# Camera-Based AI in UK Classrooms: The State of Evidence

Research Unit

July 2026

## 1 Introduction

Camera-based AI in education can be used for three different purposes: supervision of in-person examinations, analysis of student behaviour or engagement during teaching, and administrative functions such as identity verification or attendance. This paper examines publicly documented UK activity in these areas and separates operational deployments from academic research and future proposals.

No publicly verified UK school or university was found to be using AI cameras routinely to score classroom engagement or to replace human invigilators in physical examination rooms. The evidence is concentrated elsewhere. Research teams have recorded pupils in a secondary-school classroom to create engagement-labelled video, analysed collaborative behaviour from video of co-located learners, and developed models for multimodal engagement recognition. Schools have also used facial recognition for cashless catering, although these systems were not used for teaching, attendance, or examination supervision.

Human supervision remains the operating standard for school examinations. For the 2025/26 examination cycle, the Joint Council for Qualifications requires invigilators to be physically present and sets minimum staffing ratios of one invigilator for every 30 candidates in written examinations and one for every 20 candidates in on-screen tests, unless an awarding body authorises another arrangement [1]. Ofqual states that AI used as the sole form of remote invigilation is unlikely to comply with its regulations and that human involvement currently provides the best assurance of authenticity and malpractice prevention [2].

In May 2025, Colin Hughes, chief executive of AQA, formerly the Assessment and Qualifications Alliance and the UK's largest provider of academic qualifications taught in schools and colleges, publicly discussed digital invigilation as a possible next step following the introduction of digital examinations. In an official AQA post, the organisation stated that AI-supported invigilation could reduce pressure on schools and colleges, while presenting the idea as a future direction rather than an approved pilot or implementation programme [4].

Ofqual's December 2025 consultation on on-screen assessment separately identified digital tools supporting human invigilation and enhanced malpractice detection as possible security benefits, while also highlighting infrastructure, accessibility, cybersecurity, and implementation risks [3].

Education is devolved across England, Scotland, Wales, and Northern Ireland. The Department for Education guidance cited in this paper applies to England, the North Ayrshire deployment occurred in Scotland, and the school-biometric provisions of the Protection of Freedoms Act 2012 apply to England and Wales. Legal and procurement conditions therefore need to be checked within the relevant jurisdiction.

## 2 Scope and Methodology

The review covers:

- camera-supported supervision inside physical classrooms or examination rooms;
- visual analysis of student presence, attention, engagement, interaction, and conduct during teaching;
- camera-based administrative applications in education facilities, including biometric identity verification and attendance-related functions.

Remote webcam proctoring and virtual-classroom engagement detection are included only where they demonstrate a UK technical capability that may be relevant to a future physical-room application. They are not treated as evidence of classroom deployment.

Sources included peer-reviewed papers, accepted manuscripts, university repositories, official research-group pages, regulator publications, legislation, awarding-body materials, and supplier documentation. For each case, the review checked the physical setting, participants, equipment and software, purpose, reported result, and whether the work was a study, an active research programme, an operational deployment, or only a proposal. Missing technical details are stated as unavailable rather than inferred.

## 3 Research-Led Classroom AI Evidence

### 3.1 University of Strathclyde AI in Education Doctoral Research Group

The University of Strathclyde group is working on several strands of human-centred AI in education. Its public research page lists classroom engagement analysis, emotional assessment for pupils with additional support needs, teacher perspectives on inclusive AI, and speech-recognition robustness in noisy environments [5, 6].

#### 3.1.1 Classroom Engagement Analysis and Feedback

Ashwitha Devasya, Andrew Abel, Yuchen Wang, and Laibing Jia are developing *Classroom Engagement Analysis and Feedback Using Multimodal Emotion Recognition* [7]. The project page describes consent-based camera recording of classroom sessions and analysis of facial expressions, body movement, and other visual signals. The intended output is individual and class-level information that teachers can use when reviewing teaching methods and learning materials [5].

The 2025 peer-reviewed conference poster reports the first model-development stage rather than a completed classroom deployment. It uses the public DAiSEE dataset, which contains 9,068 short videos labelled for boredom, confusion, engagement, and frustration. The researchers tested a custom CNN–LSTM model and a MobileNetV2–LSTM model in Keras. Dlib was used for face detection, and the comparison also examined whether broader body and contextual information improved classification. The custom CNN–LSTM reached 97.7% accuracy on two-second video segments and performed better than the larger MobileNetV2-based model [7].

