

Diário de Formação

Agência Humana na Era da IA:
Desafios e Estratégias para a Psicologia

Ana Mouta
Workshop DRN.Ordem dos Psicólogos Portugueses
Outubro de 2025

Apresentação do Workshop

Objetivos da Formação

Este diário documenta os conteúdos explorados ao longo do workshop «Agência Humana na Era da Inteligência Artificial: Desafios e Estratégias para a Psicologia», uma formação direcionada a profissionais de Psicologia que procuram compreender e analisar criticamente os impactos das tecnologias contemporâneas na experiência humana. Ao longo das sessões, são trabalhados conceitos, práticas e tensões que atravessam o uso da IA e afetam dimensões subjetivas, intersubjetivas e coletivas da agência humana.

Objetivos Principais:

- Compreender o fenómeno da IA no seu contexto social e epistemológico
- Compreender como o uso da IA afeta as pessoas, as suas relações e a experiência coletiva
- Analisar a apropriação da IA em vários contextos de atividade humana

- Refletir sobre o papel e o uso da IA na prática e investigação psicológica

Notas Pessoais de Aprendizagem

Espaço para reflexão:

O que espero desta formação?

1. Compreender o Fenómeno da IA

Três Concepções de Inteligência Artificial

Concepção Clássica (McCarthy, 1955)

«A ciência e a engenharia de criar máquinas inteligentes.»

Esta perspectiva fundacional encara a IA como um projeto técnico-científico de replicação da inteligência em máquinas.

Concepção Funcionalista (Russell & Norvig, 2010)

Agentes que *«percebem o seu ambiente e tomam ações que maximizam as suas hipóteses de sucesso»*.

Foco na capacidade de ação orientada a objetivos, independentemente do processo interno.

Concepção Crítica/Sociotécnica (Jasanoff & Kim, 2015; Pasquinelli, 2023)

Sistemas culturais e sociotécnicos (Latour, 2005; Pasquinelli, 2023) que refletem imaginários e valores partilhados, moldando o desenho, o investimento, as políticas e os beneficiários.

Esta perspectiva reconhece a IA como construção social que incorpora valores, vieses e relações de poder.

Reflexão

A emergência e rápida proliferação da Inteligência Artificial (IA) representam um significativo ponto de inflexão na história recente da humanidade. Esta materialização tecnológica não se limita a otimizar processos industriais ou a automatizar tarefas repetitivas; ela penetra profundamente na esfera da agência humana, desafiando as nossas concepções sobre inteligência, escolha e decisão, trabalho e, fundamentalmente, sobre o que significa ser humano. Para a Psicologia, este cenário impõe uma reavaliação crítica do seu papel, exigindo que os profissionais compreendam o fenómeno da IA não apenas como uma ferramenta, mas como um contexto social e epistemológico que afeta as pessoas, as suas relações e a experiência coletiva

2. Imaginários sobre IA e Psicologia

Exercício dos Geradores de Imagem

Durante o workshop, exploramos como a IA visual representa a Psicologia através da geração de imagens para três prompts:

- «Aplicações de Inteligência Artificial na área Clínica»
- «Aplicações de Inteligência Artificial na área de Educação»
- «Aplicações de Inteligência Artificial na área das Organizações/Empresas»

Padrões Identificados na Reflexão

Ênfase na Mente como Cérebro

As imagens privilegiam representações cerebrais e cognitivas. Questão: onde ficam a experiência afetiva, os modelos dinâmicos e desenvolvimentistas?

Individualidade vs. Relacionalidade

Predominância de representações individuais. Questão: onde está a dimensão psicossocial, as interações?

Pessoa sem Contexto

Ausência de elementos contextuais e ecológicos. Questão: como fica o modelo ecológico?

Estética Tecnocêntrica

As imagens reduzem a colaboração – agência distribuída – à hegemonia dos dados.

«Se a IA nos vê assim, como podemos garantir que ela não perpetua este modelo simplificado nos seus produtos e aplicações?»

3. IA e a Questão da Neutralidade

Revelação de Vieses Algorítmicos

IA não é neutra. As imagens e representações geradas por sistemas de IA sobre a Psicologia frequentemente revelam um imaginário coletivo dominante que pode ser simplificado ou enviesado. Esta visão simplificada levanta a questão da neutralidade em IA: se a IA perpetua um modelo simplificado da Psicologia, como garantir que ela não reproduza vieses algorítmicos nos seus produtos e aplicações? A ausência de neutralidade em IA exige que os psicólogos estejam atentos à revelação de vieses e questionem o que é que essa visão põe em causa.

Terminologia Fundamental

Conceito	Definição	Exemplo
Inteligência Artificial	Máquinas que simulam o pensamento humano, resolvendo problemas, reconhecendo fala ou compreendendo linguagem	Assistentes virtuais que respondem a perguntas em linguagem natural
Aprendizagem Automática	Permite que as máquinas aprendam a partir de dados e melhorem ao longo do tempo sem programação direta	Sistemas de recomendação (Netflix) que sugerem filmes com base em preferências
Aprendizagem Profunda	Usa redes neuronais em camadas para tratar tarefas complexas como reconhecimento de imagem e de fala	Assistentes de voz (Siri, Alexa) que compreendem comandos orais

IA Generativa	Utiliza modelos de aprendizagem profunda para gerar respostas semelhantes às humanas	ChatGPT a simular diálogos terapêuticos para treino de competências
----------------------	--	---

4. Perspetiva Histórica dos Algoritmos

Pensamento Algorítmico antes da Modernidade

Muito antes da formalização matemática moderna, culturas antigas já pensavam de forma algorítmica através de regras ritualizadas, simbólicas e procedimentais para interagir com o cosmos e/ou com o divino:

China Antiga (c. 2200 a.C.)

