



COUNSELLING
TUTOR



The AI & Therapy Critical Thinking Matrix

An Ethical Evaluation Framework for Counsellors and Psychotherapists


This matrix was developed by Kenneth Kelly for counsellors, psychotherapists, and supervisors working in the United Kingdom context. AI- and digital-based tools are emerging rapidly in therapeutic settings. As a counsellor or psychotherapist, you need a reliable, practitioner-centred method to assess whether any given tool honours client safety, data protection, consent, therapeutic integrity, and current regulatory guidance. This matrix gives you exactly that.

Why Therapists Need a Critical Thinking Approach to AI

AI systems and other digital mental health tools are entering the therapeutic landscape faster than most practitioners can realistically track. That pace can create uncertainty:

- Is this tool safe for my clients?
- Does it comply with United Kingdom General Data Protection Regulation (UK GDPR)?
- Will it strengthen or weaken the therapeutic relationship?
- Where does my professional responsibility begin and end when using AI?






In the United Kingdom, many apps, web platforms and AI tools used around mental health are now described as Digital Mental Health Technologies (DMHTs).

Digital Mental Health Technologies are digital and software products designed to support mental health and wellbeing, and some of them may be regulated as medical devices when they are intended to diagnose, monitor, treat or prevent mental health conditions. The Medicines and Healthcare products Regulatory Agency (MHRA) has issued guidance for both the general public and clinicians to help them ask clear questions about what a mental health app or technology is for, what evidence supports it, how safe it is, and how to raise concerns.

This matrix provides a clear, structured way to reflect on those questions before adopting any digital or AI-enabled tool. It is designed to sit alongside MHRA guidance on Digital Mental Health Technologies and help you translate those expectations into everyday clinical decision-making.

What This Matrix Helps You Do

- **Protect client confidentiality and autonomy** in a rapidly changing digital landscape.
- **Evaluate claims made by AI tools and Digital Mental Health Technologies with clarity**, rather than relying on marketing language or assumptions.
- **Align your decisions with UK GDPR, the Data Usage and Accountability Act (DUAA)**, MHRA guidance on Digital Mental Health Technologies, and professional ethical frameworks, reducing risk for both client and therapist.
- **Safeguard the therapeutic relationship**, ensuring technology never replaces or dilutes human connection.
- **Maintain professional accountability**, even when tools introduce automation or suggestions.

- 
- **Make confident, informed choices** that withstand supervision, ethical scrutiny, regulatory expectations, and client questioning.

A Tool Designed Specifically for UK Practice

This matrix was developed by Kenneth Kelly for counsellors, psychotherapists, and supervisors working in the UK context, where expectations around data protection, client consent, relational ethics, and digital regulation differ from many international frameworks.

It is designed to be:

- **Practical** - something you can use immediately.
- **Non-technical** - no specialist digital or regulatory knowledge required.
- **Ethically grounded** - aligned with UK professional standards and current UK guidance on Digital Mental Health Technologies.
- **Client-centred** - prioritising therapeutic integrity over convenience or novelty.

WHAT THE MATRIX COVERS

The Ethical Domains

The full Critical Thinking Matrix explores **ten key areas** that therapists should consider before adopting or recommending any AI or digital tool, including **Digital Mental Health Technologies (DMHTs)**. Each domain supports ethical, relational, and legally compliant practice.



1. Intended Purpose & Regulatory Status

Clarifying what the tool is for, who it is for, and whether it may count as a regulated Digital Mental Health Technology or medical device.

2. Confidentiality & Data Privacy

How data is stored, processed, accessed, secured, and protected.

3. Informed Consent

Whether clients understand what the tool does and how their data is used.

4. Therapeutic Relationship & Depersonalisation

Ensuring technology does not undermine human connection or clinical judgement.

5. Bias & Discrimination

How the tool identifies, mitigates, or inadvertently reinforces bias.

6. Professional Accountability & Responsibility

Clarifying where the therapist remains responsible for decisions and outcomes.

7. Technical Reliability & Validity

Whether there is evidence supporting the tool's reliability, safety, and limitations.

8. Client Autonomy

Whether the tool supports or diminishes client choice, agency, and empowerment.

