



“Regulating Artificial Intelligence for Surveillance” Background Guide

Chair: Kahini Gandhi

Co-Chair: Jiya Jaiswal

NOTE FROM THE CHAIR

Dear Delegates,

My name is Kahini Gandhi, and I am a junior at Wissahickon High School! This is my third year being a part of Model UN, and I am looking forward to being your chair for this committee.

At Wissahickon, I'm the co-president of our Model UN club and have attended the Washington Area Model United Nations Conference (WAMUNC). At WAMUNC, I took part in a general assembly called the "Commonwealth Games 2026," where I represented the United Kingdom. This took place right before Cherry Blossom season, so we got to see some blooming flowers before we left. We also had a "Fun MUN" before the Closing Ceremony, where we sang to a committee across the hall and witnessed people doing the Milk Mile. The friends I made there are some of the people I interact with regularly, and I hope you experience something similar during KNIGHTMUN!

Other than Model UN, at our school, I volunteer through Key Club, partake in Indian Cultural Association activities, and enjoy learning the basics of investing through the Young Investors Society. Outside of school, I learn the drums and serve in the Civil Air Patrol—a civilian auxiliary of the Air Force—where I am a cadet commander with the rank of Cadet 2nd Lieutenant. My hobbies include painting, photography, poetry, and reading. If you have any good reading recommendations, let me know!

My goal is to moderate this committee in a manner led by you, the delegates, and to provide you with a challenge. Our general assembly topic is "Regulating Artificial Intelligence for Surveillance," and I'm very excited to meet you all at the conference!

Sincerely,

Kahini Gandhi

USG of Publications & Media

KNIGHTMUN 2026

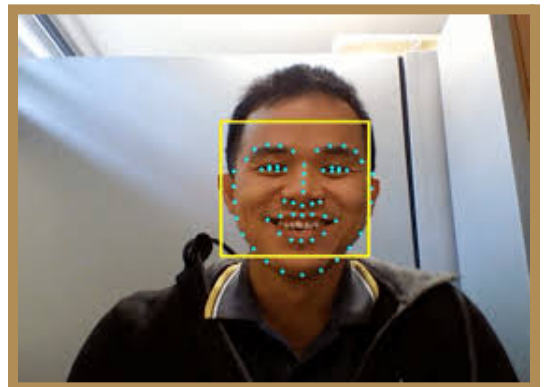
Introduction to AI Surveillance

Artificial intelligence (AI) refers to computer systems designed to perform tasks that normally require human thinking, like recognizing patterns, making decisions, or predicting outcomes. Modern AI does not work the way human brains do. Instead, it learns by processing enormous amounts of data and finding statistical patterns within it. Feed an AI system millions of photos of faces, and it will learn to distinguish one face from another. Feed it years of financial records, and it will learn to predict who is likely to default on a loan.

This process is called machine learning, and it is the engine behind almost every AI surveillance tool in use today. The key thing to understand about machine learning is that these systems are only as good as the data they are trained on. If the training data is biased – for example, if it contains far more photos of white male faces than of women or people of other ethnicities – the system's predictions will be less accurate for the groups that were underrepresented in training. This is one of the most documented and consequential flaws in AI surveillance systems deployed today.

Facial Recognition Technology

Facial recognition technology (FRT) works in three basic steps. First, a camera captures an image or video frame containing a human face. Second, the AI system analyzes the face and converts it into a unique numerical code, sometimes called a faceprint, based on measurable features like the distance between the eyes, the width of the nose, and the shape of the jawline. Third, that faceprint is compared against a database of stored faceprints to find a match.



This process can happen in two different ways.

One-to-one matching verifies that a person is who they claim to be, which is how your phone unlocks when it scans your face and how airports confirm that a traveler matches their passport photo. One-to-many matching is far more powerful and far more controversial: it takes a face captured in public and searches an entire database potentially containing millions of entries to identify who that person is. This is the form of facial recognition used in mass surveillance.

Modern FRT systems can process faces in real time, meaning a camera connected to a facial recognition system can identify everyone who walks past it as they walk past it. When networked across an entire city, as in parts of China, where hundreds of millions of cameras are in operation, this creates the ability to track any individual's movements throughout an entire urban area, reconstruct where they have been, and flag when they appear somewhere unexpected.

The accuracy of these systems has improved dramatically in recent years but remains

uneven across different populations. A landmark 2019 study by the U.S. National Institute of Standards and Technology (NIST) tested 189 facial recognition algorithms and found that most performed significantly worse on female faces, older faces, and the faces of Black and Asian individuals compared to white male faces. For some algorithms, the false positive rate incorrectly identifying someone as a match, was up to 100 times higher for Black women than for white men.

