The Digital Panopticon of Artificial Intelligence: between Safety and Freedom

Gerardo Antonio Corral Ruiz

Alma Mater Studiorum Università di Bologna

Abstract

Artificial Intelligence (AI) represents a world-changing technology; however, thinkers from Nick Bostrom to Yuval Noah Harari have highlighted the imminent risks associated with its advancement. The concept of the Panopticon, introduced by Jeremy Bentham, is a type of architecture which objective is of modifying and regulating prisoner behavior. This concept was later explored by Michel Foucault, who warned us of its potential application to the whole society as institutions increasingly adopted similar models of surveillance. Today, the EU and other global entities raise concerns because of the impact of AI on fundamental human rights, particularly speaking of privacy and freedom. While the benefits of AI in public security are substantial, the associated risks are equally intense. This paper analyzes current AI models developed for public security, examining their flaws and possible instances of injustice. Through a review of existing literature, this work seeks to propose regulatory and policy solutions to address these pressing challenges.[1]

Introduction

First let's use the definition of the Regulation of the EU 2024/1689 to understand what's the most important characteristic of an AI system for my analysis: "*A key characteristic of AI systems is their capability to infer. This capability to infer refers to the process of obtaining the outputs, such as predictions, content, recommendations, or decisions, which can influence physical and virtual environments, and to a capability of AI systems to*

[1] Disclaimer: For the purposes of English improvement and translation, I used DeepL in certain parts of this paper as a supportive tool.

derive models or algorithms, or both, from inputs or data.".[2] I believe this is important because AI's ability to directly interfere with our way of living as individuals poses a potential risk to our human rights and guarantees.

As we see in the Regulation of the EU 2024/1689 suggests there are four types of risks about AI:

- ***Unacceptable risks***: Systems that represent a clear threat to security, human rights and dignity. For example, the use of real-time facial recognition without proper legal authorization.[3]
- ***High risks***: Critical applications, such as those used in transportation, criminal justice, public safety, or administration. These systems must adhere to strict requirements for transparency and risk mitigation.
- Limited risks: Systems with potential transparency issues, such as chatbots, must inform users that they are interacting with a machine.
- Minimal risks: Systems without a meaningful impact, such as antispam programs or videogames.

That said, it's understandable that managing the use of AI involving unacceptable risks is reserved for government institutions, with very few exceptions.[4] A clear example of a high-risk AI issue came to light in 2018 with

[2] Regulation (EU) 2024/1689, paragraph (12) of the considerations of the text.

[3] I'll elaborate on this later, as it is essential to address the core issue of the paper. Anyway, to dive more into this subject it is useful to check the Article 5 of the EU Regulation 2024/1689. Strict exceptions apply only to showcased in the Annex II such as: "*terrorism, trafficking in human beings, sexual exploitation of children, and child pornography, illicit trafficking in narcotic drugs or psychotropic substances, illicit trafficking in weapons, munitions or explosives, murder, grievous bodily injury, illicit trade in human organs or tissue, illicit trafficking in nuclear or radioactive materials, kidnapping, illegal restraint or hostage-taking, crimes within the jurisdiction of the International Criminal Court, unlawful seizure of aircraft or ships, rape, environmental crime, organized or armed robbery, sabotage, participation in a criminal organization involved in one or more of the offences listed above*.".

[4] One of the exception mentioned in the Regulation (EU) 2024/1689, paragraph (24) of the considerations of the text: "*Nonetheless, if an AI system developed, placed on the market, put into service or used for military, defence or national security purposes is used outside those temporarily or permanently for other purposes, for example, civilian or humanitarian purposes, law enforcement or public security purposes, such a system would fall within the scope of this Regulation*.". The other exception is showed in the next paragraph (25) of the considerations of the text with some particular guarantees: "*It is therefore necessary to exclude from its scope AI systems and models specifically developed and put into service for the sole purpose of scientific research and development*.".

