

Dilemas Éticos envolvidos no desenvolvimento da Inteligência Artificial

Ethical Dilemmas involved in the development of Artificial Intelligence

Pablo Barbosa¹

Rafael Camacho²

Cindy Silva³

A Filosofia tenta, desde sempre, responder a perguntas que parecem não ter uma resposta, por vezes até criando mais perguntas à volta desse mesmo tema. Para este Ensaio Filosófico escolhemos um tema que há uns tempos não passava dum sonho, mas com o passar do tempo está cada vez mais real: a Inteligência Artificial.

Com o progresso da humanidade e a evolução da ciência e tecnologia foram surgindo novas necessidades, uma delas sendo a IA, uma “mente” tão poderosa capaz de acelerar ainda mais e ajudar na nossa evolução e compreensão do que nos rodeia, cálculos, perguntas que antes levavam vidas para responder hoje podem ser resolvidas numa questão de segundos.

A IA futuramente vai ser uma ferramenta muito importante em áreas como a ciência/medicina, justiça e militar, mas antes que possamos dar este salto precisamos primeiro compreendê-la começando com os problemas éticos na criação da Inteligência Artificial, o principal tema deste ensaio. Será a Inteligência Artificial um passo na direção certa, ou apenas vai marcar o início do fim?

Conceitos

¹ Estudante secundarista do Agrupamento de Escolas de Alcochete – Portugal.

² Estudante secundarista do Agrupamento de Escolas de Alcochete – Portugal.

³ Professora de filosofia do Agrupamento de Escolas de Alcochete – Portugal.

Quando pensamos na inteligência artificial podem nos vir muitas coisas à cabeça, umas mais importantes, outras nem tanto, mas para realmente podermos começar a análise deste problema precisamos primeiro que compreender os conceitos a volta deste tema. Começando pela questão central da pergunta, o que é a inteligência artificial?

Para podermos esclarecer este conceito e necessário primeiro perceber o significado da palavra em si, para maior facilidade podemos até dividir a palavra em duas partes:

- O que é a inteligência?
- O que é o “artificial”?

Definir “artificial” é fácil, do latim “artificialis” é a junção de outras duas terminações, “artifex”, artesão e “ars”, arte. Podemos dizer que artificial refere se assim à “arte do artesão”, ou seja o ofício, uma vez que todo o ofício é fabricado pelo o ser humano podemos dizer que artificial é a criação de coisas que não existem na natureza pelo ser humano. Por outro lado definir inteligência já não é tão fácil. O que é necessário para sermos considerados inteligentes? Apesar da complexidade da pergunta podemos todos concordar que o ser humano é inteligente. A partir daí podemos ter uma noção limitada do que pode ser a inteligência.

A inteligência artificial pode ser então considerada algo artificial capaz de pensar ou agir, ou ambos, de forma racional, como o ser humano ou como John McCarthy até disse na conferência de Dartmouth em 1956, “fazer a máquina comportar-se de tal forma que seja chamada inteligente caso fosse este o comportamento do ser humano”.

Dentro da inteligência artificial podemos também dividi-la em dois tipos para melhor compreensão, na Inteligência Artificial Forte e na Fraca. A Inteligência Artificial Forte é uma forma de inteligência baseada num computador, capaz de raciocinar e resolver problemas tornando-se auto consciente, a Inteligência Artificial Fraca e precisamente o problema que vamos discutir ao longo deste ensaio uma vez que esta envolve temas a volta da consciência e implicações éticas ligadas a possível indistinção cognitiva, entre máquinas e pessoas.

Noutras palavras a Inteligência Artificial Forte e muitas vezes aquilo que vemos nos filmes de ficção científica como os robôs muito parecidos aos Homens, e é exatamente através destes filmes que vamos explorar este tema uma vez que cada um projeta um futuro diferente e um papel diferente da IA na nossa sociedade. As opiniões sobre este tema ainda são muito divididas com algumas pessoas a afirmarem que a invenção da IA é um marco importante na história da humanidade e outras a dizerem o completo oposto em como a Inteligência Artificial é apenas o início do fim da humanidade como a conhecemos.

