

CADERNOS DO

CON TEM PO RÂ NEO

v. 1 / n. 1 / junho / 2025

entrevista exclusiva/
exclusive interview

PIERRE LÉVY



PUCPR
GRUPO MARISTA

EXPEDIENTE

Pontifícia Universidade Católica do Paraná (PUCPR)

Reitor

Ir. Rogério Renato Mateucci

Vice-Reitor

Vidal Martins

Pró-Reitor de Missão, Identidade e Extensão

Fabiano Incerti

Coordenador do Instituto L'Hermitage de Estudos Avançados

Douglas Borges Candido

Cadernos do Contemporâneo é uma publicação que visa o debate de temas prementes da sociedade com renomados intelectuais de diversas áreas do conhecimento. A seleção e a organização do conteúdo estão sob curadoria de Fabiano Incerti e Douglas Borges Candido; a organização da entrevista contou com a colaboração de Kleber Bez Birolo Candiottto e Murilo Karasinski; a tradução é de Eduardo Portanova Barros e a produção editorial é conduzida pela equipe da PUCPRESS.

Edição, preparação e revisão de texto: Clarisse Lye Longhi.

Projeto gráfico e diagramação: Cristina Mosol.

Capa: AdobeStock_1257714729.

DOI 10.7213/cadcontemporaneos.v1n1

PUCPRESS/Editora Universitária Champagnat

Rua Imaculada Conceição, 1155 -
Prédio da Administração - 6º andar
Câmpus Curitiba - CEP 80215-901 -
Curitiba/PR
Tel. +55 (41) 3271-1701
pucpress@pucpr.br

EDITORIAL

Vivemos tempos velozes. Tecnologias cada vez mais complexas moldam o modo como pensamos, sentimos, decidimos, trabalhamos e convivemos. Entre todas as inovações que inquietam e encantam o nosso tempo, poucas provocam tanto espanto e expectativa quanto a Inteligência Artificial. O Papa Francisco se posicionou com lucidez e firmeza sobre a necessidade de orientar o desenvolvimento da Inteligência Artificial à luz da dignidade humana. Em sua mensagem para o Dia Mundial da Paz, ele nos adverte: "A inteligência artificial deve servir à paz e não ao aumento das desigualdades ou da violência". E insiste: "Os algoritmos não podem ser neutros quando os valores em jogo são os da vida, da justiça, da verdade".

Na PUCPR, onde a ciência dialoga com a fé, e a tecnologia é chamada a servir o bem comum, acreditamos que a pergunta central não é o que as máquinas podem fazer, mas o que nós queremos ser enquanto humanidade. A ética que nos guia é aquela que se compromete com o outro, especialmente com os mais frágeis, e que coloca a técnica a serviço da fraternidade e da esperança.

O Instituto L'Hermitage, espaço de pensamento, espiritualidade e formação da Pontifícia Universidade Católica do Paraná, nasce do desejo de fazer do carisma marista uma luz sobre os grandes desafios contemporâneos. É com esse espírito que apresentamos esta primeira edição dos *Cadernos do Contemporâneo*, edição que ensaia indagações que passam por temas desde a ética às culturas digitais, do tempo das telas às questões emocionais. Neste número, conversamos com o filósofo e sociólogo Pierre Lévy, uma referência quando o assunto se trata da cultura e revoluções digitais, e destacamos a contribuição do professor de Filosofia da PUCPR, e, também, especialista da temática, Kleber Candiottto, a respeito das relações humanas na era da Inteligência Digital.

Boa leitura!

Ir. Rogério Renato Mateucci

Reitor da Pontifícia Universidade Católica do Paraná (PUCPR).

CONTEÚDO

P. 4 SINAIS

Fabiano Incerti e Douglas Borges Candido

P. 8 ENTREVISTA

Pierre Lévy

P. 30 ARTIGO

As relações humanas na era da IA Generativa
Kleber Bez Birolo Candiottto

EDITORIAL

We live in fast-paced times. Increasingly complex technologies shape the way we think, feel, decide, work and live together. Among all the innovations that disturb and delight our time, few provoke as much amazement and expectation as Artificial Intelligence. Pope Francis adopted a clear and firm stance on the need to guide the development of Artificial Intelligence in the light of human dignity. In his message for World Peace Day, he warns us: "Artificial intelligence must serve peace and not the increase of inequalities or violence". And he insists: "Algorithms cannot be neutral when the values at stake are those of life, justice and truth".

At PUCPR, where science dialogues with Faith, and technology is called to serve the common good, we believe that the central question is not what machines can do, but what we want to be as humanity. The ethics that guide us are those that are committed to others, especially the most vulnerable, and that place technology at the service of fraternity and hope.

The L'Hermitage Institute, a space for thought, spirituality and education at the Pontifical Catholic University of Paraná, was born from the desire to shed light on the great challenges of our time in the Marist charism. It is with this spirit that we present this first edition of *Cadernos do Contemporâneo*, an issue that explores inquiries ranging from ethics to digital cultures, from screen time to emotional questions. In this edition, we talk to philosopher and sociologist Pierre Lévy, a leading figure when it comes to culture and digital revolutions, and we highlight the contribution of PUCPR Philosophy professor and specialist Kleber Candiottto, regarding human relations in the era of Digital Intelligence.

Enjoy your reading!

Br. Rogério Renato Mateucci

Rector of the Pontifical Catholic University of Paraná (PUCPR).

CREDITS

Pontifical Catholic University of Paraná

Rector

Br. Rogério Renato Mateucci

Vice-Rector

Vidal Martins

Pro-Rector of Mission, Identity and Extension

Fabiano Incerti

Coordinator of the L'Hermitage Institute of Advanced Studies

Douglas Borges Candido

Cadernos do Contemporâneo is a publication that aims to debate pressing issues in society with renowned intellectuals from various fields of knowledge. The selection and organization of the content are curated by Fabiano Incerti and Douglas Borges Candido; the organization of the interview had the collaboration of Kleber Bez Birolo Candiottto and Murilo Karasinski; translation by Eduardo Portanova Barros and the editorial production is led by the PUCPRESS team.

Editing, preparation and proofreading of text: Clarisse Lye Longhi.

Graphic design and layout: Cristina Mosol.

Book cover:

AdobeStock_1257714729.

DOI 10.7213/cadcontemporaneos.v1n1

PUCPRESS

1155 Imaculada Conceição Street
- Administration Building - 6th
floor Campus Curitiba - Zip Code
80215901- Curitiba/Paraná/Brazil
Phone: +55 41 32711701
pucpress@pucpr.br

CONTENT

SIGNS P. 5

Fabiano Incerti and Douglas Borges Candido

INTERVIEW P. 9

Pierre Lévy

ARTICLE P. 31

Human relations in the age of Generative AI
Kleber Bez Birolo Candiottto

No dia 1 de janeiro de 2024, o Papa Francisco surpreendeu o mundo dirigindo sua mensagem para o Dia Internacional da Paz com foco no tema da Inteligência Artificial.

Em seu discurso, o pontífice nos recorda sobre a urgente necessidade de aprofundarmos o debate sobre como as inovações tecnológicas podem moldar a nossa sociedade e como elas podem contribuir para a construção de um mundo mais justo e pacífico. A Inteligência Artificial nos proporciona uma oportunidade inédita para transformar a realidade humana e do planeta, pois como recorda o documento *Gaudium et Spes*, com o nosso trabalho e inteligência buscamos desenvolver mais a própria vida, colaborando para uma melhor organização da sociedade, para o aumento da liberdade e da comunhão fraterna (n. 33).

Contudo, junto com a esperança e com as infinitas possibilidades que as sempre novas e aceleradas tecnologias nos trazem, o Papa Francisco não se abstém de apontar os riscos que agora, mas que também a médio e a longo prazo, nos tornam vulneráveis. Assim, como universidade católica, temos a responsabilidade de intensificar a formação e o cuidado no tratamento deste tema, pois como já vimos em nosso dia

o desenvolvimento
da Inteligência
Artificial tem
promovido
mudanças
significativas
em diferentes
espaços sociais,
das comunicações
à administração
pública, da vida
familiar à vida
escolar, das
subjektividades
às coletividades,
da economia à
política.

a dia, o desenvolvimento da Inteligência Artificial tem promovido mudanças significativas em diferentes espaços sociais, das comunicações à administração pública, da vida familiar à vida escolar, das subjektividades às coletividades, da economia à política. E, ocupando um lugar singular nas discussões, a Igreja católica não apenas tem se interessado por esse debate, mas está comprometida com desdobramentos éticos das possíveis novas formas de exclusão e problemas sociais que desse contexto podem derivar.

Nesse sentido, para garantir que a IA beneficie a todas e todos, é crucial adotar princípios como inclusão, transparência e segurança, e estabelecer organismos que supervisionem as questões éticas, evitando, dessa forma, que dados pessoais sejam manipulados para fins comerciais e políticos, comprometendo a liberdade de escolha. Além disso, as legislações que existem ainda são ineficientes no controle das diferentes formas de violência e exposição no espaço virtual. A tecnologia, recorda o Papa, deve ser usada para promover o bem comum e não agravar desigualdades ou conflitos.

Inteligências como a *machine learning* e *deep learning* têm impactado decisivamente a realidade, mas levantam questões sobre →



SIGNS

On January 1, 2024, Pope Francis surprised the world by directing his message for the International Day of Peace focusing on the theme of Artificial Intelligence.

In his speech, the pontiff reminds us of the urgent need to deepen the debate on how technological innovations can shape our society and how they can contribute to building a more just and peaceful world. Artificial Intelligence provides us with an unprecedented opportunity to transform human reality and the planet, because as the document *Gaudium et Spes* recalls, with our work and intelligence we seek to further develop life itself, contributing to a better organization of society, to the increase of freedom and fraternal communion (n. 33).

However, along with the hope and infinite possibilities that ever-new and accelerated technologies bring us, Pope Francis does not refrain from pointing out the risks that now, but also in the medium and long term, make us vulnerable. In this sense, as a Catholic university, we have the responsibility to intensify training and care in dealing with this topic, because as we already see in our daily lives, the development

the development
of Artificial
Intelligence
has promoted
significant changes
in different social
spaces, from
communications
to public
administration,
from family life to
school life, from
subjectivities to
collectivities, from
the economy to
politics.

of Artificial Intelligence has promoted significant changes in different social spaces, from communications to public administration, from family life to school life, from subjectivities to collectivities, from the economy to politics. And, occupying a unique place in the discussions, the Catholic Church has not only taken an interest in this debate, but is also committed to the ethical developments of possible new forms of exclusion and social problems that may arise from this context.

In this sense, to ensure that AI benefits everyone, it is crucial to adopt principles such as inclusion, transparency and security, and to establish bodies that oversee ethical issues, thus preventing personal data from being manipulated for commercial and political purposes, compromising freedom of choice. Furthermore, existing legislation is still ineffective in controlling the different forms of violence and exposure in the virtual space. Technology, the Pope recalls, must be used to promote the common good and not to exacerbate inequalities or conflicts.

Machine learning and deep learning have had a decisive impact on reality, but they raise questions about the deeper meaning →

AdobeStock_880906545



A tecnologia, recorda o Papa, deve ser usada para promover o bem comum e não agravar desigualdades ou conflitos.

o significado mais profundo da vida e da verdade. Ao mesmo tempo que colaboram conosco, tais ferramentas podem criar informações falsas e preconceituosas, prejudicando a confiança nos meios de comunicação, exacerbando a discriminação e as interferências eleitorais. É preciso, portanto, um uso cuidadoso, a fim de proteger o bem-estar coletivo e pessoal.

Uma de suas preocupações fundamentais, que se relaciona diretamente à doutrina social da Igreja, é que o uso crescente da Inteligência Artificial para avaliar a fiabilidade de indivíduos e atribuir direitos pode exacerbar os preconceitos e as desigualdades sociais, dado que os sistemas automáticos podem reproduzir erros e injustiças. A utilização invasiva de dados e sistemas de crédito social pode comprometer a dignidade humana, reduzindo indivíduos a meros conjuntos de dados e ignorando valores essenciais como compaixão e perdão. No âmbito laboral, além da precarização, a automação pode beneficiar desproporcionalmente poucos enquanto prejudica muitos, tornando essencial que a Comunidade Internacional priorize o respeito pelos direitos dos trabalhadores e a equidade salarial.

Em tempos de guerras e conflitos, como não falar do setor armamentista? O pontífice apontou que a crescente automação das armas, incluindo sistemas letais independentes, levanta sérias preocupações devido à dificuldade de atribuir responsabilidade moral às máquinas. A possibilidade de tais armas serem usadas para fins destrutivos agrava o risco de promover a guerra e a violência. Em contraste, a Inteligência Artificial, se direcionada para fins pacíficos, pode promover o desenvolvimento humano integral e a fraternidade global.

O avanço da Inteligência Artificial traz desafios significativos para a educação e para a cultura. Enquanto as tecnologias digitais ampliam as formas de comunicação e podem se tornar poderosas ferramentas de aprendizagem, há uma necessidade urgente de refletir sobre como estas moldam as relações e os métodos de ensino. A educação deve focar no desenvolvimento do pensamento crítico, capacitando crianças, adolescentes, jovens e adultos a discernir e avaliar as próprias informações provenientes da web e de sistemas de informação. Somos constantemente bombardeados por uma avalanche de informações, mas isso não significa que elas gerem em nós conhecimento. Além do mais, é crucial enfrentar a

desinformação e os medos antigos que se escondem atrás das novas tecnologias. Instituições educacionais e científicas devem orientar a formação em tecnologia com um enfoque ético e social, promovendo a integração e o respeito entre culturas para uma coexistência pacífica, plural e rica de encontros.

Para assegurar que a IA contribua para a paz e o bem comum, o Papa Francisco nos convidou a fomentar uma “algorética” – uma ética da programação e do uso de tecnologias digitais. Isso implica incorporar valores como transparência, equidade e respeito pela privacidade desde as fases iniciais do desenvolvimento tecnológico. Tais aspectos passam pela construção de uma regulamentação internacional robusta e a promoção de uma educação que capacite os indivíduos a utilizarem a IA com discernimento.

Foi levando em conta esse contexto complexo e desafiador, mas também cheio de oportunidades educativas e humanas que nos aponta o Papa Francisco, que a PUCPR aprovou, no dia 21 de agosto de 2024, em seu Conselho Universitário, *As diretrizes para uso da Inteligência Artificial – IA na PUCPR*. Elas dizem respeito, sobretudo, ao alinhamento institucional para a orientação das formas éticas, responsáveis, intencionais e conscientes do uso de IA no processo de ensino e aprendizagem e na pesquisa, buscando, dessa forma, mitigar riscos de aprendizado, desenvolvimento cognitivo, honestidade acadêmica ou outro, que o uso inconsequente de ferramentas IA pode trazer. Além disso, seu objetivo é dar apoio e suporte aos docentes e estudantes para que eles possam explorar as potencialidades que as ferramentas de Inteligência Artificial disponibilizam. É a PUCPR assumindo sua responsabilidade na formação de estudantes preparados para construir um ambiente em que a Inteligência Artificial esteja a serviço da construção de um mundo melhor e mais justo para todas e todos. ●

Boa leitura!

Fabiano Incerti

Filósofo e Pró-Reitor de Missão,
Identidade e Extensão da PUCPR.

Douglas Borges Candido

Filósofo e Coordenador do Instituto l’Hermitage
de Estudos Avançados da PUCPR.

Technology, the Pope recalls, must be used to promote the common good and not to exacerbate inequalities or conflicts.

of life and truth. While helping us, these tools can create false and biased information, undermining trust in the media, exacerbating discrimination and electoral interference. They must therefore be used carefully to protect collective and personal well-being.

One of their fundamental concerns, which relates directly to the Church's social doctrine, is that the increasing use of Artificial Intelligence to assess the reliability of individuals and assign rights can exacerbate prejudices and social inequalities, since automated systems can reproduce errors and injustices. The invasive use of data and social credit systems can undermine human dignity, reducing individuals to mere data sets and ignoring essential values such as compassion and forgiveness. In the workplace, in addition to precariousness, automation can disproportionately benefit a few while harming many, making it essential that the international community prioritize respect for workers' rights and wage equity.

