Cureus Journal of Computer Science
https://doi.org/10.7759/s44389-026-00146-3

Review Article

# A Review of Privacy, Ethics, and Governance Challenges Associated With Camera-Based AI Surveillance in Primary Schools

Misbah Zulfiqar [1] , Umair Iqbal [2,]

1. *Primary School Education, Independent Researcher, Gillieston Heights, AUS*
2. *Applied Artificial Intellgence, University of Newcastle, Newcastle, AUS*

Received: April 19, 2026 | Review began: May 10, 2026 | Review ended: June 25, 2026 | Published: July 01, 2026



## Abstract

The increasing integration of artificial intelligence (AI) into educational environments has led to the emergence of camera-based monitoring systems that combine visual data capture with advanced computer vision techniques. In primary school settings, these systems are used for purposes such as attendance automation, behaviour tracking, and safety monitoring. While such applications are often justified in terms of efficiency and security, their deployment raises significant concerns regarding privacy, ethics, and the rights of children. This paper performs a structured review to examine the implications of camera-based AI in primary education through a critical and interdisciplinary lens. It analyses the unique vulnerabilities of children as data subjects, emphasising their limited capacity for informed consent and the long-term consequences of continuous surveillance on their development and autonomy. The study further evaluates the extent to which existing regulatory frameworks, including the General Data Protection Regulation (GDPR), the Children's Online Privacy Protection Act (COPPA), and the Australian Privacy Act 1988, adequately address the risks associated with biometric data collection in school environments. By synthesising insights from surveillance studies, data protection law, and education ethics, the paper identifies key gaps in current governance approaches and proposes a set of guiding principles for responsible deployment. It argues that the use of camera-based AI in primary schools must be approached with heightened caution, ensuring that technological innovation does not undermine the fundamental rights, dignity, and developmental well-being of children.

**Categories:** Regulatory Compliance and Governance, Security and Privacy, Image Processing and Analysis

**Keywords:** ai in education, camera based monitoring, computer vision, child privacy, ethics of surveillance, data protection, educational technology governance, student autonomy, ethical ai, ai surveillance

## Introduction And Background

Artificial intelligence (AI) is reshaping the institutional landscape of education in ways that were scarcely conceivable a decade ago. From adaptive learning platforms to automated administrative systems, schools are increasingly integrating AI-driven tools into their operational and pedagogical frameworks. Among the most consequential developments in this domain is the deployment of camera-based AI systems, which combine closed-circuit television (CCTV) infrastructure

**How to cite this article:**
Zulfiqar M, Iqbal U (July 01, 2026) A Review of Privacy, Ethics, and Governance Challenges Associated With Camera-Based AI Surveillance in Primary Schools. Cureus J Comput Sci 3 : es44389-026-00146-3. DOI https://doi.org/10.7759/s44389-026-00146-3

with advanced computer vision algorithms to monitor students, track behaviour, automate attendance, and respond to perceived security threats [1,2]. Advocates of these technologies emphasise their potential to enhance campus safety, streamline administrative tasks, and generate data that may support safeguarding outcomes.

However, the adoption of such systems within primary school environments raises profound questions that extend well beyond their technical capabilities. Primary schools occupy a unique position in society. They are not merely instructional institutions but formative environments in which children develop their identities, social relationships, and foundational understandings of trust, authority, and community. The introduction of pervasive camera-based surveillance into these spaces therefore carries implications that differ qualitatively from those associated with surveillance in adult public or commercial settings. Children, unlike adults, cannot meaningfully consent to the collection of their biometric and behavioural data, nor can they be expected to fully comprehend or contest the conditions under which that data is gathered, stored, and used [3,4].

The rapid proliferation of camera-based AI in educational settings has outpaced the development of ethical norms and regulatory frameworks capable of governing it. Existing data protection legislation, including the European Union's General Data Protection Regulation (GDPR), the United States Children's Online Privacy Protection Act (COPPA), and Australia's Privacy Act 1988, was not designed with AI-driven school surveillance specifically in mind, and significant ambiguities remain regarding how these instruments apply to the collection of children's biometric data in educational contexts [5,6]. This regulatory gap creates conditions in which schools may deploy powerful surveillance technologies with inadequate understanding of the risks involved or the protections that ought to accompany them.

This review is motivated by three interconnected research questions. First, what are the specific privacy and ethical risks associated with the deployment of camera-based AI systems in primary schools? Second, how adequately do current legal and regulatory frameworks address those risks, particularly in relation to the rights of children? Third, what ethical principles and governance guidelines are needed to ensure that, where such technologies are deployed, they are done so in a manner consistent with the best interests of children and the values of democratic, rights-respecting societies? In pursuing these questions, this paper draws on scholarship from surveillance studies, education technology ethics, data protection law, and developmental psychology, seeking to produce a synthesised analysis that is both theoretically grounded and practically oriented.

**Research gap**

Despite the rapidly expanding literature on AI in education and surveillance technologies, a clear and targeted synthesis of the specific challenges associated with camera-based AI in primary school environments remains limited. Existing research tends to address adjacent domains in isolation, including learning analytics, higher education surveillance, or general data protection issues, without adequately accounting for the unique developmental, legal, and ethical context of primary education.

First, much of the technical and policy literature focuses on adult populations or higher education settings, where individuals possess greater autonomy and capacity for informed consent. This leaves a critical gap in understanding how continuous biometric surveillance affects younger children, whose cognitive and emotional development renders them particularly vulnerable to long-term psychological and social consequences. Second, while legal analyses have examined frameworks such as the GDPR, COPPA, and national privacy laws, there is insufficient integration between legal scholarship and educational ethics. Specifically, limited work has systematically evaluated whether existing regulatory mechanisms are normatively adequate when applied to AI-driven biometric monitoring in compulsory school environments. Third, although concerns regarding algorithmic bias, surveillance normalisation, and data commodification have been widely discussed in broader AI ethics literature, their combined implications for primary school settings remain underexplored. Research on algorithmic bias, surveillance capitalism, and data-driven schooling highlights critical risks, yet few studies synthesise these concerns into a unified governance perspective tailored to primary education.

**How to cite this article:**
Zulfiqar M, Iqbal U (July 01, 2026) A Review of Privacy, Ethics, and Governance Challenges Associated With Camera-Based AI Surveillance in Primary Schools. Cureus J Comput Sci 3 : es44389-026-00146-3. DOI https://doi.org/10.7759/s44389-026-00146-3

This review addresses these gaps by providing a structured and integrative analysis of camera-based AI surveillance in primary schools, combining insights from surveillance studies, legal frameworks, developmental psychology, and education ethics. In doing so, it moves beyond fragmented discussions toward a coherent framework for assessing the risks, limitations, and governance requirements of such technologies in child-centred educational environments.

## Review

### Methodology

This study employs a systematic review methodology following the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) framework [7]. Given that the subject spans multiple disciplines, including computer science, legal studies, education policy, ethics, and developmental psychology, a structured multi-database search strategy was adopted to ensure comprehensive and replicable coverage of the relevant literature.

Electronic searches were conducted across five academic databases: Scopus, ACM Digital Library, IEEE Xplore, PsycINFO, and ERIC. These databases were selected to capture the interdisciplinary character of the field, ensuring coverage of technical, social science, psychological, and legal literature. Search strings were constructed using Boolean operators combining terms from three conceptual domains: surveillance technology ("AI surveillance", "facial recognition", "computer vision", "emotion detection", "behaviour tracking", "biometric data"); educational setting ("school", "classroom", "education", "K-12", "primary school", "children"); and normative concerns ("ethics", "privacy", "consent", "GDPR", "child rights", "bias", "algorithmic discrimination"). Searches were not date-restricted at the search stage; date-based exclusion criteria were applied during screening.

In addition to database searches, backward and forward citation chasing was performed on all included sources, and the reference lists of key review articles were examined manually. Authoritative non-peer-reviewed sources, including regulatory guidance from the Information Commissioner's Office (ICO), the Office of the Australian Information Commissioner (OAIC), the Federal Trade Commission (FTC), and the UN Committee on the Rights of the Child, were also included where directly relevant to the research questions. Given the limited availability of directly comparable empirical studies, the review necessarily incorporates evidence from adjacent domains, including general surveillance studies, social psychology, and higher education contexts. These sources were included on the basis of conceptual relevance, with careful consideration given to their applicability and limitations when extrapolated to primary school environments.

**How to cite this article:**
Zulfiqar M, Iqbal U (July 01, 2026) A Review of Privacy, Ethics, and Governance Challenges Associated With Camera-Based AI Surveillance in Primary Schools. Cureus J Comput Sci 3 : es44389-026-00146-3. DOI https://doi.org/10.7759/s44389-026-00146-3

