UNIVERSIDADE ESTADUAL PAULISTA JÚLIO DE MESQUITA FILHO
FACULDADE DE ARQUITETURA, ARTES E COMUNICAÇÃO
PROGRAMA DE PÓS-GRADUAÇÃO EM COMUNICAÇÃO LINHA DE PESQUISA:
PRODUÇÃO DE SENTIDO NA COMUNICAÇÃO MIDIÁTICA

Fabio Cardoso

"NOTHING, FOREVER": A PRÁXIS ENUNCIATIVA E A PRODUÇÃO DE SENTIDO NA GERAÇÃO AUTÔNOMA DO AUDIOVISUAL POR INTELIGÊNCIA ARTIFICIAL

Bauru
2024

FABIO CARDOSO

"Nothing, Forever": a práxis enunciativa e a produção de sentido na geração autônoma do audiovisual por inteligência artificial

Tese de Doutorado apresentada ao Programa de pós-graduação em Comunicação da Faculdade de Arquitetura, Artes e Comunicação da Universidade Estadual Paulista "Júlio de Mesquita Filho", campus Bauru, para a obtenção do título de doutor, sob a orientação da Professora Doutora Ana Silvia Lopes Davi Médola. Linha de Pesquisa 2: Produção de Sentido na Comunicação Midiática.

Bauru

2024

Cardoso, Fabio
"*Nothing, Forever*": A práxis enunciativa e a produção de sentido na geração autônoma do audiovisual por inteligência artificial / Fabio Cardoso, 2024
153 p. : il.

Orientadora: Ana Sílvia Lopes Davi Médola

Tese (Doutorado) - Universidade Estadual Paulista (UNESP), Faculdade de Arquitetura, Artes, Comunicação e Design, Bauru, 2024

1. Inteligência Artificial Generativa. 2. Semiótica. 3. Narrativa Contínua. 4. Nothing, Forever. I. Título.

Impacto potencial desta pesquisa

Esta pesquisa sobre narrativas audiovisuais autônomas feitas com inteligência artificial impacta os ODS 4 (Educação de Qualidade), ODS 8 (Trabalho Decente e Crescimento Econômico) e ODS 10 (Redução das Desigualdades). Busca capacitar tecnologicamente e garantir o uso consciente desta tecnologia, e contribuir para o crescimento e adoção sustentável da inteligência artificial no campo do audiovisual. Também promove a transparência e a ética no uso da IA.

Potential impact of this research

This research on autonomous audiovisual narratives made with artificial intelligence impacts SDG 4 (Quality Education), SDG 8 (Decent Work and Economic Growth), and SDG 10 (Reduced Inequalities). It aims to technologically empower and ensure conscientious use of this technology, contributing to the sustainable growth and adoption of artificial intelligence in the audiovisual field. It also promotes transparency and ethics in the use of AI.

Impacto potencial de esta investigación

Esta investigación sobre narrativas audiovisuales autónomas realizadas con inteligencia artificial impacta en los ODS 4 (Educación de Calidad), ODS 8 (Trabajo Decente y Crecimiento Económico) y ODS 10 (Reducción de las Desigualdades). Busca capacitar tecnológicamente y garantizar el uso consciente de esta tecnología, contribuyendo al crecimiento sostenible y la adopción de la inteligencia artificial en el campo audiovisual. También promueve la transparencia y la ética en el uso de la IA.

UNIVERSIDADE ESTADUAL PAULISTA

Câmpus de Bauru

ATA DA DEFESA PÚBLICA DA TESE DE DOUTORADO DE FABIO CARDOSO, DISCENTE DO PROGRAMA DE PÓS-GRADUAÇÃO EM COMUNICAÇÃO, DA FACULDADE DE ARQUITETURA, ARTES, COMUNICAÇÃO E DESIGN - CÂMPUS DE BAURU.