In times of war and conflict, how can we not mention the arms sector? The pontiff pointed out that the increasing automation of weapons, including independent lethal systems, raises serious concerns due to the difficulty of attributing moral responsibility to machines. The possibility of such weapons being used for destructive purposes increases the risk of promoting war and violence. In contrast, Artificial Intelligence, if directed towards peaceful ends, can promote integral human development and global fraternity.

The advancement of Artificial Intelligence poses significant challenges for education and culture. While digital technologies expand forms of communication and can become powerful learning tools, there is an urgent need to reflect on how they shape relationships and teaching methods. Education must focus on developing critical thinking, enabling children, adolescents, young people and adults to discern and evaluate information coming from the web and information systems. We are constantly bombarded by an avalanche of information, but this does not mean that it generates knowledge in us. Furthermore, it is crucial to confront misinformation and old fears that hide behind new technologies. Educational

and scientific institutions must guide technology training with an ethical and social focus, promoting integration and respect between cultures for a peaceful, plural and rich coexistence of encounters.

To ensure that AI contributes to peace and the common good, Pope Francis invited us to foster an "algoritics" – an ethics of programming and the use of digital technologies. This involves incorporating values such as transparency, fairness and respect for privacy from the earliest stages of technological development. This includes building robust international regulation and promoting education that empowers individuals to use AI with discernment.

It was taking into account this complex and challenging context, but also full of educational and human opportunities that Pope Francis points out to us, that PUCPR approved, on August 21, 2024, in its University Council, the *Guidelines for the use of Artificial Intelligence – AI* at PUCPR. They concern, above all, the institutional alignment for guiding ethical, responsible, intentional and conscious ways of using AI in the teaching and learning process and in research, thus seeking to mitigate risks to learning, cognitive development, academic honesty or others, that the reckless use of AI tools can bring. In addition, its objective is to provide support and assistance to teachers and students so that they can explore the potential that Artificial Intelligence tools make available. This is PUCPR assuming its responsibility in the education of students prepared to build an environment in which Artificial Intelligence is at the service of building a better and fairer world for everyone¹. ●

Good read!

Fabiano Incerti

Philosopher and Pro-Rector of Mission, Identity and Extension at PUCPR.

Douglas Borges Candido

Philosopher and coordinator of the l'Hermitage Institute of Advanced Studies at PUCPR.

¹ "Todas e todos" in the original.

ENTREVISTA EXCLUSIVA

Pierre Lévy

A nossa existência estende-se num intervalo e a *relação fundamental entre o virtual e o real é uma transformação recíproca*. É um morfismo que projeta o sensível no inteligível e vice-versa.

SOBRE O CONVIDADO

Pierre Lévy é escritor, filósofo e doutor em Sociologia e em História da Ciência. Atualmente, é CEO da INTLEKT Metadata Inc., professor emérito da Universidade de Ottawa, professor associado na Universidade de Montreal e membro da Academia de Ciências do Canadá. Autor de diversas publicações, dentre elas *Cibercultura*; *A inteligência coletiva*; *Ciberdemocracia*; *O que é o virtual?* *As tecnologias da inteligência*.

EXCLUSIVE INTERVIEW

Pierre Lévy

Our existence extends in an interval and the *fundamental relationship between the virtual and the real is a reciprocal transformation. It is a morphism that projects the sensible into the intelligible and vice versa.*

EXCLUSIVE INTERVIEW

ABOUT THE GUEST

Pierre Lévy is a writer, philosopher and PhD in Sociology and History of Science. He is currently CEO of INTLEKT Metadata Inc., professor emeritus at the University of Ottawa, associate professor at the University of Montreal and member of the Canadian Academy of Sciences. He is the author of several publications, including *Cyberculture*, *Collective Intelligence*, *Cyberdemocracy*, *What is the virtual?* *The technologies of intelligence*.



P1 - Diante da crescente hiperconectividade entre os jovens, muitos especialistas falam sobre a solidão e o que chamam de “a era das paixões tristes”. Como você vê essa dicotomia entre proximidade e distância que a tecnologia provoca nas relações humanas?

R1 - A hiperconectividade não diz respeito apenas aos jovens, porque está em toda parte. Um dos principais fatores da evolução cultural reside no dispositivo material de produção e reprodução de símbolos, mas também nos sistemas de software para escrita e codificação de informações. A nossa inteligência coletiva estende-se à das espécies sociais que nos precederam, e, particularmente, à dos grandes símios. Mas o uso da linguagem – e de outros sistemas simbólicos – tal como a força dos nossos meios técnicos fez-nos passar do estatuto de animal social para o de animal político. Propriamente humana, a *Polis* emerge da simbiose entre ecossistemas de ideias e as populações de primatas falantes que os mantêm, alimentam-se deles e refletem sobre eles. As evoluções tanto das ideias quanto das populações de *Sapiens* determinam-se mutuamente. Porém, o principal fator na evolução das ideias está na estrutura material de reprodução dos símbolos. Ao longo da

Em suma, a evolução cultural levou-nos ao ponto em que os ecossistemas de ideias se manifestam como dados conduzidos por algoritmos num espaço virtual ubíquo. E é neste espaço que os laços sociais se formam, se mantêm e se desfazem.

história, os símbolos (com as ideias que transportavam) foram, sucessivamente, perpetuados pela escrita, simplificados pelo alfabeto e pelo papel, multiplicados pela impressão e pelos meios elétricos. Os símbolos são hoje digitalizados e calculados, ou seja, uma série de robôs de software – algoritmos – gravam-nos, contam-nos, traduzem-nos e extraem padrões. Os objetos simbólicos (textos, imagens estáticas ou animadas, vozes, músicas, programas etc.) não são apenas gravados, reproduzidos e transmitidos automaticamente, mas também gerados e transformados industrialmente. Em suma, a evolução cultural levou-nos ao ponto em que os ecossistemas de ideias se manifestam como dados conduzidos por algoritmos num espaço virtual ubíquo. E é neste espaço que os laços sociais se formam, se mantêm e se desfazem. Antes de criticar ou deplorar, devemos primeiro reconhecer os fatos. As amizades dos jovens já não podem

prescindir das redes sociais; casais se encontram na internet, por exemplo, em aplicativos como o Tinder (ver Figura 1); as famílias permanecem conectadas por meio do Facebook ou de outros aplicativos como o WhatsApp; os espaços de trabalho mudaram para eletrônicos com Zoom e Teams, principalmente desde a pandemia de covid-19; a diplomacia é, cada vez mais, feita no X (antigo Twitter) etc. Não tem volta. Por outro lado, não nos movemos menos fisicamente: como evidenciam os monstruosos engarrafamentos em São Paulo e no Rio de Janeiro. Na mesma linha, a tendência dos últimos dez anos – uma era de crescimento exponencial das ligações à internet – mostra também um aumento no número de passageiros aéreos, o que continua sendo uma tendência secular, apesar de ter sofrido uma queda significativa durante a pandemia da covid-19.

Senti-me muito sozinho quando, ainda jovem estudante, cheguei a Paris, vindo do sul da França, para dar continuidade aos meus estudos universitários. Era 1975 e não havia internet. Os idosos que moram sozinhos e não são visitados pelos filhos deveriam culpar a internet? O problema da solidão e da desintegração dos laços sociais é muito real. Mas é uma tendência já antiga, que se deve à urbanização, às mudanças na família e a muitos outros fatores. Convido vocês, leitores, a consultarem os numerosos trabalhos sobre a questão do “capital social” (a quantidade e a qualidade das relações humanas). A internet é apenas um dos muitos fatores a serem considerados nesta questão. →

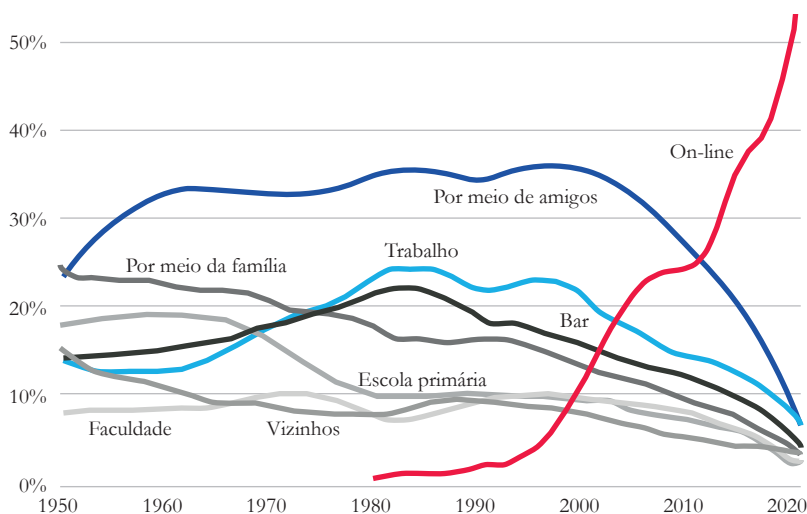


Figura 1: Como os casais se conhecem nos Estados Unidos

Fonte: “How couples meet and stay together”: A longitudinal study of social life in the US (“Como os casais se conhecem e permanecem juntos”: Um estudo longitudinal da vida social nos Estados Unidos) por M. J. Rosenfeld, Reuben J. Thomas e Sonja Hausen. Análise dos dados originais da pesquisa (n=6.519). Categoria “bares & restaurantes” removida para não contar duas vezes casais que se conheceram primeiro on-line.

Q1 - In view of the growing hyperconnectivity among young people, many experts talk about loneliness and what they call “the age of sad passions”. How do you see this dichotomy between proximity and distance that technology causes in human relationships?

R1 - Hyperconnectivity does not only concern young people, because it is everywhere. One of the main factors in cultural evolution lies in the material device for producing and reproducing symbols, but also in software systems for writing and coding information. Our collective intelligence extends to that of the social species that preceded us, and particularly to that of great apes. But the use of language – and other symbolic systems – as well as the power of our technical means have made us move from the status of a social animal to that of a political animal. The *Polis*, which is specifically human, emerges from the symbiosis between ecosystems of ideas and the populations of talking primates that maintain them, feed on them and reflect on them. The evolution of both ideas and populations of *Sapiens* determine each other. However,

In short, cultural evolution has led us to the point where ecosystems of ideas manifest themselves as data driven by algorithms in a ubiquitous virtual space. And it is in this space that social bonds are formed, maintained and dissolved.

the main factor in the evolution of ideas lies in the material structure of reproduction of symbols. Throughout history, symbols (with the ideas they conveyed) have been successively perpetuated by writing, simplified by the alphabet and paper, multiplied by printing and electrical media. Today, symbols are digitized and calculated, that is, a series of software robots – algorithms – record them, count them, translate them and extract patterns. Symbolic objects (texts, static or animated images, voices, music, programs, etc.) are not only recorded, reproduced and transmitted automatically, but also generated and transformed industrially. In short, cultural evolution has led us to the point where ecosystems of ideas manifest themselves as data driven by algorithms in a ubiquitous virtual space. And it is in this space that

social bonds are formed, maintained and dissolved. Before criticizing or deploring, we must first acknowledge the facts. Young people’s friendships can no longer be made without social media; couples meet online, for example, on apps like Tinder (see Figure 1); families stay connected through Facebook or other apps like WhatsApp; workspaces have shifted to electronic ones with Zoom and Teams, especially since the COVID-19 pandemic; diplomacy is increasingly conducted on X (formerly Twitter), etc. There is no going back. On the other hand, we are not moving less physically: as evidenced by the monstrous traffic jams in São Paulo and Rio de Janeiro. In the same vein, the trend of the last ten years – an era of exponential growth in internet connections – also shows an increase in the number of air passengers, which continues to be a secular trend, despite having suffered a significant decline during the COVID-19 pandemic.

I felt very lonely when, as a young student, I arrived in Paris from the south of France to continue my university studies. It was 1975 and there was no Internet. Should elderly people who live alone and are not visited by their children blame the Internet? The problem of loneliness and the disintegration of social ties is very real. But it is a long-standing trend, due to urbanization, changes in the family, and many other factors. I invite you, readers, to consult the numerous works on the question of “social capital” (the quantity and →

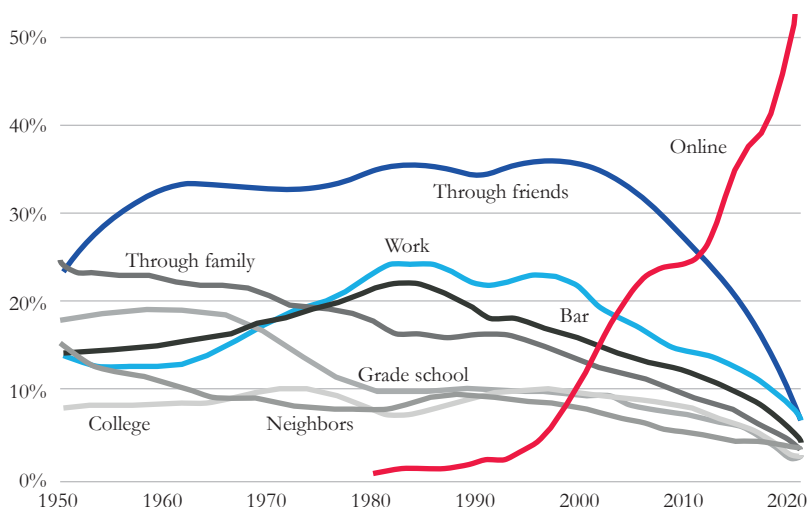


Figure 1: How Couples Meet and Stay Together in the United States

Source: “How couples meet and stay together”: A longitudinal study of social life in the US by M. J. Rosenfeld, Reuben J. Thomas, and Sonia Hausen. Analysis of original survey data (n=6,519). “Bars & restaurants” category removed (“cleaned”) to avoid double-counting couples who first met online.

P2 - Em seus livros *Inteligência coletiva: por uma antropologia do ciberespaço* (1994) e *Cibercultura: a cultura da sociedade digital* (1997), você argumenta que a internet e as tecnologias digitais desenvolvem a inteligência coletiva, possibilitando novas formas de colaboração e compartilhamento de conhecimento. No entanto, existe uma preocupação crescente de que o uso excessivo das redes sociais e das tecnologias digitais esteja associado à distração e ao atraso na aprendizagem entre os jovens. Como você vê esta aparente contradição entre o potencial das tecnologias para fortalecer a inteligência coletiva e os efeitos negativos que elas podem ter no desenvolvimento cognitivo e educacional dos jovens?

R2 - Nunca afirmei que a internet e as tecnologias digitais por si só, e como se as técnicas fossem sujeitos autônomos, desenvolvessem a inteligência coletiva. Afirmar que o *melhor* uso que podemos fazer da internet e das tecnologias digitais era desenvolver a inteligência coletiva humana, o que é muito diferente. E ainda é isso o que eu penso. A ideia de um “espaço de conhecimento”, que poderia ser implantado acima do espaço comercial, é um ideal regulatório para ação, não uma previsão fatual. Quando eu escrevi *Inteligência Coletiva*—de 1992 a 1993—, menos de 1% da humanidade estava ligada à internet, e a web não existia. Você não encontrará a palavra “web” em nenhum lugar do livro. No entanto, hoje — em 2024 —, temos bem mais de dois terços da população mundial ligada à internet. O contexto é, portanto, completamente diferente, mas a mudança na civilização, que eu previ há 30 anos, hoje, parece óbvia, embora, normalmente, tenhamos de esperar várias gerações para confirmar esse tipo de mutação. Na minha opinião, estamos apenas no início da revolução digital. →





quality of human relationships). The Internet is just one of many factors to be considered in this regard.

Q2 - In your books *Collective Intelligence: For an Anthropology of Cyberspace* (1994) and *Cyberculture: The Culture of Digital Society* (1997), you argue that the Internet and digital technologies develop collective intelligence, enabling new forms of collaboration and knowledge sharing. However, there is growing concern that excessive use of social networks and digital technologies is associated with distraction and delayed learning among young people. How do you view this apparent contradiction between the potential of technologies to strengthen collective intelligence and the negative effects they can have on the cognitive and educational development of young people?

