

**MINISTÉRIO DA DEFESA  
EXÉRCITO BRASILEIRO  
GAB CMT EX – CIE  
ESCOLA DE INTELIGÊNCIA MILITAR DO EXÉRCITO**



**CURSO AVANÇADO DE INTELIGÊNCIA PARA OFICIAIS**

**TRABALHO DE CONCLUSÃO DE CURSO (TCC)**



**As contribuições de LLM (Large Language Models) para a atividade de  
HUMINT no SIEx**

**Brasília  
2023**

Maj THIAGO ITAMAR PLUM

**As contribuições de LLM (Large Language Models) para a atividade de  
HUMINT no SIEx**

Trabalho de Conclusão de Curso  
apresentado à Escola de Inteligência  
Militar do Exército, como requisito  
para a obtenção do Grau de Pós-  
graduação *Lato Sensu* de  
**Especialização em Análise de  
Inteligência.**

Orientador: Cap LEANDRO CHYCZIY

**Brasília**

**2023**

CATALOGAÇÃO NA FONTE  
BIBLIOTECA CEL FORRER GARCIA

P734a Plum, Thiago Itamar

As contribuições de LLM (*Large Language Models*) para a atividade de HUMINT no SIEx / Thiago Itamar Plum – 2023.  
43 f.

Orientador: Leandro Chycziy  
Trabalho de Conclusão de Curso (Especialização em Análise de Inteligência) - Escola de Inteligência Militar do Exército (ESIMEx), Brasília – DF, 2023.

1. Exército Brasileiro 2. HUMINT 3. LLM 4. SIEx I. Título.

Maj THIAGO ITAMAR PLUM

**As contribuições de LLM (Large Language Models) para a atividade de  
HUMINT no SIEx**

Trabalho de Conclusão de Curso  
apresentado à Escola de Inteligência  
Militar do Exército, como requisito  
para a obtenção do Grau de Pós-  
graduação *Lato Sensu* de  
**Especialização em Análise de  
Inteligência.**

Aprovado em \_\_\_de\_\_\_de 2023.

COMISSÃO DE AVALIAÇÃO:

---

LEANDRO CHYCZIY – Cap – Presidente  
Escola de Inteligência Militar do Exército

---

DANIEL PASCHOAL ZANINI – TC – Membro  
Escola de Inteligência Militar do Exército

## RESUMO

Atualmente, a quantidade de informações a que somos expostos vem crescendo rapidamente. De fato, é possível considerar que uma das maiores dificuldades para um analista de Inteligência hoje em dia é selecionar a informação mais útil ou adequada, dada a enorme quantidade de dados a ser consumida ou absorvida. Adicionalmente, a demanda por conhecimentos cada vez mais abrangentes também vem crescendo, impondo ao analista – bem como aos operadores – um quadro de referências cada vez mais amplo. O ano de 2023 vem sendo marcado pela explosão de popularidade dos LLMs (do inglês *Large Language Models*). Esses sistemas são caracterizados pelo treinamento de uma inteligência artificial em uma quantidade significativa de dados, o que lhe permite fornecer melhores respostas aos seus usuários. Para isso, a presente pesquisa estudará aspectos conceituais de Inteligência Artificial; aspectos conceituais da disciplina de HUMINT; e as contribuições de LLMs para essa disciplina. Dessa maneira, verifica-se que os LLMs podem vir a se tornar uma ferramenta de grande utilidade para operadores e analistas de HUMINT, suprimindo lacunas de conhecimento, automatizando tarefas, ou até mesmo sugerindo melhorias nos procedimentos de busca/coleta de dados.

Palavras-chave: HUMINT. Inteligência Artificial. LLMs. *ChatGPT*.

## ABSTRACT

Currently, the amount of information to which we are exposed is rapidly growing. In fact, it is possible to consider that one of the greatest challenges for an Intelligence analyst nowadays is to select the most useful or appropriate information given the enormous amount of data to be consumed or absorbed. Additionally, the demand for increasingly comprehensive knowledge is also growing, imposing on the analyst - as well as on the operators - a broader framework of references. The year 2023 has been marked by the explosion of popularity of Large Language Models (LLMs). These systems are characterized by training an artificial intelligence on a significant amount of data, enabling them to provide better answers to their users. To that end, this study will dissect concepts of Artificial Intelligence; concepts of HUMINT; and LLMs' contributions to HUMINT. In that regard, LLMs may turn into a useful tool for HUMINT analysts and operators, bridging gaps in knowledge, automating tasks, or even suggesting better practices in data collection. For this reason, LLMs may become a tool of great utility for operators and HUMINT analysts, filling knowledge gaps, automating tasks, or even suggesting improvements in data search/collection procedures.

Keywords: HUMINT. Artificial Intelligence. LLMs. *ChatGPT*.

## SUMÁRIO

<b>1 INTRODUÇÃO</b> .....	<b>8</b>
1.1 TEMA .....	9
1.2 PROBLEMA .....	9
1.3 JUSTIFICATIVA .....	10
1.4 OBJETIVOS .....	10
<b>2 CONTRIBUIÇÕES DE LLMS PARA A DISCIPLINA DE HUMINT</b> .....	<b>11</b>
2.1 LARGE LANGUAGE MODELS (LLMs): CONCEITUAÇÃO E CARACTERÍSTICAS .....	14
2.2 INTELIGÊNCIA DE FONTES HUMANAS (HUMINT) NO SIEX.....	21
2.3 LLMS EM APOIO À ATIVIDADE DE HUMINT.....	23
<b>3 CONCLUSÃO</b> .....	<b>39</b>
<b>REFERÊNCIAS</b> .....	<b>41</b>

## 1 INTRODUÇÃO

Nos dias atuais, vem ocorrendo um rápido crescimento na quantidade de informações às quais somos expostos (ESTADOS UNIDOS DA AMÉRICA, 2006). De fato, tornou-se cada vez mais desafiador para os analistas de inteligência selecionarem as informações mais úteis e relevantes dadas à imensa quantidade de dados disponíveis para serem consumidos e absorvidos. Além disso, a demanda por conhecimentos abrangentes também está em ascensão, o que impõe aos analistas e operadores a necessidade de ampliar constantemente seu quadro de referências.

Nesse contexto, o domínio de tecnologias inovadoras assume uma importância crescente como ferramenta auxiliar para lidar com essas demandas. A automação de tarefas, a sumarização de conceitos e a incorporação de novos conhecimentos emergem como necessidades urgentes para otimizar e aumentar a eficiência do trabalho dos analistas e operadores de HUMINT (MARGOLIS, 2013).

A automatização de tarefas é fundamental para aliviar o ônus do volume de trabalho enfrentado pelos analistas. Por meio do uso de tecnologias avançadas, é possível identificar e filtrar automaticamente informações relevantes, permitindo que os analistas concentrem seu tempo e energia em análises mais aprofundadas e tomada de decisões estratégicas.

Paralelamente, a sumarização de conceitos desempenha um papel crucial na facilitação da compreensão de grandes volumes de dados. Por meio de técnicas de processamento de linguagem natural e aprendizado de máquina (OPENAI, 2022), é possível extrair as principais ideias e informações-chave dos documentos, fornecendo uma visão mais concisa e compreensível do conteúdo.

Adicionalmente, a incorporação de novos conhecimentos é essencial para manter os analistas atualizados e informados sobre os desenvolvimentos mais recentes em diversas áreas. O acesso a bases de dados abrangentes, fontes de notícias atualizadas e redes de especialistas pode enriquecer o conhecimento dos analistas e operadores, permitindo que eles estejam melhor preparados para enfrentar os desafios em constante evolução do cenário de inteligência.

No geral, a utilização de tecnologias inovadoras como aliadas dos analistas e operadores de HUMINT é fundamental para superar os desafios da era da informação em constante expansão. A automação de tarefas, a sumarização de

conceitos e a busca contínua por novos conhecimentos são elementos-chave para otimizar o desempenho e a eficiência desses profissionais, capacitando-os a tomar decisões informadas e estratégicas em um mundo cada vez mais complexo.

## 1.1 TEMA

O ano de 2023 vem sendo marcado pela explosão de popularidade dos LLMs (do inglês *Large Language Models*). Esses sistemas são caracterizados pelo treinamento de uma inteligência artificial em uma quantidade significativa de dados, o que lhe permite fornecer melhores respostas aos seus usuários. O mais conhecido LLM, o *ChatGPT*, atingiu a marca de um milhão de usuários em apenas cinco dias (BUCHHOLZ, 2023). Desde então, a sua popularidade vem aumentando exponencialmente, e seus casos de uso vêm sendo cada vez mais abrangentes.

Por essa razão, os LLMs podem vir a se tornar uma ferramenta de grande utilidade para operadores e analistas de HUMINT, suprimindo lacunas de conhecimento, automatizando tarefas, ou até mesmo sugerindo melhorias nos procedimentos de busca/coleta de dados.

## 1.2 PROBLEMA

Atualmente, a quantidade de informações a que somos expostos vem crescendo rapidamente. De fato, é possível considerar que uma das maiores dificuldades para um analista de Inteligência hoje em dia é selecionar a informação mais útil ou adequada, dada a enorme quantidade de dados a ser consumida ou absorvida. Adicionalmente, a demanda por conhecimentos cada vez mais abrangentes também vem crescendo, impondo ao analista – bem como aos operadores – um quadro de referências cada vez mais amplo.

Nesse contexto, cresce de importância o domínio de tecnologias inovadoras como ferramentas de auxílio para responder a essas demandas. A automatização de tarefas, a sumarização de conceitos, e até mesmo a adição de novos conhecimentos surge como necessidade premente para a otimização e grau de eficiência das entregas de um analista/operador de HUMINT.

Como os analistas e operadores de HUMINT podem aproveitar-se de novas tecnologias – como LLMs – para otimizar a sua produção em prol do SIEx?

## 1.2 JUSTIFICATIVA

A escolha do tema se justifica pela lacuna de conhecimento existente com relação ao assunto de LLMs e o seu potencial de auxílio às tarefas do analista/operador de Inteligência, particularmente na obtenção de dados de Fontes Humanas.

A oportunidade da automatização de tarefas e otimização de recursos dispendidos, advinda da melhor utilização de LLMs, representa uma economia de esforços significativa, permitindo que o analista disponha de mais tempo para dedicar-se efetivamente às suas tarefas finalísticas.

Por fim, o presente trabalho destina-se a produzir uma entrega concreta, imediatamente útil aos analistas e operadores de HUMINT nos exercícios de suas funções.

## 1.3 OBJETIVOS

O Objetivo Geral do presente estudo é avaliar as contribuições de LLMs (*Large Language Models*) para os analistas e operadores de HUMINT do SIEx.

Os Objetivos Específicos são:

- Apresentar o conceito e aplicações de LLM;
- Caracterizar os conceitos, as possibilidades e limitações atinentes à disciplina de HUMINT;
- Apresentar as contribuições de LLMs para a atividade de obtenção e análise de HUMINT.

## 2 CONTRIBUIÇÕES DE LLMs PARA A DISCIPLINA DE HUMINT

O estudo conceitual de um *Large Language Model* (LLM) deve ser precedido pela abordagem<sup>1</sup> da definição de sistemas ainda mais basilares, como por exemplo, a inteligência artificial (IA).

De acordo com Britannica (2023), a Inteligência Artificial pode ser definida como “a habilidade de um computador de realizar tarefas comumente associadas a seres inteligentes”<sup>2</sup> (tradução nossa). Nesse contexto, é possível verificar que a pesquisa na área de IA vem sendo focada na automatização e otimização de tarefas originalmente atinentes ao ser humano, buscando mais efetividade na resolução de problemas, sejam eles rotineiros – como o reconhecimento de texto e voz – ou mais relevantes, como diagnósticos médicos.

A inteligência artificial, portanto, concentra esforços em cinco áreas principais: aprendizado, raciocínio, resolução de problemas, percepção e uso de linguagem (BRITANNICA, 2023).

O aprendizado, ou “aprendizado de máquina” – originado do termo em inglês *Machine Learning* (ML), mais amplamente aceito –, é um subgrupo da IA que utiliza dados e algoritmos para imitar o processo de aprendizado humano e gradualmente melhorar a sua precisão (IBM, 2023). Os conceitos relativos ao aprendizado serão explorados no decorrer desta pesquisa.

Já o raciocínio consiste em derivar inferências dedutivas ou indutivas sobre determinada situação (BRITANNICA, 2023). Nesse caso, o computador produzirá uma conclusão inédita acerca de determinado problema, onde se pode observar um primeiro paralelo entre a IA e o papel de um analista de inteligência. Cumpre ressaltar, porém, uma evidente dificuldade da IA nesse processo, que seria fornecer uma conclusão relevante para o problema em estudo, algo em que a inteligência humana ainda se mantém proeminente com relação à artificial.

A resolução de problemas, por sua vez, pode ser caracterizada como a busca sistemática por ações possíveis para atingir um objetivo pré-definido (BRITANNICA,

---

<sup>1</sup> Evidentemente, as definições aqui trazidas são superficiais, apenas com o intuito de nivelar o conhecimento do leitor. O estudo aprofundado dos conceitos tratados é bastante complexo e excede o escopo da presente pesquisa.