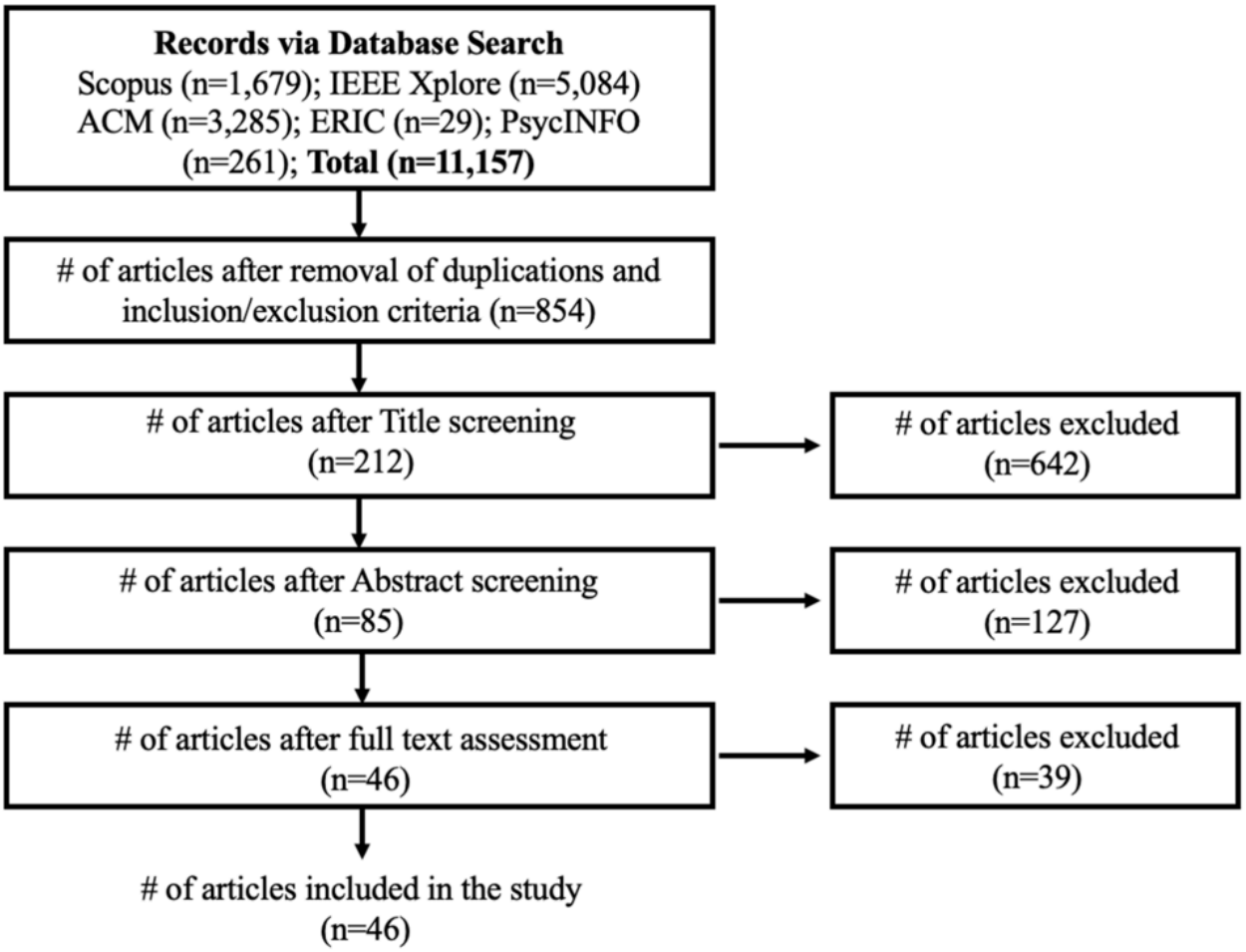


**FIGURE 1: PRISMA flow diagram illustrating the study selection process.**

PRISMA, Preferred Reporting Items for Systematic reviews and Meta-Analyses

The database search yielded approximately 11,157 records across all sources (see Figure *1*). After removing duplicates and applying inclusion and exclusion criteria, 854 unique records remained for title and abstract screening. At Stage 1, 212 records were excluded as irrelevant to the research questions, leaving 270 articles for full-text assessment. During Stage 2, a further 127 articles were excluded for reasons such as lack of relevance to primary school contexts, absence of an ethical or governance focus, or insufficient methodological rigour. The final set comprised 46 sources, including peer-reviewed articles, legal analyses, and authoritative regulatory documents.

Following de-duplication, a two-stage screening process was applied. In Stage 1, titles and abstracts were screened against the criteria set out in Table *1* by the primary reviewer, with a random 20% sample independently assessed to check consistency. In Stage 2, full-text review was conducted for all sources that passed Stage 1 screening. Disagreements were resolved through iterative discussion and reference to the inclusion criteria. Seminal theoretical works published before 2000, including Westin [8] on informational privacy, Foucault [9] on disciplinary institutions, and Nissenbaum [10] on contextual integrity, were retained on the grounds that they provide foundational conceptual frameworks without which the literature cannot be adequately interpreted. Their inclusion was determined at the full-text review stage on the basis of demonstrable relevance and sustained citational presence in the contemporary literature.

**How to cite this article:**
Zulfiqar M, Iqbal U (July 01, 2026) A Review of Privacy, Ethics, and Governance Challenges Associated With Camera-Based AI Surveillance in Primary Schools. Cureus J Comput Sci 3 : es44389-026-00146-3. DOI https://doi.org/10.7759/s44389-026-00146-3

To ensure the reliability and relevance of the included literature, a qualitative assessment of source credibility and contribution was conducted. Peer-reviewed journal articles were prioritised, and sources were evaluated based on their methodological clarity, theoretical contribution, and relevance to the research questions. Legal and regulatory documents were assessed based on their authoritative status and applicability to children's data protection and AI governance. Given the interdisciplinary nature of the review, formal scoring tools were not applied; instead, a structured narrative assessment approach was used to ensure consistency across diverse source types.

| Inclusion criteria | Exclusion criteria |
|---|---|
| Peer-reviewed journal articles, book chapters, and authoritative regulatory documents | Grey literature, blog posts, or non-peer-reviewed commentary without institutional affiliation |
| Published between 2000 and 2024 (seminal pre-2000 works retained on theoretical grounds) | Studies focused exclusively on higher education or adult workplace surveillance |
| Studies addressing AI, computer vision, biometric data, or digital surveillance in K-12 or primary school contexts | Works not available in English |
| Works addressing ethical, legal, psychological, or policy dimensions of educational technology | Technical system papers with no ethical, social, or policy component |
| Legal and regulatory texts with direct relevance to children's data protection | Studies predating widespread digital surveillance (before 1995) unless foundational theoretical works |

**TABLE 1: Inclusion and exclusion criteria.**

## Technology overview

Camera-based AI systems in educational settings are built upon the integration of physical imaging hardware with machine learning algorithms capable of processing visual data to identify, classify, and respond to persons, objects, and behaviours in real time or near-real time. Understanding their basic architecture is necessary to appreciate the privacy and ethical risks they generate, and to evaluate the adequacy of governance responses.

*Facial recognition systems* identify individuals by mapping facial geometry and comparing the resulting biometric templates against stored databases. In schools, these systems have been deployed to automate student attendance, control access to buildings, and verify identities at entry points [11,12]. The accuracy of such systems is sensitive to a range of environmental and demographic factors, including lighting conditions, image resolution, age, and race, with established research demonstrating substantially lower accuracy rates for children, women, and individuals from non-white ethnic groups [13,14].

*Emotion detection technologies* extend this capability by attempting to infer affective states, such as happiness, distress, attention, or boredom, from subtle variations in facial expression, gaze direction, and head posture [15]. The scientific validity of emotion detection systems has been questioned on fundamental grounds by a substantial body of psychological research, with Barrett et al. [16] demonstrating in a major review that facial actions do not reliably or consistently signal discrete emotional states across individuals, cultures, or contexts. The deployment of emotion detection in schools therefore rests on contested empirical foundations.

*Behaviour tracking systems* analyse movement patterns, body language, and spatial proximity across time to identify anomalies or predefined events. Attendance automation combines facial recognition with scheduling data to generate attendance records without manual roll calls. These systems operate through one of two principal architectural models: edge AI, which processes video data locally on-device or on-site; or cloud-based systems, which transmit footage or derived data to remote servers for processing. Edge AI reduces latency and limits external data transmission, while cloud-based systems offer greater computational power but introduce additional risks of interception, unauthorised access, and third-party data use [17,18].

Behaviour-tracking systems analyse movement patterns, body language, and spatial proximity over time to identify anomalies or predefined events. Attendance automation integrates facial recognition with scheduling data to generate attendance records without manual roll calls. These systems typically operate through one of two principal architectural models: edge AI, which processes video data locally on-device or on-site; or cloud-based systems, which transmit footage or derived data to remote servers for processing. Edge AI reduces latency and limits external data transmission, whereas cloud-based systems provide greater computational power but introduce additional risks of interception, unauthorised access, and third-party data use [17,18].

While these systems are often discussed collectively under the umbrella of camera-based AI surveillance, it is important to distinguish between them in terms of technical maturity, evidentiary support, and associated risk profiles. Facial recognition systems are relatively more mature and widely deployed, with well-documented performance characteristics and known biases, particularly across demographic groups. Behaviour tracking systems occupy an intermediate position, often relying on contextual interpretation of movement patterns, with varying levels of validation depending on the specific application. In contrast, emotion detection technologies remain highly contested, both in their underlying scientific assumptions and in their empirical validity, with substantial disagreement regarding their ability to reliably infer internal emotional states. These differences are significant because they imply that risks are not uniform across technologies, and governance responses must be tailored to the specific characteristics and limitations of each system rather than applied uniformly.

*Technical Foundations and Limitations of AI Surveillance Systems*

Camera-based AI systems in educational settings are typically built upon deep learning architectures, particularly convolutional neural networks (CNNs), which are designed to extract hierarchical features from visual data. In facial recognition systems, CNN-based models generate feature embeddings by mapping facial images into high-dimensional vector spaces, where similarity metrics are used to match individuals against stored biometric templates. These systems are trained on large datasets, and their performance depends heavily on the diversity and quality of training data.

Emotion detection systems extend this approach by applying supervised learning techniques to classify facial expressions into predefined emotional categories. These models rely on labelled datasets where facial landmarks, micro-expressions, and gaze patterns are correlated with emotional states. However, such models often assume a universal mapping between facial expressions and internal emotions, an assumption that has been challenged in psychological literature.

Behaviour tracking systems, in contrast, frequently combine computer vision with temporal modelling techniques such as recurrent neural networks (RNNs) or transformer-based architectures to analyse sequences of movement over time. These systems identify patterns, anomalies, or predefined behavioural indicators (e.g., inattentiveness or proximity violations) based on spatial and temporal features extracted from video streams.

