



A Santa Sé

PAPA FRANCISCO PARTICIPA DA SESSÃO DO G7 SOBRE INTELIGÊNCIA ARTIFICIAL

[13-15 de junho de 2024]

DISCURSO DO PAPA FRANCISCO

Borgo Egnazia (Itália)

Sexta-feira, 14 de junho de 2024

[Multimídia]

Um instrumento fascinante e tremendo

Prezadas Senhoras, Ilustres Senhores!

Dirijo-me hoje a vós, Líderes do Fórum Intergovernamental do G7, com uma reflexão sobre os efeitos da inteligência artificial no futuro da humanidade.

«A Sagrada Escritura atesta que Deus deu aos homens o seu Espírito a fim de terem “sabedoria, inteligência e capacidade para toda a espécie de trabalho” (Ex 35, 31)» [1]. A ciência e a tecnologia são, por conseguinte, produtos extraordinários do nosso potencial criativo, como seres humanos [2].

Pois bem, a inteligência artificial emerge precisamente do uso deste potencial criativo que Deus nos deu.

Como é sabido, trata-se de um instrumento extremamente poderoso, utilizado em muitos domínios da atividade humana: da medicina ao mundo do trabalho, da cultura à comunicação, da educação à política. E é já legítimo supor que o seu uso influenciará cada vez mais a nossa forma de viver, as nossas relações sociais e, no futuro, até mesmo a maneira como concebemos a nossa identidade enquanto seres humanos [3].

No entanto, o tema da inteligência artificial é frequentemente percebido como ambivalente: por um lado, entusiasmo pelas possibilidades que oferece; por outro, gera temor pelas consequências que deixa antever. A este respeito, pode dizer-se que todos nós somos, embora em graus diferentes, atravessados por duas emoções: ficamos entusiasmados quando imaginamos os progressos que podem advir da inteligência artificial, mas ao mesmo tempo amedrontados quando constatamos os perigos inerentes ao seu uso [4].

Além disso, não podemos duvidar que o advento da inteligência artificial represente uma verdadeira revolução cognitivo-industrial que contribuirá para a criação de um novo sistema social caracterizado por complexas transformações epocais. Por exemplo, a inteligência artificial poderia permitir uma democratização do acesso ao conhecimento, o progresso exponencial da investigação científica, a possibilidade de delegar às máquinas os trabalhos exaustivos; mas ao mesmo tempo, ela poderia trazer consigo uma maior injustiça entre nações desenvolvidas e nações em vias de desenvolvimento, entre classes sociais dominantes e classes sociais oprimidas, colocando em perigo a possibilidade de uma “cultura do encontro” em favor de uma “cultura do descarte”.

O alcance dessas complexas transformações está obviamente ligado ao rápido desenvolvimento tecnológico da própria inteligência artificial.

Este vigoroso avanço tecnológico torna a inteligência artificial, simultaneamente, um *instrumento fascinante e tremendo*, e exige uma reflexão à altura da situação.

Nesse sentido, talvez se possa partir da constatação de que a inteligência artificial é, antes de tudo, *um instrumento*. E é natural afirmar que os benefícios ou danos que trará dependerão do modo como é utilizado.

Isso é certamente verdade, pois foi assim para cada ferramenta construída pelo ser humano desde o início dos tempos.

Essa nossa capacidade de construir utensílios numa quantidade e complexidade sem paralelo entre os seres vivos, faz-nos falar de uma *condição tecno-humana*: o ser humano sempre manteve uma relação com o ambiente mediada pelas ferramentas que ia produzindo. Não é possível separar a história do homem e da civilização da história desses instrumentos. Houve quem quisesse ler em tudo isto uma espécie de insuficiência, um *deficit* do ser humano, como se, por causa dessa carência, fosse obrigado a dar vida à tecnologia [5]. Na verdade, um olhar atento e objetivo mostra-nos o contrário. Vivemos uma condição de ulterioridade em relação ao nosso ser biológico; somos seres inclinados para fora-de-nós-mesmos, ou melhor, somos radicalmente abertos ao além. Aqui tem início a nossa abertura aos outros e a Deus; daqui nasce o potencial criativo da nossa inteligência em termos de cultura e beleza; e por fim, daqui se origina a nossa capacidade técnica. A tecnologia é, assim, uma marca desta nossa ulterioridade.