Aos 25 dias do mês de março do ano de 2024, às 14:00 horas, por meio de Videoconferência, realizou-se a defesa de TESE DE DOUTORADO de FABIO CARDOSO, intitulada **"Nothing, Forever": a práxis enunciativa e a produção de sentido na geração autônoma do audiovisual por inteligência artificial.**. A Comissão Examinadora foi constituida pelos seguintes membros: Professora Associada ANA SILVIA LOPES DAVI MEDOLA (Participação Presencial) do(a) Departamento de Comunicação Social / Faculdade de Arquitetura, Artes. Comunicação e Design de Bauru, Professor Doutor ALAN CÉSAR BELO ANGELUCI (Participação Presencial) do(a) Escola de Comunicações e Artes / Universidade de São Paulo, Professor Doutor CARLOS HENRIQUE SABINO CALDAS (Participação Presencial) do(a) Programa de Mestrado Profissional em Propriedade Intelectual e Transferência de Tecnologia para Inovação / Universidade do Estado de Minas Gerais, Prof. Dr. JOÃO CARLOS MASSAROLO (Participação Virtual) do(a) Departamento de Artes e Comunicação / Universidade Federal de São Carlos, Professor Associado JOÃO PEDRO ALBINO (Participação Presencial) do(a) Departamento de Computação / Faculdade de Ciências de Bauru. Após a exposição pelo doutorando e arguição pelos membros da Comissão Examinadora que participaram do ato, de forma presencial e/ou virtual, o discente recebeu o conceito final:_ APROVADO_. Nada mais havendo, foi lavrada a presente ata, que após lida e aprovada, foi assinada pelo(a) Presidente(a) da Comissão Examinadora.

Professora Associada ANA SILVIA LOPES DAVI MEDOLA

Antes de criar inteligências artificiais, por que nós não fazemos algo sobre a nossa estupidez natural?

Steve Polyak

AGRADECIMENTOS

Essa tese foi elaborada em um período conturbado, tanto de minha vida, quanto da própria humanidade. Passaram-se dois anos de uma pandemia global, de problemas sérios de saúde (tanto de minha parte, quanto da saúde física e mental da sociedade vigente), de questões políticas deploráveis, de mudanças significativas de vida. Agradecer a alguém, nomeando-o, seria uma injustiça com as inúmeras pessoas que, de uma forma ou de outra, contribuíram para a conclusão deste trabalho. Muitos torceram por mim, me motivaram, tentaram me levantar quando eu estava tombando. Meu agradecimento sincero a todos eles.

CARDOSO, Fabio. **Nothing, Forever: Geração autônoma de um produto audiovisual por inteligência artificial.** 153 f. Tese (Doutorado em Comunicação) – Faculdade de Arquitetura, Artes, Comunicação e Design, Universidade Estadual Paulista "Julio de Mesquita Filho", Bauru, 2024.

RESUMO

O impacto transformador da inteligência artificial (IA) em múltiplas áreas do conhecimento, entre as quais a comunicação, está presente em diversos segmentos do ecossistema midiático, motivando o surgimento de novas formas de produção e consumo. Esta tese examina a experiência de *Nothing, Forever*, um produto audiovisual concebido sob este cenário, que apresenta uma estrutura de produção e fruição temporal infinita, sendo realizada integralmente por IA em todas as etapas, desde o roteiro até a exibição, utilizando tecnologias de inteligência artificial generativa. Inicialmente, o trabalho apresenta a transição da IA de uma concepção predominantemente teórica, limitada a contextos acadêmicos e narrativas de ficção científica, para sua aplicação prática abrangente, impulsionada por avanços tecnológicos em processamento de dados, computação em nuvem e metodologias científicas. No contexto de incorporação da IA no âmbito da comunicação, destaca-se como a tecnologia promove uma reconfiguração paradigmática na produção e disseminação de informação, com um enfoque particular na IA generativa, abrangendo o potencial destas ferramentas para criar conteúdos comparáveis aos gerados por humanos, bem como suas implicações na criação de produtos como os audiovisuais. Observou-se no processamento e na geração de linguagem natural por meio dos *Large Language Models* (LLM) um ponto de inflexão na evolução histórica da interação entre computação e linguística, desde as primeiras teorias sobre a capacidade dos computadores de simular processos cognitivos humanos até a implementação prática de sistemas avançados de LLM, como o *Chat GPT*. A partir da articulação de postulações teóricas de Chomsky e Saussure sobre a natureza da linguagem e sua aquisição, identificou-se como essas teorias se relacionam com os princípios subjacentes ao design e à funcionalidade dos LLM. A análise de *Nothing, Forever* sob o arcabouço teórico da semiótica discursiva, notadamente na categoria contínuo *versus* descontínuo, demonstra uma projeção temporal que desafia a práxis enunciativa de uma narrativa audiovisual no que tange às segmentações dos processos produção e modos de fruição. Isso porque a sitcom apresenta uma estrutura temporal que difere das concepções tradicionais de narrativa. Marcada pela entrada e saída dos espectadores a qualquer momento da exibição ininterrupta de conteúdo, presta a uma determinada forma de segmentação em unidades de significação sob a ótica da semiótica discursiva. Esta segmentação destaca a alternância entre continuidade *versus* descontinuidade, onde momentos de engajamento contínuo são interrompidos por pausas ou mudanças na narrativa, enriquecendo a experiência do espectador. Diante dessas características de discursivização e estabelecimento de contrato enunciativo a tese propõe denominar tal experiência como "Narrativa Generativa Contínua" para designar conteúdos audiovisuais gerados por inteligência artificial e apresentados de maneira contínua, circunscrevendo um novo modo de compreender e categorizar a realização deste tipo de produção mediada por IA.

