

# Counselling and AI

Unit 3 - Preparing for the future.

*Richard Miller*

# Objectives

- Finding your AI perspective.
- Scenarios
- The tests
- Conducting a risk assessment
- Consider suggestions of better practice



# What we've learned so far

## Quick refresher

- AI is likely to become widely used both inside and outside of sessions.
- AI tools are not automatically ethically aligned.
- Some AI tools are powerfully persuasive and may have unintended harmful effects.
- As counsellors, we can take a leading role in shaping how AI is used in our profession.

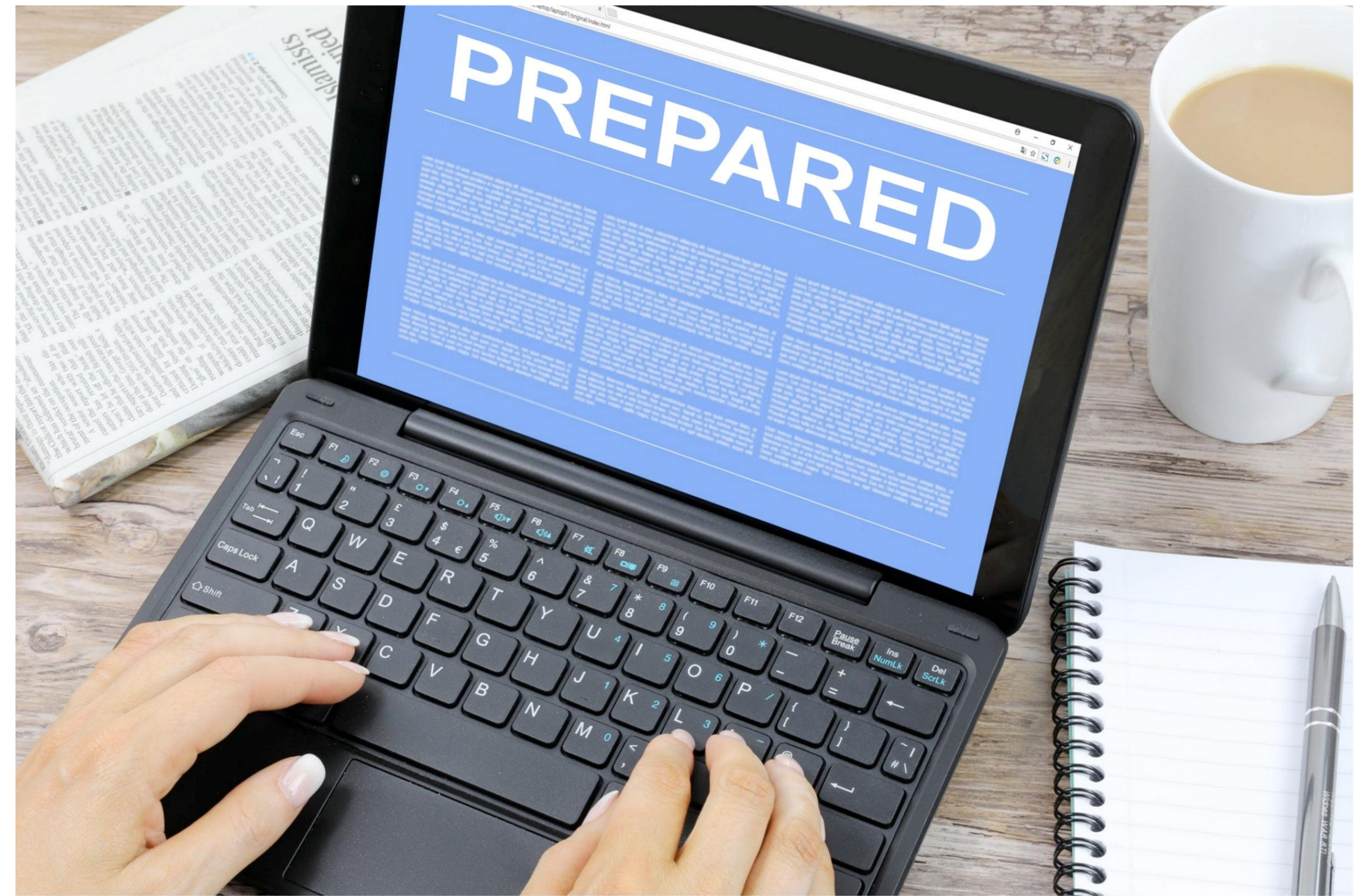
# The tough module

## Saving the best for last

- I'm aware you are still likely processing the prior sessions.
- This is a lot to take in.
- This module will challenge you to do *even more* in-depth thinking.
- Some of this might feel quite hard, overly serious and a tad bleak.
- I would rather we take AI seriously now so we can enjoy it later.



