AI and Ethics
https://doi.org/10.1007/s43681-022-00194-0

ORIGINAL RESEARCH

Reconsidering the regulation of facial recognition in public spaces

Sara Solarova[1] · Juraj Podroužek[1] · Matúš Mesarčík[2] · Adrian Gavornik[1] · Maria Bielikova[1]

Received: 31 March 2022 / Accepted: 14 June 2022

Abstract
This paper contributes to the discussion on effective regulation of facial recognition technologies (FRT) in public spaces. In response to the growing universalization of FRT in the United States and Europe as merely intrusive technology, we propose to distinguish scenarios in which the ethical and social risks of using FRT are unattainable from other scenarios in which FRT can be adjusted to improve our everyday lives. We suggest that the general ban of FRT technologies in public spaces is not an inevitable solution. Instead, we advocate for a risk-based approach with emphasis on different use-cases that weighs moral risks and identifies appropriate countermeasures. We introduce four use-cases that focus on presence of FRT on entrances to public spaces (1) Checking identities in airports (2) Authorisation to enter office buildings (3) Checking visitors in stadiums (4) Monitoring passers-by on open streets, to illustrate the diverse ethical and social concerns and possible responses to them. Based on the different levels of ethical and societal risks and applicability of respective countermeasures, we call for a distinction of public spaces between semi-open public spaces and open public spaces. We suggest that this distinction of public spaces could not only be helpful in more effective regulation and assessment of FRT in public spaces, but also that the knowledge of different risks and countermeasures will lead to better transparency and public awareness of FRT in diverse scenarios.

Keywords Facial recognition · AI regulation · Public spaces · Semi-open public spaces · AI ethics · Countermeasures · Airports · Office building · Stadiums · Open streets · Trustworthy AI · Transparency

1 Introduction

The proliferation of biometric systems, and specifically facial recognition technologies in our everyday lives, has prompted the need to foster a public discussion regarding the associated societal and ethical concerns. At the very moment, the deployment of facial recognition technologies (FRT) in the United States and Europe finds itself amid the debate between a complete prohibition or severe restriction on its general use in public spaces. Specifically, in the EU, following the recent proposal of Artificial Intelligence Act (AIA) [1] the public outcry has resulted in the growing call for the general ban of FRT in public spaces, highlighting fundamental human rights violations [2–6]. This universalisation tends to reduce the complexities of FRT to unattainable human rights abuses, which sees only one solution: retreat.

Biometric facial recognition is a form of artificial intelligence system that incorporates the automated extraction, digitisation, and comparison of the geometric and spatial distribution of facial features to identify individuals and create a faceprint [7, 8]. These faceprints or digital templates are stored in the biometric system to be either a) verified (1:1) against the enrolment template, or b) identified (1: N) against a database of other templates. In this paper, we will focus on the regulation of remote facial recognition

✉ Sara Solarova
sara.solarova@kinit.sk

Juraj Podroužek
juraj.podrouzek@kinit.sk

Matúš Mesarčík
matus.mesarcik@kinit.sk

Adrian Gavornik
adrian.gavornik@kinit.sk

Maria Bielikova
maria.bielikova@kinit.sk

[1] Kempelen Institute of Intelligent Technologies, Bratislava, Slovakia

[2] Comenius University in Bratislava, Faculty of Law and Kempelen Institute of Intelligent Technologies, Bratislava, Slovakia

Published online: 11 July 2022

technology that serves to identify natural persons at a distance against a reference database containing other biometric data. Nonetheless, we suggest that most of our arguments will hold for any other uses of facial recognition technology including non-remote FRT systems.

The establishment of an identity that is secure and convenient has become critical, thus prompting the need for reliable authentication techniques in the wake of networking, communication, and mobility [9]. Biometric facial recognition technologies offer such establishments based on our unique characteristics and many other benefits. Data processed by facial recognition systems have been used for numerous years in the area of freedom, security and justice safeguarding (supra)national security [10]. The unprecedented advancements in AI-powered FRT extend beyond the purposes of national security hence necessitating a separate discussion to which we have an ambition to contribute. Following the work of Almeida et al. [11] and Moraes et al. [12] on possible regulatory and ethical approaches for FRT deployment predominantly in law enforcement, our aim is to contribute to the debate of effective regulation of FRT by demonstrating that a distinction of public spaces between semi-open public spaces and open public spaces allows for more clarity and transparency in upcoming discussions. Our addition also lies in the assessment of various ethical and societal risks and the subsequent illustration of specific countermeasures including their level of applicability to demonstrate the differences between use-cases in which FRT could be deployed in various public spaces.

We believe that trustworthy facial recognition technology can be achieved only if its development and use is intertwined with ethics at its core. We elaborate on this idea from several perspectives. First, we support regulation of FRT that weighs the ethical and societal risks and identifies appropriate countermeasures based on a risk-based approach. Second, we hold that this risk-based approach should not only be applied to facial recognition technology in general but should take into account a set of possible use-cases in which FRT could be deployed and used. And third, we suggest that analysing the impacts of FRT for specific use-cases will allow us not only to better identify and categorise ethical and societal risks but also to propose more appropriate countermeasures.

For this purpose, we outline and discuss some of the most vocalised concerns regarding FRT, considering the overarching ethical principles and human values of transparency, fairness, robustness, privacy, and human agency. We propose a non-exhaustive set of use-cases, from which the risks and its mitigations can be applied in a more general discussion about effective regulation of FRT, and AI in general.

It should be also mentioned that we shall not refrain from analysing the impacts of facial recognition technologies in all kinds of spaces including online spaces, even if proposed legislation is focusing merely on the physical environment. This topic, however, requires more elaborated discussion and is beyond the scope of our paper. Nonetheless, we are convinced that the ethical issues of facial recognition in online spaces are equally important and deserve further attention.

