

Você pode (ou deve) usar

IAS GENERATIVAS PARA FAZER PESQUISA CIENTÍFICA?

Can you (or should you) use

GENERATIVE AIs FOR SCIENTIFIC RESEARCH?

A ilustração acima representa graficamente como as IAs generativas são alimentadas com um *prompt* (o ponto azul à esquerda), ativando o funcionamento de uma rede neural, que processa esses dados de entrada (as várias linhas convergentes no centro), para então produzir uma ampla variedade de resultados diversificados e originais (os pontos coloridos à direita)

The illustration above depicts how generative AIs are fed with a prompt (the blue dot on the left), activating the functioning of a neural network, which processes this input data (the many lines converging in the center), and then produces a wide variety of diverse and original results (the colorful dots on the right)

Por/By: Guilherme Profeta
Fotos/Photos: Fernando Rezende

Em junho de 2023, na edição de número 11 da revista Uniso Ciência, uma reportagem discutiu como o surgimento de Inteligências Artificiais generativas, capazes de produzir toda sorte de ilustrações a partir de *prompts* simples, poderia (ou não) redefinir a própria definição de arte, e como as legislações ainda se encontram defasadas em relação às possibilidades tecnológicas atuais. No ano seguinte, em junho de 2024, uma segunda reportagem discutiu se e como o ChatGPT — provavelmente o produto mais famoso dentre as IAs generativas — já foi plenamente incorporado às práticas contemporâneas, bem como as suas potencialidades e limitações, além dos riscos envolvidos em seu uso em diversos contextos, inclusive os educacionais. Nesta terceira reportagem da mesma **SÉRIE**, a discussão está centrada na utilização de IAs na pesquisa científica, uma possibilidade para a qual a academia vem se atentando e que, da mesma forma que oferece uma série de novas possibilidades, parece requerer, também, uma série de cuidados.

Back in June 2023, in issue #11 of the Science @ Uniso magazine, there was a story that discussed how the emergence of generative Artificial Intelligences, capable of producing all kinds of illustrations from simple prompts, could (or could not) redefine the very definition of art, and how current legislation still lags behind today’s technological possibilities. The following year, in June 2024, a second story discussed whether and how ChatGPT—most likely the most famous product among generative AIs—has already been fully incorporated into contemporary practices, as well as its potential and limitations, in addition to the risks involved in its use in various contexts, including education. In this third story of the same **SERIES**, the discussion is focused on the use of AIs as part of scientific research, a possibility that academia has been paying attention to, and which offers a range of new possibilities, but not without requiring a series of precautions.

PESQUISAS RECÉM-SAÍDAS DA FICÇÃO CIENTÍFICA

Não é particularmente raro que a ficção científica especulativa — aquela que se baseia em vislumbrar cenários hipotéticos além do mundo histórico atual — preveja certas mudanças de ordem tecnológica e/ou social antes que elas se tornem realidade. Com a recente (re)popularização das IAs — para ler mais sobre a história desse tipo de tecnologia, consulte as reportagens anteriores —, parece que algumas dessas lucubrações próprias da ficção científica começam a tomar corpo, com exemplos de projetos bastante fronteiriços, por assim dizer, despontando aqui e acolá.

São projetos de pesquisa baseados em IA que prometem extrapolar substancialmente os limites do conhecimento de seus respectivos campos, como aquele conduzido pelo professor doutor David Gruber, que atua na área de Biologia Marinha na City University of New York (CUNY), além de outras instituições de ensino e pesquisa também em Nova Iorque. Gruber está à frente do projeto CETI (*Cetacean Translation Initiative*, ou, numa tradução livre, Iniciativa de Tradução de Cetáceos), criado em 2020 com o objetivo de se utilizar de IA para identificar padrões em sistemas de comunicação animal, mais especificamente baleias cachalote, de modo a traduzir essa forma de comunicação não humana para linguagens humanas — e vice-versa, potencialmente.

A proposta de Gruber se baseia em métodos de tradução via IA que já funcionam para línguas humanas, por meio dos quais é possível comparar padrões entre dois idiomas distintos e traduzi-los diretamente, sem a necessidade de um elemento intermediário, como um dicionário ou algo semelhante a uma Pedra de Roseta (que contém um mesmo texto escrito numa língua desconhecida e noutra língua conhecida, servindo como referência para decifrar a primeira). No caso das baleias cachalote, que se comunicam por meio de cliques sonoros, a IA já consegue prever os próximos cliques numa mesma sequência com algo em torno de 99% de precisão. Esses cliques

PIONEERING RESEARCH THAT RESEMBLES SCI-FI

It is not particularly rare for speculative science fiction—which is based on envisioning hypothetical scenarios beyond the current historical world—to anticipate certain technological and/or social changes before they become a reality. With the recent (re)popularization of AIs—to read more on the history of this type of technology, refer to the previous stories—it seems that some of these speculations typical of science fiction are beginning to take shape, with examples of cutting-edge projects on the very frontier of knowledge emerging here and there.

These are AI-based research projects that promise to substantially push the boundaries of knowledge in their respective fields, such as the one led by professor David Gruber, a marine biologist working at the City University of New York (CUNY), as well as other educational and research institutions in New York. Gruber is at the forefront of the CETI project (which stands for Cetacean Translation Initiative), created in 2020 with the goal of using AI to identify patterns in animal communication systems, specifically in the “speech” of sperm whales, in order to translate this form of non-human communication into human languages—and vice versa, potentially.

