



Bruxelas, 8.4.2019
COM(2019) 168 final

**COMUNICAÇÃO DA COMISSÃO AO PARLAMENTO EUROPEU, AO
CONSELHO, AO COMITÉ ECONÓMICO E SOCIAL EUROPEU E AO COMITÉ
DAS REGIÕES**

Aumentar a confiança numa inteligência artificial centrada no ser humano

1. INTRODUÇÃO – A ESTRATÉGIA EUROPEIA PARA A INTELIGÊNCIA ARTIFICIAL

A inteligência artificial (IA) tem potencial para transformar o nosso mundo num mundo melhor: pode melhorar os cuidados de saúde, reduzir o consumo de energia, tornar os automóveis mais seguros e permitir aos agricultores uma utilização mais eficiente da água e dos recursos naturais. A IA pode ser utilizada para prever as alterações ambientais e climáticas e melhorar a gestão dos riscos financeiros, e fornece os instrumentos para o fabrico, com menos resíduos, de produtos adaptados às nossas necessidades. A IA pode igualmente ajudar a detetar fraudes e ameaças à cibersegurança, e permite que as autoridades de aplicação da lei lutem mais eficientemente contra a criminalidade.

A IA pode beneficiar toda a sociedade e a economia. É uma tecnologia estratégica que está a ser desenvolvida e utilizada a um ritmo acelerado em todo o mundo. No entanto, a IA também traz consigo novos desafios para o futuro do trabalho e levanta questões de ordem jurídica e ética.

A Comissão publicou, em abril de 2018, uma estratégia europeia¹ para fazer face a estes desafios e tirar o melhor partido das oportunidades oferecidas pela IA. A estratégia coloca as pessoas no centro do desenvolvimento da IA — uma **IA centrada no ser humano**. Trata-se de uma abordagem em três vertentes para reforçar a capacidade industrial e tecnológica da UE e a integração da IA em toda a economia, preparar as mudanças socioeconómicas e garantir um quadro ético e jurídico adequado.

Para concretizar a estratégia da IA, **a Comissão desenvolveu, em conjunto com os Estados-Membros, um plano coordenado sobre IA²**, que apresentou em dezembro de 2018, de modo a criar sinergias, agregar dados — a matéria-prima das aplicações de IA — e aumentar os investimentos conjuntos. O plano tem por objetivo fomentar a cooperação transfronteiriça e mobilizar todos os intervenientes para o aumento dos investimentos públicos e privados para, **pelo menos, 20 mil milhões de EUR** anuais, durante a próxima década³. A Comissão duplicou os seus investimentos em IA no Horizonte 2020 e prevê investir 1 000 milhões de EUR anuais no Horizonte Europa e no Programa Europa Digital, apoiando, nomeadamente, os espaços de dados comuns nos setores da saúde, dos transportes e da indústria transformadora, bem como grandes instalações de experimentação, como hospitais e infraestruturas inteligentes para veículos automatizados, e uma agenda de investigação estratégica.

Para aplicar tal agenda estratégica comum de investigação, inovação e implantação, a Comissão intensificou o seu **diálogo com todas as partes interessadas pertinentes**, designadamente da indústria, dos institutos de investigação e das autoridades públicas. O novo programa Europa Digital será essencial também para tornar a IA acessível às pequenas e médias empresas em todos os Estados-Membros, através de polos de inovação digital, instalações de ensaio e experimentação acrescidas, espaços de dados e programas de formação.

Graças à sua reputação no que se refere à segurança e à elevada qualidade dos produtos, a abordagem ética da IA pela Europa reforça a confiança dos cidadãos no desenvolvimento digital e visa criar uma vantagem competitiva para as empresas europeias de IA. O objetivo da presente comunicação é lançar uma fase de pilotagem abrangente que envolva todas as

¹ COM(2018) 237.

² COM(2018) 795 final.

³ Para ajudar a alcançar este objetivo, a Comissão propôs, para o próximo período de programação, 2021-2027, que a União afete ao investimento na IA pelo menos 1 000 milhões de EUR por ano de financiamento proveniente dos programas Horizonte Europa e Digital.

partes interessadas, a fim de testar a aplicação prática de orientações éticas para o desenvolvimento e a utilização da IA.

2. CRIAR CONFIANÇA NA IA CENTRADA NO SER HUMANO

A estratégia europeia para a IA e o plano coordenado tornam claro que **a confiança é um pré-requisito para garantir uma abordagem da IA centrada no ser humano**: A IA não é um fim em si mesmo, mas sim um instrumento que tem de servir as pessoas com o objetivo último de aumentar o bem-estar humano. Para tal, **deve ser assegurada a fiabilidade da IA**. Os valores em que as nossas sociedades se baseiam devem ser plenamente integrados na forma como a IA evolui.