9. Regulatory & Ethical Alignment

Compliance with UK GDPR, the Data Usage and Accountability Act (DUAA), MHRA guidance on Digital Mental Health Technologies, and relevant professional frameworks.

10. Implementation, Monitoring & Review

Practical integration into practice, supervision considerations, reporting concerns, and ongoing evaluation.





HOW TO USE THIS IN PRACTICE

This matrix is intended to be a practical decision-making companion, not a technical manual. Practitioners typically use it in several ways:

1. Before Adopting a New Tool

Use the matrix as a structured pre-assessment to ensure the tool supports client safety, ethical practice, professional standards, and current MHRA expectations for Digital Mental Health Technologies where applicable.

2. In Supervision

Discuss areas of uncertainty, risk, judgement, or boundary considerations with your supervisor, including questions about whether a particular Digital Mental Health Technology may be acting as a medical device and how this affects your responsibilities.

3. During Organisational Decision-Making

If you're part of a service, training organisation, or multidisciplinary team, the matrix provides a shared language for evaluating tools and aligning policies with data protection law, professional ethics and MHRA guidance on Digital Mental Health Technologies.

4. For Ongoing Review

All tools and mental health apps update frequently. Revisiting the matrix helps ensure continued compliance, ethical alignment, and awareness of new guidance or safety information over time.

5. For Client Conversations

Where appropriate, the matrix can support clear, transparent explanations that help clients make informed decisions about the tools you recommend or use together. You may also signpost clients and carers to

MHRA's public guidance on choosing and using mental health apps, so they can apply similar critical questions in their own digital lives.

This tool is not about blocking innovation – it is about ensuring technology enhances, rather than diminishes, therapeutic work.

THE CRITICAL THINKING MATRIX

You can use the following matrix to evaluate whether a tool or app is ethical, clinically appropriate, and aligned with UK GDPR, Digital Mental Health Technologies guidance, and data protection requirements.

Intended Purpose & Regulatory Status

- **What is the tool explicitly claiming to do** (for example, general wellbeing, psychoeducation, mood tracking, or diagnosis / treatment / relapse prevention for a specific condition)? Are these claims clearly and accessibly described?
- **If the tool claims to diagnose, predict, monitor, treat or prevent mental health conditions or symptoms**, does the developer indicate whether it is considered a Digital Mental Health Technology with a medical purpose and, if so, whether it qualifies as software as a medical device under MHRA rules?
- **Is there clear information about the intended user group** (for example, adults, young people, children, specific diagnostic groups), and is this compatible with your client population?
- **Does the stated purpose match how you and your clients are likely to use the tool in practice**, or might your use drift into a higher-risk or different regulatory category (for example, using a wellbeing app as if it were a diagnostic tool)?



Confidentiality & Data Privacy

- **Does the tool comply with UK GDPR?** This includes aspects like data storage, processing, and access. Consider whether the tool has clear data retention policies and procedures for data breaches.
- **Does the tool have robust security measures to protect client data?** Evaluate encryption protocols, access controls, and any security certifications the provider describes.
- **Does the tool's data handling align with professional ethical guidelines for counsellors and psychotherapists in the UK?** Professional bodies such as the British Association for Counselling and Psychotherapy (BACP), the UK Council for Psychotherapy (UKCP) and the National Counselling and Psychotherapy Society (NCPS) may have specific guidance on data protection in therapy.
- **Is the data stored within the UK or transferred internationally?** If data is transferred internationally, are there adequate safeguards in place to ensure UK GDPR compliance?

Informed Consent

- **Does the tool facilitate obtaining informed consent from clients?** Can clients easily understand how the tool works and how their data is used?
- **Does the tool address the complexity of AI and digital processing in a way that clients can understand?** Consider whether the tool uses clear and accessible language to explain its processes.
- **Does the tool allow clients to control their data and withdraw consent?** Ensure clients have options to access, modify, or delete their data.
- **For children and young people, does the tool provide age-appropriate explanations** of what it does, how data is used, and what to do if something feels unsafe or confusing, in addition to any parental or carer consent that may be required?

