ETHICAL GOVERNANCE IN THE AGE OF ARTIFICIAL INTELLIGENCE: CHALLENGES AND OPPORTUNITIES FOR THE WESTERN BALKANS

Marco MARSILI*

Abstract. *This paper examines the ethical and political implications of Artificial Intelligence (AI) in governance, with a focus on the Western Balkans in the broader context of European integration and democratic transformation. AI technologies increasingly shape decision-making, policy implementation, and public discourse, offering opportunities for efficiency, transparency, and citizen participation. Yet they also raise significant risks of surveillance, algorithmic bias, and reduced accountability, particularly in regions facing fragile democratic institutions, ethnic diversity, and complex security challenges. Drawing on political philosophy, ethical theory, and contemporary debates in digital governance, this study develops a normative framework for ethical AI adoption grounded in democratic values and human rights. Empirical references to governance practices in the Western Balkans, alongside comparative insights from Central and Eastern Europe, highlight both regional vulnerabilities and potential pathways for responsible AI integration. The paper argues for cross-border cooperation, civil society involvement, and policy harmonization with EU standards as essential conditions for building ethical AI governance. By addressing both risks and opportunities, it contributes to debates on democratization, security management, and the role of emerging technologies in shaping the future of governance.*

Keywords*: Artificial Intelligence; Ethical Governance; Western Balkans; Democratization; European Integration*

INTRODUCTION

The accelerating integration of Artificial Intelligence (AI) into governance structures raises pressing ethical and political questions, especially in transitional and geopolitically sensitive regions such as the Western Balkans. In societies where democratic institutions are still consolidating and where the legacy of conflict and authoritarian governance remains palpable, the adoption of AI technologies may bring both opportunities for modernization and risks of reinforcing existing inequalities or enabling new forms of control. As Floridi (2019) has argued, AI poses unique challenges to the "infosphere", demanding ethical frameworks that transcend technical issues and address broader concerns about autonomy, dignity, and justice. In this sense, the Western Balkans offer an important case study for assessing how ethical governance of AI can be aligned with the region's democratization, European integration, and security imperatives.

* Assistant Professor PhD, Department of Philosophy and Cultural Heritage University of Venice, Italy, E-mail: info@marcomarsili.it.

The European Union (EU) has sought to position itself as a normative leader in the global governance of AI, developing regulatory instruments such as the proposed *Artificial Intelligence Act* (AI Act) and the *Digital Services Act* (DSA Act), which aim to balance innovation with the protection of fundamental rights. Yet, as Jobin, Ienca, and Vayena (2019) demonstrate in their comparative survey of global AI ethics guidelines, a proliferation of principles does not automatically translate into effective governance, particularly in contexts where institutional capacity is weak or political will is fragmented. The Western Balkans, situated at the crossroads of European, Russian, and Chinese spheres of influence, illustrate this tension vividly. Here, ethical governance of AI cannot be reduced to a checklist of values, but must be embedded within broader debates on sovereignty, democratic legitimacy, and the region's path toward EU integration.

The adoption of AI in governance promises certain advantages. Machine learning (ML) systems can enhance transparency in public administration, improve efficiency in service delivery, and strengthen mechanisms of accountability by detecting corruption or maladministration (Crawford, 2021). For societies in transition, where trust in institutions is often fragile, such technological affordances could support democratic consolidation. However, these opportunities are paralleled by profound risks. Algorithmic decision-making may replicate or even exacerbate existing social biases, erode privacy, and enable new forms of surveillance (Zuboff, 2019). Moreover, in a region marked by ethnical and confessional diversity, AI-driven governance could unintentionally magnify fault lines if not carefully managed.

From a normative perspective, the ethical governance of AI intersects with fundamental freedoms and rights. A previous study of the historical and contemporary trajectories of freedom of thought, emphasizes how digital technologies reshape not only the exercise of individual liberty but also the collective conditions of democratic life (Marsili, 2025a). This reflections resonate with the challenges faced by the Western Balkans, where fragile democratic institutions must contend with the dual pressures of digital transformation and external geopolitical competition. Ethical governance of AI in this context requires a commitment to safeguarding freedom of thought and information while resisting authoritarian temptations of control that emerging technologies may enable.

The geopolitical dimension is equally critical. As Nemitz (2018) has argued, the governance of digital technologies is not merely a technical or economic matter but a constitutional one, with implications for sovereignty, legitimacy, and human rights. In the Western Balkans, this dimension is magnified by external actors competing for influence. China's Digital Silk Road and Russia's information operations highlight how AI and digital infrastructures can become tools of soft power and hybrid warfare (Marsili, 2021). For states in the region, the challenge is not only to adopt ethical governance frameworks domestically but also to resist or navigate external pressures that may undermine democratic standards.

This paper seeks to contribute to these debates by examining the ethical and political implications of AI governance in the Western Balkans through the dual lenses of political philosophy and empirical developments in the region. It argues that while AI offers opportunities to strengthen transparency, efficiency, and democratic participation, it also carries significant risks of reinforcing authoritarian practices and deepening social divides. Addressing these challenges requires a governance approach that is both ethically robust and politically sensitive, attuned to the

specificities of the Western Balkans' democratic trajectories and their aspirations toward European integration.

The analysis unfolds in four sections. The first provides a theoretical and ethical framework, situating AI governance within broader debates on ethics, democracy, and human rights. The second examines the political and institutional context of the Western Balkans, highlighting opportunities and vulnerabilities in the adoption of AI technologies. The third explores case studies of AI applications in governance across the region, including digital identity systems, predictive policing, and e-government platforms. The fourth assesses policy implications, with particular attention to how regional cooperation and alignment with EU standards can ensure that AI development supports, rather than undermines, democratic governance. The paper concludes by offering recommendations for strengthening ethical governance in the Western Balkans, drawing on both theoretical insights and practical considerations.

Ultimately, this study aims to show that the ethical governance of AI is not merely a technical or regulatory concern, but a deeply political one, tied to questions of identity, legitimacy, and integration. For the Western Balkans, engaging critically with AI technologies is not only a matter of technological adoption but also of safeguarding democratic futures in an increasingly digital and contested global order.