1.1 Facial recognition in public spaces

The debate about regulation of FRT in Europe is grounded mainly in the context of *public spaces*. This notion is also crucial in the subsequent debate in favour of or against the deployment of FRT. According to the AIA [1] which aims to put forward the horizontal regulation of AI in general, public space is defined as "any physical space accessible to the public, irrespective of private or public ownership" (Article 3 [13]). This definition embraces not only all open and freely accessible spaces like streets or town squares but all spaces that are in principle accessible for the public including spaces like airports or commercial buildings such as office buildings, retail centres or entertainment centres [14], which are also under some circumstances accessible, despite belonging to private entities. The view is also supported by Recital 9 of AIA classifying streets, relevant parts of government buildings and most transport infrastructure, spaces such as cinemas, theatres, shops, and shopping centres under the notion of public space. From the legal perspective, these areas in which FRT can be deployed are considered indistinguishable because they fall into the broader category of public spaces. However, from the perspective of ethical and societal risks that stem from deploying FRT and their possible mitigations, these areas are not equivalent and warrant a more nuanced differentiation.

Differentiating between public spaces is crucial in many aspects. For example, it provides supplemental information on the data set under investigation, i.e., the set of people that move across a space. Such differentiation between the sets of people brings into the forefront the fact that movement and presence in some of the public spaces already carries a certain expectation of one's identity being checked. Also, there are public spaces with a relatively stable and predictable set of people that need to be recognised (entrances to stadiums or company premises to an extent). And as we will demonstrate later, the set of ethical and societal risks alongside effective countermeasures to these risks may vary between different kinds of public spaces as well.

For this reason, we propose a distinction between open public spaces and semi-open public spaces [15, 16], which is more sensitive to different levels of accessibility of public spaces. We regard open public space as a physical space that is publicly accessible without further specific social selection. This does not mean, however, that the use of such a space is not subject to certain rules governing its utilisation and that such a space does not have its owner or

administrator. Frequently open public spaces are owned or managed by state institutions or public administration. On the other hand, semi-open public spaces represent a physical area that might be owned by either private or state entities. Accordingly, and unlike open public spaces, semi-open public spaces are not necessarily a world of strangers [16, 17]. Interactions that happen in semi-open public spaces are more structural, involving interpersonal networks of people that attend such spaces, adding certain private character to them particularly lacking in public spaces [18]. Therefore, this character might impose a certain degree of regulation of social accessibility, such as the expectation to be identified in order to be granted a further entry.

We believe that the knowledge of this differentiation has the potential to improve the transparency and clarity of debate on effective regulation of FRT. It should also increase the awareness of the public regarding specific purposes and contexts for their biometric data processing and debunk the universalization of FRT in public spaces as inherently intrusive. Unfortunately, this differentiation is only barely present in current and forthcoming regulation of FRT.

1.2 Current legislative trends

The current legislative trends in the United States and Europe showcase the growing hesitancy towards facial recognition use in public spaces, resulting in severe restrictions, possibly leading towards a general ban on its use in general. Back in 2018, the Supreme Court of the United States in Carpenter v. the United States, ruled that an individual has a legitimate expectation of privacy in the record of their physical movements, meaning that unless there is a secured warrant from an independent judge based on probable cause to believe one has committed a crime, an individual should not be monitored or surveilled for law enforcement purposes [19]. Following this, California became the first state to ban FRT in law enforcement in 2019 [20], enacting the California Consumer Privacy Act [21] which went into effect the same year as the Illinois Biometric Information Privacy Act, (BIPA), restricting the collection/use of the biometric data of its residents [22].

As a result, various cities across the US framed facial recognition technologies in violative terms. Cities in California (e.g., Berkeley and San Francisco) banned facial recognition technology in 2019 in law enforcement following the concerns of mass surveillance [11], and digital rights groups aim to fulfil the same aim in New York [23]. Portland followed the example by extending the ban of FRT for all city departments and private retailers such as hotels or restaurants [20]. From an individual perspective, in *Patel v. Facebook* [24] the court ruled in favour of 23 plaintiffs claiming that Facebook and its tagging suggestions directly invaded their private affairs and concrete interests under BIPA [25].

Concerning the EU legal framework, the discussion on FRT regulation is in addition to the AIA proposal framed mainly by the General Data Protection Regulation (GDPR) [22] and Law Enforcement Directive (LED) [26] that are both part of the European Data Protection Package. GDPR applies to all processing of personal data in the Member states, including private and public sector. Personal data qualify as information relating to an identified or identifiable natural person. The definition is broad and also covers information allowing indirect identification of individuals. This in practice means that the EU data protection regime is applicable to every processing operation involving information that may be theoretically linked to data subjects [27].

Biometric data in GDPR are defined as personal data that result from technical processing of physical, physiological, or behavioural characteristics of a natural person (Article 4 [15]) [22]. In general, photographs and processing thereof, shall not be considered a special category of personal data, and they are considered biometric data only when processed through a specific technology permitting unique identification or authentication of a natural person (Recital 51) [22]. Biometric data, if used for the purpose of uniquely identifying a natural person, shall be considered as a special category of personal data (Article 9 [1]) [22]. For processing of special categories of personal data GDPR stipulates stricter conditions that a controller or a processor is obliged to follow. Furthermore, GDPR sets rules on data storage and the limitation thereof, in which the data shall be kept in a form allowing for the identification of the data subject for no longer than the originally specified purpose for which the data have been processed (Article 1 [5e]) [22]. Accordingly, after the purposes have been achieved, the data shall be deleted or anonymized. This applies to both photographs and vectors used for biometric processing of the data subjects in question provided that they qualify as personal data.

