Online Safety Policy

1 Policy statement

Coventry & Warwickshire Chamber Training (CWCT) is committed to providing a safe, inclusive and enabling digital environment. Technology is integral to learning, assessment, communication and business activity. Its benefits will be promoted while foreseeable risks are identified, reduced and responded to promptly.

Online safety is a safeguarding responsibility, not solely an IT function. The organisation will take a whole-provider approach that combines leadership, education, professional conduct, appropriate technical controls, proportionate monitoring, clear reporting routes and a child-centred response. For this policy, "child" means any person under 18. The same protective principles will be applied, as appropriate, to adult learners and adults at risk.

Core expectation

Nobody should dismiss an online concern because it occurred away from the premises, on a personal device, through a private account, or outside normal learning or working hours. If it may affect safety, welfare, learning, employment, professional boundaries or the organisation, it must be considered and reported.

2 Purpose and objectives

- protect learners and staff from online harm, abuse, exploitation, bullying, harassment, fraud and unlawful activity;
- establish clear expectations for safe, ethical and lawful use of devices, systems, social media and AI;
- ensure concerns are recognised, recorded, reported, escalated and reviewed consistently;
- define governance, safeguarding, curriculum, IT and individual responsibilities;
- maintain appropriate filtering and monitoring without unreasonably restricting learning;
- support digital literacy, critical thinking and informed decision-making; and
- provide assurance that online safety arrangements are reviewed and improved.

3 Scope

This policy applies to conduct and activity involving:

- CWCT networks, Wi-Fi, accounts, email, learning platforms, cloud services, hardware and software;
- CWCT-owned, leased or managed devices;
- personal devices used for learning or work, including under bring-your-own-device arrangements;
- remote, blended, off-the-job and workplace learning;
- social media, messaging, collaboration, video-conferencing and file-sharing services;
- generative AI, embedded AI features, chatbots, image/audio/video generators and automated decision-support tools; and
- online conduct outside CWCT systems where there is a safeguarding, professional, legal, reputational or learning-related impact, including social media.

4 Legal and statutory framework

This policy should be read and applied alongside current legislation, statutory guidance and organisational procedures, including:

- Keeping Children Safe in Education 2026 (KCSIE), in force from 1 September 2026;
- Working Together to Safeguard Children;
- Department for Education filtering and monitoring standards and generative AI product safety standards;
- UK GDPR and the Data Protection Act 2018;
- Computer Misuse Act 1990;
- Communications Act 2003 and Malicious Communications Act 1988;
- Online Safety Act 2023;
- Counter-Terrorism and Security Act 2015, including the Prevent duty;
- Equality Act 2010;
- copyright, intellectual property, defamation, harassment and confidentiality law; and
- CWCT safeguarding, Prevent, data protection, staff conduct, learner conduct, complaints, whistleblowing, disciplinary and business continuity arrangements.

Where statutory guidance or legislation changes, the latest requirements take precedence and this policy will be reviewed.

5 Understanding online safety risk

CWCT uses four areas of online risk to support assessment, education and response:

Area	Examples
Content	Illegal, harmful, discriminatory, extremist, misleading or age-inappropriate material; pornography; self-harm content; misinformation; AI-generated content; manipulated images, audio or video.
Contact	Grooming; coercion; harassment; unsolicited contact; adults posing as peers; extremist recruitment; contact by criminals; phishing and social engineering.
Conduct	Cyberbullying; abusive posts; sharing private information; impersonation; non-consensual image sharing; creating or distributing deepfakes; academic dishonesty; unsafe AI use.
Commerce	Scams; fraud; gambling; financial exploitation; in-app purchases; identity theft; fake job or training offers; cryptocurrency and money-mule recruitment.

Risks commonly overlap. An apparently isolated concern may indicate wider safeguarding need, exploitation, coercion, abuse, criminality or risk to others. Staff must avoid viewing online and offline harm as separate issues.

6 Roles and responsibilities

6.1 Board / governing body

- approve this policy and maintain strategic oversight of online safety;
- ensure safeguarding leaders have appropriate authority, time, training and resources;
- seek assurance that filtering and monitoring are appropriate, effective and reviewed at least annually, with records retained;
- understand the provider's risk profile, including AI-enabled risks and online exploitation;
- receive suitable trend, incident and assurance information without unnecessary personal data; and
- ensure actions arising from audit, incidents or review are completed.