The published material does not report a completed recording campaign in a Strathclyde classroom. It documents a model trained on an existing dataset and a planned route toward consent-based classroom collection. The project is therefore an active research programme, not an operational monitoring system.

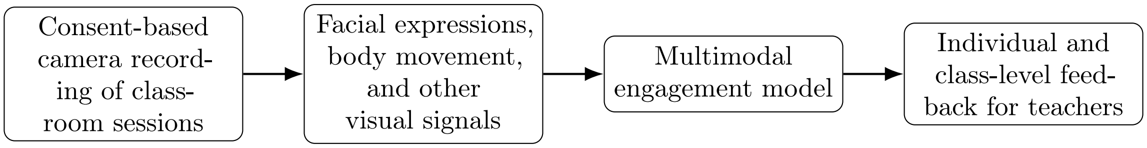


Figure 1: Intended Strathclyde classroom-engagement workflow, summarised from the research-group description [5]. Original Research Unit schematic.

### 3.1.2 Other Strathclyde Projects and Related Publications

A second project, *AI-Based Emotional Assessment for Students with Additional Support Needs*, is supervised by Andrew Abel and Laibing Jia. It proposes multimodal analysis of facial micro-expressions, observable behaviour, physiological signals such as heart rate, and audio or video data to support pupils with additional support needs in Scottish classrooms [5, 6]. The public project description outlines the intended research direction but does not report a completed classroom trial or an evaluated operational system.

The group also lists *AI in Inclusive Education: Teacher's Perspective*, which examines teachers' perceptions of AI, implementation challenges, and professional-development needs in inclusive classrooms in Scotland. The project concerns adoption and educational practice rather than camera-based student monitoring [5].

A separate project, *Analysis of Automatic Speech Recognition Systems Response to the Lombard Effect, and Increasing System Robustness to Noisy Environments*, evaluates speech-recognition models under noisy conditions. The project notes that video may be incorporated when audio alone is insufficient, but its primary focus is robust speech processing rather than classroom-engagement analysis [5].

In addition to the projects listed on the research-group website, members of the group published *Affective State Recognition in Online Learning: A Deep Learning Approach with Data Augmentation* at IEEE SMAP 2025. The study concerns affect recognition in online-learning video and demonstrates the group's technical work in this area. It is a separate publication rather than one of the physical-classroom projects described above, and it does not provide evidence of deployment in a school or lecture room [8].

## 3.2 Peer Data Labelling System

Parsonage, Horton, and Read introduced the Peer Data Labelling System (PDLS) in 2023 as a participatory method for creating labelled classroom-engagement data [9]. The research was conducted at Ribblesdale High School and involved 45 pupils across two studies. Twenty-two pupils aged 11 to 15 participated in the data-collection study, while 23 pupils aged 11 to 12 took part in the second study.

During the first study, pupils worked in pairs and exchanged roles halfway through the activity. One pupil completed a computer-based cryptography task while a camera recorded the learner. The second pupil observed the learner without seeing the task screen and used an online form to mark each change between two categories: *engaged* and *disengaged*. The labels were therefore recorded alongside the classroom video rather than being added retrospectively by external annotators [9].

The study produced 22 recordings, of which 17 were usable. These provided 2 hours, 27 minutes, and 32 seconds of labelled video and 221,300 labelled image frames. Most of the material was classified as engaged. The second study examined pupils' confidence in peer judgements and their acceptance of a possible automated engagement-monitoring system. Participants expressed greater confidence in human peer judgements than in the proposed automated alternative [9].

The researchers published a follow-up study in 2024, *Evaluating the Effectiveness of the Peer Data Labelling System (PDLS)*, which assessed the reliability of the labels created through the original method [10]. Algorithmic validation approaches did not produce sufficiently consistent classifications. Human reviewers reached agreement with the pupil observers for most of the material labelled as engaged, while identifying disengagement remained more difficult and required further refinement.

The PDLS data were also used in Graham Parsonage's 2024 doctoral thesis, *Behind the Wizard's Curtain: Designing and Developing Intelligent Systems for use in Educational Contexts* [11]. Chapter 8 describes an engagement classifier based on a Long Short-Term Memory model. The model processed sequences from the PDLS material and produced a binary classification of the recorded pupil as engaged or disengaged, which was written onto the video output. The thesis presents the implementation as a feasibility demonstration rather than a production-ready classroom system and acknowledges the need for further development and validation.