LED stipulates rules on processing personal data (including biometric data) for law enforcement purposes. In short, these include prevention, investigation, detection or prosecution of criminal offences or the execution of criminal penalties by the competent authorities who were vested the power to carry out such activities. Pursuant to Article 10 of LED, the processing of special categories of data, such as biometric data, is allowed only where strictly necessary [26]. LED represents *lex specials* to GDPR meaning that if a specific rule for processing for law enforcement purposes exists, GDPR does not apply. The directive in question contains several specifications and differences compared to GDPR including different data subjects (victims, informants, suspects etc.), obligation of logging or exclusion of data subject's rights. More flexibility also applies in terms of principles of data minimisation and purpose limitation.

The latter is essential considering processing of biometric data for law enforcement purposes including FRT technology in public spaces.

The proposal of AIA, which should serve as horizontal regulation of AI systems, introduces a ban of FRT in public spaces by law enforcement for specific exceptions as ruled by member states (Article 5 [2–4]) [1]. The rest of FRT technology used for real-time and post remote biometric identification of natural persons is classified as a high-risk area triggering requirements as prescribed by AIA (Article 5 [2–4]) [1]. Nonetheless, such a restrictive proposal as laid out by AIA is deemed insufficient in public debates, driven by the European Data Protection Supervisor and European Data Protection Board [4]. Following that, the European Parliament passed with an overwhelming majority a non-binding resolution, which calls for a complete ban on facial recognition, in public spaces, but also extends the ban on predictive policing techniques and private facial recognition databases [28]. This position is further supported by several non-governmental organisations including the European Digital Rights organisation or ALLAI [5, 29]. There is even a petition with the same goal, dismissing any use of FRT, mostly arguing against corporate and state surveillance. Also, French security law on the use of facial recognition systems in the public spaces was subject to significant criticism, putting forward the concerns on the privacy of citizens affected by possible state surveillance [30].

The direction of decisions outlined above suggest an inclination of future legislation in Europe and US towards a general ban on FRT in public spaces. Within this approach the regulation tends to prefer prohibition instead of introducing countermeasures which can mitigate the most severe risks.

1.3 Trustworthy facial recognition

Regulatory law is not the only way of securing the development and use of FRT. In recent years the position of ethics and self-regulation practices in general have become more apparent. In some cases, this massive proliferation of ethics in regulation of Information and Communication Technology (ICT) and Artificial Intelligence have even given rise to serious doubts concerning the "ethification" [31] and "ethics bluewashing" [32] to name only two of them. Nonetheless, the role of ethics in AI regulation could still be crucial, particularly in the state of policy vacuums [33], which is present in most debates on AI and standardised FRT regulation [11].

One of the key concepts in AI ethics is trustworthiness of AI systems, which comprises the idea that the AI systems should be lawful, ethical, and secure [34]. It holds that every AI-based technology, and FRT is not an exception, that has an ambition to be deemed trustworthy should conform to the common ethical principles and values [34]. In recent years, there have been many initiatives that aimed to set and translate the principles of trustworthy AI in practice, among which the Ethics Guidelines for Trustworthy AI and its Assessment List for Trustworthy AI (ALTAI) delivered by EC expert group AI HLEG [35] are one of the best known and widely accepted, at least in Europe. Most of these initiatives are grounded on universal human values and moral principles that all AI systems should meet to be deemed as ethical or trustworthy [36]. The next sections will address some of these shared concepts, namely transparency, fairness, privacy, robustness, and human agency. Table 1 illustrates most repeated and vocalised ethical concerns regarding facial recognition technology in public spaces regarding selected principles and values for trustworthy and ethical AI.

1.3.1 Transparency

One of the most known objections towards biometric facial systems pertain to the problems with their transparency and the lack of awareness of their presence. Transparency of FRT in public space should be the first requirement we put on its systems to be able to properly analyse other ethical and societal issues.

The proliferation of such systems in public spaces without proper awareness could also enhance the chilling effects, or the continuous curation of one's action [37], resulting in people altering their behaviours due to the feeling that they are constantly being watched and analysed. Based on the recent experience of using FRT to enhance the efforts towards social control [38], it is easy to imagine dystopian scenarios where private companies or non-democratic regimes start massive surveillance of their employees or citizens. And even the public awareness of such scenarios can create a society heavily paralysed and in fear of conducting any non-conform behaviour.

Table 1 Examples of ethical and societal concerns in FRT

Overarching principles and values for trustworthy AI	Examples of ethical and societal concerns in FRT
Transparency	Chilling effects, interpretability and explainability
Privacy and data governance	Forced recognition, data control
Technical robustness and safety	False positives, false negatives
Diversity, non-discrimination, and fairness	Underrepresentation, social exclusion
Human agency and oversight	Mute individuals, over-reliance

Transparency has also many other facets concerning AI explainability [39]. We should understand not only what is important for AI models used in facial recognition in general but also to some degree the reasons for a particular recognition, e.g., when a FRT system is not capable of recognizing specific identity [13]. On the other hand, there are also risks of fully transparent and explainable FRT, because higher interpretability and explainability could lead to higher vulnerability to security incidents [39].

1.3.2 Privacy and data governance

Debates about the use of FRT in public spaces often revolve around the loss of privacy. If such technologies are deployed, people fear that it would become almost impossible to avoid being scanned or monitored and therefore would find it difficult to maintain any sense of privacy. If most of the people captured by the biometric system are unintended captured persons and passers-by or their data are being processed for other purposes than formerly declared, there should be persuasive rationale as to why they are exposed to intrusive practices.

But privacy does not solely concern the unwarranted intrusion. From an ethical position, an individual has privacy when protected from intrusion, interference, and information access by others [40]. Therefore, also informational privacy, or the users' ability to control the flow of their personal information (right to informational self-determination), shall also be addressed when deploying FRT technologies [8]. The proliferation of surveillance capitalism [37] together with the omnipresent online profiling of users advocates the need to address such concerns regarding privacy in physical space.