6.2 Executive Director and senior leaders

- implement this policy across curriculum, apprenticeship, commercial and corporate activity;
- ensure roles are allocated between safeguarding, IT, data protection, curriculum and HR;
- approve organisational AI tools and risk controls;
- ensure staff and learners receive induction and updates; and
- support proportionate action where expectations are breached.

6.3 Designated Safeguarding Lead (DSL)

- take lead responsibility for online safety and online safeguarding concerns;
- ensure reporting routes are known, accessible and acted upon;
- work with IT support to review filtering and monitoring alerts and effectiveness;
- assess risk, record decisions and make referrals to children's social care, police, Prevent or other agencies where required;
- ensure victims receive appropriate support and alleged perpetrators are managed fairly and safely;
- maintain awareness of generative AI, deepfake, exploitation and emerging online harms; and
- provide safeguarding analysis and assurance to senior leaders and the board.

6.4 IT lead / managed service provider

- maintain secure systems, access controls, patches, backups, filtering and monitoring arrangements;
- ensure controls cover relevant devices, users, networks, cloud services and AI-generated formats as far as technically practicable;
- test, document and report on the operation of controls;
- provide alerts and information promptly to authorised safeguarding staff;
- preserve evidence and logs in line with retention, privacy and incident-management requirements; and
- avoid making safeguarding judgments in isolation.

6.5 Curriculum, quality and line managers

- embed age- and role-appropriate online safety, digital literacy and AI literacy;
- model safe practice and challenge unsafe or inappropriate use;
- plan safe use of technology and approved AI tools;
- make reasonable adjustments and avoid assuming digital competence; and
- report concerns and support implementation within their teams.

6.6 All staff, learners and users

- follow this policy and associated acceptable-use requirements;
- keep credentials secure and use multi-factor authentication where provided;
- use only approved systems and tools for organisational data;
- report concerns, mistakes, suspected compromise and harmful content promptly;
- preserve evidence without forwarding or repeatedly viewing harmful material; and
- co-operate with proportionate investigation and safeguarding action.

7 Acceptable use of IT, networks and devices

7.1 Permitted and expected use

- educational research, online learning, assessment, projects, presentations and collaborative work;
- work-related communication through approved, open and traceable channels;
- secure access to authorised files, systems and services;
- social media for an approved educational or business purpose under appropriate supervision;
- personal devices for learning where permitted, secured and used in accordance with this policy; and
- reasonable personal use by staff only where permitted by organisational rules and without interference with duties, security or reputation.

7.2 Prohibited use

- accessing, creating, requesting, storing or sharing illegal, abusive, pornographic, extremist, discriminatory or otherwise harmful material;
- bullying, threatening, harassing, stalking, doxxing, humiliating or impersonating another person;
- creating or sharing sexualised, intimate or degrading manipulated content, including AI-generated deepfakes;
- attempting unauthorised access, bypassing security/filtering, installing unapproved software or altering system settings;
- sending anonymous or deceptive communications, spam, malware or phishing content;
- copying, plagiarising or presenting another person's or an AI system's work as one's own;
- using personal or untraceable channels to communicate with learners where an approved channel is required;
- sharing passwords, confidential information, personal data or images without authority;
- recording, photographing or live-streaming others without a lawful and approved purpose; and
- using technology for fraud, gambling, criminal activity, radicalisation, exploitation or any activity that places another person at risk.

8 Generative artificial intelligence

8.1 Key risks

Generative AI can support learning and productivity, but outputs can be inaccurate, biased, fabricated, unsafe or unsuitable. AI services may retain prompts and uploaded content, expose personal or confidential data, reproduce protected material, enable impersonation or deception, and generate convincing harmful content at scale.