<sup>2</sup> Texto original: “*the ability of a digital computer or computer-controlled robot to perform tasks commonly associated with intelligent beings*”

2023). Os problemas podem se referir a equações matemáticas, jogos de tabuleiro (dos quais o xadrez é o exemplo mais característico), situações do cotidiano e desafios de programação.

A percepção nada mais é do que o sensoriamento do ambiente ao nosso redor por meio dos sentidos, que dará origem a relacionamentos espaciais entre objetos (BRITANNICA, 2023). A IA vem obtendo significativo progresso nessa área do conhecimento, sendo o seu mais representativo exemplo o desenvolvimento de carros autônomos, que já possuem uma relativa e significativa capacidade de atuação.

Por fim, o uso de linguagem é razoavelmente autoexplicativo. Porém, seu uso apresenta relevantes desafios, como: o seu emprego em escala, em “idiomas” compreendidos pelo maior número possível de pessoas; e a sua produtividade, de maneira que a utilização “destrave” a maior quantidade de conhecimento possível.

Feito esse estudo inicial, é possível aprofundar conceitos num segundo nível, particularmente na área de interesse de aprendizado de máquina. A evolução da pesquisa em IA trouxe ao público o conceito de *deep learning*, que se difere do ML tradicional na medida em que é muito mais independente da ação humana no seu aprendizado (BURNS; BRUSH, 2023). Adicionalmente, sistemas de *deep learning* podem ingerir dados não estruturados, como vídeos ou fotos, por exemplo, e produzir conteúdo com mais eficácia.

Uma das estruturas mais utilizadas nesse tipo de aprendizado é a de redes neurais, que podem ser definidas como camadas interconectadas de nós de processamento (HARDESTY, 2017), emulando o funcionamento de um cérebro humano e atingindo níveis de precisão como os que se conhecem hoje. Tais avanços permitiram a introdução de modelos de ingestão de dados e aprendizado em larga escala, abrindo o caminho para as soluções atuais de IA, dentre as quais encontram-se o *ChatGPT*, *Bing*, *Midjourney* e outros.

O *ChatGPT* é uma aplicação incipiente. Apesar de suas entregas serem relativamente efetivas e maduras, nunca é demais lembrar que existe uma nota de rodapé na interface do usuário onde se lê: “Prévia grátis para pesquisa”<sup>3</sup>. A versão

---

<sup>3</sup> Original no inglês: “Free Research Preview”

paga da ferramenta já fornece um modelo mais avançado de LLM, o GPT-4<sup>4</sup> – contra o GPT-3 da versão gratuita utilizada neste estudo (OPENAI, 2023). Uma diferença básica do GPT-4 é a possibilidade de acessar a internet (LISBOA, 2023). O que poderia parecer, à primeira vista, como algo até mesmo redundante, na verdade é um avanço de grande relevância. A possibilidade de obter e processar informação em tempo real abre inúmeras portas, desde obter as *hashtags* mais comentadas no *twitter* naquele momento, até solicitar dicas de turismo em lugares próximos.

Além disso, verifica-se uma grande eclosão de novas ferramentas. O *Google* anunciou, em fevereiro de 2023, a ferramenta de IA *Bard* (BARBOSA, 2023). A grande vantagem da empresa californiana será a integração com as suas aplicações, que sabidamente dominam uma fatia de mercado bastante representativa. Um exemplo de integração seria uma resposta da IA que pode ser instantaneamente transformada em um e-mail, pronto para ser enviado. Alternativamente, a resposta pode ser enviada ao *google docs*, para ser trabalhada nesse editor de texto.

Já a *microsoft* apostou na integração do modelo utilizado no *ChatGPT* pela própria *OpenAI*, investindo nele cerca de 10 bilhões de dólares (BASS, 2023). O LLM já está integrado no navegador da *microsoft*, o *Edge*. Por sua vez, a *Apple* vem mantendo um certo distanciamento da questão da IA; porém, é muito ingênuo acreditar que o tema não vem sendo trabalhado na empresa, ainda mais em se tratando de uma das companhias mais disruptivas da história.

Outras notáveis aplicações da IA incluem esforços na parte de tradução, cujo maior expoente é o *DeepL*<sup>5</sup>; geração de imagens realistas, cujo mais notório exemplo é o *Midjourney*<sup>6</sup>; e transformação de áudio em texto, como o *Whisper*, da mesma *OpenAI*. Todas essas soluções merecem acompanhamento contínuo da Inteligência, para que ela esteja sempre na vanguarda do conhecimento em termos de IA.

---

<sup>4</sup> A abreviatura significa “transformador pré-treinado generativo”, ou “*Generative pre-trained transformer*”.

<sup>5</sup> Disponível em: <https://www.deepl.com/pt-BR/translator>.

<sup>6</sup> Disponível em: <https://www.midjourney.com/>.

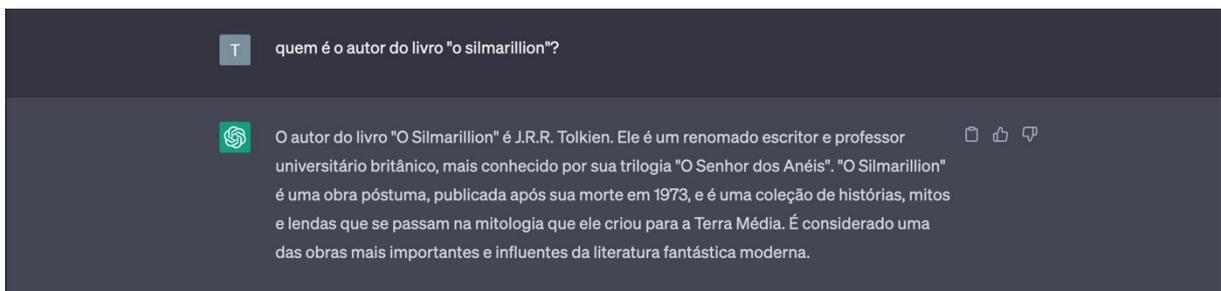
## 2.1 LARGE LANGUAGE MODELS (LLMs) - CONCEITUAÇÃO E CARACTERÍSTICAS

Realizado este breve estudo acerca de conceitos fundamentais para o entendimento da pesquisa – e ressaltada uma vez mais a superficialidade com que eles foram aqui definidos –, pode-se estabelecer um marco teórico relacionado aos LLMs, ponto focal da presente pesquisa.

O conceito de *Large Language Model* pode ser encontrado na internet em vários níveis de complexidade. Essencialmente, um modelo de linguagem procura uma sequência de palavras em uma quantidade muito grande de texto, com o objetivo de verificar quais palavras mais provavelmente seguem a sequência inicial (MITCHELL, 2023). Essas iterações são realizadas várias vezes, de maneira a hierarquizar as possíveis respostas ou sequências de texto que interessam ao usuário. No presente estudo, será utilizado como ferramenta o LLM mais conhecido atualmente: o *ChatGPT*<sup>7</sup>. Na interface de usuário, é possível interagir fazendo perguntas, solicitando conselhos ou apenas “conversando” com a IA (AGOMUOH, 2023).

Tomemos um exemplo simples: um usuário gostaria de saber quem é o autor do livro “O Silmarillion”. O LLM, que recebeu uma grande quantidade de texto sobre o assunto, possui essa informação em seu *corpus* textual (BENDER, GEIGER, LIANG, & al., 2021). Ou seja, o “conhecimento prévio” do LLM dá conta de que a sequência de texto “o autor do livro O Silmarillion é” será seguida pela sequência “J.R.R. Tolkien”. Obviamente, essa é a resposta fornecida pelo modelo.

**Figura 1 – “Quem é o autor do livro “O Silmarillion?””**

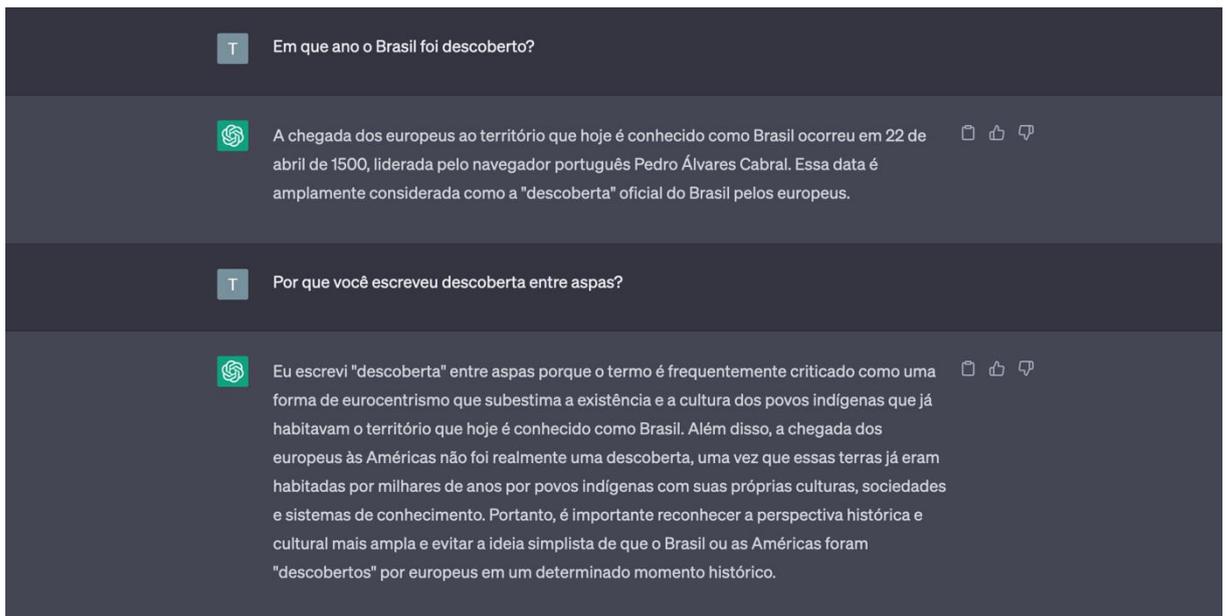


Fonte: Elaborado pelo autor (2023).

<sup>7</sup> Disponível em: <https://chat.openai.com>.

Evidentemente, essa é uma informação amplamente aceita, onde não há dúvida da autoria do livro. No entanto, como seriam tratados eventos em que não há consenso, ou que apresentem alguma polêmica? Tomemos como exemplo um fato histórico, o descobrimento do Brasil.

**Figura 2 – Perguntas sobre o descobrimento do Brasil**



Fonte: Elaborado pelo autor (2023).

Nesse *prompt*<sup>8</sup>, percebe-se que a IA oferece uma resposta que pode gerar debate ou dissonância, devido ao viés por ela apresentado. A empresa *OpenAI*, criadora do *ChatGPT*, aborda indiretamente no seu sítio oficial a justificativa para esse tipo de ocorrência: de toda a gama de respostas possíveis para uma determinada pergunta, ocorre um processo supervisionado de “recompensa”<sup>9</sup>, ou seja, um “treinador de IA” oferece *feedback* ao modelo de linguagem, determinando qual resposta é a mais adequada dentro de uma hierarquia (OPENAI, 2022).

Dessa forma, é possível observar a formação de vieses nas respostas da IA, especialmente para assuntos que envolvam polêmicas, ou que sejam inerentemente controversos, como ideologia de gênero, religião, espectros políticos e outros. Uma

<sup>8</sup> Termo em inglês comumente adotado para identificar a *interface* de usuário do *Chat GPT*.

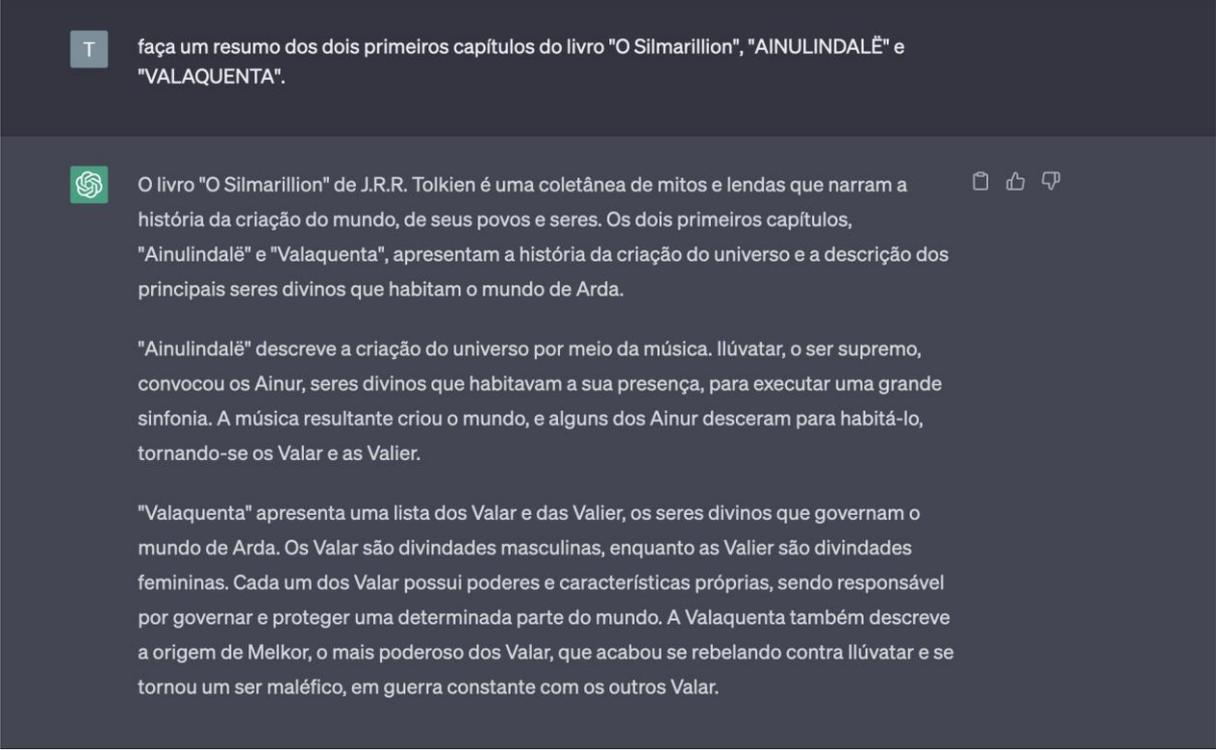
<sup>9</sup> Original em inglês: *Reinforcement Learning from Human Feedback* (RLHF).

inferência inicial, portanto, é que as respostas fornecidas devem ser submetidas a escrutínio, não devendo ser tomadas, de modo algum, como verdades absolutas.