Social Scoring Systems

Social scoring uses AI to aggregate data about how individuals behave and translates that data into a numerical score that affects their access to opportunities and services. The concept draws on technologies that already exist in everyday life. Credit scores, for instance, are a simple form of behavioral scoring. Social scoring extends them far beyond financial history to encompass a much broader picture of a person's life.

China's Social Credit System is the most widely discussed example, though it is important to understand that it is not a single unified system. It is a collection of overlapping programs run by different city and provincial governments, as well as the national government, each with somewhat different rules and focuses. Some components track corporate compliance with regulations. Others track individual financial behavior. Some city-level programs have experimented with tracking things like traffic violations, court order compliance, or even how people behave on public transport. People flagged by these systems can face restrictions including being banned from purchasing airline or train tickets, being excluded from certain schools or government jobs, or having their names published on public blacklists.

The data feeding into social scoring systems typically comes from multiple sources combined together: government records, financial databases, court records, surveillance cameras, internet activity, and increasingly, data purchased or obtained from private companies. The AI system synthesizes all of this information to generate a score or classification. One of the most significant concerns about these systems is opacity, where individuals often cannot find out exactly what data was used to calculate their score, why it changed, or how to challenge a decision made on the basis of it.

How These Technologies Work Together

Facial recognition and social scoring are most powerful and most concerning when they operate together as part of an integrated surveillance infrastructure. A camera network using facial recognition can log every time a specific individual appears in a public space. That location and movement data can then feed into a social scoring system alongside data from other sources. The result is a system that does not just watch people but builds detailed behavioral

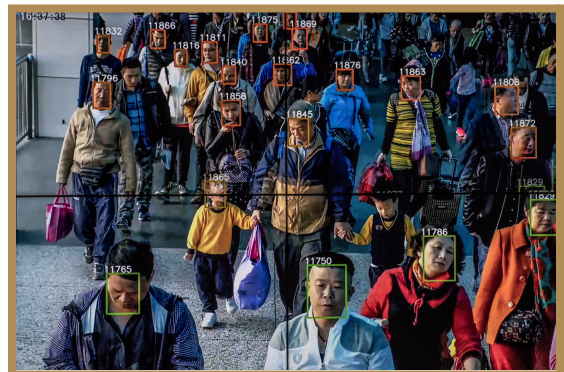
profiles over time, tracking not just what people do but predicting what they are likely to do and adjusting their access to services and opportunities accordingly.

This kind of integrated surveillance infrastructure has been described by some researchers as a fundamental shift in the nature of state power. Historically, governments could monitor some people some of the time. AI surveillance, at scale, makes it technically possible to monitor most people most of the time without requiring large numbers of human observers.

Why Does This Matter Now?

AI surveillance is not a future concern, but it is happening today, at a massive scale, in countries around the world. Consider these facts:

- As of 2024, facial recognition cameras are in use in public spaces in at least 109 countries, according to research by Comparitech.
- In the United States, multiple cities have documented cases where people were wrongly arrested based on faulty facial recognition matches. In nearly every documented case, the wrongly identified person was Black.
- Authoritarian governments have used AI surveillance tools to monitor and suppress ethnic and religious minorities. China's surveillance of the Uyghur Muslim minority in Xinjiang has been described by the UN as potentially constituting crimes against humanity.
- Companies based in democratic countries have sold surveillance technology to governments with documented records of human rights abuse, often with limited oversight or accountability.



At the same time, it would be inaccurate to say that all AI surveillance is bad or that governments have no legitimate reasons to use these tools. Law enforcement agencies argue that facial recognition has helped solve serious crimes, locate missing children, and prevent terrorist attacks.

International and Regional Framework

Universal Declaration of Human Rights (1948)

The UDHR is the most fundamental human rights document in the world. Adopted by the UN General Assembly in 1948, it establishes the basic rights and freedoms that all people are entitled to, regardless of where they live. Several articles are directly relevant to AI surveillance. The UDHR was written before computers existed, but UN human rights bodies have confirmed that its protections apply fully in the digital age. A government that uses AI to secretly monitor its citizens' political opinions or track who attends religious services is violating the spirit and letter of the UDHR.