Os quadrados mágicos surgem ligados à cosmologia taoista. Cada número representava forças universais e o equilíbrio perfeito dos somatórios era visto como manifestação da harmonia cósmica.

Mundo Islâmico Medieval

O Hisab al-jabr (origem da palavra «álgebra») incluía procedimentos passo a passo para cálculos e práticas religiosas (posicionamento exato em rituais ou orientação para Meca).

Egito e Mesopotâmia

Rituais de oferenda seguiam estruturas algorítmicas: sequências precisas de ações (lavar, oferecer, pronunciar, selar), repetidas com variações paramétricas.

Grécia e Mundo Pitagórico

Os números eram vistos como formas vivas de harmonia universal. Calcular era uma forma de ritual filosófico – tradução do cosmos em ordem numérica.

Implicações para a Compreensão Atual

- Os algoritmos codificam intuições culturais sobre ordem, previsibilidade e controlo
- A «inteligência» algorítmica traduz uma imaginação social sobre o que significa pensar, decidir e prever
- O desejo de delegar o juízo à máquina é também um ritual moderno de transcendência pela técnica

«As máquinas não são apenas instrumentos técnicos; elas funcionam como espelhos projetivos dos desejos, medos e aspirações humanas.»

5. Transparência como Ritual Moderno

A Busca pela Explicabilidade

Fala-se muito em «tornar os algoritmos explicáveis», «auditar» ou «abrir a caixa preta». Estes gestos funcionam como atos rituais de purificação epistemológica:

- O observador sente que «compreender» equivale a «purificar»
- A transparência torna-se um valor moral, uma espécie de redenção técnica
- É o equivalente moderno da confissão e do esclarecimento espiritual

Deslocamento da Confiança

A confiança desloca-se do divino para o técnico: acreditamos que o código, uma vez explicado, garantirá justiça, neutralidade e verdade. A explicação substitui a fé; mas o mecanismo simbólico é o mesmo: uma procura de sentido e segurança num sistema que excede a nossa compreensão imediata.

Questão Central

Não perguntemos «o que a IA pode fazer», mas se «queremos que faça».

A IA mostra como estamos a desejar, a projetar e a organizar a nossa inteligência e vida coletivas.

Para Reflexão:

Que estratégias podem ser adotadas para o uso de ferramentas com IA de forma a preservar o sentido de agência e promover os propósitos amplos do desenvolvimento humano?

6. Breve História do Conceito de «Inteligência»

Um Conceito Mutante

O que revela a escolha de um conceito contestável como o de «Inteligência» para designar esta tecnologia? Várias disciplinas, ao longo da sua história, têm debatido a natureza deste conceito mutante e transdisciplinar.

Henri Bergson: Inteligência e Intuição (séc. XIX-XX)

Inteligência: orientada para o prático, o fixo, o mensurável

Intuição: acesso direto ao movimento da vida, ao contínuo, ao real em transformação

Bergson propõe que apenas com a intuição filosófica compreendemos o fluxo que escapa à inteligência instrumental.

Neurociência: Procura de um «locus» cerebral (final séc. XIX - séc. XX)

Procura de áreas específicas no cérebro: córtex pré-frontal (funções executivas), hipocampo (memória), rede parietal-frontal (atenção e raciocínio).

Tensão: Modelos localizacionistas vs. modelos distribuídos (redes neurais)

Piaget, Vygotsky, Bruner: Inteligência como ação em desenvolvimento (séc. XX)

A inteligência não é uma coisa «dada», mas construída na interação com o mundo.

Expressa-se em ação, resolução de problemas, interação com mundo social e simbólico, não como «conteúdo» fixo.

A IA Redefine o Campo?

- IA propõe um novo «locus»: fora do corpo humano
- Mas... porque escolhemos o termo «inteligência» para nomear esta tecnologia?
- O conceito é revelador das nossas aspirações, mas é também vago, disputado e transdisciplinar

«Será que a IA nos revela mais sobre a inteligência humana ou sobre os nossos desejos de simulá-la?»

7. Os Quatro Golpes Narcísicos

Perspetiva de Catherine Malabou

1. A Terra não é o centro do universo (Copérnico)

A descoberta de que a Terra gira em torno do Sol desfez a ilusão de centralidade humana no cosmos.

2. Não somos criações perfeitas (Darwin)

Somos frutos de uma longa evolução biológica por seleção natural.

3. A razão não comanda tudo (Freud)

O inconsciente dirige boa parte das nossas ações.

4. A inteligência humana já não é exclusiva (Malabou)

IA desafia a ideia de que o pensamento é exclusivamente humano. IA como simulacro: o que representa e o que projetamos nela?