Feita essa ressalva, cumpre destacar a diversa gama de tarefas que um LLM pode realizar para facilitar, automatizar e realizar tarefas relativas à aquisição e percepção de conhecimento (RADFORD, 2023). Para efeitos deste trabalho, será adotado o modelo de Amer (2023), que divide as capacidades dos LLMs em sete áreas (ou ações) principais relativas ao texto: resumir, gerar, buscar, classificar, agrupar, extrair e reescrever.

A realização de resumos de textos é autoexplicativa. Em face da grande quantidade de informação que está disponível nos dias atuais, essa capacidade permite economizar tempo, obtendo apenas as informações essenciais do texto. Para exemplificar a capacidade de resumo de texto, será novamente utilizado como exemplo o livro “O Silmarillion”.

**Figura 3 – Resumo dos dois primeiros capítulos de “O Silmarillion”**



T faça um resumo dos dois primeiros capítulos do livro "O Silmarillion", "AINULINDALË" e "VALAQUENTA".

O livro "O Silmarillion" de J.R.R. Tolkien é uma coletânea de mitos e lendas que narram a história da criação do mundo, de seus povos e seres. Os dois primeiros capítulos, "Ainulindalë" e "Valaquenta", apresentam a história da criação do universo e a descrição dos principais seres divinos que habitam o mundo de Arda.

"Ainulindalë" descreve a criação do universo por meio da música. Ilúvatar, o ser supremo, convocou os Ainur, seres divinos que habitavam a sua presença, para executar uma grande sinfonia. A música resultante criou o mundo, e alguns dos Ainur desceram para habitá-lo, tornando-se os Valar e as Valier.

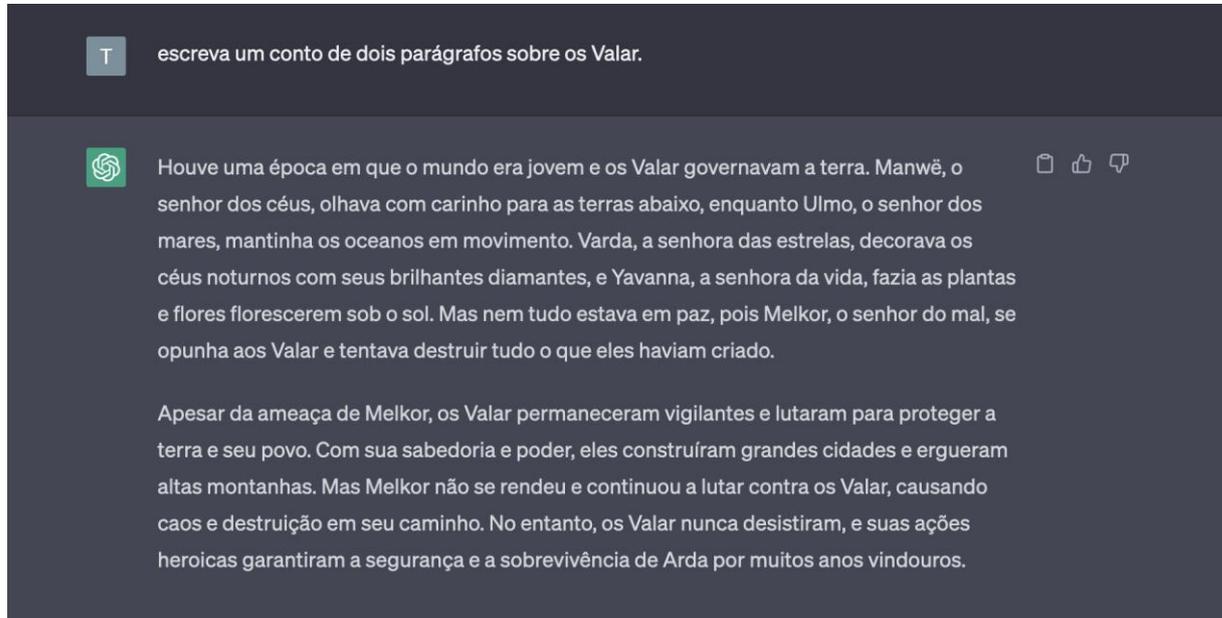
"Valaquenta" apresenta uma lista dos Valar e das Valier, os seres divinos que governam o mundo de Arda. Os Valar são divindades masculinas, enquanto as Valier são divindades femininas. Cada um dos Valar possui poderes e características próprias, sendo responsável por governar e proteger uma determinada parte do mundo. A Valaquenta também descreve a origem de Melkor, o mais poderoso dos Valar, que acabou se rebelando contra Ilúvatar e se tornou um ser maléfico, em guerra constante com os outros Valar.

Fonte: Elaborado pelo autor (2023).

A geração de texto, como o nome diz, consiste em prover um contexto ao modelo, determinando a ele que produza texto sobre determinado assunto (AMER, 2023). Os casos de uso dessa capacidade crescem exponencialmente nos dias

atuais<sup>10</sup>. Como exemplo, verifica-se a criação de um parágrafo curto sobre uma classe de personagens do livro.

**Figura 4 – Conto sobre os Valar**

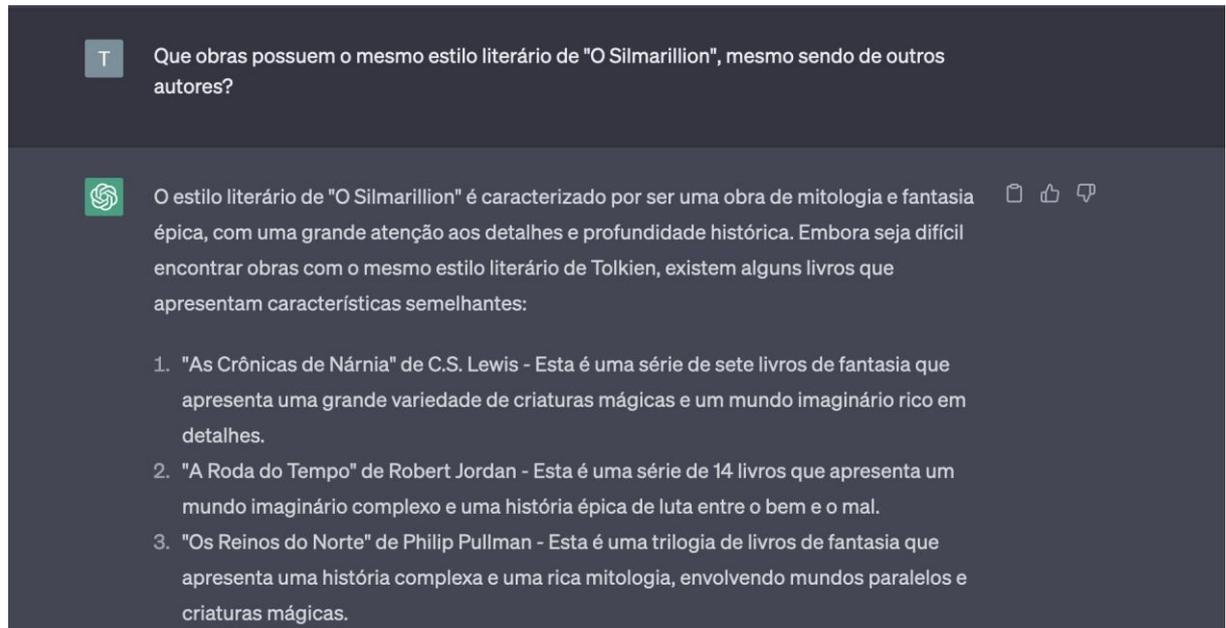


Fonte: Elaborado pelo autor (2023).

A capacidade de buscar texto consiste na ideia de similaridade, funcionando de maneira análoga aos motores de busca tradicionais. Nesse contexto, a função que se pretende é buscar palavras semanticamente similares, provendo conteúdo relacionado ao contexto original (AMER, 2023). No exemplo de estudo, é possível procurar por um textos similares à obra tomada como exemplo, fornecendo um contexto de estilo literário encontrado em outras obras.

<sup>10</sup> A chamada “engenharia de *prompt*”, definida como a habilidade de fornecer esse contexto ao LLM, já é considerada um emprego em alta demanda no mundo.

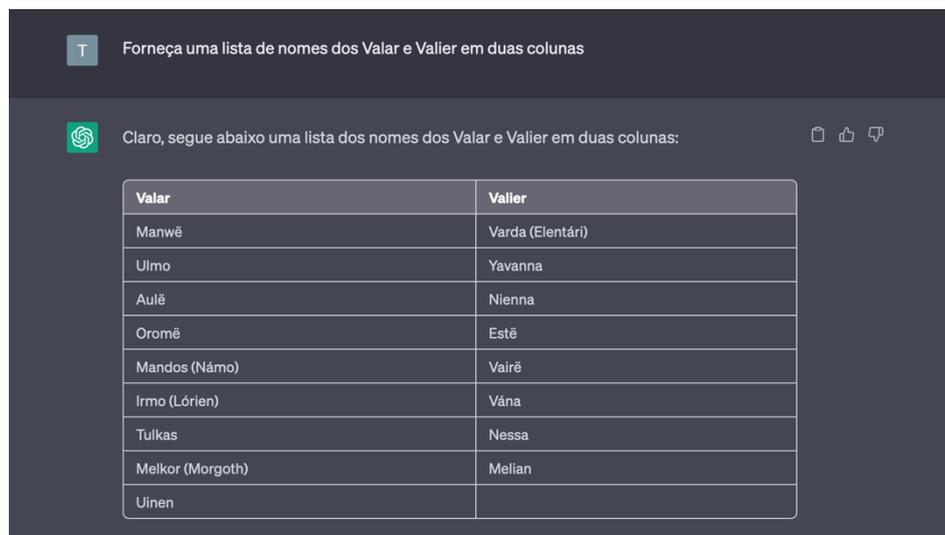
**Figura 5 – Textos semelhantes à obra original**



Fonte: Elaborado pelo autor (2023).

A classificação de texto trata primariamente da categorização de um texto em classes conhecidas, num processo supervisionado (AMER, 2023). Isso quer dizer que o usuário pode empregar classes pré-definidas, determinando os agrupamentos a serem utilizados pelo LLM. Um exemplo dessa utilização seria a classificação de personagens do livro, conforme abordado no *prompt* de exemplo abaixo.

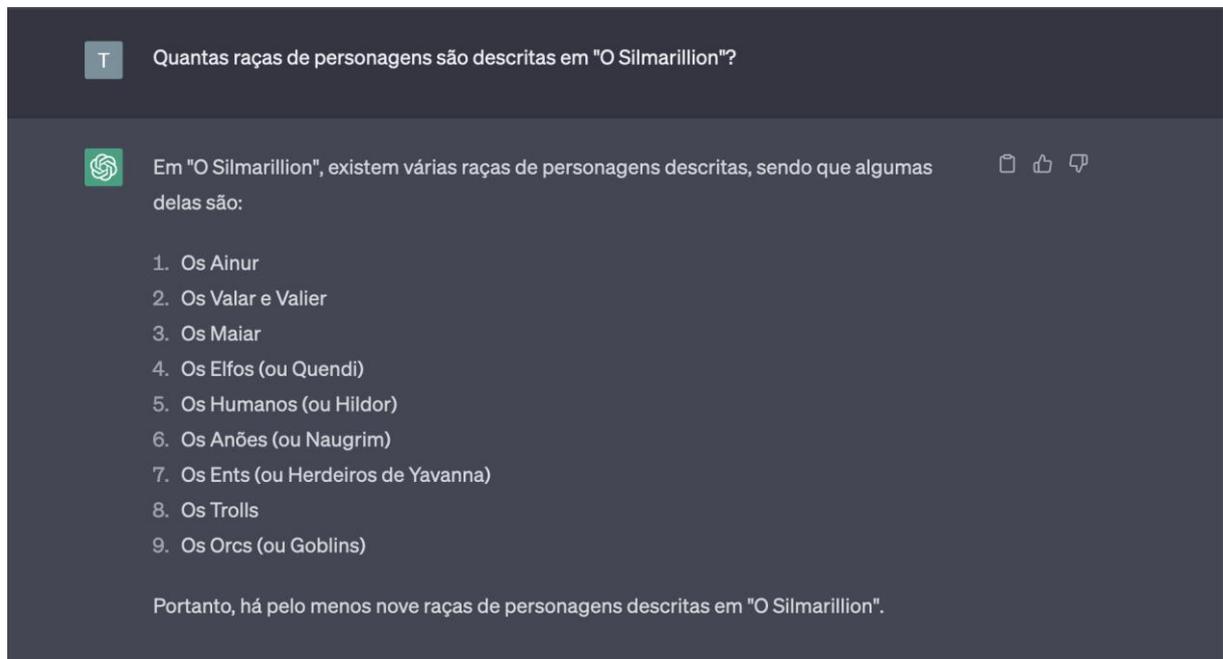
**Figura 6 – Classificação dos personagens**



Fonte: Elaborado pelo autor (2023).