1. THEORETICAL AND ETHICAL FRAMEWORK

The ethical and political governance of Artificial Intelligence cannot be understood in isolation from broader theoretical debates on democracy, human rights, and the digital transformation of society. AI is not merely a technical tool but a socio-technical system that reflects, amplifies, and sometimes distorts the values and structures of the societies in which it is embedded. Any assessment of AI governance in the Western Balkans must therefore be situated within a normative framework that considers the role of technology in shaping political legitimacy, social cohesion, and the protection of fundamental rights.

At the most general level, the debate surrounding AI governance resonates with the philosophical concern over the relationship between human autonomy and external forms of control. Floridi (2019) has emphasized that the digital revolution is transforming the "infosphere"—the environment in which human life is increasingly lived—requiring an ethics that goes beyond traditional questions of responsibility or harm to encompass questions of agency, dignity, and justice. For Floridi, the ethical challenge lies not only in how AI makes decisions, but in how it reconfigures the conditions under which humans make their own. This shift demands a governance model that ensures AI systems respect and promote human flourishing rather than undermine it.

A related concern is the proliferation of ethical guidelines and principles without effective implementation. Jobin, Ienca, and Vayena (2019) documented the global landscape of AI ethics guidelines, highlighting broad convergence around principles such as transparency, justice, and accountability. Yet these principles are often aspirational, lacking mechanisms of enforcement or adaptation to specific political contexts. The Western Balkans illustrate this gap: while governments may adopt the rhetoric of ethical AI, the institutional weaknesses, corruption, and contested sovereignties characteristic of the region risk turning ethical principles into mere symbolic gestures. As such, the theoretical framework for analyzing AI

governance here must be attentive to the interplay between normative ideals and political realities.

The democratic implications of AI governance form a central dimension of this debate. As Nemitz (2018) has argued, the governance of digital technologies is not a technocratic question but a constitutional one, touching on the distribution of power and the protection of fundamental rights. In regions where democratic institutions are fragile, AI systems can either reinforce democratic accountability or provide new instruments of authoritarian control. The distinction depends on the values embedded in design, the transparency of implementation, and the safeguards for citizen participation. This makes the Western Balkans a critical site for examining how AI interacts with democratic consolidation and European integration.

Recent scholarship is particularly relevant here. Earlier research in *Libertà di pensiero nell'era digitale* (Marsili, 2025a) traces the genealogy of freedom of thought from antiquity to the digital age and shows how technological transformations continuously reshape individual autonomy and collective deliberation. In this perspective, the digital age presents both unprecedented opportunities for pluralism and significant risks of manipulation, surveillance, and erosion of critical reasoning. His analysis suggests that the ethical governance of AI must begin from the protection of cognitive freedom, understood not merely as the absence of coercion but as the active cultivation of conditions for informed and autonomous judgment. This perspective is vital for societies in the Western Balkans, where democratic cultures are still developing and where disinformation campaigns have repeatedly undermined public trust in institutions.

The question of freedom in the digital era also intersects with security. Earlier research in *Dal campo di battaglia al metaverso: la nuova frontiera delle operazioni militari* (Marsili, 2025b) shows how emerging technologies blur the boundaries between civilian and military domains, transforming cyberspace into a contested battlefield of narratives and perceptions. While his focus is on defense and hybrid warfare, the implications for AI governance are clear: technologies initially designed for efficiency or security can be repurposed for manipulation and control. In transitional societies with fragile democratic institutions, this risk is amplified. Ethical governance must therefore anticipate not only the intended uses of AI but also its potential misuses in the domains of propaganda, surveillance, and social control.

Earlier research on Europe's vulnerabilities in the context of the Russo-Ukrainian war further underscores this point. In his analysis of the EU's geopolitical challenges (2023a), he emphasizes how information operations exploit democratic weaknesses and social divisions, undermining trust and resilience. Transposed to the debate on AI, this perspective highlights the necessity of embedding governance within a broader framework of cognitive resilience. It is not enough to regulate algorithms in the abstract; societies must cultivate the democratic capacities—such as media literacy, pluralism of voices, and participatory oversight—that allow citizens to critically engage with technologies and resist manipulation.

From an ethical standpoint, the Western Balkans present a particularly complex environment. The region's history of ethnical and confessional diversity creates both an opportunity and a challenge for AI governance. On one hand, AI systems designed to promote inclusivity and pluralism could help strengthen social cohesion and support democratic integration. On the other, if algorithms are trained on biased data or deployed without sensitivity to historical grievances, they risk

exacerbating divisions. As Crawford (2021) reminds us, AI is never neutral; it reflects the social, political, and economic structures that produce it. For the Western Balkans, ethical governance means ensuring that AI systems are not merely imported from external actors but adapted to the region's specific social realities.

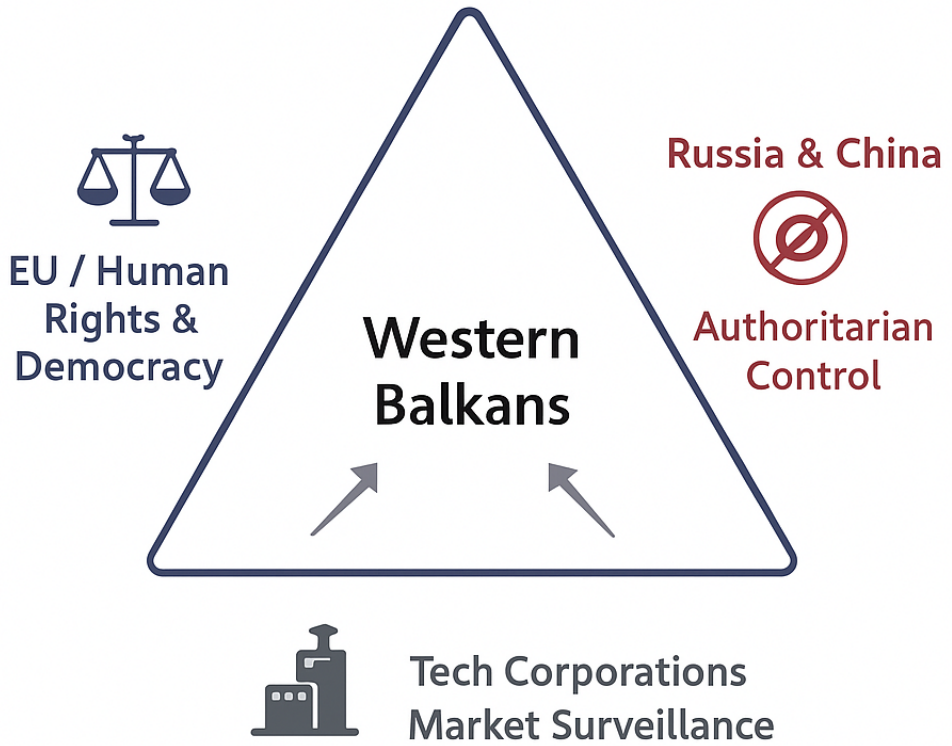
This raises the issue of external influence. Both Russia and China have actively engaged in promoting digital infrastructures in the Western Balkans, often framing their involvement as support for modernization while embedding alternative governance logics. In this sense, the governance of AI becomes a question of geopolitical alignment: whether the region will adopt European models grounded in human rights and democratic values, or drift toward more authoritarian templates. Zuboff's (2019) concept of "surveillance capitalism" provides a useful warning here: when technologies are governed primarily by logics of control or profit extraction, they undermine democratic accountability and erode the very freedoms they promise to expand. The Western Balkans, caught between competing spheres of influence, must therefore be especially vigilant in adopting governance frameworks that resist both authoritarian and exploitative trajectories.