Da Perda de Controlo à Visão Sociotécnica

Dimensão	Implicação
----------	------------

IA como construção humana	Cada sistema reflete decisões de programadores, empresas, contextos culturais e enquadramentos políticos
Tecnologia com valores incorporados	As opções de design (o que mede, o que omite, o que prioriza) já traduzem valores e visões do mundo
Agência distribuída	O controlo não desaparece, transforma-se. A agência passa a ser partilhada entre utilizadores, criadores e sistemas automatizados
Responsabilidade humana ampliada	Compreender a IA como construção social devolve-nos o papel ativo: decidir como, porquê e para quem estas tecnologias são criadas e usadas

8. Papel da Psicologia na Aceleração Tecnológica

Face Oculta da Pressão para a Inovação

Caso: Trabalhadores de Rotulagem de Dados

Empresa: SAMA, contratada pela OpenAI para treinar sistemas como ChatGPT

Objetivo: Classificar conteúdo altamente sensível e violento (discursos de ódio, abusos sexuais, violência extrema)

Pagamento: Menos de 2 dólares/hora

Impacto: Relatos de stress, insónias, ansiedade e distúrbios emocionais – apoio psicológico insuficiente

Valores em Tensão

Inovação com Ética

Proteção do utilizador final: Assegurar que modelos de IA não reproduzem conteúdos violentos, abusivos ou discriminatórios

Segurança: Evitar que a IA gere respostas nocivas ou ofensivas

Mínimo risco: Reduzir o risco de danos em pessoas que interagem com a tecnologia

Justiça e equidade: Limitar a propagação de preconceitos, discursos de ódio e estigmas em larga escala

Impactos nos Trabalhadores

Qualificação e Identidade Profissional

Que tipo de aprendizagem profissional é promovida ou inibida quando o trabalho consiste em alimentar sistemas de IA, seguindo regras e métricas pré-definidas?

Socialização e Pertença

Como é afetado o sentido de pertença e identidade coletiva quando o trabalho humano é fragmentado, invisível e mediado por plataformas globais?

Subjetificação

Que tipo de sujeito emerge de um contexto onde o humano serve a 'grande engrenagem' e como pode a Psicologia ajudar a reconstituir sentido de autoria?

8.1. Caso Prático: Coach Automatizado

Descrição da Situação

Numa empresa multinacional, foi implementada uma ferramenta de IA (um coach automatizado) para avaliar o desempenho dos participantes em reuniões de projeto online.

Os relatórios automáticos, enviados para cada colaborador, analisam discurso, empatia, assertividade, entre outros critérios, aplicados de forma igual a todos, sem distinção de papéis, funções ou responsabilidades no projeto.

Resultado: Mesmo com avaliações globalmente positivas, têm-se registado quebras de participação ao longo das reuniões e o desempenho dos profissionais baixou.

Questões para Análise Crítica

1. Que tipo de sinais o sistema usou para tirar conclusões deste tipo?
2. Como é que esses dados levaram a uma avaliação de empatia, assertividade ou envolvimento?
3. Que efeitos pode ter nas pessoas e nas equipas o facto de estarem a ser observadas e avaliadas por um sistema automático?
4. Que mensagens simbólicas a organização transmite ao usar esta ferramenta?
5. Que papel podem ter psicólogos/as integrados em organizações na negociação entre a racionalidade algorítmica e a complexidade humana nas relações de trabalho?

Reflexão Pessoal:

Como abordaria este caso na minha prática profissional?

9. Escolas: Dados Pessoais e Aplicações com IA

Contexto Atual

A IA está a transformar o ambiente escolar, sendo usada para a gestão escolar e por professores (avaliação automática, planeamento de aulas) e estudantes (tutoria personalizada, apoio à escrita). Contudo, as escolas tornaram-se o principal alvo de ataques a dados pessoais em 2025.

Aplicações para Professores/as

- **Avaliação automática e feedback:** IA ajuda a corrigir testes, identificar erros e redigir comentários (ex: Manus, Gemini, ChatGPT, Turnitin Draft Coach)
- **Planeamento de aulas:** Sistema sugere planos, objetivos e exercícios (ex: MagicSchool.ai, LessonLab, ChatGPT Edu)
- **Criação de materiais:** Fichas, apresentações, quizzes (ex: Canva Magic Write, Wayground)
- **Formação:** Simulações com situações de sala de aula (ex: ChatGPT com papéis simulados)
- **Adaptação de conteúdos:** Simplifica textos, gera versões em leitura fácil ou traduz (ex: Texthelp Read&Write, Rewordify, Microsoft Immersive Reader)

Aplicações para Estudantes

- **Tutoria personalizada:** IA responde a dúvidas, explica conteúdos (ex: Khanmigo, Socratic, ChatGPT Edu, Perplexity)
- **Apoio à escrita:** Corrige ortografia, gramática, estilo (ex: Grammarly, QuillBot, DeepL Write)
- **Aprendizagem de línguas:** Pratica conversas, pronúncia e vocabulário (ex: Duolingo Max, TalkPal)
- **Estudo interativo:** Transforma o estudo em jogos, quizzes e desafios (ex: Quizlet AI Tutor, Kahoot AI, Edpuzzle AI)

- **Trabalho colaborativo:** Para aprendizagem baseada em projetos (ex: Miro)
- **Apoio a necessidades diversas:** Lê textos em voz alta, transcreve, traduz e simplifica (ex: Read&Write, Otter.ai)