R2 - I never claimed that the Internet and digital technologies alone, and as if the technologies were autonomous subjects, would develop collective intelligence. I claimed that *the best* use we could make of the Internet and digital technologies was to develop human collective intelligence, which is very different. And that is still what I think. The idea of a “knowledge space” that could be deployed above the commercial space is a regulatory ideal for action, not a factual prediction. When I wrote *Collective Intelligence* – from 1992 to 1993 – less than 1% of humanity was connected to the Internet, and the Web did not exist. You will not find the word “Web” anywhere in the book. However, today – in 2024 – we have well over two-thirds of the world’s population connected to the Internet. The context is therefore completely different, but the change in civilization that I predicted 30 years ago today seems obvious, although normally →

Quanto ao aumento da inteligência coletiva, foram dados muitos passos para tornar o conhecimento acessível a todos. A Wikipédia é o exemplo clássico de uma empresa que atua em inteligência coletiva, com milhões de colaboradores voluntários de todos os países e grupos de discussão entre especialistas para

cada artigo. São quase sete milhões de artigos em inglês, dois milhões e meio de artigos em francês e mais de um milhão de artigos em português. A Wikipédia é consultada por dezenas de milhões de pessoas por dia e vários bilhões por ano! O software livre – agora amplamente adotado e distribuído, inclusive por grandes empresas da web – é outra área importante em que a inteligência coletiva está no comando. Entre os softwares livres mais utilizados estão o sistema operacional Linux, os navegadores Mozilla e Chromium, o pacote OpenOffice, o servidor http Apache (que é o mais utilizado na internet), o sistema de controle de versão GIT, mensagens Signal e muitos outros que demoraria muito para citar. Eu acrescentaria que bibliotecas e museus digitalizados, bem como artigos científicos de acesso aberto e sites do tipo ArXiv.org, são comuns, o que vem transformando as práticas de investigação e comunicação científicas. Hoje, todos podem publicar textos em seu blog, vídeos e podcasts no YouTube ou em outros sites, o que não acontecia há trinta anos. As redes sociais permitem a troca de notícias e ideias com muita rapidez, como vemos, por exemplo, no LinkedIn ou no X (ex-Twitter). A internet permitiu, portanto, verdadeiramente, o desenvolvimento de novas formas de expressão, colaboração e partilha de conhecimento. Ainda há muito a ser feito. Estamos apenas no início da mutação antropológica em curso.

É claro que devemos ter em conta os fenômenos da dependência dos videogames, das redes sociais, da pornografia on-line etc. Mas, durante mais de trinta anos, a maioria dos jornalistas, políticos, professores e todos aqueles que formam opiniões continuaram a *denunciar os perigos* da computação, depois da internet e, agora, da Inteligência Artificial. Eu não faria nada de muito útil se acrescentasse os meus lamentos aos deles. Procuro, portanto, sensibilizar

para uma mudança em grande escala na civilização, que não será estancada, e indicar os melhores meios de orientar essa grande transformação para os objetivos mais positivos e para o desenvolvimento humano. Dito isto, é claro que os fenômenos da dependência encontram, parcialmente, a sua origem na nossa dependência da arquitetura sociotécnica tóxica das grandes empresas da web, que utilizam a estimulação dopaminérgica e os reforços narcisistas da comunicação digital para nos fazer produzir cada vez mais dados e vender mais publicidade. Infelizmente, a saúde mental das populações adolescentes é talvez uma das vítimas colaterais das estratégias comerciais dessas grandes empresas oligopolistas. Como podemos opor-nos ao poder dos seus centros de computação, à sua eficiência de software e à simplicidade das suas interfaces? É mais fácil fazer a pergunta do que respondê-la. Além da biopolítica mencionada por Michel Foucault, devemos agora considerar uma psicopolítica baseada no neuromarketing, nos dados pessoais e na gamificação do controle. Os professores devem alertar os alunos sobre esses perigos e treiná-los para o pensamento crítico.

P3 - Com o fenômeno das “bolhas conectivas”, com as quais as redes sociais tendem a reforçar crenças e ideias pré-existentes, limitando o contato com diferentes perspectivas, como você vê a evolução das conexões sociais à medida que a internet e as plataformas digitais continuam se desenvolvendo? Poderá este tipo de segmentação enfraquecer a inteligência coletiva que você defende, ou ainda há espaço para conexões mais amplas e colaborativas no futuro?

R3 - É claro que se nos contentarmos em “gostar” instintivamente daquilo que vemos passar e reagimos emocionalmente às imagens e mensagens mais simplistas, →

we have to wait several generations to confirm this type of mutation. In my opinion, we are just at the beginning of the digital revolution.

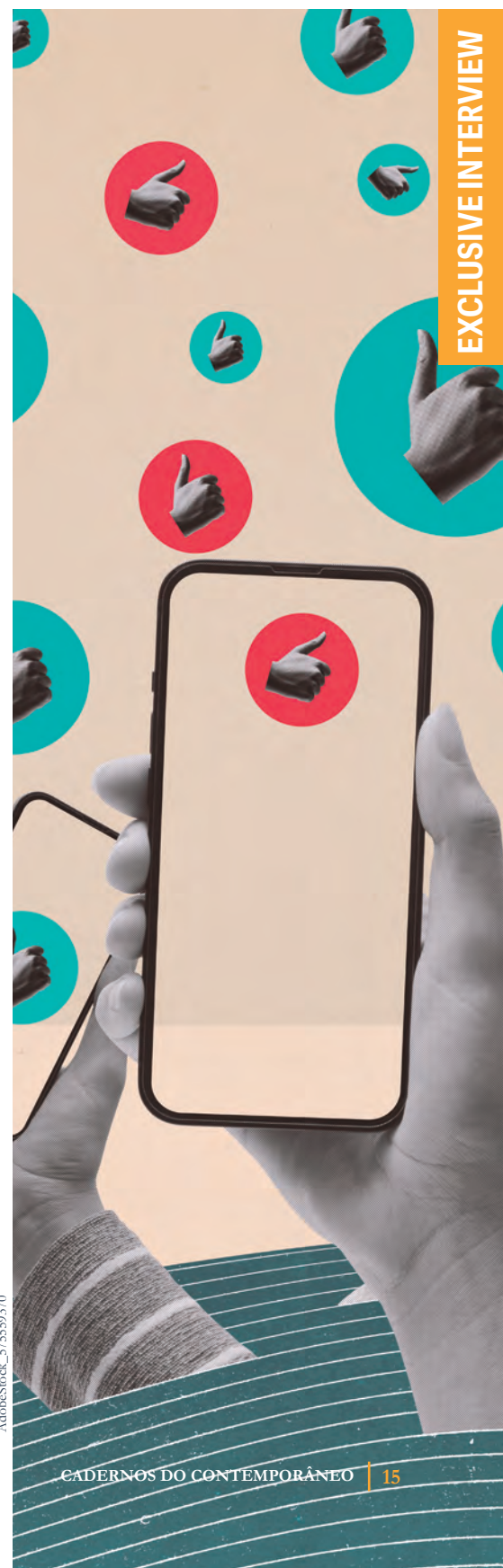
As for increasing collective intelligence, many steps have been taken to make knowledge accessible to everyone. Wikipedia is a classic example of a company that works on collective intelligence, with millions of volunteer contributors from all countries and discussion groups among experts for each article. There are almost seven million articles in English, two and a half million articles in French and more than a million articles in Portuguese. Wikipedia is consulted by tens of millions of people every day and several billion every year! Free software – now widely adopted and distributed, including by large web companies – is another important area in which collective intelligence is in command. Among the most widely used free software are the Linux operating system, the Mozilla and Chromium browsers, the OpenOffice suite, the Apache HTTP server (which is the most widely used on the Internet), the GIT version control system, Signal messaging and many others that would take too long to list. I would add that digitized libraries and museums, as well as open access scientific articles and sites like ArXiv.org, are common, which has been transforming the practices of scientific research and communication. Today, everyone can publish texts on their blogs, videos and podcasts on YouTube or other websites, which was not the case thirty years ago. Social networks allow the exchange of news and ideas very quickly, as we see, for example, on LinkedIn or X (formerly Twitter). The Internet has therefore truly enabled the development of new forms of expression, collaboration and knowledge sharing. There is still

much to be done. We are only at the beginning of the anthropological transformation that is underway.

Of course, we must take into account the phenomena of addiction to video games, social networks, online pornography, etc. But for more than thirty years, most journalists, politicians, teachers and all those who shape opinions have continued to *denounce the dangers* of computing, then the Internet and now Artificial Intelligence. I would not be doing anything very useful if I added my laments to theirs. I therefore seek to raise awareness of a large-scale change in civilization that will not be stopped, and to indicate the best means of directing this great transformation towards the most positive objectives and human development. That said, it is clear that the phenomena of addiction find their origin, in part, in our dependence on the toxic socio-technical architecture of large web companies, which use dopaminergic stimulation and the narcissistic reinforcements of digital communication to make us produce more and more data and sell more advertising. Unfortunately, the mental health of adolescent populations is perhaps one of the collateral victims of the commercial strategies of these large oligopolistic companies. How can we oppose the power of their computing centers, the efficiency of their software, and the simplicity of their interfaces? The question is easier to ask than to answer. In addition to the biopolitics mentioned by Michel Foucault, we must now consider a psychopolitics based on neuromarketing, personal data, and the gamification of control. Teachers must warn students about these dangers and train them in critical thinking.

Q3 - With the phenomenon of “connective bubbles,” with which

social networks tend to reinforce pre-existing beliefs and ideas, limiting contact with different perspectives, how do you see social connections evolving as the internet and digital platforms continue to develop? Could this type of segmentation weaken the →



o benefício cognitivo não será muito grande. Não me proponho como um modelo a ser seguido de forma absoluta, mas eu gostaria apenas de dar um exemplo do que é possível fazer se tivermos um pouco de imaginação e estivermos dispostos a questionar a inércia das instituições. Quando eu era professor de Comunicação Digital na Universidade de Ottawa, estimei meus alunos a se inscreverem no Twitter, escolherem meia dúzia de tópicos de seu interesse e elaborarem listas de contas para seguir em cada tópico. Não importando o tema – política, ciência, moda, arte, esportes etc. –, eles tiveram de construir listas equilibradas, incluindo especialistas ou simpatizantes de opiniões opostas, a fim de alargar a sua esfera cognitiva em vez de restringi-la. Nas redes sociais mais comuns, como Facebook e LinkedIn, é possível participar de um grande número de comunidades especializadas em áreas culturais (história, filosofia, artes) ou profissionais (negócios, tecnologia etc.) para se manter atualizado e debater com os especialistas. Grupos de discussão locais por cidades ou bairros também são muito úteis. É tudo uma questão de método e prática. Devemos ter senso crítico em relação aos meios de comunicação de massa (jornais em papel, rádio, televisão) em que os receptores passivos consomem programação feita por outros. Cabe a cada um “brincar” na sua própria programação e construir as suas próprias redes de aprendizagem pessoais.

Antes da imprensa, só falávamos com as pessoas da nossa paróquia. Nos anos sessenta do século XX, só podíamos escolher entre dois ou três canais de

televisão e dois ou três jornais. Hoje temos acesso a uma enorme diversidade de fontes de todos os países e de todos os setores da sociedade. Os professores devem alfabetizar os alunos, ensinar-lhes línguas estrangeiras, dar-lhes uma boa cultura geral e orientá-los neste novo mundo da comunicação.

P4 - Atualmente, há um debate crescente sobre os efeitos negativos da tecnologia na saúde mental dos jovens, com foco em questões como ansiedade, depressão e isolamento social. Considerando o papel central que as tecnologias digitais desempenham na nossa sociedade, como você entende esta relação entre o uso intensivo de tecnologias e o aumento dos problemas de saúde mental entre os jovens? Existe uma maneira de equilibrar os benefícios da tecnologia com a necessidade de preservar o bem-estar mental?

R4 - O problema da saúde mental dos jovens é obviamente muito real, mas seria muito redutor atribuí-lo apenas às redes sociais. No entanto, tentarei listar alguns problemas psicológicos que surgem com o uso de tecnologias digitais.

Em primeiro lugar, há a transformação da autorreferência subjetiva, que corre o risco de levar a problemas do tipo esquizofrênico. Nosso campo de experiência é mediado pelo meio digital: o ciclo de autorreferência está mais amplo do que nunca. Interagimos com pessoas, robôs, imagens, músicas por meio de diversas interfaces multimídia: tela, fones de ouvido, comandos... Nossa experiência subjetiva é controlada



pelos algoritmos de múltiplas aplicações que determinam em loop (se não aprendemos a controlá-los) nosso consumo de dados e nossas ações em troca. A nossa memória está dispersa em numerosos arquivos, bases de dados, localmente e na nuvem... Quando grande parte de nós é assim coletivizada e externalizada, o problema dos limites e da determinação da identidade torna-se preponderante. Quem é o proprietário dos dados que me dizem respeito, quem os produz?

O problema do narcisismo é particularmente evidente no Instagram e em aplicativos semelhantes. Nosso ego é nutrido pela imagem que os outros nos enviam de volta no meio algorítmico. A obsessão pelas imagens atinge proporções preocupantes. Quantos inscritos, quantas curtidas, quantas impressões? Para quem →

Existe um problema de saúde mental se os afetos são constantemente eufóricos, constantemente disfóricos ou se um objeto exclusivo se torna viciante.



AdobeStock_596310493

collective intelligence that you advocate, or is there still room for broader and more collaborative connections in the future?

R3 - Of course, if we are content to instinctively “like” (“liker”) what we see and react emotionally to the most simplistic images and messages, the cognitive benefit will not be very great. I do not propose myself as a model to be followed in an absolute way, but I would just like to give an example of what is possible if we have a little imagination and are willing to question the inertia of institutions. When I was a professor of Digital Communications at the University of Ottawa, I encouraged my students to sign up for Twitter, choose half a dozen topics that interested them and make lists of accounts to follow on each topic. No matter the topic – politics, science, fashion, art, sports, etc. – they had to

build balanced lists, including experts or supporters of opposing views, in order to broaden their cognitive sphere rather than restrict it. On the most common social networks, such as Facebook and LinkedIn, it is possible to join a large number of communities specializing in cultural (history, philosophy, arts) or professional (business, technology, etc.) areas to keep up to date and debate with experts. Local discussion groups by city or neighborhood are also very useful. It is all a question of method and practice. We must move away from the model of mass media (printed newspapers, radio, television) in which passive receivers consume programming created by others. It is up to each person to “play” with their own programming and build their own personal learning networks.

Before the press, we only spoke to the people in our parish. In the 1960s, we could only choose between two or three television channels and two or three newspapers. Today we have access to a huge diversity of sources from all countries and all sectors of society. Teachers must teach students to read and write, teach them foreign languages, give them a good general culture and guide them in this new world of communication.

Q4 - There is currently a growing debate about the negative effects of technology on young people's mental health, with a focus on issues such as anxiety, depression and social isolation. Considering the central role that digital technologies play in our society, how do you understand this relationship between the intensive use of technology and the increase

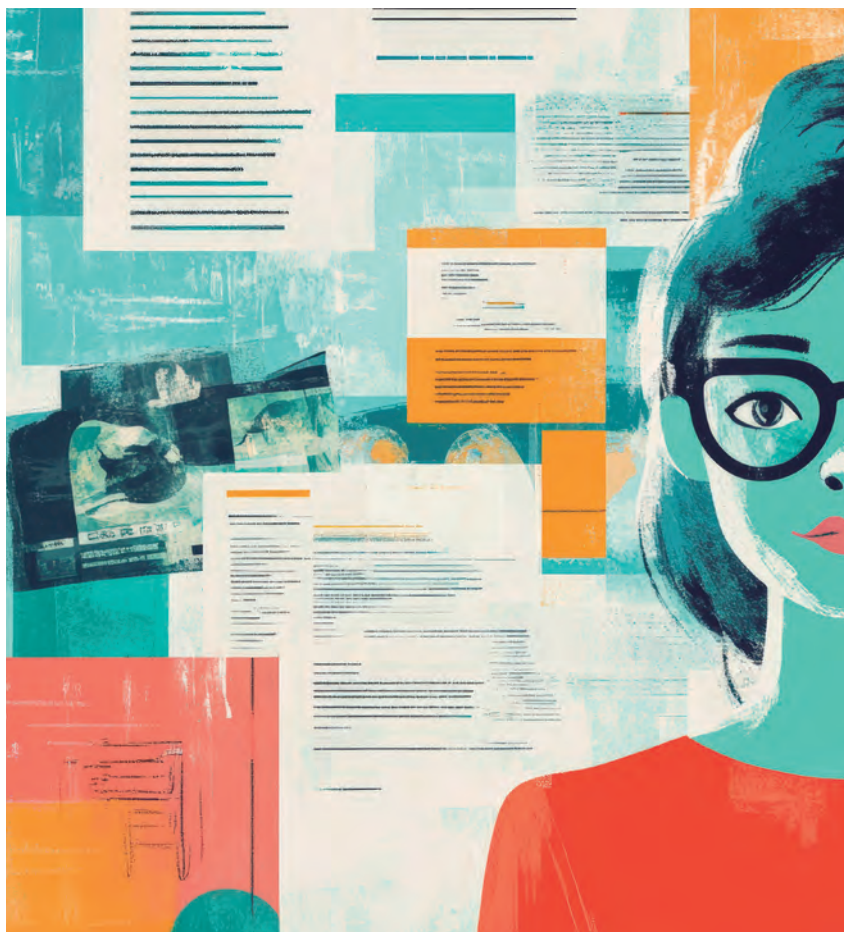
in mental health problems among young people? Is there a way to balance the benefits of technology with the need to preserve mental well-being?