In theoretical terms, then, the ethical governance of AI in the Western Balkans can be understood as a contest between three logics: the logic of human rights and democracy, represented by the EU and supported by normative scholarship in this field; the logic of authoritarian control, exemplified by the technological models promoted by Russia and China; and the logic of market-driven surveillance, epitomized by global tech corporations and analyzed by Zuboff and Crawford. Navigating these competing logics requires not only sound regulation but also a robust ethical framework capable of integrating normative principles with empirical realities.

The Western Balkans' path toward European integration provides a critical opportunity to anchor AI governance in the first of these logics. Yet it is underscored in earlier research (Marsili, 2025a) that freedom of thought and democratic deliberation cannot be taken for granted; they must be actively protected and cultivated. Ethical governance of AI, therefore, is not simply about compliance with European regulations but about nurturing the democratic culture that makes such compliance meaningful. This involves embedding human-rights assessments into technology adoption, ensuring transparency and accountability in decision-making processes, and fostering regional cooperation to resist external pressures.

In conclusion, the theoretical and ethical framework for AI governance in the Western Balkans must combine normative commitments with political realism. It must recognize the transformative potential of AI while guarding against its risks, situating governance within the broader struggles for democracy, sovereignty, and integration. By drawing on philosophical insights (Floridi, 2019), global ethical guidelines (Jobin et al., 2019), critical analyses of digital capitalism (Zuboff, 2019; Crawford, 2021), and region-specific contributions (Marsili, 2023a; 2025a; 2025b), this framework provides the conceptual grounding for the subsequent sections of this paper. The ethical governance of AI in the Western Balkans is not a peripheral issue but a central question for the region's democratic future and its place in Europe.

Figure 1. Competing logics of AI governance in the Western Balkans: democratic-human rights logic (EU), authoritarian logic (Russia and China), and market-driven surveillance logic (global tech corporations).



2. AI AND GOVERNANCE IN THE WESTERN BALKANS: POLITICAL AND INSTITUTIONAL CONTEXT

The Western Balkans represent a region where democratic consolidation remains fragile and institutions are often caught between the aspirations of European integration and entrenched patterns of competitive authoritarianism. As Florian Bieber (2018, p. 338) demonstrates, governments across the region frequently combine electoral processes with illiberal practices, including media control, selective rule of law, and restricted civic space. In such contexts, the introduction of Artificial Intelligence into governance does not occur in a political vacuum: instead, it interacts with weak institutions and blurred boundaries between public interest and partisan control.

This institutional fragility raises profound questions about how AI might be deployed in governance. On the one hand, AI-driven tools for public administration—such as automated service delivery or predictive analytics in policymaking—could enhance efficiency and transparency. On the other hand, without strong safeguards, these same tools may reinforce existing power asymmetries, enabling new forms of surveillance or discrimination. Earlier research (Marsili, 2023b, p. 27) highlights that compliance with international law and ethical standards is essential if technological innovation is to align with democratic values.

The European Union has positioned itself as a key normative actor in promoting ethical AI. The AI Act entyered into force in July 2024 emphasize transparency, accountability, and human-centric approaches. Yet, in the Western Balkans, these principles encounter the persistent gap between policy frameworks and actual practice. The adoption of EU standards is often rhetorical, while implementation lags due to political resistance or administrative incapacity. In such an environment, AI risks becoming another instrument of elite consolidation rather than a driver of democratic deepening.

Ultimately, the governance challenge in the Western Balkans is not merely about introducing advanced technologies but about ensuring that their deployment strengthens, rather than undermines, democratic legitimacy. Bridging this gap requires embedding AI governance within broader reforms on the rule of law, institutional accountability, and civil society participation.

3. AI, DEMOCRACY, AND FUNDAMENTAL RIGHTS: RISKS AND OPPORTUNITIES

Artificial Intelligence presents a profound paradox for democratic governance in the Western Balkans. On the one hand, AI offers opportunities to modernize public services, enhance transparency, and support evidence-based policymaking. On the other hand, its deployment in fragile democracies poses significant risks for human rights, privacy, and political pluralism.

At the European level, the emerging regulatory framework on AI is explicitly grounded in the protection of fundamental rights. The EU has set out a vision for "trustworthy AI" anchored in legality, ethical robustness, and social benefit. This vision is operationalized in the AI Act, which categorizes AI systems by risk level and seeks to prevent their misuse in ways that could endanger democratic values. This approach resonates with broader EU commitments to human dignity, equality, and data protection, enshrined in the *Charter of Fundamental Rights of the European Union* (CFR) proclaimed in 2000.

Yet, as scholars have noted, the region's hybrid political orders amplify the risks of misapplication. As Florian Bieber (2018, p. 345) argues, weak institutions and clientelist practices in the Western Balkans enable ruling elites to capture democratic mechanisms while maintaining a façade of compliance. Within such systems, the use of AI-driven surveillance, predictive policing, or algorithmic decision-making could easily be weaponized to monitor dissent, manipulate public discourse, and marginalize opposition voices.