No entanto, o uso das nossas ferramentas nem sempre está orientado exclusivamente para o bem. Mesmo que o ser humano sinta interiormente uma vocação para além de si mesmo e para o conhecimento, vivido como instrumento de bem ao serviço dos irmãos, das irmãs e da *casa comum* (cf. *Gaudium et spes*, 16), isso nem sempre acontece. Ao contrário, não poucas vezes, precisamente graças à sua liberdade radical, a humanidade perverteu os fins do seu ser, transformando-se em inimiga de si mesma e do planeta [6]. O mesmo pode acontecer com os instrumentos tecnológicos. Somente se for garantida a sua vocação ao serviço do homem, os instrumentos tecnológicos revelarão não apenas a grandeza e a dignidade única do ser humano, mas também o mandato que este recebeu de “cultivar e guardar” (cfr. *Gn 2,15*) o planeta e todos os seus habitantes. Falar de tecnologia é falar sobre o que significa ser humano e, portanto, sobre aquela nossa condição única entre liberdade e responsabilidade, ou seja, é falar de ética.

Na verdade, quando os nossos antepassados afiavam pedras de sílex para fazer facas, usavam-nas tanto para cortar a pele dos vestuários quanto para se matarem uns aos outros. O mesmo se pode dizer de outras tecnologias muito mais avançadas, como a energia produzida pela fusão de átomos, como ocorre no Sol, que certamente poderia ser utilizada para produzir energia limpa e renovável, mas também para reduzir o nosso planeta a um monte de cinzas.

A inteligência artificial, no entanto, é um instrumento ainda mais complexo. Quase diria que se trata de um instrumento *sui generis*. Assim, enquanto o uso de uma ferramenta simples (como a faca) está sob o controlo do ser humano que a utiliza e o seu bom uso depende somente deste, a inteligência artificial, ao contrário, pode adaptar-se autonomamente à tarefa que lhe é atribuída e, se for projetada dessa forma, fazer escolhas independentes do ser humano para alcançar o objetivo estabelecido [7].

Convém sempre recordar que a máquina pode, sob algumas formas e com estes novos meios, produzir escolhas algorítmicas. O que a máquina faz é uma escolha técnica entre várias possibilidades e baseia-se ou em critérios bem definidos ou em inferências estatísticas. Pelo contrário, o ser humano não só escolhe como, no seu coração, é capaz de decidir. A decisão é um elemento que poderíamos definir como o mais estratégico de uma escolha e requer uma avaliação prática. Frequentemente, na difícil tarefa de governar, somos chamados a tomar decisões com consequências para muitas pessoas. A esse respeito, a reflexão humana sempre falou de sabedoria, a *phronesis* da filosofia grega e, pelo menos em parte, a sabedoria da Sagrada Escritura. Diante dos prodígios das máquinas, que parecem saber escolher de forma independente, devemos ter bem claro que a decisão deve ser sempre deixada ao ser humano, mesmo sob os tons dramáticos e urgentes com que, às vezes, se apresenta na nossa vida. Condenaríamos a humanidade a um futuro sem esperança se retirássemos às pessoas a capacidade de decidir sobre si mesmas e sobre as suas vidas, obrigando-as a depender das escolhas das máquinas. Precisamos de garantir e proteger um espaço de controle significativo do ser humano sobre o processo de escolha dos programas de inteligência artificial: está em jogo a própria dignidade humana.

Permitam-me insistir precisamente sobre este tema: num drama como o dos conflitos armados, é urgente repensar o desenvolvimento e o uso de dispositivos como as chamadas “armas autónomas letais”, a fim de banir a sua utilização, começando desde já pelo compromisso efetivo e concreto de introduzir um controlo humano cada vez mais significativo. Nenhuma máquina, em caso algum, deveria ter a possibilidade de optar por tirar a vida a um ser humano.

Acresce que o bom uso, pelo menos das formas avançadas de inteligência artificial, não estará totalmente sob o controlo nem dos utilizadores nem dos programadores que, no momento da conceção, definiram os seus objetivos originais. E isso é tanto mais verdadeiro quanto é altamente provável que, num futuro não distante, os programas de inteligência artificial possam comunicar diretamente entre si para melhorar o seu desempenho. E se, no passado, os seres humanos que moldaram ferramentas simples viram a sua existência moldada por elas – a faca permitiu-lhes sobreviver ao frio, mas também desenvolver a arte da guerra – agora que moldaram um instrumento complexo, verão este último moldar ainda mais a sua existência [8].

O mecanismo básico da inteligência artificial

Gostaria agora de me deter brevemente sobre a complexidade da inteligência artificial. Na sua essência, a inteligência artificial é uma ferramenta projetada para a resolução de um problema e funciona através de um encadeamento lógico de operações algébricas, realizadas sobre categorias de dados, que são comparados para descobrir correlações, melhorando o seu valor estatístico, graças a um processo de autoaprendizagem, baseado na busca de novos dados e na auto-modificação dos seus procedimentos de cálculo.