Despite these technical capabilities, all three categories of systems exhibit significant limitations. First, model accuracy is highly sensitive to training data bias. Datasets used to train facial recognition and emotion detection systems have historically underrepresented children and diverse demographic groups, leading to systematic disparities in error rates [13,14]. Second, model generalisation across environments is limited: variations in lighting, camera angle, occlusion (e.g., masks, movement), and classroom dynamics can significantly degrade performance in real-world school settings. Third, the interpretability of these systems remains limited. Many deep learning models operate as “black boxes,” making it difficult for school administrators or educators to understand how specific outputs or classifications are generated. This

lack of transparency complicates accountability, particularly when automated decisions affect student records or behavioural assessments. Finally, technical bias is not limited to data imbalance but is also embedded in model design choices, label construction, and optimisation objectives. In emotion detection systems in particular, the imposition of predefined emotional categories reflects culturally and scientifically contested assumptions, introducing normative bias into ostensibly objective systems [16].

These technical limitations are directly relevant to the ethical and governance concerns discussed in subsequent sections. Inaccurate or biased outputs are not merely technical failures; in school contexts, they can translate into misidentification, unjust disciplinary actions, or distorted assessments of student well-being, thereby amplifying the risks associated with AI surveillance.

**Privacy risks in primary schools**

*Surveillance of Minors and Informational Privacy*

The academic literature on children's privacy in digital environments has expanded substantially over the past decade, driven by the proliferation of educational technology and the increasing datafication of childhood [3,19]. Westin's [8] foundational definition of privacy as the claim of individuals to determine for themselves when, how, and to what extent information about them is communicated to others remains influential, though scholars have increasingly questioned whether this liberal, autonomy-centred conception adequately addresses the specific circumstances of children [4,20].

Livingstone [4] argues that children's rights frameworks, and in particular the United Nations Convention on the Rights of the Child (UNCRC), establish a distinct normative basis for children's privacy that supplements rather than displaces adult-centred privacy theory. Under Article 16 of the UNCRC, children have the right to protection from arbitrary or unlawful interference with their privacy, a right that the UN Committee on the Rights of the Child [21] has interpreted as applying with full force to digital environments. The Committee's General Comment No. 25, issued in 2021, explicitly addresses AI and digital data collection in relation to children, providing authoritative interpretive guidance that has not yet been incorporated into the domestic legislation of most jurisdictions.

The concept of dataveillance, the systematic monitoring of personal data to regulate or manage behaviour [22], has been extensively theorised in relation to adult populations but has received increasing application to educational settings. Williamson [1] coins the term "data-driven schooling" to describe the emergence of educational governance practices oriented around the collection, analysis, and use of student data, arguing that these practices reconfigure the school as a site of quantification rather than formation. This reorientation has profound implications for the kind of educational experience children receive and for the institutional values that school environments transmit.

Surveillance in primary schools is further complicated by what Foucault [9] identified, in the context of prisons, hospitals, and schools, as the disciplinary function of visibility. Foucault's analysis of the Panopticon, a design in which observed individuals cannot determine whether they are being watched at any given moment, has been extensively applied to surveillance technologies that create similar effects of permanent visibility without continuous active monitoring [23,24]. In primary school contexts, the Panoptic effect of camera-based AI may be particularly pronounced because children lack the social and cognitive resources to contextualise or resist its disciplinary implications.

*Contextual Integrity and the Breach of Educational Norms*

Nissenbaum's [10] theory of contextual integrity has become one of the most widely cited frameworks in privacy scholarship and has received extensive application in educational technology contexts [20,25,26]. The theory holds that privacy violations occur not simply when information about an individual is disclosed, but when that disclosure violates the norms of information flow appropriate to the context in which the information was originally shared. Applied to schools, the theory implies that children share their physical presence, emotional expressions, and behavioural patterns within a context governed by norms of trust, learning, care, and developmental support. The extraction of that same information into algorithmic databases and surveillance profiles represents a fundamental breach of contextual norms, regardless of the stated purpose of the collection.

**How to cite this article:**
Zulfiqar M, Iqbal U (July 01, 2026) A Review of Privacy, Ethics, and Governance Challenges Associated With Camera-Based AI Surveillance in Primary Schools. Cureus J Comput Sci 3 : es44389-026-00146-3. DOI https://doi.org/10.7759/s44389-026-00146-3

Prinsloo and Slade [25] extend the contextual integrity framework to learning analytics, arguing that even when data collection appears to serve pedagogical goals, the translation of classroom behaviour into algorithmic profiles violates the contextual norms within which that behaviour occurred. Their analysis suggests that the key question is not whether the data is used for educational purposes but whether the flow of information matches the norms of the context in which it was shared. Camera-based AI that converts a child's facial expression into an emotional data point, or their movement across a classroom into a behaviour profile, represents precisely the kind of contextual integrity violation Nissenbaum's framework identifies.

Drachsler and Greller [26] raise a related concern about what they term the "consent paradox" in educational data: students, and in primary school contexts their parents, are often required to consent to data collection as a condition of school attendance, rendering consent structurally coercive. This concern has direct legal implications under the GDPR, which requires that consent be freely given and has specifically addressed the coercive character of consent obtained in the context of public services [27]. The Swedish Data Protection Authority's 2019 ruling on facial recognition attendance in schools turned precisely on this issue, finding that consent obtained in a compulsory school environment could not satisfy the freely given standard [28].

*Function Creep and Surveillance Capitalism*

Function creep, the expansion of a technology's use beyond its original purpose through incremental and often unnoticed extensions [23,29], represents one of the most consistently identified risks in the literature on educational AI surveillance. Koops [29] provides a systematic taxonomy of function creep in surveillance systems, distinguishing between purpose creep (using data for purposes beyond those stated at collection), audience creep (sharing data with entities not originally envisaged), and scope creep (expanding the range of data collected as the technical infrastructure is established). All three forms are documented in educational technology contexts.

Zuboff's [30] theory of surveillance capitalism provides a structural explanation for why function creep in educational data systems is not merely a risk but a predictable consequence of the commercial incentives governing technology provision. Zuboff argues that the extraction and commodification of behavioural data is the foundational logic of the dominant technology sector, and that educational institutions adopting AI systems from commercial vendors are incorporating this logic into their own governance. The implications for primary schools are significant: vendor contracts may include provisions permitting data sharing for product improvement or advertising that schools neither fully understand nor adequately scrutinise.

Williamson et al. [31] extend this analysis to identify what they call the "platform university" logic permeating educational technology provision, arguing that schools adopting AI platforms are participating in broader processes of digital commodification that transform student data into tradeable assets. While their analysis is oriented toward higher education, its implications for primary schools, where the power asymmetry between institutions and students is even more pronounced, are, if anything, more concerning. The relative absence of digital literacy and advocacy resources in primary school communities makes effective resistance to these dynamics particularly difficult.

*Data Security and Breach Risks*

Biometric data occupies a unique position in data security risk assessment because, unlike passwords or addresses, it cannot be changed if compromised. Solove [32] argues that privacy self-management frameworks, which ask individuals to protect themselves through consent choices, are structurally inadequate for addressing biometric data risks because the long-term consequences of biometric compromise cannot be assessed at the time of initial data collection. This argument has particular force in the context of children, whose biometric data collected in primary school may remain identifiable and exploitable throughout their adult lives.

The security literature documents a consistent pattern of inadequate data protection in educational technology systems. Verizon's annual Data Breach Investigations Report has repeatedly identified the education sector as one of the most frequently breached industries, with limited IT security capacity and large quantities of sensitive personal data creating

attractive targets [33]. The introduction of biometric surveillance systems into this sector substantially elevates the sensitivity of the data at risk.

**Ethical concerns**

*Utilitarian Analysis*

From a utilitarian perspective, the central question is whether the aggregate benefits of camera-based AI surveillance in schools outweigh the aggregate harms. Proponents argue that such systems can improve school safety, identify distressed students, and reduce administrative burdens, thereby increasing the time teachers can dedicate to instruction. In some contexts, these technologies are also framed as enabling earlier intervention in student well-being and enhancing institutional responsiveness to safeguarding risks. However, the empirical evidence supporting these benefits remains uneven and context-dependent. In addition, utilitarian calculation in this domain is significantly complicated by the difficulty of quantifying harms that are diffuse, probabilistic, and long-term, including psychological effects on children, chilling effects on behaviour and expression, and the normalisation of surveillance as a feature of social life [34].

Crawford [35] argues that the instrumental framing of AI benefits systematically underestimates distributional and structural harms. An AI attendance system may produce aggregate efficiency gains, but if its error rates are concentrated among students from particular ethnic groups, the aggregate benefit calculation conceals serious injustice. Eubanks [36] makes a parallel argument in relation to automated decision-making systems targeting marginalised communities, demonstrating that the populations bearing the costs of algorithmic systems are rarely the populations who design, commission, or benefit from them. In primary schools, children from minority groups, who are disproportionately subject to misidentification by facial recognition and emotion detection systems [13], may bear costs that are invisible in aggregate benefit assessments.

*Deontological Analysis*

Deontological ethical frameworks hold that certain actions are intrinsically wrong regardless of their consequences, because they violate fundamental duties or treat persons as means rather than ends. From a Kantian perspective, the continuous biometric monitoring of children without meaningful consent treats them as objects of administrative management rather than as persons with inherent dignity and developing autonomy [37,38]. O'Neill's influential reformulation of Kantian ethics around the concept of principled action emphasises that genuine respect for persons requires not merely that their interests be weighed but that they be treated as agents capable of rational deliberation. Children are developing rational agents, and educational environments ought to support that development rather than subordinate it to institutional monitoring imperatives.

The concept of children's rights as a distinct moral and legal category, grounded in their developmental vulnerability and the social obligations this generates, provides an additional deontological basis for scepticism about AI surveillance in primary schools. The UN Convention on the Rights of the Child establishes that actions concerning children shall be governed by the best interests of the child as a primary consideration (Article 3). The deployment of camera-based AI surveillance in primary schools must, on this standard, be justified specifically in relation to children's interests, not merely in relation to institutional efficiency or adult safety preferences. The literature reviewed does not identify evidence that such child-centred justification has been systematically provided in any documented deployment.