Why should  
we prepare for  
the future?



# Detrimental outcome scenarios

## What we can prevent happening

- **Race to the bottom for privacy standards.** As AI tools become integrated into counselling contexts, there could be a competitive push to gather more personal data to improve AI effectiveness. This could lead to a gradual erosion of confidentiality standards.
- **Reliance on AI diagnostics** - If AI becomes skilled at diagnosing, or assessing, mental health conditions it may generate a trend to overly rely on these tools. In this scenario the nuanced understanding and human judgment of counsellors may become undervalued. This may lead towards misdiagnosis, unhelpful standardisation and normalisation of oversimplified interventions.
- **Quality vs quantity in delivery** - AI may allow counselling services to be delivered more efficiently to a larger number of clients however, this may create a scenario where the focus shifts to serving more clients at the expense of quality of care as services seek to maximise client turnover.
- **Homogenisation of approaches.** - AI systems may start converging on widely applicable therapeutic techniques at the expense of diverse or personalised approaches. This could lead to situations where client needs are not addressed because the AI driven methods favour 'one-size-fits-all' solutions. The approaches used by the AI will be informed by the data it is trained on, therefore more widely used approaches will have larger training datasets, and may therefore be more present in AI tools.
- **Competitive pressures.** - As AI tools become more sophisticated, counsellors may feel pressured to adopt specific AI technologies to stay competitive, even if they believe these tools are not in the best interest of their clients, or if their judgement disagrees with the AI's approach.
- **Economic incentives vs client wellbeing** - As the cost of compute reduces, AI tools may be deployed based primarily on their ability to reduce costs rather than their efficacy in improving client outcomes.

# Beneficial outcomes

## What we could promote happening

- **Enhanced Personalised care** - AI tools are developed which can analyse vast amounts of data and create highly personalised interventions. In this scenario competition between counselling service providers may move towards the most tailored and effective support and may improve client outcomes.
- **Superior training and supervision** - AI-driven simulations and training programs become a benchmark in the profession. Counsellors and therapists compete to train with advanced systems leading to a higher overall standard of practice and more skilled professionals in the field.
- **Ethical AI development race** - As public awareness and demand of ethical AI's grow, organisations start to compete to develop the most ethically sound and transparent AI tools. This leads to advancements in ethical AI, setting high standards for privacy, consent and client rights.
- **Quality-Driven market dynamic** - As the cost of therapy approaches £0 the market competition may shift from competing on cost to focusing on quality of service. AI tools that help achieve meaningful and enduring improvements in client wellbeing become highly sought after which generates further innovation and investment in evidence-based counselling.

Why should / prepare for the future?