A inteligência artificial está assim projetada para resolver problemas específicos, mas, para quem a utiliza, é frequentemente irresistível a tentação de tirar deduções gerais, mesmo antropológicas, a partir das soluções específicas que propõe.

Um bom exemplo é o uso de programas projetados para ajudar os magistrados nas decisões relativas à concessão de prisão domiciliária a reclusos que cumprem pena num estabelecimento prisional. Neste caso, pede-se à inteligência artificial que preveja a probabilidade de reincidência no crime cometido por um condenado a partir de categorias preestabelecidas (tipo de crime, comportamento na prisão, avaliação psicológica, etc.), permitindo que a inteligência artificial tenha acesso a categorias de dados relativos à vida privada do preso (origem étnica, nível de instrução, linha de crédito, etc.). O uso de tal metodologia – que por vezes corre o risco de delegar *de facto* numa máquina a última palavra sobre o destino de uma pessoa – pode implicitamente trazer consigo a referência aos preconceitos inerentes às categorias de dados utilizados pela inteligência artificial.

Estar classificado num determinado grupo étnico ou, mais prosaicamente, ter cometido anos antes uma infração menor (por exemplo, não ter pago uma multa de estacionamento) influenciará

a decisão sobre a concessão da prisão domiciliária. Porém, o ser humano está sempre em evolução e é capaz de surpreender com as suas ações, algo que uma máquina não pode ter em consideração.

Deve-se ainda observar que aplicações semelhantes à mencionada sofrerão uma aceleração pelo facto dos programas de inteligência artificial estarem cada vez mais dotados com a capacidade de interagir diretamente com os seres humanos (*chatbots*), mantendo conversas com eles e estabelecendo relações de proximidade, frequentemente, muito agradáveis e reconfortantes, uma vez que esses programas de inteligência artificial serão projetados para aprender a responder, de forma personalizada, às necessidades físicas e psicológicas dos seres humanos.

Esquecer que a inteligência artificial não é outro ser humano e que não pode propor princípios gerais, é muitas vezes um erro grave que decorre ou da profunda necessidade de os seres humanos encontrarem uma forma estável de companhia ou dum pressuposto subconsciente, isto é, do pressuposto de que as observações conseguidas mediante um mecanismo de cálculo sejam dotadas de qualidades de certeza indiscutível e de universalidade inquestionável.

No entanto, este pressuposto é arriscado, como demonstra o exame dos limites intrínsecos do próprio cálculo. A inteligência artificial usa operações algébricas a realizar segundo uma sequência lógica (por exemplo, se o valor de X for superior ao de Y, multiplica X por Y; caso contrário, divide X por Y). Este método de cálculo – o chamado “algoritmo” – não tem nem objetividade nem neutralidade [9]. Sendo baseado na álgebra, só pode examinar realidades formalizadas em termos numéricos [10].

Além disso, não se deve esquecer que os algoritmos projetados para resolver problemas muito complexos são tão sofisticados que é difícil, para os próprios programadores, compreenderem exatamente como conseguem alcançar os seus resultados. Esta tendência para a sofisticação corre o risco de se acelerar consideravelmente com a introdução de computadores quânticos que não funcionarão de acordo com circuitos binários (semicondutores ou microchips) mas segundo as leis, bastante complexas, da física quântica. Por outro lado, a introdução contínua de microchips, cada vez mais eficientes, já se tornou uma das causas da dominância do uso da inteligência artificial por parte das poucas nações que com ela estão equipadas.

Sofisticadas ou não, a qualidade das respostas que os programas de inteligência artificial fornecem depende, em última análise, dos dados que usam e da forma como são estruturados.

Por fim, gostaria de assinalar um último campo, no qual emerge claramente a complexidade do mecanismo da chamada inteligência artificial generativa (*Generative Artificial Intelligence*). Ninguém duvida que existem hoje magníficos instrumentos de acesso ao conhecimento, permitindo até o *self-learning* e o *self-tutoring* numa infinidade de áreas. Muitos de nós ficámos impressionados com as aplicações, facilmente disponíveis online, para redigir um texto ou

produzir uma imagem sobre qualquer tema ou assunto. Particularmente fascinados por esta perspectiva são os estudantes, que as utilizam de forma desproporcionada quando precisam de preparar trabalhos escolares.