*Virtue Ethics and Educational Environments*

Virtue ethics, associated with the Aristotelian tradition and its contemporary revivals [39,40], asks what kind of character and what kind of communities our practices express and cultivate. Applied to education, the tradition asks what virtues a school is modelling and what kind of human beings its practices are forming. A school that deploys pervasive AI surveillance invites reflection on the compatibility of that infrastructure with educational values of trust, care, intellectual courage, and the development of moral agency.

Noddings' [40] ethics of care offers a particularly relevant lens, arguing that educational institutions ought to be organised around relationships of authentic care between teachers and students, relationships characterised by attention, responsiveness, and mutual trust. A surveillance infrastructure premised on monitoring and algorithmic assessment of children's emotional and behavioural states is structurally in tension with care ethics, which requires that the other be known through relationship rather than through data. The introduction of emotion-detection systems into classrooms arguably substitutes algorithmic inference for the relational knowledge that care ethics regards as essential to good teaching.

*Algorithmic Bias, Psychological Harm, and Power*

The empirical literature on algorithmic bias in facial recognition and emotion detection systems is now substantial and consistent in its findings. Buolamwini and Gebru's [13] Gender Shades study demonstrated error rate differentials of up to 34 percentage points between the best-performing demographic group and the worst-performing group (darker-skinned females) across three commercial facial analysis systems. Raji and Buolamwini [14] found that subsequent industry efforts to improve accuracy reduced but did not eliminate these differentials. The implications for primary schools are direct: AI systems that misidentify children from minority ethnic groups at higher rates will produce administrative records and behavioural flags that disproportionately affect those children, with potentially lasting consequences for their educational trajectories.

The psychological effects of surveillance on children have been investigated from multiple disciplinary perspectives. Foucault's [9] theoretical account of the normalising function of surveillance has been supported by empirical research in social psychology demonstrating that awareness of being observed increases self-monitoring, reduces spontaneity, and promotes conformity to perceived norms. In children, who are at sensitive developmental stages for identity formation, social learning, and the development of authentic self-expression, these effects may be significant, although current evidence is largely theoretical or extrapolated from adjacent domains and has not yet been consistently demonstrated in primary school AI surveillance contexts. Marx [41] argues that the harm of surveillance is not merely informational but experiential: being watched changes how one inhabits social space, independently of whether the surveillance data is ever used.

The power imbalance inherent in school surveillance deserves explicit analysis. Children in primary schools cannot exit the institution, cannot meaningfully consent to its surveillance practices, cannot access or contest the data it holds about them, and cannot organise collectively to resist surveillance policies. This structural powerlessness places primary school children in a position of radical vulnerability relative to AI surveillance systems, a vulnerability that is further intensified when, as is frequently the case, surveillance systems are procured and managed by school administrators without meaningful involvement of teachers, parents, or community representatives [2].

**Legal and regulatory landscape**

The governance of camera-based AI in schools operates within a complex and fragmented regulatory landscape. Table *2* provides a comparative overview of the three principal regulatory frameworks examined in this review.

| Dimension | GDPR (EU) | COPPA (USA) | Privacy Act 1988 (Australia) |
|---|---|---|---|
| Child age threshold | 13-16 years (member state variation) | Under 13 | No specific threshold |
| Biometric data classification | Special category; explicit consent required | Personal information; verifiable parental consent | Sensitive information; enhanced obligations |
| Data minimisation | Explicit principle (Article 5) | Implied; no explicit rule | APP 3: reasonably necessary |
| School consent authority | Schools may consent only for educational purposes; freely given standard applies | Schools may act as agents for parents for school-authorised use | No specific provision; general consent principles apply |
| Facial recognition in schools | Presumptively prohibited; DPA enforcement confirmed (Sweden, 2019) | Not explicitly addressed | Not explicitly addressed; reform pending |
| Enforcement record | Active; significant penalties imposed | Moderate; FTC enforcement actions | Limited; reform under review |

**TABLE 2: Comparative regulatory framework.**

COPPA, Children's Online Privacy Protection Act; FTC, Federal Trade Commission; GDPR, General Data Protection Regulation

*The General Data Protection Regulation*

The GDPR, in force since May 2018, is the most comprehensive and influential data protection framework and has served as a global reference point for regulatory reform in many jurisdictions. The GDPR establishes a range of principles directly relevant to school AI surveillance: personal data must be processed lawfully, fairly, and transparently; collected for specified, explicit, and legitimate purposes; and limited to what is necessary for those purposes [42]. The Regulation classifies biometric data processed for the purpose of uniquely identifying a natural person as special category data subject to the most stringent processing conditions, including explicit consent (Article 9).

The GDPR's treatment of children's consent is set out in Article 8, which establishes that where processing is based on consent and directed at a child, that consent is valid only where the child is at least 16 years of age, or at least 13 where member states provide for a lower threshold. Children below that threshold require parental or guardian consent. The Article 29 Working Party's guidance on consent has clarified that consent in institutional contexts, where individuals have limited ability to refuse without adverse consequences, cannot satisfy the freely given standard. This guidance has direct and significant implications for facial recognition and biometric surveillance in schools, where attendance is compulsory and the power to withhold consent is therefore structurally absent.

The Swedish Data Protection Authority's 2019 ruling on facial recognition attendance in schools represents the most significant enforcement action to date in this domain. The Authority found that the school had processed biometric data without lawful basis, that parental consent could not be considered freely given in a compulsory school context, and imposed a 200,000 Swedish kronor fine. Mantelero [43] argues that this ruling signals a broader principle: that

institutional contexts characterised by power asymmetry and compulsory participation are structurally incompatible with consent-based processing of biometric data, and that alternative lawful bases, such as legitimate interests, are similarly unavailable given the severity of the privacy intrusion.

Recent regulatory developments in the EU have added further layers of relevant governance. The EU Artificial Intelligence Act, adopted in 2024 [44], classifies real-time remote biometric identification in publicly accessible spaces as a high-risk AI system subject to strict conformity requirements, and prohibits its use by law enforcement without judicial authorisation. While the Act does not explicitly address school deployments, its risk classification framework and the prohibitions it establishes for biometric AI in public spaces are highly relevant to the interpretation of permissible school deployments.

*COPPA and the United States Regulatory Framework*

The Children's Online Privacy Protection Act of 1998 establishes specific requirements for the collection of personal information from children under 13 in online contexts, including verifiable parental consent, clear privacy notices, data minimisation, and restrictions on disclosure to third parties [45]. While COPPA's original framing was oriented toward commercial online services, its principles have been applied to school technology contexts by the FTC, and the Family Educational Rights and Privacy Act (FERPA) provides additional protections for educational records.

Reidenberg et al. [46] conducted an empirical examination of school cloud computing contracts and found systemic inadequacy in the privacy protections afforded to student data, with most contracts granting vendors rights to use student data for commercial purposes without meaningful restriction. Their findings illustrate the gap between formal regulatory requirements and actual contractual practice in educational technology procurement, a gap that camera-based AI deployments are likely to replicate or amplify. Polonetsky and Jerome [47] argue that COPPA's consent framework is structurally inadequate for addressing school surveillance because it locates decision-making authority in parents rather than considering children as rights-bearing subjects whose own interests must be independently assessed.

The United States has not enacted a comprehensive federal data protection framework comparable to the GDPR, and the regulation of AI in educational settings remains fragmented across federal and state levels. Several states, including Illinois, with its Biometric Information Privacy Act (BIPA), and New York, have enacted more specific legislation restricting the use of biometric data in schools, providing examples of more targeted regulatory approaches [48]. The Illinois BIPA, which requires written consent prior to the collection or use of biometric data and provides a private right of action for violations, has been characterised by Peppet [48] as the most protective biometric privacy statute in the United States, and has been applied in several cases involving educational technology providers.

*The Australian Privacy Act and Reform Agenda*

The Privacy Act 1988 [49], together with the Australian Privacy Principles (APPs), establishes the primary framework governing the handling of personal information by government agencies and private sector organisations. APP 3 requires personal information to be collected only by lawful and fair means and only if reasonably necessary; APP 11 requires reasonable security measures to prevent misuse, loss, and unauthorised access [50]. The Act currently lacks specific provisions addressing biometric data collection from children in schools.

The Attorney-General's Department's 2023 Privacy Act Review Report identifies children's privacy as a priority area, recommending enhanced protections including a statutory tort for serious invasions of privacy, strengthened children's privacy code provisions, and specific regulation of children's data in digital services [50]. These recommendations are significant but have not yet been enacted. The Australian framework's reliance on administrative enforcement through the OAIC, without a private right of action, substantially limits its deterrent effect and leaves individuals, including children and their families, without effective remedial options.

Across all three frameworks, common themes emerge. Consent requirements, though structurally similar, are inconsistently applied to biometric and AI-generated data in school contexts. Data minimisation principles, which require that only the data necessary for a specified purpose be collected, are frequently honoured in their formal statement but difficult to enforce against complex AI systems that generate rich derivative data as a byproduct of their operation.

Children's rights, while acknowledged in all three frameworks, receive enhanced protection in the GDPR, weaker structural protection in COPPA, and are only partially addressed by the Australian Privacy Act in its current form. Taken together, the existing regulatory landscape provides an important but insufficient foundation for governing camera-based AI in primary schools.

## Case studies

### *AI Classroom Monitoring in China*

The deployment of AI classroom monitoring systems in China has been among the most extensively documented applications of educational surveillance technology globally. Systems deployed across multiple provinces have combined facial recognition with behaviour analysis to monitor student attention levels, emotional states, and classroom engagement in real time, with data fed to teachers and administrators on dashboards [51,52]. The Hangzhou No. 11 High School's "smart classroom behaviour management system" attracted significant media and scholarly attention, deploying cameras that tracked facial expressions and classified students into seven emotional categories, updating every 30 seconds.