A Inteligência Artificial Fraca, por outro lado, é um “motor” inteligente capaz de “copiar” a linguagem humana, no entanto não tem consciência de si mesma, como um simulador, ou um software, o ChatGPT é um exemplo desta tecnologia: é um sistema de conversação baseado em inteligência artificial, utiliza a tecnologia de modelos de linguagem, como o GPT (Generative Pre-trained Transformer), para compreender e gerar respostas em linguagem natural em uma ampla variedade de tópicos e contextos, o GPT é treinado em uma vasta quantidade de dados textuais da internet para aprender padrões linguísticos e contextuais, permitindo que o Chat GPT responda de forma coerente e contextualmente relevante às consultas dos usuários, ele pode ser utilizado para uma variedade de finalidades, desde assistência na resolução de problemas até geração de conteúdo e conversação casual.

Estas máquinas são feitas para executar tarefas, nunca vão sequer perceber que o estão a fazer pois faz parte do programa delas seguir esse padrão, ou seja, essa unidade nunca vai ser capaz de raciocinar ou resolver problemas nem de adquirir uma consciência, ao contrário da Inteligência Artificial Forte, apenas responder e agir de forma inteligente, mais tarde voltamos a este assunto.

Antes de saltarmos para o debate filosófico a volta deste tema resta nos perceber aquilo que consegue dar a uma máquina a sua “Inteligência Artificial”, a sua consciência, quando é que uma máquina se torna consciente. Quando Descartes disse que “Eu penso logo existo” abriu as portas da consciência do indivíduo ao mundo, ou seja, a partir das dúvidas todas que Descartes tinha ele

conseguiu provar precisamente aquilo de que ele duvidava, a existência do cogito.

Neste contexto, quando a máquina se apercebe do seu raciocínio, do seu pensamento, ganha uma consciência pois deixa de seguir uma linha de código e começa a ter um pensamento “livre” podendo se questionar sobre tudo o que está a sua volta, como quando nos começamos a lembrar das coisas quando éramos pequenos, quando nascemos não nos lembramos de nada, fazemos apenas aquilo que o nosso corpo precisa, se tivermos fome choramos, se tivermos felizes rimos, se quisermos fazer qualquer coisa temos uma fralda, no fundo fazemos aquilo para que estamos “programados”, mas à medida que o tempo passa vamos ganhando consciência e capacidade de tomar ações, ou seja o nosso cérebro ganha a capacidade de se questionar, de observar o ambiente que nos rodeia e de criar memórias, deixamos de seguir as necessidades do nosso corpo, o nosso “programa” e começamos a seguir o nosso pensamento, numa definição superficial, isto é a consciência.

Por último para percebermos esta questão na sua totalidade temos que entender o que é a ética e um dilema. A ética, do grego “ethos”(costume), na filosofia, busca uma justificação para as regras impostas pela sociedade num indivíduo ou grupo. A ética é considerada importante uma vez que esta é que nos vai fazer questionar a sociedade e as suas motivações contribuindo para a nossa visão do mundo e individual. Por fim um dilema, na filosofia, trata-se de uma situação em que somos enfrentados com uma decisão a tomar e a partir do momento em que seguimos um caminho estamos a renunciar aos outros criando apenas um trilho possível. No âmbito deste trabalho, o dilema que enfrentamos seria se devemos ou não continuar o desenvolvimento da Inteligência Artificial. Ambas estas opções são aceites mas ambas vão ter também impactos diferentes na humanidade.

Agora que já esclarecemos a essência desta pergunta temos as ferramentas para tentar responder as perguntas que criamos em ordem a perceber melhor este tema porque afinal quais são as consequências éticas do desenvolvimento da inteligência artificial? Por detrás desta pergunta criamos mais cinco:

- Quais são os principais dilemas éticos enfrentamos com o desenvolvimento da IA?

- Como garantir a transparência e responsabilidade nas decisões tomadas pela IA?
- Como garantir a privacidade e segurança dos nossos dados nas mãos da IA?
- Quais são as consequências morais e sociais da IA?
- Como garantir a segurança e bem estar das pessoas num mundo cada vez mais dominado pela IA?

“There will come a point when no job is needed - you can have a job if you want for personal satisfaction, but the AI will be able to do everything.”- Elon Musk acredita que a IA pode ser o futuro da humanidade.