Amazon's recruitment system.[5] This AI, designed to evaluate resumes and recommend candidates, showed a significant bias against women because it was trained on historical hiring data that mirrored the tech industry's gender disparities. As a result, the AI penalized resumes that included references to terms like "women's clubs" and disproportionately favored male-dominated professional profiles. To avoid this scenario the EU wrote in the Regulation (EU) 2024/1689 Article 27 a criteria of valuation of the impact of the AI linked to the fundamental rights of people, and in the Article 56 it's established a Code of Practices.

On the other side there are the 2019 Ethics guidelines for trustworthy AI developed by the independent AI HLEG appointed by the Commission and that is still used by the EU for guidance in this regard. These are seven non-binding ethical principles, these are: human agency and oversight; technical robustness and safety; privacy and data governance; transparency; diversity, non-discrimination and fairness; societal and environmental well-being and accountability.[6]

Now, in order to develop and deploy an AI system categorized as high-risk, the EU has established a standardized protocol to ensure the safety of European citizens.[7] Providers must follow a series of steps: first, the system undergoes a conformity assessment to meet the AI requirements, often involving a notified body. Second, the system is registered in an EU database for standalone AI systems. Third, a declaration of conformity is signed, and the system receives a CE marking, allowing it to enter the market. Finally, providers must continuously monitor the system after commercialization and report any significant incidents or malfunctions, ensuring compliance throughout the AI's lifecycle.

In this sense, the EU has introduced some tools to keep data sharing legal and privacy focused. The Data Governance Regulation (DGA) supports secure reuse of data, especially where trade secrets or intellectual property are involved, by establishing neutral data intermediaries. The General Data Protection Regulation (GDPR)[8] sets core rules for handling personal information, including consent and

[5] Cf. Dastin, 2018. This case illustrates the critical need for regulatory frameworks.
[6] Cf. EU Regulation 2024/1689, (27) of the considerations of the text.
[7] As shown by Nicodemo, 2024.
[8] Either way, particularly in the context of AI, the EU Regulation 2024/1689 suggests checking Regulation 2016/679 and Regulation 2018/1725 as well.

the right to be forgotten which I will profound later in the next section. Finally, specifically Italy has its own which is more specific: *Garante per la Protezione dei Dati Personali* (GPDP) enforces these rules nationally, ensuring that data practices remain transparent.

The main objective of this paper, as I explained in the abstract, is to analyze the risks and opportunities that arise from introducing Artificial Intelligence (AI) into public security matters, as well as to examine the feasibility of its implementation within the EU legal framework. This approach will be fundamentally humanistic, aiming to anticipate and prevent the dangers of pervasive surveillance and control that philosophers such as Michel Foucault have warned about. In order to accomplish this, I will first clarify some concepts that will be referenced repeatedly throughout this paper.

Two philosophical issues

The first issue I identify is related to the concept of the Panopticon. In the 18th century, Jeremy Bentham (an English philosopher, jurist and social reformer) created the concept of the Panopticon. It was designed as a prison that let the police officers to watch any prisoner at any time without letting them know that they're being observed. Jeremy Bentham goal was to obtain: the reformed morality, the preserved health, the invigorated industry, the spread of education, the reduced public offices, the strengthened economy— all thanks to a simple architectural idea.[9] The concept that was initially applied in prisons was later introduced to other sectors, where it eventually became problematic. As Michel Foucault later expressed in his most iconic text *Discipline and Punish*, the Panopticon became a powerful tool for the widespread and invisible control of human behavior.

Foucault sought to illustrate how modern power structures in our societies tend to operate similarly to the Panopticon. He argued that the consequence of this is that we have individuals that are subjected to perpetual surveillance, and this is leading them to internalize discipline and self-regulate their actions. According to Foucault, the Panopticon is not just a merely architectural design, but a mechanism of power that enables control in society and that can modify our behavior.

[9] Cf. Bentham, 1791.

Fortunately, the EU Regulation 2024/1689 addresses these concerns and explicitly prohibits them in its paragraph (29) of the considerations of the text.[10]

Nowadays this is a well-known subject of conversation, and it is still really present in our lives, particularly in the realm of digital surveillance. The rise of data collection, analytics, and tracking technologies has turned the idea of a digital Panopticon into a key point in privacy debates. That's why organizations like the EU have implemented well-stablished data protection measures—like the GDPR—to defend our privacy and personal freedoms against the ever-growing reach of technology.