Critical analysis of these deployments must attend to both their technical and structural dimensions. Technically, the systems rely on emotion detection algorithms whose scientific validity is contested [16], raising the prospect that the emotional data generated is not merely intrusive but unreliable. Structurally, the systems were deployed in an institutional context characterised by intense academic competition and high-stakes examinations, creating incentives for students to perform attentiveness rather than develop genuine engagement. Liang et al. [51] argue that the social credit logic underpinning many Chinese surveillance systems, in which behavioural compliance is rewarded and non-compliance penalised, maps directly onto educational AI monitoring, producing an environment oriented toward behavioural conformity rather than intellectual development.

The Chinese case also illustrates the normalisation concern identified in the theoretical literature. [52] reports that many students in surveilled classrooms did not initially object to the systems and in some cases endorsed them as fair mechanisms for demonstrating effort. This acceptance, however, may itself reflect the disciplinary effects of surveillance, consistent with Foucault's [9] analysis of how systems of visibility produce subjects who internalise their own monitoring.

### *Facial Recognition Attendance Systems in Western Contexts*

The Swedish case examined by the Data Protection Authority in 2019 provides a landmark regulatory reference point for facial recognition attendance in Western educational contexts. The Anderstorp High School in Skelleftea had deployed a facial recognition system for attendance monitoring across 22 students, framing it as a pilot study. The DPA's ruling that the processing was unlawful established several important principles: that biometric data processing in schools requires explicit consent satisfying the freely given standard; that this standard cannot be met in compulsory educational contexts; and that the educational purpose pursued does not constitute a sufficient justification for biometric processing in the absence of lawful basis.

Andrejevic and Selwyn [12] contextualise this case within the broader landscape of automated identification in education, arguing that the efficiency framing of facial recognition attendance obscures the fundamental transformation it represents in the relationship between students and institutions. Manual attendance-taking, however inconvenient, involves a social interaction between teacher and student; facial recognition substitutes algorithmic identification for that interaction, removing from the process both the possibility of discretion and the opportunity for relational connection. This is not merely an aesthetic concern: research in educational psychology consistently links teacher-student relationship quality to student engagement, well-being, and academic outcomes.

In the United States context, the Lockport City School District in New York attracted significant attention for its 2020 deployment of facial recognition technology in K-12 schools, becoming the first school district in the state to do so. The deployment generated immediate regulatory scrutiny from the New York State Education Department, which imposed a

moratorium on AI facial recognition in schools pending regulatory review. A subsequent independent evaluation found that the system had been deployed without adequate consideration of bias risks, without meaningful community consultation, and without legal analysis sufficient to satisfy state privacy requirements.

*Behaviour Tracking and Learning Analytics Platforms*

Behaviour tracking tools have been marketed to schools under various descriptors, including student well-being platforms and classroom management systems. Williamson [31] provides an extensive analysis of platforms such as ClassDojo, used in millions of classrooms globally, arguing that they constitute commercial surveillance infrastructure embedded in educational contexts. ClassDojo collects longitudinal data on student behaviour, social-emotional learning indicators, and family engagement, aggregating this data into profiles that may follow students across their educational careers. Williamson argues that the platform's framing of behaviour tracking as positive reinforcement for growth mindset development conceals its character as a commercial data collection operation.

Regan and Jesse [53] examine the ethical dimensions of personalised learning platforms that use continuous behavioural data to adapt content and pacing to individual students, arguing that while the educational goal is legitimate, the surveillance infrastructure it requires is disproportionate and creates serious risks of discriminatory tracking. Their analysis draws attention to the feedback loop between behavioural data and curricular assignment: systems that track student behaviour to personalise learning may simultaneously channel students into differentiated pathways based on algorithmic assessments that replicate existing inequalities. This concern has particular salience in primary schools, where early educational tracking can have lifelong consequences for educational and economic trajectories.

**Risk-benefit tradeoff**

A rigorous assessment of camera-based AI in primary schools requires systematic analysis of both the potential benefits attributed to these systems and the risks they generate. The following framework, presented in Figure *2*, attempts to organise this analysis in a manner that is useful for decision-makers while being honest about the significant uncertainties involved. The framework in Figure *3* organises the principal benefits and risks associated with camera-based AI in primary schools. These should be understood as frameworks for structured deliberation rather than as definitive empirical findings, given the limited and often contested evidence base for many claimed benefits.

The risk matrix in Figure *4* reveals a consistent pattern: the most probable risks associated with camera-based AI in primary schools are also among the most severe, while the benefits with the strongest evidentiary support tend to be incremental and do not require the collection of biometric or continuous behavioural data. This asymmetry is significant for policy and governance purposes. It suggests that the deployment of camera-based AI in primary schools cannot be justified on a straightforward cost-benefit analysis, but rather requires a far more demanding framework that treats children's rights and dignity as constraints upon, rather than factors to be weighed against, operational considerations.

A conceptual framework for decision-making in this domain might therefore distinguish between uses of camera-based technology that are capable of satisfying rigorous ethical and legal conditions, such as basic CCTV for physical security with strict data minimisation and retention limits, and uses that are presumptively impermissible without compelling and demonstrable justification, including facial recognition, emotion detection, and continuous behavioural tracking.

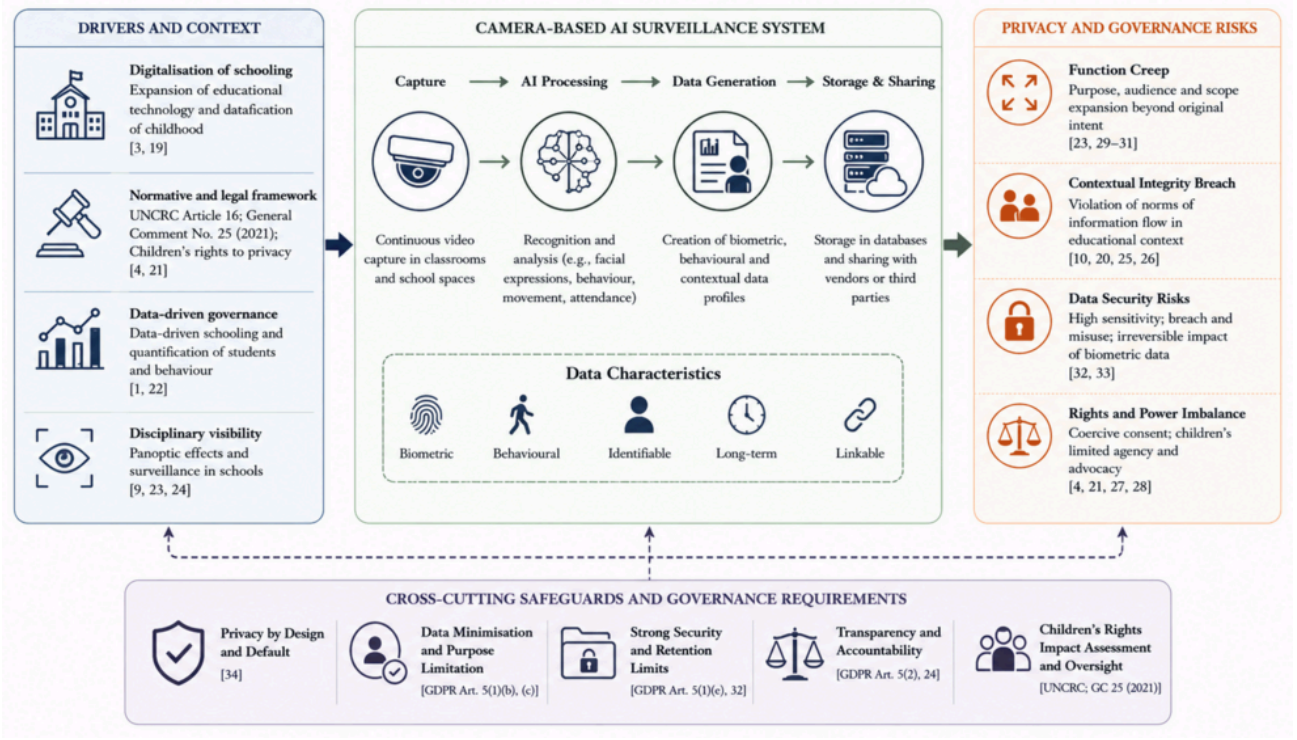


**FIGURE 2: Illustration of privacy risks related to use of camera-based AI in primary school settings.**

| Claimed Benefit | Associated Risk | Evidence Strength | Risk Severity |
|---|---|---|---|
| Enhanced campus safety | Normalisation of surveillance; chilling effect on behaviour | Limited | High |
| Automated attendance | Biometric data collection without meaningful consent | Moderate | High |
| Early distress detection | Inaccuracy of emotion detection; misclassification of children | Contested | High |
| Reduced admin burden | Function creep; data repurposed beyond original scope | Moderate | Medium |
| Real-time behavioural insight | Algorithmic bias; discriminatory impact on minority students | Limited | High |
| Deterrence of misconduct | Psychological harm; anxiety and self-censorship in children | Limited | Medium |

**FIGURE 3: Risk-benefit tradeoff framework.**

**How to cite this article:**
Zulfiqar M, Iqbal U (July 01, 2026) A Review of Privacy, Ethics, and Governance Challenges Associated With Camera-Based AI Surveillance in Primary Schools. Cureus J Comput Sci 3 : es44389-026-00146-3. DOI https://doi.org/10.7759/s44389-026-00146-3

| | Low Severity | Medium Severity | High Severity |
|---|---|---|---|
| **High Likelihood** | Unnecessary data retention | Function creep and data misuse | Biometric breach; discrimination |
| **Medium Likelihood** | Administrative inefficiency | Chilling effect on student expression | Normalisation of child surveillance |
| **Low Likelihood** | System downtime | Parental complaint without regulatory consequence | Third-party data sharing without consent |

Cells shared by overall risk level: red = critical, orange = high; green = moderate; teal = low

**FIGURE 4: A combination of severity and likelihood in the form of a risk matrix.**

## Proposed ethical framework and guidelines

Drawing on the theoretical analysis, empirical literature, and regulatory review in preceding sections, this review proposes seven principles constituting a governance framework for camera-based AI in primary schools. These principles draw on and extend existing frameworks (see Figure *5*), including Cavoukian's [54] Privacy by Design, the European Data Protection Board's guidelines on processing of biometric data, and the emerging child-centred AI design literature.