It is noted in earlier research (Marsili, 2023, p. 43) that the ethical deployment of disruptive technologies depends on strong legal frameworks and an institutional culture of accountability. Without these safeguards, technological innovation risks reproducing existing asymmetries of power rather than democratizing access to information and services. The geopolitical dimension further complicates matters, since earlier research on global strategic competition (Marsili, 2025a, pp. 312–315) shows that digital infrastructures are also arenas of influence where external powers project narratives and exert control. The Western Balkans, already vulnerable to external influence campaigns, could thus see AI becoming an additional channel for authoritarian entrenchment.

Civil society emerges as a crucial counterbalance. Comparative studies indicate that civic actors have played a key role in advocating for human rights and democratic

oversight in transition contexts (Rakar & Kolarič, 2020, p. 132). Their involvement in monitoring algorithmic governance, promoting digital literacy, and ensuring transparency in AI procurement processes may help mitigate risks. However, civil society organizations in the Western Balkans often lack resources and political access, which makes EU support and cross-border partnerships indispensable.

Ultimately, the challenge is not only technological but normative: AI governance must be designed to reinforce democratic legitimacy, not erode it. The Western Balkans' trajectory towards European integration will depend on whether AI becomes a tool of authoritarian consolidation or a vehicle for embedding democratic values in digital governance.

4. AI, DEMOCRACY, AND CIVIL SOCIETY IN THE WESTERN BALKANS

The relationship between artificial intelligence, democratic governance, and civil society in the Western Balkans can only be understood against the backdrop of what Florian Bieber (2018) has identified as "competitive authoritarianism" in the region. Despite formal commitments to democratization and European integration, governments across the Balkans often display authoritarian practices: the concentration of executive power, the erosion of independent media, and the instrumentalization of institutions. Within this fragile democratic environment, the deployment of AI technologies—whether in the field of digital governance, policing, or electoral management—raises profound questions. AI can reinforce transparency and citizen engagement, but it can also consolidate power asymmetries, particularly where oversight institutions and civil society lack sufficient autonomy or resources (Bieber, 2018, p. 340).

The risks of AI misuse are evident in Serbia, where the government has promoted a strategy of "digitalization" as part of its modernization agenda. On the one hand, investments in e-government services and digital infrastructure are often presented as evidence of alignment with EU standards. On the other, reports by watchdog organizations have highlighted concerns about the large-scale deployment of facial recognition systems in Belgrade, often without transparent regulatory frameworks or adequate privacy safeguards. In a country where the media environment is polarized and civil society organizations are frequently targeted by delegitimizing narratives, such technologies can easily be appropriated for political surveillance and social control rather than citizen empowerment. Civil society actors in Serbia have raised these issues before EU institutions, underlining the need for conditionality mechanisms to tie candidate status more explicitly to guarantees of digital rights (Bieber, 2018, p. 344).

Bosnia and Herzegovina presents a different but equally telling case. Here, the fragmented political system, marked by ethnic divisions and weak central authority, complicates the creation of unified governance frameworks. AI-driven technologies are still in their infancy, but international actors, including the EU and the United Nations Development Programme (UNDP), have begun supporting projects that use data analytics for public administration reform and anti-corruption monitoring. While such projects are promising, the lack of strong legal oversight mechanisms raises questions about accountability. Civil society organizations often operate in a contested political space, but they have nonetheless been crucial in advocating for transparency and pushing for stronger protections of personal data. This is consistent with earlier

research (Marsili, 2023b, p. 37), according to which emerging technologies can serve democratic ends only when embedded within a normative framework that privileges human rights and international legal standards over short-term political expediency.

Civil society, therefore, could play a decisive role in shaping the normative frameworks for AI across the region. Independent watchdogs, think tanks, and advocacy groups can demand transparency in the procurement of AI systems, ensure that impact assessments are carried out, and promote the inclusion of human rights clauses in contractual agreements with technology providers. The EU has already experimented with such approaches in its *Digital Services Act* and *Artificial Intelligence Act*, and extending these principles to candidate countries could help build "cognitive resilience" against both domestic authoritarian tendencies and foreign disinformation campaigns. Moreover, drawing on regional cooperation platforms, civil society actors can exchange best practices and establish cross-border monitoring initiatives, a necessity in a space where political transitions are interlinked and vulnerabilities are often exploited transnationally.

Ultimately, the democratic implications of AI in the Western Balkans cannot be separated from the broader struggle over the rule of law and European integration. If AI is adopted without proper safeguards, it risks deepening existing governance deficits, reinforcing executive dominance, and undermining fragile civic spaces. If, however, it is harnessed within transparent, participatory, and rights-based frameworks, it could empower citizens and strengthen democratic institutions. The challenge lies in ensuring that the latter path is chosen, and that civil society is not only a passive observer but an active architect of ethical AI governance in the region.

5. REGIONAL COOPERATION, EUROPEAN INTEGRATION, AND AI

The governance of artificial intelligence in the Western Balkans cannot be addressed in isolation from the region's European integration process. Since the 2003 Thessaloniki Agenda, the EU has repeatedly affirmed that the Western Balkans belong within the European family, provided they meet the standards of democracy, rule of law, and fundamental rights. Today, digital transformation and AI governance are increasingly becoming part of these standards, as reflected in the EU's *Digital Agenda for the Western Balkans* (2018) and the AI Act. For aspiring member states, alignment with these frameworks is not only a technical requirement but also a political signal of their readiness to embrace European norms.

Regional cooperation plays a key role in this context. Initiatives such as the Regional Cooperation Council (RCC) and the Berlin Process have provided platforms to coordinate digital strategies and build capacity for technological governance. For instance, the RCC has launched programs promoting digital literacy and cross-border e-services, thereby laying the groundwork for harmonized AI adoption. Yet, as several scholars have pointed out, regional cooperation in the Western Balkans often suffers from a lack of institutionalization and political will, with states prioritizing bilateral disputes or domestic political gains over collective progress (Bieber, 2018, p. 349). This structural weakness risks creating fragmented regulatory environments that could be exploited by external actors, including global technology firms and non-EU powers.