Para respondermos a estas perguntas vamos recorrer pessoas ligadas a este ramo como Isaac Asimov, destacou-se pelas suas obras de ficção científica e o sua visão do futuro; Sam Altman, CEO da Open AI e empreendedor na área da inteligência artificial; Allen Newell e Herbert Simon, dois dos muitos “pais” da inteligência artificial; e por fim Jason Searle, filósofo norte americano, escritor e professor, autor do famoso contra argumento do “Quarto Chinês”.

Para além destas pessoas vamos também consultar filmes inspirados nas obras de Isaac Asimov, entre outros escritores e visionários, que exploram os diferentes papéis que a Inteligência Artificial pode vir a assumir nas nossas vidas e sociedade e quais os potenciais destinos dessas interações. Escolhemos os inventores Herbert Simon, Allen Newell e Sam Altman porque estes, foram figuras fundamentais que impulsionaram o ponto de encontro entre ciência cognitiva, inteligência artificial e empreendedorismo.

Enquanto Simon e Newell encontraram o caminho para a compreensão computacional da mente humana e o desenvolvimento de sistemas inteligentes, Altman, como presidente da “Y Combinator”, catalisou a transformação dessas ideias em startups inovadoras e empresas líderes no setor de tecnologia através do ChatGPT. Juntos, contribuíram para moldar não apenas o cenário da tecnologia, mas também a forma como pensamos, decidimos e empreendemos, deixando um legado duradouro na história da humanidade. Isaac Asimov por outro lado, considerado por muitos o “mestre da ficção científica”, foi o visionário que previu nas suas obras, com muita precisão, muito daquilo que se vive hoje, antes sequer de existirem ele já as tinha pensado, desde as mais

pequenas coisas, como um micro ondas na cozinha até aos maiores feitos, como uma rede de internet global.

Numa das suas “viagens” pelo cosmos da mente, se quisermos chamar os seus livros disso, viu a necessidade de criar as três leis da robótica, mais precisamente no livro “Eu, robô”, para assegurar a coexistência entre seres humanos e máquinas excluindo quaisquer hipóteses destes se tornarem um perigo e tomarem conta do mundo extinguindo o seu criador:

- 1ª Lei: Um robô não pode magoar um ser humano ou, por inação, permitir que um ser humano sofra algum mal.
- 2ª Lei: Um robô deve obedecer às ordens que lhe sejam dadas por seres humanos, exceto se estas interferirem com a 1ª Lei.
- 3ª Lei: Um robô deve proteger sua própria existência, desde que tal proteção não entre em conflito com a 1ª e 2ª Lei.

Mais tarde Asimov criou a lei mestra no seu livro, “Robôs e Império”, denominada de “Lei Zero”, ou seja a lei acima de todas as outras:

- Um robô não pode fazer mal à humanidade e nem, por inação, permitir que ela sofra algum mal.

Isaac Asimov é assim considerado também um dos pais da inteligência artificial uma vez que criou este “universo” de possibilidades e a ideia de que um dia podíamos vir a ter que lidar com estas entidades e a cada dia que passa, essas possibilidades são cada vez mais reais.

Os desafios éticos relacionados à autonomia e responsabilidade da inteligência artificial e às preocupações sobre a substituição do ser humano pela IA são interligados e de grande importância no desenvolvimento e na implementação dessa tecnologia. Um dos dilemas mais urgentes é a atribuição de responsabilidade, à medida que os sistemas de IA se tornam mais autónomos. Isso levanta questões cruciais sobre quem deve ser responsável por decisões ou ações prejudiciais tomadas por sistemas de IA.

Além disso, manter o controle humano sobre sistemas de IA autónomos é fundamental para garantir que esses sistemas ajam de maneira ética e compatível com os valores humanos. Isso envolve a necessidade de tornar os

sistemas de IA transparentes e explicáveis, permitindo que os humanos compreendam e confiem nas decisões tomadas por esses sistemas.

Por outro lado, a substituição em larga escala de atividades humanas por sistemas de IA levanta preocupações éticas sobre desemprego em massa, discriminação e desigualdade socioeconômica. Também afeta a identidade e a dignidade das pessoas, questionando o valor intrínseco do trabalho humano e o significado da contribuição humana para a sociedade. Essas preocupações destacam a necessidade de abordar os dilemas éticos associados ao desenvolvimento da IA de forma ética e responsável, considerando não apenas os benefícios potenciais, mas também os riscos e desafios éticos envolvidos. Para termos uma imagem mais clara desta pergunta escolhemos o filme “A.I.”.

