Bourne Alliance Multi Academy Trust

Policy document for: **Deep fakes**

Updated: Spring 2026

This policy forms part of the Trust Safeguarding and Child Protection policy and reflects the latest KCSIE guidance. Keeping children safe in education 2026 (comes into force from 1 September 2026) See page 6 for full details

KCSIE 2026 explicitly identifies **AI-generated and digitally manipulated imagery** as a safeguarding risk linked to sexual violence, coercion, child-on-child abuse, harmful sexual behaviour (HSB), CSE, CCE, domestic abuse, radicalisation, fraud and image-based sexual abuse (IBSA).
Deepfakes are now a **day-to-day safeguarding responsibility**.

What is a Deepfake?

Deepfakes are digitally manipulated or AI-generated images, audio or video that depict a person saying or doing something they never did. They may be created using:

- photo or video editing software
- deepfake apps and generators
- AI text-to-image tools
- face-swap filters
- synthetic avatars

Deepfake technology is **highly accessible** to children online. Deepfakes can:

- impersonate real people
- create pseudo-images that look like genuine photographs
- depict sexual acts that never occurred
- fabricate criminal behaviour
- create false evidence
- mimic voices, gestures and facial expressions

Deepfake avatars and influencers (e.g., Lil Miquela, Shudu) are widely followed by children who may not realise they are not real.

Since 2024, it has been a **criminal offence** to create a deepfake of an adult to cause alarm, humiliation or distress — even if there is no intent to share it. Child images are already protected under existing legislation.

Deepfakes as Image-Based Sexual Abuse (IBSA)

Under KCSIE 2026, deepfakes are classified as **image-based sexual abuse**, even when:

- the child never posed
- the image is entirely fake
- the child is fictional or synthetic
- the creator “didn’t intend harm”

Deepfakes may include:

- sexualised synthetic images
- nudification (AI removal of clothing)
- sexualised deepfake videos
- sexualised images of staff
- sexualised images of fictional minors

Creating or sharing sexualised deepfakes of anyone under 18 is a **criminal offence** under the Online Safety Act and Sexual Offences Act.

Deepfakes as Child-on-Child Abuse and Harmful Sexual Behaviour (HSB)

Creating, sharing or threatening to share deepfakes of peers is:

- child-on-child abuse
- harmful sexual behaviour
- sexual harassment
- coercion
- bullying
- online sexual abuse

Deepfakes must be treated with the **same level of safeguarding urgency as real sexual images**.

Risks for Children as Viewers of Deepfakes

Children may experience harm when viewing deepfakes, including:

Difficulty distinguishing real from fake

Children (and adults) struggle to recognise deepfakes. This increases vulnerability to:

- misinformation
- political manipulation
- extremist narratives
- commercial exploitation

Radicalisation (Prevent Duty)

Deepfakes can:

- normalise extremist attitudes
- spread conspiracy theories
- manipulate political views
- simulate extremist grooming

These are **Prevent concerns** requiring DSL involvement.

Exposure to pornography

Children are accessing pornography earlier than ever. Deepfake pornography:

- is non-consensual
- is misogynistic
- distorts understanding of sex and relationships
- normalises harmful behaviour

Psychological harm

Deepfakes can cause:

- trauma
- distress
- fear
- confusion
- desensitisation to violence or manipulation

SEND vulnerability

Children with SEND — particularly those in Aspire — are:

- less able to detect deepfakes
- more trusting of digital content
- more socially isolated
- more vulnerable to manipulation

Risks for Children as Victims of Deepfakes

Children may experience severe harm when deepfakes are created of them:

Reputational damage and stigma

Deepfakes may depict children in:

- sexual acts
- criminal behaviour
- demeaning or violent scenarios

This can lead to:

- bullying
- isolation
- social stigma
- sextortion
- long-term reputational harm

Emotional harm and trauma

Removal of deepfakes is almost impossible. Children may experience:

- overwhelming distress
- shame
- fear
- long-term trauma

Staff must avoid victim-blaming and empower children to seek help.

Sextortion and blackmail

Deepfakes can be used to:

- threaten children
- extort money
- extort sexual content
- coerce silence

The National Crime Agency has issued warnings to schools about rising sextortion cases.

Grooming

Deepfake video calls or avatars can:

- impersonate children
- impersonate trusted adults
- be used to groom children

Normalisation of harmful behaviour

Deepfakes may:

- depict violence
- depict non-consensual acts
- reinforce misogyny
- perpetuate unrealistic body ideals
- encourage harmful stereotypes

Deepfakes Targeting Staff

Deepfakes may target staff through:

- sexualised synthetic images
- nudification
- impersonation
- false allegations
- fabricated audio or video

These incidents must be treated as **urgent safeguarding concerns**.

Whole-School Education

Digital literacy underpins teaching across the Trust. Children are taught:

- how deepfakes are created
- why deepfakes are harmful
- that deepfakes are not victimless
- that deepfakes disproportionately target women and girls
- that deepfakes are a form of misogyny
- how to critically evaluate digital content
- how to seek help safely

A new AI curriculum being introduced in the mainstream schools (2026/27) teaches:

- legal and ethical issues
- risks of creating deepfakes
- critical thinking
- media literacy
- consent
- online safety

"Just claiming an image is not real" is **not a valid defence**.
SEND-specific teaching ensures children understand their increased vulnerability.

Staff Guidance and Responsibilities

Staff must:

- teach media literacy using national frameworks
- deliver RSE content on consent, pornography and harmful content
- teach online safety including deepfakes
- use UKCIS Education for a Connected World
- link deepfakes to nudes/semi-nudes guidance
- understand KCSIE 2026 expectations

All staff must know:

- deepfake pornography is illegal
- deepfakes are a form of abuse
- how to report concerns on CPOMS
- to speak directly to a DSL immediately
- that deepfake threats must be treated as urgently as physical abuse

DSL Responsibilities

DSLs will:

- complete an immediate risk assessment
- consider links to other vulnerabilities (SEND, exploitation, mental health)
- determine whether the incident is IBSA, HSB, CSE, CCE, domestic abuse or Prevent
- liaise with Police, Family Help and Prevent
- ensure safe information sharing
- monitor patterns through caseload reviews
- ensure filtering systems detect deepfake tools
- ensure staff training is current
- ensure records transfer when a child moves school

Intervention Protocols

Schools will provide access to:

- **Childline Report Remove**
- **Internet Watch Foundation (IWF)**
- **Take It Down** (anonymous removal tool)

Tools will be:

- accessible on school websites
- displayed on posters
- unblocked by filtering systems

Reporting to DSLs

Staff must report immediately if:

- any deepfake or pseudo-image is created, shared or viewed
- any threat, joke or reference to deepfakes is made
- any pupil expresses fear or concern about deepfakes
- any staff member is targeted
- any pupil is experimenting with harmful deepfake tools

All incidents must be recorded on CPOMS using the child's exact words.