The published evidence therefore extends beyond dataset creation. It includes an evaluation of the peer-generated labels and an experimental machine-learning classifier trained within the same doctoral research programme. No independent study was identified that reused the PDLS dataset to validate or deploy a camera-based engagement-recognition system in another educational setting.

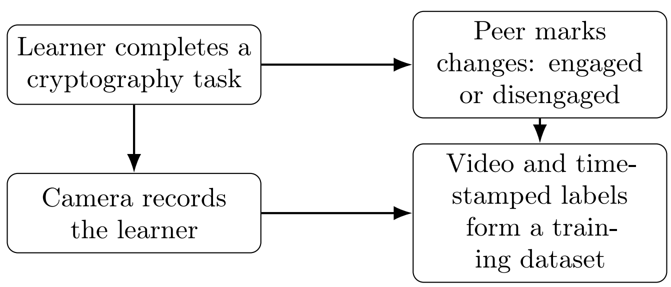


Figure 2: PDLS experiment structure, based on Parsonage et al. (2023) [9]. Original Research Unit schematic.

### 3.3 Video Analytics for Collaborative Problem Solving

Cukurova, Zhou, Spikol, and Landolfi (2020) investigated whether video analytics could be used to model collaborative problem-solving behaviour in a co-located physical learning environment [12]. The study involved 18 engineering students—17 men and one woman—with an average age of 20. The students were selected from a class of 30 and were organised into six groups of three.

Each group completed three sequential open-ended design activities: prototyping an interactive toy, constructing a colour-sorting machine, and building an autonomous car. This design could have produced 18 group-task recordings. After recordings with insufficient usable data were excluded, the final dataset contained 15 videos and approximately 500 minutes of recorded collaborative activity [12]. The paper does not specify the number, model, or exact placement of the cameras. It reports that recording quality was affected in some sessions because cameras were too close or too far from participants, did not capture all interactions, or were moved during the activities.

Two open-source computer-vision tools were used to process the videos. FaceNet, a facial-recognition system that converts detected faces into numerical representations, was used to associate each visible participant with an anonymised student identifier. OpenPose, a deep-learning framework for detecting the position of multiple human bodies in images and video, identified 18 body key points for each visible participant, including points corresponding to the face, shoulders, elbows, wrists, and other parts of the body [12].

The researchers used these detected points to calculate eight features. Four measured how frequently a participant's face, right hand, left hand, or both hands appeared in the frame. Four measured spatial relationships: the distance between an individual's hands, between participants' bodies, between their faces, and between group members' hands. These features were extracted automatically from the videos using Python [12].

Two researchers then manually coded the recorded activity in 30-second intervals. They classified four observable behaviours: making, watching, speaking, and listening. Agreement between the two coders was 98%, with an ordinal Krippendorff's alpha of 0.912. The researchers also rated each group as demonstrating high, medium, or low collaborative problem-solving competence, reaching an inter-rater alpha of 0.876 [12].

Decision-tree models were developed in RapidMiner to test whether the automatically extracted video features could predict the manually coded behaviours and the groups' collaborative problem-solving competence. The mean classification accuracies were 63.75% for making, 64.58% for watching, 63.75% for speaking, and 60.00% for listening. The model predicting overall collaborative problem-solving competence achieved an accuracy of 62.50%. It performed best when identifying groups rated as having low competence, reaching 75.00% precision and 100.00% recall for that category [12].

Distances between participants' faces and hands were among the most informative features for predicting behaviour. By contrast, the frequency with which faces appeared in the frame contributed less to predicting overall collaborative competence. The authors selected decision trees because their rules could be inspected by teachers and learners, unlike less transparent deep-learning models. However, the reported accuracy remained limited, and missing or incomplete camera data reduced the reliability of the predictions [12].

The study demonstrates the use of camera recordings and interpretable machine-learning models in a physical collaborative-learning environment. It remained a controlled research experiment and did not report adoption as an operational classroom-monitoring product.

## 4 Adjacent UK Research in Remote and Virtual Learning

The projects in this section also use cameras, facial movement, gaze, or body-pose estimation to analyse learning behaviour. Unlike the studies above, they were developed for remote, online, or virtual-learning environments rather than through cameras recording groups of learners in a shared physical classroom. They are included as evidence of UK technical capability, not as evidence of in-class deployment.