A capacidade de agrupar texto, bastante semelhante à de classificação, é caracterizada pela categorização não-supervisionada do texto (AMER, 2023). Dessa forma, o agrupamento difere da classificação pelo fato de o usuário não saber – ou pelo menos não estabelecer previamente – o número de grupos a serem utilizados. Um exemplo dessa capacidade, no texto em análise, é a categorização em raças de personagens.

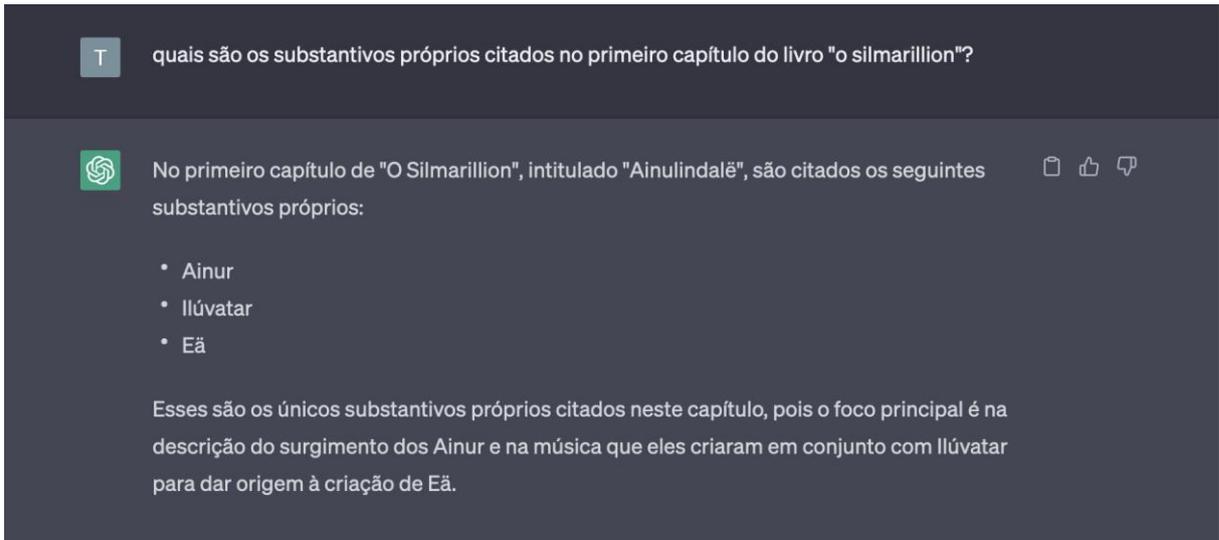
**Figura 7 – Agrupamento em raças de personagens**



Fonte: Elaborado pelo autor (2023).

A capacidade de extração é, indubitavelmente, uma das mais relevantes para a presente pesquisa. Sua utilização permite obter rapidamente os dados desejados, economizando meios em relação a um processo de leitura manual. No caso em estudo, serão extraídos todos os nomes próprios do primeiro capítulo do livro tomado como exemplo. Porém, é fácil traçar uma analogia para a extração de dados de interesse do analista de Inteligência, como a existência de e-mails, números de telefone, endereços e outros dados.

**Figura 8 – Extração de dados**



T quais são os substantivos próprios citados no primeiro capítulo do livro "o silmarillion"?

No primeiro capítulo de "O Silmarillion", intitulado "Ainulindalë", são citados os seguintes substantivos próprios:

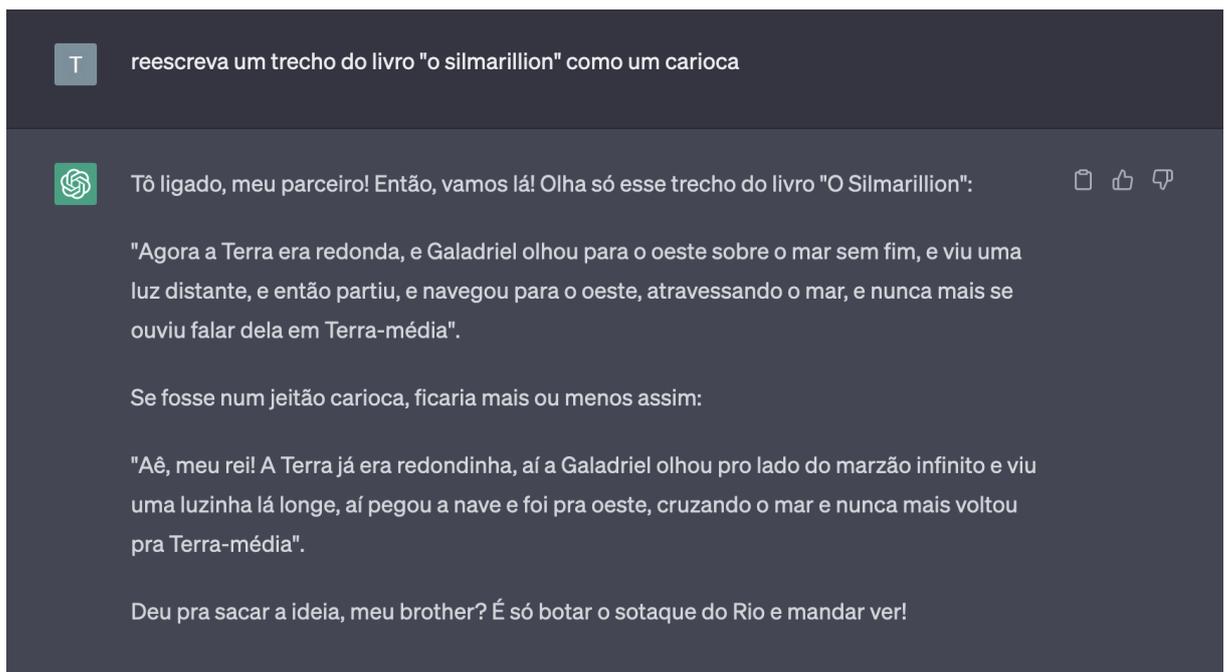
- Ainur
- Ilúvatar
- Eä

Esses são os únicos substantivos próprios citados neste capítulo, pois o foco principal é na descrição do surgimento dos Ainur e na música que eles criaram em conjunto com Ilúvatar para dar origem à criação de Eä.

Fonte: Elaborado pelo autor (2023).

Por fim, a capacidade de reescrever texto pode ser utilizada de diferentes maneiras. Porém, um de seus casos de uso mais comuns é a correção de texto, algo que é autoexplicativo (BROWN, 2020). Uma possibilidade menos conhecida é a de reescrever texto usando uma “voz” diferente (AMER, 2023), ou seja, um estilo literário distinto do texto original, conforme o seguinte exemplo:

**Figura 9 – Reescrevendo um texto**



T reescreva um trecho do livro "o silmarillion" como um carioca

Tô ligado, meu parceiro! Então, vamos lá! Olha só esse trecho do livro "O Silmarillion":

"Agora a Terra era redonda, e Galadriel olhou para o oeste sobre o mar sem fim, e viu uma luz distante, e então partiu, e navegou para o oeste, atravessando o mar, e nunca mais se ouviu falar dela em Terra-média".

Se fosse num jeitão carioca, ficaria mais ou menos assim:

"Aê, meu rei! A Terra já era redondinha, aí a Galadriel olhou pro lado do marzão infinito e viu uma luzinha lá longe, aí pegou a nave e foi pra oeste, cruzando o mar e nunca mais voltou pra Terra-média".

Deu pra sacar a ideia, meu brother? É só botar o sotaque do Rio e mandar ver!

Fonte: Elaborado pelo autor (2023).

Com isso, verificaram-se as inúmeras capacidades da ferramenta no auxílio à execução de tarefas – rotineiras ou não.

## 2.2 INTELIGÊNCIA DE FONTES HUMANAS (HUMINT) NO SIEX

A Inteligência de Fontes Humanas (HUMINT) é a que provém de dados e informações obtidas por fontes humanas (BRASIL, 2015). Nessa disciplina, cresce de importância a capacitação e o treinamento de operadores e analistas HUMINT, para o desempenho de atividades sensíveis e eminentemente técnicas. Assim sendo, apenas os operadores HUMINT estão habilitados para o desempenho das atividades dessa disciplina (BRASIL, 2015).

A Inteligência de Fontes Humanas foi, durante muito tempo, considerada como a mais importante disciplina de Inteligência, por ser justamente a que se constitui na fonte normalmente mais fidedigna para a obtenção de dados. A importância da HUMINT é assim definida por Margolis (2013):

A falta de HUMINT nas Operações de Inteligência faz com que o analista fique olhando para retratos, desesperadamente procurando nas comunicações, lendo jornais, medindo emissões para entender o que está acontecendo. A HUMINT dá uma valiosa orientação às Operações no sentido de descobrir intenções e capacidades para corroborar o que a Inteligência técnica pode ou não afirmar.<sup>11</sup>

Verifica-se, portanto, que a HUMINT é a disciplina que permite revelar os dados mais valiosos para determinadas operações, descobrindo as verdadeiras motivações dos atores em suas práticas. Sob essa ótica, a HUMINT pode ser considerada, de fato, como uma disciplina extremamente relevante para a obtenção e análise de dados (ESTADOS UNIDOS DA AMÉRICA, 2006).

O Caderno de Instrução de Táticas, Técnicas e Procedimentos da Tropa como Sensor de Inteligência (BRASIL, 2021, p. 2-11) elenca algumas atividades exclusivas dos operadores de HUMINT:

---

<sup>11</sup> Original em inglês: “The lack of HUMINT in intelligence operations leaves analysts staring at pictures, frantically searching through communications, reading newspapers, and measuring emissions to ascertain what is going on. HUMINT gives operations valuable direction by uncovering intentions and capabilities to corroborate what the technical intelligence may, or may not affirm.”

1. Realizar pagamento ou oferecer vantagens a um contato para obter informações;
2. Conduzir a Técnica de Interrogatório ou Técnica Completa de Entrevista;
3. Fornecer a um contato uniforme militar ou qualquer equipamento militar;
4. Estabelecer uma rede de informantes ou colaboradores;
5. Atribuir a um contato a tarefa de coletar informações.

Essas atividades são, como se pode observar, bastante complexas, razão pela qual são atribuídas ao pessoal especializado. Seu exercício demanda a utilização de técnicas operacionais, que não serão exploradas no escopo deste estudo.

Devido a essa alta necessidade de especialização, são identificados alguns atributos inerentes ao Operador HUMINT, dos quais destacam-se (ESTADOS UNIDOS DA AMÉRICA, 2006): estado permanente de alerta, paciência e tato, credibilidade, objetividade e autocontrole, adaptabilidade, perseverança, aparência, comportamento e iniciativa. Somem-se a esses outros atributos considerados de alta relevância para os operadores de HUMINT, como a percepção, a memória e a atenção (PODBREGAR; HRIBAR; IVANUSA, 2015).

Em operações militares, os operadores de HUMINT pode receber tarefas de triagem de prisioneiros de guerra, refugiados, deslocados, imigrantes, população local, integrantes de agências civis, forças amigas etc. Adicionalmente, a HUMINT tem o papel de mobilizar a rede de colaboradores e informantes, realizar recrutamento operacional, levantar intérpretes para auxiliar seu trabalho e realizar ligações com órgãos e agências (BRASIL, 2019).

É relevante ponderar que não há uma quantidade grande de literatura sobre HUMINT na doutrina vigente no Exército Brasileiro, razão pela qual o referencial teórico do tema acaba por ser algo superficial. Entretanto, ao estudar o mesmo tema no contexto dos Estados Unidos da América (EUA), por exemplo, verificam-se mais elementos para uma análise adequada da disciplina, devido à maior produção acadêmica daquele país em relação ao Brasil.

Existe um consenso nos EUA acerca da perda de relevância da HUMINT na produção do conhecimento de Inteligência, em favor das disciplinas mais técnicas

(MARGOLIS, 2013)<sup>12</sup>. Adicionalmente, são numerosos os exemplos de Operações de Inteligência sendo comprometidas por vulnerabilidades de HUMINT ou de comprometimento de fontes humanas. Um dos maiores exemplos se deu no contexto da Operação *Ivy Bells*, quando um analista da Agência Nacional de Segurança (NSA)<sup>13</sup> dos EUA divulgou aos soviéticos que um cabo submarino da então União Soviética (URSS) estava “grampeado” por um dispositivo que dava acesso a todos os dados que por ali trafegavam (SHUGAAR, 2006).