Palavras-chave: Inteligência artificial generativa. Semiótica. Narrativa Generativa Contínua. *Nothing, Forever.*

CARDOSO, Fabio. **Nothing, Forever: Autonomous generation of an audiovisual product by artificial intelligence.** 153 p. Thesis (PhD in Communication) – Faculty of Architecture, Arts, Communication and Design, São Paulo State University "Julio de Mesquita Filho", Bauru, 2024.

ABSTRACT

The transformative impact of Artificial Intelligence (AI) in multiple areas of knowledge, including communication, is present in various segments of the media ecosystem, motivating the emergence of new forms of production and consumption. This thesis examines the experience of *Nothing, Forever*, an audiovisual product conceived under this scenario, which presents a structure of infinite temporal production and fruition, being entirely realized by AI in all stages, from scriptwriting to exhibition, using generative artificial intelligence technologies. Initially, the work presents the transition of AI from a predominantly theoretical conception, limited to academic contexts and science fiction narratives, to its comprehensive practical application, driven by technological advances in data processing, cloud computing, and scientific methodologies. In the context of incorporating AI into the realm of communication, it highlights how the technology promotes a paradigmatic reconfiguration in the production and dissemination of information, with a particular focus on generative AI, encompassing the potential of these tools to create content comparable to that generated by humans, as well as its implications in the creation of products such as audiovisuals. A turning point in the historical evolution of the interaction between computing and linguistics was observed in the processing and generation of natural language through *Large Language Models* (LLMs), from the earliest theories about computers' ability to simulate human cognitive processes to the practical implementation of advanced LLM systems, such as *Chat GPT*. By articulating theoretical postulations of Chomsky and Saussure regarding the nature of language and its acquisition, it was identified how these theories relate to the underlying principles of LLM design and functionality. The analysis of *Nothing, Forever* under the theoretical framework of discursive semiotics, notably in the category of continuous *versus* discontinuous, demonstrates a temporal projection that challenges the enunciative praxis of an audiovisual narrative regarding the segmentation of production processes and modes of fruition. This is because the sitcom presents a temporal structure that differs from traditional narrative conceptions. Marked by the entry and exit of viewers at any time during the uninterrupted exhibition of content, it lends itself to a certain form of segmentation into units of meaning from the perspective of discursive semiotics. This segmentation highlights the alternation between continuity versus discontinuity, where moments of continuous engagement are interrupted by pauses or changes in the narrative, enriching the viewer's experience. Given these characteristics of discursivities and the establishment of enunciative contracts, the thesis proposes to denominate such an experience as "Continuous Generative Narrative" to designate audiovisual content generated by artificial intelligence and presented continuously, circumscribing a new way of understanding and categorizing the realization of this type of IA-mediated production.