A União baseia-se nos **valores do respeito pela dignidade humana, liberdade, democracia, igualdade, Estado de direito e respeito pelos direitos humanos**, nomeadamente os direitos das minorias⁴. Estes valores são comuns às sociedades de todos os Estados-Membros em que prevalecem o pluralismo, a não discriminação, a tolerância, a justiça, a solidariedade e a igualdade. Além disso, a **Carta dos Direitos Fundamentais da UE** reúne num único texto todos os direitos individuais, cívicos, políticos, económicos e sociais reconhecidos aos povos da UE.

A UE dispõe de um **quadro regulamentar sólido**, que estabelecerá a norma mundial de uma IA centrada no ser humano. O Regulamento Geral sobre a Proteção de Dados assegura um elevado nível de proteção dos dados pessoais e exige a aplicação de medidas para assegurar a proteção de dados desde a conceção e por defeito⁵. O Regulamento relativo ao livre fluxo de dados não pessoais elimina os obstáculos à livre circulação de dados não pessoais e assegura o tratamento de todas as categorias de dados em toda a Europa. O Regulamento Cibersegurança, recentemente adotado, ajudará a reforçar a confiança no mundo em linha; a proposta de Regulamento Privacidade Eletrónica⁶ visa igualmente este objetivo.

Porém, a IA coloca novos desafios, uma vez que permite que as máquinas «aprendam», e tomem e apliquem decisões sem intervenção humana. Em breve, este tipo de funcionalidade tornar-se-á padrão em muitos tipos de bens e serviços, desde os telefones inteligentes até aos automóveis automatizados, aos robôs e às aplicações em linha. Contudo, as decisões tomadas por algoritmos podem resultar de dados incompletos, não sendo, por conseguinte, fiáveis, podem ser alvo de ciberataques, e podem ser tendenciosos ou simplesmente errados. A aplicação indevida da tecnologia à medida que se desenvolve conduziria, por conseguinte, a resultados problemáticos, bem como à relutância dos cidadãos em aceitarem ou utilizarem esta tecnologia.

Em vez disso, a tecnologia da IA deve ser desenvolvida de forma que coloque as pessoas no seu centro e seja, assim, digna da confiança do público. Tal implica que as aplicações de IA devem não só ser coerentes com a legislação, como também respeitar os princípios éticos e assegurar que a sua aplicação evita danos não intencionais. A diversidade em termos de

⁴ A UE é também parte na Convenção da ONU sobre os Direitos das Pessoas com Deficiência.

⁵ Regulamento (UE) 2016/679. O Regulamento Geral sobre a Proteção de Dados (RGPD) garante a livre circulação de dados pessoais na União. Contém disposições sobre a tomada de decisões unicamente com base no tratamento automatizado, incluindo a definição de perfis. As pessoas em causa têm o direito de serem informadas da existência de processos de decisão autónomos e de receberem informações relevantes sobre a lógica subjacente à tomada automatizada de decisões, bem como sobre o significado e as consequências previstas do seu tratamento. Têm também o direito, em tais casos, de obterem intervenção humana, de exprimirem o seu ponto de vista e de contestarem a decisão.

⁶ COM(2017) 10.

género, raça ou etnia, religião ou crença, deficiência e idade deve ser assegurada em todas as fases do desenvolvimento da IA. As aplicações de IA devem capacitar os cidadãos e respeitar os seus direitos fundamentais. Devem ter como objetivo aumentar as capacidades das pessoas, não substituí-las, e permitir o acesso de pessoas com deficiência.

Por conseguinte, são necessárias **orientações éticas** que se baseiem no quadro regulamentar existente e que sejam aplicadas por criadores, fornecedores e utilizadores da IA no mercado interno, estabelecendo condições de concorrência equitativas em todos os Estados-Membros. É por esta razão que a Comissão criou um **Grupo de peritos de alto nível sobre a IA**⁷, que representa um vasto leque de partes interessadas, e foi incumbido da elaboração de orientações éticas no domínio da IA, bem como da preparação de um conjunto de recomendações para uma política de IA mais ampla. Ao mesmo tempo, foi criada a **Aliança Europeia para a IA**⁸, uma plataforma aberta com mais de 2700 membros, para dar um maior contributo para o trabalho do Grupo de peritos de alto nível sobre a IA.