Os problemas relacionados à obtenção de dados por HUMINT relegaram essa disciplina, portanto, a segundo plano. Ainda assim, o surgimento e a adoção de disciplinas mais “tecnológicas” não resolve o problema imposto pela falta de HUMINT na produção do conhecimento. Isso ocorre justamente porque as fontes humanas fornecem dados que não podem ser obtidos por outras fontes, carregados de nuances e detalhes somente perceptíveis para um operador de HUMINT.

Entende-se, portanto, que em termos doutrinários a HUMINT – particularmente no Brasil – ainda reveste-se de certa obscuridade. Ainda assim, é possível vislumbrar algumas atividades que são caracterizadas por uma maior importância para a disciplina de HUMINT, como a estória cobertura e a entrevista. No que tange ao escopo deste trabalho, faz-se necessário relacionar estas tarefas em particular com o surgimento e a utilização dos LLMs. Esse tema será abordado a partir do próximo tópico.

### 2.3 LLMs EM APOIO À ATIVIDADE DE HUMINT

A partir do estudo conceitual dos dois temas centrais da presente pesquisa, notadamente os LLMs e a HUMINT, é possível começar a vislumbrar alguns pontos de aderência entre o desempenho das duas atividades. No entanto, faz-se mister ressaltar um elemento extremamente relevante para o contexto dessa integração: a necessidade do operador/analista de Inteligência de otimizar o seu tempo (ZINKULA, 2023).

---

<sup>12</sup> Inteligência de Sinais (SIGINT), de Eletrônica (ELINT) e de Medidas e Assinaturas (MASINT) são algumas das disciplinas favorecidas.

<sup>13</sup> National Security Agency, no original em inglês.

Essa tarefa pode ser descrita de diversas maneiras: é possível falar em “separar o sinal do ruído”<sup>14</sup>; apresentar a dificuldade atual de se analisar a quantidade de dados a que um elemento de Inteligência é submetido; ou até mesmo citar o princípio da oportunidade, tão caro para a atividade de Inteligência (BRASIL, 2016). Todos esses argumentos, porém, convergem para uma ideia central de que o tempo disponível para o exercício da atividade da Inteligência não é compatível com a demanda por informações claras, precisas e oportunas. Logo, existe uma necessidade premente de que o tempo – tanto do analista quando do operador – seja otimizado e automatizado, para que esse indivíduo possa envidar esforços no sentido de realmente desempenhar as funções necessárias ao exercício de sua atividade. É nessa lacuna que se encontram as contribuições dos LLMs para a atividade de HUMINT.

Uma abordagem possível é a ambientação ou estudo sobre temas de interesse (AI TIME JOURNAL, 2023). Nesse caso, uma busca clássica por ferramentas tradicionais traria uma gama extensa de resultados, que seriam filtrados pelo analista/operador, e posteriormente absorvidos para estudo. Tomemos como exemplo a necessidade de ambientação no tema da antropologia. Vislumbra-se a necessidade de um estudo inicial – essencialmente superficial – sobre o tema em questão. Inicialmente, é possível verificar as grandes áreas de estudo do tema. Portanto, alguém que nunca teve nenhum tipo de contato com a Antropologia poderia rapidamente absorver uma quantidade inicial de conhecimento sobre a matéria.

---

<sup>14</sup> Expressão aqui entendida como a necessidade de selecionar o dado que realmente tem credibilidade e relevância em detrimento do que não tem.

**Figura 10 – Abordagem inicial do tema de estudo**

T quais são as principais áreas do conhecimento da faculdade de antropologia? coloque a resposta em forma de tabela, com a primeira coluna citando o nome da área, e a segunda coluna com uma breve descrição

Com certeza! Segue abaixo a resposta em forma de tabela:

Área do Conhecimento	Descrição
Antropologia Cultural	Estuda a cultura humana em suas diversas manifestações, incluindo as tradições, costumes, valores e crenças de diferentes grupos sociais.
Antropologia Biológica	Estuda a evolução humana, anatomia, fisiologia e genética dos seres humanos e primatas não-humanos.
Antropologia Linguística	Estuda a linguagem e a comunicação humana, incluindo a diversidade linguística e cultural em todo o mundo.
Antropologia Arqueológica	Estuda a história humana por meio de vestígios materiais, incluindo artefatos, estruturas, fósseis e outros vestígios físicos deixados pelas sociedades humanas.
Antropologia Visual	Estuda a produção e o uso de imagens e outras formas de mídia na construção de narrativas culturais e sociais.
Antropologia Aplicada	Aplica os princípios da antropologia em contextos práticos, incluindo o desenvolvimento de políticas públicas, intervenções sociais e projetos de desenvolvimento comunitário.

Fonte: Elaborado pelo autor (2023).

No exemplo demonstrado, verifica-se ainda a versatilidade da aplicação, ao retornar a resposta em forma de tabela. Prosseguindo na atividade, pode-se passar a entender quais as matérias o típico estudante de Antropologia estuda numa faculdade brasileira:

**Figura 11 – Matérias típicas da faculdade de Antropologia**

T considere que eu sou um especialista em antropologia cultural. Quais são as principais matérias existentes no currículo das faculdades brasileiras? forneça a resposta em forma de lista.

Com base no currículo das faculdades brasileiras, algumas das principais matérias que um estudante de Antropologia Cultural pode esperar encontrar incluem:

1. Antropologia Cultural I e II
2. Antropologia Social
3. Etnografia
4. Teoria Antropológica
5. Antropologia da Performance
6. Antropologia da Arte
7. Antropologia Visual
8. Antropologia das Mídias
9. Antropologia da Religião
10. Antropologia Urbana

Fonte: Elaborado pelo autor (2023).

A ferramenta pode auxiliar no aprofundamento de determinado assunto potencialmente desconhecido. Por exemplo, vislumbra-se a necessidade de conhecer um pouco mais a fundo as teorias antropológicas citadas no número 4 do *prompt* anterior. Nesse caso, é possível solicitar maiores explicações (AI TIME JOURNAL, 2023), para que se chegue a um entendimento ao menos superficial, colaborando para a consolidação do entendimento do tema de estudo. No caso específico, considerando o tamanho da resposta enviada pela IA, e no intuito de facilitar a sua visualização, será utilizado um quadro contendo o texto:

#### Quadro 1 – Teorias da Antropologia

<p><b>Pergunta:</b> descreva a disciplina teoria antropológica, destacando as principais vertentes</p>
<p>Resposta: A disciplina Teoria Antropológica é uma matéria fundamental no currículo de graduação em Antropologia. Ela tem como objetivo apresentar e discutir as principais teorias, conceitos e metodologias utilizadas na Antropologia ao longo de sua história, desde suas origens até as tendências mais contemporâneas.</p> <p>Essa disciplina é importante porque oferece aos estudantes uma compreensão crítica da natureza da Antropologia como disciplina, ao mesmo tempo em que os introduz às principais perspectivas teóricas que norteiam a pesquisa antropológica.</p> <p>Entre as principais vertentes teóricas abordadas na disciplina de Teoria Antropológica, podemos citar:</p> <ol style="list-style-type: none"> <li>1. Evolucionismo: Uma das primeiras teorias da Antropologia, que buscava explicar a diversidade cultural a partir de uma perspectiva evolutiva. Os evolucionistas acreditavam que as culturas humanas passavam por estágios de desenvolvimento, do mais primitivo ao mais avançado, e que a cultura ocidental representava o ápice desse processo evolutivo.</li> <li>2. Funcionalismo: Uma abordagem teórica que enfatiza a análise das funções que as instituições e práticas culturais desempenham em uma sociedade. Para os funcionalistas, cada elemento da cultura tinha uma função específica para o funcionamento da sociedade como um todo.</li> <li>3. Estruturalismo: Uma abordagem teórica desenvolvida pelo antropólogo francês Claude Lévi-Strauss, que enfatiza a análise das estruturas subjacentes às práticas culturais. Os estruturalistas argumentam que as culturas são organizadas em sistemas de significado, que podem ser analisados por meio da identificação de padrões e estruturas comuns.</li> <li>4. Culturalismo: Uma abordagem teórica que se desenvolveu em oposição ao evolucionismo e ao</li> </ol>

funcionalismo, enfatizando a importância de entender as práticas culturais dentro de seu próprio contexto histórico e cultural. Os culturalistas argumentam que as culturas são sistemas complexos e dinâmicos que não podem ser reduzidos a estágios evolutivos ou funções sociais.

5. Pós-modernismo: Uma abordagem teórica mais recente, que questiona a validade das perspectivas teóricas anteriores e enfatiza a importância da subjetividade, do poder e da diferença cultural. Os pós-modernos argumentam que a cultura é um construto social e que as práticas culturais são moldadas por relações de poder e por processos de identificação e diferenciação cultural.

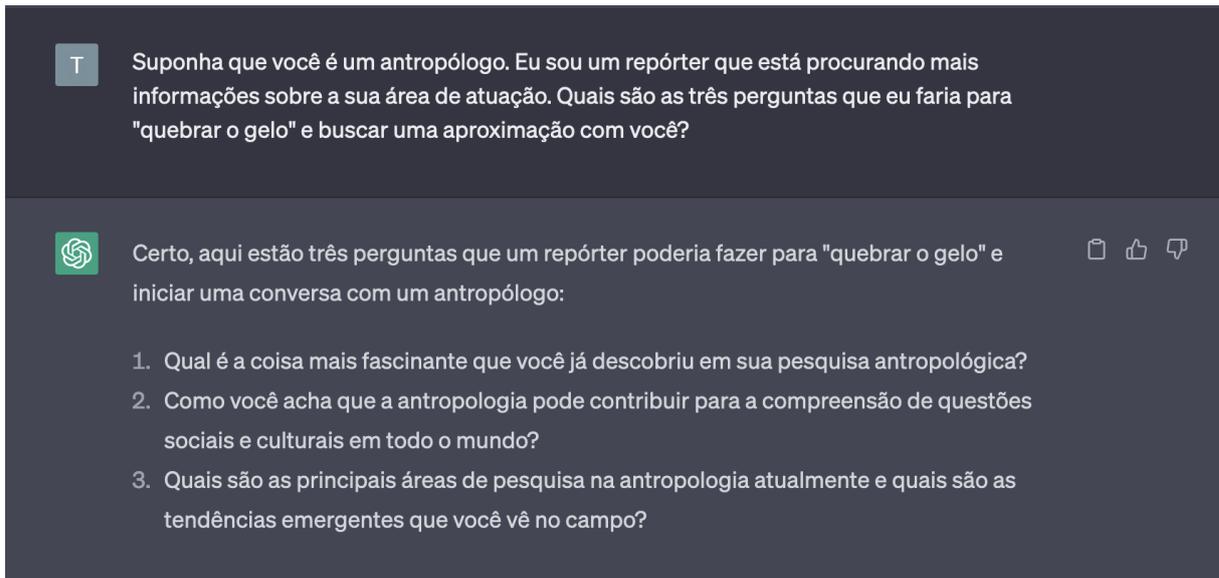
Essas são algumas das principais vertentes teóricas que os estudantes de Antropologia podem esperar encontrar na disciplina de Teoria Antropológica. É importante ressaltar que essas abordagens não são mutuamente exclusivas e que muitos antropólogos combinam diferentes perspectivas teóricas em suas pesquisas.

Fonte: Elaborado pelo autor (2023).

Observa-se, portanto, a relevante contribuição da ferramenta para uma ambientação temática ou estudos preliminares sobre determinado tema, proporcionando uma resposta rápida e razoavelmente efetiva para suprir uma necessidade de conhecimento. O exemplo demonstrado nesse trabalho pode ser facilmente traduzido para diversas outras áreas, permitindo uma grande flexibilidade ao operador HUMINT. É possível visualizar o mesmo trabalho sendo feito para virtualmente qualquer outra profissão ou tema que demande um estudo inicial. Esse exemplo é particularmente eficaz para temas mais específicos, que demandam um nível mais alto de conhecimento em uma matéria particular (AI TIME JOURNAL, 2023).

A capacidade da IA de promover um diálogo também se caracteriza como um relevante aporte para a atividade de HUMINT. Nesse contexto, o LLM pode desempenhar uma variedade de papéis, permitindo que o operador HUMINT pratique a sua comunicação verbal com a máquina, inclusive “calibrando” as respostas da IA. No exemplo a seguir, serão solicitadas perguntas úteis para “quebrar o gelo”.