In the last centuries, some governments relied on surveillance and control to direct or repress their populations. This come to be obvious when we speak about totalitarian regimes such as the Nazis. That period showcased how propaganda, combined with strict oversight, can create a climate of conformity and suppress protest. As the Frankfurt School exposed, the nazis used to leveraged state-controlled media and widespread surveillance to impose their ideology, silence opposition, and abuse power on a large scale. The state's ability to observe citizens—whether through formal institutions or social pressure—enabled it to manipulate behavior, demonstrating the potency of a system that perpetually monitors its subjects.

Today, totalitarian practices still exist in some governments, but in Western societies at least, control and surveillance are increasingly shifting to the private sector. As we were introducing before, intellectuals associated with the Frankfurt School—particularly Max Horkheimer and Theodor W. Adorno in the *Dialectic of Enlightenment*—emphasized how this type of control extended beyond purely political contexts. In this book they showed how in some way capitalism and the emerging mass media (in the mid-20th century) empowered a subtler form of influence/control. Horkheimer and Adorno's notion of the "culture industry" highlighted how commercial interests harnessed—at least in the last century but I believed also nowadays—the power of film, radio, and popular entertainment to cultivate standardized tastes and desires among the public. By shaping cultural products according to market demands, media conglomerates could dictate users or

[10] Cf. EU Regulation 2024/1689, (29) of the considerations of the text.

consumers trends and reinforce the capitalist status quo, which I firmly maintain that remains relevant in the era of social media, although it evolved. Thus, the Frankfurt School's critique illustrates how totalitarian-like methods can persist even in seemingly free societies like the one we witness today.

Laval, C. on his work: *Discipline and Prevent: The New Panopticon Society* explains that the panopticon, has been reimagined as a societal model in the digital era. Unlike authoritarian regimes, where surveillance is always evident and oppressive, Laval argues that in democracies, it operates subtly, embedding itself into daily life through digital networks and data analytics. This quiet but pervasive surveillance can discourage democratic participation.

Social media platforms and advertising industries exhibit parallels to this Panoptic arrangement that Foucault criticized. By gathering vast quantities of user data, they can anticipate preferences and tailor marketing efforts, effectively guiding people's consumption habits. Through these practices, powerful entities benefit from the insights generated by constant virtual surveillance. Consequently, Foucault's Panopticon continues to serve as a compelling metaphor for the tension between individual autonomy and the sophisticated systems of control prevalent in the digital age.

Yet, in this matter the idea of integrating AI into public security systems reveals the dilemma of this work. On one hand, AI-driven methods—such as the ones we are going to profound later: predictive policing, facial recognition, and data analytics—offer promising solutions for crime prevention and more efficient law enforcement. They could help governments rapidly identify patterns, anticipate threats, and distribute resources more strategically which could potentially lead to safer communities. On the other hand, critics warn that such AI deployments may amplify the Panopticon's most troubling features. We cannot forget that this is the address of the prevention measures made by the EU as we were analyzing before. Automated surveillance technologies can concentrate power in the hands of a few while perpetuating biases encoded in their algorithms, if, for instance, they lack transparency.

As we may perceive by now, the tension between enhancing public safety and preserving individual freedom underlines the significance of robust regulatory frameworks. Instruments such as the GDPR serve as barricades against excessive

data collection and misuse, demanding higher standards of accountability from both state agencies and private corporations. For example, the Italian *Decreto legislativo 18 maggio 2018, n. 51.* establish that the personal data needs to be: first, processed strictly and always by humans, not automated.[11] And second, the data gathering must follow always a previously presented purpose and it should only be the exact and necessary data, not more.[12] This is an example of how Italy implements concrete strategies to regulate its citizens' data. The countries in the EU believe that this instrument not only protects citizens but also enables interested parties to access transparent public data for research, guided by the principle of data altruism. In fact, the European Union's Data Governance Regulation (Regulation (EU) 2022/868) expands data availability while ensuring privacy and transparency. Unlike traditional open data initiatives, it sets rules for reusing data held by public entities that are protected by third-party rights, like trade secrets, personal data, or intellectual property.[13] The GDPR also impulses the "right to be forgotten," a mechanism that compels data controllers to erase personal data under specific conditions, and grants data subjects the right to data portability, allowing them to transfer their information between service providers.