Em dezembro de 2018, o Grupo de peritos de alto nível sobre a IA publicou um primeiro projeto das orientações éticas. Na sequência de uma consulta às **partes interessadas**⁹ e de **reuniões com representantes dos Estados-Membros**¹⁰, o Grupo de peritos apresentou à Comissão, em março de 2019, um documento revisto. Nas suas observações até à data, as partes interessadas reagiram favoravelmente à natureza prática das orientações e às orientações concretas que oferecem aos criadores, fornecedores e utilizadores da IA sobre a forma de garantir a fiabilidade.

2.1. Orientações para uma IA de confiança elaboradas pelo Grupo de peritos de alto nível sobre a IA

As orientações elaboradas pelo Grupo de peritos de alto nível sobre a IA, a que a presente comunicação se refere¹¹, baseiam-se, em especial, no trabalho realizado pelo Grupo Europeu de Ética para as Ciências e as Novas Tecnologias e pela Agência dos Direitos Fundamentais.

As orientações determinam que, para se alcançar uma «IA de confiança», são necessárias três componentes: 1) Conformidade com a legislação, 2) Respeito dos princípios éticos e 3) Robustez.

Com base nestas três componentes e nos valores europeus definidos na secção 2, as orientações identificam sete requisitos essenciais que as aplicações de IA devem respeitar para serem consideradas de confiança. As orientações incluem igualmente uma lista de avaliação para ajudar a verificar se estes requisitos são cumpridos.

Os sete requisitos principais são os seguintes:

⁷ <https://ec.europa.eu/digital-single-market/en/high-level-expert-group-artificial-intelligence>

⁸ <https://ec.europa.eu/digital-single-market/en/european-ai-alliance>

⁹ Esta consulta resultou em comentários de 511 organizações, associações, empresas, institutos de investigação, particulares e outros. O resumo das reações recebidas está disponível em: https://ec.europa.eu/futurium/en/system/files/ged/consultation_feedback_on_draft_ai_ethics_guidelines_4.pdf

¹⁰ Os trabalhos do grupo de peritos foram recebidos de forma positiva pelos Estados-Membros, tendo as conclusões do Conselho adotadas em 18 de fevereiro de 2019 registado, *inter alia*, a futura publicação das orientações éticas e apoiado os esforços da Comissão para introduzir uma abordagem ética da UE na cena mundial: <https://data.consilium.europa.eu/doc/document/ST-6177-2019-INIT/pt/pdf>

¹¹ <https://ec.europa.eu/futurium/en/ai-alliance-consultation/guidelines#Top>

- Iniciativa e controlo por humanos
- Robustez e segurança
- Privacidade e governação dos dados
- Transparência
- Diversidade, não discriminação e equidade
- Bem-estar societal e ambiental
- Responsabilização

Embora estes requisitos se destinem a ser aplicados a todos os sistemas de IA em diferentes setores e indústrias, para uma aplicação concreta e proporcionada, deve ser tido em conta o contexto específico em que são aplicados mediante uma abordagem baseada no impacto. A título de exemplo, a sugestão, por uma aplicação de IA, de um livro inadequado para leitura é muito menos perigosa do que o diagnóstico errado de um cancro, pelo que a primeira pode ser objeto de uma supervisão menos rigorosa.

As orientações elaboradas pelo Grupo de peritos de alto nível sobre a IA não são vinculativas e, como tal, não criam novas obrigações jurídicas. Todavia, muitas das disposições existentes (e frequentemente de domínio ou utilização específicos) do direito da União refletem já um ou mais destes requisitos essenciais; por exemplo, a segurança, as normas sobre proteção dos dados pessoais, a privacidade e a proteção do ambiente.

A Comissão congratula-se com o trabalho do Grupo de peritos de alto nível sobre a IA e considera-o um contributo valioso para o seu processo de elaboração de políticas.

2.2. Requisitos fundamentais para uma IA de confiança

A Comissão apoia os seguintes requisitos fundamentais para uma inteligência artificial de confiança, que se baseiam nos valores europeus. Incentiva as partes interessadas a aplicarem os requisitos e a testarem a lista de avaliação que os torna operacionais, a fim de criar um clima de confiança para o êxito do desenvolvimento e da utilização da IA. A Comissão congratula-se com as reações das partes interessadas, úteis para avaliar se esta lista de apreciações apresentada nas orientações carece de maior ajustamento.

