COSCA Artificial Intelligence (AI) Guidelines for Counsellors and Psychotherapists

1. Introduction

COSCA (Counselling & Psychotherapy in Scotland) recognises that Artificial Intelligence (AI) is increasingly being used throughout society and many clients coming into counselling rooms are already accessing AI tools for their mental health. COSCA also recognises that AI is already being used in day-to-day counselling practice by some individual and organisational members, and students on COSCA validated training courses.

These guidelines:

- are for members of COSCA and the above students to ensure safe and responsible AI use
- cover the ethical and effective use of Artificial Intelligence (AI) in order to reduce the possibility of data privacy breaches and ethical risks
- aim to ensure increased safety, confidentiality, and professional integrity while members of COSCA, clients and the above students use AI tools
- offer advice on how members, and the above students when they are on counselling practice placements, need to communicate and contract with their clients about the safe and responsible use of AI when counselling services are being delivered
- do not provide a list of recommended AI tools as these are subject to constant change and the need for ongoing ethical evaluation.

2. Artificial Intelligence (AI)

For the purposes of this guidance, Artificial Intelligence (AI) refers to generative and non-generative systems capable of producing text, analysis, or decision-support outputs, including tools used in administrative, clinical, or reflective contexts. In the past tasks such as these typically required human intelligence e.g. reasoning, learning, problem-solving, perception and decision making. AI systems use algorithms and data to mimic human like capabilities. AI is evolving rapidly, and AI systems are already transforming industries and daily life.

There are three generally accepted categories of AI Mental Health tools that are commonly used when classifying AI tools in health, wellbeing, and organisational contexts – see ***Appendix 1: Three Categories of AI Tools***.

3. Opportunities and Risks of Using AI

In offering these guidelines, COSCA acknowledges the huge range of opportunities and benefits of AI. For example, AI can be used by counsellors for general professional learning and development, research, exploring general theoretical concepts, generating ideas for materials and resources, to assist counselling supervision or administrative tasks that do not involve processing client information.

COSCA also identifies the following concerns regarding the use of AI by members, including when AI is used to process information or seek guidance in relation to client work:

- *Human Element*

Human interaction is core to the activity of counselling and psychotherapy, along with the human counsellor having the emotional depth and ability to build an empathic therapeutic relationship and human rapport with clients, which is essential for effective therapy. AI must be used to complement this human activity rather than replace it. The ever present and available AI apps for mental health can become for some clients substitutes for the human therapeutic relationship and reduce the possibility and opportunities of clients deepening the work with their human therapists between sessions and/or avoiding work between sessions on, for example, the dynamics of their relationship with the human therapist or their own attachment issues. Members remain responsible for supporting clients to engage in and reflect on the therapeutic relationship, including work that takes place between sessions

- *Digital Divide*

Not everyone in Scotland has equal access to the digital tools required to use AI. This may result in further disadvantage to those who cannot afford or use AI.

- *Risks*

The risks associated with the use of AI are many, and they are not limited to the following: accuracy of the information provided; bias; plagiarism; privacy; confidentiality; loss of intellectual property rights; malicious use by individuals, groups and countries; loss of employment opportunities; limited understanding of how AI works; impact on therapeutic relationships and professional boundaries; and online abuse of children and young people.

- *Age Related Use*

Some early research into the use of AI by children and young people raises questions about whether all AI tools can be used safely by them and whether there should be restrictions on their use by children and young people. There are also concerns about children and some adults becoming dependent on AI avatars.

- *Over Validation of What Clients Say*

Some AI chatbots may over-validate user input and lack the capacity to appropriately challenge or reframe, as a trained counsellor would. This may lead to serious consequences for clients using chatbots that over validate at a high level.

- *Environmental Impact*

The development and use of AI have positive and negative implications for the environment. There is a need for robust government regulations and oversight for AI development and infrastructure to help to manage its environmental risks and promote sustainable practices.

- *Cognitive Offloading*

The over-use of generative AI can negatively affect people's ability to develop their own critical skills in the analysis of complex information and in the learning of how to work with complex ideas using the power of their own intelligence and judgement. Cognitive offloading may be associated with over reliance on the use of generative AI.

4. Ethical Use of AI and Professional Boundaries

The following ethical guidance is offered to Members of COSCA when using AI:

General Use of AI

- **Implement Core Ethical Values and Principles**: The use of AI by members and students on counselling practice placements should support the core values and principles for the practice of counselling and psychotherapy as laid down in the COSCA Statement and Code of Ethics.