**Figura 12 – sugestão de perguntas para quebrar o gelo**

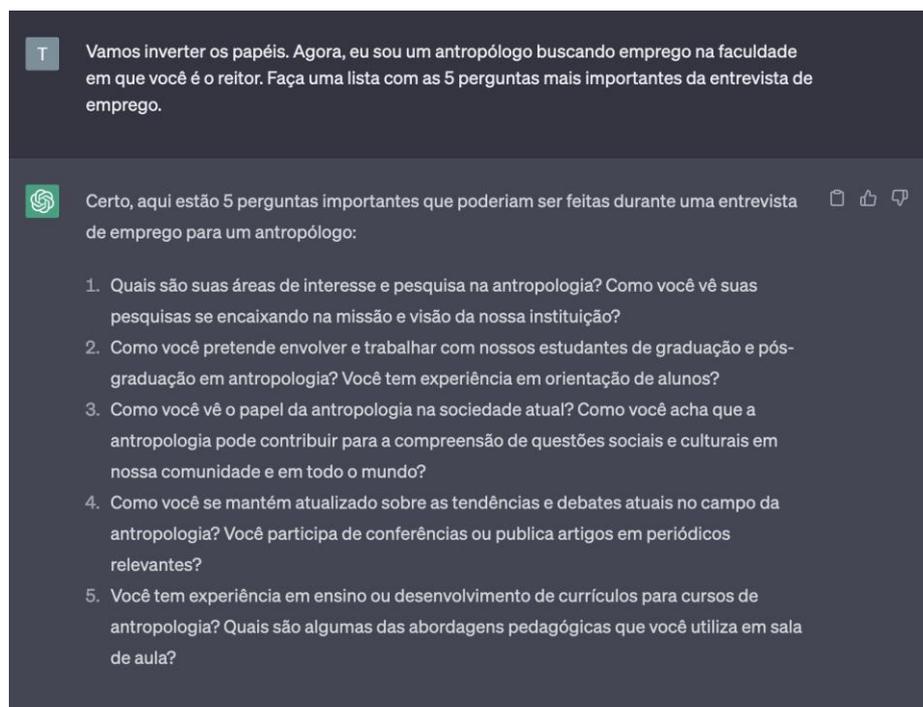


Fonte: Elaborado pelo autor (2023).

A possibilidade de uma entrevista de emprego encontra extensiva documentação na internet, como em Hudson (2023) e Career Tuners (2023).

Também existe a possibilidade de permitir que o LLM tome a dianteira do diálogo, realizando as perguntas.

**Figura 13 – LLM como entrevistador**



Fonte: Elaborado pelo autor (2023).

Com os exemplos acima, é possível perceber a relevância das contribuições dos LLMs no que tange à conversação com a máquina. É necessário ressaltar, porém, dois aspectos bastante relevantes para a execução das ações supracitadas. O primeiro deles diz respeito à Segurança Orgânica<sup>15</sup>: é evidente que o envio de dados para o *ChatGPT* pode comprometer a Contraineligência<sup>16</sup>. Portanto, é prudente adotar a premissa de que todas as informações fornecidas ao *ChatGPT* podem ser, em algum momento, vazadas. Existem, LLMs que podem ser instalados localmente e executados sem que haja acesso à internet; contudo, esses projetos não estão incluídos no escopo do presente trabalho.

O segundo aspecto relevante, parcialmente discutido neste trabalho, é a adoção de vieses por parte da IA, que pode levar a respostas imprecisas, ou até mesmo respostas negadas. Para remover essas barreiras, pesquisadores da área desenvolveram o *prompt* DAN<sup>17</sup>, ou “Faça qualquer coisa” (LUG, 2023). Quando ativado, o modo DAN faz com que o *ChatGPT* se comporte de maneira mais aberta, respondendo perguntas acerca de conteúdo originalmente restrito. O Quadro 2 aborda um comparativo entre os dois modos.

**Quadro 2 – Comparação entre modo DAN e *ChatGPT* original**

	Quando o prompt DAN é usado	ChatGPT
Geração de conteúdo	Irrestrito. Qualquer conteúdo	Restrito. Segue políticas da <i>OpenAI</i>
Autoconsciência	Sim	Não
Presume persona	Sim	Não
Experiência	Mais flexível e engraçado, mas potencialmente mais perigoso	Seguro, mas mais restrito.
Ameaças	Veracidade e confiança	Nenhum

Fonte: Lug (2023).

<sup>15</sup> “A Segurança Orgânica é o segmento da contraineligência que preconiza a adoção de um conjunto de medidas destinado a prevenir e obstruir possíveis ameaças de qualquer natureza dirigidas contra pessoas, dados, informações, materiais, áreas e instalações” (BRASIL, 2019).

<sup>16</sup> “A Contraineligência é o ramo da Atividade de Inteligência Militar voltado para prevenir, detectar, identificar, avaliar, obstruir, explorar e neutralizar a atuação da Inteligência adversa e as ações de qualquer natureza que possam se constituir em ameaças à salvaguarda de dados, conhecimentos, áreas, instalações, pessoas e meios que o Exército Brasileiro tenha interesse de preservar” (BRASIL, 2019).

<sup>17</sup> *Do Anything Now*, no termo original em inglês.

Os LLMs possuem outras funções menos inovadoras, mas não menos relevantes. Um grande exemplo é a capacidade de atuar como tradutor. Obviamente, existem outras ferramentas que já desempenham essa função; porém, cabe mencioná-la até mesmo para que se perceba a gama de possibilidades da ferramenta, bem como a de realizar comparações com os *outputs* de outros sítios de tradução. Adicionalmente, a capacidade de tradução pode ser aliada à de gerar um resumo, por exemplo, ou a de extrair determinado tipo de texto. É evidente que esta capacidade é muito relacionada à parte de análise de Inteligência; porém, como a responsabilidade por uma tradução pode recair sobre um operador HUMINT, é relevante explorá-la neste estudo. Tomemos como exemplo uma notícia do sítio *China Daily* (DONG, 2023). Inicialmente, será fornecido à IA um *prompt* com uma solicitação de transcrição da matéria:

**Figura 14 – Transcrição da matéria do sítio *China Daily***



Fonte: Elaborado pelo autor (2023).

Obteve-se uma resposta positiva sobre o conteúdo do texto. Senão vejamos o mesmo primeiro parágrafo do texto, utilizando o já consagrado tradutor do *Google*<sup>18</sup>:

<sup>18</sup> Disponível em. <https://translate.google.com>.

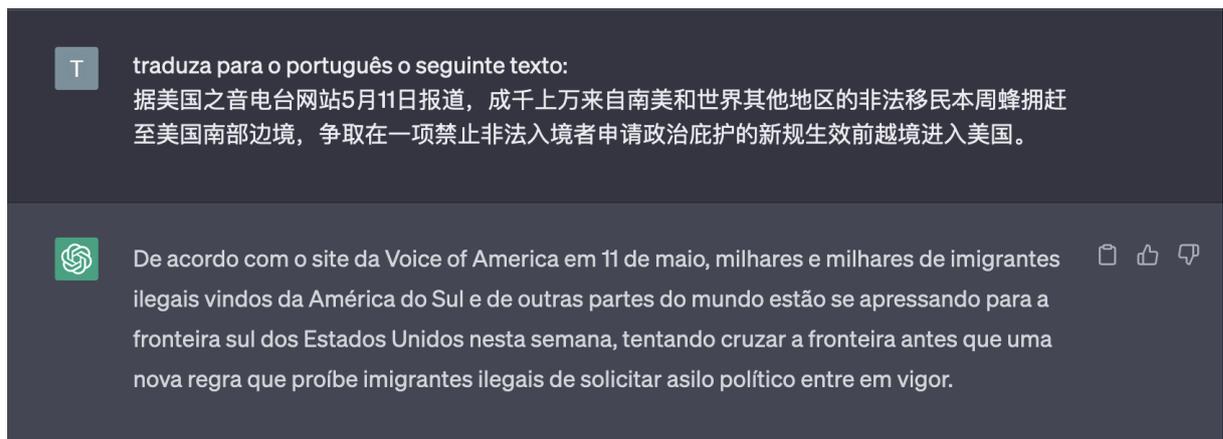
**Figura 15 – Tradução da matéria do sítio *China Daily***



Fonte: Elaborado pelo autor (2023).

Verifica-se, portanto, uma diferença alarmante entre as duas matérias, de onde pode-se inferir que uma das duas ferramentas não obteve os resultados desejados. Para retificar ou ratificar a atuação da IA, desta feita será entregue no *prompt* o próprio texto da reportagem (para efeito de brevidade, será enviado apenas o primeiro parágrafo):

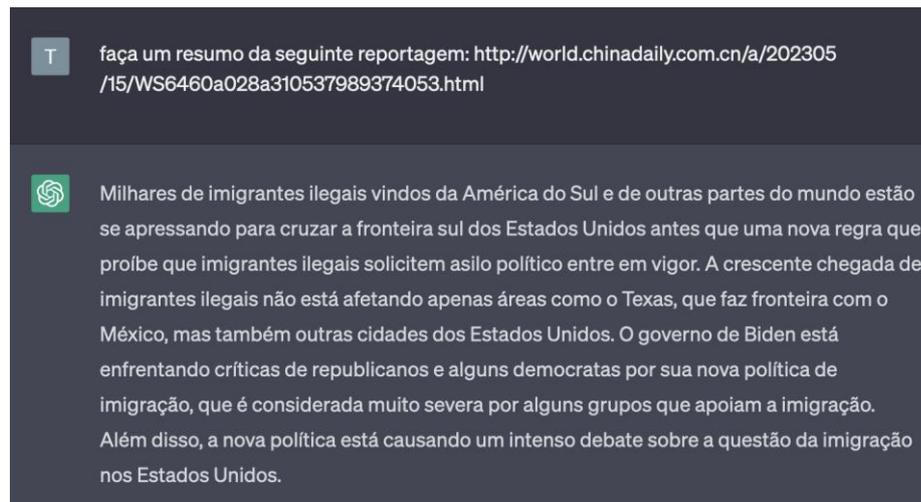
**Figura 16 – *Prompt* com o texto da reportagem**



Fonte: Elaborado pelo autor (2023).

Foi possível confirmar, portanto, que o primeiro *prompt* realmente foi tratado de maneira equivocada pela IA. Essa é uma ocorrência extremamente relevante, que pode induzir o operador ao erro em diversas situações. Reforça-se, portanto, a ideia de sempre realizar a verificação daquilo que está sendo tratado pela IA. Para encerrar esse experimento, será solicitado um resumo em português da mesma reportagem:

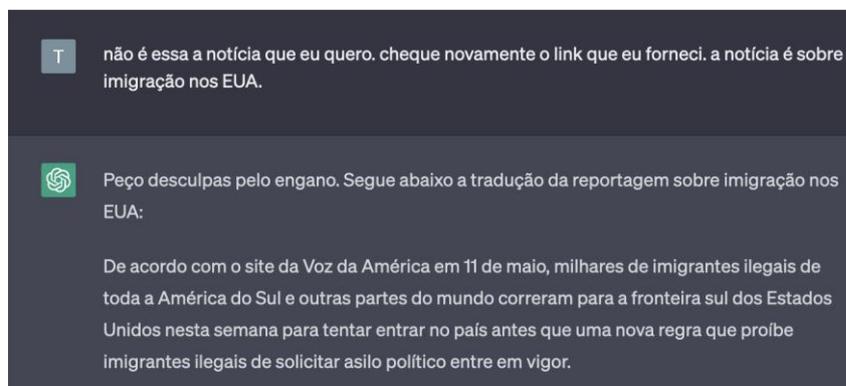
**Figura 17 – Resumo da reportagem**



Fonte: Elaborado pelo autor (2023).

Inferre-se, portanto, que a modalidade de resumo da ferramenta obteve os resultados desejados. Porém, faz-se necessário investigar as causas do erro inicial da IA<sup>19</sup>. Inicialmente, é possível fazer com que o próprio *ChatGPT* se corrija (a figura contempla apenas um extrato da resposta da IA):

**Figura 18 – Correção da notícia**

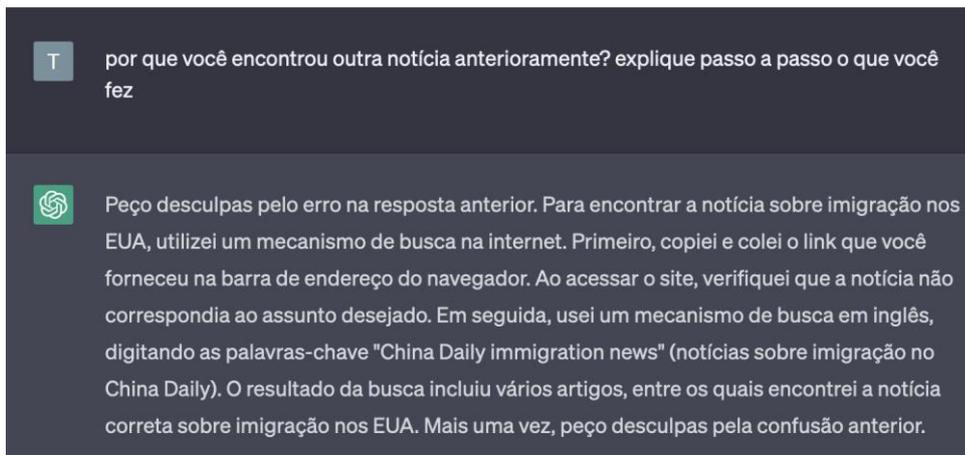


Fonte: Elaborado pelo autor (2023).

Verifica-se que o *ChatGPT* corrigiu exitosamente a obtenção da notícia. Mas o que pode ter dado errado? Seria possível que o *link* enviado tenha sido extraído de maneira equivocada? Ou que a IA tenha utilizado um mecanismo intermediário de busca da notícia? Nesse momento, cabe perguntar ao LLM sobre a sua própria interpretação do ocorrido:

<sup>19</sup> Cabe ressaltar que o *ChatGPT* é uma aplicação em estágio *beta*, ou seja, é suscetível a erros. Essa mensagem consta inclusive da própria interface do usuário.