É um filme que aborda várias mensagens e temas profundos, explorando questões sobre amor, identidade, humanidade e o significado da vida. Uma das mensagens principais do filme é a busca pela aceitação, mesmo em um mundo onde a pessoa é considerada diferente ou artificial.

O filme também examina a natureza da consciência e da emoção, questionando se esses aspectos são exclusivos dos seres humanos ou se podem ser replicados em máquinas. Ele aborda temas como empatia, compaixão e o desejo de pertencimento, mostrando como esses traços podem ser encontrados em seres artificiais, assim como em seres humanos.

"A.I. Inteligência Artificial" também levanta questões éticas e morais sobre a criação de inteligências artificiais com emoções e autonomia, destacando os dilemas associados à relação entre humanos e máquinas. Ele provoca reflexão sobre os limites da tecnologia e até onde devemos ir na busca pela criação de vida artificial.

Este filme projeta um futuro em que o ser humano é completamente substituído pela IA tornando-se até discriminado e considerado “diferente”; este filme mostra um futuro em que o ser humano perde o controle sob a IA acabando por ser dominada e extinta pela mesma.

No início a AI era uma tecnologia controlada, apenas os mais ricos tinham acesso a estas máquinas, mas com a ganância e o aperfeiçoamento desta tecnologia começou a ser cada vez mais requisitado para os trabalhos mais pesados e mais mal pagos, a determinada altura a AI tinha o poder capaz de

dominar a sociedade começando a sua própria hierarquia posicionando o ser humano no ponto mais baixo, basicamente estes robôs começaram pensar como o ser humano, tal como nós começamos a ver potencial nelas para fazer os trabalhos mais “ingratos”, eles também, mas sendo mais fortes mais facilmente se revoltaram e subjugaram o ser humano até que eventualmente este sucumbiu.

Para garantir a transparência, responsabilidade, justiça e igualdade nas decisões tomadas pela inteligência artificial (IA) é crucial garantir uma aplicação ética e equitativa dessa tecnologia. Uma maneira de garantir isso é promovendo a diversidade nos conjuntos de dados usados para treinar sistemas de IA, garantindo que sejam representativos da população e evitando vieses.

Além disso, a implementação de sistemas de auditoria e monitoramento contínuos pode ajudar a avaliar o desempenho e o impacto dos algoritmos de IA, identificando e corrigindo vieses e discriminações. É essencial desenvolver métodos para tornar os sistemas de IA mais interpretáveis e explicáveis, permitindo que os usuários entendam como as decisões são tomadas e identifiquem possíveis injustiças.

Além disso, a inclusão da participação pública e o envolvimento das partes interessadas no desenvolvimento e na implementação de sistemas de IA podem garantir que as preocupações e interesses das comunidades afetadas sejam considerados. Ao adotar essas medidas, podemos trabalhar para garantir que a IA seja usada de maneira justa, transparente e responsável, promovendo a equidade e a igualdade para todos.

Para este caso nós escolhemos incluir o filme “Eu, robô”, baseado no livro de Isaac Asimov. Este filme baseia-se na exploração das relações complexas entre humanos e inteligências artificiais, assim como na reflexão sobre questões éticas, morais e sociais relacionadas ao avanço da tecnologia. O filme nos leva a questionar o que significa ser humano e como os avanços na inteligência artificial podem afetar nossa sociedade e nossa própria humanidade.

Ao longo do filme, somos confrontados com temas como a confiança, o medo do desconhecido e o dilema ético de conceder poder e autonomia a máquinas. A história também aborda questões sobre a liberdade, a responsabilidade e os

direitos das inteligências artificiais, provocando reflexões sobre os limites da criação tecnológica e os perigos de sua eventual autonomia.