I. Iniciativa e controlo por humanos

Os sistemas de IA devem ajudar os indivíduos a fazerem escolhas melhores e mais informadas, em conformidade com os seus objetivos. Devem agir como facilitadores de uma sociedade próspera e equitativa, apoiando a atividade humana e promovendo os **direitos fundamentais**, e não reduzindo, limitando ou guiando a autonomia humana. O **bem-estar geral do utilizador** deve estar no centro da funcionalidade do sistema.

O controlo humano contribui para garantir que um sistema de IA não prejudica a autonomia humana nem causa outros efeitos adversos. Em função do sistema específico baseado em inteligência artificial e da sua área de aplicação, devem ser assegurados os graus adequados de **medidas de controlo**, incluindo a capacidade de adaptação, a

exatidão e a explicação dos sistemas baseados em inteligência artificial¹². A **supervisão** pode ser efetuada através de mecanismos de governação, como a garantia de uma abordagem que respeite os princípios da intervenção humana, da supervisão humana e da detenção do controlo por humanos¹³ («*human-in-the-loop*», «*human-on-the-loop*» ou «*human-in-command*»). Deve garantir-se que as autoridades públicas têm capacidade para exercer os seus poderes de supervisão em conformidade com os seus mandatos. Dado que todos os outros fatores são iguais, quanto menor for o grau de supervisão de um ser humano sobre um sistema de IA, maior é a necessidade de ensaios mais aprofundados e de uma governação mais rigorosa.

II. Robustez e segurança

Uma IA de confiança exige que os algoritmos sejam seguros, fiáveis e suficientemente robustos para lidar com erros ou incoerências durante todas as fases do ciclo de vida do sistema de IA e para lidar adequadamente com resultados errados. Os sistemas de IA têm de ser **fiáveis**, suficientemente **resilientes** contra os ataques não dissimulados e tentativas mais subtis de manipular dados ou os próprios algoritmos, e devem assegurar um **plano de recurso** em caso de problemas. As suas decisões devem ser **exatas**, ou, pelo menos, refletir corretamente o seu nível de precisão, e os seus resultados devem ser **reprodutíveis**.

Além disso, os sistemas de IA devem integrar mecanismos de proteção e de segurança desde a conceção para garantir que são **comprovadamente seguros** em todas as fases, tendo em conta a segurança física e mental de todas as partes envolvidas. Tal inclui a minimização e, sempre que possível, a reversibilidade de consequências ou de erros não intencionais na operação do sistema. Devem ser criados processos destinados a clarificar e avaliar os riscos potenciais associados à utilização de sistemas de IA em diferentes áreas de aplicação.

III. Privacidade e governação dos dados

A privacidade e a **proteção de dados** devem ser garantidas em **todas as fases** do ciclo de vida do sistema de IA. Os registos digitais de comportamento humano podem permitir aos sistemas de IA inferir não só as preferências dos indivíduos, a sua idade e o seu género, como também a sua orientação sexual, religião ou opinião política. A fim de permitir que as pessoas confiem no tratamento de dados, há que assegurar-lhes o pleno controlo sobre os seus próprios dados e a não utilização dos dados que lhes dizem respeito para as prejudicar ou discriminar.

Além da proteção da privacidade e dos dados pessoais, devem ser cumpridos requisitos

¹² O Regulamento Geral sobre a Proteção de Dados dá aos indivíduos o direito de não estarem sujeitos a uma decisão baseada exclusivamente no tratamento automatizado se tal produzir efeitos jurídicos sobre os utilizadores ou os afetar de forma similar (artigo 22.º do RGPD) .

¹³ A intervenção humana deve ocorrer em cada ciclo de decisão do sistema, o que, em muitos casos, não é nem possível nem desejável («*human-in-the-loop*» - HITL). A supervisão humana refere-se à capacidade de intervenção humana durante o ciclo de conceção do sistema e de monitorização do funcionamento do sistema («*human-on-the-loop*» - HOTL). O conceito de «detenção de controlo por humanos» refere-se à capacidade de supervisionar a atividade global do sistema de IA (incluindo o seu impacto mais vasto a nível económico, social, jurídico e ético) e a capacidade de decidir quando e como utilizar o sistema em qualquer situação específica («*human-in-command*» - HIC). Este conceito pode incluir a decisão de não utilizar um sistema de IA numa determinada situação, de estabelecer níveis de apreciação humana durante a utilização do sistema e de garantir a capacidade de anular uma decisão tomada pelo sistema.

para garantir sistemas de IA de elevada qualidade. A qualidade dos conjuntos de dados utilizados é fundamental para o desempenho dos sistemas de IA. Os dados recolhidos estes podem refletir enviesamentos socialmente construídos ou conter inexatidões, erros e lapsos. Esta questão deve ser abordada antes da formação de um sistema de IA com qualquer conjunto de dados. Além disso, deve ser assegurada a **integridade** dos dados. Os processos e conjuntos de dados utilizados devem ser testados e documentados em cada fase, como as do planeamento, da formação, dos ensaios e da implantação. Este requisito deve aplicar-se igualmente aos sistemas de IA que não sejam desenvolvidos internamente, mas adquiridos noutros locais. Por último, o **acesso** aos dados deve ser adequadamente regulado e controlado.