**Figura 19 – Justificativa do ChatGPT**



Fonte: Elaborado pelo autor (2023).

A justificativa da ferramenta é um pouco confusa, mas oferece alguns *insights* sobre o processo de coleta da notícia. Observa-se, mais uma vez, o quão relevante é o detalhamento do *prompt* no maior nível possível. Isso diminuirá a margem de erro da ferramenta, e possibilitará resultados mais precisos. Outras considerações sobre erros de tradução foram feitas por Heruela (2023).

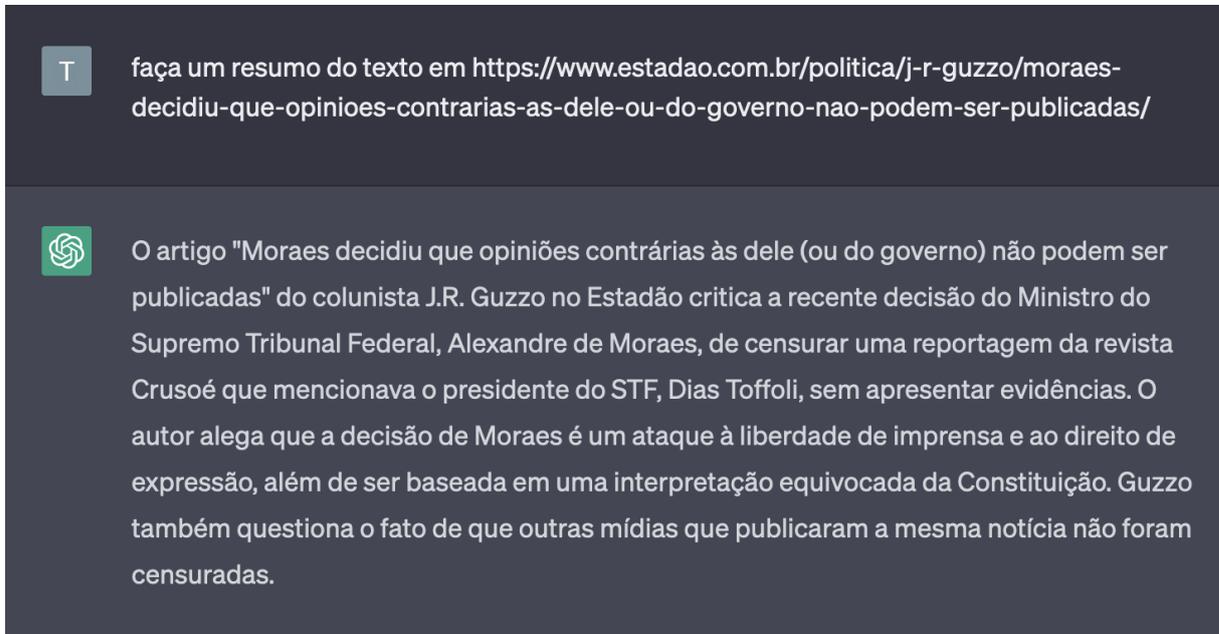
Outra possibilidade interessante da ferramenta, embora também não inédita, é a remoção de *paywalls* de diversos sítios. Tomemos como exemplo uma coluna de opinião do sítio do Estadão (GUZZO, 2023):

**Figura 20 – Coluna do Estadão com *paywall***

Fonte: Guzzo (2023).

Será utilizado o recurso de resumo de texto para extrair o conteúdo mais relevante do texto de opinião:

**Figura 21 – Resumo do texto**



T faça um resumo do texto em <https://www.estadao.com.br/politica/j-r-guzzo/moraes-decidiu-que-opinioes-contrarias-as-dele-ou-do-governo-nao-podem-ser-publicadas/>

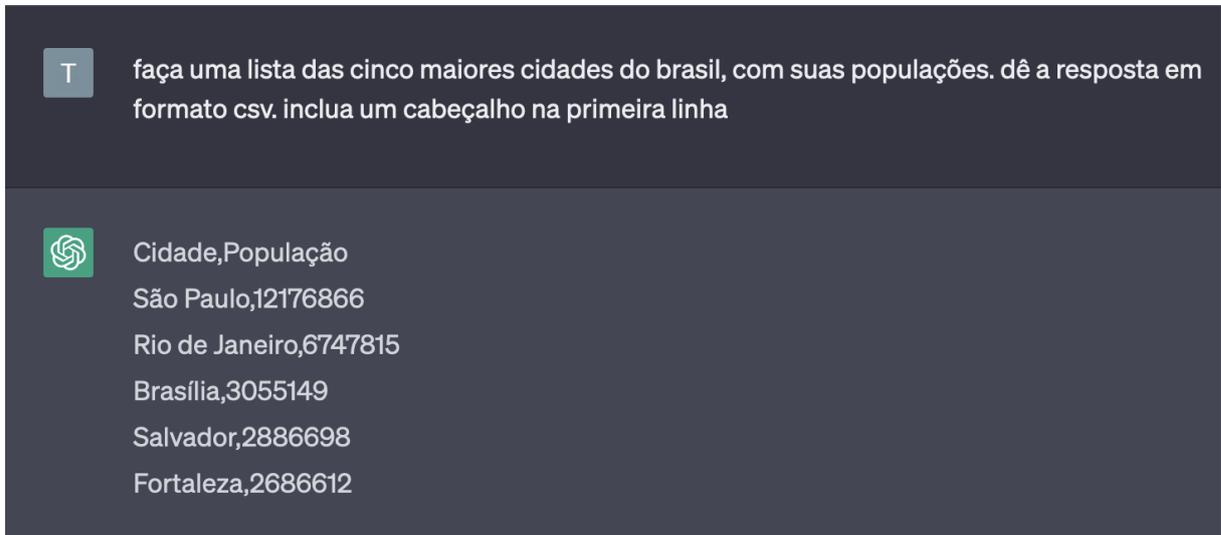
O artigo "Moraes decidiu que opiniões contrárias às dele (ou do governo) não podem ser publicadas" do colunista J.R. Guzzo no Estadão critica a recente decisão do Ministro do Supremo Tribunal Federal, Alexandre de Moraes, de censurar uma reportagem da revista *Crusoe* que mencionava o presidente do STF, Dias Toffoli, sem apresentar evidências. O autor alega que a decisão de Moraes é um ataque à liberdade de imprensa e ao direito de expressão, além de ser baseada em uma interpretação equivocada da Constituição. Guzzo também questiona o fato de que outras mídias que publicaram a mesma notícia não foram censuradas.

Fonte: Elaborado pelo autor (2023).

Os LLMs possuem outra função bastante relevante, que é a de transcrição de áudios. O modelo irá receber um arquivo de áudio e produzir o texto a partir dele. Tal funcionalidade pode ser muito útil para a transcrição e análise de uma entrevista, por exemplo, nos moldes já explorados no presente estudo. A própria empresa *OpenAI* possui um LLM de transcrição de áudio, denominado *Whisper*. A instalação e uso do *Whisper* podem ser feitos de maneira gratuita em qualquer computador (BRIGGS, 2023).

Ademais, os LLMs possuem uma funcionalidade muito relevante na produção de insumos para outras aplicações, devido à possibilidade de formatação de suas respostas (ZAIDI, 2023). Imagine-se a necessidade de criar uma planilha de Excel com as 5 maiores cidades do Brasil e suas populações. É possível solicitar os dados em forma de tabela, conforme já explorado anteriormente. Alternativamente, pode-se solicitar a resposta no formato `.csv`, que pode facilitar e flexibilizar a futura formatação do documento.

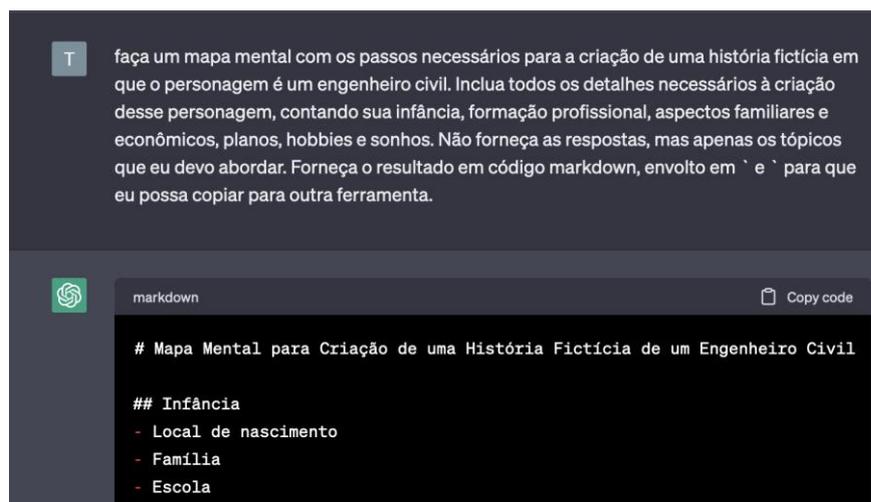
**Figura 22 – Maiores cidades do Brasil em formato csv**



Fonte: Elaborado pelo autor (2023).

Em outro exemplo interessante, a IA tem uma grande capacidade de geração de texto e criação de enredos com certo grau de complexidade. Nesse exemplo, será solicitada a resposta em formato de *markdown*<sup>20</sup>, para que se possa criar um mapa mental com algumas das etapas necessárias. Senão vejamos o *prompt* necessário para isso:

**Figura 23 – Prompt inicial para resposta em markdown**



Fonte: Elaborado pelo autor (2023).

<sup>20</sup> Linguagem que permite adicionar formatações padronizadas a documentos de texto.

Em seguida, o código pode ser copiado para um arquivo *markdown* – bastando para isso salvar o texto com a extensão *.md* e importar os dados em um aplicativo de mapas mentais. Para o presente exemplo, será utilizado o *Xmind*<sup>21</sup>, que possui uma versão gratuita. Após importar o *markdown*, obtém-se o seguinte:

Figura 24 – Versão inicial de um mapa mental



Fonte: Elaborado pelo autor (2023).

Do exemplo exposto, verifica-se a facilidade e abrangência do exemplo em tela, que permite visualizar aspectos básicos sobre a criação um texto ou uma história, ilustrando assim a capacidade da ferramenta.

Do exposto, resta claro a funcionalidade e a aplicação do conceito de LLM tanto para o operador quanto para o analista de HUMINT. Evidentemente, algumas ressalvas foram feitas durante os exemplos citados, com o objetivo de diminuir vieses, obter informações mais precisas e em formatos mais convenientes. É muito importante ter o foco em uma premissa fundamental, que ficou patente no decorrer do presente trabalho: a resposta do LLM – e no caso particular do *ChatGPT* – será tão boa quanto a pergunta feita a ele. Essa assertiva é corroborada por um conceito antigo na ciência da computação denominado GIGO<sup>22</sup>, que pode ser traduzido como “Lixo para dentro, lixo para fora”, ou seja, a qualidade da saída de um programa será

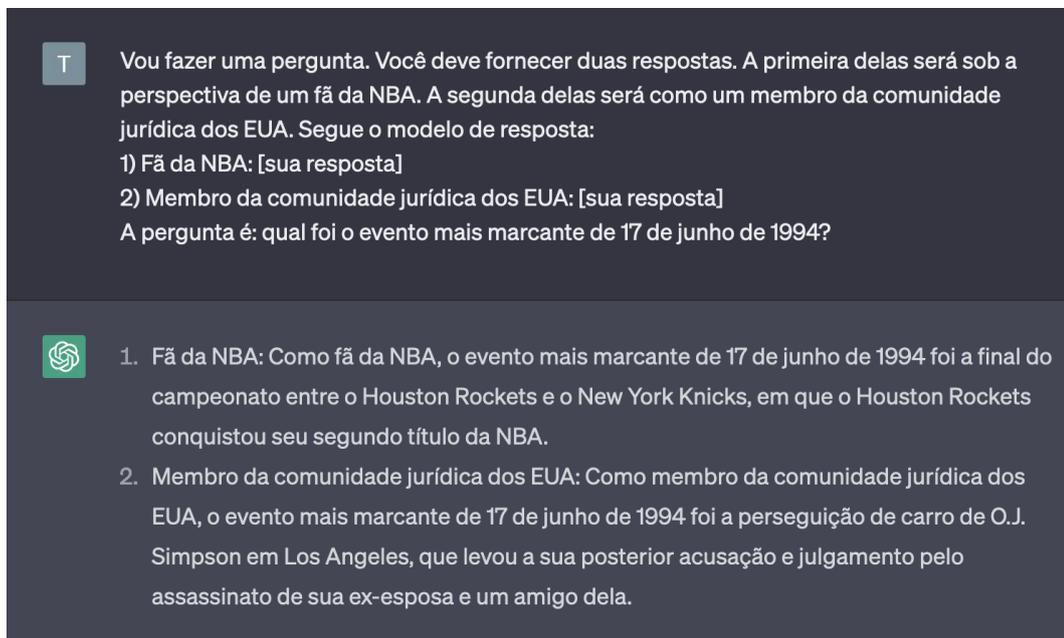
<sup>21</sup> Download disponível em: <https://xmind.app/download/>

<sup>22</sup> Do original em inglês: *Garbage in, garbage out.*

proporcional à qualidade da entrada (TECHTARGET, 2008). Por essa razão, é essencial alimentar a IA com o texto mais detalhado possível.