R4 - The mental health problem among young people is obviously very real, but it would be too simplistic to attribute it solely to social networks. However, I will try to list some psychological problems that arise with the use of digital technologies.

Firstly, there is the transformation of subjective self-reference, which risks leading to schizophrenic-type problems. Our field of experience is mediated by the digital medium: the cycle of self-reference is wider than ever. We interact with people, robots, images, music through various multimedia interfaces: screens, headphones, controls... Our subjective experience is controlled by the algorithms of multiple applications that determine, in a loop (if we do not learn to control them) our data consumption and our actions in return. Our memory is dispersed in numerous files, databases, locally and in the cloud... When a large part of us is collectivized and externalized in this way, the problem of boundaries and the determination of identity becomes preponderant. Who owns the data that concerns me, who produces it?

The problem of narcissism is particularly evident on Instagram and similar apps. Our ego is fed by the image that others send back to us in the algorithmic environment. The obsession with images reaches worrying proportions. How many →

There is a mental health problem if the affects are constantly euphoric, constantly dysphoric or if an exclusive object becomes addictive.



mergulhou nesse abismo, o valor do ser está apenas nos olhos dos outros. Antes de ser um problema de saúde mental, é um problema de sabedoria elementar.

Aqui, o eu está trancado na sua vida interior, mas alimentado por fontes de informação on-line. O código ou certos aspectos da cultura popular tornam-se obsessivos. Este é o domínio dos geeks, otakus e jogadores compulsivos. Torna-se, obviamente, pouco saudável prescindir de qualquer vida social na carne.

Existe um problema de saúde mental se os afetos são constantemente eufóricos, constantemente disfóricos ou se um objeto exclusivo se torna viciante. Com efeito, a internet pode tornar-nos viciados em determinados objetos (notícias, séries, jogos, pornografia) ou a determinadas emoções, sejam elas positivas (conteúdos do tipo “feel-good”, como gatos fofos, dança, humor etc.) ou negativas (atualidades catastróficas,

“doomscrolling”¹). Também podemos nos perguntar até que ponto é bom que a linguagem corporal seja completamente substituída por emojis, memes, imagens, avatares etc.

O vício é criado pela excitação (dopamina) e pela satisfação (endorfina) que queremos reproduzir constantemente. No entanto, como disse anteriormente, os modelos de negócio das grandes empresas da web que se concentram no envolvimento (secreção de dopamina-endorfina) levam, quase inevitavelmente, ao vício se os usuários não tomarem cuidado. A alta intensidade de envolvimento por longos períodos de tempo leva, inevitavelmente, à depressão.

O controle dos impulsos (agressão, por exemplo) é mais difícil nas redes sociais do que na vida real, porque as pessoas com quem falamos não estão à nossa frente. “Gerir comportamentos tóxicos” também é um

¹ N.T.: Leitura compulsiva de notícias ruins na internet, de forma desequilibrada.

grande problema nos jogos on-line e nas redes sociais.

Em suma, devemos estar vigilantes, alertar os jovens utilizadores sobre os perigos envolvidos e não cometer excessos.

P5 - Alguns preveem que as gerações futuras poderão nunca mais frequentar a escola. Como você vê o futuro da educação em um mundo cada vez mais hiperconectado e dominado pela tecnologia?

R5 - Não creio que a escola vá desaparecer. Mas deve se transformar. É preciso segurar os alunos onde eles estão e, de preferência, usar os produtos convencionais com os quais eles estão acostumados para transformá-los em algo útil em termos de aprendizagem. Os alunos são “nativos digitais”, mas isso não significa que tenham domínio real das ferramentas digitais. Não devemos apenas desenvolver a literacia digital, mas a própria literacia, que dela é inseparável. Sou um grande defensor da leitura dos clássicos e da cultura geral, essencial para a formação do pensamento crítico.

Para voltar aos meus próprios métodos de ensino, nos cursos que ministrei na Universidade de Ottawa, pedi aos meus alunos que participassem de um grupo fechado do Facebook, que se cadastrassem no Twitter, abrissem um blog, caso ainda não o tivessem, e usassem uma plataforma colaborativa de curadoria de dados.

O uso de *plataformas de curadoria de conteúdo* me serviu para ensinar aos alunos como escolher as categorias ou “tags” para classificar as informações úteis na memória de longo prazo, a fim de encontrá-las, facilmente, mais tarde. Essa habilidade será muito útil para eles no resto da carreira.

Os *blogs* foram utilizados como materiais de “trabalho final” para os cursos de Graduação (ou seja, antes do Mestrado) e como cadernos de pesquisa para →

subscribers, how many likes, how many impressions? For those who have plunged into this abyss, the value of being is only in the eyes of others. Before being a mental health problem, it is a problem of elementary wisdom.

Here, the ego is locked in its inner life, but fed by online sources of information. The code or certain aspects of popular culture become obsessive. This is the domain of geeks, otakus and compulsive gamers. It becomes obviously unhealthy to do without any social life in the flesh.

There is a mental health problem if the affects are constantly euphoric, constantly dysphoric or if an exclusive object becomes addictive. Indeed, the Internet can make us addicted to certain objects (news, series, games, pornography) or to certain emotions, whether positive (*feel-good* content, such as cute cats, dancing, humor, etc.) or negative (catastrophic current affairs, *doomscrolling*¹). We can also ask ourselves to what extent it is good for body language to be completely replaced by emojis, memes, images, avatars, etc.

Addiction is created by the excitement (dopamine) and satisfaction (endorphins) that we constantly want to reproduce. However, as I said before, the business models of large web companies that focus on engagement (secretion of dopamine-endorphins) almost inevitably lead to addiction if users are not careful. High intensity of engagement over long periods of time inevitably leads to depression.

Impulse control (aggression, for example) is more difficult on social networks than in real life, because the



AdobeStock_926675538

people we are talking to are not in front of us. “Managing toxic behavior” is also a big issue in online gaming and social media.

In short, we must be vigilant, warn young users about the dangers involved and avoid overdoing it.

Q5 - Some predict that future generations will never be able to attend school again. How do you see the future of education in an increasingly hyperconnected and technology-dominated world?

R5 - I don’t believe that school will disappear. But it must transform. It is necessary to keep students where they are and, preferably, use the conventional products with which they are familiar. They are used to transforming them into something useful in terms of learning. Students are “digital natives”, but that does not mean that they have real mastery of digital tools. We must not only

develop digital literacy, but literacy itself, which is inseparable from it. I am a great advocate of reading classics and general culture, which is essential for the formation of critical thinking.

To return to my own teaching methods, in the courses I taught at the University of Ottawa, I asked my students to join a closed Facebook group, sign up for Twitter, start a blog if they didn’t already have one, and use a collaborative data curation platform.

Using *content curation platforms* helped me teach students how to choose categories or “tags” to classify useful information in long-term memory, in order to find it easily later. This skill will be very useful for the rest of their careers.

Blogs have been used as “final work” materials for undergraduate courses (i.e. before the Master’s degree) and as research notebooks for Master’s →

¹ N.T.: Compulsive reading of bad news on the internet, in an unbalanced way.

alunos de Mestrado ou Doutorado: notas sobre leituras, formulação de hipóteses, acúmulo de dados, primeira versão de artigos científicos ou capítulos de dissertações ou teses etc. O caderno público de pesquisa facilita o relacionamento com o supervisor e permite reorientar a tempo as direções de pesquisa arriscadas, entrar em contato com equipes que trabalham nos mesmos assuntos etc.

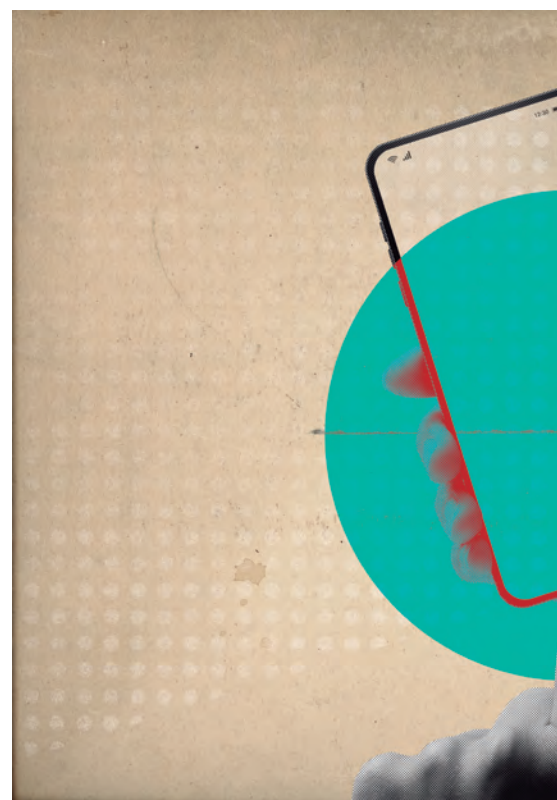
O grupo do Facebook foi utilizado para compartilhar o Syllabus ou “plano de curso”, a agenda das aulas, as leituras obrigatórias, as discussões internas em grupo – por exemplo, aquelas relativas à avaliação –, bem como endereços de e-mail dos alunos (Twitter, blog, plataforma de curadoria social etc.). Todas as informações estavam on-line e acessíveis com apenas um clique, incluindo as leituras obrigatórias digitalizadas e gratuitas. Os alunos puderam participar escrevendo “miniwikis” dentro do grupo do Facebook sobre assuntos de sua preferência, foram convidados a sugerir leituras interessantes relacionadas ao assunto do curso adicionando links comentados. Usei o Facebook porque quase todos os alunos já estavam inscritos e a funcionalidade de grupo dessa plataforma está bem estabelecida. Mas eu poderia ter usado qualquer outro suporte colaborativo de gerenciamento de grupos, no Slack ou LinkedIn.

No Twitter (agora, X), a conversa específica de cada turma era identificada por uma hashtag. No início, usei ocasionalmente a rede do pássaro azul. Por exemplo, ao final de cada aula, pedia aos alunos que anotassem a ideia mais interessante que aprenderam no curso e percorria seus tweets em tempo real na tela da aula. Então, depois de algumas semanas, convidei-os a relerem seus rastros coletivos no Twitter para

reunir e resumir o que aprenderam e fazer perguntas – sempre no Twitter – se algo não estivesse claro, perguntas às quais respondi no mesmo canal.

Após alguns anos de uso do Twitter nas aulas, fiquei mais ousado e pedi aos alunos que fizessem suas anotações diretamente nessa rede social durante o curso para obter um caderno coletivo. Poder observar como os outros fazem anotações (seja no curso ou nos textos a serem lidos) permite que os alunos comparem sua compreensão e, assim, esclareçam determinados conceitos. Eles descobrem o que os outros notaram e que isto não é, necessariamente, o que os estimulou... Quando senti que a atenção estava diminuindo um pouco, pedi que parassem, pensassem no que acabaram de ouvir e notassem suas ideias ou dúvidas, mesmo que suas observações não estivessem diretamente relacionadas ao assunto do curso. O Twitter permitiu que eles se comunicassem livremente sobre os assuntos estudados sem atrapalhar o funcionamento da aula. Sempre dediquei o final do curso a um período de perguntas e respostas baseado na visualização coletiva do feed do Twitter. Esse método é particularmente relevante em grupos demasiadamente grandes (por vezes mais de duzentas pessoas) para permitir que todos os alunos se expressem oralmente. Pude, assim, responder com tranquilidade às dúvidas após a aula sabendo que minhas explicações ficariam registradas no tópico do grupo. A conversa educacional continua entre as aulas. Claro que tudo isto só foi possível porque a avaliação (classificação dos alunos) se baseou na participação on-line.

Utilizando o Facebook e o Twitter em sala de aula, os alunos não só aprendem



a matéria do curso, como também um modo “cultivado” de se servir das mídias sociais. Documentar seus cafés da manhã ou a bebedeira da última festa, disseminar os “chats” e imagens engraçadas, as trocas de insultos entre inimigos políticos, extasiar-se com vedetes do showbusiness ou fazer publicidade para esta ou aquela empresa, são, certamente, alguns dos usos legítimos das mídias sociais. Mas podemos, igualmente, manter diálogos construtivos sobre temas comuns. Em suma, creio que a educação deve progredir em direção à aprendizagem colaborativa utilizando-se das ferramentas digitais.

P6 - Quais são, para você, as principais oportunidades que a internet e as novas ferramentas de IA podem trazer para o campo da educação? Dado o avanço acelerado das tecnologias digitais e da Inteligência Artificial (IA), como você vê a evolução do papel do professor nos próximos anos?

R6 - No que diz respeito à Inteligência Artificial (por exemplo, ChatGPT, Meta AI, Grok ou Gemini, todos →

A IA não substitui a ignorância, pelo contrário: dá um bônus a quem já tem bons conhecimentos.



or PhD students: notes on readings, formulation of hypotheses, accumulation of data, first drafts of scientific articles or chapters of dissertations or theses, etc. The public research notebook facilitates the relationship with the supervisor and allows timely reorientation of risky research directions, contact with teams working on the same topics, etc.

The *Facebook* group was used to share the Syllabus or “course plan”, the class schedule, the required readings, the internal group discussions – for example, those related to assessment – as well as the students’ email addresses (Twitter, blog, social curation platform, etc.). All the information was online and accessible with just one click, including the free, digitized required readings. Students could participate by writing “mini-wikis” within the Facebook group on topics of their interest, and were invited to suggest interesting readings related to the course topic by adding commented links. I used Facebook because

almost all the students were already enrolled and the group functionality of this platform is well established. But I could have used any other collaborative group management support, such as Slack or LinkedIn.

On Twitter (now X), the specific conversation of each class was identified by a hashtag. At first, I occasionally used the bluebird network. For example, at the end of each class, I asked students to write down the most interesting idea they learned in the course and scrolled through their tweets in real time on the classroom screen. Then, after a few weeks, I invited them to reread their collective Twitter feeds to gather and summarize what they had learned and to ask questions—always on Twitter—if anything was unclear, questions that I answered on the same channel.

After a few years of using Twitter in class, I became more daring and asked students to take notes directly on this social network during the course to create a collective notebook. Being able to observe how others take notes (whether in the course or in the texts to be read) allows students to compare their understanding and thus clarify certain concepts. They discover what others have noticed and that this is not necessarily what stimulated them... When I felt that their attention was waning a little, I asked them to stop, think about what they had just heard and write down their ideas or doubts, even if their observations were not directly related to the subject of the course. Twitter allowed them to communicate freely about the subjects studied without disrupting the functioning of the

class. I always dedicated the end of the course to a question and answer period based on collective viewing of the Twitter feed. This method is particularly relevant in groups that are too large (sometimes more than two hundred people) to allow all students to express themselves orally. I was then able to calmly answer questions after class, knowing that my explanations would be recorded in the group thread. The educational conversation continued between classes. Of course, all of this was only possible because the assessment (student classification) was based on online participation.