Gruber’s proposal is based on AI translation methods that already work for human languages, through which it is possible to compare patterns between two distinct languages and translate them directly, without the need for an intermediary element, such as a dictionary or something akin to a Rosetta Stone (which contains the same text written in an unknown language and another known language, serving as a reference for deciphering the first). In the case of sperm whales, which communicate through sound clicks, AI can already predict the next clicks in the same sequence with about 99% accuracy. These clicks comprise more

SE VOCÊ GOSTOU DESTA REPORTAGEM...

Continue lendo as reportagens anteriores, parte da mesma série sobre Inteligência Artificial:

IF YOU ENJOYED THIS STORY...

Keep reading the previous stories, part of the same series on Artificial Intelligence:



“Inteligência artificial: pode a tecnologia redefinir os limites da arte?” (jun./2023); siga o link pelo **QR code** para acessar



“Artificial Intelligence: will technology redefine the boundaries of art?” (June/2023); use the QR code to follow the link



“Que fim levou o ChatGPT?” (jun./2024); siga o link pelo **QR code** para acessar



“What happened to ChatGPT?” (June/2024); use the QR code to follow the link

somam mais de 150 sequências identificadas (algo semelhante ao que chamaríamos de fonemas) e os cientistas já perceberam que eles não apenas são adquiridos socialmente, mas também apresentam diferenças regionais (como dialetos). O próximo passo é compreender os contextos específicos em que essas sequências são utilizadas, para então, um dia — quem sabe —, aprender a se comunicar com as baleias no “idioma” das próprias baleias. Você pode ouvir mais sobre essa pesquisa (em inglês) no *podcast* de divulgação científica *Big Brains*, da Universidade de Chicago, no **EPISÓDIO** de 30 de maio de 2024.

Outro exemplo é a pesquisa conduzida no projeto SETI (*Search for Extraterrestrial Intelligence*, ou, em tradução livre para o português, Busca por Inteligência Extraterrestre). Não é à toa que os nomes são parecidos; o CETI com C é muito mais recente do que o SETI, mas, de certa forma, a intenção é a mesma: comunicar-se com formas de vida não humana. A diferença é que, no caso do SETI com S, as formas de vida são hipotéticas e, diferentemente das baleias cachalote, estão muito longe de nós — muito longe *mesmo!* —, em outros planetas.

O SETI vem “ouvindo” o céu desde o fim da década de 1950, com o intuito de identificar sinais (ou transmissões) de rádio com características diferentes daquelas que normalmente seriam encontradas no espaço sideral, indicando a existência de civilizações inteligentes fora do planeta Terra. Dificuldades nesse método incluem o fato de esses sinais em potencial estarem chegando de qualquer direção, em qualquer frequência, o que significa que, se não houver um telescópio apontado e calibrado para ouvir o sinal certo no momento certo, é bastante provável que ele seja perdido para sempre. Além disso, há a questão de que, para que seja possível identificar esses sinais em meio a todo o ruído de fundo, faz-se necessário ter pesquisadores identificando anomalias nesses padrões, o que implica em analisar uma quantidade (literalmente) astronômica de dados de uma vez só, algo que seria humanamente impossível. É aí que entram as IAs fazendo o que elas fazem de

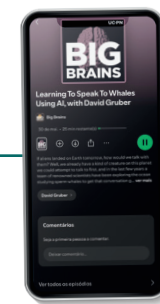
than 150 identified sequences (similar to what we would call phonemes), and scientists have already observed that they are not only socially acquired, but also exhibit regional differences (like dialects). The next step is to understand the specific contexts in which these sequences are used, so that one day—who knows—we might learn to communicate with whales in their own “language.” You can hear more about this research (in English) on the University of Chicago’s science outreach podcast titled Big Brains, in the **EPISODE** from May 30, 2024.

Another example is the research conducted in the SETI project (Search for Extraterrestrial Intelligence). It is no coincidence that the names are similar; the CETI with a C is much more recent than the SETI, but in a way, the intention is the same: to communicate with non-human forms of life. The difference is that, in the case of SETI with an S, the forms of life are hypothetical and, unlike sperm whales, are very far away from us—very far indeed!—, on other planets.

SETI has been “listening” to the sky since the late 1950s, with the aim of identifying radio signals (or transmissions) that have characteristics different from those typically found in outer space, thus indicating the existence of intelligent civilizations beyond Earth. Challenges with this method include the fact that these potential signals could be coming from any direction, at any frequency, which means that if there isn’t a telescope pointed and calibrated to listen to the right signal at the right moment, it is quite likely that it will be lost forever. Additionally, there is the issue that, in order to identify these signals amid all the background noise, researchers need to identify anomalies in these patterns, which involves analyzing an astronomical amount of data at once (quite literally), something that would be humanly impossible. Here come the AIs then, doing what they do best: identifying patterns.

melhor: identificar padrões. Até o início de 2024, uma dessas IAs, desenvolvida por um estudante da Universidade de Toronto chamado Peter Ma, já havia identificado oito sinais em potencial. Você pode ler mais sobre essa iniciativa numa **REPORTAGEM** de Emma Woollacott publicada pela BBC (também em inglês) em 21 de fevereiro de 2024.