IV. Transparência

Deve ser assegurada a **rastreabilidade** dos sistemas de IA; é importante registar e documentar as decisões tomadas pelos sistemas, bem como todo o processo (incluindo uma descrição da recolha e rotulagem de dados e uma descrição do algoritmo utilizado) que deu origem às decisões. Por este motivo, deve ser prestada, na medida do possível, uma **explicação** do processo de tomada de decisões algorítmicas, adaptado às pessoas envolvidas. Devem ser prosseguidos os trabalhos de investigação em curso com vista ao desenvolvimento de mecanismos de explicação. Além disso, devem estar disponíveis explicações sobre o grau de influência e orientação de um sistema de IA no processo de tomada de decisão organizacional, as escolhas de conceção do sistema, bem como a justificação da sua implantação (garantindo, assim, não só a transparência dos dados e do sistema, mas também a transparência do modelo empresarial).

Por último, é importante **comunicar** adequadamente as capacidades e limitações do sistema de IA às diferentes partes interessadas envolvidas, de uma forma adequada ao caso em concreto. Além disso, os sistemas de IA devem ser identificáveis como tal, assegurando que os utilizadores sabem que estão a interagir com um sistema de IA e que pessoas são responsáveis por este sistema.

V. Diversidade, não discriminação e equidade

Os conjuntos de dados utilizados pelos sistemas de IA (tanto para a aprendizagem como para o funcionamento) podem ser afetados pela inclusão inadvertida de modelos de má conduta, incompletos e de má governação. A subsistirem, tais enviesamentos poderiam conduzir à discriminação (in)direta. Podem resultar danos também da exploração intencional de preconceitos (dos consumidores) ou da concorrência desleal. Além disso, o modo de desenvolvimento dos sistemas de IA (por exemplo, a forma de escrita do código de programação de um algoritmo) também pode ser tendencioso. Estas preocupações devem ser abordadas desde o início do desenvolvimento do sistema.

A constituição de **equipas de conceção diversas** e a criação de mecanismos que garantam a **participação**, em especial dos cidadãos, no desenvolvimento da IA também podem ajudar a dar resposta a estas preocupações. É aconselhável consultar as partes interessadas que possam ser direta ou indiretamente afetadas pelo sistema ao longo de todo o seu ciclo de vida. Os sistemas de IA devem ter em conta toda a gama de capacidades, competências e requisitos humanos, bem como garantir a acessibilidade através de uma abordagem de conceção universal, a fim de assegurar a igualdade de acesso das pessoas com deficiência.

VI. Bem-estar societal e ambiental

Para que a IA seja de confiança, há que ter em conta o seu impacto sobre o **ambiente e outros seres sencientes**. Idealmente, todos os seres humanos, incluindo as gerações futuras, devem beneficiar da biodiversidade e de um ambiente habitável. Devem, por conseguinte, ser incentivadas a sustentabilidade e a **responsabilidade ecológica** dos sistemas de IA. O mesmo se aplica às soluções de inteligência artificial que abordam domínios de interesse global, como, por exemplo, os Objetivos de Desenvolvimento Sustentável das Nações Unidas.

Além disso, o impacto dos sistemas de IA deve ser considerado não apenas numa perspetiva individual, mas também na perspetiva de **toda a sociedade**. A utilização de sistemas de IA deve merecer uma atenção especial em situações relacionadas com o processo democrático, incluindo a tomada de decisões políticas e os contextos eleitorais. Além disso, há que ter em conta o **impacto social** da IA. Embora os sistemas de IA possam ser utilizados para reforçar as competências sociais, podem contribuir igualmente para a sua deterioração.

VII. Responsabilização

Devem ser criados mecanismos para assegurar a responsabilidade e a responsabilização dos sistemas de IA, bem como dos seus resultados, tanto antes como depois da sua aplicação. A **auditabilidade** dos sistemas de IA é fundamental neste contexto, uma vez que a avaliação dos sistemas de IA por auditores internos e externos, bem como a disponibilidade de tais relatórios de avaliação, contribuem fortemente para a fiabilidade da tecnologia. A auditabilidade externa deve ser assegurada, em especial, nas aplicações que afetam os direitos fundamentais, incluindo as aplicações críticas para a segurança.