# Preparing yourself for AI

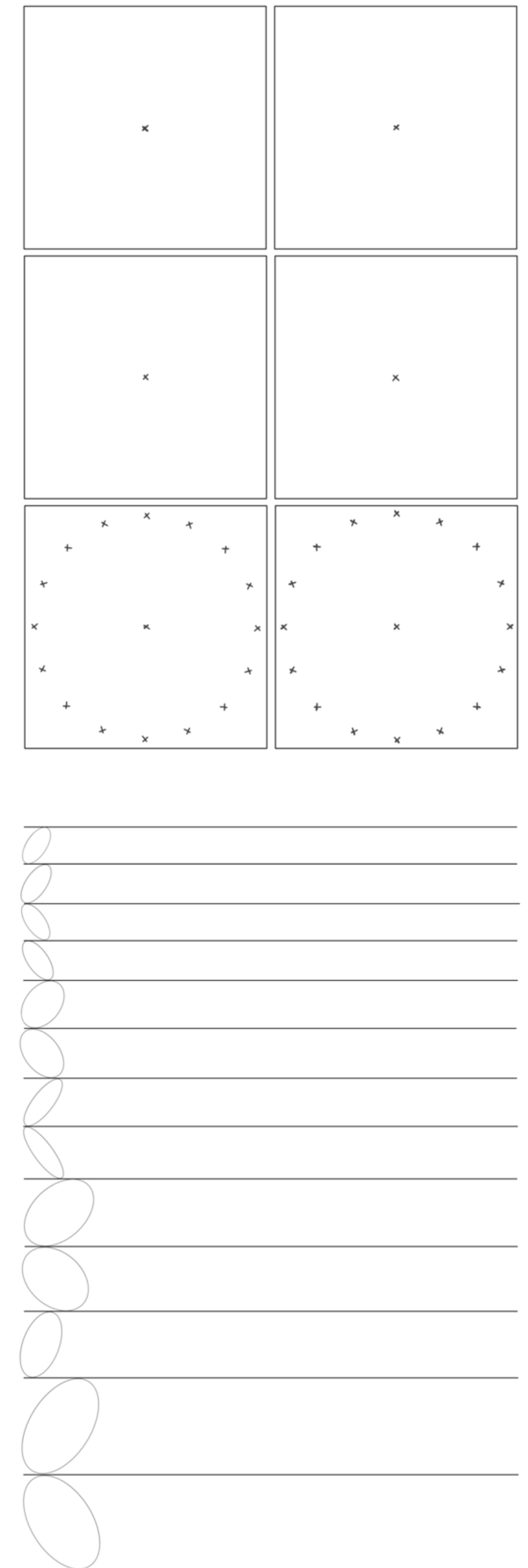
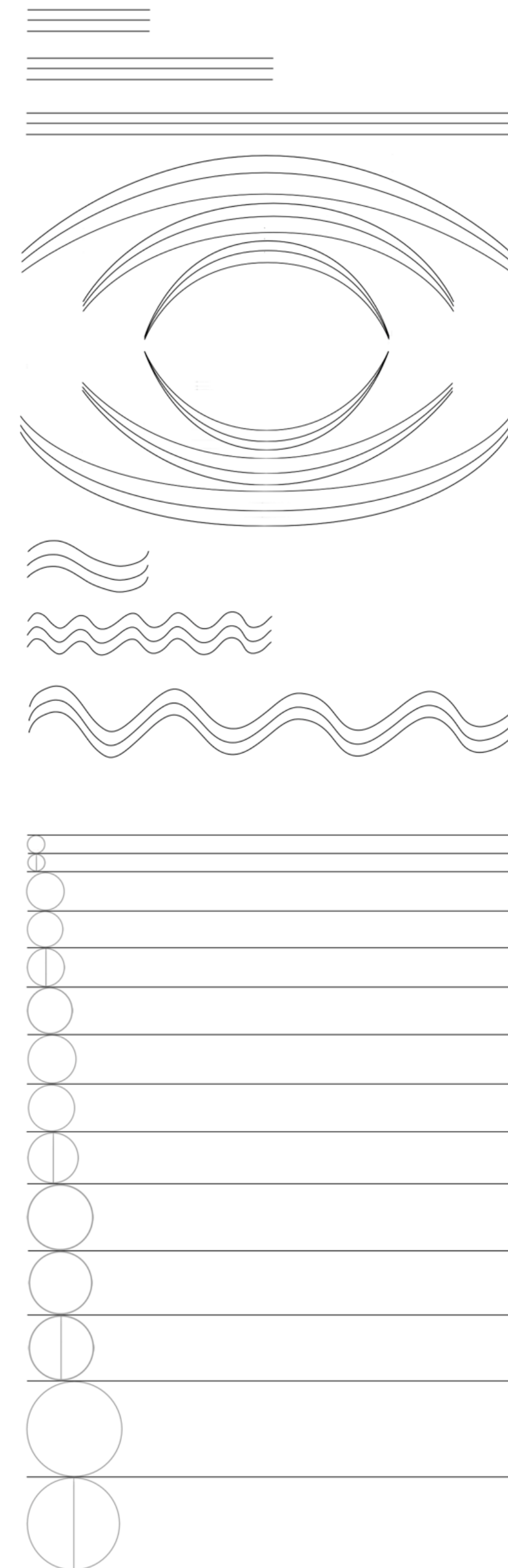
## Establishing your ethical approach

- What AI tools you choose to use, if any, should be guided by your understanding of the technology and your own ethical rationale.
- If you do not determine your own view on AI, you will likely be influenced by changes in technology, industry and social norms. This isn't necessarily bad.
- Counsellors can be responsible for thinking about how AI should be used in counselling contexts. Otherwise, we are leaving it for the tech sector to decide the future of our profession.
- Your idea of progress may be very different from others.
- Your level of preparedness should reflect the amount you intend to use AI.