No entanto, estes alunos, frequentemente muito mais preparados e habituados ao uso da inteligência artificial do que os seus professores, esquecem que a chamada inteligência artificial generativa, em sentido estrito, não é propriamente “generativa”. Na verdade, busca nos *big data* informações, elaborando-as segundo o estilo que lhe foi solicitado. Nem desenvolve conceitos nem análises novas. Repete as que encontra, dando-lhes uma forma apelativa. E quanto mais uma noção ou uma hipótese se repete, mais a considera legítima e válida. Em vez de “generativa”, ela é “reforçadora”, no sentido de que reorganiza os conteúdos existentes, contribuindo para consolidá-los, muitas vezes sem verificar se contêm erros ou preconceitos.

Deste modo, não só se corre o risco de legitimar *fake news* e de reforçar a vantagem de uma cultura dominante, mas igualmente de minar o processo educativo *in nuce*. A educação, que deveria fornecer aos estudantes a possibilidade de uma reflexão autêntica, corre o risco de se reduzir a uma repetição de noções que, cada vez mais, serão consideradas incontestáveis, simplesmente por causa da sua contínua repetição [11].

Colocar novamente a dignidade da pessoa no centro em vista de uma proposta ética comum

Uma observação mais geral deve-se agora acrescentar ao que já foi dito. A era de inovação tecnológica, que estamos atravessando, é acompanhada por uma conjuntura social particular e sem precedentes: sobre os grandes temas da vida social, torna-se cada vez mais difícil encontrar consensos. Mesmo em comunidades caracterizadas por uma certa continuidade cultural, surgem frequentemente debates acesos e confrontos que dificultam a produção de reflexões e soluções políticas comuns orientadas para a procura do que é bom e justo. Além da complexidade de visões legítimas que caracterizam a família humana, surge um fator que parece unir essas diferentes instâncias. Regista-se como que uma perda ou, pelo menos, um eclipse do sentido do humano e uma aparente insignificância do conceito de dignidade humana [12]. Parece que se está a perder o valor e o significado profundo de uma das categorias fundamentais do Ocidente: a categoria de pessoa humana. E assim, nesta era em que os programas de inteligência artificial questionam o ser humano e as suas ações, é precisamente a fraqueza do *ethos* ligado à perceção do valor e da dignidade da pessoa humana que corre o risco de ser o maior *vulnus* na implementação e no desenvolvimento destes sistemas. Efetivamente, não devemos esquecer que nenhuma inovação é neutra. A tecnologia nasce com um propósito e, com o seu impacto na sociedade humana, representa sempre uma forma de ordem nas relações sociais e uma disposição de poder, permitindo a uns realizar determinadas ações, enquanto a outros impede de concretizar outras. Esta dimensão constitutiva de poder da tecnologia inclui sempre, de uma maneira mais ou menos explícita, a visão do mundo de quem a criou e desenvolveu.

O mesmo se aplica aos programas de inteligência artificial. Para que estes sejam instrumentos de construção do bem e de um amanhã melhor, devem estar sempre orientados ao bem de cada ser humano. Devem ter uma inspiração ética.

Com efeito, a decisão ética é aquela que tem em conta não apenas os resultados de uma ação, mas também os valores em jogo e os deveres que deles derivam. Por isso, congratulei-me com a assinatura da *Rome Call for AI Ethics* [13], na cidade de Roma em 2020, e com o seu apoio a essa forma de moderação ética dos algoritmos e programas de inteligência artificial, que designei de “algor-ética” [14]. Num contexto plural e global, em que se apresentam também sensibilidades diferentes e hierarquias plurais nas escalas de valores, poderia parecer difícil encontrar uma única hierarquia de valores. Mas na análise ética, também podemos recorrer a outros tipos de instrumentos: se tivermos dificuldade em definir um único conjunto de valores globais, podemos encontrar alguns princípios comuns com os quais enfrentar e resolver eventuais dilemas ou conflitos da vida.

Foi por esta razão que nasceu a *Rome Call*: no termo “algor-ética” condensam-se uma série de princípios que se revelam uma plataforma global e plural capaz de encontrar o apoio de culturas, religiões, organizações internacionais e grandes empresas, que são protagonistas deste desenvolvimento.

A política de que precisamos

Porque é inerente ao seu mecanismo fundamental, não podemos esconder o risco concreto de que a inteligência artificial limita a visão do mundo a realidades expressáveis em números e encerradas em categorias pré-concebidas, excluindo o contributo de outras formas de verdade e impondo modelos antropológicos, socioeconómicos e culturais uniformes. O paradigma tecnológico incarnado pela inteligência artificial corre, pois, o risco de dar lugar a um paradigma muito mais perigoso, que identifiquei como “paradigma tecnocrático” [15]. Não podemos permitir que um instrumento tão poderoso e indispensável como a inteligência artificial reforce tal paradigma; pelo contrário, devemos fazer da inteligência artificial precisamente um baluarte contra a sua expansão.