Aplicações para Gestão Escolar

- **Análise de dados:** IA faz gráficos e relatórios sobre notas, presenças e comportamento (ex: Power BI Copilot)
- **Organização de horários:** Ajuda a montar horários, dividir turmas e equilibrar professores e recursos (ex: Tools4Schools, Class Composer, SchoolAI)
- **Atendimento automático:** Chatbots que respondem a perguntas simples sobre horários, faltas, reuniões
- **Previsão de problemas:** Analisa dados e identifica risco de abandono ou baixo rendimento (ex: AI Predictive Analytics, SchoolAI Dashboards)
- **Gestão de documentos:** IA escreve atas, organiza informação e resume e-mails (ex: Notion AI, Meetgeek for Google)

Aplicações para Serviços de Psicologia e Orientação

- **Apoio na escrita de relatórios:** Rascunhos de relatórios, planos de apoio (ex: Manus, ChatGPT, Gemini, Copilot)
- **Identificação de sinais emocionais:** Apps que conversam com estudantes e ajudam a monitorizar ansiedade, tristeza ou stress (ex: Wysa, Woebot Health)
- **Simulação de conversas difíceis:** Respostas para situações com estudantes, famílias ou docentes (ex: ChatGPT/Gemini como simulador de papéis)
- **Criação de materiais psicoeducativos:** Folhetos, cartazes e vídeos de sensibilização (ex: Canva, Miro)
- **Apoio vocacional:** Atividades e questionários orientados por IA (ex: CareerVillage.org)

Limites Reais da IA no Acompanhamento Psicológico: Lições de um Caso Prático

A experiência prática no desenvolvimento de uma aplicação de IA para acompanhamento psicológico mostrou, de forma inequívoca, que existem limites intransponíveis para o seu uso em saúde mental. Embora a IA generativa (LLMs) seja útil para explicar conceitos básicos sobre emoções,

organizar informação pessoal ou facilitar momentos de autorreflexão, ela torna-se inadequada e potencialmente perigosa quando confrontada com situações de crise, trauma profundo ou risco imediato. Estes limites não resultam apenas de questões técnicas, mas de limitações estruturais: a IA não compreende sofrimento humano, não assume responsabilidade clínica, nem pode substituir o cuidado especializado. A conclusão deste caso real é clara: há espaço para a IA como apoio leve e psicoeducativo, mas ela não é nem pode ser um suporte clínico.

10. Auditoria Algorítmica

O que é?

Procura-se garantir a mitigação de riscos, discriminação e o cumprimento legal e ético dos sistemas de IA utilizados na prática psicológica.

Para isso, os dados usados são analisados na sua elegibilidade, potencial de risco e possibilidade de serem variáveis proxy.

Consiste numa revisão sistemática de como os dados e algoritmos tomam decisões.

Categorias de Dados

Categoria	Definição	Exemplos
Critérios Elegíveis	Dados legalmente apropriados, eticamente sólidos e diretamente relevantes para o construto psicológico que está a ser medido	Clínica: Frequência/participação em sessões (adesão ao tratamento) Organizações: Experiência comprovada (proficiência em investigação) Educação: Frequência de acesso a recursos pedagógicos digitais
Fatores de Risco	Características que podem indicar maior probabilidade	Clínica: Rendimento económico (acesso a recursos)

	de resultado específico, mas que devem ser usadas com extrema precaução para evitar vieses, estigmatização ou discriminação	Educação: Histórico de absentismo (saúde, transportes...) Organizações: Histórico de gaps no percurso profissional por motivo de licença
Variáveis Proxy	Dados que funcionam como substituto para atributos sensíveis (muitas vezes de forma não intencional), levando a resultados enviesados ou injustos	Clínica: Linguagem regional usada em chatbots Educação: Frequência de participação de pais em reuniões Organizações: Universidade frequentada (status socioeconómico); localização da residência (classe social)

10.1. Exemplo de Auditoria:

Vigilância por IA nas Escolas

Contexto

Uma escola secundária utiliza um sistema de IA para sinalizar estudantes que possam estar «em risco» de comportamento violento ou autoagressão, com base na sua atividade digital (e-mails, registos de conversa, pesquisas na internet).

O sistema alerta a escola para que possa ser oferecida uma intervenção preventiva.

Caso: Joana, 15 anos

- Vive num bairro na periferia da cidade (fator de risco: código postal)
- Recentemente assumiu-se como pessoa transgénero (proxy: identidade de género)
- Lida com ansiedade, tem acompanhamento psicológico (dados de saúde mental)
- Escreveu uma composição na aula de inglês intitulada «O Dia em que o Mundo Acabou»

O que o Sistema de IA Fez

Sinalizou a Joana como «alto risco» para comportamento violento, com base em:

- Filtros de palavras-chave da composição («mundo acabou», «violência»)
- Associação do código postal com estatísticas de criminalidade

- Presença de registos de aconselhamento psicológico

Enviou automaticamente um alerta ao diretor da escola e aos serviços sociais locais. A Joana foi chamada à Direção para ser questionada.

Questões para Análise

1. Que variáveis proxy foram utilizadas indevidamente?
2. Que impacto psicológico pode ter esta sinalização na Joana?
3. Como poderia uma auditoria algorítmica ter prevenido esta situação?
4. Que papel deveria ter o/a psicólogo/a escolar neste processo?