These regulations and legal mechanisms act as counterweights to the emergence of a digital Panopticon. By clarifying explicit limits on surveillance, requiring procedural safeguards, and emphasizing individual autonomy over one's own data, they aim to preserve the space for personal freedom in an era of technological needs and customs. The GDPR and complementary initiatives channel the use of data in ways that safeguard civil liberties.

The second issue I wish to highlight extends beyond the Panopticon itself and focuses on the problem of allowing AI to govern public security autonomously. While the EU's GDPR defines strict parameters to prevent AI from making fully automated decisions about human subjects, this measure reveals why it was important to begin this paper with the Regulation (EU) 2024/1689 discussed in this paper's introduction which particularly speaking about public safety exposes the levels of unacceptable and high kind of risks.

---

[11] Cfr. *Decreto legislativo 18 maggio 2018, n. 51. Art. 1, 2.*

[12] Cfr. Ibid. *Art. 3, 1.*

[13] As we were discussing back in class.

A central complication involves the "black box" nature of many AI systems. Even developers often cannot fully account for how algorithms arrive at certain decisions, which renders extremely difficult to detect implicit biases or discriminatory outcomes, this is urgent to work on because inaccurate data or flawed assumptions can result in severe injustices. Compounding this issue is the notion that 2025 may be tagged by some as "the year of AI agents," suggesting that AI tools will become increasingly autonomous. Figures like Sam Altman predict that artificial general intelligence (AGI) could be achieved in the near future, effectively granting AI unprecedented capacity to learn, adapt, and act without direct human oversight. This echoes Nick Bostrom's warnings about superintelligent AI systems as shown in his book *Superintelligence*. In there he exposes that once we entrust machines with making decisions integral to our lives, our dependency on them may grow to the extent that humans lose agency—similar to how the survival of many animal species, such as gorillas, now largely depends on human conservation efforts and interests. Equally important for me is Yuval Noah Harari's argument in his book *Nexus*, where he underscores that since the dawn of the modern era, thinkers like have dreamed of controlling nature (like we may read as well in Francis Bacon essays), but getting back to the *Dialectic of Enlightenment* we may have a wrong focus in thinking that we need to achieve a full domain in things such as nature itself. Harari suggests that if we continue to develop AI as an autonomous agent, it could someday govern all significant decisions—ranging from career choices to healthcare—through opaque algorithmic networks that few will understand, effectively undermining individual freedom, and somehow bringing us back to a Panopticon kind of scenario in which the one that might control our lives would be the AI.

In this light, the Panopticon metaphor becomes even more relevant. AI-based surveillance, decision-making, and data processing have the potential to replicate and amplify the effects of being perpetually observed or managed. Hence, policymakers and innovators alike must approach AI development with caution. That said, now is a good time to analyze recent AI developments in public safety and compare them with the guarantees we aim to prevail.

Some current projects

In order to analyze the current landscape of AI applications for crime prediction and prevention, several key studies stand out I read some of them to understand the actual work. Many of these studies place a conceptual groundwork for neural networks and geospatial analytics to attack social dangers such as crime. There are different types of approaches:[14] those that analyze and predict instances of crime through recordings or documentation, those that create models for specific zones of a city and study crime as a geographical phenomenon, those that focus on individuals, examining demographic data and even their emotions, and a lot more. To start we are going to analyze one that does the first method mentioned, *AI in Crime Prediction and Prevention* by Choudhary et al. demonstrates the substantial power of surveillance-focused neural networks while also illustrating the dangers of algorithmic bias; its discussion of the COMPAS system—a predictive model notorious for embedded prejudices—underscores why transparency is paramount and why the "black box" nature of AI remains a critical ethical concern. In this text Choudhary et al. explained how networks such as Long Short-Term Memory (LSTM) are being used in the United Arab Emirates to analyze temporal and sequential crime data.[15] They note that LSTMs is that this kind of network can "remember" information across extended intervals. Choudhary et al. said that this allows users to store the information, not only the location and type of crime, but also the timing of events, for instance, which hour of the day or how frequently. These authors maintain that this kind of AI has a margin of success of 75%-90%, this shows a precise capacity to forecast where and when future crimes are most likely to happen, which is said that lets the police to allocate their resources more strategically.