1.3.3 Technical robustness and safety

One of the other significant ethical issues concerning the negative effects of FRT is its inaccuracy considering the erosion of safety and human dignity from false positives and false negatives. False positives in identification processes happen when an individual's facial template is incorrectly identified via running it through the biometric database [12]. This involves situations when an individual is being incorrectly flagged as suspicious, or a potential threat. False negatives refer to the outcome when a facial template is not linked to the corresponding person in the database, thus not correctly identifying the target. This accounts for situations where an individual fails to authenticate when entering a building or any other secured premises.

1.3.4 Diversity, non-discrimination, and fairness

There have been numerous cases where black, Asian and minority ethnic (BAME) populations have been disproportionately misidentified or not identified via FRT. Facial-scan technologies were proven to fail while identifying dark-skinned individuals mostly due to the training data being optimised for lighter-skinned users, in which white populations are disproportionately more represented. Buolamwini and Gebru's research [41] displayed profound inaccuracies in gender identification that relies on a person's skin colour. In Brazil, 90.5% of the people who were arrested resulting from the deployment of FRT for public security purposes, were black [12]. These results are often described as technological limitations; however, they are anchored in social and racial bias towards BAME populations. If FRT failures to identify a person correctly happen in a structural pattern, they can affect one's mental health, particularly self-esteem and self-respect [42–45].

Biometric facial recognition systems can also exacerbate bodily social sorting phenomena, where such technologies can expose hidden or sensitive information that could be later exploited by stakeholders to hierarchize populations [43]. In this fashion FRT can be deployed to track minority movements and reinforce their exclusion from the rest of society [46].

1.3.5 Human agency and oversight

With the widespread use of FRT, human bodies can increasingly serve as gateways to physical and virtual spaces [47]. They will function as a password, carrying substantial information that will progressively redefine the ontology of the human body [48]. This manifests itself in bypassing the human mind and directly moving to communication with the body that would be expected to provide objective information sufficient to make a decision. There are concerns that biometric technologies which prefer body over mind could produce mute individuals who may not be obligated or sometimes not allowed to consent, participate, or voice themselves [43].

Human autonomy can also be affected by over-reliance on machine-based decisions. That includes the risk that people will rely too much on machine decisions and do not use their own reasoning and capabilities. Such over-reliance could undermine the idea that it should be people who make decisions that affect other people's lives and who take responsibility for these decisions.

1.4 Contextualisation of FRT in public spaces

All the ethical and societal concerns mentioned above should be deemed legitimate. Yet we suggest that their relevance

heavily depend on the specific use-cases for which FRT systems are deployed. To support this position, we introduce four specific use-cases where FRT has been or will be deployed in the near future, namely (1) Checking identities in airports (2) Authorisation to enter office buildings (3) Checking visitors in stadiums (4) Monitoring passers-by on open streets. In each use-case, we outline examples of ethical and societal risks considering various stakeholders that might be reasonably affected by FRT and examples of possible countermeasures.

It must be mentioned that this list is non-exhaustive and focuses only on a small subset of use-cases considering FRT systems deployed at entrances to public spaces. These examples serve as an illustration of how different use-cases can vary in depth of ethical concerns and the degree to which they can be dealt with. Deeper insight into use-cases, where the amount of risks and availability of effective countermeasures might vary depending on e.g., the exact point where FRT is deployed, would be valuable for future discussion but is beyond the scope of this paper.

1.4.1 Checking identities in airports

Facial recognition technologies have been continuously used in airports all over the world. In Germany, multiple airports have used a system which integrates FRT for identity verification [20]. This technology can spot illegal attempts to enter a country at a more efficient and precise level, thus increasing the levels of public security maintaining public order and simultaneously increasing the comfort of the passengers to decrease the waiting time. Moreover, responding to the COVID-19 threats, FRT can help to address the need for contactless security checks amidst the pandemic crisis, reducing the transmission of pathogens at the airport. The use of FRT in airports has also one of the highest levels of support among the public and experts [49].

1.4.1.1 Chilling effects The expectation to be verified at the airport is apparent for all its visitors. In some cases of digital onboarding, individuals must be identified to be permitted to enter the airport's premises. Yet the extent of such biometric identification should still be properly, and clearly explained and other alternatives should be available for people who exercise their right to not be processed by facial recognition systems.

1.4.1.2 Forced recognition Nevertheless, we recognize the concern that a person's position during an airport check might become asymmetrical, particularly when she may be distressed that refusing to undergo the biometric identification control might raise suspicion from the side of the authorities. In these situations, societal pressure might nudge people to subject themselves to biometric identification even if they preferred the alternative. Therefore, it is inevitable to ensure that the alternative to biometric identification not only exists but also it should guarantee equal standing and reliability to the biometric one. Hereby, an individual can make a free choice without having to be concerned about consequences thereof.

1.4.1.3 Social exclusion Even though there is a significant improvement of FRT accuracy considering the demographic features in recent years, the presence of unfair biases is still inevitable to address. The risk that socially biased systems can result in perpetuated social inequalities would be considerable at the airports when the systems will single out specific individuals or groups of people for increased harassment or searches [50].

1.4.1.4 False positives, false negatives The occurrence of false-positive and false-negatives and the ethical concerns thereof can be undertaken by guaranteeing the right to obtain human intervention on the part of the controller and to express one's own point of view and to contest the decision of facial recognition system should there be a suspicion of erroneous automated outcome. The risk of false positives could increase the discomfort of passengers and can have a negative impact on their dignity. On the other hand, the occurrence of false negatives can heavily endanger the security of the whole area and it is one of the biggest issues to be dealt with. Guaranteeing a human intervention would be also useful as a part of fallback procedures in case of malfunctioning of FRT or of serious doubts about its accuracy in specific cases.