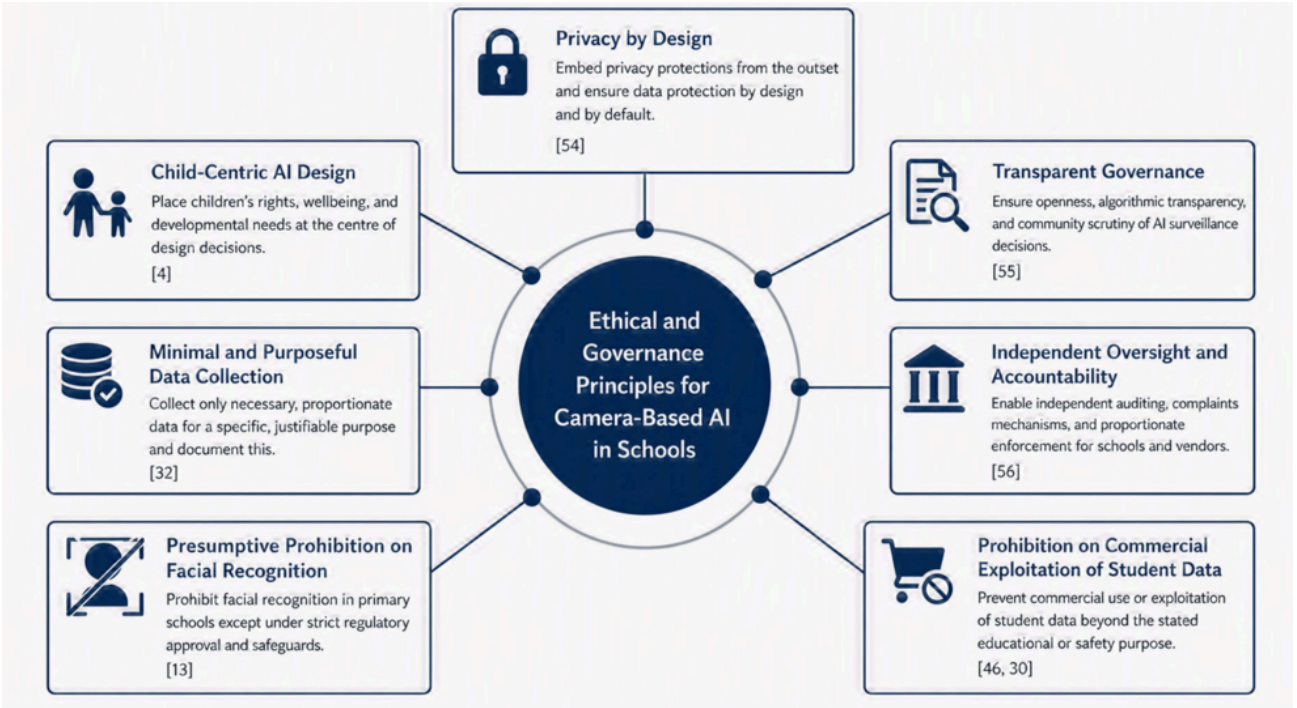


**FIGURE 5: Key ethical principles for the camera-based AI in primary schools.**

### *Privacy by Design*

Cavoukian's [54] Privacy by Design framework requires that privacy protections be embedded in the architecture of systems from the outset of their development, rather than added as compliance measures after design is complete. Applied to school camera-based AI, this principle entails that systems be designed to collect the minimum data necessary for their stated purpose; to process data on-device where technically feasible; to apply automatic deletion

schedules proportionate to operational needs; and to make privacy-protective configurations the default. Procuring schools and regulatory bodies should require demonstrated compliance with privacy-by-design standards, including third-party auditing, as a condition of deployment approval. The principle is consistent with GDPR requirements for data protection by design and by default (Article 25) and with the EU AI Act's requirements for high-risk AI systems.

*Child-Centric AI Design*

UNICEF's Policy Guidance on AI for Children argues for a child-centric approach to AI design that places children's rights and developmental needs, rather than adult or institutional preferences, at the centre of design decisions. This principle requires consultation with child development experts, education specialists, and children themselves in the design process; assessment of the emotional, psychological, and developmental implications of monitoring prior to deployment; and evaluation of systems against their impact on children's well-being, autonomy, and sense of safety. Livingstone [4] argue that child-centric digital design requires moving beyond the framing of children as objects of protection to recognise them as rights-bearing subjects with agency and interests that must be actively represented in design and governance processes.

*Minimal and Purposeful Data Collection*

The data minimisation principle, present in the GDPR (Article 5(1)(c)), COPPA, and Australian Privacy Principle 3, is extended here to require active justification for each category of data collected. Schools deploying camera-based AI should be required to document, for each system component, the specific educational or safety purpose it serves, why that purpose cannot be achieved through less intrusive means, and what data collection is proportionate to that purpose alone. This documentation should be subject to independent review before deployment and at regular intervals thereafter. Solove [32] argues that data minimisation requirements are only effective if they are operationalised through concrete documentation and verification requirements, rather than left as abstract principles subject to self-assessment.

*Presumptive Prohibition on Facial Recognition*

In light of the severity of risks associated with facial recognition, the absence of demonstrated educational benefits proportionate to those risks, the established inaccuracy of facial recognition systems for children and minority groups, and the impossibility of genuine consent in mandatory school environments, a presumptive prohibition on facial recognition in primary schools is warranted. This position is consistent with the regulatory findings of the Swedish DPA (2019), the New York State Education Department moratorium (2020), and the Algorithmic Justice League's recommendations regarding biometric AI in educational settings [13]. Where schools seek to deploy facial recognition for a compelling and specific purpose, such deployment should require explicit regulatory approval, independent ethical review, demonstrated proportionality, and robust safeguards including strict data retention limits and prohibition on data sharing with third parties.

*Transparent Governance*

Decisions about the adoption, configuration, and operation of camera-based AI in schools must be made through processes that are open to scrutiny by parents, students where developmentally appropriate, teachers, and the broader community. Schools should publish clear and accessible documentation of any AI surveillance systems in operation, including the purposes they serve, the data they collect, and the safeguards that apply. Algorithmic transparency, including publication of the algorithmic systems used, their performance metrics across demographic groups, and the criteria triggering automated alerts, is a necessary component of meaningful governance [55]. Governance structures should include mechanisms for ongoing review and for community-initiated reassessment of deployments that give rise to concerns.

*Independent Oversight and Accountability*

Independent oversight bodies, whether existing data protection authorities or newly established education technology regulators, should have the power to audit school AI deployments, investigate complaints, require modifications or cessation of non-compliant systems, and impose proportionate sanctions. Accountability mechanisms should extend to

technology vendors as well as to schools, given the commercial interests that drive the promotion of AI surveillance products in educational markets. The model of co-regulatory oversight, combining mandatory registration of school AI deployments with proactive auditing powers and reactive complaints investigation, is consistent with emerging international regulatory practice and with recommendations from the UN Special Rapporteur on the Right to Privacy [56].

*Prohibition on Commercial Exploitation of Student Data*

Data generated by AI surveillance systems in schools must not be available for commercial exploitation, advertising targeting, or any purpose unrelated to the specific educational or safety function for which it was collected. Reidenberg et al. [46] demonstrate that vendor contracts in educational technology routinely include provisions permitting data use for product improvement and advertising that schools neither fully understand nor adequately scrutinise. This principle requires both contractual safeguards in procurement agreements and regulatory enforcement capable of deterring violation. Zuboff's [30] analysis of surveillance capitalism suggests that contractual safeguards alone are insufficient given the structural incentives of commercial AI providers, and that regulatory prohibition with meaningful enforcement sanctions is the only reliable mechanism for preventing the commodification of children's educational data.

## Discussion

The systematic review of the literature on camera-based AI surveillance in primary schools reveals a field characterised by a significant imbalance between technological capability and governance readiness. However, this assessment should be interpreted in light of the relatively limited empirical evidence base, with much of the literature adopting theoretical, normative, or policy-oriented approaches rather than reporting direct observational or experimental findings. However, it is important to acknowledge that much of the evidence informing these concerns is derived from theoretical and extrapolative analyses rather than direct empirical observation in primary school environments. As a result, while the identified risks are plausible and well-supported conceptually, their real-world manifestation and magnitude may vary across contexts and implementations. However, given the limited availability of large-scale empirical studies in primary school contexts, these findings may not be uniformly generalisable across different educational systems, technological implementations, and socio-cultural environments.

A recurrent theme across the literature is the framing of surveillance as care. In the Chinese classroom monitoring context, in the behaviour-tracking platforms deployed in Western schools, and in the marketing of facial recognition for student safety purposes, AI surveillance is consistently presented as a technology of protection and support rather than as a technology of control. This rhetorical framing is significant because it makes critical scrutiny of surveillance systems more socially costly, aligning opposition to surveillance with indifference to children's well-being rather than with concern for their rights. At the same time, it should be recognised that the literature does not uniformly reject these technologies and, in some cases, identifies potential operational and safeguarding benefits, although these are often less systematically studied than associated risks.

The literature also reveals a systematic tendency to underestimate the long-term and cumulative effects of surveillance on children's development. Individual studies documenting the immediate psychological effects of surveillance, such as increased anxiety or reduced spontaneity, capture only a portion of the relevant harm. The normalisation of surveillance as a feature of institutional life, the habituation to biometric identification, and the incorporation of surveillance compliance into children's developing self-concepts are effects that operate over years or decades and may not be visible in short-term research designs. Lyon [23] argues that the deep cultural consequences of surveillance normalisation, including the erosion of expectations of privacy and the attenuation of the social norms that constrain institutional power, are precisely the effects that deserve the most sustained scholarly attention. Importantly, claims regarding psychological and developmental impacts remain primarily inferred rather than directly established through longitudinal or experimental studies in primary school environments and should therefore be interpreted with appropriate caution.