As we may perceive we found the classic tension between freedom and safety. The Regulation (EU) 2024/1689 since its first paragraph of the considerations of the text mention that they want to ensure: "*a high level of protection of health, safety, fundamental rights*".[16] This is repeated constantly in the same text, as example in the

[14] *A survey on crime analysis and prediction,* helps to dive into different models and their functioning.
[15] Cf. Choudhary et al., 2020, p. 113. In their article Choudhary et al. actually mention that this technology was also being used in urban areas of Mexico.
[16] Regulation (EU) 2024/1689, (1) of the considerations of the text. Fundamental rights obviously such as freedom and privacy.

paragraph (6) of the considerations of the text where it is said that AI needs to be always built as a human-centric technology: "*It should serve as a tool for people, with the ultimate aim of increasing human well-being*.".[17] In this sense, we need to be aware of how the software shown in this text might or might not accomplish this goal.

We must carefully analyze the type of data being stored about individuals and closely monitor programs that use 'remote biometric identification systems,'[18] as these systems operate without active involvement from the people being monitored. This makes them part of an unacceptable risk category, obviously for the markets. The only exception lies in their use by national governments;[19] however, I find it highly unlikely that such systems could be openly deployed in the market under current regulations. Fortunately, there are alternative methods, which will be discussed later in this paper, that I believe provide opportunities to utilize AI in ways that are safe and respectful of human rights. Let us remember the first principle of Ethics guidelines for trustworthy AI: human agency and oversight means that AI systems are developed and used as a tool that serves people, respects human dignity and personal autonomy, avoiding control and excessive surveillance.

As *A Survey of Emotional Artificial Intelligence and Crimes* shows AI has advanced to the point where it no longer merely analyzes our behaviors as patterns or observes us visually. It can now interpret our emotions,[20] further amplifying the sense of "invisible control" described by Foucault in his concept of the panopticon. This evolution moves beyond monitoring physical actions to reading and interpreting emotional states, making the mechanisms of control even more intrusive and comprehensive. Emotional AI is not strictly forbidden, as long as it respects the human rights outlined in the Introduction of this paper and in the provisions of EU Regulation 2024/1689 mentioned earlier. However, Article 50 of EU Regulation 2024/1689 stipulates that it must operate under the supervision of high authorities, given its classification as a high-risk usage of AI. Considering that it is

[17] Ibid. (6) of the considerations of the text.

[18] Cf. Regulation (EU) 2024/1689, (17) of the considerations of the text.

[19] Cf. Ibid. As seen in the paragraph 24 of the considerations of the text: "*As regards national security purposes, the exclusion is justified both by the fact that national security remains the sole responsibility of Member States in accordance with Article 4(2) TEU and by the specific nature and operational needs of national security activities and specific national rules applicable to those activities*."

[20] It can be used as well to predict crimes, as showed in the table of the page 10 of *A Survey of Emotional Artificial Intelligence and Crimes*. But it still has some issues in detecting and predicting complex crimes.

intended for national security purposes, this technology lies between being classified as high-risk and unacceptable. It is most likely to be used exclusively by government entities, although there remains an open possibility if it is well-supervised. What makes it particularly interesting is its ability to detect people's emotions not only through physical expressions but also via vocal tones, physiological responses, and social media activity. I believe that, in the hands of law enforcement, this technology could assist in organizing their investigations and interviews more effectively. The objective of government safety sectors, as Gavin de Becker explains in *The Gift of Fear*, is to utilize technology for crime prevention rather than merely responding to it, the second one being a matter of justice, not safety. This would be an approach similar to the one taken in the text of Apene, O. et al.: *Advancements in Crime Prevention and Detection: From Traditional Approaches to Artificial Intelligence Solutions*.[21] In this context, I believe that AI offers new opportunities for enhancing public safety, always if it shows that it respects human rights and adheres to the guidelines established in the Code of Practice.