- fabricated information, references, evidence or citations ("hallucinations");
- bias, stereotyping or discriminatory outcomes;
- loss of confidentiality, privacy or intellectual property through prompts and uploads;
- over-reliance that weakens learning, professional judgment or independent thinking;

- academic or assessment malpractice and uncertain authorship;
- unsafe advice or inappropriate content;
- prompt injection, malicious links, data exfiltration or embedded harmful instructions;
- creation of deepfakes, scams, grooming material, misinformation or extremist propaganda; and
- unequal access, accessibility barriers and digital exclusion.

8.2 Conditions for appropriate use

- Only tools approved by CWCT may be used for CWCT work, learning or data. Embedded AI features are included in this requirement
- Users must follow the purpose, age, account, privacy and licensing restrictions of the service
- Personal, special-category, safeguarding, commercially sensitive, assessment-secure or confidential information must not be entered into a public or unapproved AI tool
- AI must support, not replace, professional judgment, teaching, assessment, safeguarding decisions or direct engagement with learners
- The user remains accountable for checking accuracy, bias, relevance, tone, accessibility, copyright and safety before using an output
- Material use of AI must be acknowledged where required by an assessment, awarding organisation, customer, publication or CWCT instruction
- AI-generated evidence must not be submitted as proof of competence, experience or knowledge unless the assessment rules expressly permit the defined use
- Users must not use AI to impersonate, deceive, harass, exploit, discriminate, evade controls or create harmful or illegal content
- AI outputs concerning safeguarding must be treated as unverified. Concerns must be referred to the DSL and handled through approved procedures

8.3 Expectations for staff

- complete any required AI and data protection training;
- carry out an appropriate risk assessment before introducing an AI-enabled product or materially new use;
- use human review for consequential outputs and never delegate safeguarding decisions to AI;
- design learning activities that make permitted and prohibited AI use explicit;
- consider learners' age, SEND, vulnerability, language, digital skills and access;
- retain suitable evidence of prompts/outputs where required for quality assurance, assessment integrity or incident review; and
- report unexpected harmful outputs, data exposure, bias or tool failure.

8.4 Expectations for learners

- use AI only when expressly permitted and within the boundaries set by staff or assessment rules;
- state how AI was used when asked;
- check outputs against reliable sources and do not invent references;
- protect their own and others' information;
- submit work that demonstrates their own knowledge, skills and behaviours; and
- seek help if an AI interaction produces harmful, disturbing, coercive or suspicious content.

AI safeguarding principle

An AI tool is not a trusted adult, safeguarding professional, confidential reporting route or authoritative source. Learners and staff must use CWCT's reporting routes for concerns and emergencies.

9 Deepfakes, synthetic media and manipulated content

A deepfake is synthetic or manipulated image, audio or video content that convincingly portrays a person saying or doing something they did not say or do. This includes fully generated and altered content, face or voice cloning, and AI-created intimate images.

- Creating, requesting, possessing or sharing deepfakes to bully, humiliate, sexualise, defraud, threaten, impersonate or damage another person is prohibited and may be a safeguarding or criminal matter
- Staff must treat suspected AI-generated intimate or indecent content involving a child as a safeguarding concern, whether or not a real image was used to create it
- Recipients must not forward, download, edit, screenshot unnecessarily or repeatedly view suspected harmful content. They should preserve the device or link where safe, record how the concern came to attention and report immediately to the DSL
- CWCT will not require a learner to prove that content is fake before support is offered. The immediate focus will be safety, impact, evidence preservation and removal/reporting options
- Where authenticity matters, staff will use corroboration, source checking and specialist advice rather than relying on appearance alone
- Public statements, contact with platforms, parents/carers, employers, police or external agencies will be coordinated by authorised staff

10 Social media, messaging and professional boundaries

- Staff must maintain appropriate professional boundaries and must not connect with or privately message learners through personal social media accounts
- Communication with learners and employers must use approved, open and traceable channels. Personal messaging or disappearing-message functions must not be used unless formally authorised with safeguards
- Staff must keep personal and professional accounts appropriately separated, apply privacy settings and make clear when personal views are their own
- No user may post confidential information, learner data, assessment material, internal incidents, images or recordings without authority and a lawful basis
- Posts must not be discriminatory, defamatory, threatening, bullying, harassing, misleading or likely to bring CWCT into disrepute
- Users should be alert to fake profiles, location sharing, oversharing, account takeover, malicious links, coordinated harassment and social engineering
- Use of social media outside working or study hours may still be addressed where it creates a safeguarding concern, breaches professional standards or has a material connection with CWCT

11 Online exploitation and harmful contact

Online exploitation may involve grooming, coercion, manipulation, deception or abuse for sexual, criminal, financial, extremist or other harmful purposes. It can begin online and move offline or technology may be used to control an existing relationship.