Os **potenciais impactos negativos** dos sistemas de IA devem ser identificados, avaliados, documentados e minimizados. A utilização de avaliações de impacto facilita este processo. Estas avaliações devem ser proporcionais à dimensão dos riscos que os sistemas de IA representam. Os **compromissos** entre os requisitos, muitas vezes inevitáveis, devem ser abordados de uma forma racional e metodológica, devendo ser tidos em conta. Por último, prever-se mecanismos acessíveis, que garantam uma **reparação adequada**, para as ocorrências de impactos adversos e injustos.

2.3. Próximas etapas: uma fase de pilotagem que envolva as partes interessadas à escala mais ampla possível

A obtenção de um consenso sobre estes requisitos essenciais para os sistemas de IA constitui um primeiro marco importante para a elaboração de orientações para uma IA ética. Como próxima etapa, a Comissão assegurará que estas orientações podem ser testadas e aplicadas na prática.

Para o efeito, a Comissão lançará uma fase de pilotagem específica, destinada a obter reações estruturadas das partes interessadas. Este exercício centrar-se-á, em especial, na lista de avaliação que o grupo de peritos de alto nível elaborou para cada um dos requisitos essenciais.

Este trabalho terá duas vertentes: i) uma fase de pilotagem das orientações que envolve as partes interessadas que desenvolvem ou utilizam a IA, incluindo as administrações públicas; ii) um processo contínuo de consulta das partes interessadas e de sensibilização dos Estados-

Membros e de diferentes grupos de partes interessadas, incluindo a indústria e os setores dos serviços.

- (i) A partir de junho de 2019, todas as partes interessadas e particulares serão convidados a testar a lista de avaliação e a comunicar informações sobre a forma de a melhorar. Além disso, o grupo de peritos de alto nível para a IA procederá a uma análise aprofundada com as partes interessadas do setor privado e do setor público, a fim de recolher informações mais pormenorizadas sobre as possibilidades de aplicação das orientações numa vasta gama de domínios de aplicação. Todas as observações sobre a funcionalidade e viabilidade das orientações serão avaliadas até ao final de 2019.
- (ii) Paralelamente, a Comissão organizará outras atividades de sensibilização, oferecendo aos representantes do Grupo de peritos de alto nível sobre a IA a oportunidade de apresentarem as orientações às partes interessadas dos Estados-Membros, incluindo a indústria e os setores dos serviços, e de lhes proporcionar uma oportunidade suplementar de comentarem e contribuírem para as orientações relativas à IA.

A Comissão terá em conta o trabalho do grupo de peritos em matéria de ética no domínio da condução conectada e automatizada¹⁴, assim como o trabalho com projetos de investigação financiados pela UE no domínio da IA e com as parcerias público-privadas relevantes para a aplicação dos requisitos essenciais¹⁵. Por exemplo, a Comissão apoiará, em coordenação com os Estados-Membros, o desenvolvimento de uma base de dados comum de imagens de saúde dedicadas inicialmente às formas mais comuns de cancro, de modo a que os algoritmos possam ser treinados para o diagnóstico de sintomas com uma precisão muito elevada. Do mesmo modo, a cooperação entre a Comissão e os Estados-Membros permite a abertura de um número crescente de corredores transfronteiriços para a realização de ensaios de veículos conectados e automatizados. As orientações devem ser aplicadas nestes projetos e testadas, e os resultados serão tidos em conta no processo de avaliação.

A fase de pilotagem e a consulta das partes interessadas beneficiarão da contribuição da Aliança Europeia da Inteligência Artificial e da AI4EU, a plataforma de IA a pedido. O projeto AI4EU¹⁶, lançado em janeiro de 2019, reúne algoritmos, ferramentas, conjuntos de dados e serviços para ajudar as organizações, em especial as pequenas e médias empresas, a aplicarem soluções de inteligência artificial. A Aliança Europeia da Inteligência Artificial, juntamente com a AI4EU, continuará a mobilizar o ecossistema da IA em toda a Europa, tendo também em vista orientar as diretrizes de ética da IA e promover o respeito pela IA centrada no ser humano.

No início de 2020, com base na avaliação das reações recebidas durante a fase de pilotagem, o **Grupo de peritos de alto nível sobre a IA procederá à revisão e à atualização das orientações**. Com base nessa análise e na experiência adquirida, a **Comissão avaliará os resultados e proporá eventuais etapas seguintes**.