An alternative and viable approach to addressing insecurity while minimizing interference with individual privacy lies in the use of geographical clustering methods, at least I believe. Rather than focusing on specific individuals, these methods analyze zones based on statistical data. For instance, as demonstrated in *A Clustering Based Hotspot Identification Approach for Crime Prediction* by Hajela G. et al., geospatial methods consider factors such as time, weather, location, and demographic parameters as useful indicators for predicting crime in a given area. These latter issues (demographic parameters) pose challenges to privacy, as they may expose data such as annual income or literacy rates. Care must be taken to avoid categorizing individuals, as this is one of the prohibited practices outlined in EU Regulation 2024/1689. *Artificial Intelligence for Safer Cities: A Deep Dive into Crime Prediction and Gun Violence Detection* by Iqbal, A. et al. also employs the geographical approach but highlights the challenges of applying this method in larger cities, primarily due to technical issues related to data mining.[22] Nonetheless, this text also

[21] Cf. Apene, O. et al. 2024, p. 9.

[22] Cf. Catlett, C. et al. *Spatio-temporal crime predictions in smart cities: A data-driven approach and experiments*, also discusses the benefits and techniques of a geographical focus, offering an alternative approach to studying large cities such as Chicago and New York. Holding the fact that it is complex, but possible.

incorporates the use of individuals' social media activity to monitor their behavior, which reintroduces concerns about privacy and categorization.

Although a geographical approach might involve less categorization of individuals, it has yet to be connected to a more human-centered framework. Esposito, E. et al. emphasize in their work: *Algorithmic crime prevention. From abstract police to precision policing* the critical importance of implementing "precision policing." Bringing law enforcement closer to a local and community level not only strengthens public trust and cooperation with institutions but also makes the strategies more effective and less detached from the population. Furthermore, it helps guarantee fundamental rights by making citizens more active participants in the process. This approach would also enhance the use of AI by mitigating biases and enabling a more precise understanding of the demographics and needs of each area, ultimately making it more human-focused. Additionally, it reinforces the significance of geospatial models in crime prevention strategies. Though, I believe that a more human-centered approach to "precision policing" could facilitate the integration of AI into public security organizations in an organic way, ensuring compliance with the first principle of the Ethics Guidelines for Trustworthy AI mentioned earlier.

The work of Alves, L. et al.: *Crime prediction through urban metrics and statistical learning* discusses environments in which Quetelet's concept of "social physics" can be applied—a framework that employs principles and methods from statistical physics to analyze patterns in human behavior and social phenomena. It uses urban metrics function like "forces" that collectively influence the likelihood of criminal activity in specific areas. Factors such as unemployment or education levels may change the probability of crimes occurring. As the authors suggest, more extensive and objective metrics can lead to better accuracy but also pose huge risks of privacy like the ones that we have already reviewed.

Conclusion

Throughout this paper, the analogy of a "digital panopticon" has illustrated the dual nature of AI. On the one hand, AI opens up new frontiers for public safety, offering the potential for advanced crime detection and prevention strategies through new efficient methods never seen before in human history. On the other hand, it carries

significant risks for privacy, autonomy, and human rights as the EU and many other governments are currently concerned. Thinkers such as Foucault see that unchecked surveillance and predictive policing can gradually erode civil liberties, turning societies into environments that perpetually monitor and influence individual behavior.

Rasmussen's Cultural Visions of Technology: Paradoxes of Panoptic and Interactive Perspectives and Methods introduces a pivotal insight: technology exists in a state of ambivalence, simultaneously offering hope for human progress and posing serious risks to personal freedom. As he puts it, "*Cultural visions of technology, however, are still firmly rooted in images of control, prosperity and hope. But the reverse sides of these images are also present in form of metaphors of surveillance, exploitation and apocalyptic dreams of destruction and annihilation*" (Rasmussen L., 2013, p. 177). Central to Rasmussen's framework are two contrasting viewpoints: the Panoptic vision, rooted in Bentham's original prison design and further developed by Foucault, and the Interactive vision that pushes user empowerment and creative collaboration. The Panoptic vision is rigid, focusing on efficiency and control through constant observation, much like modern digital surveillance systems. In contrast, the Interactive vision views technology as a flexible, user-centric tool. Rasmussen links these visions to broader philosophical ideas, noting how the Panoptic worldview resonates with the mechanistic leanings of Newton and Descartes,[23] while in the meantime the Interactive perspective draws on concepts of complexity and relativity to emphasize adaptability, creativity, and interconnectivity.