The review also identifies a significant gap in the literature regarding the perspectives of children themselves on AI surveillance in schools. The great majority of the sources reviewed adopt adult perspectives, whether those of scholars, regulators, or institutional actors. The limited research that has sought children's views on school surveillance suggests that children's attitudes are complex, contextually sensitive, and not reducible to simple acceptance or rejection [4].

Regan and Jesse [53] call for participatory research designs that incorporate children's voices into governance frameworks, arguing that child-centric AI governance requires not merely that adults consider children's interests but that children are meaningfully included in deliberative processes.

A critical limitation of this review is its reliance on sources published in English, which underrepresents the extensive deployment of AI surveillance in educational settings in China, India, the Gulf states, and other contexts where significant activity is occurring but where publication in accessible peer-reviewed venues remains limited. The geographic skew of the included literature towards European and North American contexts means that the governance frameworks draw primarily on liberal democratic regulatory traditions, and their applicability in other political and legal contexts warrants further investigation.

**Review limitations**

Several limitations of this review should be acknowledged. First, although the study adopts a structured systematic review methodology following PRISMA principles, the analysis is limited to sources available in English. This introduces a potential language bias and may underrepresent relevant research and policy developments in regions where AI surveillance in education is actively deployed, including China, India, and Middle Eastern contexts. Second, the interdisciplinary scope of the review, spanning computer science, law, education, and ethics, necessarily involves synthesising diverse forms of evidence with varying methodological standards. While this breadth is essential to address the research questions, it also limits the capacity to perform quantitative meta-analysis or directly compare findings across domains.

Third, the review relies predominantly on theoretical, normative, and secondary literature, with relatively limited availability of empirical studies directly examining the real-world impact of camera-based AI surveillance on primary school students. While this reflects the current state of the field, it constrains the ability to establish causal relationships between AI surveillance and specific psychological, behavioural, or educational outcomes. The absence of robust quantitative data, longitudinal studies, and controlled empirical designs limits the strength of conclusions and reduces the generalisability of findings across diverse educational contexts. As a result, the claims advanced in this review should be interpreted as analytically grounded and theoretically supported rather than empirically definitive. Fourth, the rapidly evolving nature of AI technologies and regulatory frameworks means that some aspects of the analysis may become outdated as new policies, technical capabilities, and deployment practices emerge. In particular, ongoing developments such as the implementation of the EU AI Act and national privacy reforms may significantly reshape the governance landscape in the near future.

Another limitation concerns the heterogeneity and indirect nature of the evidence base. Many of the claims advanced in this review, particularly those relating to psychological, behavioural, and developmental impacts, are derived from adjacent domains such as digital surveillance, social psychology, and higher education contexts, rather than directly from primary school AI surveillance implementations. As a result, the evidentiary chain often involves analogy and extrapolation rather than direct observation. While such interdisciplinary inference is necessary given the emerging nature of the field, it introduces uncertainty regarding the transferability and specificity of these findings to primary school environments.

A further limitation relates to potential selection and framing bias within the reviewed literature and its interpretation. While this review acknowledges potential benefits of camera-based AI in primary schools, including improved safety, administrative efficiency, and early identification of student well-being concerns, the analysis places greater emphasis on privacy, ethical, and governance risks. This reflects both the normative focus of much of the existing literature and the critical lens adopted in surveillance and data protection scholarship. However, this asymmetry may result in the underdevelopment of benefit-oriented perspectives and an overrepresentation of risk-based interpretations. Future research would benefit from more balanced empirical assessments that systematically evaluate both the advantages and limitations of these technologies in real-world educational settings.

**How to cite this article:**
Zulfiqar M, Iqbal U (July 01, 2026) A Review of Privacy, Ethics, and Governance Challenges Associated With Camera-Based AI Surveillance in Primary Schools. Cureus J Comput Sci 3 : es44389-026-00146-3. DOI https://doi.org/10.7759/s44389-026-00146-3

Finally, although the review incorporates case studies to contextualise key issues, these examples are limited in number and may not fully capture the diversity of implementation practices across different educational systems and socio-political contexts. These limitations highlight the need for further empirical, longitudinal, and regionally diverse research, as well as ongoing reassessment of governance frameworks as technologies and regulatory environments evolve.

## Conclusions

This paper has examined the deployment of camera-based AI surveillance systems in primary schools from ethical, legal, and governance perspectives. It has argued that such systems are associated with serious and, in some cases, potentially irreducible privacy and ethical risks. However, these conclusions are derived primarily from theoretical, legal, and interdisciplinary analysis, reflecting the current state of the literature rather than a comprehensive body of empirical evidence. As such, while the risks identified are well-supported conceptually, their precise magnitude, distribution, and long-term effects require further empirical validation. The analysis has demonstrated that camera-based AI in primary schools raises concerns that are qualitatively different from those associated with AI surveillance in adult or commercial contexts. The developmental vulnerability of children, their inability to provide meaningful consent, the formative character of the educational environment, and the power asymmetries inherent in institutional schooling combine to create conditions in which the presumptive burden of justification for AI surveillance should be high and the standards for permissible deployment correspondingly stringent. The proposed risk matrix illustrates that the most probable harms associated with these technologies, including biometric data breaches, discriminatory misidentification, function creep, and the normalisation of surveillance, are also among the most severe, while the evidentiary basis for the claimed benefits is often weak and contested. The proposed ethical framework offers a set of principles, including privacy by design, child-centric AI design, minimal data collection, a strong presumption against the use of facial recognition technologies in primary schools unless compelling, context-specific justification and robust safeguards can be demonstrated, transparent governance, independent oversight, and prohibition on commercial data use, that together constitute a foundation for responsible governance in this domain. These principles are not a counsel of technological abstinence; they are an argument that where technology is deployed in environments serving children, the standards of justification, design, and accountability must be commensurate with the rights and vulnerabilities of those children.

Future research should prioritise the development of empirically grounded studies, including longitudinal investigations, case-based evaluations, and user-centred research involving students, educators, and parents. Such work is essential to substantiate theoretical concerns, assess real-world impacts, and refine governance frameworks in ways that are responsive to lived educational contexts. At the same time, it is important to emphasise that the conclusions of this review are necessarily grounded in conceptual, legal, and interdisciplinary analysis rather than extensive empirical validation. While the risks identified are supported by established theories of surveillance, bias, and child development, the current lack of longitudinal and context-specific empirical studies limits the ability to definitively assess the scale and variability of these impacts. These policy-oriented recommendations are grounded in precautionary reasoning informed by ethical and legal analysis, rather than direct causal evidence, and should therefore be understood as normative guidance subject to further empirical validation. Future research incorporating real-world deployments, behavioural data, and user-centred perspectives will be essential to refine and validate these conclusions. Accordingly, the conclusions of this review should be understood as analytically robust but contingent upon a still-developing empirical evidence base, particularly with respect to causality and generalisability. The call for stricter regulation of camera-based AI in primary schools reflects not a rejection of technological innovation, but a recognition that current evidence, while compelling in its ethical and legal implications, remains incomplete from an empirical standpoint. Strengthening this evidence base is therefore a critical priority for future interdisciplinary research, alongside the continued development of governance frameworks that protect the rights and well-being of children. The conclusions presented in this review reflect a precautionary and analytically grounded interpretation of the current evidence base. However, important areas including psychological and developmental impacts, the effectiveness of specific interventions, and the realisation of claimed benefits remain insufficiently supported by direct empirical research in primary school contexts. As a result,

**How to cite this article:**
Zulfiqar M, Iqbal U (July 01, 2026) A Review of Privacy, Ethics, and Governance Challenges Associated With Camera-Based AI Surveillance in Primary Schools. Cureus J Comput Sci 3 : es44389-026-00146-3. DOI https://doi.org/10.7759/s44389-026-00146-3

some conclusions should be interpreted as theoretically plausible and normatively justified rather than empirically confirmed. Addressing these gaps through longitudinal, context-specific, and user-centred research represents a critical priority for future work in this field.

## Additional Information

### Author Contributions

All authors have reviewed the final version to be published and agreed to be accountable for all aspects of the work.

**Acquisition, analysis, or interpretation of data:** Umair Iqbal, Misbah Zulfiqar

**Critical review of the manuscript for important intellectual content:** Umair Iqbal

**Supervision:** Umair Iqbal

**Concept and design:** Misbah Zulfiqar

**Drafting of the manuscript:** Misbah Zulfiqar

### Disclosures

**Conflicts of interest:** In compliance with the ICMJE uniform disclosure form, all authors declare the following: **Payment/services info:** All authors have declared that no financial support was received from any organization for the submitted work. **Financial relationships:** All authors have declared that they have no financial relationships at present or within the previous three years with any organizations that might have an interest in the submitted work. **Other relationships:** All authors have declared that there are no other relationships or activities that could appear to have influenced the submitted work.

### Data Availability Statements

Data sharing is not applicable. This work is theoretical; no datasets were generated or analyzed.

## References

1. Williamson B: Big Data in Education: The Digital Future of Learning, Policy and Practice. SAGE Publications, London; 2017. 10.4135/9781529714920
2. Selwyn N: Should Robots Replace Teachers? AI and the Future of Education. Polity Press, Cambridge; 2019.
3. Lupton D, Williamson B: The datafied child: The dataveillance of children and implications for their rights. New Media & Society. 2017, 19:780-794. 10.1177/1461444816686328
4. Livingstone S: Children's use of the internet: Reflections on the emerging research agenda. New Media & Society. 2003, 5:147-166. 10.1177/1461444803005002001
5. Mantelero A: AI and Big Data: A blueprint for a human rights, social and ethical impact assessment. Computer Law & Security Review. 2018, 34:754-772. 10.1016/j.clsr.2018.05.017
6. Bygrave LA: Data Privacy Law: An International Perspective. Oxford University Press, Oxford; 2014. 10.1093/acprof:oso/9780199675555.001.0001
7. Page MJ, McKenzie JE, Bossuyt PM, et al.: The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. BMJ. 2021, 372:n71. 10.1136/bmj.n71
8. Westin AF: Privacy and freedom. Washington and Lee Law Review. 1968, 25:166-170.
9. Foucault M: Discipline and punish: the birth of the prison (an excerpt). Coronavirus, Psychoanalysis, and Philosophy: Conversations on Pandemics, Politics and Society. Castrillón F, Marchevsky T (ed): Routledge, London; 2021. 4. 10.4324/9781003150497-3
10. Nissenbaum H: Privacy as contextualiIntegrity. Washington Law Review. 2004, 79:119.