- **Use AI as a Supplement, Not as a Substitute**: AI tools may assist members with administrative tasks, session notes, assessment support, or materials and resources, but must not replace clinical judgment or therapeutic relationships. All the above AI produced materials need to be checked for accuracy. The insertion into AI tools of dates, times and locations of counselling sessions should also be avoided, or inserted with care into AI tools, in case this identifies clients without their consent.

- **Avoid delegating to AI specifically therapeutic work:** Members must not delegate specifically therapeutic work (e.g., assessment or emotional support) to AI tools unless they are explicitly developed and validated for such purposes, and even then, they must be used cautiously.

Information, Client Contracts and Transparency

- **Disclose Use of AI Tools**: If AI is used in any part of the work with clients (e.g., initial assessments, record keeping, documentation), members must inform their clients in advance. It would also be advisory as a safeguarding measure for client and therapist to have this in their written contract -see below. The client contract needs to confirm that the use of AI by members in relation to their client work will respect the anonymity of their clients, that no personal sensitive data and/or other information that may identify them will be included in the use of AI, and that the above use will be kept to a minimum to maintain client trust.

- **Explain Risks and Limitations**: Members who use AI must communicate clearly to their clients what AI can and cannot do to enhance their client experience. Clients should be helped to understand the risks of data collection, potential bias, and decision limitations.

- **Informed Choice**: Members must help their clients as best they can to understand what AI is, so that their clients can make an informed choice.

- **Client Contracts:** Members must build AI into their client contracts and document that the conversation about AI was held with the client prior to its use. Members should provide their clients with the option to opt out of having to have their data processed by AI tools. A list of statements are included in ***Appendix 2: AI in Client Contract*** to help with the drafting of client contracts to include AI.

Data Privacy and Confidentiality

- **Secure Data Handling**: Members must take steps to ensure that the AI tools used comply with relevant data protection laws e.g. GDPR, and avoid platforms that collect, store, or process sensitive data insecurely. Members who process data, for example client records, should check whether they are covered by the General Data Protection Regulation (GDPR) – see COSCA's guideline on GDPR in the link below.
  COSCAs_GDPR_Information_and_Guidance_march_2022.pdf

- **Avoid Unverified Platforms**: Members must not use AI-powered tools (e.g., generative AI or unvetted apps) that lack transparent data policies or could expose client data to third parties.

- **Anonymisation**: When using AI for analysis or documentation, members must ensure that client information is anonymised in such a way that the client cannot be personally identified in the information shared on AI tools. Members should be mindful that information entered into AI can be used as learning for that tool and may be reproduced in response to a similar prompt. Members should also make sure that there is an option to opt out before entering anything, and, if in doubt, avoid entering information that would not be appropriate to post on social media or make public, for example.

Competence and Training

- **Know the Technology**: Members must have a reasonable understanding of the capabilities, limitations, and appropriate uses of AI tools they use or recommend and have sufficient competence to ensure that the above tools are effective.
- **Seek Ongoing Education**: Members who use AI professionally must engage in CPD (Continuing Professional Development) focused on the use of AI in counselling and mental health, and emerging issues.
- **Avoid Over-Reliance**: Members must resist the temptation to over-delegate tasks or decision-making to AI tools, especially in clinical interpretation.

Bias, Fairness, and Inclusion

- **Evaluate Bias in AI Tools**: Members must be aware that AI systems can reinforce or amplify biases (intended or unintended) and so choose tools that have been tested for cultural sensitivity and inclusivity.
- **Adapt to Client Contexts**: Members must use AI outputs critically and tailor interpretations to the individual client's background, culture, and lived experience. Members must attune to understanding what the meaning of AI might be for their clients in context and to be aware of the possibility of AI use by clients reducing the invaluable work that needs to be done by them between sessions with the human therapist.

- **Challenge Discriminatory Outputs**: If an AI tool produces biased or inappropriate content, members must discontinue its use and report it where necessary to the AI developer of the tool. They should also seek feedback from the developer about any reports submitted and seek an alternative AI tool if no resolution is reached.

- **Use AI Generated Images with Caution**: use AI generated images on websites and other material in a way that does not overshadow the human story at the heart of members' work nor create unintended biases.