### 4.1 EngageSense

Irfan, Patel, and Hassan of London Metropolitan University published *EngageSense* in 2025 for real-time engagement detection in virtual classrooms [13]. The system used laptop webcams rather than fixed cameras in a physical classroom.

Dlib's Histogram of Oriented Gradients detector and a linear Support Vector Machine were used for face detection. A convolutional neural network classified gaze as left, right, or centre, while OpenPose MobileNetV1 estimated 19 body key points. Gaze and pose were combined over one-second video intervals to classify the learner as fully engaged, partially engaged, or not engaged [13].

The custom eye dataset contained 4,453 augmented images from 10 participants. The paper reports 99.5% accuracy for gaze-direction classification and approximately 90% accuracy for the combined engagement classification. The authors identify the small participant group as a limitation and propose a larger and more diverse dataset in future work.

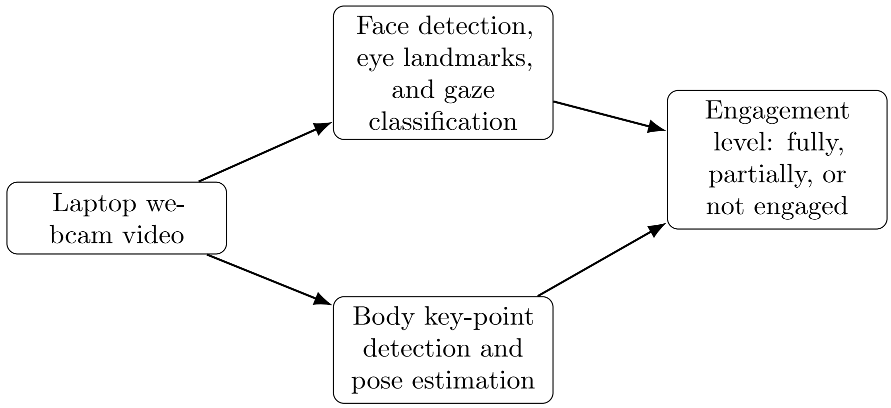


Figure 3: EngageSense processing pipeline, summarised from Irfan, Patel, and Hassan (2025) [13]. Original Research Unit schematic.

EngageSense provides a functioning UK-developed pipeline for gaze- and pose-based classification, but its evaluation was small and remote. It should not be described as a camera system installed in a physical classroom.

### 4.2 PUZZLED: Facial-Movement Analysis of Student Struggle

Linson, Xu, English, and Fisher (2022) developed the PUZZLED prototype to identify group-level student struggle during asynchronous video-lecture viewing [15]. The study involved 10 university students. Participants watched two recorded computer-vision lectures on their own devices while laptop webcams captured their faces. They also clicked an on-screen control when they felt they were struggling and later completed retrospective self-reports. Independent observers reviewed the recordings and annotated possible struggle-related states.

The prototype integrated pre-trained models from Intel's Open Model Zoo for face detection, facial-landmark localisation, gaze estimation, head-pose estimation, and eye-state detection. The analysis focused particularly on changes in vertical gaze direction and mapped computer-vision outputs, self-reports, and observer annotations onto a common lecture timeline [15].

The pilot found that upward gaze-aversion events often aligned with difficult lecture segments and with manual struggle annotations, although the authors emphasised the small sample, browser and interface problems during data collection, and individual differences in facial behaviour. PUZZLED was designed as teacher-facing feedback: the timeline was intended to show where a cohort encountered difficulty so instructors could revise content or provide support. It was not a physical-classroom deployment and did not automate sanctions or assessment decisions.

## 5 Camera-Based Administrative Deployments in Schools

### 5.1 North Ayrshire Council School Canteens

North Ayrshire Council introduced facial recognition in nine Scottish school canteens on 6 September 2021, covering 2,569 pupils. The ICO records identify CRB Cunninghams as the supplier and data processor [16].

A catering employee manually activated the till camera. It captured one still image of the pupil and attempted to match that image with a stored biometric template. A successful match opened the pupil's online account, after which the employee approved or declined the transaction. Where several possible matches appeared, the employee selected the correct account manually [16].

The system was transaction-triggered rather than continuous live facial recognition. CRB Cunninghams' product description also states that point-of-sale staff activate the camera and that the system stores a mathematical facial template in a closed-loop database [17]. The public documents do not disclose the camera model, matching algorithm, independent accuracy test, or false-match rate.