Siga os links pelos *QR codes* para ouvir o *podcast* (no Spotify) e ler a reportagem:



EXEMPLOS DE PESQUISAS EM ANDAMENTO NA UNISO

Na Universidade de Sorocaba (Uniso), o professor doutor Thiago Simon Marques, coordenador do curso de graduação em Ciências Biológicas e professor do Programa de Pós-Graduação em Processos Tecnológicos e Ambientais, é um dos pesquisadores que vem conduzindo pesquisas semelhantes — não no sentido de usar a IA para desvendar as particularidades de linguagem de determinadas espécies não humanas, mas para monitorar a ocorrência dessas espécies no espaço. No caso de seu estudo, a espécie em questão é a anta, também conhecida como tapir (*Tapirus terrestris*), um mamífero muito comum na região de Sorocaba, que foi muito caçado historicamente e é muito

By early 2024, one of these AIs, developed by a University of Toronto student named Peter Ma, had already identified eight potential signals. You can read more about this initiative in a **STORY** by Emma Woollacott published by the BBC (also in English) on February 21, 2024.

To listen to the podcast (on Spotify) and read the story, use the QR codes to follow the links:



EXAMPLES OF ONGOING RESEARCH BEING HELD AT UNISO

At Uniso, professor Thiago Simon Marques, coordinator of the university’s undergraduate program in Biological Sciences, as well as a professor at Uniso’s graduate program in Technological and Environmental Processes, is one of the researchers conducting similar studies—not in the sense of using AI to unravel the language specifics of certain non-human species, but to monitor the occurrence of these species in physical spaces. In his study, that species is the tapir (*Tapirus terrestris*), a mammal that is very common in the region of Sorocaba, which has been heavily hunted in the past, and is very important

importante para semear as florestas do estado de São Paulo.

“Monitorar a biodiversidade nem sempre é uma tarefa simples e, hoje em dia, a IA tem auxiliado muito nesse processo de tentar entender padrões, sejam padrões de distribuição ou identificação de espécies”, explica Marques. “Não raro nós precisamos saber quais espécies ocorrem em determinado lugar, porque, no mundo real, onde os recursos são limitados, nem sempre há dinheiro para conservar todas as áreas. Assim, áreas onde ocorrem mais espécies acabam sendo mais destinadas para essa finalidade de conservação. No caso das antas, a exemplo das baleias cachalote, devemos lembrar que elas também são animais que emitem sons característicos. O projeto que estamos desenvolvendo usa gravadores autônomos, posicionados no Parque Zoológico Municipal Quinzinho de Barros, em Sorocaba, por meio dos quais estamos catalogando esses sons e, a partir daí, vamos treinar um modelo de IA para o reconhecimento dos sons de maneira automatizada. Então, depois de treinar a IA em cativeiro, que é um ambiente controlado, a ideia é colocar gravadores em ambiente natural para detectar o som das antas de maneira autônoma, assim registrando a localização desses animais. Com esses bancos de dados bem catalogados, nós vamos conseguir saber de uma maneira muito mais apurada por onde a biodiversidade está espalhada, e isso é um ganho muito interessante quando pensamos em conservação biológica.”

Marques destaca que isso otimiza custos, porque é possível monitorar determinadas áreas sem ter necessariamente a presença de um biólogo em campo. Hoje, principalmente com a *Starlink* — a rede global de satélites da empresa SpaceX, que pode oferecer internet em áreas bastante isoladas —, é possível fazer isso em áreas muito remotas, como a floresta amazônica, por exemplo, além de ser possível usar a mesma tecnologia até mesmo para monitorar atividades de caça e desmatamento, se as IAs estiverem calibradas para reconhecer o ruído de tiros ou motosserras.

Outro exemplo de pesquisa envolvendo IA para a otimização de recursos em pesquisa é o trabalho de conclusão de curso do estudante Paulo Henrique Urbano Leite, do curso de graduação em

when it comes to seeding the forests of the state of São Paulo.

“Monitoring biodiversity is not always a simple task, and nowadays, AI has greatly assisted in this process of trying to understand patterns, whether they are patterns of distribution or species identification,” Marques explains. “It is often necessary to know which species occur in a given area because, in the real world, where resources are limited, there is not always enough money to conserve all areas. Thus, areas with more species tend to be prioritized for conservation purposes. In the case of tapirs, like sperm whales, we should remember that they also emit characteristic sounds. The project we are developing uses autonomous recorders positioned in the Quinzinho de Barros Municipal Zoo, in Sorocaba, through which we are cataloging these sounds, and from there, we will train an AI model for automated sound recognition. So, after training the AI in captivity, which is a controlled environment, the idea is to place recorders in a natural environment to autonomously detect the sounds of tapirs, thereby recording the locations of these animals. With these well-cataloged databases, we will be able to know much more accurately where biodiversity is spread, which is a very interesting advantage when we think about biological conservation.”

Marques emphasizes that this optimizes costs because it is possible to monitor certain areas without necessarily having a biologist on-site. Today, especially with Starlink—the global satellite network from SpaceX, which can provide internet access in very isolated areas—it is possible to do this in very remote regions, such as the Amazon rainforest, for example. Additionally, the same technology can even be used to monitor hunting and deforestation activities, if the AIs are calibrated to recognize the noise of gunshots or chainsaws.