1.4.1.5 Data control Being able to get information on the amount of time for which the video footage and images will be stored in a system does not only address awareness issues but also the privacy concerns. Henceforth, the system could be designed to delete personal data from the airport system after the take-off. Following this approach, we can mitigate the risk of ambiguity with tracking the duration of biometric mass surveillance.

1.4.2 Authorisation to enter office buildings

With the use of FRT, owners of these buildings can improve security, identify unauthorised access, and make the entry process seamless and comparatively faster than requiring identification with an ID card. Security passes or ID cards can be stolen, duplicated, or borrowed, which might compromise the security of a particular facility. Facial biometric identification will significantly decrease such a risk as biometric identifiers are more difficult to obstruct.

1.4.2.1 Chilling effects Most of the ethical risks and respective countermeasures available for FRT deployed on entrances of company premises are similar to the use of FRT at the airport checks. As in the airport scenario, most visitors have an expectation they will be subject to verification checks. But visible and clear information on the use of FRT within given space to safeguard the awareness should be provided.

1.4.2.2 Forced recognition For people who choose to not be identified by FRT, a separate entrance for conventional access should be available to address the concern of autonomy, regardless of whether they are employees or the general public attending the place.

1.4.2.3 Data control The set of identities is relatively stable. Companies are expected to have a database of their employees or people who are allowed to enter their premises beyond the general public space, such as the reception area, typically accessible to the wider public without the need for identification. This means that even when a biometric data are taken and processed for the purposes of identification against a larger database of people, and the access is denied, the data could be erased automatically and not stored for future purposes. This way, the concern of intrusion can be addressed, namely by decreasing the amount of time for which the biometric data are stored.

1.4.3 Checking visitors on stadiums

Stadiums are known for hosting a larger number of people who are emotionally charged during a particular event. As a result, violence and fights frequently occur, often resulting in casualties. Facial recognition technology can help to identify the people on the blacklist thereby preventing them from entering the stadium. The automated process of entry decreases the waiting time for attendees of an event, improving the overall experience. The benefits of using FRT can improve the overall sense of security, such as by preventing people with previous instances of violent behaviour in stadiums from entering or enhancing the convenience of getting in.

1.4.3.1 Chilling effects This use of FRT should also be properly and clearly communicated, so that the individuals are informed about the processes of FRT within the stadiums.

1.4.3.2 Forced recognition Separate entrances should be provided for people who exercise their right for their data to not be processed by facial recognition systems.

1.4.3.3 False positives and false negatives The ethical concerns addressing the erroneous decision of denying an individual the entry to the stadium (i.e., false positive) can be mitigated by human oversight. Such an approach can minimise the potential concern of accuracy.

1.4.3.4 Data control As in previous use-cases, we expect the set of people possibly identified to be approximately stable, given the capacity of the stadium and the number of bought tickets. Unless the individual is on the blacklist based on their prior violent conduct, the biometric data of the attendees should be stored for a limited time after the event deemed absolutely necessary for maintaining public safety and order. The software in place does not need to store images of anyone besides the blacklisted people and the data stored in the internal stadium's system must not be connected to the internet or any other system, which minimises the possibility of being hacked [51].

1.4.4 Monitoring passers-by on open streets

The deployment of facial recognition technologies in open streets is not an imaginary scenario. In the Netherlands, Poland and Germany, this technology was not put in place based on evidentiary and imminent threats but more on precautionary or deterrence measures. These account to improving the overall sense of security, such as preventing criminality, finding missing children or terrorists. The statistics of success, namely in Cologne, where thousands of people were analysed daily showed that less than 0.1% of what was recorded yielded in probative values. This brings into question the proportionality and the continuous justification of deployment of facial recognition by state authorities to monitor the population [52].

Arguing for facial recognition in open streets becomes more difficult particularly with respect to the right to freedom from criminal investigation or unreasonable monitoring when there is an absence of prior evidence of any violation of the laws. These rights are upheld in liberal democracies, and it is agreed that states have no right to engage in selective monitoring or seeking process of wrongdoings of citizens when there is no reasonable doubt of suspicious behaviour [8].

1.4.4.1 Chilling effects The set of identities on open streets and town squares is not as stable as in previous use-cases and these people do not in general expect to be recognised by entering them. That raises the questions about providing an effective and fair way to inform people about face recognition before it even happens, e.g., with notifications sent into their smartphones. On a more societal level, the uncontrolled use of FRT in open streets can discourage citizens to

partake in certain events (protests, rallies, demonstrations), hence limiting the potential of participatory democracy.

1.4.4.2 Forced recognition The ability to provide an opt-out scenario in open streets is hard to maintain, given it is an open space in which it is almost impossible to differentiate movements and possibly entries, in which the set of people is dynamic and not a priori defined. In most cases it could be quite unrealistic to avoid the recognition by providing separate entrances and more conventional means to checking the identities. These measures would also contravene with the primary purposes of FRT deployment for this use-case.

1.4.4.3 Mute individuals If FRT deployed on entrances to open streets and town squares would scan citizens by default, the body could easily surpass the position of our minds, where the body will be the main provider of objective data seen as the reliable communicator. This way, people can perceive their faces and bodies to have a more significant voice than their minds thus affecting the core of the democracy which depends on people participating in public debates with their minds and actual voices [43].

1.5 Proposed solutions

First, it has to be stated that many of the concerns and possible risks mentioned above can be mitigated by effective countermeasures already supported by the existing European legal framework e.g., GDPR. GDPR also requires explicit consent with processing sensitive personal data (including biometric data for identification purposes) if other exceptions for processing sensitive data are not applicable. And in case of data subjects not consenting to be subject to facial recognition systems, other alternatives could be offered, e.g., a separate entrance with no automated recognition [22].