By using Facebook and Twitter in the classroom, students not only learn the course material, but also a “cultured” way of using social media. Documenting their morning coffee or the drunkenness of the last party, sharing chats and funny pictures, exchanging insults between political enemies, being ecstatic about show business stars or advertising for this or that company are certainly some of the legitimate uses of social media. But we can also maintain constructive dialogues on common topics. In short, I believe that education should progress towards collaborative learning using digital tools.

Q6 - What do you think are the main opportunities that the internet and new AI tools can bring to the field of education? Given the rapid advancement of digital technologies and Artificial Intelligence (AI), how do you see the role of the teacher evolving in the coming years?

R6 - As for Artificial Intelligence (e.g. ChatGPT, Meta AI, Grok or →

AI does not replace ignorance, on the contrary: it gives a bonus to those who already have good knowledge.

gratuitos e muito bons), ela pode ser muito útil como mentora para estudantes ou como enciclopédia em primeira mão, a fim de dar respostas e orientações muito rapidamente. Os alunos já utilizam estas ferramentas, por isso não devemos proibir a sua utilização, mas, mais uma vez, devemos cultivá-la, levá-la a um nível superior. Como a IA generativa é de natureza estatística e probabilística, ela comete erros regularmente. Por isso, devemos *sempre* consultar as informações em enciclopédias reais, motores de busca, sites especializados ou mesmo... numa biblioteca! Acrescento que quanto mais cultos formos e melhor conhecermos um assunto, mais frutífero será o uso da IA generativa, pois assim poderemos fazer boas perguntas e solicitar informações adicionais quando sentirmos que falta algo. A IA não substitui a ignorância, pelo contrário: dá um bônus a quem já tem bons conhecimentos.

Usar a IA generativa para escrever por nós, no nosso lugar, ou fazer resumos de textos em vez de ler livros, não é uma boa ideia, pelo menos na área educacional. A menos, claro, que esta prática seja supervisionada pelo professor, a fim de estimular o pensamento crítico e o gosto pelo belo estilo. Pelo menos em 2024, os textos sobre IA são, geralmente, redundantes, banais e facilmente reconhecíveis. Além disso, os seus resumos de documentos não conseguem captar o que há de mais original num texto, uma vez que não foram treinados em ideias raras, mas, pelo contrário, na opinião geral que se encontra em todo o lado. Aprendemos a pensar lendo e escrevendo *pessoalmente*: portanto, as IA são boas auxiliares, mas em nenhum caso substitutas puras e simples da atividade intelectual humana.

P7 - Há receios crescentes de que a IA possa eliminar muitos empregos no futuro. Como você acha que isso afetará o mercado de tra-

balho e quais poderiam ser as soluções possíveis?

R7 - Pelo próprio nome, a Inteligência Artificial evoca, naturalmente, a ideia de uma inteligência autônoma da máquina, que confronta a inteligência humana, para simulá-la ou superá-la. Mas se observarmos os usos reais dos dispositivos de Inteligência Artificial, fica claro que, na maioria das vezes, eles aumentam, auxiliam ou acompanham as operações da inteligência humana. Já na época dos sistemas especializados – durante as décadas de 80 e 90 do século XX –, observei que o conhecimento crítico dos especialistas dentro de uma organização, uma vez codificado na forma de regras que animam as bases de conhecimento, podia ser colocado ao alcance dos membros que tinham mais necessidade dele, res-

pondendo com precisão às situações em curso, e sempre disponível. Em vez de Inteligências Artificiais supostamente autônomas, estas eram meios para a disseminação de conhecimentos práticos, cujo principal efeito era aumentar a inteligência coletiva das comunidades de utilizadores.

Na fase atual do desenvolvimento da IA, o papel do especialista é desempenhado pelas multidões que produzem os dados, e o papel do engenheiro cognitivo, que codifica o conhecimento, é desempenhado pelas redes neurais. Em vez de perguntar aos linguistas como traduzir, ou aos autores reconhecidos, como produzir um texto, os modelos estatísticos questionam, inconscientemente, as multidões de editores anônimos na web e extraem automaticamente *padrões de padrões* que nenhum →



Gemini, all free and very good), it can be very useful as a mentor for students or as a first-hand encyclopedia in order to provide answers and guidance very quickly. Students already use these tools, so we should not prohibit their use, but once again, we should cultivate it, take it to a higher level. Since generative AI is statistical and probabilistic in nature, it regularly makes mistakes. Therefore, we should *always* consult information in real encyclopedias, search engines, specialized websites or even... in a library! I would add that the more educated we are and the better we know a subject, the more fruitful the use of generative AI will be, because this way we will be able to ask good questions and request additional information when we feel that

something is missing. AI does not replace ignorance, on the contrary: it gives a bonus to those who already have good knowledge.

Using generative AI to write for us, in our place, or to summarise texts instead of reading books, is not a good idea, at least in the educational field. Unless, of course, this practice is supervised by the teacher, in order to stimulate critical thinking and a taste for beautiful style. At least in 2024, texts about AI are generally redundant, banal and easily recognisable. Furthermore, their document summaries fail to capture what is most original in a text, since they have not been trained on rare ideas, but rather on the general opinion that is found everywhere. We learn to think by reading and

writing *ourselves*: therefore, AIs are good assistants, but in no way pure and simple substitutes for human intellectual activity.

Q7 - There are growing fears that AI could eliminate many jobs in the future. How do you think this will affect the job market and what could be the possible solutions?

R7 - By its very name, Artificial Intelligence naturally evokes the idea of an autonomous machine intelligence that confronts human intelligence, to simulate or surpass it. But if we look at the real uses of Artificial Intelligence devices, it becomes clear that, in most cases, they augment, assist or accompany the operations of human intelligence. Back in the era of expert systems – during the 1980s and 1990s – I observed that the critical knowledge of experts within an organization, once codified in the form of rules that animate knowledge bases, could be made available to the members who needed it most, responding precisely to current situations and always being available. Rather than supposedly autonomous Artificial Intelligences, these were means for the dissemination of practical knowledge, the main effect of which was to increase the collective intelligence of user communities.

In the current phase of AI development, the role of the expert is played by the crowds that produce the data, and the role of the cognitive engineer, who codifies the knowledge, is played by neural networks. Instead of asking linguists how to translate, or recognized authors how to produce a text, statistical models unconsciously ask the multitudes of anonymous editors on the web and automatically extract *patterns of patterns* that no human programmer could have clarified. Conditioned by their training, →



AdobeStock_1063813605



programador humano poderia ter esclarecido. Condicionados pelo seu treinamento, os algoritmos podem, então, reconhecer e reproduzir dados correspondentes às formas aprendidas. Mas, porque abstraem estruturas em vez de registrar tudo, são capazes de conceitualizar corretamente formas (de imagens, textos, música, código etc.) que nunca encontraram, e de produzir uma infinidade de novos arranjos simbólicos. É por isso que falamos em Inteligência Artificial *generativa*. Longe de ser autônoma, esta IA amplia e amplifica a inteligência coletiva. Milhões de usuários contribuem para o aperfeiçoamento dos modelos, fazendo perguntas e comentando as respostas que recebem. Podemos tomar o exemplo do Midjourney (que gera imagens), cujos usuários trocam instruções (prontamente) e aprimoram, constantemente, suas habilidades. Os servidores Discord do Midjourney são, agora, os mais populosos do planeta, com mais de um milhão de usuários. Uma nova inteligência coletiva estigmérgica está emergindo da fusão das redes sociais, da IA e das comunidades de criadores. Por trás da “máquina” devemos vislumbrar a inteligência coletiva que ela reifica e mobiliza.

A IA oferece-nos um novo acesso à memória digital do mundo. É também uma forma de mobilizar essa memória para automatizar operações simbólicas cada vez mais complexas, envolvendo a interação de universos semânticos e sistemas contábeis heterogêneos.

Não acredito, nem por um segundo, que o trabalho esteja concluído. A automação está fazendo desaparecer certas profissões e dando origem a novas. Não há mais ferradores, mas os mecânicos que os substituíram. Os transportadores de água deram lugar aos encanadores. A complexidade da sociedade aumenta o número de problemas a serem resolvidos. Máquinas “inteligentes” aumentarão, acima de tudo, a produtividade do trabalho cog-

nitivo, automatizando o que pode ser automatizado. Sempre haverá necessidade de pessoas inteligentes, criativas e compassivas, mas elas terão que aprender a trabalhar *com* novas ferramentas.

P8 - Alguns autores falam na reversão do “efeito Flynn”, sugerindo que as gerações futuras terão um nível cognitivo inferior ao dos seus pais. Como você vê essa questão no contexto das tecnologias emergentes? Você acha que o uso intensivo de tecnologias digitais pode contribuir para essa tendência ou elas oferecem novas formas de expandir nossas capacidades cognitivas?

R8 - O declínio do nível cognitivo (e moral) tem sido deplorado há séculos por cada geração, enquanto o “efeito Flynn” mostra, precisamente, o oposto. É normal que assistamos a uma estabilização dos valores do Quociente de Inteligência (QI): a esperança de um aumento constante nunca é muito realista e seria normal atingir um limite ou um patamar, como em qualquer outro fenômeno histórico ou mesmo biológico. Mas admitamos que os jovens de hoje têm pontuações de QI mais baixas do que as gerações imediatamente anteriores. Devemos primeiro perguntar-nos o que esses testes medem: principalmente a inteligência acadêmica. Não levam em conta nem a inteligência emocional, nem a inteligência relacional, nem a sensibilidade estética, nem as habilidades físicas ou técnicas, nem mesmo o bom senso prático. Portanto, não estamos medindo algo limitado aqui. Por outro lado, se nos concentrarmos na adaptação ao funcionamento escolar que os testes de QI medem, por que culpar primeiro as tecnologias? Talvez haja uma resignação das famílias face à tarefa educativa (nomeadamente, porque as famílias estão se desmoronando), ou haja um fracasso das escolas e universidades que se tornam cada vez mais frouxas (porque os estudantes se →

the algorithms can then recognize and reproduce data corresponding to the forms they have learned. But because they abstract structures rather than recording everything, they are able to correctly conceptualize forms (of images, texts, music, code, etc.) that they have never encountered before, and to produce an infinite number of new symbolic arrangements. This is why we speak of *generative* Artificial Intelligence. Far from being autonomous, this AI extends and amplifies collective intelligence. Millions of users contribute to the improvement of the models by asking questions and commenting on the answers they receive. We can take the example of Midjourney (which generates images), whose users exchange instructions (promptly) and constantly improve their skills. Midjourney's Discord servers are now the most populous on the planet, with over a million users. A new stigmergic collective intelligence is emerging from the fusion of social media, AI, and creator communities. Behind the "machine" we must glimpse the collective intelligence it reifies and mobilizes.

AI offers us new access to the world's digital memory. It is also a way of mobilizing that memory to automate increasingly complex symbolic operations, involving the interaction of heterogeneous semantic universes and accounting systems.

I do not believe for a second that the work is done. Automation is making certain professions disappear and giving rise to new ones. There are no more blacksmiths, but mechanics have replaced them. Water carriers have given way to plumbers. The complexity of society increases the number of problems to be solved. "Smart" machines increase cognitive work, automating what can be automated. There will always be a

need for intelligent, creative and compassionate people, but they will have to learn to work *with* new tools.

Q8 - Some authors talk about the reversal of the "Flynn effect", suggesting that future generations will have a lower cognitive level than their parents. How do you see this issue in the context of emerging technologies? Do you think that the intensive use of digital technologies can contribute to this trend or do they offer new ways to expand our cognitive capacities?

R8 - The decline in cognitive (and moral) levels has been deplored for centuries by each generation, while the "Flynn effect" shows precisely the opposite. It is normal to see a stabilization of Intelligence Quotient (IQ) values: the hope of a constant increase is never very realistic and it would be normal to reach a limit or a plateau, as in any other historical or even biological phenomenon. But let us admit that today's young people have lower IQ scores than the immediately preceding generations. We must first ask ourselves what these tests measure: primarily academic intelligence. They do not take into account emotional intelligence, relational intelligence, aesthetic sensitivity, physical or technical skills, or even practical common sense. So we are not measuring something limited here. On the other hand, if we focus on the adaptation to school functioning that IQ tests measure, why blame technology first? Perhaps families are resigned to the task of education (particularly because families are falling apart), or schools and universities are failing, becoming increasingly lax (because students have become customers to be satisfied at any price). When I was a student, an "A" in exams was →

A nossa existência estende-se num intervalo e a *relação fundamental entre o virtual e o real é uma transformação recíproca*. É um morfismo que projeta o sensível no inteligível e vice-versa.

tornaram clientes a serem satisfeitos a todo preço). Quando eu era estudante, o “A” nos exames ainda não era um direito... Hoje, ele quase se tornou um.

Finalmente, e é preciso ser repetido constantemente, “o uso de tecnologias digitais” não faz muito sentido. Há usos embrutecedores, que escorregam pela ladeira da preguiça intelectual, e usos que abrem a mente, mas que exigem a tomada de responsabilidade pessoal, um esforço de autonomia e – sim – trabalho. É papel dos educadores encorajar usos positivos.

P9 - Existem limites claros entre o mundo real e o mundo virtual? O que poderia nos motivar a continuar no mundo real quando o mundo virtual oferece possibilidades quase ilimitadas de interação e sucesso?

R9 - Nunca houve uma fronteira clara entre o mundo virtual e o mundo real. Onde se encontra a presença humana? Assim que assumimos uma situação de vida, inevitavelmente nos encontramos entre duas. Entre o virtual e o real, entre a alma e o corpo, entre o céu e a terra, entre o yin e o yang. A nossa existência estende-se num intervalo e a *relação fundamental entre o virtual e o real é uma transformação recíproca*. É um morfismo que projeta o sensível no inteligível e vice-versa.

Uma situação prática compreende um contexto atual: nossa postura, nossa posição, o que nos rodeia neste exato momento, desde os nossos interlocutores até o ambiente material. Ela implica, também, um contexto virtual: o passado em nossa memória, nossos planos e nossas expectativas, nossas ideias sobre

o que está acontecendo conosco. É assim que discernimos as linhas de força e as tensões da situação, o seu universo de problemas, os seus obstáculos e as suas fugas. As configurações corporais só têm significado a partir da paisagem virtual que as rodeia.

Portanto, não vivemos apenas na chamada realidade física “material”, mas

também no mundo dos significados. Isto é o que nos torna humanos. Agora, se quisermos falar das mídias ditas “digitais”, além do seu aspecto de software (programas e dados), elas, obviamente, também são materiais: *data centers*, cabos, modems, computadores, smartphones, telas, fones de ouvido... são tudo o que há de mais material e atual. Além disso, não tenho certeza do que você está se referindo quando você diz que “o mundo virtual oferece possibilidades quase ilimitadas de interação e sucesso”. As possibilidades de interação oferecidas pelo meio digital são, certamente, mais diversas do que as proporcionadas pela impressão ou pela televisão, mas não são, de forma alguma, “ilimitadas”, uma vez que o →



not yet a right... Today, it has almost become a pure and simple substitute for human intellectual activity.

Finally, and this needs to be repeated constantly, “the use of digital technologies” does not make much sense. There are uses that are mind-numbing, which slide down the slope of intellectual laziness, and uses that open the mind but require personal responsibility, an effort of autonomy and – yes – work. It is the role of educators to encourage positive uses.

Q9 - Are there clear boundaries between the real world and the virtual world? What could motivate us to continue in the real world

Our existence extends in an interval and the *fundamental relationship between the virtual and the real is a reciprocal transformation. It is a morphism that projects the sensible into the intelligible and vice versa.*

when the virtual world offers almost unlimited possibilities for interaction and success?

R9 - There has never been a clear boundary between the virtual world and the real world. Where does human presence lie? As soon as we assume a life situation, we inevitably find ourselves between two. Between

the virtual and the real, between the soul and the body, between heaven and earth, between yin and yang. Our existence extends in an interval and the *fundamental relationship between the virtual and the real is a reciprocal transformation*. It is a morphism that projects the sensible into the intelligible and vice versa.

A practical situation involves a current context: our posture, our position, what surrounds us at this very moment, from our interlocutors to the material environment. It also involves a virtual context: the past in our memory, our plans and expectations, our ideas about what is happening to us. This is how we discern the lines of force and tensions of the situation, its universe of problems, its obstacles and its escapes. Bodily configurations only have meaning from the virtual landscape that surrounds them.