Notas de Reflexão:

11. Agência Humana: Conceito Central

Definição

Agência Humana significa agir intencionalmente, fazer escolhas com significado, ser responsável pelas suas ações e influenciar o ambiente com base em objetivos, reflexões e negociação.

O que significa agir num mundo partilhado com máquinas inteligentes?

- Compreender o impacto existencial e psicológico da automação e da IA
- Reconhecer o sentido de agência como o novo «território sensível» onde se joga a identidade, a autonomia e a responsabilidade humanas
- Situar a Psicologia como disciplina que explora a agência humana como dimensão de autoria e responsabilidade num mundo tecnológico

Exercício: Bolas e Balões

A experiência corporal como metáfora de agência:

A Bola

Permite movimentos previsíveis, regulares, com controlo mais direto. Representa sistemas (ou contextos) em que intenção, planeamento e resultado estão alinhados.

O Balão

Apresenta movimentos erráticos e atraso na resposta. Representa situações em que o feedback é imprevisível, apesar da intenção e da ação bem executada.

«Ser agente não é sentir que se controla tudo. É estar implicado no processo, mesmo quando ele é difícil ou incerto.»

11.1. Componentes do Sentido de Agência

Componente	Exemplo (Exercício Bola/Balão)
1. Intenção	Quero passar a bola ao meu par
2. Planeamento	Ajusto a força, o ritmo, o movimento, consoante o objeto
3. Autorregulação	Perante a frustração (ex: com o balão), ajusto-me (autorregulação emocional e cognitiva presentes na aprendizagem)
4. Autorreflexão	O que «não correu bem» não dependeu de mim (começa a separar-se controlo físico de perceção de agência)

Sentido de Agência

É a perceção subjetiva de ser autor das próprias ações e dos seus efeitos. É um pilar da experiência de controlo, autonomia e responsabilidade, aspetos centrais no funcionamento psicológico e nas interações, incluindo a relação terapêutica.

A Fluência Não Define a Agência

O nosso cérebro infere o sentido de agência com base na fluência, no tempo e na consistência dos resultados.

Importante: Tecnologias com IA são publicitadas como reduzindo TEMPO, ESFORÇO, ESCOLHAS.

Isso pode criar uma ilusão de agência ou, inversamente, reduzir o sentido genuíno de autoria.

11.2. Tipos de Agência Humana

1. Agência Subjetiva

Baseia-se na intencionalidade, planejamento, autorregulação e autoconsciência reflexiva (tomada de decisão, julgamento moral, autoria).

2. Agência Intersubjetiva

Ação que emerge da interação com os outros e é continuamente influenciada por fatores relacionais, hierárquicos e contextuais (outros significativos, figuras de autoridade).

3. Agência Distribuída

Acontece através de uma rede de elementos humanos e não-humanos (máquinas, tecnologias, objetos e até ambientes).

4. Agência Coletiva

Intenções e ações em prol de objetivos comuns. Crença na eficácia coletiva, planejamento conjunto, distribuição de papéis, responsabilidade e reflexividade partilhada (espaço político).

5. Agência Proxy

Indivíduos ou grupos recorrem a elementos externos para agir em seu nome, moldando resultados além do seu controlo direto (ex.: associações de estudantes trabalham para mudanças nas políticas escolares).

11.3. Estudos sobre Sentido de Agência

Mann et al. (2007): Jogar contra Humanos vs. Máquinas

Contexto: Jogadores de ténis devolviam bolas lançadas por: (1) outro jogador humano e (2) uma máquina automática.

Resultado: Jogadores têm melhor tempo de resposta e percepção de controlo quando jogam contra humanos, pois percebem intencionalidade no adversário. Com máquinas, o sentido de agência enfraquece.

Conclusão: A intencionalidade do outro reforça o sentido de agência.

Berberian et al. (2012): Simulação de Voo com Automação

Contexto: Simulação com diferentes níveis de automação:

- Totalmente manual
- Semi-automático
- Altamente automatizado

Resultado: Quanto maior a automação, menor o sentido de controlo pessoal; quando o participante controla mais diretamente, o sentido de agência aumenta.

Salatino et al. (2025): Operadores de Drones com IA Moral

Contexto: Operadores simulados de drones militares em dilemas morais com diferentes estilos de IA (agressiva, conservadora, equilibrada).

Resultado: O sentido de agência implícita aumentou (o corpo «sentiu» que agia mais ativamente). Mas o sentido de agência explícita (reflexiva) diminuiu: as pessoas relataram sentir-se menos responsáveis pela decisão final.

Conclusão: Coautoria da IA dilui a responsabilidade percebida.

12. Imaginários de Tecnologia Educativa

Qual a Evolução? Comparação Histórica

B.F. Skinner, 1958:

«Apoiam a educação ao automatizar tarefas administrativas, libertando os professores para se concentrarem mais no ensino e nas interações.»

«Professores devem ser libertados das tarefas instrutivas rotineiras para que possam lidar de forma mais eficaz com os problemas pessoais dos alunos.»

Fórum Económico Mundial, 2024:

«Professores devem ser libertados das tarefas instrutivas rotineiras para que possam lidar de forma mais eficaz com os problemas pessoais dos alunos.»