Depersonalisation & Therapeutic Relationship

- **Could the tool lead to a depersonalised therapeutic experience?** Consider whether the tool prioritises human interaction and the therapist’s role.
- **Could the tool negatively impact the therapeutic relationship?** Assess whether the tool might undermine trust, rapport, or the client’s sense of autonomy.
- **Does the tool allow for flexibility and personalisation in therapy?** Ensure the tool does not impose rigid processes that hinder the therapist’s clinical judgement.
- **Does the tool risk being experienced as a replacement for human therapeutic support** (for example, conversational agents presented as “your therapist”), and how will you frame its role so that it clearly supplements rather than replaces the therapeutic relationship?

Bias & Discrimination


- **Does the tool mitigate potential biases in its algorithms?** Consider whether the tool has undergone bias audits and fairness assessments.
- **Has the tool been tested on diverse client groups, including the ages, cultural backgrounds, and identities you commonly work with?** Ensure the tool does not perpetuate existing biases or discriminate against certain groups.
- **Has the tool been evaluated for the specific age range and population you work with** (for example, children, adolescents, adults, older adults), rather than assuming that findings in one group automatically apply to another?
- **Does the tool allow for adjustments based on individual client characteristics?** The tool should not rely solely on algorithmic outputs and should allow for therapist adjustments.

Professional Accountability & Responsibility

- **Does the tool support the therapist's professional judgement and decision-making?** Ensure the tool does not replace the therapist's role as the primary decision-maker.
- **Does the tool provide clear guidelines for accountability in case of adverse outcomes?** Consider who is responsible if the tool's recommendations lead to negative consequences (for example, the provider, the organisation, the practitioner).
- **Does the tool promote ethical practice and adherence to professional standards?** It should align with ethical guidelines and not encourage practices that violate professional boundaries.
- **If a client experiences harm, technical failure, or concerning content** when using the tool (for example, unsafe advice, self-harm prompts or serious glitches), do you know how to document this, respond clinically, and report it to the appropriate bodies (for example, via MHRA reporting routes, the organisation's risk system, and the app provider)?

Technical Reliability & Validity

- **Is there clear, accessible evidence supporting the tool's safety, reliability and effectiveness** for its stated purpose and user group (for example, independent evaluations, published studies, or structured real-world assessments)?
- **Are claims of benefit proportionate to the evidence**, or does the marketing suggest outcomes (for example, "cures anxiety", "instant therapy") that go beyond what is supported? Be wary of exaggerated or absolute claims.
- **Does the tool provide transparent information about its limitations?** Ensure the tool does not overstate its capabilities or guarantee specific outcomes.

- 
- **Is the tool regularly updated and maintained to ensure accuracy?** Check for updates, bug fixes, and improvements based on user feedback and evolving scientific knowledge.
 - **Where the tool appears to qualify as a Digital Mental Health Technology with a medical purpose, has the developer indicated how it has been characterised** and, if relevant, evaluated under UK medical device rules (for example, as software as a medical device, and with what level of risk classification)?

Impact on Client Autonomy

- **Could the tool manipulate client decision-making or undermine their autonomy?** Consider whether the tool uses persuasive techniques or limits client choices.
- **Does the tool promote client empowerment and self-determination?** Ensure the tool supports the client's active participation and informed decision-making in their therapy.
- **Does the tool provide clients with options and control over their engagement with the tool?** Clients should be able to adjust settings, access information, and disengage from the tool as needed.

Regulatory & Ethical Guidelines

- **Does the tool comply with evolving regulatory and ethical guidelines for AI and Digital Mental Health Technologies in therapy?** Stay updated on relevant regulations and professional standards for using digital tools in mental health.
- **Does the tool address emerging ethical concerns related to AI in healthcare?** Be aware of potential issues like algorithmic bias, data privacy, and the impact of AI on human relationships.
- **Does the tool provide resources for therapists to stay informed about ethical and regulatory considerations?** Look for features such as links to

ethical guidelines, access to expert opinions, or forums for discussion.