# Things we could do with AI

If you are feeling tech-savvy

- Generate worksheets.
- Automate all of our emails.
- Use AI to generate blog posts.
- Use AI for SEO to promote their website.
- Discuss generic ethical dilemmas with a chatbot.
- Use AI to teach yourself more about DBT.
- Add an AI assistant to your webpage.



# Even more applications....

- Upload recordings of your sessions to transcribe it so you can share it with your clients.
- Use AI during a session to generate bespoke homework materials.
- Use AI for supervision when you can't see your supervisor and want to discuss how to support your most fragile clients.
- Upload all of your case notes to ask the AI for suggestions on how best to help your clients.
- Offer your clients an AI chatbot version of yourself that they can speak to while you are on holiday.

**Starting to feel uncomfortable yet?**

# It could get quite grim....

- You use an AI model to replicate your voice and image and use it to automatically conduct sessions with multiple clients simultaneously without informing your clients.
- You let an AI write all your case notes and then don't read them to check if they are correct.
- You let your clients have access to an AI chatbot version of yourself any time they want. You then don't check what the client talks about with the bot.
- You get the client to sign a confidentiality agreement that says you use AI but don't explain the significance of this clearly.
- You upload all of the information you can collect on your clients, including recordings of sessions, and use this to train your own AI model.
- You use a chatbot to have a romance with a simulated version of a client you have a crush on, to help you get over it....
- You use AI to write your accreditation submission.
- You don't tell your supervisor about how you use AI because you are worried you might get in trouble.

# Pop your thinking hats on

**Honour the experience and knowledge you've earned over the years you've trained and practiced.**

**Take the next few minutes to reflect on the following questions.**

**Think hard.**

**Jot down your answers if you can.**

# Finding your view on AI

## Reflective questions.

- What are your general views on AI's role in counselling?
- Do you see a role for AI in your practice? If so, how would you use it?
- How comfortable are you with using AI for admin tasks?
- How comfortable are you with using AI for client-related activities?
- What boundaries, if any, would you want to place on the use of AI in your practice?

# Reflective questions part 2.

- How important is it for you to discuss AI use in supervision or to have specific training in this area?
- Would you consider sharing an AI usage policy with clients? Why or why not?
- What considerations do you have around using AI directly with a client in session?
- Do you feel a responsibility to shape or challenge the development of AI in the counselling profession?
- What training or CPD would help you feel competent to use AI, if at all?

# Reflective questions part 3.

- What excites you the most about AI?
- Can you picture how it might help you improve/worsen your practice?
- Can you picture how it might improve/worsen your client's wellbeing?
- In your practice, what tasks would you most like an AI to help you with?
- Does the future feel more exciting because of the potential of AI?

You may need more time to determine your own views.

These questions were intended to help you begin the process of finding your view.

After today, keep thinking about it.