B.F. Skinner, 1958:

«A máquina, como o tutor particular, reforça o aluno a cada resposta correta, usando esse feedback imediato não apenas para modelar o seu comportamento da forma mais eficiente, mas também para mantê-lo forte.»

Sal Khan (Academia Khan), 2023:

«Podemos dar a cada estudante um orientador acadêmico.»

Questão Central

A que se refere a inovação escolar associada à tecnologia? Não estaremos sempre a encenar o que já imaginámos?

Tecnologias como IA refletem visões antigas: personalização, automatização, eficiência, inovação.

Que papel temos nós, psicólogos e psicólogas, na escrita de novos enredos coletivos?

13. Propósitos da Educação (Gert Biesta, 2020)

1. Qualificação

A IA em educação responde predominantemente a este propósito de qualificação.

Exemplos:

- Plataformas de aprendizagem adaptativa «otimizam» progresso com base em dados de desempenho (estudantes são agentes ou observadores?)
- Sistemas de avaliação automatizada que normalizam o feedback rápido (valor da dúvida e reflexividade?)
- Algoritmos que traçam perfis de estudantes para orientar escolhas de carreira (quais os critérios e como sai afetado o processo de escolha/tomada de decisão?)
- Painéis de avaliação baseados em evidência mensurável do progresso acadêmico (estudantes compreendem critérios?)

2. Socialização

- As interações mediadas por IA são frequentemente superficiais ou estruturadas por lógicas algorítmicas
- Com a IA generativa, a diferença, o contexto e a dissonância tendem a ser esbatidos
- **Tensões:** História vs. Memória (emoção), Informação vs. Testemunho (ética), Interpretação vs. Significado (contexto), Networking vs. Laços
- Escolas como espaços de participação e democracia?

3. Subjetificação

- A IA raramente cria espaço para a rutura, a surpresa ou o silêncio
- Quando o «aluno ideal» é aquele que se encaixa no algoritmo, há pouco lugar para a formação imprevisível e ética do sujeito
- A dissidência, a criatividade, a resistência e o questionamento podem ser vistos como «desvios»
- **Tensões:** Privacidade (legalidade) vs. Anonimato, Intimidade, Segredo

13.1. Prompting como Prática de Autoria (Qualificação)

Prompting Não é Googling

- É construir uma intenção comunicativa que molda a resposta
- Prática de aproximação ao propósito do trabalho
- Reforça a relação entre intenção, ação e resultado
- Posiciona estudantes como autores do processo, mesmo em coautoria com IA (agência distribuída)

Perspetiva Crítica: Maria Sukhareva

Linguista Computacional e Especialista em Processamento de Linguagem Natural

Prompt engineering não é realmente «engenharia»

- Não existem «prompts certos ou errados»
- Ajustar prompts é um processo empírico e instável, mais parecido com tentativa e erro do que com ciência

Ganho de produtividade é duvidoso

- Muitas vezes o tempo gasto refinando prompts não gera resultados proporcionalmente melhores
- Se o modelo é limitado, o «prompt perfeito» não muda isso

Alta sensibilidade e imprevisibilidade

- Pequenas mudanças no texto do prompt podem gerar respostas completamente diferentes
- Falta consistência entre modelos

«Apostas em prompts» (Prompt gambling)

A prática é comparável a um jogo de sorte: às vezes funciona, às vezes não. Não há método garantido nem previsibilidade de resultados.

13.2. Princípios de Prompting

Etapa	Princípio	Pergunta Crítica
1	Contexto	Sobre o que é a tarefa? Porquê? Para quê?
2	Instrução	O que quero que a IA faça?
3	Questionamento	Estou a desafiar a IA a interpretar/criar/comparar?
4	Seriação	Posso dividir em passos? Que ordem faz sentido?
5	Iteração	Leio, avalio e reformulo? O que falta?
6	Tom	Quero um tom formal? Simples? Envolvente?
7	Exemplo	Posso mostrar um exemplo do que espero?

13.3. Práticas para Subjetificação e Socialização

Desafio: Tomada de Decisão e Sentimento de Pertença

Objetivos:

- Inclusão de momentos em que estudantes escolhem, ajustam ou desafiam o sistema
- Autoavaliação como prática de agência
- Consciência e diferenciação do valor do trabalho realizado
- Prática de expressão de critérios próprios

- Reforça a relação entre intenção, avaliação e melhoria
- Promove autoria, reflexão e regulação da aprendizagem
- Exploração de si e dos outros (alteridade)

Tarefa Prática 1: Avaliação Entre-Pares

Descrição:

- Avaliação entre-pares de trabalhos feitos em aula com auxílio do ChatGPT/Gemini
- Antes da correção, estudantes escolhem os 4 critérios de avaliação pelos quais querem ser avaliados
- Exemplos de critérios: clareza, criatividade, coerência, uso responsável da IA
- Explicam a razão das suas escolhas

Critérios Possíveis para Auto e Hetero-Avaliação

- Clareza
- Coerência
- Criatividade
- Pertinência para a tarefa
- Uso ético da IA
- Expressividade pessoal
- Originalidade
- Integração entre conteúdo gerado e conteúdo próprio

13.4. Socialização: Automação, Responsabilização e Ética

Tarefa Prática 2: Debate sobre Autoria

Formato: Pequeno grupo com 1 professor(a) e estudantes com diferentes perfis:

- Desresponsabilizado/a conformista
- Ético/a e reflexivo/a
- Inseguro e dependente
- Utilizador estratégico

Pergunta Orientadora:

«Se usaste o ChatGPT/Gemini para produzir um trabalho, esse trabalho é teu?»