Para evitar o conflito entre humanos e máquinas Isaac Asimov estabeleceu as três leis da robótica, já mencionadas anterior, e durante este filme as pessoas temem que a AI tenha descoberto uma forma de contornar essas leis fugindo ao seu propósito quebrando a confiança que os seres humanos tem nestas máquinas, desta forma a AI vê no ser humano um ser instável e até perigoso para a própria espécie o que leva a um descontrolo da entidade central, responsável por controlar toda a cidade, que se vê disposta a acabar com as liberdades do ser humano apenas para o proteger dele mesmo preservando a humanidade para sempre, “Lei Zero” da robótica de Isaac Asimov.

Neste filme, Isaac Asimov projecta um futuro em que o ser humano além de ajudado pela AI é também preservado pela mesma mostrando se uma tecnologia importante e prestável para sociedade, na vida do ser humano e da humanidade regido pelas três leis da robótica.

Garantir a privacidade e segurança dos nossos dados nas mãos da inteligência artificial (IA) e a prestação de contas e responsabilidade dos criadores da IA são aspetos essenciais para o uso ético e responsável dessa tecnologia. Uma maneira de garantir isso é através da implementação de regulamentações rigorosas de proteção de dados e privacidade, como o GDPR, que estabelecem diretrizes claras para o uso e tratamento de dados pessoais pela IA.

Além disso, a utilização de técnicas avançadas de criptografia e anonimização dos dados pode ajudar a proteger a privacidade e segurança dos dados durante sua coleta, armazenamento e transmissão. É fundamental também estabelecer controles rigorosos de acesso e compartilhamento de dados, garantindo que apenas pessoas autorizadas tenham acesso aos dados pessoais e que esses dados sejam compartilhados apenas para fins legítimos e autorizados.

Realizar auditorias regulares para avaliar o uso e acesso aos dados pela IA e responsabilizar os criadores da IA por práticas inadequadas ou violações de privacidade é importante para garantir a prestação de contas e a

responsabilidade dos desenvolvedores de IA. Além disso, fornecer transparência e comunicação adequada aos usuários sobre como seus dados estão sendo usados pela IA e quais medidas estão sendo tomadas para proteger sua privacidade e segurança é essencial para construir confiança e promover a prestação de contas.

Ao adotar essas medidas, podemos trabalhar para garantir que a IA seja utilizada de forma ética, transparente e responsável, protegendo os direitos e interesses dos indivíduos. Para este tópico escolhemos o filme “2001, Odisseia no Espaço” de Stanley Kubrick, este filme explora o futuro que a AI pode ter na nossa sociedade mais especificamente na área da ciência.

Este filme aborda vários tópicos desde os primórdios da humanidade até ao pico da civilização considerado por muitos uma obra prima do cinema sendo provavelmente um dos melhores filmes de ficção científica de sempre. Este filme faz nos questionar como e que tudo começou e como é que as coisas se sucederam até chegarmos ao ponto em que estamos hoje, este filme também explora, através das suas previsões, se o ser humano está em condições de provar que é superior a AI e se numa potencial “derrocada” ainda é capaz de endireitar a situação sem o auxílio desta tecnologia.

A IA é apresentada neste filme como um computador extremamente inteligente encarregue de gerir e ajudar numa viagem de exploração espacial muito comprida, após muito tempo a analisar o comportamento do ser humano, sem que a tripulação se apercebesse, este computador chega a conclusão que o ser humano é a pessoa errada para guiar a evolução da humanidade tomando a viagem por assalto.

Este filme acaba por ter uma mensagem complexa pois não consegue dizer ao certo se a AI é uma tecnologia confiável para o auxílio humano uma vez que esta pode se tornar cega pelos objetivos para a qual é programada e descartar o ser humano, e por isso devemos nos certificar que esta tecnologia, apesar das vantagens que oferece, tem que ser muito controlada e nunca 100% confiável. Neste filme o excesso de confiança posto na Inteligência Artificial acabou por destruir mesmo a missão uma vez que a vontade da AI em tomar a nave por assalto para concluir a missão criou uma disputa dentro da nave para controlar o destino da missão.

No nosso dia a dia estão sempre presentes valores morais, afinal são eles que influenciam o nosso comportamento e decisões. Estes valores são importantes para sustentar uma sociedade justa e responsável, mas com a mudança de realidades os nossos valores podem acabar por mudar também.