The council paused the system on 22 October 2021 and confirmed that facial templates and backups had been deleted by 11 November 2021. The ICO identified likely infringements concerning lawful basis, transparency, retention, and the Data Protection Impact Assessment [16].

### 5.2 Chelmer Valley High School

Chelmer Valley High School in Essex introduced CRB Cunninghams facial recognition for cashless catering in March 2023. The public record does not name the camera model or provide independent performance testing [18].

The school had used fingerprint recognition for catering since 2016. It introduced facial recognition without completing the required Data Protection Impact Assessment before processing began and treated pupils as enrolled unless parents opted out. The ICO found that this was not valid affirmative consent and noted that many pupils were old enough to exercise their own data rights [18].

## 6 Regulatory Position and Practical Implications

The Department for Education updated its guidance on children's biometric information in June 2026. It applies to schools and colleges in England and covers automated biometric recognition, including fingerprint and facial recognition [19].

The guidance says facial recognition will often be inappropriate for functions such as school-lunch payments where less intrusive alternatives are available. It also states that live facial recognition is not appropriate in schools and colleges because it would be difficult to justify as fair, necessary, proportionate, and lawful under the UK GDPR [19]. In England and Wales, the Protection of Freedoms Act 2012 requires parental notification and consent arrangements, requires processing to stop if the pupil objects, and requires a reasonable non-biometric alternative [20].

A consent form alone does not make continuous classroom recording lawful. An institution still has to establish a clear purpose, lawful basis, necessity, proportionality, transparency, data minimisation, security, retention limits, and suitable rights for pupils. High-risk video or biometric processing requires a Data Protection Impact Assessment before collection begins [19, 21].

Consent-based academic recording can be approved under research ethics and data-protection controls, as in the PDLS study. Continuous identification, attention inference, or automated decisions would require a stronger justification and closer safeguards. An examination-room product would also need human oversight under the current Ofqual position and independent testing of false positives, accessibility, bias, cyber security, and technical failure.

## 7 Conclusion

The UK evidence does not support a claim of broad deployment of AI cameras for classroom engagement tracking or in-person examination invigilation. The verified school deployments concern facial recognition for canteen payments, and both attracted regulatory intervention.

The academic evidence is more substantial. Strathclyde has an active programme directed at multimodal classroom engagement, with initial model results published from the DAiSEE dataset and plans for consent-based classroom recording. PDLS produced engagement-labelled video in a real secondary-school classroom. UCL-led work used OpenPose and interpretable models to analyse co-located collaborative learning. EngageSense produced a real-time gaze- and pose-based pipeline for virtual classrooms, while PUZZLED mapped webcam-derived facial movement and self-reported struggle onto a lecture timeline. Neither remote study establishes physical-classroom deployment.

Camera-supported invigilation remains a future proposal. Human invigilators are still required under current operating rules. AQA has publicly discussed a human-supervised camera model, while Ofqual permits exploration of digital tools but rejects AI-only remote invigilation under the present evidence base.

The viable direction for a UK pilot is therefore narrow: human-supervised decision support, a clearly defined educational or assessment purpose, prior ethical and data-protection approval, limited data collection and retention, transparent information for participants, and independent evaluation in realistic rooms.

## References


[1] Joint Council for Qualifications. (2025). *Instructions for Conducting Examinations, 1 September 2025 to 31 August 2026*. `https://www.jcq.org.uk/wp-content/uploads/sites/2/2025/10/Print-version-JCQ-Instructions-for-conducting-examinations-2025_6_FINAL.pdf`

[2] Ofqual. (2024). *Ofqual's Approach to Regulating the Use of Artificial Intelligence in the Qualifications Sector*. `https://www.gov.uk/government/publications/ofquals-approach-to-regulating-the-use-of-artificial-intelligence-in-the-qualifications-sector/ofquals-approach-to-regulating-the-use-of-artificial-intelligence-in-the-qualifications-sector`

[3] Ofqual. (2025). *Regulating on-screen assessment: Consultation*. Published 11 December 2025. `https://www.gov.uk/government/consultations/regulating-on-screen-assessment/regulating-on-screen-assessment`

[4] AQA. (2025). *How AI Can Revolutionize Invigilation for Exams*. Official LinkedIn post. `https://www.linkedin.com/posts/aqa_use-ai-cameras-to-invigilate-gcses-says-activity-7329031039761149952-yWw-`