E é exatamente aqui que a ação política é urgente, como lembra a Encíclica *Fratelli tutti*. É certo que «muitos possuem uma má noção da política, e não se pode ignorar que frequentemente, por trás deste facto, estão os erros, a corrupção e a ineficiência de alguns políticos. A isto vêm juntar-se as estratégias que visam enfraquecê-la, substituí-la pela economia ou dominá-la por alguma ideologia. E contudo poderá o mundo funcionar sem política? Poderá encontrar um caminho eficaz para a fraternidade universal e a paz social sem uma boa política?» [16].

A nossa resposta a estas últimas perguntas é: não! A política é necessária! Quero repetir nesta ocasião que «perante tantas formas de política mesquinhas e fixadas no interesse imediato [...], a

grandeza política mostra-se quando, em momentos difíceis, se trabalha com base em grandes princípios e pensando no bem comum a longo prazo. O poder político tem muita dificuldade em assumir este dever num projeto de nação e, mais ainda, num projeto comum para a humanidade presente e futura» [17].

Prezadas Senhoras, Ilustres Senhores!

Esta minha reflexão sobre os efeitos da inteligência artificial no futuro da humanidade leva-nos assim a considerar a importância de uma “política sã” para olharmos o nosso futuro com esperança e confiança. Como já disse noutros lugares, «a sociedade mundial tem graves carências estruturais que não se resolvem com remendos ou soluções rápidas meramente ocasionais. Há coisas que devem ser mudadas com reajustamentos profundos e transformações importantes. E só uma política sã poderia conduzir o processo, envolvendo os mais diversos setores e os conhecimentos mais variados. Desta forma, uma economia integrada num projeto político, social, cultural e popular que vise o bem comum pode “abrir caminho a oportunidades diferentes, que não implica frear a criatividade humana nem o seu sonho de progresso, mas orientar esta energia por novos canais” (*Laudato si'*, 191)» [18].

É precisamente este o caso da inteligência artificial. Cabe a todos nós fazer um bom uso dela, e cabe à política criar as condições para que essa boa utilização seja possível e frutuosa.

Obrigado.

[1] Francisco, *Mensagem para a LVII Jornada Mundial da Paz* (1/II/2024), 1.

[2] Cf. *ibid.*

[3] Cf. *ibid.*, 2.

[4] Esta ambivalência foi já percebida pelo Papa São Paulo VI no seu *Discurso aos funcionários do “Aloysianum de Gallarate”* (19/VI/1964).

[5] Cf. A. Gehlen, *L'uomo. La sua natura e il suo posto nel mondo*, Milano 1983, 43.

[6] Cf. Francisco, Carta enc. *Laudato si'* (24/VI/2015), 102-114.

[7] Cf. Francisco, *Mensagem para a LVII Jornada Mundial da Paz* (1/II/2024), 3.

[8] As intuições de Marshall McLuhan e de John M. Culkin são particularmente pertinentes para as consequências do uso da inteligência artificial.

[9] Cf. Francisco, *Discurso aos participantes da Plenária da Pontifícia Academia para a Vida* (28/II/2020).

[10] Cf. Francisco, *Mensagem para a LVII Jornada Mundial da Paz* (1/II/2024), 4.

[11] Cf. Francisco, *Mensagem para a LVII Jornada Mundial da Paz* (1/II/2024), 3 e 7.

[12] Cf. Dicastério para a Doutrina da Fé, Declaração *Dignitas infinita* sobre a dignidade humana (2/IV/2024).

[13] Cf. Francisco, *Discurso aos participantes da Plenária da Pontifícia Academia para a Vida* (28/II/2020).

[14] Cf. Francisco, *Discurso aos participantes no Congresso sobre a dignidade dos menores no mundo digital* (14/XI/2019); Cf. *Discurso aos participantes da Plenária da Pontifícia Academia para a Vida* (28/II/2020).

[15] Para uma exposição mais ampla, remeto para a minha Carta Encíclica *Laudato si'* sobre o cuidado da casa comum de 24 de maio de 2015.

[16] Carta Enc. *Fratelli tutti* sobre a fraternidade e a amizade social (3/X/2020), 176

[17] *Ibid.*, 178.

[18] *Ibid.*, 179.