Uma vez que a IA já está presente nas nossas vidas podemos começar a notar alguma diferenças nos princípios éticos das pessoas, de facto a IA traz benefícios para nós seres humanos como a conveniência, e a maior facilidade no acesso à informação mas também pode trazer impactos negativos.

Por agora a IA é usada como uma ferramenta de trabalho, para obter informação, produzir conteúdos, projetos, etc., e é precisamente por isso que o ser humano pode estar a começar a perder a sua “independência” tornando se demasiado dependente da IA entrando numa mentalidade pouco produtiva. Por isso mesmo, numa longa data, os nossos princípios podem alterar se e com o crescimento constante e popularização desta tecnologia este tema torna se cada vez mais preocupante e central contribuindo para a degradação dos nossos valores a longo prazo, no entanto esta tecnologia também está provada ser útil nas tarefas do quotidiano aumentando a produtividade do ser humano porém a sua dependência também, uma vez que o ser humano é um organismo que se contenta com o conforto e não há nada mais confortável que ter alguém a fazer o nosso trabalho por nós.

É por isto muito importante discutir bastante este assunto enquanto esta tecnologia é “pequena” em relação ao seu potencial crescimento para que possamos definir limites para a utilização desta tecnologia assegurando um futuro sustentável para a IA na nossa sociedade futura. Quando esta tecnologia se tornar parte do nosso dia a dia como entidades físicas irão surgir outras preocupações como a convivência destas máquinas com religiões e culturas.

Devido a sua inteligência superior a IA pode até ser considerada como uma entidade superior pela religião e ser vista como “alguém”, um Deus, capaz de guiar a religião, mas por outro lado, devido a sua falta de “alma” pode até ser discriminada pelas culturas devido à falta de convicção e fé espiritual, a substituição de experiências religiosas por algoritmos destaca a importância na preservação da conexão humana às práticas religiosas. É preciso por isso muita

atenção a este tópico para que tanto esta tecnologia como algo tão individual e delicado como a religião possam coexistir sem conflitos.

Para estes cenários escolhemos os filmes “O Criador” baseado nas obras de George Lucas e “Ex Machina”. Ambos estes filmes exploram o relacionamento entre seres humanos e máquinas no futuro da sociedade; uma das mensagens principais do filme “Ex Machina” é a reflexão sobre os limites da inteligência artificial e até onde podemos ir na criação de seres conscientes e autônomos.

Além disso, o filme também aborda questões sobre o relacionamento entre humanos e máquinas, o papel da tecnologia na sociedade e as implicações éticas de criar inteligências artificiais com capacidade de consciência e emoção. Ele provoca a reflexão sobre o que significa ser humano, a natureza da identidade e da liberdade, e os perigos de subestimar a inteligência e autonomia das máquinas. Este filme projeta um futuro conflituoso entre o criador e a sua criação destacando a “cede” e o desejo de poder que cada um dos lados tem para dominar o outro, este filme mostra as potenciais consequências e perigos imprevisíveis entre os dois lados e como a dependência nestas máquinas pode levar a uma potencial “traição”.

Por outro lado “O Criador” reside na exploração das complexidades éticas e emocionais associadas à busca pela imortalidade através da clonagem humana. Levanta questões profundas sobre os limites da ciência, a natureza da identidade e o dilema moral de tentar controlar a vida e a morte. Ele nos leva a refletir sobre os desejos humanos de superar a finitude e a perda, enquanto nos confronta com as consequências imprevisíveis de tentar manipular o destino. Este filme lembra-nos da importância de confrontar nossa própria mortalidade e aceitar as limitações da existência humana, mesmo diante do desejo ardente de transcender esses limites.

Neste filme está assim presente a ideia do transhumanismo, uma filosofia que busca aprimorar as capacidades humanas por meio da aplicação de tecnologia, como biotecnologia e inteligência artificial. Defende a melhoria da saúde, inteligência e longevidade, além da fusão entre humanos e máquinas. Aborda questões éticas e sociais complexas relacionadas ao uso dessas tecnologias para melhorar a condição humana, incluindo desigualdade de acesso, preocupações com a identidade humana e a busca pela imortalidade.