11.1 Indicators may include

- secretive or rapidly changing online relationships;
- unexplained gifts, money, devices, travel, debt or multiple accounts;
- pressure to share images, move to encrypted platforms, recruit others, receive or transfer money, or undertake “favours”;
- threats to expose images or information, blackmail or sextortion;
- contact from fake recruiters, influencers, romantic interests, gangs or adults presenting as peers;
- sudden withdrawal, distress, fear, attendance change or decline in engagement linked to online activity;
- extremist, misogynistic, hateful, incel or violent content and communities;
- instructions to evade family, employer, provider, police or platform controls; and
- use of AI-generated identities, voice cloning or deepfake content to gain trust or apply pressure.

11.2 Response

Staff must listen, avoid blame, avoid promising confidentiality, record the learner’s words accurately and report promptly to the DSL. Staff must not investigate by contacting a suspected perpetrator or setting up covert communication. Immediate danger, suspected crime or risk of significant harm must be escalated through emergency and safeguarding procedures.

12 Cyberbullying, harassment and child-on-child harm

- CWCT will not tolerate bullying, harassment, violence, discrimination, sexual harassment or abuse through digital technology
- Reports will be taken seriously even where content has been deleted, is anonymous, occurred outside CWCT, or involves learners from different settings
- The response will consider the needs and safety of the person harmed, the alleged perpetrator and other affected learners
- Measures may include safeguarding support, risk assessment, preservation of evidence, platform reporting, removal requests, education, restorative or disciplinary action, employer involvement and referral to external agencies
- Retaliation, victim-blaming and repeated circulation of harmful content are prohibited

13 Filtering and monitoring arrangements

13.1 Principles

Filtering is preventative and seeks to block illegal, inappropriate or potentially harmful content. Monitoring is reactive and identifies concerning activity or produces alerts for review. Neither replaces staff vigilance, education, professional curiosity or safeguarding reporting.

13.2 CWCT arrangements

- Filtering will be risk-based and proportionate to users, devices, curriculum needs, locations and modes of delivery. It should cover web links and digital content, including text, image, audio, video and AI-generated content, as far as technically practicable.
- Monitoring may include technical monitoring, system and security logs, alerts, audit trails and active supervision of screens in learning environments.
- Access to monitoring data will be restricted to authorised personnel and handled lawfully, proportionately and securely.
- Alerts will be triaged promptly against an agreed safeguarding and IT process. Safeguarding interpretation and action will be led by the DSL or authorised deputy.
- Attempts to bypass controls, including use of proxies, unauthorised VPNs, personal hotspots or alternative accounts, are prohibited.
- Reasonable educational access will be maintained. Requests to unblock or reclassify content will follow a documented approval process with safeguarding oversight where relevant.
- Controls will be considered for remote and cloud use and for personal devices connected to CWCT networks. Where CWCT cannot technically monitor a personal device, users remain bound by this policy and staff must apply appropriate supervision and learning controls.

13.3 Annual and event-driven review

The DSL, IT lead and senior leader will review filtering and monitoring at least annually and after a serious incident, material system change or significant change in learner need or risk. A record will be retained.

Review area	Minimum evidence
Roles and escalation	Named leads, deputies, alert routing, out-of-hours arrangements and decision authority.
Technical coverage	Networks, devices, accounts, remote/cloud access, categories blocked, monitoring capability and known gaps.
Effectiveness	Test results, sample alerts, response times, false positives/negatives, bypass attempts and incidents.
Safeguarding fit	Learner age/profile, SEND and vulnerability, curriculum needs, emerging harms, AI-generated content and local risks.
Privacy and access	Lawful basis, transparency, access permissions, retention, security and audit trail.
Actions	Owner, deadline, priority, completion status and board assurance.