A IA ética é uma proposta vantajosa para todos. A garantia do respeito dos valores fundamentais não só é essencial por si mesma, como também facilita a aceitação pelo público

¹⁴ Ver a Comunicação da Comissão «Rumo à mobilidade automatizada: uma estratégia da UE para a mobilidade do futuro», COM (2018) 283.

¹⁵ No âmbito do Fundo Europeu de Defesa, a Comissão elaborará também orientações éticas específicas para a avaliação das propostas de projetos no domínio da inteligência artificial para a defesa.

¹⁶ <https://ec.europa.eu/digital-single-market/en/news/artificial-intelligence-ai4eu-project-launches-1-january-2019>

e aumenta a vantagem competitiva das empresas europeias da IA através da criação de uma marca de IA de confiança, centrada no ser humano e conhecida pela ética e pela segurança dos produtos. Essa vantagem tem por base, de um modo mais geral, a forte reputação das empresas europeias no fornecimento de produtos seguros de elevada qualidade. A fase de pilotagem contribuirá para assegurar que os produtos da IA cumprem esta promessa.

2.4. Rumo a orientações internacionais para a ética no domínio da IA

Os debates internacionais em matéria de ética no domínio da IA intensificaram-se após a presidência japonesa do G7 ter colocado o tema no topo da agenda em 2016. Tendo em conta as interligações internacionais do desenvolvimento da IA em termos de circulação de dados, desenvolvimento algorítmico e investimentos em investigação, **a Comissão prosseguirá os seus esforços para que a abordagem da União seja aceite à escala mundial e para a formação de um consenso sobre uma IA centrada no ser humano**¹⁷.

O trabalho realizado pelo Grupo de peritos de alto nível sobre a IA, mais especificamente a lista de requisitos e o processo de compromisso com as partes interessadas, proporciona à Comissão um contributo adicional valioso para os debates internacionais. A União Europeia pode assumir um papel de liderança na elaboração de orientações internacionais em matéria de IA e, se possível, num mecanismo de avaliação conexo.

Por conseguinte, a Comissão vai:

Reforçar a cooperação com parceiros que partilhem uma visão idêntica:

- explorando as possibilidades de convergência com os projetos de orientações éticas de países terceiros (por exemplo, Japão, Canadá e Singapura) e, com base no grupo de países que partilham as mesmas ideias, preparar um debate mais amplo, apoiado por ações de execução do Instrumento de Parceria para a cooperação com países terceiros¹⁸;
- explorando formas de contribuição das empresas de países terceiros e das organizações internacionais para a fase de pilotagem das orientações através de testes e validação.

Continuar a desempenhar um papel ativo nos debates e iniciativas internacionais:

- contribuindo para fóruns multilaterais como o G7 e o G20;
- encetando diálogos com países terceiros e organizando reuniões bilaterais e multilaterais para criar um consenso sobre a IA centrada no ser humano;

¹⁷ A alta representante da União para os Negócios Estrangeiros e a Política de Segurança tenciona efetuar, com o apoio da Comissão, consultas nas instâncias das Nações Unidas, no Painel Mundial de Tecnologia e noutros fóruns multilaterais, e, em particular, coordenar propostas destinadas a fazer face a estes complexos desafios em matéria de segurança.

¹⁸ Regulamento (UE) n.º 234/2014 do Parlamento Europeu e do Conselho, de 11 de março de 2014, que cria um Instrumento de Parceria para a cooperação com países terceiros (JO L 77 de 15.3.2014, p. 77). Por exemplo, o projeto previsto sobre «Uma aliança internacional para uma abordagem centrada no ser humano da inteligência artificial» facilitará iniciativas conjuntas com parceiros que partilham as mesmas ideias, a fim de promover orientações éticas e de adotar princípios comuns e conclusões operacionais. Este projeto permitirá que a UE e os países com as mesmas ideias debatam as conclusões operacionais resultantes das orientações éticas sobre IA propostas pelo grupo de peritos, a fim de alcançarem uma abordagem comum. Além disso, prevê o acompanhamento da adoção de tecnologias de IA a nível mundial. Por último, o projeto prevê a organização de atividades de diplomacia pública em eventos internacionais, por exemplo, do G7, do G20 e da Organização de Cooperação e de Desenvolvimento Económicos.

- contribuindo para as atividades de normalização pertinentes em organizações internacionais de desenvolvimento de normas, a fim de promover esta visão; -
- reforçando a recolha e a divulgação de conhecimentos sobre políticas públicas, em conjunto com as organizações internacionais relevantes.