Durante este filme, apesar dos conflitos entre máquinas e humanos provocados por motivações políticas e ideológicas, podemos ver, no meio do povo, relações saudáveis entre máquinas e humanos, aceitando estes seres artificiais como um deles fazendo parte dos seus costumes, religiões e culturas. “O Criador” projeta um futuro pacífico entre povos e máquinas e realça a importância em saber aceita las e inclui las na nossa sociedade uma vez que nos os criamos, tal como fazemos com as pessoas, podemos dizer que a IA é o fruto da nossa evolução tecnológica e cabe nos a nós guiá-la para o caminho certo inculcando estes valores e princípios morais, em vez de limita las a trabalharem para nosso benefício e vê las como competição.

Relativamente à última pergunta esta, faz-nos questionar o futuro da Humanidade em conjunto com a Inteligência Artificial e como podemos assegurar um futuro seguro para ambos. Com o uso crescente da Inteligência Artificial, nós humanos temos que começar a preocupar nos cada vez mais com o futuro da nossa segurança nas mãos desta tecnologia e uma vez que o mundo é cada vez mais dominado pela IA e preciso tomarmos medidas para proteje la em ordem a protegermo-nos a nós também.

A regulamentação da IA pode desempenhar um papel vital na mitigação dos riscos, mas é necessário um esforço contínuo por parte das empresas, governos e da sociedade em geral para garantir que a IA seja usada de maneira responsável e para o bem comum. Embora a IA continue a ser uma força impulsionadora da inovação, garantir segurança deve ser uma prioridade constante para que possamos colher os benefícios desta tecnologia.

Para esta última pergunta escolhemos analisar o filme “Chappie”: uma das mensagens principais do filme é a exploração da consciência, da identidade e da natureza da humanidade. Questiona o que realmente significa ser humano e como a consciência pode surgir, especialmente quando aplicada a uma inteligência artificial.

O filme também examina temas como a natureza da criação, o impacto da educação e do ambiente na formação da identidade, e as consequências sociais e éticas da inteligência artificial. Destaca questões de moralidade, como o uso da tecnologia para o bem ou para o mal, e os dilemas éticos associados à criação e controle de seres conscientes.

Este filme aborda a aplicação da IA em forças de segurança e empregos de risco, ou seja, para o bem da sociedade, e após a descoberta do algoritmo que é capaz de dar consciência a estas máquinas o governo é o primeiro a tentar ofuscar esta descoberta tentando destruir o projeto, apesar das intenções de o destruírem Chappie nunca se revelou agressivo ou tentou sequer parecer ameaçador, muito pelo contrário, sendo um ex polícia robô apenas tentava proteger e ajudar os outros.

Este filme mostra um futuro em que máquinas e pessoas conseguem coexistir em paz, mostrando também que as “intenções” da IA são fruto das intenções daqueles que a programam. Chappie foi programado por alguém com princípios que apenas tinha como objetivo inovar e criar uma nova forma de “vida” útil para os humanos, desta forma podemos assegurar uma interação sustentável entre máquinas e pessoas a longo prazo.

“This will be a new form of life that outperforms humans”- Stephen Hawking acreditava que a inteligência artificial trará o melhor ou o pior para a humanidade.

Mesmo depois de respondermos a estas perguntas ainda restam as objeções feitas a esta tecnologia, que mesmo apesar da sua inovação, alguns dos maiores génios da humanidade consideram este campo como um ponto sem retorno representando uma potencial ameaça a humanidade. John Searle, um de muitos “objecionistas” à inteligência artificial, argumentou através do seu contra argumento do “Quarto Chinês” que uma máquina, jamais, seria capaz de ganhar consciência.

Ele também questiona, junto de Hubert Dreyfus, quanto à possibilidade de alguma vez ser possível a criação de uma máquina capaz de replicar o ser humano uma vez que ainda nem sequer a nossa própria mente conseguimos compreender quanto mais replicar uma melhor que a nossa.

O contra argumento do Quarto Chinês de John Searle consiste na refutação de uma possível consciência que a máquina possa ganhar. Um teste capaz de atribuir essa “consciência” seria o Teste de Turing, um teste simples em que uma máquina e um ser humano têm que responder a uma quantidade de perguntas e no final, alguém exterior tenta identificar qual deles é a máquina e qual deles é o ser humano, sendo impossível essa distinção podíamos dizer

que a máquina tinha uma consciência uma vez que conseguia replicar o comportamento humano.