# Finding your attitude to AI

**We don't all have to see AI the same way.**

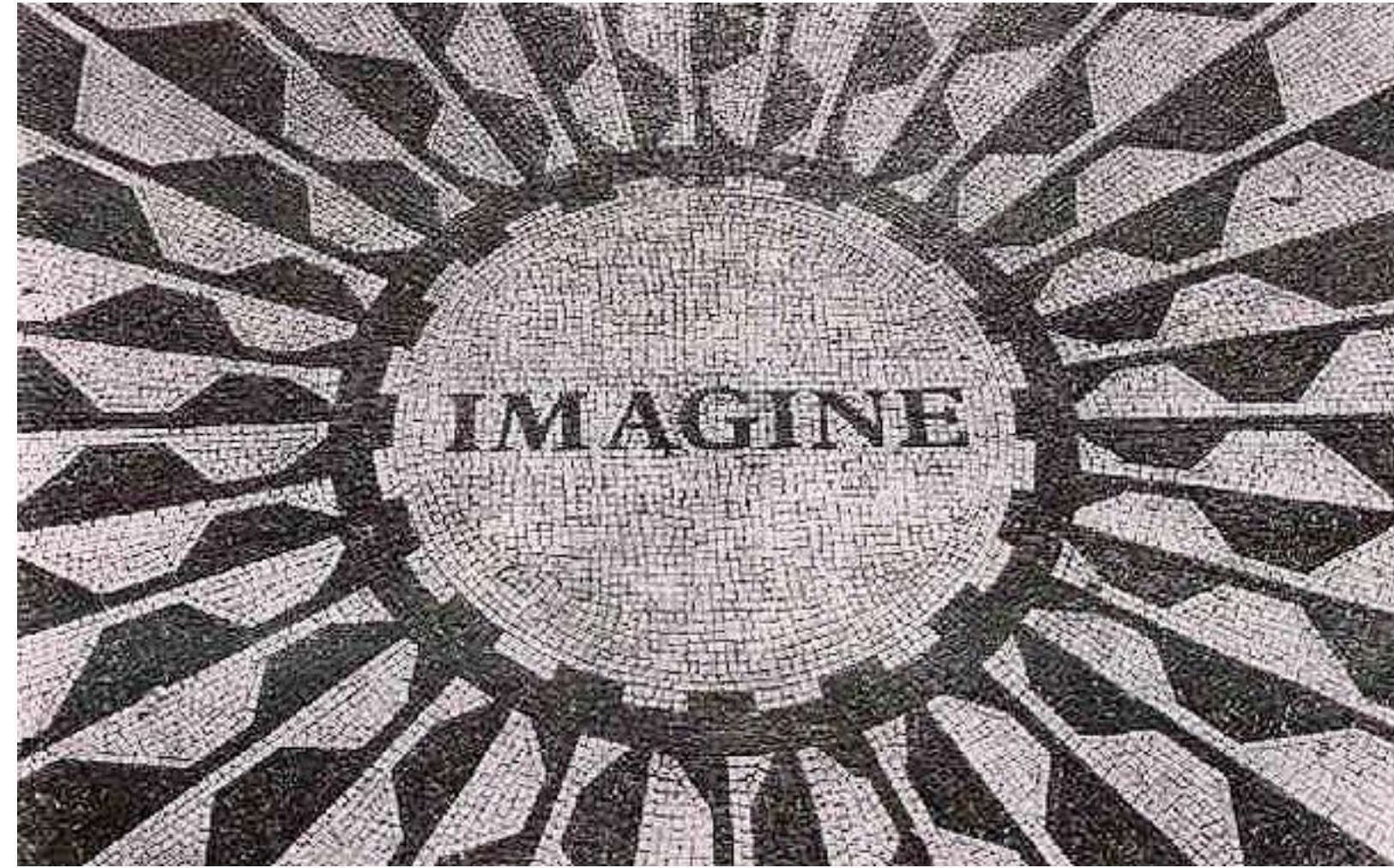
- **Secure** - You would not use AI in any way in any part of your practice because you believe it's use is unethical.
- **Defensive** - You prefer to generally avoid using AI, unless in specific, pre-planned purposes. You would not ever use it in your client work.
- **Cautious** - You would use AI from time to time to help with various administrative tasks and will only occasionally use it to reflect on your work with clients in pre-planned circumstances.
- **Balanced** - You use AI regularly to help manage your administrative processes, you may dialogue with an AI between sessions to reflect on your work, manage case notes, and generate ideas for supporting your practice. You have a risk assessment in place, and routinely discuss AI with your supervisor and have training and insurance.

# Finding your attitude to AI

**We don't all have to see AI the same way.**

- **Balanced aggressive** - You would use AI frequently in nearly every aspect of your administrative processes and automate much of your work. You offer AI-based services to your clients and use AI in-session with their informed consent. You have a risk assessment, have published your AI usage policy publicly, and have a dedicated AI complaints process. You have training and insurance and regularly seek CPD in this area.
- **Aggressive** - You believe not using AI is unethical and routinely use AI to determine optimal methods to support clients. You offer clients access to AI chatbots trained on your own practice, use AI for resolving ethical dilemmas and actively contribute to training AI models to help improve their efficacy. You have conducted a risk assessment, publicly published your AI policy, have a complaints process, and are part of a research network or advocacy group which actively promotes AI services.