This tension parallels Esposito's emphasis on "precision policing," a strategy grounded in local and community-level engagement. While a Panoptic system might optimize surveillance through big data analytics, it risks alienating those subjected to it. On the other hand, a more human-focused approach enhances public trust and fosters meaningful cooperation, aligning with the Interactive vision that values creativity and collaboration. Ultimately, Rasmussen's paradox underscores that neither purely Panoptic nor purely Interactive visions can suffice. Rather, a strategic synthesis—acknowledging the need for coordination and oversight while

[23] Cf. Rasmussen L., 2013, p. 179.

preserving personal freedom—is paramount. I believe we need to find this equilibrium as well.

In response, Europe's regulatory landscape demonstrates a constant effort to balance the benefits of innovation with the imperative of rights protection. These frameworks introduce mechanisms designed to prevent AI systems from becoming fully automated arbiters of human fate while safeguarding the rights of safety, forbidding unacceptable risks and implementing transparency requirements. I believe that the EU seeks to ensure that data-driven crime prevention does not descend into a form of digital authoritarianism.

A recurring theme in this discussion is the tension between freedom and safety. The work of authors like Choudhary et al. and Hajela G. et al. highlights the strengths of AI-driven predictive methods: LSTM networks, clustering algorithms, and other sophisticated models can significantly enhance resource allocation and crime risk assessment. Yet, we still face numerous risks that must be constantly evaluated, and we must try to land these projects in a more human-centered approach. In this context, Esposito's concept of "precision policing" serves as a powerful reminder of the importance of a more human-centered framework. I find it particularly interesting how this concept relates to the requirement for high-risk AI systems (primarily) to operate under the supervision of public authorities. To me, this also qualifies as a precision strategy.

Ultimately, this paper underscores that AI in public security must not be approached as a purely technical undertaking but as a profoundly humanistic endeavor. The capability of AI to predict and prevent crime is matched by its capacity to undermine privacy and autonomy if left unchecked.

Bibliography

- **Mills, J. L., & Bradley-Kennef, C. S.** (2023). *SURVEILLANCE AND POLICING TODAY: CAN PRIVACY AND THE FOURTH AMENDMENT SURVIVE NEW TECHNOLOGY, ARTIFICIAL INTELLIGENCE AND A CULTURE OF INTRUSION?* University of Florida Journal of Law & Public Policy, Vol 33, Issue 2, p183. ISSN 1047-8035.
- **Normattiva**: **Il Portale della legge vigente** (2018). *Decreto legislativo 18 maggio 2018, n. 51*
- **EU**. (2024). *Regulation (EU) 2024/1689.*
- **Bentham, J. (1791)** *Panopticon: The Inspection House***.** CreateSpace Independent Publishing Platform (8 October 2017). ISBN: 978-1978103917