Another example of research involving AI for resource optimization is the final graduation project of Paulo Henrique Urbano Leite, a student from Uniso’s undergraduate program in Computer



Foto/Photo: Stanislav (Adobe Stock)



Foto/Photo: sdecoret (Adobe Stock)

De animais marinhos a extraterrestres: IAs estão sendo utilizadas para identificar padrões na comunicação de baleias cachalote e em sinais de rádio que podem ter origem em civilizações alienígenas

From marine animals to extraterrestrials: AIs are being used to identify patterns in sperm whales’ communication and in radio signals that may originate from alien civilizations



Foto/Photo: Ammit (Adobe Stock)

Uma pesquisa desenvolvida na Uniso está usando IA para identificar o som das antas (*Tapirus terrestris*), para então monitorar de forma autônoma a sua ocorrência na natureza

A study developed at Uniso is using AI to identify the sound of tapirs (*Tapirus terrestris*) in order to autonomously monitor their occurrence in the wild

Engenharia da Computação. Em suas palavras, ele buscou na IA — especificamente nas IAs generativas, ou seja, aquelas capazes de gerar conteúdos originais (como textos, imagens, sons, vídeos, código etc.) — uma forma de resolver um “problema clássico da microscopia eletrônica de varredura”, que é a ausência de cor nas imagens produzidas por esse tipo de microscópio.

Conforme explica o professor mestre Denicezar Ângelo Baldo, coordenador do curso de graduação em Engenharia da Computação, os microscópios eletrônicos de varredura (MEV) são um tipo de microscópio que se utiliza de feixes de elétrons, em vez de luz, para varrer a superfície de uma amostra. Por isso as imagens são geradas em preto e branco, o que faz com que a **COLORIZAÇÃO** tenha de ser feita posteriormente, não raro manualmente, caso algum tipo de análise ou divulgação

Engineering. In his words, he intended to use AI—specifically generative AIs, meaning those capable of producing original content (such as texts, images, sounds, videos, code, etc.)—as a way of solving a “classic problem of scanning electron microscopy,” which is the lack of color in the images produced by this type of microscope.

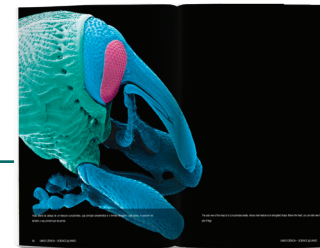
As explained by professor Denicezar Ângelo Baldo, coordinator of Uniso’s undergraduate program in Computer Engineering, scanning electron microscopes (SEM) are a type of microscope that uses electron beams instead of light to scan the surface of a sample. Therefore, the images are generated in black and white, which means that **COLORIZATION** has to be done afterwards, often manually, if any kind of analysis or outreach activity requires colorful

necessite dessa especificidade. Dependendo da complexidade da imagem, o processo pode ser bastante trabalhoso e demorado. Na Uniso, muitas pesquisas nas áreas das Ciências Farmacêuticas e/ou da Engenharia de Materiais, por exemplo, se utilizam do MEV para caracterização físico-química de objetos diversos.

“Imagens coloridas precisam de três canais de cor”, explica Leite. “Com isso em mente, a ideia desse projeto é criar um mecanismo automático para mapear uma imagem em escala de cinza e torná-la colorida. Esse processo, quando feito de forma manual, é muito trabalhoso. Enquanto estava preparando os dados para o treinamento de uma IA que pudesse fazer isso, eu levei, em média, 30 minutos para colorizar cada imagem. Quando você tem muitas imagens de uma só vez, o tempo e o custo para fazer isso podem ser bastante expressivos. Nossa proposta é fazê-lo de forma completamente autônoma: o pesquisador faz a imagem no MEV e a passa por um modelo de rede neural generativa que torna a imagem colorida de forma totalmente instantânea, produzindo aqueles canais de cor que estão faltando.”

images. Depending on the complexity of the image, the process can be quite laborious and time-consuming. At Uniso, many research projects in the fields of Pharmaceutical Sciences and/or Materials Engineering, for example, utilize SEM for the physicochemical characterization of various objects.

“Colored images require three color channels,” Leite explains. “With this in mind, the idea of this project is to create an automatic mechanism to map a grayscale image and turn it into color. This process, when done manually, is very labor-intensive. While preparing the data to train an AI that could perform this task, it took me, on average, 30 minutes to colorize each image. When you have many images at once, the time and cost to do this can be quite significant. Our proposal is to do it completely autonomously: the researcher produces the image with the SEM, and passes it through a generative neural network model that instantly turns the image into color, thus creating those missing color channels.”