Am I prepared to use AI?



Picture an AI tool you might be curious to try out in your counselling practice.

Write down what it is.

We are going to put your AI tool through some simple tests to help us think clearly about whether it is good to use or not.

# Test 1

- Do I have suitable training to use this? Can I evidence that this training is sufficient for me to practice well?

# Test 2

- Am I insured to use this AI tool in my client work? What would my insurer's stance on this be, and would I be covered if something went wrong?

# Test 3

- Have I tested this AI tool sufficiently so that I can trust it to work as intended?

# Test 4

- Do I know what might go wrong with the tool? Do I have a plan if something goes wrong while using it?

# Test 5

- Has my client consented to the use of this AI tool, and do they understand the probability of a confidentiality breach occurring?

# Test 6

- If, at some point in future, I am no longer able to use this AI tool, do I have a plan for what I would do instead?

# Test 7

- Have I discussed what I am doing with my supervisor? Do I have an AI policy which details how and why I use it?

# Test 8

- If someone wanted to see how I use AI safely, do I have a rationale or a written risk assessment I can give them which shows my working?

# Practice risk assessment

# Scenario 1.

## *A new AI chatbot on the website.*

- You work with a local charity and are conscious that clients sometimes contact the service on evenings and weekends when nobody is there to respond.
- You've tried running an 'on call' system, but it has been unpopular with staff, and upon reflecting on self care, you decide to drop the on call process to handle new enquiries.
- You've heard about a new AI chatbot service which is free and quite easy to add to your website. The chatbot can speak to clients and automatically book them in for an appointment.

# How would you risk assess this?

Would you even think to?

Would you assume your manager handles that stuff?

# Does this chatbot sound okay to use?

- What information does the chatbot collect, and how does it store it?
- What if the client uses the chatbot for therapy?
- What boundaries might the clients and counsellors need to have protected?
- At what point is the client asked if they give consent to use the AI?
- In what way might the chatbot unintentionally harm the client, the organisation, or wider society?

“Eh, dunno, I’ll ask BACP if I’m allowed to use it....but they might not like it, so maybe I shouldn’t ask them?”

An AI innovator’s dilemma

# Taking individual responsibility for our practice

## Be a force for good.

- Don't assume someone else has properly thought this through, ask for a risk assessment.
- Remember that you live in a world with very persuasive AI tools which may have been used to market the AI tool you considering using.
- Remember, you practice in a country with relatively little legislation or authorities which can approve/disprove your actions.
- But we really do want to help people.
- Using AI is not evil or bad, but we need to empower ourselves to use it well.

# There is no one standard risk assessment

You need to make your own

- As counsellors we deeply care about people, their emotions, their stories and greatly value confidentiality.
- We routinely work with ethical dilemmas.
- We deeply care about the authenticity of human connection.
- We are the first generation of counsellors with access to AI tools, the standards we set ourselves now will influence future use substantially.

# How to bake a risk assessment

## Suggested ingredients

- Create a list of possible harms to clients, counsellors, the organisation and wider society.
- A list of remedies - what actions do you take to mitigate any risk.
- Include a process which attempts to track how likely a risk is.
- Create a list of possible benefits to clients, counsellors, the organisation and wide society.
- A section reflecting on confidentiality, data protection and clients rights.



# Our chatbot - brief risk assessment

- Client risks: data privacy, inappropriate referrals, inaccurate information being shared, inappropriate/pseudo therapy dialogues being offered by chatbot. Absence of consent.
- Counsellor risks: Clients being automatically booked in at short notice, inaccurate, inappropriate or misleading information being presented to counsellor by chatbot.
- Likelihood. Most likely problems? Clients using the bot to book in multiple sessions, clients engaging in extended chat with realistic AI.
- Benefits: clients can book themselves in easily. Counsellors can enjoy time off at the weekend.
- Actions to mediate - If client doesn't want to use chatbot, offer alternative. Before chatting with bot, seek informed client consent. Give counsellors the right to consent to appointments before finalising the booking. Create and publish AI policy on website, include appropriate usage guidelines for clients, create complaints process which clients can use if needed.