International Covenant on Civil and Political Rights (1966)

The ICCPR is a legally binding treaty that most countries in the world have ratified, meaning they have formally agreed to be bound by it. Unlike the UDHR, which is a declaration of principles, the ICCPR creates actual legal obligations for the countries that sign it. Article 17 is the most important for this topic. It prohibits "arbitrary or unlawful interference" with a person's privacy. The UN Human Rights Committee, the expert body that monitors compliance with the ICCPR, has said that any government surveillance must meet three tests to be consistent with Article 17:

1. Legality: It must be authorized by a clear, publicly available law.
2. Necessity: It must be necessary to achieve a legitimate goal.
3. Proportionality: It must not go further than needed to achieve that goal.

Mass surveillance programs that monitor entire populations without individual suspicion have a very hard time meeting these tests.

UN General Assembly Resolutions on Privacy in the Digital Age

In 2013, following major revelations about mass surveillance programs run by the United States and its allies, the General Assembly passed Resolution 68/167, the first UN resolution specifically addressing digital privacy. It affirmed that "the same rights that people have offline must also be protected online." Since then, the GA has passed several follow-up resolutions that have added increasingly specific language about surveillance technologies and AI.

These resolutions have called on Member States to

- I. Review their surveillance laws for compliance with international human rights law
- II. Establish independent oversight bodies for intelligence and surveillance activities
- III. Ensure that individuals have legal recourse when their privacy is violated
- IV. Refrain from mass surveillance of communications

UNESCO Recommendation on the Ethics of Artificial Intelligence (2021)

In 2021, all 193 UNESCO Member States unanimously agreed on the first global framework for AI ethics. This recommendation covers many aspects of AI, but on surveillance specifically it calls on Member States to consider moratoriums (temporary bans) on the use of AI tools for real-time identification of people in public spaces where these tools pose serious human rights risks. A moratorium would mean pausing the use of these technologies until adequate safeguards are in place.

The Global Digital Compact (2024)

The Global Digital Compact was agreed upon at the UN Summit of the Future in September 2024. It is one of the most recent and comprehensive statements of what UN Member States collectively believe about digital governance. The Compact affirms that human rights apply fully in digital spaces and calls for international cooperation on AI governance, including the development of common AI safety standards. It represents the starting point for what many hope will eventually become a binding international AI governance framework.

European Union AI Act (2024)

The EU AI Act is the world's first comprehensive law specifically regulating AI systems. It classifies AI applications by the level of risk they pose and sets rules accordingly. Crucially, it prohibits the use of real-time facial recognition in public spaces for law enforcement purposes, with only very narrow exceptions.

The EU AI Act is not a UN document and only directly applies in EU countries. However, because any company selling AI products in Europe must comply with it, experts expect it to raise standards for AI systems globally, much like how European car safety regulations have influenced car manufacturing worldwide.

Role of the International System

Office of the High Commissioner for Human Rights (OHCHR)

OHCHR is the main UN office responsible for promoting and protecting human rights. In 2021, OHCHR published a major report calling for a moratorium on the sale and use of AI surveillance tools until adequate human rights protections are in place. The report documented cases where AI surveillance had been used to target journalists, political dissidents, and minority communities. OHCHR has also developed guidance for businesses on their responsibilities when developing or selling AI products that affect human rights.

Special Rapporteur on Freedom of Expression

This expert has documented how AI surveillance creates an effect on free speech and a free press. When people know they are being watched or even just believe they might be, they

often change their behavior. They avoid attending protests, stop communicating with certain people, or self-censor their opinions online. The Special Rapporteur has noted that this effect is particularly damaging for journalists, activists, and members of minority communities who already face discrimination.

International Telecommunication Union (ITU)

The ITU is the UN agency that sets technical standards for telecommunications and digital technologies. It has been working on AI governance frameworks and runs the "AI for Good" platform, which promotes the beneficial use of AI while addressing its risks. For developing countries that lack the technical expertise to evaluate AI surveillance systems before purchasing them, the ITU can be an important source of guidance and capacity building.

Non-Governmental Organizations (NGOs)

Civil society organizations have been essential in documenting AI surveillance abuses and pushing for international accountability. Organizations like Human Rights Watch, Amnesty International, Access Now, and Privacy International publish detailed investigations into how governments are using surveillance technologies. Their research provides crucial evidence that shapes UN debates and media coverage.

Export of Surveillance Systems



One of the most urgent dimensions of this issue is the global trade in surveillance technologies. AI surveillance tools are manufactured in some countries and sold to governments in others, often with very little oversight of how those tools will be used.

Research by the Carnegie Endowment for International Peace found that AI surveillance technology has been sold by companies in democratic countries to authoritarian governments

around the world. Chinese technology companies, in particular, have exported surveillance infrastructure to dozens of countries across Africa, Asia, and Latin America. This has allowed governments that lack the technical capacity to build their own surveillance systems to rapidly acquire capabilities that they then use to monitor political opponents, ethnic minorities, and journalists.