The EU's conditionality mechanisms remain a double-edged sword. On one hand, they create strong incentives for candidate states to harmonize their digital governance frameworks with EU acquis. On the other, the slow pace of accession negotiations and perceived double standards undermine the credibility of

conditionality, sometimes fueling nationalist narratives that frame EU requirements as externally imposed constraints. In this environment, the integration of AI regulation into accession negotiations could either accelerate reforms or entrench skepticism, depending on how it is implemented. Civil society once again becomes crucial here, not only to monitor compliance but also to communicate to citizens the democratic and practical benefits of aligning with EU digital norms.

Recent scholarship highlights that AI governance in transitional societies must address both the opportunities of technological modernization and the risks of entrenching existing political pathologies. It is emphasized in earlier research (Marsili, 2025a, p. 214) that digital technologies, including AI, carry a dual potential: they can serve as instruments of emancipation, enhancing democratic participation, or as tools of domination, enabling surveillance and information manipulation. This ambiguity is particularly acute in the Western Balkans, where unresolved statehood issues, ethnic divisions, and fragile institutions intersect with the diffusion of digital technologies. Embedding AI within the broader framework of European solidarity—described in earlier research as a "geo-strategic challenge" of integration (Marsili, 2025a)—may therefore provide not only a normative anchor but also a geopolitical imperative.

The cross-border dimension of AI is equally important. Disinformation campaigns, cyber operations, and algorithmically amplified hate speech often spill across national boundaries, exacerbating tensions in a region already prone to nationalist rhetoric. Coordinated regional responses, supported by EU institutions, could mitigate these risks by establishing joint monitoring mechanisms, shared databases of AI applications, and regional oversight bodies. Such approaches resonate with recent calls in European security studies for a more "holistic" understanding of security that integrates human rights, digital resilience, and democratic accountability (Marsili & Hughes, 2024, pp. 21–22).

In conclusion, regional cooperation and European integration provide both opportunities and constraints for AI governance in the Western Balkans. If approached strategically, they can transform AI into a driver of democratic consolidation and cross-border trust-building. If neglected, however, they risk entrenching fragmentation, dependency, and vulnerability to authoritarian appropriation of digital technologies.

6. NATIONAL CASE STUDIES: AI GOVERNANCE IN SERBIA, BOSNIA AND HERZEGOVINA, AND ALBANIA

The trajectory of artificial intelligence governance in the Western Balkans reflects not only regional trends but also the specific political, institutional, and socio-economic realities of each state. Examining Serbia, Bosnia and Herzegovina, and Albania offers insights into how differing degrees of institutional strength, EU integration, and civil society engagement shape the ethical and democratic implications of AI adoption.

Serbia presents one of the most paradoxical contexts in the region. On the one hand, it has positioned itself as a regional leader in digital transformation, launching the Digital Serbia Initiative and promoting ICT as a driver of economic modernization. On the other hand, Serbia has faced increasing criticism for democratic backsliding, media capture, and weak rule of law (Bieber, 2020, p. 13). This duality raises serious concerns regarding AI adoption. Without robust safeguards, AI-enabled technologies may be used to reinforce executive dominance, expand state surveillance, and restrict media freedom.

Furthermore, Serbia has developed close ties with non-EU actors, particularly China, through initiatives such as the Safe City surveillance systems implemented by Huawei in Belgrade. These projects raise ethical concerns regarding data privacy, transparency, and accountability, especially in a country with limited independent oversight institutions. While EU accession remains formally on the agenda, the reliance on non-EU digital infrastructures risks undermining Serbia's regulatory alignment with European standards.

Bosnia and Herzegovina illustrates the challenges of adopting AI in a highly fragmented and ethnically divided political system. With two entities, ten cantons, and a weak central government, policy coordination in digital transformation is often inconsistent and politically contested. This institutional fragmentation has hindered the development of a coherent national AI strategy.

Civil society and international actors, including the EU Delegation and UNDP, have attempted to fill this gap by supporting digital governance reforms and pilot projects in e-governance. However, the lack of political consensus and the persistence of ethnonationalist rivalries limit the transformative potential of such initiatives. Earlier research (Marsili, 2023b, p. 34) suggests that emerging disruptive technologies may amplify pre-existing vulnerabilities when governance institutions lack both coherence and legitimacy. In Bosnia and Herzegovina, AI could either reinforce existing ethnic divisions—through the manipulation of online discourse and disinformation—or provide opportunities for cross-entity cooperation if framed as a neutral tool for modernization.

In contrast, Albania has demonstrated stronger reform momentum in the domain of digital transformation. Its *Digital Agenda 2022–2026* explicitly references the integration of AI and data-driven governance into public administration reforms, with a focus on transparency and service delivery. Albania's rapid progress in digitizing public services through the e-Albania platform has been lauded by the European Commission as a significant step toward modernization (European Commission, 2022, p. 45).

Nevertheless, challenges remain. Limited institutional capacity, high levels of corruption, and insufficient data protection frameworks risk undermining public trust in AI-driven governance. At the same time, Albania's stronger alignment with EU digital policies and its clear commitment to accession provide a framework within which reforms can be consolidated. This distinguishes Albania from Serbia, where digital modernization is often decoupled from democratic reform, and from Bosnia, where political fragmentation stalls systemic progress.

The comparison highlights three distinct trajectories. Serbia demonstrates how AI adoption can proceed in the absence of democratic safeguards, with risks of authoritarian appropriation. Bosnia and Herzegovina underscores the difficulties of fragmented governance, where institutional weakness impedes coherent AI strategies. Albania shows the potential of EU-driven reform momentum, though its progress remains fragile in the face of corruption and limited administrative capacity.

Across these cases, the ethical governance of AI depends not only on technical capacity but also on political will, institutional coherence, and the credibility of European integration as a transformative anchor. Earlier joint research (Marsili and Hughes, 2024, p. 21) observes that security and human rights frameworks must evolve together if digital technologies are to serve democratic rather than authoritarian ends.

7. CIVIL SOCIETY AND NON-STATE ACTORS IN AI GOVERNANCE IN THE WESTERN BALKANS

While state institutions remain the primary arena for setting policies on digital transformation and AI, the role of civil society, academia, and private sector actors is increasingly important in shaping the ethical, political, and social outcomes of AI adoption in the Western Balkans. In contexts characterized by fragile institutions and limited democratic accountability, these non-state actors often fill governance gaps, provide oversight, and mobilize societal debates on technology and rights.