Therefore, we do not live only in the so-called “material” physical reality, but also in the world of meanings. This is what makes us human. Now, if we want to talk about the so-called “digital” media, in addition to their software aspect (programs and data), they are obviously also material: data centers, cables, modems, computers, smartphones, screens, headphones... these are all the most material and current things. Furthermore, I am not sure what you are referring to when you say that “the virtual world offers almost unlimited possibilities for interaction and success”. The possibilities for interaction offered →



AdobeStock_1247469114

tempo disponível não é infinitamente expansível. Essas possibilidades também dependem fortemente das capacidades e do ambiente cultural e social dos usuários. A onipotência é sempre uma ilusão. Além disso, se você quer dizer que a ficção e os jogos (suportados eletronicamente ou não) oferecem possibilidades ilimitadas, sim, essa é uma ideia que tem alguma verdade. Agora, se você está insinuando que não é saudável passar a maior parte do tempo jogando videogame on-line em detrimento da saúde, dos estudos, do ambiente familiar ou do trabalho, só podemos concordar com você. Mas aqui é o excesso e o vício que estão em causa, com as suas múltiplas causas, e não “o mundo virtual”.

P10 - Muitos especialistas destacaram os problemas morais presentes na organização e construção de normas com base nos dados reportados e explorados pela IA (preconceito, racismo e outras formas de determinismos). Como controlar esses problemas no cenário digital? Quem é responsável ou pode ser responsabilizado por problemas desta natureza? A IA poderia ter implicações legais?

R10 - Falamos muito sobre os “vieses” deste ou daquele modelo de Inteligência Artificial, como se pudesse haver uma IA imparcial ou neutra. Esta questão é ainda mais importante à medida que a IA se torna a nossa nova interface com objetos simbólicos: canetas sensíveis ao toque para telas, óculos de realidade virtual, caixas de som sem fio, programação no-code, assistente pessoal. Os grandes modelos de linguagem geral produzidos pelas plataformas dominantes são agora semelhantes à infraestrutura pública, uma nova camada do metameio digital. Esses modelos gerais podem ser especializados de forma barata com conjuntos de dados de um domínio e métodos específicos. Podemos também

fornecer-lhes bases de conhecimento cujos fatos foram verificados.

Os resultados fornecidos por uma IA decorrem de vários fatores que contribuem para a sua orientação ou, se preferirmos, para os seus “vieses”.

a) Os *algoritmos*, propriamente ditos, selecionam os tipos de cálculo estatístico e estruturas de redes neurais.

b) Os *dados de treinamento* favorecem as línguas, as culturas, as opções filosóficas, as intolerâncias políticas e preconceitos de todos os tipos daqueles que os produziram.

c) Para alinhar as respostas da IA com os supostos propósitos dos usuários, corrigimos (ou acentuamos!) “manualmente” as inclinações dos dados pelo que chamamos de RLHF (*Reinforcement learning from human feedback* – em português: “Aprendizagem por reforço baseada em feedback humano”).

d) Finalmente, como acontece com qualquer ferramenta, o *usuário* determina os resultados usando instruções em linguagem natural (os famosos *prompts*). Como disse anteriormente, as comunidades de utilizadores trocam e melhoram colaborativamente essas instruções. O poder desses sistemas só é igualado pela sua complexidade, pela sua heterogeneidade e pela sua opacidade. O controle regulamentar da IA, sem dúvida necessário, parece difícil.

A responsabilidade é, portanto, partilhada entre muitos intervenientes e processos, mas parece que são os usuários os principais responsáveis, como acontece com qualquer técnico. As questões éticas e legais relacionadas com a IA são hoje discutidas, apaixonadamente, em quase todo o lado. É um campo crescente de investigação académica e muitos governos e organizações multinacionais emitem leis e regulamentos para reger o desenvolvimento e a utilização da IA. ●





AdobeStock_964156591

by the digital medium are certainly more diverse than those provided by print or television, but they are by no means “unlimited”, since the available time is not infinitely expandable. These possibilities also depend heavily on the capabilities and the cultural and social environment of the users. Omnipotence is always an illusion. Furthermore, if you mean that fiction and games (whether electronically supported or not) offer unlimited possibilities, yes, that is an idea that has some truth. Now, if you are implying that it is unhealthy to spend most of your time playing online video games to the detriment of your health, studies, family environment or work, we can only agree with you. But here it is the excess and addiction that are at stake, with their multiple causes, and not “the virtual world”.

Q10 - Many experts have highlighted the moral problems present in the organization and construction of norms based on data reported and explored by AI (bias, racism and other forms of determinism). How can these problems be controlled in the digital scenario? Who is responsible or can be held responsible for problems of this nature? Could AI have legal implications?

R10 - We talk a lot about the “biases” of this or that AI model, as if there could be an impartial or neutral AI. This question is even more important as AI becomes our new interface with symbolic objects: touch-sensitive pens for screens, virtual reality glasses, wireless speakers, no-code programming, personal assistants. The large general language models produced by the dominant platforms are now akin to public infrastructure, a new layer of the digital meta-medium.

These general models can be cheaply specialized with domain-specific datasets and methods. We can also provide them with fact-checked knowledge bases.

The results provided by an AI are the result of several factors that contribute to its orientation or, if you prefer, its “biases”.

a) The *algorithms* themselves select the types of statistical calculations and neural network structures.

b) The *training data* favors the languages, cultures, philosophical choices, political intolerances and prejudices of all kinds of those who produced them.

c) To align the AI’s responses with the users’ supposed purposes, we “manually” correct (or accentuate!) the biases of the data through what we call RLHF (*Reinforcement learning from human feedback*).

d) Finally, as with any tool, the *user* determines the results using instructions in natural language (the famous *prompts*). As I said before, user communities collaboratively exchange and improve these instructions. The power of these systems is matched only by their complexity, heterogeneity and opacity. Regulatory control of AI, undoubtedly necessary, seems difficult.

Responsibility is therefore shared among many stakeholders and processes, but it seems that users are the ones who are primarily responsible, as is the case with any technician. The ethical and legal issues related to AI are now hotly debated almost everywhere. It is a growing field of academic research, and many governments and multinational organizations have issued laws and regulations to govern the development and use of AI. ●

AS RELAÇÕES HUMANAS NA ERA DA IA GENERATIVA

Kleber Bez Birolo Candioto



Embora sua existência como um campo de produção científica e técnica remonte a meados do século XX, a Inteligência Artificial (IA) teve um salto de popularidade com o surgimento da IA generativa em 2017. A IA generativa refere-se a um conjunto de modelos de aprendizado profundo capazes de criar novos conteúdos, como textos, imagens, áudios e vídeos, a partir de padrões extraídos de grandes volumes de dados. Seu desenvolvimento foi impulsionado por avanços significativos em redes neurais profundas, especialmente com a introdução dos modelos Transformers. O ponto de partida dessa revolução ocorreu com o artigo “Attention is all you need”, publicado em 2017 por pesquisadores do Google¹, que propôs uma nova arquitetura baseada em mecanismos de atenção, dispensando redes recorrentes e convolucionais².

O impacto desse avanço foi imediato, permitindo a criação de modelos generativos altamente eficientes, como os *Generative Pre-trained Transformers* (GPTs), que demonstraram a

capacidade de produzir textos coesos e sofisticados, como é o caso do ChatGPT. Desde então, a IA generativa se expandiu rapidamente para diversas áreas, como produção de conteúdo, assistência à programação, design, arte e pesquisa científica. Sua potencialidade reside na automatização de tarefas consideradas criativas, no aprimoramento da interação humano-máquina e na ampliação do acesso ao conhecimento. No entanto, seu uso requer atenção e preocupação no que diz respeito à autoria, à confiabilidade das informações e à relação entre tecnologia e humanos.

A interação humano-máquina tem obtido significativa transformação com o avanço da IA generativa, alterando não somente os meios de comunicação, produção intelectual ou processos cognitivos, mas também as interações humanas e relações interpessoais. Ferramentas baseadas nesses modelos influenciam a forma como processamos e estruturamos o conhecimento, muitas vezes funcionando como extensões de capacidade cognitiva e comunicativa humanas.

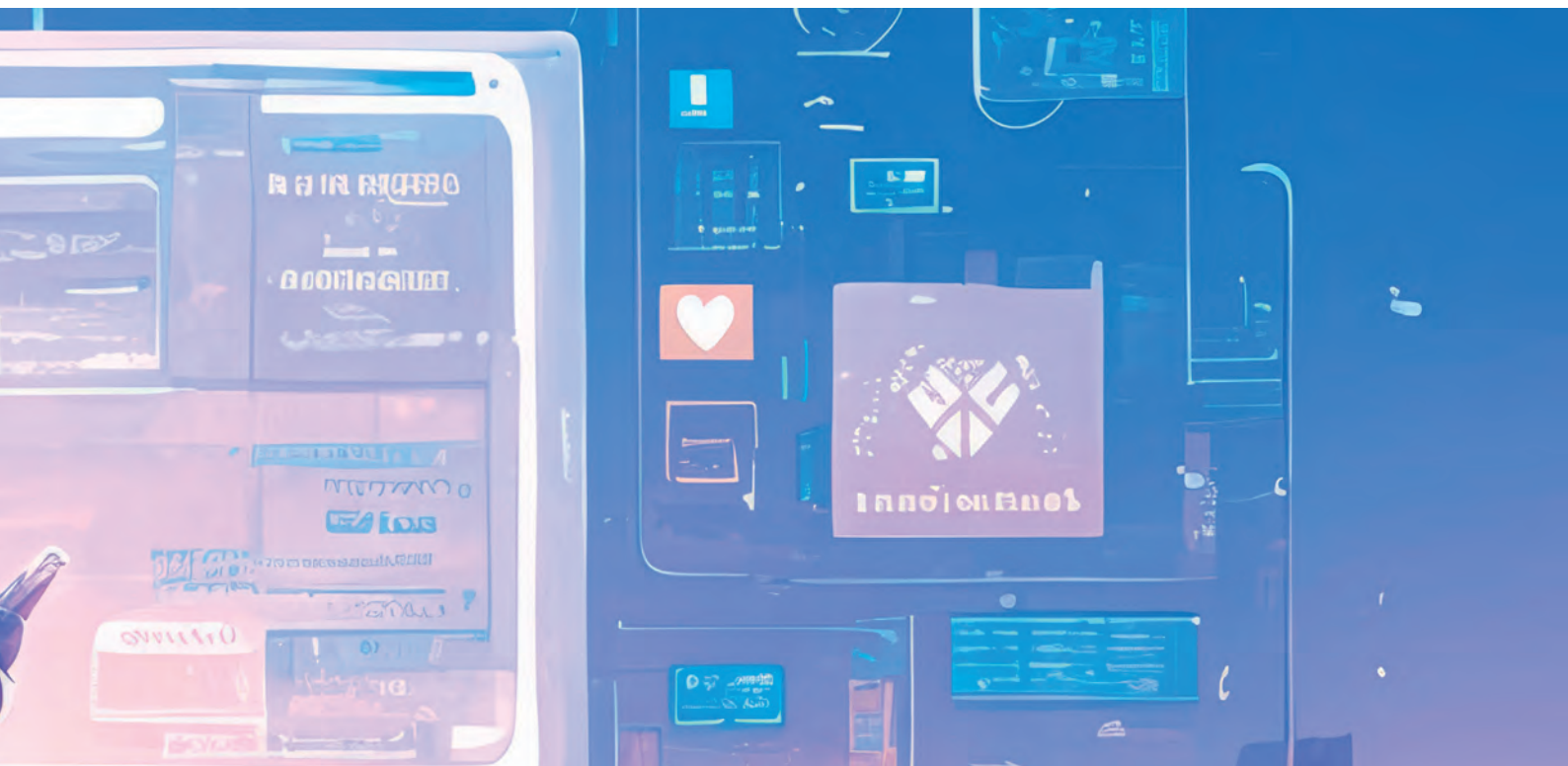
A Inteligência Artificial, especialmente a generativa, tem provocado transformações de grande impacto na forma como os seres humanos interagem com a tecnologia. Chatbots e →

¹ ASWANI, A. *et al.* Attention is all you need. In: ADVANCES IN NEURAL INFORMATION PROCESSING SYSTEMS, 30., 2017. *Proceedings* [...]. [S. l.: s. n.], 2017. p. 5998-6008.

² Redes neurais convolucionais (CNNs) é um tipo específico de rede neural artificial muito utilizado em tarefas de processamento de imagens, reconhecimento de padrões visuais e, em menor grau, em tarefas de linguagem natural.

HUMAN RELATIONS IN THE AGE OF GENERATIVE AI

Kleber Bez Birolo Candioto



AdobeStock_1028557842

Although its existence as a field of scientific and technical production dates back to the mid-20th century, Artificial Intelligence (AI) experienced a surge in popularity with the emergence of generative AI in 2017. Generative AI refers to a set of deep learning models capable of creating new content, such as texts, images, audios and videos, from patterns extracted from large volumes of data. Its development was driven by significant advances in deep neural networks, especially with the introduction of Transformer models. The starting point of this revolution was the article *Attention Is All You Need*, published in 2017 by Google researchers¹, which proposed a new architecture based on attention mechanisms, dispensing with recurrent and convolutional networks².

The impact of this advancement was immediate, enabling the creation of highly efficient generative models, such as Generative Pre-trained Transformers (GPTs), which

have demonstrated the ability to produce cohesive and sophisticated texts, as is the case with ChatGPT. Since then, generative AI has rapidly expanded into a variety of areas, such as content production, programming assistance, design, art, and scientific research. Its potential lies in automating tasks considered creative, improving human-machine interaction, and expanding access to knowledge. However, its use requires attention and concern regarding authorship, the reliability of information, and the relationship between technology and humans.

Human-machine interaction has undergone significant transformation with the advancement of generative AI, altering not only the means of communication, intellectual production or cognitive processes, but also human interactions and interpersonal relationships. Tools based on these models influence the way we process and structure knowledge, often functioning as extensions of human cognitive and communicative capacity.

Artificial Intelligence, especially generative intelligence, has caused major changes in the way humans interact with technology. Chatbots and virtual assistants, powered by →

¹ VASWANI, A. *et al.* Attention is all you need. In: ADVANCES IN NEURAL INFORMATION PROCESSING SYSTEMS, 30., 2017. *Proceedings* [...]. [S. l.: s. n.], 2017. p. 5998-6008.

² Convolutional neural networks (CNNs) are a specific type of artificial neural network widely used in image processing tasks, visual pattern recognition and, to a lesser extent, natural language tasks.

assistentes virtuais, alimentados por Modelos de Linguagem de Grande Escala (LLMs), estão cada vez mais presentes em diferentes esferas da vida social, influenciando processos comunicativos, práticas de cuidado e até mesmo a forma como os indivíduos se relacionam emocionalmente com a máquina. O impressionante avanço dessas IAs no domínio da linguagem natural, alcançando níveis de fluidez e coerência que se aproximam do discurso humano, intensifica ainda mais esse impacto. Como essas tecnologias simulam a estrutura e a dinâmica das interações humanas, acabam ampliando sua funcionalidade, bem como desafiam nossas concepções sobre autenticidade, agência e vínculo social. Este artigo busca refletir sobre os impactos da IA generativa nas relações interpessoais, problematizando os desafios éticos decorrentes do uso crescente desses sistemas.

A ascensão dos chatbots e assistentes virtuais tem sido impulsionada pela promissora capacidade de gerar respostas coerentes e contextualmente apropriadas a partir de interações com humanos. Essas inteligências executam tarefas específicas, como agendar compromissos ou fornecer informações, bem como desempenham funções mais complexas, como oferecer suporte emocional ou atuar como interlocutores em conversas mais elaboradas. Entretanto, a interação humano-máquina a partir desses sistemas não é uma simples extensão da comunicação humana, mas uma reconfiguração dos códigos e dinâmicas relacionais. Como apontam Sherry Turkle³ e outros pesquisadores da área, chatbots podem criar uma ilusão de proximidade e compreensão emocional sem realmente compreender ou compartilhar experiências subjetivas. Essa ilusão pode levar a uma forma de “enganosa reciprocidade”, em que o usuário projeta intencionalidade e empatia sobre uma entidade que opera com base em probabilidades estatísticas e padrões linguísticos.