Existing legislation could also help to manage possible tensions in values and principles listed above. A clear example of such countermeasures could be illustrated in the use of FRT in stadiums. The enhanced perception of security and potential prevention of violence could decrease the sense of individual privacy. In these cases, countermeasures of clear rules on the duration of the biometric faceprints storage or the awareness of the use of such technologies can mitigate the concerns of privacy loss. Almeida et al. [11], Moraes et al. [12], Smith and Miller [8] discussed potential shortcoming with the existing regulations and proposed regulatory and legislative developments to mitigate such risks. This paper does not continue the debate towards specific new regulation, instead we problematize the term public space and argue that for each use-case in which FRT could be deployed, the value tensions must be addressed individually and precisely, because these value tensions do not have to be automatically present in other scenarios.

It is of most importance to realise that the commonly addressed concerns and countermeasures do not always have to be universal or applicable in every single encounter with FRT but are rather context specific. Table 2 shows that at least in areas of transparency, privacy, and human agency the applicability of countermeasures to respective ethical and societal risks can be less effective considering the last use-case.

We suggest that for monitoring passers-by on open streets the effective countermeasures could be much harder to reach due to the nature of this space itself. The distinctive and certain amount of private character of semi-open public spaces permits changes to the facilitation of such space where the individuals are aware of and understand the borders of where this space begins and ends. Also, the interactions in semi-open public spaces are more predictable and therefore provide more alternatives for conventional forms of access or identity management. However, open-public spaces represent the areas available for people to freely navigate them, exercise their right to assembly and express their opinions [53]. At the same time, they also constitute the world of strangers [16], where interpersonal interactions do not occur as frequently as in semi-open public spaces [52] which are defined by the sense of commonality and habitual engagement of people who are part of the space [18]. That leads us to the need for clear distinction of public spaces into semi-open public, represented by first three use-cases, and open-public spaces. We are convinced that more nuanced distinction of public spaces could help better understand the risks and applicable countermeasures of FRT and contribute to more effective regulation of this technology.

For FRT, as for any new technology, it should hold that prohibiting its deployment is reasonable when the technology from its nature yields untenable risks that cannot be possibly balanced out by any countermeasures. However, if the nature of a given technology is not inherently unethical or wrong, we ought to be able to manage the specificities of such technology and functionalities within individual use-cases instead of prohibiting it in general. From this perspective the distinction between public spaces would be of utmost importance in marking these red lines not only for deployment and use of FRT systems but also for AI regulation in general.

Additionally, we would like to emphasise the idea that facial recognition affects various direct and indirect stakeholders [54] for whom it can constitute different sets of problems and solutions. It is therefore crucial that these stakeholder groups are thoroughly identified and engaged with during all stages of development, deployment, and use of the facial recognition systems and for every use-case. All stakeholders should be also capable of articulating concerns

Table 2 Examples of ethical risks considering FRT and different applicability of its countermeasures

Examples of risks	Examples of countermeasures	Applicability of countermeasures
People will be not aware about FRT purpose and aims	Inform people about the use of FRT technologies before entering the area and explain what a person can expect before opt-in	UC1 Airports—Good UC2 Companies—Good UC3 Stadiums—Good **UC4 Streets—Poor**
People will be forced to be analysed by FRT	Provide separate entrances for conventional access	UC1 Airports—Good UC2 Premises—Good UC3 Stadiums—Good **UC4 Streets—Poor**
Individual data will be stored and used for other purposes	Decrease the amount of time for which biometric personal data can be stored and require detailed logs of data processing	UC1 Airports—Good UC2 Premises—Good UC3 Stadiums—Good UC4 Streets—Good
FRT will undermine human autonomy and the right to be heard in public places	Ensure that biometric identification will serve as an alternative to the traditional forms of identification, not its complete substitute	UC1 Airports—Good UC2 Premises—Good UC3 Stadiums—Good **UC4 Streets—Poor**

Bold refers to the emphasis on the inability to mitigate the ethical risks associated with open-public spaces

and embracing the benefits so that the future design of such technology will dedicate time and effort to mitigate these ethical and societal concerns.

Modern governance is experiencing a socio-technical shift [55] in which the public can decide to reap the benefits and control the risks. The general ban would, unfortunately, in the long-term, weaken efforts for novel approaches to AI and new technologies in general through reducing the complexity of its applications into inherently violative consequences. Outright prohibition will also vitiate any effort of the affected groups to collectively discuss how to ensure both—technical innovation and respective moral requirements. At the same time, the providers should stay accountable for mitigations of risks that these systems could pose to society. We believe that the combination of legal and regulatory mechanisms [12] including soft law, oversight bodies, regulatory sandboxes, guidelines or internal assessments and external audits [11] can increase accountability of FRT, and its providers and simultaneously build trustworthiness from the population itself.

1.6 Conclusion

This paper discussed the contemporary utilisation of biometric facial recognition technologies. With such, we introduced the rationales as to beneficiary use-cases of trustworthy FRT that can improve security, efficiency and in some cases adapt to the public health crisis and make certain spaces even safer. Yet, we also highlighted the problematic aspects of facial recognition used in public spaces and introduced the examples of ethical and societal risks that could infringe upon values and principles widely accepted in liberal democracies.

We have shown that there could be significant differences not only between ethical and societal risks but also on the level of effective countermeasures for different use-cases when entering public spaces. With such, we aimed to present a much-needed distinction for future regulation of these technologies. Thus, instead of the general prohibition of facial recognition, we prefer articulating specific requirements for FRT use-cases. As a result, we propose the differentiation of public spaces into two categories, semi-open public and open-public spaces as they yield use-cases which should be regulated differently.

Acknowledgements This work is partially supported by The Ministry of Education, Science, Research and Sport of the Slovak Republic under the Contract No. 0827/2021.

Declarations

Conflict of interest Authors are reporting that Innovatrics, company developing facial recognition technology is one of the partners of Kempelen Institute of Intelligent Technologies.