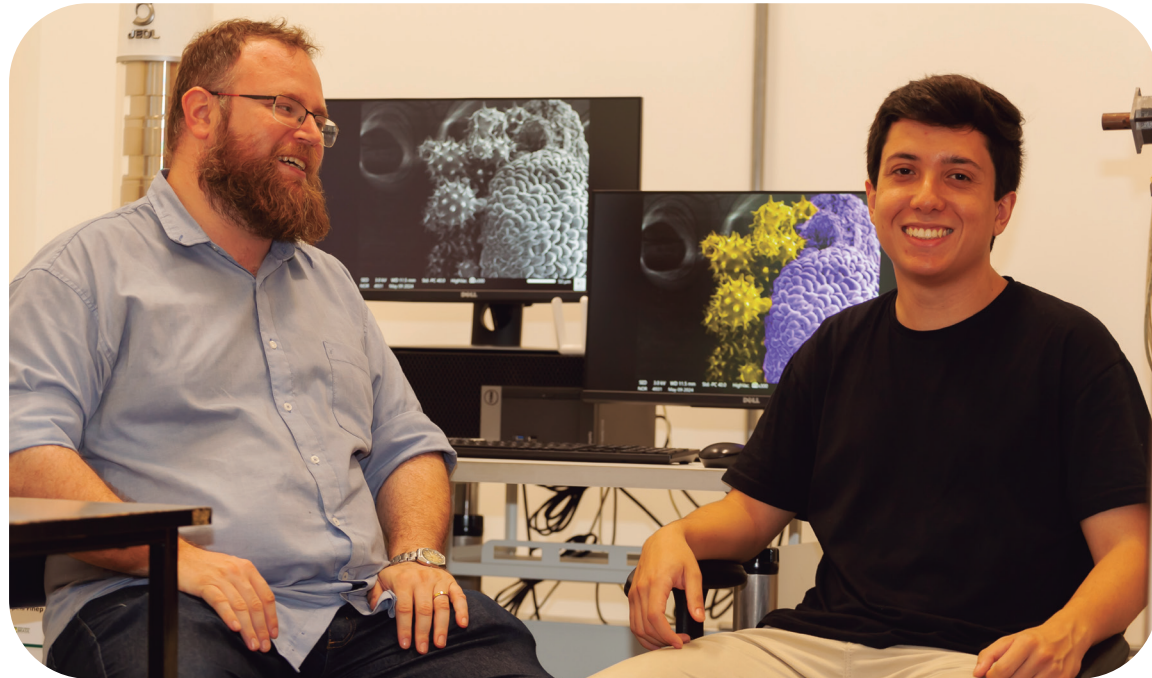
Exemplos de imagens do MEV colorizadas manualmente podem ser conferidos ao longo da edição de número 10 (dez./2022) da revista Uniso Ciência; siga o link pelo QR code para acessar a edição completa

Examples of manually colorized SEM images can be found throughout issue number 10 (Dec./2022) of the Science @ Uniso magazine; use the QR code to follow the link and get access to the full issue



Na imagem que aparece ao fundo da foto na página 18, é possível ver, à esquerda, uma imagem de grãos de pólen, em preto e branco, da forma como ela foi obtida pelo MEV. À direita, pode ser conferida a mesma imagem depois de ela ter sido colorizada automaticamente pela IA desenvolvida por Leite. O processo leva, em média, sete segundos por imagem — contra os 30 minutos que o pesquisador levaria para fazê-lo manualmente. Dependendo do *hardware* em que o sistema estiver rodando, esse tempo pode cair para 2 segundos, segundo o estudante. Considerando que essas imagens coloridas funcionam não apenas para a

In the image that appears in the background of the photo on page 18, it is possible to see, on the left, a grayscale image of pollen grains, as obtained by the SEM. On the right, the same image can be viewed after it has been automatically colorized by the AI developed by Leite. The process takes, on average, seven seconds per image—compared to the 30 minutes it would take the researcher to do it manually. Depending on the hardware on which the system is running, this time can drop to 2 seconds, according to the student. Considering that these color images serve not only for outreach



À esquerda, o professor mestre Denicezar A. Baldo, coordenador do curso de graduação em Engenharia da Computação da Uniso; à direita, o estudante de graduação Paulo Henrique U. Leite; ao fundo, imagem gerada pelo microscópio eletrônico de varredura (MEV) da Uniso e colorizada por meio de IA

On the left, professor Denicezar A. Baldo, coordinator of Uniso's undergraduate program in Computer Engineering; on the right, the undergraduate student Paulo Henrique U. Leite; in the background, an image generated by a scanning electron microscope (SEM) at Uniso, and its colorized version, which was created by AI

divulgação, mas também para facilitar o processo de análise — uma vez que identificar estruturas coloridas é muito mais fácil do que em preto e branco —, Leite defende que esse tipo de IA pode ser compreendido como uma ferramenta de auxílio à pesquisa. “Isso porque ela permite delegar à IA uma tarefa que demandaria tempo do cientista, de modo que esse pesquisador pode continuar focado em seu trabalho, enquanto a IA opera a colorização”, ele conclui.

E QUANDO A PESQUISA ENVOLVE SERES HUMANOS?

Das pesquisas já mencionadas nesta reportagem, a única que faz uso efetivo de IAs generativas é a de Leite — ainda que, se um dia estudos posteriores àqueles mencionados anteriormente viessem a gerar sons ou sinais originais para se comunicar com baleias, ou antas (ou quem sabe até alienígenas), é plausível imaginar que isso seria feito por meio de IAs desse tipo.

but also to facilitate the analysis process—since identifying colored structures is much easier than in black and white—Leite argues that this type of AI can be understood as a research aid tool. “This is because it allows delegating a task that would take time from the scientist to the AI, so that researchers can continue to focus on their work while the AI handles the colorization,” he concludes.