11. Garvie C, Bedoya A, Frankle J: The Perpetual Line-Up: Unregulated Police Face Recognition in America. Georgetown Law Center on Privacy and Technology, Washington, DC; 2016.

12. Andrejevic M, Selwyn N: Facial recognition technology in schools: critical questions and concerns. Learning, Media and Technology. 2020, 45:115-128. 10.1080/17439884.2020.1686014

13. Buolamwini J, Gebru T: Gender shades: Intersectional accuracy disparities in commercial gender classification. Proceedings of Machine Learning Research. 2018, 81:1-15.

14. Raji ID, Buolamwini J: Actionable auditing: Investigating the impact of publicly naming biased performance results of commercial AI products. AIES '19: Proceedings of the 2019 AAAI/ACM Conference on AI, Ethics, and Society. 2019, 429-435. 10.1145/3306618.3314244

15. McDuff D, Kaliouby R, Kodra E, Picard R: Measuring voter's candidate preference based on affective responses to election debates. 2013 Humaine Association Conference on Affective Computing and Intelligent Interaction, Geneva, Switzerland. 2013, 369-374. 10.1109/ACII.2013.67

16. Barrett LF, Adolphs R, Marsella S, Martinez AM, Pollak SD: Emotional expressions reconsidered: Challenges to inferring emotion from human facial movements. Psychological Science in the Public Interest. 2019, 20:1-68. 10.1177/1529100619832930

17. Shi W, Cao J, Zhang Q, Li Y, Xu L: Edge computing: Vision and challenges. IEEE Internet of Things Journal. 2016, 3:637-646. 10.1109/jiot.2016.2579198

18. Satyanarayanan M: The emergence of edge computing. Computer. 2017, 50:30-39. 10.1109/mc.2017.9

19. Mascheroni G: Datafied childhoods: contextualising datafication in everyday life. Current Sociology. 2018, 68:798-813. 10.1177/0011392118807534

20. Barocas S, Nissenbaum H: Big data's end run around anonymity and consent. Privacy, Big Data, and the Public Good: Frameworks for Engagement. Lane J, Stodden V, Bender S, Nissenbaum H (ed): Cambridge University Press, Cambridge; 2014. 44-75. 10.1017/CBO9781107590205.004

21. The Office of the High Commissioner for Human Rights: General comment No. 25 (2021) on children's rights in relation to the digital environment. (2021). https://www.ohchr.org/en/documents/general-comments-and-recommendations/general-comment-no-25-2021-childrens-rights-r....

22. Clarke R: Information technology and dataveillance. Communications of the ACM. 1988, 31:498-512. 10.1145/42411.42413

23. Lyon D: Surveillance After Snowden. Polity Press, Cambridge; 2015.

24. Ganascia J-G: The generalized sousveillance society. Social Science Information. 2010, 49:489-507. 10.1177/0539018410371027

25. Prinsloo P, Slade S: Student vulnerability, agency and learning analytics: An exploration. Journal of Learning Analytics. 2016, 3:159-182. 10.18608/jla.2016.31.10

26. Drachsler H, Greller W: Privacy and analytics: it's a DELICATE issue a checklist for trusted learning analytics. Proceedings of the Sixth International Conference on Learning Analytics & Knowledge. ACM, New York, NY; 2016. 89-98. 10.1145/2883851.2883893

27. European Commission: Article 29 Working Party: Guidelines on Consent under Regulation 2016/679. (2018). https://www.google.com/url?sa=t&source=web&rct=j&opi=89978449&url=https://ec.europa.eu/newsroom/article29/redirection....

28. Swedish Data Protection Authority: Supervision pursuant to the General Data Protection Regulation (EU) 2016/679 - facial recognition used to monitor the attendance of students. (2019). https://www.imy.se/globalassets/dokument/beslut/facial-recognition-used-to-monitor-the-attendance-of-students.pdf.

29. Koops BJ: The concept of function creep. Law, Innovation and Technology. 2021, 13:29-56. 10.1080/17579961.2021.1898299

30. Zuboff S: The Age of Surveillance Capitalism: The Fight for a Human Future at the New Frontier of Power. PublicAffairs, New York, NY; 2019.

31. Williamson B, Bayne S, Shay S: The datafication of teaching in higher education: Critical issues and perspectives. Teaching in Higher Education. 2020, 25:351-365. 10.1080/13562517.2020.1748811

32. Solove DJ: Introduction: Privacy self-management and the consent dilemma. Harvard Law Review. 2013, 126:1880-1903.

33. Verizon: 2026 Data Breach Investigations Report. (2026). https://www.verizon.com/business/en-au/resources/reports/dbir/?cmp=knc:ggl:ac:vbg:intl:AUResourcesDBIR&ds_cid=2203743....


**How to cite this article:**
Zulfiqar M, Iqbal U (July 01, 2026) A Review of Privacy, Ethics, and Governance Challenges Associated With Camera-Based AI Surveillance in Primary Schools. Cureus J Comput Sci 3 : es44389-026-00146-3. DOI https://doi.org/10.7759/s44389-026-00146-3

34. Acquisti A, Brandimarte L, Loewenstein G: Privacy and human behavior in the age of information. Science. 2015, 347:509-514. 10.1126/science.aaa1465

35. Crawford K: Atlas of AI: Power, Politics, and the Planetary Costs of Artificial Intelligence. Yale University Press, New Haven, CT; 2021. 10.2307/j.ctv1ghv45t

36. New America: Automating Inequality: How High-Tech Tools Profile, Police, and Punish the Poor. (2018). https://www.newamerica.org/insights/automating-inequality/.

37. Kant I: Groundwork of the Metaphysics of Morals. Cambridge University Press, Cambridge; 2012. 10.1017/CBO9780511919978

38. O'Neill O: Autonomy and Trust in Bioethics. Cambridge University Press, Cambridge; 2009. 10.1017/CBO9780511606250

39. MacIntyre A: After Virtue: A Study in Moral Theory. University of Notre Dame Press, Notre Dame; 2007. 10.2307/jj.21995806

40. Noddings N: Caring: A Feminine Approach to Ethics and Moral Education. University of California Press, Berkeley, CA; 1984.

41. Marx GT: Surveillance studies. International Encyclopedia of the Social and Behavioral Sciences (Second Edition). 2015, 733-741. 10.1016/B978-0-08-097086-8.64025-4

42. European Union (2016) Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 (General Data Protection Regulation). Official Journal of the European Union, L 119. (2018). https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:32016R0679.

43. Mantelero A: Beyond data. Beyond Data: Human Rights, Ethical and Social Impact Assessment in AI. T.M.C. Asser Press, The Hague; 2022. 36:1-43. 10.1007/978-94-6265-531-7_1

44. EU AI Act. (2024). https://www.artificial-intelligence-act.com/Artificial_Intelligence_Act_Links.html.

45. Federal Trade Commission: Children's Online Privacy Protection Act. (1998). https://www.ftc.gov/legal-library/browse/statutes/childrens-online-privacy-protection-act.

46. Reidenberg J, Russell NC, Kovnot J, Norton TB, Cloutier R, Alvarado D: Privacy and Cloud Computing in Public Schools. Fordham Center on Law and Information Policy, New York, NY; 2013.

47. Polonetsky J, Jerome J: Student Data: Trust, Transparency, and the Role of Consent. Future of Privacy Forum, Washington, DC; 2014.

48. Peppet SR: Regulating the Internet of Things: First steps toward managing discrimination, privacy, security, and consent. Texas Law Review. 2014, 93:85-178.

49. Australian Government. Office of the Australian Information Commissioner: Australian Privacy Principles guidelines. (2025). https://www.oaic.gov.au/privacy/australian-privacy-principles/australian-privacy-principles-guidelines.

50. Australian Government. Attorney-General's Department: Privacy Act Review Report. (2023). https://www.ag.gov.au/rights-and-protections/publications/privacy-act-review-report.

51. Liang F, Das V, Kostyuk N, Hussain MM: Constructing a data-driven society: China's social credit system as a state surveillance infrastructure. Policy & Internet. 2018, 10:415-453. 10.1002/poi3.183

52. South China Morning Post: China students 'stressed' by dozens of surveillance cameras in university class told devices are digital media teaching aids. (2024). https://www.scmp.com/news/people-culture/trending-china/article/3254479/china-students-stressed-surveillance-cameras-....

53. Regan PM, Jesse J: Ethical challenges of edtech, big data and personalized learning: twenty-first century student sorting and tracking. Ethics and Information Technology. 2019, 21:167-179. 10.1007/s10676-018-9492-2

54. Cavoukian A: Privacy by Design: The 7 Foundational Principles: Implementation and Mapping of Fair Information Practices. Information & Privacy Commissioner of Ontario, Ontario, Canada; 2009.

55. Diakopoulos N: Accountability in algorithmic decision making. Communications of the ACM. 2016, 59:56-62. 10.1145/2844110

56. The Office of the High Commissioner for Human Rights: Special Rapporteur on the right to privacy. (2018). https://www.ohchr.org/en/special-procedures/sr-privacy.


**How to cite this article:**
Zulfiqar M, Iqbal U (July 01, 2026) A Review of Privacy, Ethics, and Governance Challenges Associated With Camera-Based AI Surveillance in Primary Schools. Cureus J Comput Sci 3 : es44389-026-00146-3. DOI https://doi.org/10.7759/s44389-026-00146-3