Civil society organizations (CSOs) in the region have traditionally played a crucial role in democratization processes, monitoring elections, and advocating for human rights. In recent years, their mandate has expanded to include digital rights, data protection, and freedom of expression online. For instance, in Serbia, organizations such as Share Foundation have actively documented cases of digital surveillance and advocated for stronger data protection legislation. These initiatives underscore the potential of CSOs to act as watchdogs against the misuse of AI in surveillance or political manipulation.

However, civil society faces structural obstacles, including political pressure, limited funding, and shrinking civic space. Bieber (2020, p. 67) notes that governments across the region have used legal and financial instruments to weaken independent NGOs, thereby reducing their capacity to engage effectively in digital governance debates. This raises concerns that AI governance could evolve as a predominantly state-controlled domain, with minimal participatory input.

Universities and research institutes in the Western Balkans, often supported by EU-funded programs such as Horizon Europe or Erasmus+, have become important sites for research on AI ethics, digital governance, and cybersecurity. These institutions not only produce technical expertise but also contribute normative frameworks, often aligned with EU standards. In Albania, for example, the University of Tirana has partnered with European counterparts to integrate AI ethics into curricula, creating new epistemic communities that promote democratic standards in technological adoption.

The need for interdisciplinary approaches to emerging disruptive technologies, particularly those that cross the boundaries between law, security, and ethics, is emphasized in earlier research (Marsili, 2023b, p. 28). Such epistemic communities can serve as crucial intermediaries between policymakers, international organizations, and local stakeholders, helping to adapt global AI debates to regional specificities.

The private sector, particularly tech companies and start-ups, also plays an important role in AI adoption. While local companies in the Western Balkans remain relatively small, international actors—including multinational corporations and Chinese digital infrastructure providers—exercise disproportionate influence. This dynamic is particularly visible in Serbia, where Chinese companies such as Huawei provide AI-driven surveillance systems. In contrast, Albania's closer alignment with EU partners has encouraged partnerships with European firms promoting digital services under GDPR-compliant standards.

International organizations, notably the European Union, Council of Europe, and NATO, act as both funders and norm-setters. Through accession negotiations, monitoring reports, and technical assistance, the EU in particular has become a key driver of digital governance reforms. The interplay between EU pressure and local civil society advocacy creates a normative environment where ethical concerns about AI—

privacy, transparency, accountability—can be articulated, even if implementation remains uneven.

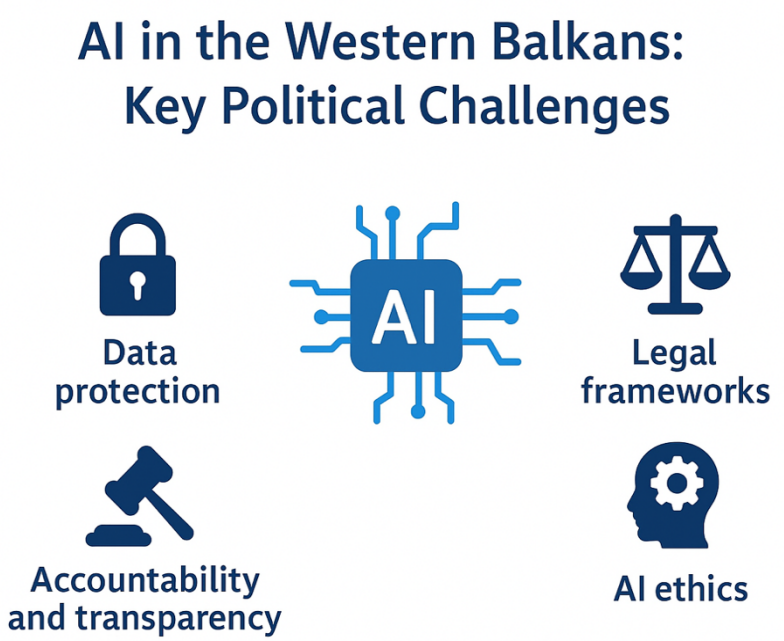
The involvement of civil society and non-state actors offers opportunities to democratize AI governance, ensuring that technological innovation aligns with human rights and democratic norms. However, this potential is contingent upon the preservation of civic space, academic freedom, and independent media. In authoritarian-leaning contexts, these actors risk marginalization or co-optation. Conversely, in reform-oriented settings such as Albania, they may become critical partners in bridging the gap between European digital standards and domestic realities.

As Marsili and Hughes (2024, p. 21) argue in the broader security domain, human security frameworks require inclusive, multi-actor participation to be effective. This logic applies equally to AI governance: without civil society engagement, governance risks being reduced to a technocratic exercise, detached from democratic accountability.

8. AI, SECURITY, AND DISINFORMATION: GEOPOLITICAL DIMENSIONS IN THE WESTERN BALKANS

The governance of artificial intelligence in the Western Balkans cannot be understood solely through a domestic lens. The region occupies a sensitive geopolitical position between the European Union, Russia, China, and Turkey, all of which project influence through economic, political, and technological channels. In this context, AI technologies intersect with issues of security, hybrid threats, and disinformation, reinforcing existing fragilities in governance and democratic institutions.

Figure 2. Geopolitical influence in AI and disinformation governance in the Western Balkans: external pressures from the EU, Russia, and China.



The Western Balkans have long been a fertile ground for disinformation campaigns, both domestically generated and externally driven. Research has documented how Russia, in particular, has deployed media influence and online propaganda to destabilize pro-European narratives and amplify nationalist or anti-Western sentiments (Miskimmon et al., 2017, p. 154). With the growing adoption of AI-driven content production and distribution, these campaigns acquire new capacities for scale, personalization, and emotional manipulation. Automated bots and generative AI tools make it increasingly difficult for citizens to distinguish between authentic and fabricated content, deepening mistrust in institutions and the media.

It has been noted in earlier research (Marsili, 2021, p. 162) that Russian influence strategies in contested neighborhoods often rely on information warfare to erode societal cohesion and undermine integration projects.In the Western Balkans, such tactics intersect with pre-existing ethnic and political divisions, raising the risk that AI-driven disinformation could exacerbate polarization and hinder democratic consolidation.