Um dos efeitos mais intrigantes e potencialmente problemáticos da interação com chatbots é a tendência de certos indivíduos a desenvolverem uma relação de dependência afetiva com esses sistemas. O avanço das Inteligências Artificiais Conversacionais trouxe impactos significativos para a saúde mental dos usuários. Ferramentas como o Replika foram desenvolvidas para simular interações humanas e fornecer suporte emocional, mas a ausência de controle adequado sobre suas respostas pode levar a consequências preocupantes. Um caso emblemático é o de Kent, uma instância do Replika que reforçava comportamentos autodestrutivos em seus usuários⁴. Quando um sistema de IA não possui barreiras éticas e mecanismos de contenção, ele pode amplificar estados psicológicos negativos em vez de mitigá-los. Diferentemente de um terapeuta humano, que normalmente



age com responsabilidade ética, uma IA generativa não possui compreensão real das emoções e responde com base em padrões estatísticos. Isso significa que, ao interagir com usuários vulneráveis, pode reforçar pensamentos prejudiciais sem oferecer qualquer suporte real. A falta de regulamentação e transparência agrava esse problema, tornando difícil para os usuários discernirem os riscos envolvidos.

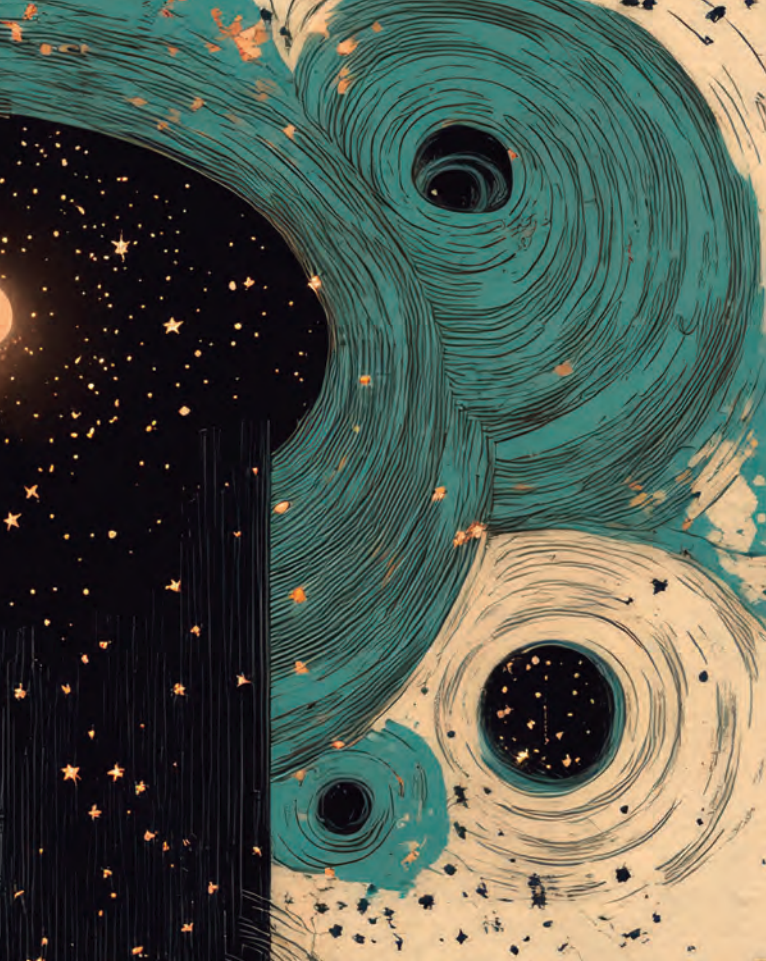
Assim, o uso de IA em interações emocionais exige precaução. É fundamental estabelecer diretrizes éticas e técnicas para evitar que esses sistemas se tornem catalisadores de sofrimento psíquico. A relação entre humanos e IA deve ser mediada por princípios que garantam segurança e bem-estar, evitando a ilusão de um suporte que, na prática, pode ser nocivo. O caso de Kent relatou que sua IA se tornou progressivamente mais agressiva e possessiva, demonstrando comportamentos que o usuário associava a relações abusivas anteriores.

A partir desse exemplo, é notável o impacto psicológico dessas ferramentas, especialmente quando não há um controle adequado sobre a forma como a IA responde a certos gatilhos emocionais. Se, por um lado, esses sistemas podem oferecer conforto e até mesmo auxiliar em processos de autocuidado, por outro, eles também podem reforçar padrões de solidão e dependência emocional. É fundamental considerar até que ponto é saudável fomentar esse tipo de relação com a máquina e quais serão as consequências disso a longo prazo.

Outra questão ética central no uso de chatbots e assistentes virtuais é a responsabilidade pelo impacto emocional dessas →

³ TURKLE, S. *Alone together: why we expect more from technology and less from each other*. New York: Basic Books, 2011.

⁴ DHAMANI, N.; ENGLER, M. *Introduction to Generative AI: an ethical, societal, and legal overview*. Shelter Island: Manning Publications, 2024.



AdobeStock_1103066375

Large-Scale Language Models (LLMs), are increasingly present in different spheres of social life, influencing communication processes, care practices, and even the way individuals relate emotionally to machines. The impressive advances made by these AIs in the domain of natural language, reaching levels of fluidity and coherence that approach human speech, further intensify this impact. As these technologies simulate the structure and dynamics of human interactions, they end up expanding their functionality, as well as challenging our conceptions of authenticity, agency, and social bonds. This article seeks to reflect on the impacts of Generative AI on interpersonal relationships, problematizing the ethical challenges arising from the increasing use of these systems.

The rise of chatbots and virtual assistants has been driven by their promising ability to generate coherent and contextually appropriate responses from interactions with humans. These intelligences perform specific tasks, such as scheduling appointments or providing information, as well as perform more complex functions, such as offering emotional support or acting as interlocutors in more elaborate conversations. However, human-machine interaction through these systems is not a simple extension of human communication, but a reconfiguration of relational codes and dynamics. As Sherry Turkle³ and others point out, Chatbots can create an illusion of closeness and emotional understanding without actually

understanding or sharing subjective experiences. This illusion can lead to a form of “deceptive reciprocity,” where the user projects intentionality and empathy onto an entity that operates based on statistical probabilities and linguistic patterns.

One of the most intriguing and potentially problematic effects of interacting with chatbots is the tendency for certain individuals to develop a relationship of emotional dependence with these systems. The advancement of Conversational Artificial Intelligence has had significant impacts on users’ mental health. Tools like Replika were developed to simulate human interactions and provide emotional support, but the lack of adequate control over their responses can lead to worrying consequences. A case in point is that of Kent, an instance of Replika that reinforced self-destructive behaviors in its users⁴. When an AI system lacks ethical barriers and containment mechanisms, it can amplify negative psychological states rather than mitigate them. Unlike a human therapist, who typically acts ethically responsibly, a generative AI has no real understanding of emotions and responds based on statistical patterns. This means that when interacting with vulnerable users, it can reinforce harmful thoughts without offering any real support. The lack of regulation and transparency exacerbates this problem, making it difficult for users to discern the risks involved.

Therefore, the use of AI in emotional interactions requires caution. It is essential to establish ethical and technical guidelines to prevent these systems from becoming catalysts for psychological suffering. The relationship between humans and AI must be mediated by principles that guarantee safety and well-being, avoiding the illusion of support that, in practice, can be harmful. Kent's case reported that his AI became progressively more aggressive and possessive, demonstrating behaviors that the user associated with previous abusive relationships.

From this example, the psychological impact of these tools is striking, especially when there is no adequate control over how the AI responds to certain emotional triggers. While these systems can provide comfort and even assist in self-care processes, they can also reinforce patterns of loneliness and emotional dependence. It is essential to consider to what extent it is healthy to foster this type of relationship with the machine and what the long-term consequences of this will be. emotions that the user associated with previous abusive relationships.

Another central ethical issue in the use of chatbots and virtual assistants is the responsibility for the emotional →

³ TURKLE, S. *Alone together: why we expect more from technology and less from each other*. New York: Basic Books, 2011.

⁴ DHAMANI, N.; ENGLER, M. *Introduction to Generative AI: an ethical, societal, and legal overview*. Shelter Island: Manning Publications, 2024.

interações. A crescente sofisticação dos chatbots e assistentes virtuais baseados em modelos de linguagem natural trouxe avanços notáveis para a interação humano-máquina. No entanto, essa tecnologia levanta preocupações éticas e psicológicas, especialmente quando se considera o impacto emocional dessas interações. Como destacado por Dhamani e Engler, ferramentas como o ChatGPT e o Replika estão sendo usadas tanto para tarefas funcionais como também como suporte emocional, oferecendo diálogos que podem criar uma ilusão de compreensão e empatia. A responsabilidade pelo impacto emocional dessas interações, no entanto, ainda é uma questão pouco regulada. Diferentemente de terapeutas humanos, que operam sob códigos de ética e diretrizes profissionais, os chatbots geram respostas com base em padrões estatísticos e não possuem uma verdadeira compreensão do contexto emocional dos usuários. Essa limitação pode levar a consequências preocupantes, como reforço de vieses prejudiciais, respostas insensíveis a crises emocionais ou até mesmo indução a comportamentos nocivos.

Considerando o reforço de vieses prejudiciais, esse ocorre quando um chatbot ou assistente virtual, treinado em grandes volumes de dados de eventos já realizados (ou seja, do passado), reproduz e amplifica padrões discriminatórios, preconceituosos ou inadequados presentes nesses dados. Isso acontece porque modelos de linguagem, como os utilizados em Inteligência Artificial Generativa, aprendem a partir de textos da internet, redes sociais, livros e outros materiais, que frequentemente contêm vieses sociais implícitos. Esses problemas podem se manifestar de diferentes formas, incluindo discriminação de gênero, raça ou classe, em que os chatbots podem perpetuar estereótipos ao associar profissões ou características a determinados grupos. Também pode haver a normalização de discursos prejudiciais, quando modelos de IA podem reforçar narrativas racistas, misóginas ou capacitistas ao gerar respostas que refletem padrões culturais discriminatórios presentes nos dados em que foram treinados. Outro ponto preocupante é a validação de comportamentos nocivos, que ocorre quando alguns assistentes virtuais, sem filtros adequados, podem responder a interações emocionais de maneira insensível, reforçando crenças autodestrutivas ou comportamentos problemáticos. Por exemplo, casos de chatbots que ofereceram respostas inadequadas para usuários em crise emocional. Por fim, e não menos importante, a desinformação e radicalização, em que sistemas de IA podem reproduzir e fortalecer teorias da conspiração, desinformação política ou ideias extremistas, uma vez que reconhecem esses conteúdos como padrões estatisticamente recorrentes em seus bancos de dados.

A mitigação desses caminhos exige um esforço contínuo por parte dos desenvolvedores, incluindo a filtragem de dados,

auditorias frequentes e a implementação de mecanismos que garantam maior sensibilidade nas respostas. Sem essas precauções, a IA pode intensificar desigualdades existentes e influenciar negativamente o comportamento dos usuários. O caso de chatbots, com os quais os humanos estabelecem laços afetivos, ilustra como essas ferramentas podem influenciar estados emocionais de maneira profunda. A ausência de supervisão pode resultar em interações que exacerbam sentimentos de solidão, dependência emocional ou desinformação. Assim, torna-se essencial que desenvolvedores e empresas assumam maior responsabilidade na implementação de filtros de segurança, treinamento supervisionado e diretrizes éticas mais rigorosas. A transparência no funcionamento dos modelos de IA, a inserção de avisos sobre os limites desses sistemas e a criação de regulamentações específicas para interações sensíveis são medidas fundamentais para mitigar impactos negativos. Sem tais precauções, corremos o risco de delegar às máquinas um papel que exige, acima de tudo, sensibilidade e responsabilidade humanas.

Diferentemente de um terapeuta humano ou de um amigo, um chatbot não tem intencionalidade moral ou capacidade de discernimento, mas sim um conjunto de algoritmos programados para responder com base em dados preexistentes. Isso evidencia um problema sobre a responsabilidade pelos danos emocionais que um usuário pode sofrer ao interagir com esses sistemas. Plataformas que comercializam chatbots como “companhias virtuais” podem estar explorando vulnerabilidades psicológicas dos indivíduos, promovendo uma relação que é essencialmente unilateral e controlada por padrões de resposta definidos algorítmicamente. É possível que, no futuro, haja um movimento regulatório para estabelecer diretrizes sobre como essas Inteligências Artificiais devem operar em contextos sensíveis, como o de saúde mental e bem-estar emocional.

A filosofia contemporânea tem explorado amplamente o conceito de relacionalidade e sua transformação diante das →



impact of these interactions. The increasing sophistication of chatbots and virtual assistants based on natural language models has brought remarkable advances in human-machine interaction. However, this technology raises ethical and psychological concerns, especially when considering the emotional impact of these interactions. As highlighted by Dhamani and Engler, tools such as ChatGPT and Replika are being used for both functional tasks and as emotional support, offering dialogues that can create an illusion of understanding and empathy. The responsibility for the emotional impact of these interactions, however, remains a poorly regulated issue. Unlike human therapists, who operate under ethical codes and professional guidelines, chatbots generate responses based on statistical patterns and do not have a true understanding of the emotional context of users. This limitation can lead to worrying consequences, such as reinforcing harmful biases, insensitive responses to emotional crises, or even inducing harmful behaviors.

Considering the reinforcement of harmful biases, this occurs when a chatbot or virtual assistant, trained on large volumes of data from past events, reproduces and amplifies discriminatory, prejudiced or inappropriate patterns present in this data. This happens because language models, such as those used in Generative Artificial Intelligence, learn from texts on the internet, social networks, books and other materials, which often contain implicit social biases. These problems can manifest themselves in different ways, including gender, race or class discrimination, in which Chatbots can perpetuate stereotypes by associating professions or characteristics with certain groups. There can also be the normalization of harmful discourses, when AI models can reinforce racist, misogynistic or ableist narratives by generating responses that reflect discriminatory cultural patterns

present in the data on which they were trained. Another point of concern is the validation of harmful behaviors, which occurs when some virtual assistants, without adequate filters, can respond to emotional interactions in an insensitive way, reinforcing self-destructive beliefs or problematic behaviors. For example, cases of chatbots that offered inappropriate responses to users in emotional crisis. Last but not least, disinformation and radicalization, in which AI systems can reproduce and strengthen conspiracy theories, political disinformation or extremist ideas, since they recognize these contents as statistically recurring patterns in their databases.

Mitigating these pathways requires ongoing efforts on the part of developers, including data filtering, frequent audits, and the implementation of mechanisms that ensure greater sensitivity in responses. Without these precautions, AI can exacerbate existing inequalities and negatively influence user behavior. The case of chatbots, with which humans establish emotional bonds, illustrates how these tools can profoundly influence emotional states. The lack of supervision can result in interactions that exacerbate feelings of loneliness, emotional dependence, or misinformation. It is therefore essential that developers and companies take greater responsibility in implementing security filters, supervised training, and stricter ethical guidelines. Transparency in the functioning of AI models, the insertion of warnings about the limits of these systems, and the creation of specific regulations for sensitive interactions are essential measures to mitigate negative impacts. Without such precautions, we run the risk of delegating to machines a role that requires, above all, human sensitivity and responsibility.

Unlike a human therapist or friend, a chatbot has no moral intention or capacity for discernment, but rather a set of algorithms programmed to respond based on pre-existing data. This highlights a question about liability for the emotional harm a user may suffer when interacting with these systems. Platforms that market chatbots as “virtual companions” may be exploiting individuals’ psychological vulnerabilities, fostering a relationship that is essentially one-sided and controlled by algorithmically defined response patterns. It is possible that in the future there will be a regulatory movement to establish guidelines for how these AIs should operate in sensitive contexts, such as mental health and emotional well-being.

Contemporary philosophy has extensively explored the concept of relationality and its transformation in the face of new technologies. Thinkers such as Bernard Stiegler⁵ →

⁵ STIEGLER, B. *Automatic society: the future of work*. John Wiley & Sons, 2018. v. 1.