3. CONCLUSÕES

A UE assenta num conjunto de valores fundamentais e construiu um quadro regulador sólido e equilibrado para esses valores. Dada a novidade e os desafios específicos que a tecnologia da IA coloca, é necessário estabelecer, com base nesse quadro regulador, orientações éticas para o desenvolvimento e a utilização dessa tecnologia. A IA só pode ser considerada de confiança se for desenvolvida e utilizada de forma que respeite os valores éticos amplamente partilhados.

Tendo em vista este objetivo, a Comissão congratula-se com os contributos preparados pelo Grupo de peritos de alto nível sobre a IA. Com base nos requisitos essenciais para que a IA seja considerada de confiança, a Comissão irá lançar uma fase de pilotagem específica, a fim de garantir a aplicação prática das orientações éticas resultantes para o desenvolvimento e a utilização da IA. A Comissão trabalhará igualmente com todas as partes interessadas e os nossos parceiros internacionais para forjar um amplo consenso social sobre uma IA centrada no ser humano.

A dimensão ética da IA não constitui um luxo nem um acessório, sendo, antes, um elemento que terá de ser plenamente integrado no seu desenvolvimento. Empenhados que estamos numa IA centrada no ser humano e baseada na confiança, salvaguardamos o respeito pelos nossos valores sociais fundamentais e delineamos uma marca distintiva para a Europa e a sua indústria como líder na IA de vanguarda, que pode ser digna de confiança em todo o mundo.

Além disso, a fim de assegurar o desenvolvimento ético da IA na Europa, num contexto mais amplo, a Comissão segue uma abordagem global que inclui, em particular, as seguintes linhas de ação a implementar até ao terceiro trimestre de 2019:

- Criação de um conjunto de **redes de centros de excelência em investigação no domínio da IA** através do Horizonte 2020. Seleção de quatro redes, no máximo, centrada nos grandes desafios científicos ou tecnológicos, tais como os mecanismos de explicação e a interação homem-máquina avançada, que são ingredientes essenciais para uma IA de confiança.
- Criação de **redes de polos de inovação digital**¹⁹ centrados na IA, na indústria transformadora e em grandes volumes de dados.
- Em conjunto com os Estados-Membros e as partes interessadas, a Comissão iniciará discussões preparatórias para a conceção e a aplicação de **um modelo de partilha de dados e para utilização, da melhor forma possível, dos espaços de dados comuns**, com especial incidência nos transportes, nos cuidados de saúde e na indústria transformadora²⁰.

¹⁹ <http://s3platform.jrc.ec.europa.eu/digital-innovation-hubs>

²⁰ Os recursos necessários serão mobilizados a partir do Programa-Quadro Horizonte 2020 (no âmbito do qual cerca de 1,5 mil milhões de EUR são consagrados à IA para o período 2018-2020) e do seu programa sucessor Horizonte Europa, a parte digital do Mecanismo Interligar a Europa e, em especial, do futuro Programa Europa Digital. Os projetos basear-se-ão também em recursos do setor privado e dos programas dos Estados-Membros.

Além disso, a Comissão está a elaborar um relatório sobre os desafios colocados pela IA aos quadros de segurança e de responsabilidade e um documento de orientação sobre a aplicação da Diretiva em matéria de responsabilidade decorrente dos produtos²¹. Simultaneamente, a Empresa Comum Europeia para a Computação de Alto Desempenho (EuroHPC)²² desenvolverá a próxima geração de supercomputadores, uma vez que a capacidade de computação é essencial para o processamento de dados e o treino da IA, e a Europa precisa de dominar toda a cadeia de valor digital. A parceria em curso com os Estados-Membros e a indústria sobre componentes e sistemas microeletrónicos (ECSEL)²³, bem como a Iniciativa do Processador Europeu²⁴, contribuirão para o desenvolvimento de tecnologias de processadores de baixa potência para computação fiável e de alto desempenho e para computação de proximidade.

Tal como o trabalho sobre orientações éticas para a IA, todas estas iniciativas assentam numa **estreita cooperação entre todas as partes interessadas**, os Estados-Membros, a indústria, os agentes da sociedade e os cidadãos. De um modo geral, a abordagem da inteligência artificial pela Europa demonstra que a competitividade económica e a confiança da sociedade se devem basear nos mesmos valores fundamentais e reforçar-se mutuamente.

²¹ Ver a Comunicação «A inteligência Artificial para a Europa», COM (2018) 237.

²² <https://eurohpc-ju.europa.eu>

²³ www.ecsel.eu

²⁴ www.european-processor-initiative.eu