The securitization of AI technologies is another dimension shaping governance in the region. Governments facing weak legitimacy and societal fragmentation may be tempted to deploy AI for surveillance and control, framing such measures as tools for security and stability. The Serbian government's collaboration with Chinese firms to install "safe city" surveillance systems powered by facial recognition is emblematic. While these projects are justified in terms of crime prevention, they raise profound ethical concerns about privacy, proportionality, and democratic oversight.

The tension between security and liberty has a broader geopolitical dimension: Chinese technology often comes bundled with infrastructure investments through the Belt and Road Initiative, while EU counterparts promote frameworks aligned with human rights standards. This creates a dual dependency in which Western Balkan states navigate between competing technological and normative orders. Earlier research (Marsili, 2023b, p. 38) notes that emerging disruptive technologies pose not only legal but also ethical challenges, particularly when adopted in fragile democratic contexts without robust accountability mechanisms.

Disinformation campaigns in the Western Balkans have a direct impact on electoral politics and the credibility of European integration. During electoral cycles, AI-enhanced disinformation can shape public opinion, delegitimize opponents, and sow doubt about EU commitments. Bieber (2020, p. 114) underlines that illiberal actors within the region already exploit media capture and information manipulation; AI-based disinformation merely intensifies these dynamics, enabling faster and more targeted dissemination of polarizing narratives.

At the same time, the EU has recognized the vulnerability of the Western Balkans to hybrid threats. Initiatives such as the EU Hybrid Fusion Cell and the StratCom Task Forces aim to monitor and counteract disinformation, while candidate countries are encouraged to align with the EU Digital Services Act and cybersecurity frameworks. Yet, implementation remains partial, and the gap between EU standards and local practice persists.

Given the transnational nature of AI-driven security challenges, isolated national responses are insufficient. Regional cooperation, including through the Regional Cooperation Council (RCC) and NATO-supported platforms, is crucial to counter disinformation, promote digital resilience, and strengthen democratic oversight of AI technologies. Earlier joint research (Marsili and Hughes, 2024, p. 21) argues that

effective responses in the field of human security require multi-actor and multi-level engagement. This applies with even greater force in the AI domain, where the blurring of civil, military, and technological boundaries demands integrated governance frameworks that combine security imperatives with the protection of fundamental rights.

9. POLICY IMPLICATIONS AND RECOMMENDATIONS

The governance of Artificial Intelligence (AI) in the Western Balkans must be approached not only as a technical or regulatory challenge, but as a matter of political stability, democratic consolidation, and regional integration. As the European Union moves forward with a comprehensive legal framework for AI—the AI Act, complemented by the *Digital Services Act* and the *General Data Protection Regulation* (GDPR)—the Western Balkans face the dual task of aligning with this acquis while addressing domestic vulnerabilities. The "Brussels effect" (Bradford, 2020, p. 4) suggests that EU norms radiate beyond its borders, creating incentives for candidate and potential candidate countries to adopt European standards even before formal accession. This dynamic is particularly relevant for AI governance, where early adaptation can enhance both market access and political credibility.

Policy recommendations emerge at multiple levels. At the institutional level, Western Balkan states should establish national AI councils or inter-ministerial committees tasked with coordinating digital policy, ensuring interoperability with EU frameworks, and promoting accountability. These bodies should integrate civil society actors, as research shows that non-state participation increases transparency and legitimacy in transitional democracies (Kmezic, 2019, p. 117). At the regional level, the creation of a Western Balkans Task Force on AI Governance and Ethics, modeled on NATO's cooperative security arrangements, could foster peer learning, resource pooling, and resilience against disinformation campaigns.

Another crucial dimension is capacity-building. The Western Balkans continue to suffer from brain drain and weak institutional infrastructures; without targeted investment in education and digital literacy, AI systems risk being imported and deployed without adequate oversight. It is emphasized in earlier research (Marsili, 2023b, p. 39) that resilience in the digital domain depends not only on regulation, but also on embedding democratic values and human rights at the design and procurement stage of technological infrastructures. Strengthening universities, supporting fact-checking networks, and creating regional hubs of excellence in AI ethics can help mitigate dependency on external actors.

Finally, geopolitical alignment must be considered. As Serbia, Bosnia and Herzegovina, and Albania navigate relations with both Brussels and external actors such as Russia and China, AI governance becomes part of a broader strategic choice. Transparent, rights-based AI policies aligned with the EU could serve as a safeguard against authoritarian influence, reinforcing democratic trajectories while accelerating integration into European structures.

In sum, policy approaches to AI governance in the Western Balkans should not be limited to compliance with EU standards, but should aim at building cognitive resilience, empowering civil society, and securing democratic governance against external and internal threats. This requires a delicate balance between regulation, participation, and security—an approach that situates AI governance at the core of the region's European future.

CONCLUSION

The trajectory of Artificial Intelligence governance in the Western Balkans reveals the complexity of introducing disruptive technologies into societies marked by democratic fragility, contested institutions, and geopolitical vulnerability. From the theoretical perspective, the ethical dilemmas of AI are inseparable from broader questions of legitimacy, accountability, and the preservation of fundamental rights. As noted in critical scholarship on technology and governance, the challenge lies not only in regulating machines, but in safeguarding the moral and political fabric of democratic life itself.

The regional context underscores this tension. Serbia, Bosnia and Herzegovina, and Albania—while at different stages of political consolidation and European integration—share similar vulnerabilities: weak institutional capacity, exposure to disinformation, and persistent risks of state capture. The empirical analysis demonstrates how the uneven adoption of digital infrastructures intersects with governance gaps, producing both opportunities for modernization and risks of democratic backsliding. Civil society and non-state actors emerge as indispensable intermediaries, ensuring that technological governance is not monopolized by political elites but remains responsive to societal needs.

At the same time, the transnational dimension cannot be ignored. AI governance in the Western Balkans is conditioned by external pressures, from the "Brussels effect" of EU regulation to the geopolitical influence of Russia and China, which often promote alternative technological and normative models. In this sense, AI is not only a domestic governance issue but a geopolitical battleground where competing visions of sovereignty, democracy, and human rights collide.