WHAT ABOUT RESEARCH THAT INVOLVES HUMAN PARTICIPANTS?

Among the studies already mentioned in this story, the only one that makes effective use of generative AIs is Leite's—although, if upcoming studies were to generate original sounds or signals to communicate with whales, tapirs (or perhaps even aliens) in the future, it is plausible to imagine that this would be done by using generative AIs.

Um aspecto que todos esses estudos mencionados têm em comum, no entanto, é que eles *não* envolvem seres humanos como participantes. Isso é importante ressaltar porque, quando isso acontece, em qualquer tipo de estudo — sejam ensaios clínicos para o desenvolvimento de medicamentos ou vacinas, por exemplo, ou mesmo estudos focados no registro e/ou na observação de opiniões e comportamentos —, há uma série de procedimentos extras que devem ser rigorosamente seguidos. Nesse tipo de pesquisa, a utilização de IAs, especialmente as generativas, é um complicador a mais, daí a importância de envolver os **COMITÊS DE ÉTICA EM PESQUISA** com Seres Humanos nessa discussão.

One aspect that all these studies have in common, though, is that they do not involve human beings as participants. This is important to emphasize because, when that does happen, in any type of study—whether clinical trials for the development of drugs or vaccines, for example, or even studies focused on recording and/or observing opinions and behaviors—there is a series of additional procedures that must be strictly followed. In this type of research, the use of AIs, especially generative ones, adds another layer of complexity, which is why it is important to involve **RESEARCH ETHICS COMMITTEES** in this debate.

PARA SABER MAIS: OS COMITÊS DE ÉTICA EM PESQUISA COM SERES HUMANOS

No contexto da pesquisa científica brasileira, um Comitê de Ética em Pesquisa (CEP) é um órgão interdisciplinar e independente, formado por pesquisadores de diversas áreas do conhecimento, que tem por responsabilidade avaliar, aprovar e fiscalizar quaisquer projetos de pesquisa que envolva seres humanos. Esses comitês, geralmente alocados em instituições de ensino e pesquisa onde estudos desse tipo são desenvolvidos — como a Uniso —, são credenciados junto ao Ministério da Saúde e servem para garantir a segurança e o bem-estar de todos os envolvidos na pesquisa acadêmica, especialmente dos participantes. No Brasil, esses comitês trabalham principalmente conforme as diretrizes da Resolução 466/2012 do Conselho Nacional de Saúde (CNS), além de outras resoluções e/ou leis, como mais recentemente a Lei Geral de Proteção de Dados (LGPD). Todo projeto de pesquisa que envolva a coleta de dados a partir de seres humanos deve ser aprovado num CEP antes de iniciado.

TO KNOW BETTER: RESEARCH ETHICS COMMITTEES

In the context of Brazilian scientific research, a Research Ethics Committee (CEP, in the Portuguese acronym) is an interdisciplinary and independent entity, composed of researchers from various fields of knowledge, responsible for evaluating, approving, and overseeing any research projects involving human subjects. These committees, usually located in educational and research institutions where such studies are conducted—like Uniso—are accredited by the Brazilian Ministry of Health, and serve to ensure the safety and well-being of everyone involved in academic research, especially the participants. In Brazil, these committees primarily operate according to the guidelines of Resolution 466/2012 of the National Health Council (CNS), along with other resolutions and/or laws, including the recently enacted General Data Protection Law (LGPD). Every single research project that involves the collection of data from human subjects must be approved by a Research Ethics Committee before it begins.

Nestes momentos de transição, em que esse tipo de tecnologia já está incorporada em certas práticas contemporâneas, sem necessariamente estar regulada, isso pode gerar muitas dúvidas quanto ao seu uso ético. E a academia não é exceção; para os pesquisadores de diversas áreas do conhecimento (especialmente das Humanidades e das Ciências Sociais Aplicadas, em que a forma de construir conhecimento é diferente das Ciências Biológicas ou das Engenharias, por exemplo), talvez a principal delas possa ser resumida da seguinte forma: que informações um projeto de pesquisa deve minimamente incluir, se usar algum tipo de IA como parte de seus procedimentos, para ser aprovado num CEP, assim garantindo a segurança e o bem-estar de seus voluntários?

Segundo o professor doutor Diego Aparecido Carvalho Albuquerque, coordenador do Comitê de Ética em Pesquisa com Seres Humanos da Uniso (CEP-Uniso), não existe uma fórmula pronta, uma vez que cada projeto é analisado e aprovado individualmente, e também pelo fato de não existirem, ao menos ainda, resoluções específicas sobre a utilização de IAs na pesquisa. “Isso é natural; às vezes a tecnologia avança um pouco mais rapidamente do que as políticas em que nós nos baseamos”, ele diz. No entanto, existem alguns pontos de atenção: “O primeiro deles seria descrever no seu protocolo de pesquisa, que está sendo submetido ao CEP, qual é a IA que você, pesquisador, está planejando utilizar, e qual é a justificativa de ela ser utilizada, exatamente em quais etapas do processo. O CEP está aqui para a proteção integral do participante e quando você descreve que IA pretende utilizar, em qual momento e o porquê dessa utilização, isso dá clareza para que nós chequemos se essa utilização por acaso fere alguma das resoluções em que nós nos baseamos.”