14 Reporting and responding to concerns

14.1 How to report

- Learners should report to any member of staff, their Training Advisor/Assessor, or the safeguarding team using the published safeguarding contact route.
- Staff must report immediately to the DSL or deputy using the approved safeguarding recording process.
- Concerns about a member of staff must follow the allegations/low-level concerns procedure and be reported to the designated senior manager. Concerns about that person must be directed to the alternative named route.
- A data breach, cyber incident or compromised account must also be reported through the IT/data protection incident process.
- If anyone is in immediate danger or a crime is in progress, emergency services must be contacted.

14.2 Immediate actions

- Prioritise immediate safety and medical need;
- Listen calmly and take the report seriously;
- Do not blame, shame or ask the learner to confront anyone;
- Do not promise confidentiality; explain that information will be shared only with those who need it;
- Preserve evidence without further distributing harmful material;
- Record facts, words used, dates, platforms, accounts, links and actions taken; and
- Refer promptly for safeguarding assessment and external referral where required.

14.3 Incident decision-making

The response will be child-centred, timely, proportionate and informed by the nature of the content/contact/conduct/commerce risk, age, vulnerability, intent, impact, recurrence, coercion, reach, legality, risk to others and any offline connection. Disciplinary action will not replace safeguarding action.

15 Data protection, privacy and evidence

- Monitoring and investigation will be lawful, necessary, proportionate and transparent.
- Only the minimum relevant personal data will be accessed, shared or retained.
- Safeguarding records, system logs and evidence will be secured and available only to authorised people.
- Harmful content must not be copied into general records where a descriptive record and secure specialist preservation is sufficient.
- Suspected indecent images of children must not be printed, forwarded, copied or retained by staff outside approved police/safeguarding direction.
- Device searches or access to personal accounts will only occur where lawful, authorised and consistent with CWCT procedures.
- Information will be shared where necessary to protect a child or adult at risk, with decisions and rationale recorded.

16 Education, induction and training

- All staff will read KCSIE 2026 Part One and complete safeguarding induction and updates relevant to their role
- Staff development will include online risk, professional boundaries, AI and deepfakes, exploitation, filtering and monitoring, reporting, evidence handling, data protection and current local/emerging risks
- Learners will receive age- and programme-appropriate online safety education at induction and through the curriculum, reinforced when issues arise
- Teaching will develop critical evaluation of sources, misinformation recognition, privacy, consent, respectful communication, account security, scam awareness, AI literacy and routes for help
- Additional support and accessible formats will be provided where disability, SEND, language, digital exclusion, vulnerability or prior harm creates barriers
- Governors/board members will receive sufficient training to challenge and assure arrangements

17 Remote learning, workplace learning and BYOD

- The same safeguarding and professional standards apply in remote and workplace settings
- Staff must use approved accounts and platforms, suitable environments and professional communication practices
- One-to-one online contact must be planned, observable or auditable as appropriate, and consistent with staff conduct requirements
- Personal devices must have a screen lock, current security updates and secure authentication. Users must not allow others to access CWCT accounts
- CWCT data must not be stored locally on unmanaged devices unless expressly authorised and protected
- Employers supporting apprentices should report online safety concerns affecting the learner and co-operate with safeguarding plans
- Where network filtering does not extend to a learner's home, workplace or mobile data connection, CWCT will reinforce safe practice and reporting routes and use platform/device controls available to it

18 Breaches of policy

Potential breaches will be assessed according to safeguarding risk, impact, intent, recurrence and applicable staff, learner, contractual or disciplinary procedures. Action may include advice, education, restrictions, removal of access, risk management, disciplinary action, withdrawal of service, reporting to an employer, regulator, awarding organisation, platform, police or safeguarding agency. Any action will be reasonable, proportionate and consistent with legal and safeguarding duties.