The policy implications are clear. Western Balkan states must not reduce AI governance to a technical compliance exercise; rather, they must embed democratic values into the very architecture of AI deployment. This means building resilient institutions, investing in education and digital literacy, fostering regional cooperation, and ensuring meaningful participation of civil society. Alignment with European frameworks is essential, but so too is cultivating local ownership and societal trust.

Ultimately, the integration of AI into governance in the Western Balkans offers both risks and opportunities. If harnessed within an ethical, participatory, and rights-based framework, AI can enhance transparency, accountability, and regional cooperation—contributing to the democratic consolidation and European future of the region. If neglected or captured by authoritarian interests, it risks amplifying inequalities and eroding fragile democratic gains. The choice is therefore not merely technological but profoundly political: whether the Western Balkans will use AI as a tool for emancipation or as an instrument of control.

BIBLIOGRAPHY

Bieber, Florian (2018), *Patterns of competitive authoritarianism in the Western Balkans*, in *East European Politics*, vol. 34, no. 3, p. 337- 354. https://doi.org/10.1080/21599165.2018.1490272.

Bradford, Anu (2020), *The Brussels Effect: How the European Union Rules the World*, Oxford University Press, https://doi.org/10.1093/oso/9780190088583.001.0001.

Charter of Fundamental Rights of the European Union, *OJ C 326, 26.10.2012, pp. 391–407*.

Crawford, Kate (2021), *The Atlas of AI*, Yale University Press.

European Commisison (2018, 25 June), *European Commission launches Digital Agenda for the Western Balkans*, https://ec.europa.eu/commission/presscorner/detail/en/ip_18_4242.

European Commission (2022, 12 October), *Albania 2022 Report*, European Commission, Brussels, https://neighbourhood-enlargement.ec.europa.eu/albania-report-2022_en

Floridi, Luciano (2019), *The logic of information: A theory of philosophy as conceptual design*, Oxford University Press.

Jobin, Anna; Marcello Ienca; Effy Vayena (2019), *The global landscape of AI ethics guidelines*, in *Nature Machine Intelligence, vol.* 1, no. 9, p. 389–399, https://doi.org/10.1038/s42256-019-0088-2.

Kmezic, Marko (2019), *Rule of Law and Democracy in the Western Balkans: Addressing the Gap Between Policies and Practice*, Foundation BFPE for a Responsible Society, Belgrade.

Marsili, Marco (2021), *The Russian influence strategy in its contested neighbourhood*, in Holger Mölder, Vladimir Sazonov, Archil Chochia and Tanel Kerikmäe (coord.), *The Russian Federation in Global Information Warfare*, Springer, Cham, p. 149–172, https://doi.org/10.1007/978-3-030-73955-3_8.

Marsili, Marco (2023a), *Inside and beyond the Russo-Ukrainian War: The Pitfalls of the European Union*, in *Newsletter Annual of the Academy of Yuste: Reflections on Europe and Ibero-America, Vol. 3, Year 2022*, Fundación Academia Europea e Iberoamericana de Yuste, Cuacos de Yuste, p. 429-445, https://doi.org/10.5281/zenodo.8075295.

Marsili, Marco (2023b), *Military Emerging Disruptive Technologies: Compliance with International Law and Ethical Standards*, in Ignas Kalpokas & Julia Kalpokienė (coord.), *Intelligent and Autonomous: Transforming Values in the Face of Technology*, Brill, Leiden, p. 25–50, https://doi.org/10.1163/9789004547261_004.

Marsili, Marco (2025a), *Libertà di pensiero nell'era digitale*, Edizioni dell'Università Popolare (Edup), Roma.

Marsili, Marco (2025b), *Dal campo di battaglia al metaverso: la nuova frontiera delle operazioni militari*, Libreriauniversitaria.it Edizioni, Limena, https://doi.org/10.5281/zenodo.15361100.

Marsili, Marco (2025c), *The Polar Silk Road: A Euro-Sino Economic, Political and Geo-strategic Challenge*, in Teresa Freixes Sanjuán & Juan Carlos Moreno Piñero (coord.), *La Unión Europea y el Principio de Solidaridad / The European Union and the Principle of Solidarity*, Peter Lang, Brussels, p. 301–329, https://doi.org/10.3726/b22493.

Marsili, Marco and David Hughes, (2024), *Legal Framework Supporting Human Security*, in *Human Security: Frameworks, Concepts, Actors and Challenges in Relevance to NATO*, NATO Science and Technology Organization (STO), Paris, p. 19–22, https://doi.org/10.5281/zenodo.10982437.

Nemitz, Paul (2018), *Constitutional democracy and technology in the age of artificial intelligence*, in *Philosophical Transactions of the Royal Society A,* vol. 376*,* no. 2133, p. 20180089, https://doi.org/10.1098/rsta.2018.0089.

Rakar, Tatjana, and Zinka Kolarič (2020), *The Role of Civil Society Organisations in the Slovenian Welfare System during the Transition Period after 1990*, in Comparative Southeast European Studies, vol. 68, no. 2, p. 130–147, https://doi.org/10.1515/soeu-2020-0012.

Regional Cooperation Council (RCC), *Digital Trasformation*, https://www.rcc.int/flagships/7/digital-transformation.

Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC (General Data Protection Regulation). Consolidated text: http://data.europa.eu/eli/reg/2016/679/2016-05-04.

Regulation (EU) 2024/1689 of the European Parliament and of the Council of 13 June 2024 laying down harmonised rules on artificial intelligence and amending Regulations (EC) No 300/2008, (EU) No 167/2013, (EU) No 168/2013, (EU) 2018/858, (EU) 2018/1139 and (EU) 2019/2144 and Directives 2014/90/EU, (EU) 2016/797 and (EU) 2020/1828 (Artificial Intelligence Act), OJ L, 2024/1689, 12.7.2024, http://data.europa.eu/eli/reg/2024/1689/oj.

Regulation (EU) 2022/2065 of the European Parliament and of the Council of 19 October 2022 on a Single Market For Digital Services and amending Directive 2000/31/EC (Digital Services Act), *OJ L 277, 27.10.2022, pp. 1–102*, http://data.europa.eu/eli/reg/2022/2065/oj.

Zuboff, Shoshana (2019), *The age of surveillance capitalism*, PublicAffairs, New York.