O professor doutor Nobel Penteado de Freitas, docente na graduação em Ciências Biológicas e membro titular do CEP-Uniso, acrescenta que isso precisa ficar claro não só para o CEP, mas também para os próprios participantes da pesquisa: “O que é realmente importante, no caso de pesquisadores que pretendem se utilizar de IA, é deixar isso claro tanto no protocolo de pesquisa quanto no Termo de Esclarecimento Livre e Esclarecido (TCLE), que é o termo de consentimento que todos os

In these moments of transition, when this type of technology is already incorporated into certain contemporary practices but is not necessarily regulated, many ethical issues and doubts can arise. Academia is no exception; for researchers working in various fields of knowledge (especially in the Humanities and Applied Social Sciences, where the methods for validating knowledge differ from those in Biological Sciences or Engineering, for example), perhaps the main question can be summarized as follows: what information should a research project minimally include, if it uses any type of AI as part of its procedures, to be approved by a Research Ethics Committee, thereby ensuring the safety and well-being of its volunteers?

According to professor Diego Aparecido Carvalho Albuquerque, coordinator of Uniso’s Research Ethics Committee, there is no such thing as an all-in-one solution, as each project is analyzed and approved individually, and also because there are currently no specific resolutions regarding the use of AIs in research. “This is natural; sometimes technology advances a bit faster than the policies we rely on,” he says. However, there are some key points to consider: “The first would be to describe in your research protocol, which is being submitted to the ethics committee, which AI you, as the researcher in charge, plan to use at which stages of the process, and the justification for its use. The committee is here to ensure the protection of all participants, and when you describe which AI you intend to use, when, and why, this provides clarity for us to check whether this use violates any of the resolutions we rely on.”

Professor Nobel Penteado de Freitas, a faculty member working at Uniso’s undergraduate program in Biological Sciences, and also a member of Uniso’s Research Ethics Committee, adds that this needs to be clear not only to the committee, but also to the research participants themselves: “What is really important, when it comes to researchers who intend to use AI, is to make this clear both in the research protocol and in the Informed Consent Form (TCLE, in the Portuguese acronym), which is the consent document that all participants must sign. This must be presented



Da esquerda para a direita, os professores doutores Ana Cristina da Costa P. Grohs, Diego A. C. Albuquerque e Nobel Penteado de Freitas, membros do Comitê de Ética em Pesquisa com Seres Humanos da Uniso (CEP-Uniso)

From left to right, professors Ana Cristina da Costa P. Grohs, Diego A. C. Albuquerque, and Nobel Penteado de Freitas, members of Uniso’s Research Ethics Committee

participantes têm de assinar. E isso precisa aparecer de uma maneira clara, numa linguagem acessível que todo mundo possa entender. Outra coisa que ainda deve gerar bastante discussão dentro do CEP é a avaliação dos riscos, uma vez que pode haver certa insegurança em relação à proteção dos dados.”

Além disso, a professora doutora Ana Cristina da Costa Piletti Grohs, que é docente na graduação em Relações Públicas e também membro titular do CEP-Uniso, chama atenção para a potencial ocorrência de vieses, e como os pesquisadores pretendem reconhecer e abordar essa questão em particular: “Um dos nossos desafios fundamentais é entender essa diversidade de tecnologias diferentes que utilizam IA. Você tem modelos generativos, preditivos, supervisionados, não supervisionados etc. e, considerando essa variedade, uma questão muito importante é o controle que o pesquisador

clearly, and in accessible language that everyone can understand. Another issue that is likely to generate considerable discussion within the committee is the assessment of risks, as there may be some uncertainty regarding data protection.”

Additionally, professor Ana Cristina da Costa Piletti Grohs, a faculty member working at Uniso’s undergraduate program in Public Relations, and another member of Uniso’s Research Ethics Committees, highlights the potential occurrence of biases, and how researchers intend to recognize and address this particular issue: “One of our fundamental challenges is to understand this diversity of different technologies that use AI. You have generative, predictive, supervised, and unsupervised models, and so forth, and considering this variety, a very important issue is the control

tem (ou não) sobre o processamento desses dados, especialmente para evitar os vieses.” Por vieses, podem ser compreendidas as tendências ou preferências sistemáticas que afetam a geração de resultados: se os dados com que a IA foi treinada contêm maior representação de certas culturas, gêneros ou opiniões, por exemplo, ela pode gerar respostas favoráveis a certos grupos em detrimento de outros, culminando em desinformação ou discriminação.

O professor doutor Luís Roberto Momberg Albano, pesquisador na área de IA e aprendizagem de máquinas — que foi a principal fonte para a segunda reportagem desta série (em junho de 2024), discutindo as limitações do ChatGPT —, concorda. “Há dois pontos que entendo como importantes para se discutir quando pensamos na aprovação de projetos de pesquisa envolvendo IA nos CEPs”, ele diz. “A primeira coisa a se considerar é que, mesmo quando as IAs se propõem a *não* utilizar bases de dados externas, essas ferramentas já foram treinadas nessas bases originalmente. Então, quando nós pedimos para uma IA fazer qualquer tipo de criação ou análise, independentemente do grau de higiene técnica de dados a que esses modelos generativos se proponham, ela inevitavelmente vai ‘sujar’ qualquer coisa que esteja sendo gerada com os dados oriundos de suas bases. É muito improvável encontrar isenção, por isso a análise por parte do pesquisador requer muita cautela.”