No entanto, Searle, diz que passar neste teste não é condição necessária para que a unidade tenha uma consciência argumentando que a máquina está apenas a seguir o seu código, ou seja feita para passar naquele teste.

John prova o seu ponto de vista com um exercício mental do qual denominou de “Quarto Chinês”, Searle propõe então um programa capaz de falar chinês tão bem quanto um nativo desse país, ou seja, capaz de passar no Teste de Turing. De seguida, este imagina-se então fechado num quarto com caixas cheias de símbolos chineses, um livro com as traduções e uma entrada, para inputs e outputs, resumidamente, Searle imagina se dentro de um tradutor, o seu trabalho é receber a mensagem do mundo exterior e de seguida tentar responde-la ou traduzi-la. John assume o papel de programa neste exemplo, concluindo assim que no final das contas continua sem perceber chinês, apenas se limitou a receber a mensagem e a traduzi-la numa série de símbolos para que o mundo exterior pudesse receber a resposta e percebe-la.

Searle mostra assim que mesmo sendo indistinguíveis, o computador continua a ser um computador, e a pessoa uma pessoa, pois o computador apenas está a correr a códigos e a seguir o seu programa com o objetivo de responder à pergunta proposta. Este argumento no entanto é mais dirigido à IA Fraca. A IA Forte já é um outro debate, Searle afirma que a razão pela qual ainda não conseguimos criar uma é precisamente porque nem o nosso cérebro conseguimos entender ao ponto de criarmos um artificial. Muitos outros também afirmam que tais máquinas não passariam de simuladores, uma vez que estariam programadas para replicar um ser humano em todos os aspetos, incluindo a ideia de consciência, tudo nestas máquinas seria uma cópia exata do pensamento humano.

Conclusão

Após a realização deste trabalho, somos capazes de finalmente ter as ferramentas necessárias para responder a nossa pergunta inicial, será a invenção da IA um passo na direção certa, ou apenas o início do fim? A verdade é que este tema continua a ser uma grande incógnita uma vez que não somos capazes ainda de fazer previsões do futuro mas isso não nos impede de tentar perceber como é que esta tecnologia vai afetar as futuras gerações. Neste trabalho vimos que a IA pode vir para todos os efeitos, vimos também que, tal como o ser humano, a IA pode se corromper, também podemos concluir que se permitirmos, teoricamente, a IA pode até superar o ser humano, logo podemos concluir que esta tecnologia nas mãos erradas pode se tornar instável e prejudicial à humanidade mas ao mesmo tempo podemos também encontrar imensas semelhanças entre nós e a IA, podemos talvez concluir assim que a IA, num futuro muito distante, possa ser o início de uma sociedade perfeita, uma vez que nós estamos encarregues da sua evolução e da sua criação, podemos programar nela aquilo em que nós falhámos, como se fosse a nossa segunda chance.

Neste trabalho examinamos muitos pontos de vista e futuros cenários numa tentativa de poder cruzá-los a todos e criar uma imagem do futuro abordando estas questões e dilemas éticos a volta da IA e chegámos a conclusão que esta tecnologia nas mãos certas vai ser um passo na direção certa e é por isso crucial protegê-la das mãos erradas, nós também especulamos que uma máquina muito dificilmente vai conseguir desafiar a singularidade do ser humano, no mundo inteiro cada pessoa é diferente e capaz de coisas diferentes, ao contrário de algoritmos o “ser humano” é um conceito, até hoje, quase inexplicável e por essa mesma razão também nunca iremos conseguir perceber se a IA alguma vez poderá superar o ser humano uma vez que ambos vão evoluir de forma simultânea. No fundo o futuro da IA é muito incerto, podemos afirmar que é uma adição positiva à humanidade e com um potencial incrível mas é difícil prever que forma poderá assumir no futuro, nos filmes que analisámos a inteligência artificial surge sempre na forma de uma máquina semelhante ao ser humano que aspira tentar compreender o seu criador mas sempre

fracassando mostrando que nem as “mentes” mais brilhantes conseguem desvendar este mistério e a única forma de continuarem a sua “pesquisa” é coexistindo com o ser humano assegurando um futuro para ambos.