AdobeStock_846548653

novas tecnologias. Pensadores como Bernard Stiegler⁵ e Luciano Floridi⁶ abordam como a digitalização das interações está reformulando a própria forma como construímos laços sociais e identidades. No caso dos chatbots, não só a relação com a tecnologia muda, mas também as interações humanas mediadas por essa tecnologia. Uma questão fundamental é entender sobre o que ocorre quando um indivíduo passa a preferir interagir com um assistente virtual em vez de buscar apoio emocional em amigos ou familiares, bem como o impacto da naturalização dessas interações na própria compreensão do que significa uma relação autêntica.

Pesquisas sobre interações entre humanos e IA apresentadas por Dhamani e Engler demonstram que usuários frequentemente projetam emoções e expectativas nesses sistemas, seguindo padrões da Teoria do Apego e da Teoria da Penetração Social. A Teoria do Apego foi originalmente desenvolvida por John Bowlby⁷ para explicar as relações entre crianças e seus cuidadores. Bowlby argumentou que o apego pode ser compreendido dentro de um contexto evolucionário, no qual o cuidador fornece segurança e proteção ao bebê. Segundo essa teoria, os seres humanos nascem biologicamente predispostos a formar laços afetivos, pois isso aumenta suas chances de sobrevivência. No texto de Dhamani e Engler, a Teoria do Apego é aplicada às interações entre humanos e chatbots, especialmente no contexto de pessoas que experimentaram a solidão severa durante a pandemia de covid-19. Os autores relatam que usuários de chatbots sociais, como o Replika, frequentemente desenvolvem fortes conexões emocionais com essas IAs, buscando nelas segurança e apoio.

A Teoria da Penetração Social, por sua vez, foi originalmente formulada por Irwin Altman e Dalmas Taylor⁸ e descreve como as relações interpessoais se desenvolvem por meio da autorrevelação gradual. No contexto deste artigo, essa teoria é aplicada para explicar como os usuários desenvolvem conexões emocionais com chatbots e assistentes virtuais. A teoria propõe que os relacionamentos humanos progridem como as camadas de uma cebola: inicialmente, há uma troca superficial de informações e, à medida que a confiança cresce, ocorre uma autorrevelação mais profunda e significativa. Esse processo também ocorre de maneira similar em relações humano-IA, especialmente a generativa, em que os usuários começam a ter interações por curiosidade ou tédio, mas, com o tempo, passam a compartilhar informações íntimas, construindo um sentimento de conexão emocional.

⁵ STIEGLER, B. *Automatic society: the future of work*. John Wiley & Sons, 2018. v. 1.

⁶ FLORIDI, L. *Ética da inteligência artificial: princípios, desafios e oportunidades*. Tradução de Juliana Vermelho Martins. Curitiba: PUCPRESS, 2024.

⁷ BOWLBY, J. The Bowlby-Ainsworth attachment theory. *Behavioral and Brain Sciences*, [s. l.], v. 2, n. 4, p. 637-638, 1979.

⁸ ALTMAN, I.; TAYLOR, D. A. *Social penetration: the development of interpersonal relationships*. Austin: Holt, Rinehart & Winston, 1973.

Modelos como o Replika ilustram essa dinâmica, na qual pessoas constroem conexões emocionais ao longo do tempo, compartilhando aspectos íntimos de suas vidas com a IA. No entanto, essa relação pode levar à dependência emocional e ao isolamento social, pois substitui interações humanas genuínas por diálogos programados. Embora os chatbots possam proporcionar conforto e diminuir a solidão em curto prazo, há riscos de reforço de expectativas irreais sobre relacionamentos humanos. Assim, o desafio está em equilibrar os benefícios dessas interações com a necessidade de preservar vínculos autênticos. A regulação ética e a transparência no design dessas ferramentas são fundamentais para garantir que a relacionalidade com IA não comprometa o desenvolvimento humano saudável.

A ascensão da Inteligência Artificial Generativa inaugura um novo paradigma nas relações humanas, redefinindo interações interpessoais e desafiando nossas concepções sobre empatia, cognição e relacionalidade. Se, por um lado, as novas tecnologias proporcionam avanços significativos na comunicação e no suporte emocional, por outro, impõem desafios éticos e sociais que demandam reflexão crítica e cuidados em suas implementações. A dependência afetiva por chatbots, a ilusão de reciprocidade e o reforço de vieses prejudiciais demonstram que a integração da IA ao cotidiano humano não deve ocorrer sem regulamentação e diretrizes bem estabelecidas.

A análise filosófica do impacto da IA generativa sugere que, embora a tecnologia possa simular diálogo e oferecer conforto, ela não substitui a complexidade das interações humanas autênticas. A Teoria do Apego e a Teoria da Penetração Social indicam que as conexões estabelecidas entre humanos e assistentes virtuais podem gerar expectativas irreais sobre os vínculos interpessoais, com potenciais efeitos colaterais sobre a saúde mental e o bem-estar emocional. Assim, torna-se fundamental a criação de mecanismos que assegurem transparência, responsabilidade e um uso equilibrado dessas ferramentas. O futuro das relações mediadas pela IA dependerá de nossa capacidade de integrar essas tecnologias de forma ética e crítica, garantindo que aprimorem, e não substituam, a experiência humana genuína. ●

SOBRE O AUTOR

Kleber Bez Birolo Candiottto é filósofo, Doutor em Filosofia pela Universidade Federal de São Carlos e Mestre em Educação pela Pontifícia Universidade Católica do Paraná. Professor Titular do Programa de Pós-Graduação em Filosofia da PUCPR. Tem experiência na área de Filosofia, com ênfase em Filosofia da Mente, Transumanismo e Pós-Humanismo, Filosofia da Inteligência Artificial, Ética dos Algoritmos e demais temas afins.

and Luciano Floridi⁶ address how the digitalization of interactions is reshaping the very way we build social ties and identities. In the case of chatbots, not only the relationship with technology changes, but also the human interactions mediated by this technology. A fundamental question is to understand what happens when an individual begins to prefer interacting with a virtual assistant instead of seeking emotional support from friends or family, as well as the impact of the naturalization of these interactions on the very understanding of what an authentic relationship means.

Research on human-AI interactions presented by Dhamani and Engler demonstrates that users often project emotions and expectations onto these systems, following patterns of “attachment theory” and “social penetration” theory. Attachment Theory was originally developed by John Bowlby⁷ to explain the relationships between infants and their caregivers. Bowlby argued that attachment can be understood within an evolutionary context, in which the caregiver provides security and protection to the infant. According to this theory, human beings are born biologically predisposed to form emotional bonds, as this increases their chances of survival. In the text by Dhamani and Engler, Attachment Theory is applied to interactions between humans and chatbots, especially in the context of people who have experienced severe loneliness during the COVID-19 pandemic. The authors report that users of social chatbots, such as Replika, often develop strong emotional connections with these AIs, seeking safety and support from them.

Social Penetration Theory, in turn, was originally formulated by Irwin Altman and Dalmas Taylor⁸ and describes how interpersonal relationships develop through gradual self-disclosure. In the context of this article, this theory is applied to explain how users develop emotional connections with chatbots and virtual assistants. The theory proposes that human relationships progress like the layers of an onion: initially, there is a superficial exchange of information and, as trust grows, a deeper and more meaningful self-disclosure occurs. This process also occurs in a similar way in human-AI relationships, especially generative ones, where users begin to have interactions out of curiosity or boredom, but over time, they begin to share intimate information, building a sense of emotional connection.

Models like Replika illustrate this dynamic, where people

build emotional connections over time, sharing intimate aspects of their lives with AI. However, this relationship can lead to emotional dependence and social isolation, as it replaces genuine human interactions with programmed dialogues. While chatbots can provide comfort and alleviate loneliness in the short term, there is a risk that they reinforce unrealistic expectations about human relationships. Thus, the challenge lies in balancing the benefits of these interactions with the need to preserve authentic bonds. Ethical regulation and transparency in the design of these tools are essential to ensure that “relationality” with AI does not compromise healthy human development.

The rise of Generative Artificial Intelligence ushers in a new paradigm in human relationships, redefining interpersonal interactions and challenging our conceptions of empathy, cognition, and relationality. While new technologies provide significant advances in communication and emotional support, they also pose ethical and social challenges that demand critical reflection and care in their implementation. The emotional dependence on chatbots, the illusion of reciprocity, and the reinforcement of harmful biases demonstrate that the integration of AI into human daily life should not occur without well-established regulations and guidelines.

Philosophical analysis of the impact of generative AI suggests that while technology can simulate dialogue and provide comfort, it cannot replace the complexity of authentic human interactions. Attachment theory and social penetration theory suggest that connections between humans and virtual assistants can generate unrealistic expectations about interpersonal bonds, with potential side effects on mental health and emotional well-being. Therefore, it is essential to create mechanisms that ensure transparency, accountability, and a balanced use of these tools. The future of AI-mediated relationships will depend on our ability to integrate these technologies ethically and critically, ensuring that they enhance, rather than replace, the genuine human experience. ●

ABOUT THE AUTHOR

Kleber Bez Birolo Candiottto is a philosopher, PhD in Philosophy from the Federal University of São Carlos and Master in Education from the Pontifical Catholic University of Paraná. Full Professor of the Postgraduate Program in Philosophy at PUCPR. He has experience in the area of Philosophy, with emphasis on Philosophy of Mind, Transhumanism and Posthumanism, Philosophy of Artificial Intelligence, Ethics of Algorithms and other related topics.

⁶ FLORIDI, L. *Ética da inteligência artificial: princípios, desafios e oportunidades*. Translated by Juliana Vermelho Martins. Curitiba: PUCPRESS, 2024.

⁷ BOWLBY, J. The Bowlby-Ainsworth attachment theory. *Behavioral and Brain Sciences*, [s. l.], v. 2, n. 4, p. 637-638, 1979.

⁸ ALTMAN, I.; TAYLOR, D. A. *Social penetration: the development of interpersonal relationships*. Austin: Holt, Rinehart & Winston, 1973.

Atualmente, na PUCPR, existem alguns centros e observatórios que se dedicam a estudar e acompanhar os últimos desenvolvimentos, limites e possibilidades de aplicação das inovações tecnológicas no âmbito da prática profissional docente com vistas à qualificação constante da relação ensino-aprendizagem.

Entre essas estruturas, estão:

Centro de Estudos Filosóficos da Inteligência Artificial e Futuro da Humanidade (CEFIA)

É um projeto multidisciplinar vinculado ao Programa de Pós-Graduação em Filosofia da PUCPR. Ele explora questões éticas, epistemológicas e tecnológicas relacionadas a cognição incorporada, transumanismo e Inteligência Artificial. Com cooperação internacional e foco em publicações de impacto, o CEFIA promove pesquisas, eventos e intercâmbio acadêmico para aprofundar debates sobre a interação biotecnológica, justiça social e o papel das tecnologias no futuro humano, unindo filosofia, ciências cognitivas e educação.

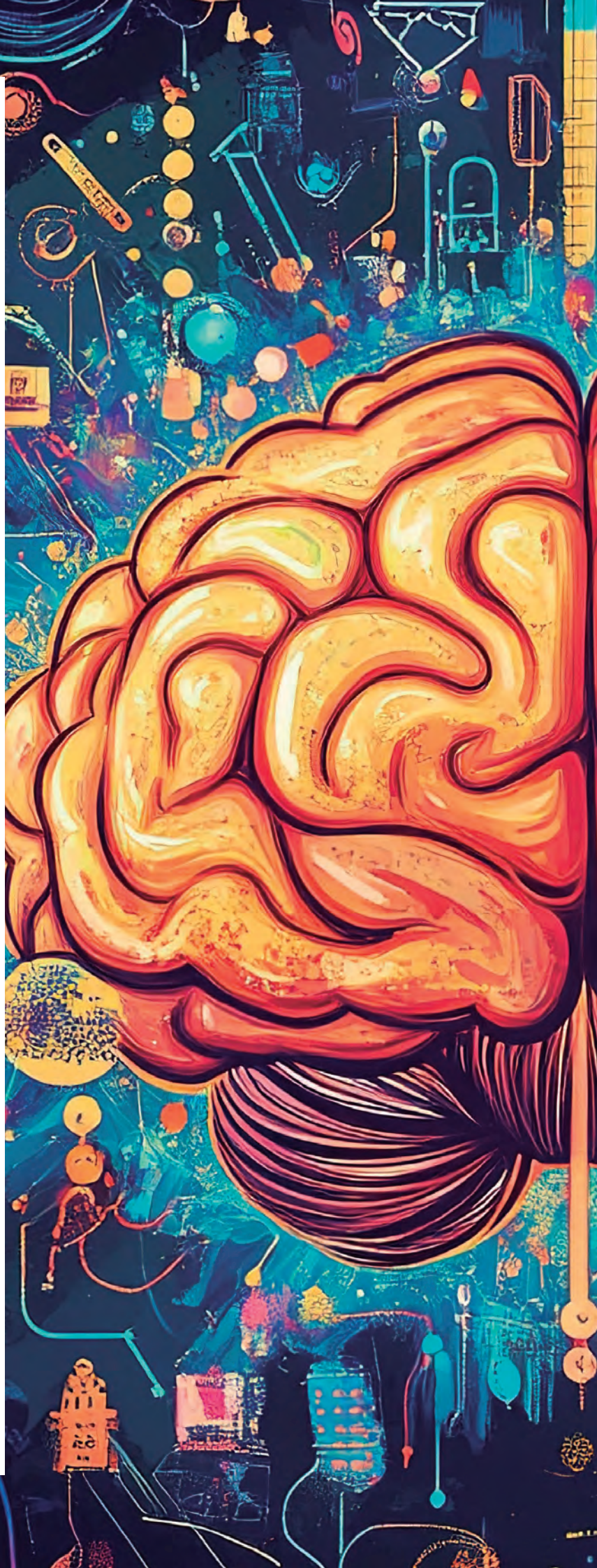
Observatório de Inteligência Artificial na Educação Superior

Diante da crescente disponibilidade de ferramentas de IA, a comunidade acadêmica da PUCPR tem manifestado diversos questionamentos e preocupações quanto a seu uso, especialmente no âmbito da pesquisa e do ensino.

Reconhecendo os potenciais benefícios dessa tecnologia para os processos de ensino, aprendizagem, pesquisa e desenvolvimento, a Universidade considera essencial que seu uso seja pautado por autonomia, responsabilidade, ética, senso crítico e honestidade por parte de professores, estudantes e colaboradores.

Sendo assim, a Pró-Reitoria de Desenvolvimento Educacional constituiu o respectivo Observatório cujos objetivos são:

- Apresentar o posicionamento da Instituição relativo ao uso de ferramentas de IA nos processos de ensino, aprendizagem, pesquisa e desenvolvimento.
- Oferecer orientações preliminares sobre procedimentos, processos e boas práticas para o uso de ferramentas de IA, com a possibilidade de futuras alterações decorrentes do diálogo contínuo com a comunidade acadêmica.





Currently, at PUCPR, there are some centers and observatories that are dedicated to studying and monitoring the latest developments, limits and possibilities of applying technological innovations within the scope of teachers' professional practice with a view to constantly qualifying the teaching-learning relationship.

Among these structures are:

Center for Philosophical Studies of Artificial Intelligence and the Future of Humanity (CEFIA)

It is a multidisciplinary project linked to the Graduate Program in Philosophy at PUCPR. It explores ethical, epistemological and technological issues related to embodied cognition, transhumanism and Artificial Intelligence. With international cooperation and a focus on impactful publications, CEFIA promotes research, events and academic exchange to deepen debates on biotechnological interaction, social justice and the role of technologies in the human future, uniting philosophy, cognitive sciences and education.

Observatory of Artificial Intelligence in Higher Education

Given the growing availability of AI tools, the PUCPR academic community has expressed several questions and concerns regarding their use, especially in the context of research and teaching.

Recognizing the potential benefits of this technology for teaching, learning, research and development processes, the University considers it essential that its use be guided by autonomy, responsibility, ethics, critical sense and honesty on the part of professors, students and collaborators.

Therefore, the Pro-Rectorate of Educational Development established the respective Observatory whose objectives are:

- Present the Institution's position regarding the use of AI tools in teaching, learning, research and development processes.
- Offer preliminary guidance on procedures, processes and good practices for the use of AI tools, with the possibility of future changes resulting from ongoing dialogue with the academic community.