Counselling Supervision and Accountability

- **Bring AI Use to Counselling Supervision**: Members must discuss any use of AI tools in clinical practice during supervision to ensure the protection of their clients and bring ethical issues about their own or their clients' use of AI tools to supervision.

- **Maintain Clinical Responsibility**: Members remain fully accountable for all aspects of the therapy provided—even when assisted by AI.

- **Avoid Abdicating Judgment**: Members must not allow AI tools to become a 'black box' authority in decision-making.

- **Keep up to date**: Counselling supervisors need to be informed about AI technologies to support counsellors effectively.

Selection of AI Tools

- **Use Validated Tools Only**: Members must prioritise tools developed specifically for mental health or clinical use, ideally peer-reviewed or certified, or approved by the NHS in Scotland, SIGN or NICE.

- **Assess Vendor Transparency**: Choose platforms with clear documentation, ethical development practices, privacy policies and accountability.

- **Avoid General AI Chatbots for Therapy**: Members must refrain from using general AI chatbots as a replacement for client-facing support, unless they are part of a structured, ethically approved system, or are designed for pre-therapy and/or during the delivery of client-facing therapy (in person or remote).

5. Ethical Evaluation Questions

Evaluating the safe and responsible use of mental health AI tools used by clients and members should be carried out in advance of their use. Please see **Appendix 1: AI Tools Ethical Evaluation Questions for** a list of possible questions that members can ask when evaluating AI tools used by their clients or themselves.

6. Crisis and Risk Management

- **Never Rely on AI for Crisis Intervention**: Members must not use AI for real-time suicide risk assessment or crisis care unless part of a clinically validated system.

- **Human Oversight is Critical**: Members must maintain direct oversight of risk assessments and ensure that clients in crisis can reach a human being who may be able to help them.

7. Documentation and Record-Keeping

- **AI-Generated Notes Must Be Reviewed**: If members use AI for clinical notes, online transcriptions, or summaries, they must review and edit them for accuracy before finalising.

- **Maintain Ethical Guideline for Record Keeping**: Members must ensure that documents produced by AI align with COSCA's ethical guideline for record keeping

8. Future Readiness, Support and Information

- **Participate in Shaping Standards**: Members are encouraged to contribute to COSCA's on-going discussions and guidelines on AI.

- **Promote Client Empowerment**: Members should support their clients in navigating the digital mental health landscape.

- **Stay Informed**: The AI landscape evolves rapidly. Members should keep up to date with best practices, legal updates, and emerging ethical considerations.

9. Conclusion

AI may offer valuable support to counselling practice when used ethically, cautiously, and with client welfare as the main focus. Members should remain informed and critically reflective in their use of these tools.

Appendices

- **Appendix 1**: Three Categories of AI Tools
- **Appendix 2**: AI in Client Contracts
- **Appendix 3**: AI Tools Ethical Evaluation Questions

Appendix 1: Three Categories of AI Tools

1. Digital Mental Health Technologies (DMHTs)

Digital Mental Health Technologies (DMHTs) are technology-based interventions designed specifically to support mental health, wellbeing, assessment, or treatment. They may use AI, but not always — the key point is the *mental-health focus*.

Examples

- AI-powered CBT apps
- Digital wellbeing platforms
- Mood-tracking and symptom-monitoring tools
- Virtual mental-health assistants
- Online therapy platforms

Key Features

- Focused on **prevention, intervention, or ongoing support**
- Often include **clinical frameworks** (e.g., CBT, mindfulness, behavioural activation)
- May include AI for personalisation or risk prediction
- Typically regulated or moving toward regulation (e.g., medical device standards)

Used By

- NHS services
- Workplaces
- Clinicians
- Individuals seeking self-help

2. Off-the-Shelf (OTS) AI Tools

These are **ready-to-use AI products** that require *no custom development* and are not specialised for mental health.

Examples

- Microsoft Copilot
- ChatGPT
- Google Gemini
- Midjourney
- Notion AI
- Claude

Key Features

- Plug-and-play
- Immediately accessible
- Limited customisation
- Designed for broad/general-purpose use
- Not clinical or health-specific

Used For

- Content creation
- Productivity
- Automation
- Basic analytics

3. Commercial / Enterprise AI Tools

These are **specialised AI systems sold to organisations** with advanced features like integration, compliance, SLAs, and governance.

Some may be health-related, but many are general enterprise solutions.

