alfred A. Knopf.

Evangelista, R., & Bruno, F. (2019). WhatsApp and political instability in Brazil: Targeted messages and political radicalization. Internet Policy Review, 8(4). https://doi.org/10.14763/2019.4.1434

Foucault, M. (1975). Surveiller et punir: Naissance de la prison. Gallimard.

Fuchs, C. (2018). Propaganda 2.0: Herman and Chomsky's propaganda model in the age of the internet, big data and social media. In J. Pedro-Carañana, D. Broudy, & J. Klaehn (Eds.), The propaganda model today: Filtering perception and awareness (pp. 71–92). University of Westminster Press.

Geissler, D., Bär, D., Pröllochs, N., & Feuerriegel, S. (2023). Russian propaganda on social media during the 2022 invasion of Ukraine. EPJ Data Science, 12, 35. https://doi.org/10.1140/epjds/s13688-023-00414-5

Gillespie, T. (2018). Custodians of the internet: Platforms, content moderation, and the hidden decisions that shape social media. Yale University Press.

Goldstein, J. A., Chao, J., Grossman, S., Stamos, A., & Tomz, M. (2024). How persuasive is AI-generated propaganda? PNAS Nexus, 3(2), pgae034. https://doi.org/10.1093/pnasnexus/pgae034

Goldstein, J. A., Sastry, G., Musser, M., DiResta, R., Gentzel, M., & Sedova, K. (2023). Generative language models and automated influence operations: Emerging threats and potential mitigations. Stanford Internet Observatory. https://arxiv.org/abs/2301.04246

Gorwa, R., Lechowski, G., & Schneiss, D. (2024). Platform lobbying: Policy influence strategies and the EU's Digital Services Act. Internet Policy Review, 13(2). https://doi.org/10.14763/2024.2.1782

Guriev, S., & Treisman, D. (2022). Spin dictators: The changing face of tyranny in the 21st century. Princeton University Press.

Haimson, O. L., & Delmonaco, D. (2025). Digital silence: The psychological impact of being shadow banned on social media. Frontiers in Psychology. https://doi.org/10.3389/fpsyg.2025.1659272

Heidegger, M. (1977). The question concerning technology, and other essays. Harper & Row.

La Cava, L., Greco, S., & Tagarelli, A. (2021). Understanding the growth of the Fediverse through the lens of Mastodon. Applied Network Science, 6(1), 1–35. https://doi.org/10.1007/s41109-021-00392-5

Lessig, L. (1999). Code and other laws of cyberspace. Basic Books.

McIntyre, L. (2018). Post-truth. MIT Press.

Noble, S. U. (2018). Algorithms of oppression: How search engines reinforce racism. New York University Press.

O'Neil, C. (2016). Weapons of math destruction: How big data increases inequality and threatens democracy. Crown.

Ribeiro, M. H., Ottoni, R., West, R., Almeida, V. A. F., & Meira, W., Jr. (2020). Auditing radicalization pathways on YouTube. Proceedings of the 2020 Conference on Fairness, Accountability, and Transparency, 131–141. https://doi.org/10.1145/3351095.3372879

Sandvig, C., Hamilton, K., Karahalios, K., & Langbort, C. (2014). Auditing algorithms: Research methods for detecting discrimination on internet platforms. Data and Discrimination: Converting Critical Concerns into Productive Inquiry, 1–23.

Stanley, J. (2018). How fascism works: The politics of us and them. Random House.

UNESCO. (2021). Recommendation on the ethics of artificial intelligence. https://unesdoc.unesco.org/ark:/48223/pf0000381137

Zuboff, S. (2019). The age of surveillance capitalism. PublicAffairs.