A segunda questão diz respeito ao que se chama de *overfit*, que é o quanto as IAs generativas costumam se adaptar especificamente para as pessoas que estão fazendo a solicitação. “Depois de certo tempo de uso desse tipo de tecnologia por uma mesma pessoa, a IA vai acabar focando nas temáticas de preferência de quem está operando a busca. Isso é inevitável”, Albano alerta. “Vou dar um exemplo: eu pesquiso muito sobre Ciência da Informação, então, se eventualmente eu pesquisar alguma coisa sobre outra área — digamos que algum conceito da Educação, por exemplo —, a IA ainda vai me apresentar esse conceito à luz da Ciência da Informação, porque ela entende que essa é a temática principal com que eu trabalho. Então vai acontecendo ali um *overfitting* (ou sobreajuste) muito grande.”

that the researcher has (or does not have) over the processing of this data, especially to avoid biases.” Biases can be understood as systematic tendencies or preferences that affect the generation of results: if the data on which the AI was trained contains a greater representation of certain cultures, genders, or opinions, for example, it may generate responses favoring certain groups to the detriment of others, culminating in misinformation or discrimination.

Professor Luís Roberto Momberg Albano, a researcher in the field of AI and machine learning—who was the main source for the second story in this series (in June 2024), back then discussing the limitations of ChatGPT—agrees. “There are two issues that I consider quite important to discuss when thinking about the approval of research projects involving AI in Research Ethics Committees,” he says. “The first thing to consider is that, even when AIs propose not to use external databases, these tools have already been trained on those databases originally. So, when we ask an AI to perform any kind of creation or analysis, regardless of the degree of technical data hygiene that these generative models claim to maintain, it will inevitably ‘pollute’ anything being generated with data from its sources. It is very unlikely to find neutrality, so the analysis performed by the researcher requires a lot of caution.”

The second issue concerns what is known as overfitting, which refers to how generative AIs tend to adapt specifically to the individuals making the requests. “After a certain period of using this type of technology by the same person, the AI will end up focusing on the themes preferred by the person operating the search. This is inevitable,” Albano warns. “Let me give an example: I do a lot of research on Information Science, so if I eventually search for something in another field—let’s say a concept from Education, for example—the AI will still present that concept through the lens of Information Science, because it understands that this is the primary theme I work with. So, a significant overfitting happens there.”

Em outras palavras, uma IA se torna gradualmente mais precisa conforme processa dados conhecidos, mas pode falhar ao ser exposta a novos cenários, tendendo a preencher novas lacunas com padrões construídos a partir das preferências identificadas no histórico de interações prévias com cada usuário. Dessa forma, não é somente a base de dados que pode “contaminar” os dados gerados, mas também o próprio usuário, e isso precisa ser considerado pelos pesquisadores quando IAs generativas forem utilizadas em qualquer etapa de um projeto de pesquisa (da formulação de parte do texto à análise de dados inseridos por meio de *prompts*).

In other words, an AI gradually becomes more accurate as it processes known data but may fail when exposed to new scenarios, tending to fill new gaps with patterns constructed from the preferences identified in the user’s previous interaction history. Therefore, it is not only the database that can “contaminate” the generated data, but also the users themselves, and researchers do need to keep that in mind when generative AIs are used at any stage of a research project (from formulating part of the text to analyzing data generated through prompts).

Diretrizes gerais/General guidelines

- A princípio, não é proibido usar IAs em projetos de pesquisa, mas, no caso de pesquisas que envolvem seres humanos, deve ficar claro ao CEP que tipo de IA será utilizada, em quais etapas do estudo e o porquê de essa utilização ser necessária.
- It is not prohibited to use AIs in research projects, however, when it comes to research that involves human subjects, it must be made clear to a Research Ethics Committee what type of AI will be used, at which stages of the study, and why this use is necessary.
- Pesquisadores devem incluir informações sobre a utilização de IAs nos termos de consentimento que todos os participantes da pesquisa têm de assinar ao concordar com a coleta de seus dados. Isso deve ficar claro antes que o voluntário concorde em participar.
- Researchers must include information about the use of AIs in the consent forms that all participants must sign when agreeing to the collection of their data. This must be clear before the volunteer agrees to participate.
- A avaliação de riscos aos participantes de uma pesquisa científica deve considerar a segurança dos dados oferecidos à IA, em conformidade com a legislação vigente.
- The risk assessment for volunteers taking part in scientific research must consider the security of the data provided to the AI, in compliance with current laws on data protection.
- Tanto os dados com que uma IA generativa foi treinada originalmente quanto o próprio usuário podem “contaminar” as respostas obtidas por um pesquisador que se usa de IA ao processar dados de uma pesquisa. Isso não pode ser ignorado.
- Responses obtained by a researcher using AI to process research data can be “contaminated” both by the original databases originally used to train the AI, as well as by the user themselves. This cannot be overlooked.
- Todo projeto de pesquisa é único e será analisado individualmente. Não existe uma fórmula que sirva para todos.
- Every research project is unique and will be analyzed individually. There is no all-in-one solution.