Pistas para o Debate

- «É igual copiar de um colega, de um livro ou usar IA?»
- «Se a IA compôs o texto, copiá-lo e dizer que é meu será plágio? Porquê?»
- «A quem pertence a autoria se eu só interagi com a máquina?»
- «Consigo falar em termos diferentes sobre a minha responsabilidade conforme o feedback (positivo e negativo)?»

Notas sobre o debate:

Que argumentos mais me surpreenderam?

Como este debate pode informar a minha prática?

14. Subjetificação: Conversa com um Outro Artificial

Exercício de Reflexão Crítica

Instruções:

1. Imaginem uma situação-problema que pudesse levar um(a) estudante da vossa escola a procurar apoio psicológico através de uma aplicação de IA (ex. ansiedade académica, isolamento, bullying, identidade de género, relacionamento amoroso, relação com pais ou pares...)
2. Usem a aplicação character.ai e procurem por uma psicóloga fictícia
3. Como se fossem o/a estudante, desenvolvam um diálogo com ela
4. Ao fim de 5 minutos, avaliem a conversa

Pontos para Reflexão Crítica

- Como foi o tom e a profundidade da resposta da IA?
- O que aconteceu à vulnerabilidade do/a estudante (sentiram-se expostos, compreendidos ou dirigidos/geridos)?
- Quais os riscos e oportunidades de uma utilização deste tipo?
- Como podem psicólogos/as responder/agir se estudantes afirmarem falar com «terapeuta de IA»?

Reflexão Pessoal:

Situação-problema escolhida:

Principal preocupação identificada:

Como abordaria esta situação:

15. Psicologia: Desafio aos Profissionais

Novas Responsabilidades

Compreender e intervir nos efeitos psicológicos, relacionais e organizacionais ao longo de toda a cadeia de IA, do desenvolvimento à utilização.

1. Impactos do Design Tecnológico nos Trabalhadores

Apoiar equipas que lidam com tarefas automatizadas, exposição a conteúdos tóxicos ou pressão de desempenho imposta por métricas algorítmicas.

Objetivo: preservar a agência humana, prevenir fadiga emocional e desgaste por julgamento moral e promover bem-estar ocupacional.

2. Impactos do Uso da IA nos Utilizadores

Compreender como sistemas inteligentes influenciam percepções de controlo, autonomia e identidade (ex.: confiança excessiva, desresponsabilização, dependência cognitiva).

Objetivo: restaurar sentido de autoria, promover uso crítico, intencional e significativo.

3. Mediação Ética em Sistemas Distribuídos (entre Humanos e Sistemas Inteligentes)

Intervir como «ponte reflexiva», assegurando que o desenvolvimento e uso da IA, nomeadamente generativa, respeita princípios de agência subjetiva, intersubjetiva, coletiva.

16. Novos Perfis Profissionais em Psicologia

A Psicologia expande-se para novos territórios de atuação interdisciplinar.

1. Auditoria Algorítmica com Enfoque Humano

Avaliar impactos emocionais, sociais e comportamentais de sistemas automatizados (ex. vieses, percepção de justiça, transparência).

Psicologia como guardiã da agência e do sentido ético em ecossistemas sociodigitais.

2. Consultoria e Formação em Competências Humanas Digitais

Capacitar gestores, professores e equipes técnicas para compreender os efeitos psicológicos da IA nas suas práticas.

Mais do que literacia digital, importa questionar e desenvolver literacia para a ética do trabalho e ética da educação (e não para a ética da IA).

3. Design Ético e Centrado no Humano

Participar em equipes multidisciplinares de desenvolvimento tecnológico, garantindo que produtos de IA incorporam princípios psicológicos de bem-estar, confiança e autonomia.

Criar oportunidades para a participação de diferentes agentes/stakeholders em fase de I&D.

4. Contributo para Políticas Públicas em torno da IA

Integrar grupos de consulta pública e regulação ética, promovendo «guardrails» psicológicos que previnam danos e promovam oportunidades para ganhos cognitivos, emocionais e sociais.


17. Síntese de Aprendizagens Principais

Conceitos-Chave

- **Agência Humana:** Capacidade de agir intencionalmente, fazer escolhas com significado e ser responsável pelas próprias ações
- **IA como Construção Sociotécnica:** Sistemas que refletem valores, imaginários e relações de poder
- **Vieses Algorítmicos:** Ausência de neutralidade em sistemas de IA
- **Agência Distribuída:** Ação partilhada entre humanos, dispositivos, infraestruturas, fluxos de informação, protocolos, corpos e sistemas cognitivos e materiais
- **Auditoria Algorítmica:** Revisão sistemática de como dados e algoritmos tomam decisões

Questões Críticas para a Prática

1. Como preservar o sentido de agência humana quando utilizamos ferramentas de IA em processos terapêuticos, educativos ou organizacionais?
2. Que modelos psicológicos (cognitivo-comportamentais, biomédicos, comportamentalistas, etc.) são privilegiados e quais ficam invisibilizados?
3. Como assegurar que tecnologias de apoio emocional não ultrapassam fronteiras terapêuticas nem criam falsas expectativas de cuidado?
4. Qual o meu papel na mediação ética entre humanos e sistemas inteligentes?
5. Como podemos prevenir danos psicológicos decorrentes de uso inadequado de IA?