- **Does the developer explain whether the tool is considered a Digital Mental Health Technology**, whether it has a medical purpose, and, if so, whether it is subject to UK medical device regulation and classification (for example, Class I, IIa, IIb or III)?
- **Do your own policies and procedures reflect up-to-date MHRA and National Institute for Health and Care Excellence (NICE) guidance on Digital Mental Health Technologies**, not just generic data protection and confidentiality requirements?

Implementation, Monitoring & Review

- **How will the tool be integrated into your practice?** Consider factors such as training requirements, data integration with existing systems, and client onboarding processes.
- **Is there an ongoing review process to monitor the tool's impact and effectiveness?** Establish procedures for evaluating client feedback, assessing therapeutic outcomes, and making adjustments to the tool's usage.
- **Does the tool provider offer support and guidance for implementation and troubleshooting?** Ensure you have access to technical assistance, training materials, and resources for addressing potential issues.
- **Are you clear about how to raise and escalate concerns about the tool's safety or performance** (for example, internal incident reporting, contacting the provider, and using MHRA routes where relevant for Digital Mental Health Technologies)?

The questions in this matrix can help UK counsellors and psychotherapists critically evaluate the ethical, clinical and legal implications of using AI tools and Digital Mental Health Technologies in their practice.



VERSION, ALIGNMENT & NOTES

Version Information

- Version: 1.1
- Release Date: January 2026
- Updated periodically as AI guidance, Digital Mental Health Technologies regulation and professional expectations evolve.

Ethical & Regulatory Alignment

This resource aligns with:

- UK GDPR and the Data Usage and Accountability Act (DUAA)
- Professional ethical frameworks (BACP, UKCP, NCPS)
- Medicines and Healthcare products Regulatory Agency (MHRA) and National Institute for Health and Care Excellence (NICE) guidance on Digital Mental Health Technologies (DMHTs)
- Principles of relational, client-centred practice

Professional Notes

- This matrix is a reflective support tool, not legal advice.
- Always consider supervision, organisational policy, and up-to-date regulatory guidance, including MHRA and NICE resources on Digital Mental Health Technologies.
- The practitioner remains responsible for clinical decision-making.

Copyright

© Kenneth Kelly, Counselling Tutor

Permission granted for UK therapists, supervisors, and training organisations to share this overview document.

References and Further Reading

British Association for Counselling and Psychotherapy (BACP). *Ethical Framework for the Counselling Professions* and digital practice guidance.

British Psychological Society (BPS). *Statements and Commentary on New MHRA Guidance for Mental Health Apps and Technologies*.

Expert Reference Group on the Use of Artificial Intelligence in Counselling and Psychotherapy. *Shared AI Charter for UK Counselling and Psychotherapy Organisations* and associated resources on ethical AI.


Information Commissioner's Office (ICO). *Guide to the UK General Data Protection Regulation (UK GDPR)* and resources on data protection in health and social care.

Kelly, K. *The AI & Therapy Critical Thinking Matrix: An Ethical Evaluation Framework for Counsellors and Psychotherapists*. Counselling Tutor.

Medicines and Healthcare products Regulatory Agency (MHRA). *Digital Mental Health Technologies – Regulation and Evaluation for Manufacturers, Healthcare Professionals and the Public*. UK Government guidance collection.

Medicines and Healthcare products Regulatory Agency (MHRA). *Digital Mental Health Technology (DMHT) – Device Characterisation and Regulatory Expectations*.

Medicines and Healthcare products Regulatory Agency (MHRA). *MHRA Issues New Guidance for People Using Mental Health Apps and Technologies*. UK Government news release.



MindEd. *Digital Mental Health Technologies (DMHT) – e-learning resources for practitioners, parents and carers.*

National Counselling and Psychotherapy Society (NCPS). *Ethical Framework and Guidance on the Use of Technology in Counselling and Psychotherapy.*

National Institute for Health and Care Excellence (NICE). *Guidance and Evidence Standards for Digital Health Technologies* (including digital mental health tools).

UK Council for Psychotherapy (UKCP). *Ethical Principles and Code of Professional Conduct* and guidance on online / digital practice.

UK Government. *Data Protection Act 2018* and *Data Usage and Accountability Act (DUAA)* – legislation governing personal data use.