# Making AI use safe - a check list

- Different counsellors may identify different risks, benefits and different actions.
- The more the merrier.
- If you have colleagues, invite them to 'sense check' your risk assessment.
- Discuss the use of the AI tool with your supervisor.
- Check with your insurer if you are covered for its use.
- Have a plan for what to do if it goes wrong, or you can no longer use it.

# Just to be clear

- A risk assessment can be a hand written list.
- It can be an excel spreadsheet.
- The depth of the risk assessment should reflect the level of risk involved. The higher the level of risk, the more carefully it should be assessed.
- I am not generally a fan of extensive risk assessments. I would not advocate for them in this situation unless I believed it could truly protect client wellbeing and trust in our profession.
- How might you risk assess the following scenarios?

# Scenario 2.

## *The high-tech AI avatar*

- You work in private practice and have been struggling to keep up with demand. You find it stressful to manage a waiting list.
- You want to take on more clients because you worry about the wellbeing of the people you can't see.
- In order to see more clients, you decide to create an AI avatar version of yourself which clients can access, even when you are busy.
- The AI avatar clones your voice and likeness accurately.
- To act like you, it needs to be trained on recordings of any old counselling sessions you have.

# Scenario 3

## *The new miracle platform we've all been waiting for*

- A busy IAPT service has been trying to find better ways to manage its admin processes.
- It set up a customer relations management software system 5 years ago, but it is clunky and unpopular with the counsellors.
- They receive an advert for a 'Gold Standard' AI powered online counselling platform which automates nearly all admin processes. It's being used by other health boards.
- It can conduct initial assessments, offer diagnoses, and give feedback to counsellors based on recordings of sessions.
- It can automatically create transcripts of sessions to give to the clients.
- It can instantly generate custom homework for students attending the service.
- Research suggests it is 99% reliable....

# The future is here

**The scenarios are already possible using today's technology.**

- The general goal is to maximise the good that an AI can do while minimising the harm it may do.
- Once we've created a risk assessment, we should review it annually.
- The AI tool you are using is likely to be cloud-based and liable to having updates which include new features you'll need to account for.
- Challenge colleagues who use AI, check if they have discussed it with their supervisor or if they have conducted a risk assessment?

# The Future

# Suggestions for trainees

- Disclose AI use if you use it to proof read, or entirely write your assignments.
- Ask your training body if they have an AI policy.
- Read sources before you cite them. Automatically generated citations may be incorrect.
- Don't assume your trainers know more about ethical AI use than you do.

# Suggestions for counsellors

- Don't type in, or upload, any identifying or sensitive information to an AI.
- Check if your insurance covers you for using AI tools, even if your employer expects you to use it as part of your job.
- Don't rely on lazy tickbox waivers. If you use AI tools with clients, inform them about what they are consenting to.
- Don't automatically defer to AI's recommendations for your practice. Trust your own judgement and use supervision well.
- Test new tools before using them with clients.

# Suggestions for supervisors

- Your supervisees might use AI without thinking about it carefully, so ask them.
- If they are using AI, ask to see their risk assessment, or offer to do one with them.
- Don't assume that a knowledgeable tech savvy supervisee has it all figured out.
- Actively ask your supervisees if they are using AI and encourage them to reflect on how it may effect their practice.
- Reflect on how vulnerable clients may be effected by AI.

# Suggestions for trainers

- Don't use AI to entirely generate your slides, at least edit them carefully.
- If you do use AI to generate ideas or content, disclose this publicly or while presenting it.
- Reflect carefully on how AI may skew your thinking by inadvertently reinforcing biases, or perpetuating falsehoods.
- Create an AI policy which covers how assessments are to be managed.
- Consider how AI tools could benefit learners fairly.

# Suggestions for policy makers

- Perform assessments of the quality of data an AI tool is trained on.
- Do not solely use other AI tools to assess other AI tools.
- Create and publish AI policies if you use AI in your organisation.
- Create a whistleblowing system where staff can raise concerns about harmful AI use.
- Create a transparent complaints process for AI related issues.