References

1. European Commission: Proposal for laying down harmonised rules on artificial intelligence (Artificial Intelligence Act) and amending certain union legislative acts (COM (2021) 206 final). European Commission. https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX:52021PC0206 (2021). Accessed 1 May 2021
2. Algorithm Watch: Open letter calling for a global ban on biometric recognition technologies that enable mass and discriminatory surveillance. Algorithm Watch. https://algorithmwatch.org/

en/open-letter-ban-biometric-surveillance/ (2017). Accessed 10 Nov 2021
3. Devich-Cyril, M.: Defund facial recognition before it's too late. https://www.theatlantic.com/technology/archive/2020/07/defund-facial-recognition/613771/ (2020). Accessed 13 Nov 2021
4. EDPB: EDPB & EDPS call for ban on use of AI for automated recognition of human features in publicly accessible spaces, and some other uses of AI that can lead to unfair discrimination. https://edpb.europa.eu/news/news/2021/edpb-edps-call-ban-use-ai-automated-recognition-human-features-publicly-accessible_en (2021). Accessed 10 Nov 2021
5. European Digital Rights: New AI law proposal calls out harms of biometric mass surveillance, but does not resolve them. https://edri.org/our-work/new-ai-law-proposal-calls-out-harms-of-biometric-mass-surveillance-but-does-not-resolve-them/ (2021). Accessed 11 Dec 2021
6. Reclaim Your Face: Reclaim Your Face-Reclaim Your Face. https://reclaimyourface.eu/ (2021). Accessed 16 Nov 2021
7. Castalvecchi, D.: Beating biometric bias. Nature **587**, 347–349 (2020). https://doi.org/10.4324/9781315779607-10
8. Smith, M., Miller, S.: The ethical application of biometric facial recognition technology. AI & Soc (2021). https://doi.org/10.1007/S00146-021-01199-9
9. Jain, A.K., Ross, A., Pankanti, S.: Biometrics: a tool for information security. IEEE Trans. Inf. Forensics Secur. **1**(2), 125–143 (2006). https://doi.org/10.1109/TIFS.2006.873653
10. Fuster, G. G., Peeters, M. N.: Person identification, human rights and ethical principles: Rethinking biometrics in the era of artificial intelligence. https://www.europarl.europa.eu/stoa/en/document/EPRS_STU(2021)697191 (2021). Accessed 20 Dec 2021
11. Almeida, D., Shmarko, K., Lomas, E.: The ethics of facial recognition technologies, surveillance, and accountability in an age of artificial intelligence: a comparative analysis of US, EU, and UK regulatory frameworks. AI and Ethics (2021). https://doi.org/10.1007/s43681-021-00077-w
12. Moraes, T.G., Almeida, E.C., de Pereira, J.R.L.: Smile, you are being identified! Risks and measures for the use of facial recognition in (semi-)public spaces. AI and Ethics **1**(2), 159–172 (2021). https://doi.org/10.1007/s43681-020-00014-3
13. Williford, J.R., May, B.B., Byrne, J.: Explainable face recognition. In: Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics), 12356 LNCS, pp. 248–263. Springer, Cham (2020). https://doi.org/10.1007/978-3-030-58621-8_15
14. Zogg, R., Roth, K.W., Brodrick, J.: Using CHP systems in commercial buildings. ASHRAE J. **47**(9), 33–36 (2005)
15. Peterson, M.: Living with difference in hyper-diverse areas: how important are encounters in semi-public spaces? Soc. Cult. Geogr. **18**, 1067–1085 (2016). https://doi.org/10.1080/14649365.2016.1210667
16. Lofland, L.H.: A World of Strangers. Basic Books, New york, Ny (1973)
17. Lofland, L.H.: Social life in the public realm: a review. J. Contemp. Ethnogr. **17**, 453–482 (1989)
18. Amin, A.: Ethnicity and the multicultural city: living with diversity. Environ Plan A **34**, 959–980 (2002)
19. Carpenter v. United States, 138 S. Ct. 2206: https://www.supremecourt.gov/opinions/17pdf/16-402_h315.pdf (2018). Accessed 12 Nov 2021
20. Kostka, G., Steinacker, L., Meckel, M.: Between security and convenience: facial recognition technology in the eyes of citizens in China, Germany, the United Kingdom, and the United States. Public Underst. Sci. **30**(6), 671–690 (2021). https://doi.org/10.1177/09636625211001555
21. California Consumer Privacy Act of 2018. California Civil Code, 1798.100 (2018) Accessed 3 Dec 2021
22. General Data Protection Regulation, 2016: https://eur-lex.europa.eu/eli/reg/2016/679/oj (2016). Accessed 27 Oct 2021
23. Conger, K., Fausset, R., & Kovaleski, S. F.: San Francisco bans facial recognition technology. The New York Times, pp. 1–3 (2019). Accessed 18 Nov 2021
24. Patel v. Facebook Inc., 290 F. Supp. 3d 948 (N.D. Cal. 2018): https://law.justia.com/cases/federal/appellate-courts/ca9/18-15982/18-15982-2019-08-08.html (2018). Accessed 19 Nov 2021
25. Biometric Information Privacy Act. 740 ILCS 14 (2018) Accessed 3 Dec 2021
26. European Parliament and Council: Directive (EU) 2016/680 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 by competent authorities for the purposes of the prevention, investigation, detection or prosecution of criminal offences or the execution of criminal penalties, and on the free movement of such data, and repealing Council Framework Decision 2008/977/JHA (2016). Accessed 8 Dec 2021
27. Court of Justice of the European Union: Case C-582/14, Patrick Breyer v Bundesrepublik Deutschland (2016). Accessed 3 Jan 2022
28. European Parliament: European Parliament Resolution on artificial intelligence in criminal law and its use by the police and judicial authorities in criminal matters. https://www.europarl.europa.eu/doceo/document/A-9-2021-0232_EN.html (2021). Accessed 17 Dec 2021
29. ALLAI: Council of Europe Calls for a Ban on Certain Facial Recognition Applications–ALLAI. https://allai.nl/council-of-europe-calls-for-a-ban-on-certain-facial-recognition-applications/ (2021). Accessed 25 Nov 2021
30. Pascu, L.: France looks to establish legal framework to deploy biometric video surveillance. https://www.biometricupdate.com/202001/france-looks-to-establish-legal-framework-to-deploy-biometric-video-surveillance (2020). Accessed 29 Dec 2021
31. Van Dijk, N., Casiraghi, S., Gutwirth, S.: The 'Ethification' of ICT. Governance artificial intelligence and data protection in the European Union. Comput. Law Secur. Rev. **43**, 1057 (2021). https://doi.org/10.1016/J.CLSR.2021.105597
32. Floridi, L.: Translating principles into practices of digital ethics: Five risks of being unethical. Philos. Stud. Ser. **144**, 81–90 (2021). https://doi.org/10.1007/978-3-030-81907-1_6
33. Moor, J.H.: The future of computer ethics: You ain't seen nothin' yet! Ethics Inf. Technol. **3**(2), 89–91 (2001). https://doi.org/10.1023/A:1011881522593
34. High-Level Expert Group on AI.: Ethics guidelines for trustworthy Artificial Intelligence. https://digital-strategy.ec.europa.eu/en/library/ethics-guidelines-trustworthy-ai (2019). Accessed 24 Nov 2021
35. European Commission, Directorate-General for Communications Networks, Content and Technology: The Assessment List for Trustworthy Artificial Intelligence (ALTAI) for self-assessment. Publications Office. https://data.europa.eu/doi/https://doi.org/10.2759/791819 (2020). Accessed 17 Nov 2021
36. Fjeld, J., Achten, N., Hilligoss, H., Nagy, A., Srikumar, M.: Principled artificial intelligence: mapping consensus in ethical and rights-based approaches to principles for AI. Berkman Klein Center Res. Publ. (2021). https://doi.org/10.2139/ssrn.3518482
37. Zuboff, S.: The Age of Surveillance Capitalism: The Fight for a Human Future at the New Frontier of Power. Profile Books, London (2016)
38. Creemers, R.: China's social credit system: an evolving practice of control. SSRN J. (2018). https://doi.org/10.2139/ssrn.3175792
39. Zhou, J., Gandomi, A.H., Chen, F., Holzinger, A.: Evaluating the quality of machine learning explanations: a survey on methods and metrics. Electronics (Switzerland) **10**(5), 1–19 (2021). https://doi.org/10.3390/electronics10050593