Currently, there is no binding international framework governing the export of AI surveillance technologies. Some proposals that have been discussed include

- I. Requiring companies to conduct human rights due diligence before selling surveillance tools to foreign governments, assessing whether those tools are likely to be used to violate rights.
- II. Creating international export controls on surveillance technology, similar to existing controls on weapons.
- III. Establishing a transparency registry where governments must disclose what surveillance technologies they have purchased and from whom.
- IV. Building capacity in developing countries so they can independently evaluate surveillance tools before purchasing them and have the regulatory infrastructure to oversee their use.

Conclusion

AI-powered surveillance is one of the defining human rights challenges of our generation. Facial recognition and social scoring systems give governments tools that can fundamentally transform the relationship between the state and its citizens for better or for worse, depending entirely on the safeguards in place and the intentions of those in power.

The international community has started to address this challenge, but progress has been slow. The General Assembly has an important role to play in setting norms, encouraging Member States to reform their domestic laws, and pushing for stronger international frameworks. As a delegate, your task is to represent your country's interests honestly while contributing to the development of policies that protect the rights of people everywhere. The best resolutions will acknowledge the legitimate security concerns that motivate surveillance programs while establishing clear, enforceable safeguards against abuse.

Questions to Guide Your Research

1. Does your country use facial recognition or social scoring systems? What laws govern their use?
2. Has your country ratified the ICCPR? Has the UN Human Rights Committee raised concerns about its surveillance practices?
3. How has your country voted on UN resolutions about privacy in the digital age?

4. Does your country export or import AI surveillance technology? To or from which countries?
5. Does your country have an independent oversight body for intelligence and surveillance activities?
6. What has your country's government said publicly about AI governance or the UNESCO AI Ethics Recommendation?
7. Have there been documented cases in your country of surveillance being used against minorities, journalists, or political opponents?

Bibliography

- I. Access Now. *Reclaim Your Face: Ban Biometric Mass Surveillance*. 2021. Available at <https://reclaimyourface.eu/>
- II. Amnesty International. *Ban Dangerous Facial Recognition Technology That Amplifies Racist Policing*. 2021. Available at <https://www.amnesty.org/en/latest/news/2021/01/ban-dangerous-facial-recognition-technology-that-amplifies-racist-policing/>
- III. Carnegie Endowment for International Peace. *AI Global Surveillance Index*. 2019. Available at <https://carnegieendowment.org/publications/interactive/ai-global-surveillance>
- IV. European Parliament and Council of the EU. *Artificial Intelligence Act*. 2024. Available at <https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX:32024R1689>
- V. Human Rights Watch. *How China's Surveillance Is Spreading Globally*. 2019. Available at <https://www.hrw.org/news/2019/11/06/how-chinas-surveillance-spreading-globally>
- VI. Office of the United Nations High Commissioner for Human Rights. *The Right to Privacy in the Digital Age*. 2021. Available at <https://www.ohchr.org/en/calls-for-input/right-privacy-digital-age>
- VII. United Nations, General Assembly. *Universal Declaration of Human Rights (A/RES/217 A (III))*. 1948. Available at <http://www.un.org/en/documents/udhr/>
- VIII. United Nations, General Assembly. *International Covenant on Civil and Political Rights (A/RES/2200 (XXI))*. 1966. Available at [http://undocs.org/en/A/RES/2200\(XXI\)](http://undocs.org/en/A/RES/2200(XXI))
- IX. United Nations, General Assembly. *The Right to Privacy in the Digital Age (A/RES/68/167)*. 2013. Available at <http://undocs.org/en/A/RES/68/167>

- X. United Nations, General Assembly. *The Right to Privacy in the Digital Age (A/RES/75/176)*. 2020. Available at <http://undocs.org/en/A/RES/75/176>
- XI. United Nations, General Assembly. *Global Digital Compact (A/RES/79/1)*. 2024. Available at <http://undocs.org/en/A/RES/79/1>
- XII. United Nations Educational, Scientific and Cultural Organization. *Recommendation on the Ethics of Artificial Intelligence*. 2021. Available at <https://unesdoc.unesco.org/ark:/48223/pf0000381137>
- XIII. U.S. National Institute of Standards and Technology. *Face Recognition Vendor Testing — Demographic Effects*. 2019. Available at <https://nvlpubs.nist.gov/nistpubs/ir/2019/NIST.IR.8280.pdf>