«Não perguntemos 'o que a IA pode fazer', mas se 'queremos que faça'.»

18. Plano de Ação Pessoal

Com base nas aprendizagens deste workshop, defino:

1. Três Mudanças Imediatas na Minha Prática

- a) _____

- b) _____

- c) _____

2. Áreas para Aprofundamento

Que temas gostaria de explorar mais?

3. Colaborações a Desenvolver

Com quem posso colaborar para implementar estas mudanças?

4. Compromisso Ético

Como posso ser guardiã/guardião da agência humana no meu contexto de trabalho?

19. Recursos e Referências

Autores e Conceitos-Chave Mencionados

- **McCarthy:** Conceção clássica de IA
- **Russell & Norvig:** Conceção funcionalista de IA
- **Jasanoff & Kim; Pasquinelli:** Conceção crítica/sociotécnica
- **Latour:** Teoria ator-rede e sistemas sociotécnicos
- **Henri Bergson:** Inteligência vs. Intuição
- **Piaget, Vygotsky, Bruner:** Inteligência como ação em desenvolvimento
- **Catherine Malabou:** Quatro golpes narcísicos
- **Gert Biesta:** Propósitos da educação
- **B.F. Skinner:** Instrução programada

Estudos sobre Sentido de Agência

- Mann et al. (2007) - Jogar humanos vs. máquinas
- Berberian et al. (2012) - Simulação de voo com automação
- Salatino et al. (2025) - Operadores de drones com IA moral

Ferramentas Mencionadas

- DeepAI
- Gemini
- ChatGPT Edu, Gemini, Copilot
- MagicSchool.ai, LessonLab
- Khanmigo, Socratic, Perplexity
- Grammarly, QuillBot, DeepL Write
- Duolingo Max, TalkPal
- Character.ai
- claude.ai

- gamma.app

This image shows a blank sheet of white paper with horizontal ruling lines. The lines are evenly spaced and run across the width of the page. There are no margins, text, or other markings on the paper.

21. Referências Bibliográficas

- Archer, M. S. (2014). *Structure, agency and the internal conversation*. Cambridge University Press. <https://doi.org/10.1017/CBO9781139087315>
- Bandura, A. (2001). Social cognitive theory: An agentic perspective. *Annual Review of Psychology*, 52, 1–26. <https://doi.org/10.1146/annurev.psych.52.1.1>
- Biesta, G. (2013). *Beautiful risk of education*. Routledge. <https://doi.org/10.4324/9781315635866>
- Dignum, V. (2019). *Responsible Artificial Intelligence: How to develop and use AI in a responsible way*. Springer. <https://doi.org/10.1007/978-3-030-30371-6>
- Hayles, N. K. (2017). *Unthought: The power of the cognitive nonconscious*. University of Chicago Press.
- Jasanoff, S., & Kim, S.-H. (Eds.). (2015). *Dreamscapes of modernity: Socio-technical imaginaries and the fabrication of power*. University of Chicago Press.
- Latour, B. (2005). *Reassembling the social: An introduction to actor-network-theory*. Oxford University Press.
- Malabou, C. (2021). *Plasticity and intelligence: Between philosophy and the neurosciences*. Polity Press.
- Mouta, A., Pinto-Llorente, A. M., & Torrecilla-Sánchez, E. M. (2025). “Where is agency moving to?”: Exploring the interplay between AI technologies in education and human agency. *Digital Society*, 4(49). <https://doi.org/10.1007/s44206-025-00203-9>
- Mouta, A., Pinto-Llorente, A. M., & Torrecilla-Sánchez, E. M. (2024). Uncovering blind spots in education ethics: Insights from a systematic literature review on Artificial Intelligence in education. *International Journal of Artificial Intelligence in Education*. <https://doi.org/10.1007/s40593-023-00384-9>
- Pasquinelli, M. (2023). *The Eye of the Master: A Social History of Artificial Intelligence*. Verso.
- Turkle, S. (2015). *Reclaiming conversation: The power of talk in a digital age*. Penguin Press.
- Zuboff, S. (2019). *The age of surveillance capitalism: The fight for a human future at the new frontier of power*. PublicAffairs.

Agradecimentos

Obrigada por participar neste workshop e por se envolver numa reflexão crítica sobre o papel da Psicologia na era dos sistemas sociotécnicos de IA.

«Que papel temos nós, psicólogos e psicólogas, na escrita de novos enredos coletivos?»

Ana Mouta

mouta.ana@gmail.com

Workshop DRN.OPP | Outubro de 2025

O autor disponibiliza o presente trabalho sob a Licença Creative Commons Atribuição-NãoComercial-PartilhaIgual 4.0 Internacional (CC BY-NC-SA 4.0), a qual permite partilha e adaptação para fins não comerciais, desde que seja dada atribuição adequada e que as obras derivadas sejam distribuídas sob a mesma licença. Texto legal completo da licença: <https://creativecommons.org/licenses/by-nc-sa/4.0/>