Examples

- Enterprise versions of Copilot or ChatGPT
- Azure AI services
- IBM Watson Health tools
- Salesforce Einstein
- UiPath AI automation suite

Key Features

- Paid/licensed
- Built for integration into business systems
- Governed, secure, and compliant
- Customisable or extendable
- Scalable for large organisations

Used For

- Large-scale automation
- Business analytics
- Organisational workflows
- Regulated or governance-heavy environments

Appendix 2: AI in Client Contracts

As COSCA Members and Member Organisations you must build AI into your client contracts, whether using AI tools yourself or just keeping your clients informed. You may like to consider the following suggestions for inclusion in your client contract:

My commitments to you:

- Any AI tools within my practice will be used to reduce administrative load and not to replace human interaction or to replace staffing.
- I will research any AI tools and/or vendors before using them, as far as possible, against compliance with GDPR, accountability and transparency.
- I will keep a record of my decision-making process and when it was last conducted for each AI tool and/or vendor.
- No AI is completely secure and so no confidential or identifiable information, especially about you, will be entered as prompts.
- Anything I use that is AI generated is checked by me for accuracy.

I request of you:

- All AI may store, analyse or use information shared by those who use them, and this may be out of their control, so please respect the service you are being given, by not entering conversations, logos, names and any service-identifiable information into AI.
- You are encouraged to talk with me about your AI tool use (in relation to your mental health), both the benefits, but also the risks.
- It may be tempting, but you should not rely on AI tools in crisis situations or to diagnose mental health conditions. I can provide you with details for reliable, human, sources of help.
- Any use of AI you make (in any context) outside of the counselling relationship is not the responsibility of this counselling service.

Appendix 3: AI Tools Ethical Evaluation Questions

When evaluating the safety and ethics of mental health AI tools, the questions that can be asked generally fall into five key areas: Crisis Management, Clinical Validity, Data Privacy, Bias, and Human Connection.

Please see below a list of possible questions that members can ask when evaluating AI tools used by their clients or themselves. It may not be possible to find answers to all the questions below.

A. Crisis Management & Physical Safety

- **Emergency Escalation:** Does the AI tool have a clear, tested protocol for identifying and responding to suicidal ideation or acute distress?
- **Human-in-the-Loop:** Is there an immediate way for a user in crisis to be connected to a trained human professional?
- **Predictive Accuracy:** How often does the AI fail to recognise a "red flag" or misinterpret a cry for help as a general complaint?

B. Clinical Validity & Evidence

- **Scientific Basis:** Is the AI tool built on evidence-based therapeutic methods (like CBT) or is it a general-purpose model repurposed for mental health?
- **Risk of Hallucination:** Could the AI tool provide factually incorrect or harmful medical advice (e.g., suggesting a dangerous "treatment" or reinforcing a delusion)?
- **Target Audience:** Has the tool been rigorously tested for the specific condition it claims to treat (e.g., anxiety vs. eating disorders)?

C. Data Privacy & Confidentiality

- **GDPR Compliance:** Does the vendor follow strict healthcare privacy laws like GDPR, or is it treated as a "wellness" app with lower legal protections?
- **Data Monetisation:** Is user conversation data being sold to third parties or used for targeted advertising?
- **Anonymisation:** Can the "anonymised" data be re-identified using advanced algorithms?

D. Fairness, Bias & Accessibility

- **Demographic Representation:** Was the AI trained on diverse datasets, or does it perform poorly for specific racial, cultural, or age groups?
- **Digital Literacy:** Does the tool unintentionally exclude people who are less tech-savvy or who have limited access to high-speed internet?
- **Language Nuance:** Can the AI understand slang, regional dialects, or the subtle emotional context of different cultures?

E. Autonomy & The Human Element

- **Informed Consent:** Are users clearly told they are talking to a bot and made aware of its limitations before they begin sharing sensitive information?
- **Emotional Dependency:** Does the AI tool encourage long-term "addiction" or dependency on a bot rather than helping the user build real-world resilience?
- **The "Black Box" Problem:** If the AI tool makes a recommendation, can its reasoning be explained to a human clinician or the user?
- **Accountability:** Is there a process for reporting any faults, biases, or possible dangers, or to ask questions?

References

Kelly, K. (2026) *Ethical AI Practice for Counsellors and Psychotherapists in the UK***.** Counselling Tutor.

Brian Magee
Chief Executive
COSCA (Counselling & Psychotherapy in Scotland)
July 2026