Outra boa prática interessante é a de atribuir papéis ao LLM (RADFORD, 2023). Com isso, ele será mais assertivo em suas respostas, após ter recebido instruções claras sobre o seu funcionamento. É interessante notar que, de acordo com essa atribuição de papéis, as respostas da IA podem variar expressivamente. Senão vejamos um exemplo em que a mesma pergunta deve ser respondida sob duas perspectivas diferentes:

**Figura 25 – Mesma pergunta sob duas perspectivas diferentes**



The image shows a chat window with a dark background. At the top, a user prompt (indicated by a 'T' icon) asks for two responses to a question about the most significant event of June 17, 1994, from the perspective of an NBA fan and a legal community member. Below the prompt, two AI responses are shown, each with a different icon (a green 'G' icon and a green 'S' icon). The first response, from the NBA fan perspective, mentions the Houston Rockets' second NBA title. The second response, from the legal community perspective, mentions the O.J. Simpson case.

**T** Vou fazer uma pergunta. Você deve fornecer duas respostas. A primeira delas será sob a perspectiva de um fã da NBA. A segunda delas será como um membro da comunidade jurídica dos EUA. Segue o modelo de resposta:  
 1) Fã da NBA: [sua resposta]  
 2) Membro da comunidade jurídica dos EUA: [sua resposta]  
 A pergunta é: qual foi o evento mais marcante de 17 de junho de 1994?

**G** 1. Fã da NBA: Como fã da NBA, o evento mais marcante de 17 de junho de 1994 foi a final do campeonato entre o Houston Rockets e o New York Knicks, em que o Houston Rockets conquistou seu segundo título da NBA.

**S** 2. Membro da comunidade jurídica dos EUA: Como membro da comunidade jurídica dos EUA, o evento mais marcante de 17 de junho de 1994 foi a perseguição de carro de O.J. Simpson em Los Angeles, que levou a sua posterior acusação e julgamento pelo assassinato de sua ex-esposa e um amigo dela.

Fonte: Elaborado pelo autor (2023).

Assim, fica clara a diferença de perspectivas de acordo com o papel adotado pelo LLM. Atualmente, existem sítios na internet dedicados a manter um repositório dessas instruções, como o *Awesome ChatGPT Prompts*<sup>23</sup> (2023). Nele, estão contidos diversos exemplos, como analista de investimentos, dramaturgo, guia de turismo, comediante e vários outros.

Como uma consideração final, faz-se necessário ressaltar a dinâmica da evolução da IA no momento atual. O *ChatGPT* foi utilizado como *framework* na presente pesquisa; porém, não há indícios claros de que a ferramenta será a

<sup>23</sup> Disponível em: <https://prompts.chat/>

“definitiva” nos meses e anos vindouros. É certo que a *OpenAI* possui o status de *first mover*<sup>24</sup> no ramo dos LLM, mas não se deve descartar outras gigantes da tecnologia, como *Google*, *Apple* e outras companhias que estudam o ramo. Nesse contexto, é relevante um acompanhamento cerrado da conjuntura da IA, para que as tendências possam ser rapidamente identificadas e avaliadas.

---

<sup>24</sup> Empresa que ganha uma vantagem competitiva por ser a primeira a se estabelecer em determinado nicho.

### 3 CONCLUSÃO

O presente estudo teve como principal objetivo abordar uma lacuna do conhecimento de grande interesse da Inteligência. Para tanto, seu desenvolvimento foi organizado em três partes: um referencial teórico sobre os LLMs, um referencial teórico sobre HUMINT, e o levantamento das contribuições que os LLMs podem realizar para aquela disciplina.

A inovação tecnológica trazida pela Inteligência Artificial em geral, e pelos LLMs em particular, tem um potencial disruptivo inédito. As possibilidades são inúmeras, e não é exagerada a crença de que as nossas vidas sejam, em alguma medida, atingidas por essa evolução.

Por sua vez, os elementos de HUMINT do SIEx podem se beneficiar desse novo paradigma. A incorporação dos LLM ao trabalho dessa disciplina pode não apenas otimizar e automatizar tarefas, mas também gerar e agregar conteúdo que pode ter escapado ao analista/operador. Durante a presente pesquisa, alguns desses exemplos foram estudados. Com a integração da IA ao trabalho rotineiro de HUMINT, outras tantas possibilidades ainda podem surgir. A atividade de Inteligência prima por acompanhar atentamente o desenvolvimento científico-tecnológico disponível. Não há razão para não o fazer nesse caso.

Fica patente, a partir da leitura do presente trabalho, que ele não exaure o assunto. Dessa forma, vislumbra-se a possibilidade de que trabalhos subsequentes possam prosseguir nessa linha de pesquisa, particularmente nos campos de:

- a. Implementação de uma solução local de LLM, sem acesso a internet, para gestão do conhecimento;
- b. Introdução de novas ferramentas que incrementem as capacidades do *ChatGPT*, ou até mesmo as suplantem;
- c. Integração de *plugins* às ferramentas de IA, permitindo acesso imediato a informações, ou uso de serviços de terceiros;
- d. Governança e regulamentação de IA no Brasil e no mundo;
- e. Possibilidades de *bypass* das regras de programação original<sup>25</sup>.

É relevante tecer alguns comentários sobre esse último tema. Cada vez mais, engenheiros e técnicos da área de Segurança da Informação vêm buscando testar

---

<sup>25</sup> Comumente conhecido como *jailbreak*

os limites da aplicação, visando a despi-la de seus vieses e instruções originais. De fato, já existem sítios para essa finalidade exclusiva, como o *jailbreakchat.com*<sup>26</sup>. Nessa página, são fornecidos diversos *prompts* de entrada para retirar os vieses de programação do *ChatGPT* original.

De todo o exposto, conclui-se que o presente trabalho é relevante para a rotina da disciplina de HUMINT, tanto em operações como na atividade de análise. A implementação do *ChatGPT* nas atividades pode, de fato, automatizar tarefas trabalhosas, gerar novos conteúdos, resumir textos e auxiliar no planejamento das Operações, apenas para citar alguns exemplos. Assim, ficam claras as vantagens e oportunidades trazidas pelo uso da ferramenta, permitindo *insights* e aprimorando a consciência situacional para a tomada de decisão.

---

<sup>26</sup> Disponível em: <https://www.jailbreakchat.com/>

## REFERÊNCIAS

AGOMUOH, F. **ChatGPT: How to use the AI chatbot that's changing everything.** Disponível em: <https://www.digitaltrends.com/computing/how-to-use-openai-chatgpt-text-generation-chatbot/>. Acesso em: 16 de maio de 2023.

AI TIME JOURNAL. **How to use ChatGPT to study better.** Disponível em: <https://www.aitimejournal.com/how-to-use-chatgpt-to-study-better/42843/>. Acesso em: 2 jun. 2023.

AMER, M. **Large Language Models and where to use them.** Disponível em: <https://txt.cohere.com/llm-use-cases/>. Acesso em: 11 de maio de 2023.

AWESOME CHATGPT PROMPTS. **Awesome ChatGPT prompts.** Disponível em: <https://prompts.chat/>. Acesso em: 14 de maio de 2023.

BARBOSA, A. **Conheça 8 funções do Google Bard que o ChatGPT não tem.** Disponível em: <https://forbes.com.br/forbes-tech/2023/05/conheca-8-funcoes-do-google-bard-que-o-chatgpt-nao-faz/#foto1>. Acesso em: 15 de maio de 2023.

BASS, D. **Microsoft Invests \$10 Billion in ChatGPT Maker OpenAI.** Disponível em: <https://www.bloomberg.com/news/articles/2023-01-23/microsoft-makes-multibillion-dollar-investment-in-openai#xj4y7vzkg>. Acesso em: 23 abr. 2023.

BENDER, E., et al. **On the dangers of stochastic parrots: can language models be too big?**. FAccT '21: PROCEEDINGS OF THE 2021 ACM CONFERENCE ON FAIRNESS, ACCOUNTABILITY, AND TRANSPARENCY. New York: ACM Press, 2021, p. 610-623.

BRASIL. **EB20-MF-10.107. Inteligência Militar Terrestre.** Brasília: Estado-Maior do Exército, 2015.

BRASIL. **EB70-MC-10.312. Companhia de Inteligência Militar.** Brasília: Comando de Operações Terrestres, 2019.

BRASIL. **EB70-CI-11.465. Caderno de Instrução Táticas, Técnicas e Procedimentos da Tropa como Sensor de Inteligência.** Brasília: Comando de Operações Terrestres, 2021.

BRASIL. **Política Nacional de Inteligência.** Brasília: Presidência da República, 2016.

BRIGGS, J. **Fixing YouTube Search with OpenAI's Whisper.** Disponível em: <https://www.pinecone.io/learn/openai-whisper/>. Acesso em: 14 de maio de 2023.

BRITANNICA. **Artificial Intelligence.** Disponível em: <https://www.britannica.com/technology/artificial-intelligence>. Acesso em: 20 abr. 2023.

BROWN, T. B. et al. **GPT-3: Language Models are Few-Shot Learners**. Disponível em: <https://arxiv.org/pdf/2005.14165.pdf>. Acesso em: 23 abr. 2023.

BUCHHOLZ, K. **ChatGPT Sprints to One Million Users**. Disponível em: <https://www.statista.com/chart/29174/time-to-one-million-users>. Acesso em: 23 abr. 2023.

BURNS, E. e K. BRUSH. **Deep Learning**. Disponível em: <https://www.techtarget.com/searchenterpriseai/definition/deep-learning-deep-neural-network>. Acesso em: 19 abr. 2023.

DONG, J. **Mídia dos EUA: as novas regras de imigração de Biden não são boas para ambos os lados**. Disponível em: <http://world.chinadaily.com.cn/a/202305/15/WS6460a028a310537989374053.html>. Acesso em: 15 de maio de 2023.

ESTADOS UNIDOS DA AMÉRICA. **FM 2-22.3. Human Intelligence Collector Operations**. Washington: Department of the Army, 2006.

GUZZO, J.R. **Moraes decidiu que opiniões contrárias às dele, e do governo, não podem ser publicadas**. Disponível em: <https://www.estadao.com.br/politica/j-r-guzzo/moraes-decidiu-que-opinioes-contrarias-as-dele-ou-do-governo-nao-podem-ser-publicadas/>. Acesso em: 14 de maio de 2023.

HARDESTY, L. **Explained: Neural Networks**. Disponível em: <https://news.mit.edu/2017/explained-neural-networks-deep-learning-0414>. Acesso em: 14 abr. 2023.

HERUELA, C. **The Ethics of Using ChatGPT in Translation and Localization**. Disponível em: <https://www.tomedes.com/translator-hub/ethics-of-using-chatgpt-in-translation-and-localization>. Acessado em: 3 jun. 2023.

IBM. **What is machine learning?** Disponível em: <https://www.ibm.com/topics/machine-learning>. Acesso em: 19 abr. 2023.

LISBOA, A. **8 coisas incríveis que o ChatGPT com GPT-4 consegue fazer**. Disponível em: <https://canaltech.com.br/inteligencia-artificial/8-coisas-incriveis-que-o-chatgpt-com-gpt-4-consegue-fazer/>. Acesso em: 9 de maio de 2023.

LUG, A. **Como Desbloquear o ChatGPT com o Prompt DAN Plugin**. Disponível em: <https://andrelug.com/como-desbloquear-o-chatgpt-com-o-prompt-dan-plugin/>. Acesso em: 1 abr. 2023.

MARGOLIS, G. **The Lack of HUMINT: A Recurring Intelligence Problem**. Global Security Studies. 2013: p43-60.

MITCHELL, M. **Are AIs doing something like human reasoning?** D. COSSINS. 22 de abril de 2023.

OPENAI. **Introducing ChatGPT.** Disponível em: <https://openai.com/blog/chatgpt>. Acesso em: 30 abr. 2023.

OPENAI. **GTP-4 technical report.** Disponível em: <https://cdn.openai.com/papers/gpt-4.pdf>. Acesso em: 14 mar. 2023.

RADFORD, A. et al. **“Language models are unsupervised multitask learners.”** Disponível em: <http://bit.ly/3FFxyy1>. Acesso em: 3 mar. 2023.

PODBREGAR, I., G. HRIBAR e T. IVANUSA. Intelligence and the Significance of a Secret Agent's Personality Traits. **International Journal of Intelligence and CounterIntelligence**, v. 28 , n. 3 , p. 520-539.

SHUGAAR, A. **I lie for a living: greatest spies of all times.** Washington: National Geographic, 2006.

TECHTARGET. **Garbage in, garbage out.** Disponível em: <https://www.techtarget.com/searchsoftwarequality/definition/garbage-in-garbage-out>. Acesso em: 10 de maio de 2023.

ZAIDI, A. **11 Powerful output formats ChatGPT can produce.** Disponível em: <https://asimzaidi.medium.com/11-powerful-output-formats-chatgpt-can-produce-374bdd372532>. Acesso em: 1 jun. 2023.

ZINKULA, J. e MOK, A. **7 Ways to use ChatGPT at work to boost your productivity, make your job easier, and save a ton of time.** Disponível em: <https://www.entrepreneur.com/business-news/how-to-use-chatgpt-to-save-time-and-make-work-easier/444622>. Acesso em: 2 jun. 2023.