19 Quality assurance, audit and board assurance

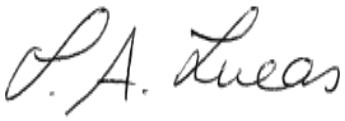
- Periodic review of incident themes, response timeliness and repeat concerns;
- Annual recorded review of filtering and monitoring;
- Sampling of safeguarding records and actions;
- Review of approved AI tools, risk assessments and reported failures;
- Learner and staff feedback on confidence, access and reporting;
- Checks on curriculum coverage, induction and staff training completion;
- Penetration, security or supplier assurance where appropriate; and
- An improvement plan with named owners and completion dates.

Board assurance questions

Are learners safer in practice? Do staff know what to do? Are technical controls appropriate and tested? Are AI and deepfake risks understood? Are incidents leading to learning and improvement? Are vulnerable learners able to report and access support?

20 Review and publication

This policy will be reviewed at least annually, and earlier where there is a relevant statutory update, new technology, serious incident, audit finding or material change to provision. The current version will be made accessible to staff and learners. Key reporting information will be communicated in accessible formats and reinforced through induction and training.



Signed: Executive Director Date: 1st September 2026

Appendix A Staff and learner quick response guide

If you encounter...	Do	Do not
Harmful or illegal content	Stop, move away from public view, record how it came to attention, report to the DSL.	Forward, download, print, repeatedly view or investigate alone.
Suspected deepfake or intimate image	Offer support, preserve available source details, report immediately.	Ask the victim to prove it is fake or circulate it for verification.
AI produces unsafe or confidential output	Stop using the tool, retain safe evidence, report to line manager/IT/DSL as relevant.	Paste more personal data into the tool or publish unchecked output.
Phishing or account compromise	Disconnect or stop interaction, report to IT, follow password/security instructions.	Reply, click further links, delete evidence or conceal the mistake.
Grooming, exploitation or blackmail	Listen, reassure, preserve messages safely, make an urgent safeguarding report.	Contact the suspected perpetrator or advise the learner to pay/comply.
Immediate danger	Contact emergency services and notify safeguarding leadership.	Rely on email or routine reporting alone.

Appendix B Permitted AI use decision check

1. Is the tool approved by CWCT and suitable for the user's age and purpose?
2. Does the use comply with awarding organisation, assessment and customer rules?
3. Will any personal, confidential, safeguarding or commercially sensitive information be entered? If yes, stop unless expressly approved and protected.
4. Can a competent person verify accuracy, bias, copyright, accessibility and safety?
5. Could the output affect a person's rights, safety, assessment, employment or access to support? If yes, ensure meaningful human decision-making and senior approval.
6. Will use be acknowledged and evidenced where required?
7. Is there a clear route to report unexpected or harmful output?

Appendix C Filtering and monitoring annual review record

Field	Record
Review date	
Reviewers and roles	
Systems and locations covered	
Learner/user profile and identified risks	
Tests completed and outcomes	
Monitoring alerts sampled and response findings	
AI-generated content capability and gaps	
Privacy, access and retention checks	
Incidents and lessons learned	
Actions, owners and deadlines	
Senior leader sign-off	
Board/governor assurance date	

Appendix D Related organisational documents

- Safeguarding and Child Protection Policy
- Prevent Policy and risk assessment
- Staff Code of Conduct
- Learner Code of Conduct and disciplinary procedure
- Data Protection and Information Security policies
- Acceptable Use Agreements for staff and learners
- Assessment Malpractice / Academic Integrity procedure
- Whistleblowing and complaints procedures
- Incident response and business continuity arrangements
- Photography, video and consent guidance
- Remote learning and BYOD arrangements

Appendix E External reference points

Keeping Children Safe in Education 2026: https://www.gov.uk/government/publications/keeping-children-safe-in-education--2

DfE filtering and monitoring core standard: https://www.gov.uk/guidance/meeting-digital-and-technology-standards-in-schools-and-colleges/filtering-and-monitoring-core-standard

Working Together to Safeguard Children: https://www.gov.uk/government/publications/working-together-to-safeguard-children--2

Generative AI product safety standards: https://www.gov.uk/government/publications/generative-ai-product-safety-standards

UK Safer Internet Centre: https://saferinternet.org.uk/

Report Remove: https://www.childline.org.uk/info-advice/bullying-abuse-safety/online-mobile-safety/report-remove/

CEOP Safety Centre: https://www.ceop.police.uk/Safety-Centre/