40. Tavani, H.T.: Ethics and Technology: Controversies, Questions, and Strategies For Ethical Computing. Wiley, Hoboken (2007)
41. Buolamwini, J., Gebru, T.: Gender shades: intersectional accuracy disparities in commercial gender classification. Proc. Mach. Learn. Res. **81**, 1–15 (2018)
42. Introna, L.D., Wood, D.: Picturing algorithmic surveillance: the politics of facial recognition systems. Surveill. Soc. **2**(23), 177–198 (2004). https://doi.org/10.24908/ss.v2i2/3.3373
43. Marciano, A.: Reframing biometric surveillance: from a means of inspection to a form of control. Ethics Inf. Technol. **21**(2), 127–136 (2019). https://doi.org/10.1007/s10676-018-9493-1
44. Murray, H.: Monstrous play in negative spaces: illegible bodies and the cultural construction of biometric technology. Commun. Rev. **10**(4), 347–365 (2007). https://doi.org/10.1080/10714420701715415
45. Waelen, R.A.: The struggle for recognition in the age of facial recognition technology. AI and Ethics (2022). https://doi.org/10.1007/s43681-022-00146-8
46. Roberts, S.R.: The biopolitics of China's "war on terror" and the exclusion of the Uyghurs. Crit. Asian Stud. **50**(2), 232–258 (2018)
47. Lyon, D.: Surveillance as Social Sorting: Privacy, Risk, and Digital Discrimination. Routledge, London (2002)
48. Van der Ploeg, I.: Biometrics and the body as information: Normative issues of the socio-technical coding of the body. In: Lyon, D. (ed.) Surveillance as Social Sorting, pp. 57–73. Taylor & Francis, London (2005)
49. Van Noorden, R.: The ethical questions that haunt facial-recognition research. Nature **587**(7834), 354–359 (2020). https://doi.org/10.1038/d41586-020-03187-3
50. Magnet, S., Rodgers, T.: Stripping for the State. Fem. Media Stud. **12**(1), 101–118 (2011). https://doi.org/10.1080/14680777.2011.558352
51. Mayhew, S.: Danish football stadium deploys Panasonic facial recognition to improve fan safety. https://www.biometricupdate.com/201907/danish-football-stadium-deploys-panasonic-facial-recognition-to-improve-fan-safety (2019). Accessed 11 Nov 2021
52. Montag, B. L., Mcleod, R., Mets, L. De, Gauld, M., Rodger, F., & Pełka, M.: The rise and rise of biometric mass surveillance in the EU. EDRi. https://edri.org/our-work/new-edri-report-reveals-depths-of-biometric-mass-surveillance-in-germany-the-netherlands-and-poland/ (2021). Accessed 23 November 2021
53. Parliament, E.: Charter of Fundamental Rights of the European Union. Office for Official Publications of the European Communities, Luxembourg (2000)
54. Friedman, B., Hendry, D.G., Borning, A.: A survey of value sensitive design methods. Found. Trends Hum. Comput. Interact. **11**(2), 63–125 (2017). https://doi.org/10.1561/1100000015
55. Rao, U., Nair, V.: Aadhaar: governing with biometrics, South Asia. J. South Asia Studies **42**(3), 469–481 (2019). https://doi.org/10.1080/00856401.2019.1595343