[5] University of Strathclyde, AI in Education Doctoral Research Group. *Research Study*. `https://haiti.cis.strath.ac.uk/research-study/`

[6] University of Strathclyde, AI in Education Doctoral Research Group. *About Us and Research Team*. `https://haiti.cis.strath.ac.uk/sample-page/`

[7] Devasya, A., Abel, A., Wang, Y., and Jia, L. (2025). *Classroom Engagement Analysis and Feedback Using MER*. BERA Annual Conference 2025, peer-reviewed poster. `https://pureportal.strath.ac.uk/en/publications/classroom-engagement-analysis-and-feedback-using-mer/`

[8] Devasya, A. and Abel, A. (2025). *Affective State Recognition in Online Learning: A Deep Learning Approach with Data Augmentation*. IEEE SMAP 2025. `https://haiti.cis.strath.ac.uk/elementor-592/`

[9] Parsonage, G., Horton, M., and Read, J. C. (2023). *The Peer Data Labelling System (PDLS): A Participatory Approach to Classifying Engagement in the Classroom.* INTERACT 2023. https://knowledge.lancashire.ac.uk/id/eprint/48659/1/pdls.pdf

[10] Parsonage, G., Horton, M., and Read, J. C. (2024). *Evaluating the Effectiveness of the Peer Data Labelling System (PDLS).* In *Artificial Intelligence in HCI*, Lecture Notes in Computer Science, vol. 14734, pp. 67–83. Springer. https://doi.org/10.1007/978-3-031-60606-9_5

[11] Parsonage, G. (2024). *Behind the Wizard's Curtain: Designing and Developing Intelligent Systems for use in Educational Contexts.* Doctoral thesis, University of Central Lancashire. https://doi.org/10.17030/uclan.thesis.00052603

[12] Cukurova, M., Zhou, Q., Spikol, D., and Landolfi, L. (2020). *Modelling Collaborative Problem-Solving Competence with Transparent Learning Analytics: Is Video Data Enough?* In *Proceedings of the Tenth International Conference on Learning Analytics & Knowledge*, pp. 270–275. Association for Computing Machinery. https://doi.org/10.1145/3375462.3375484

[13] Irfan, M., Patel, P., and Hassan, B. (2025). *EngageSense: A Hybrid Approach for Real Time Engagement Detection for Virtual Classrooms.* IEEE EDUCON 2025. https://doi.org/10.1109/EDUCON62633.2025.11016600

[14] London Metropolitan University Repository. (2025). *EngageSense: Accepted Manuscript, Creative Commons Attribution 4.0.* https://repository.londonmet.ac.uk/10102/

[15] Linson, A., Xu, Y., English, A. R., and Fisher, R. B. (2022). *Identifying Student Struggle by Analyzing Facial Movement During Asynchronous Video Lecture Viewing: Towards an Automated Tool to Support Instructors.* Artificial Intelligence in Education 2022. https://oro.open.ac.uk/84409/1/puzzled_AIED22_preprint.pdf

[16] Information Commissioner's Office. (2022). *North Ayrshire Council's Use of Facial Recognition Technology in Its Schools.* https://ico.org.uk/media2/migrated/4023847/ico-letter-to-nac-appendix.pdf

[17] CRB Cunninghams. *Facial Recognition for Cashless Catering.* https://marketing.crbcunninghams.co.uk/acton/fs/blocks/showLandingPage/a/35817/p/p-00cd/t/page/fm/0

[18] Information Commissioner's Office. (2024). *Reprimand Issued to Chelmer Valley High School.* https://ico.org.uk/media2/migrated/4030350/20240702-cvhs-reprimand.pdf

[19] Department for Education. (2026). *Protection of Biometric Data of Children in Schools and Colleges.* Updated 17 June 2026. https://assets.publishing.service.gov.uk/media/6a313437141f0690ad5fa3fb/Protection_of_biometric_data_of_children_in_schools_and_colleges.pdf

[20] UK Parliament. (2012). *Protection of Freedoms Act 2012, Part 1, Chapter 2.* https://www.legislation.gov.uk/ukpga/2012/9/part/1/chapter/2

[21] Information Commissioner's Office. *Video Surveillance Guidance for Organisations.* https://ico.org.uk/for-organisations/uk-gdpr-guidance-and-resources/cctv-and-video-surveillance/guidance-on-video-surveillance-including-cctv/