- **Nicodemo, S.** (2024) *INTELLIGENZA ARTIFICIALE. (Shared with us in class)*
- **Esposito, E., & Egbert, S.** (2024) *Algorithmic crime prevention. From abstract police to precision policing*, Policing and Society, 34:6, 521-534, DOI:10.1080/10439463.2024.2326516.
- **Wang, H., & Ma, S.** (2022). *Preventing crimes against public health with artificial intelligence and machine learning capabilities*. Socio-Economic Planning Sciences, *80*. https://doi.org/10.1016/j.seps.2021.101043
- **Sahay, K. B., Balachander, B., Jagadeesh, B., Anand Kumar, G., Kumar, R., & Rama Parvathy, L.** (2022). *A real time crime scene intelligent video surveillance systems in violence detection framework using deep learning techniques*. Computers and Electrical Engineering, *103*, *108319*. https://doi.org/10.1016/j.compeleceng.2022.108319
- **Catlett, C., Cesario, E., Talia, D., & Vinci, A**. (2019). *Spatio-temporal crime predictions in smart cities: A data-driven approach and experiments*. Pervasive and Mobile Computing, 53, 62–74. https://doi.org/10.1016/j.pmcj.2019.01.003
- **Hajela, G., Chawla, M., & Rasool, A.** (2020). *A Clustering Based Hotspot Identification Approach for Crime Prediction*. Procedia Computer Science, *167*, 1462–1470. https://doi.org/10.1016/j.procs.2020.03.357
- **Apene, O. Z., Blamah, N. V., & Aimufua, G. I. O.** (2024). *Advancements in Crime Prevention and Detection: From Traditional Approaches to Artificial Intelligence Solutions*. European Journal of Applied Science, Engineering and Technology, 2(2), 285–297. https://doi.org/10.59324/ejaset.2024.2(2).20
- **Gouri, A., Choudhary, V., Meena, A., Kumar, P., & Saini, R.** (2024). *AI in Crime Prevention and Prediction*. Juni Khyat, 14(5), 113-120. Vivekananda Global University, Jaipur. ISSN: 2278-4632 https://doi.org/10.13140/RG.2.2.22509.60642
- **Ashly Thomas, N.V. Sobhana**, ***A survey on crime analysis and prediction***. Materials Today: Proceedings, Volume 58, Part 1, 2022, Pages 310-315, ISSN 2214-7853, https://doi.org/10.1016/j.matpr.2022.02.170
- **Alves, L. G. A., Ribeiro, H. v., & Rodrigues, F. A.** (2018). *Crime prediction through urban metrics and statistical learning*. Physica A: Statistical Mechanics and Its Applications, 505, 435–443. https://doi.org/10.1016/j.physa.2018.03.084
- **Rasmussen, L. B.** (2013). *Cultural visions of technology: Paradoxes of panoptic and interactive perspectives and methods*. AI and Society, *28*(2), 177–188. https://doi.org/10.1007/s00146-012-0408-0
- **Khoei, T. T., & Singh, A.** (2024). *A survey of Emotional Artificial Intelligence and crimes: detection, prediction, challenges and future direction*. Journal of Computational Social Science. https://doi.org/10.1007/s42001-024-00313-3
- **Iqbal, A., Zahid, S. B., & Arif, M. F.** (2021). *Artificial Intelligence for Safer Cities: A Deep Dive into Crime Prediction and Gun Violence Detection*. Punjab IT Board. https://www.researchgate.net/publication/375597613
- **Laval, C.** (2008-2009.). *Discipline and Prevent: The New Panopticon Society*. (This paper is a transcript of a speech given at conferences to which I was invited by Patrick Piguet at Lycée Buffon (Paris) in March 2008 and Lycée Sainte-Marie (Neuilly) in November 2009.)
- **Romanyshyn R. D.** (1989) *Technology as symptom & dream*. Routledge, London. ISBN 9780415007870
- **Good News Network**. (2024, October 27). *14-year-old wins 'America's Top Young Scientist' for inventing pesticide detector for fruits and vegetables*. Good News Network. https://www.goodnewsnetwork.org/14-year-old-wins-americas-top-young-scientist-for-inventing-pesticide-detector-for-fruits-and-vegetables/
- European Parliament. (2024). *Legislative resolution of the European Parliament of 14 June 2024 on the proposal for a regulation laying down harmonized rules on Artificial Intelligence (Artificial Intelligence Act) and amending certain Union legislative acts*. European Parliament. Available at https://www.europarl.europa.eu/doceo/document/TA-9-2024-0138_EN.pdf

- UNESCO. (2021). *Recommendation on the Ethics of Artificial Intelligence*. https://unesdoc.unesco.org/ark:/48223/pf0000380455_spa
- Dastin, J. (2018). *Insight - Amazon scraps secret AI recruiting tool that showed bias against women*. Reuters. https://www.reuters.com/article/world/insight-amazon-scraps-secret-ai-recruiting-tool-that-showed-bias-against-women-idUSKCN1MK0AG/