# Suggestions for everyone

## Do your own risk assessments!

- Keep up to date with CPD.
- Look for Good Practice guides
- Consider cryptography or data security training
- Follow me on linked in/twitter for some reasonably priced, but eminently valuable content....
- Be generally wary of the first generation of AI products, experts and advocates, including me.
- Look for robust research, there is currently a shortage.
- Conduct research!

# Getting excited for AI

## Becoming proud of our ethical standards

- By doing the tough reflective work now, we can influence the future of our profession for the better.
- Some AI tools could be really transformative.
- New approaches to therapy might freshen up our industry.
- New research methods could help find good answers to challenges.
- You are among the first people in the country to receive training like this.
- Go boast to your colleagues about how brave you've been!

# The 8 tests were

## A reminder incase you missed them

1. Do I have suitable training to use this? Can I evidence that this training is sufficient for me to practice well?
2. Am I insured to use this AI tool in my client work? What would my insurer's stance on this be, and would I be covered if something went wrong?
3. Have I tested this AI tool sufficiently so that I can trust it to work as intended?
4. Do I know what might go wrong with the tool? Do I have a plan if something goes wrong while using it?
5. Has my client consented to the use of this AI tool, and do they understand the probability of a confidentiality breach occurring?
6. If, at some point in future, I am no longer able to use this AI tool, do I have a plan for what I would do instead?
7. Have I discussed what I am doing with my supervisor? Do I have an AI policy which details how and why I use it?
8. If someone wanted to see how I use AI safely, do I have a rationale or a written risk assessment I can give them which shows my working?

If you remember only one thing  
from today....

# 4 things you can do

- Do not use AI in therapeutic contexts unless you have training, insurance and informed client consent.
- If you decide to use AI in therapeutic contexts, take responsibility for its use. Do a risk assessment and test it before using it with your clients.
- Even if you do a risk assessment, be responsible and have a plan for what to do if it goes wrong, the client withdraws consent, or you can no longer use the AI tool.
- Don't upload identifying client information to an AI system.

Take a moment to look over the questions you've noted down today.

Is there anything else you'd like to ask?

Thank you for your attention.  
Well done for reaching the end.

# Further reading

Pierce, J. (1980). *An Introduction to Information Theory: Symbols, Signals and Noise*. Dover Publishers.

Minsky, M. (2007). *The emotion machine: Commonsense thinking, artificial intelligence, and the future of the human mind*. Simon and Schuster.

Kahneman, D. (2011). Thinking, fast and slow. *Farrar, Straus and Giroux*.

Bostrom, N. (2014) *Superintelligence: Paths, Dangers, Strategies*. Oxford: Oxford University Press.

Christian, B. (2021). *The alignment problem: How can machines learn human values?*. Atlantic Books.

Galef, J. (2021). *The scout mindset: Why some people see things clearly and others don't*. Penguin.

Aschenbrennar, L. (2024) Situational Awareness: The Decade Ahead. For our posterity. [www.forourposterity.com/situational-awareness-the-decade-ahead](http://www.forourposterity.com/situational-awareness-the-decade-ahead)

Chivers, T. (2024). *Everything is Predictable: How Bayesian Statistics Explain Our World*. Simon and Schuster.

Freeman, J. (2024). *Provide or Punish? Students' Views on Generative AI in Higher Education*. HEPI Policy Note 51. *Higher Education Policy Institute*.

Hogarth, I. (2024) Fourth Progress report. AI safety Institute. <https://www.aisi.gov.uk/work/fourth-progress-report>

EU AI Act - <https://tinyurl.com/2r67rzmd>

Scarfe, P., Watcham, K., Clarke, A., & Roesch, E. (2024). A real-world test of artificial intelligence infiltration of a university examinations system: A “Turing Test” case study. *PloS one*, 19(6), e0305354. Scarfe, P., Watcham, K., Clarke, A., & Roesch, E. (2024). A real-world test of artificial intelligence infiltration of a university examinations system: A “Turing Test” case study. *PloS one*, 19(6), e0305354.

Sufyan, N. S., Fadhel, F. H., Alkhathami, S. S., & Mukhadi, J. Y. (2024). Artificial intelligence and social intelligence: preliminary comparison study between AI models and psychologists. *Frontiers in Psychology*, 15, 1353022.

Questions? 30 minutes.