Keywords: Generative Artificial Inteligence. Semiotics. Continuous Generative Narrative. *Nothing, Forever.*

LISTA DE FIGURAS

Figura 1 - Captura da página da Twitch de exibição de Nothing, Forever 14
Figura 2 - Captura de tela do filme "Sunspring" (2016) 16
Figura 3 - Captura de tela da sitcom Nothing, Forever (2023) 16
Figura 4 - O filho de Leonardo, Gonzalo, junto a Norbert Wiener, usando a máquina "El Ajedrecista" (1922) 21
Figura 5 - Diagrama descritivo do funcionamento de uma GAN 28
Figura 6 - Os três tipos de resolução de problemas de aprendizado de máquina. 33
Figura 7 - Vídeo colorido por um algoritmo de aprendizado de máquina. O filme é Beeld en Geluid" that was shot in 1921 in the picturesq villages of Giethoorn and Hierden, de 1921. 35
Figura 8 - Imagem criada pelo software AARON. Exposta hoje no museu de computação de Boston, EUA 38
Figura 9 - - Imagem gerada pelo software DeepDream 40
Figura 10 - Abordagens de conceito de IA 42
Figura 11- Interface do software ELIZA 52
Figura 12 - Modelo algorítmico da representação de como o conceito de Transformers se aplica a um modelo de LLM 71
Figura 13 - Arquitetura do ChatGPT 72
Figura 14 - Representação cartesiana do eixo sintagma e paradigma 94
Figura 15 - Gráfico gerado pelo próprio ChatGPT para suas escolhas após a frase “O sol brilha no céu” 97
Figura 16 - Gráfico do próprio ChatGPT sobre as escolhas feitas após a frase “O sol está ausente do céu” 97
Figura 17 - Cena de Nothing, Forever 106
Figura 18 - Take da fachada do prédio de Nothing, Forever 110
Figura 19 - Imagem do apartamento de Larry 111
Figura 20 - Diferença entre resoluções de monitores e televisores 118

SUMÁRIO

INTRODUÇÃO 10

1 A INTELIGÊNCIA ARTIFICIAL 21

1.1 Aprendizado de máquina 31

1.2 Categorias de inteligências artificiais 36

1.3 Conceitos de uma inteligência artificial criativa 38

2 OS MODELOS DE LINGUAGEM NATURAL 46

2.1 Algoritmos de Modelos de Linguagem 54

2.2 Os *Large Language Models* (LLM) 56

2.3 O Processamento de Linguagem Natural 58

2.4 A NLP Estatística e os *Large Language Models* modernos 69

3 OS LLM E A GERAÇÃO DE ROTEIROS AUDIOVISUAIS 77

3.1 Produção de Sentido no conteúdo gerado por um LLM 91

4 *NOTHING, FOREVER*: A SITCOM INFINITA 104

4.1 *Nothing, Forever* como um produto audiovisual 105

4.1.1 *Nothing, Forever* sob o viés das artes 112

4.2 A subversão do fluxo 116

4.3 Contratos fiduciários entre humanos e máquinas 125

4.4 A problemática da análise diante da opacidade algorítmica 134

5 – CONSIDERAÇÕES FINAIS 144

REFERÊNCIAS 148

INTRODUÇÃO

Inicialmente delimitada aos domínios de laboratórios acadêmicos ou à esfera da literatura de ficção científica, a inteligência artificial (IA) começou a se destacar em diversificados setores da sociedade a partir do início da segunda década do século XXI. Esta trajetória ascendente é atribuível, em larga escala, ao incremento na disponibilidade e na eficácia do processamento por máquinas, à expansão da computação em nuvem distribuída, bem como ao avanço dos estudos científicos pertinentes ao campo. Numerosos produtos e funcionalidades incorporados à rotina diária recorrem à inteligência artificial, manifestando-se tanto de maneira explícita, em que a utilização é transparente e amplamente divulgada, quanto de forma implícita, onde a aplicação da IA ocorre de maneira não evidente, sem a percepção direta do usuário. Um consumidor tem acesso a aspiradores de pó robóticos, que se valem de recursos de inteligência artificial para mapear o local e conseguir limpá-lo sem esbarrar em obstáculos, instituições bancárias adquirem *softwares* munidos de recursos de IA para análise de crédito, decidindo, baseado no resultado do algoritmo, conceder ou não um empréstimo ou oferecer um produto financeiro específico, baseado em hábitos de consumo filtrados e analisados pela máquina, dentre muitos outros exemplos de produtos ou situações cotidianas em que a inteligência artificial figura como parte essencial:

> Machines, electrification, cars, computers, the Internet and smartphones among other inventions have affected our lives and shaped our work and social environment. The impact of AI technologies can be even more profound than that of both the industrial and digital revolutions put together, as it holds the potential to affect practically all tasks currently performed by humans, diminishing the amount of work left for people and increasing wealth inequalities and their free time. (MAKRIDAKIS, 2017, p. 58)

O uso crescente de recursos de inteligência artificial em produtos e serviços é um fenômeno global, e tem um impacto significativo em várias áreas do conhecimento humano, dentre elas, a comunicação (KAPLAN e HAENLEIN, 2019, p. 15-16). A integração da inteligência artificial no domínio da comunicação está induzindo uma reconfiguração paradigmática tanto na gênese quanto na disseminação de informação, e as implicações e aplicações dessa intersecção abrangem várias áreas desta ciência, como a produção de conteúdo automatizada, personalização e segmentação automática da audiência, análise sentimental e avaliação de *feedback*, monitoramento midiático, análise de tendências, publicidade

direcionada, melhorias na acessibilidade, dentre outras. Gentzkow (2018, p. 6-7) considera que o impacto da inteligência artificial na comunicação é significativo, porém, pondera que, dentre produção e demanda, uso de inteligência artificial na comunicação será mais impactante no lado da demanda, e não do lado da produção:

> There is no question that AI will have profound impacts on media markets. While automation of production may play some role, the unique properties of media goods mean the more important effects are likely to occur on the demand side. Here, there is great potential for social good, as AI can make it easier for consumers to navigate the bewildering mass of online content through search and personalized recommendations, and to identify cases where third parties are attempting to manipulate them. (GENTZKOW, 2018, p. 7)

A inserção de recursos de inteligência artificial na indústria audiovisual aumentou a eficiência de processos, abriu novas possibilidades criativas, contribuiu para a fidelização e engajamento da sua audiência e ampliou o acesso a ferramentas antes restritas a grandes produções, diminuindo a barreira financeira para a produção de conteúdo de áudio e vídeo (DU e HAN, 2021, p. 2). Atualmente, ferramentas de inteligência artificial de custo acessível estão disponíveis para uso doméstico, enriquecendo significativamente o campo da produção de conteúdo. Estas tecnologias oferecem uma diversidade de funcionalidades, como edição e corte de vídeos, automação de colorização, análise de roteiros para identificar elementos de engajamento e antecipar as reações do público, bem como sugerir ajustes para aprimorar a produção. Esses recursos potencializam os criadores de conteúdo, facilitando a democratização e aprimoramento do processo de criação audiovisual.

A aceitação dessas ferramentas de inteligência artificial, que incrementam a eficiência dos processos produtivos dentro da comunidade ligada à indústria, é notavelmente positiva. Estas tecnologias são rapidamente integradas aos fluxos de produção existentes, sendo utilizadas de maneira eficaz. Paralelamente, enquanto essas aplicações práticas da inteligência artificial consolidam sua posição, emerge um novo paradigma no cenário produtivo de conteúdo audiovisual: o advento das inteligências artificiais generativas. Este desenvolvimento representa um avanço significativo, introduzindo capacidades inovadoras de criação e modelagem de conteúdo, que devem influenciar ainda mais os métodos tradicionais de produção.

Lim et al. (2023, p. 1) caracterizam a IA generativa como uma vertente tecnológica capaz de impulsionar modelos de aprendizado de máquina a produzir conteúdos que mimetizam aqueles elaborados por seres humanos, em reação a um leque diversificado e complexo de entradas, abrangendo diferentes tipos, tais como textos, imagens e áudios. Os autores enfatizam a importância de estabelecer uma distinção precisa entre o conceito de "IA Generativa" e termos correlatos, dado que existem outras modalidades de inteligência artificial responsáveis pela geração de conteúdo, porém operando sob mecanismos distintos (ou não tão rigorosamente definidos). A título ilustrativo, mencionam as inteligências artificiais conversacionais, projetadas para simular diálogos humanos, as quais, apesar de engajarem em interações semelhantes às humanas, não geram suas respostas de maneira autônoma:

> It is also important to distinguish Generative AI with related concepts. In line with our definition, it is worth noting that Generative AI has the unique ability to not only provide a response but also generate the content in that response, going beyond the human-like interactions in Conversational AI. In addition, Generative AI can generate new responses beyond its explicit programming, whereas Conversational AI typically relies on predefined responses. However, not all Generative AI is conversational, and not all Conversational AI lacks the ability to generate content. Augmented AI models, such as *ChatGPT*, combine both Generative and Conversational AI to enhance their capabilities. They differ with Generic AI, for example, Scite (2023), which utilizes natural language processing (NLP) and machine learning to determine whether a scientific publication supports, refutes, or mentions a claim related to a specific query, as well as the citation statements reflecting that citing pattern. Notwithstanding these differences, we wish to shine the spotlight on Generative AI in this article in response to the great debate of its use in education that has been sparked by its content-generating capability. (LIM et al., 2023, p. 2)

Produtos oriundos de tecnologias de inteligência artificial generativa, tal como o *ChatGPT*[1], tem a habilidade de engajar em diálogos de uma forma que emula a interação humana. Adicionalmente, essa tecnologia consegue adaptar as conversas ao contexto e basear suas respostas em análises probabilísticas, que são resultado de um amplo treinamento com extensas bases de dados textuais, que excedem *terabytes*[2] em dimensão. Esse processo leva em

[1] *ChatGPT* é um modelo de linguagem desenvolvido pela OpenAI, baseado na arquitetura GPT-4, projetado para entender e gerar texto em linguagem natural. Lançado inicialmente em 2022, o *ChatGPT* é capaz de realizar uma ampla gama de tarefas de processamento de linguagem, incluindo responder perguntas, criar conteúdo escrito, traduzir idiomas e muito mais, com base em um vasto conjunto de dados coletados até abril de 2023. Não possui consciência ou entendimento próprio, operando inteiramente com base nos padrões e informações codificados durante seu treinamento.

[2] Terabytes (TB) são uma unidade de medida de capacidade de armazenamento digital. Um terabyte equivale a aproximadamente 1.000 gigabytes (GB). Esta unidade é comumente usada para descrever a capacidade de armazenamento de dispositivos de grande escala, como discos rígidos, sistemas de armazenamento em rede e

consideração tanto o contexto fornecido quanto o conteúdo gerado pela própria IA. Conhecidos como Grandes Modelos de Linguagem (*Large Language Models, LLM*), estes algoritmos facilitam a produção de textos detalhados, e, dentre eles, roteiros para produções audiovisuais.

Sob uma perspectiva técnica, os textos produzidos por um LLM apresentam semelhanças semânticas com aqueles criados por seres humanos, refletindo a capacidade e a versatilidade dessas ferramentas em simular processos criativos humanos de maneira eficaz. Para avaliar a existência ou não destas semelhanças – e quantizá-las, caso existam – são realizados estudos que geralmente envolvem métricas de avaliação como a ROUGE (*Recall-Oriented Understudy for Gisting Evaluation*) (LIN, 2004, p. 2), entre outras medidas específicas de semelhança semântica. Além disso, pesquisas qualitativas, como avaliações por pares ou análises de conteúdo, são frequentemente utilizadas para entender melhor as nuances da geração de texto por LLMs em comparação com a escrita humana.

A avaliação da semelhança entre textos gerados pelo *ChatGPT* e os feitos por humanos usando a métrica ROUGE envolve comparar o conteúdo dos textos gerados pela máquina com um ou mais textos de referência criados por humanos. A ROUGE é uma família de métricas que mede a qualidade da geração de texto automatizada, como a produção de resumos, em comparação com um conjunto de referências humanas. Essa avaliação é feita principalmente através da contagem de sobreposições de unidades linguísticas, como n-gramas[3], palavras e sequências de *bytes*.

> To assess the effectiveness of ROUGE measures, we compute the correlation between ROUGE assigned summary scores and human assigned summary scores. The intuition is that a good evaluation measure should assign a good score to a good summary and a bad score to a bad summary. (LIN, 2004, p. 5)

Para esta comparação, os textos gerados pelo *ChatGPT* e os textos de referência criados por humanos são coletados e preparados para análise. Isso pode envolver a tokenização[4] do texto em palavras, n-gramas ou outras unidades linguísticas relevantes. Utiliza-se um software

serviços de armazenamento em nuvem, indicando que podem conter uma vasta quantidade de dados, incluindo milhões de arquivos de texto, milhares de fotos de alta resolução, ou centenas de horas de vídeo em alta definição.

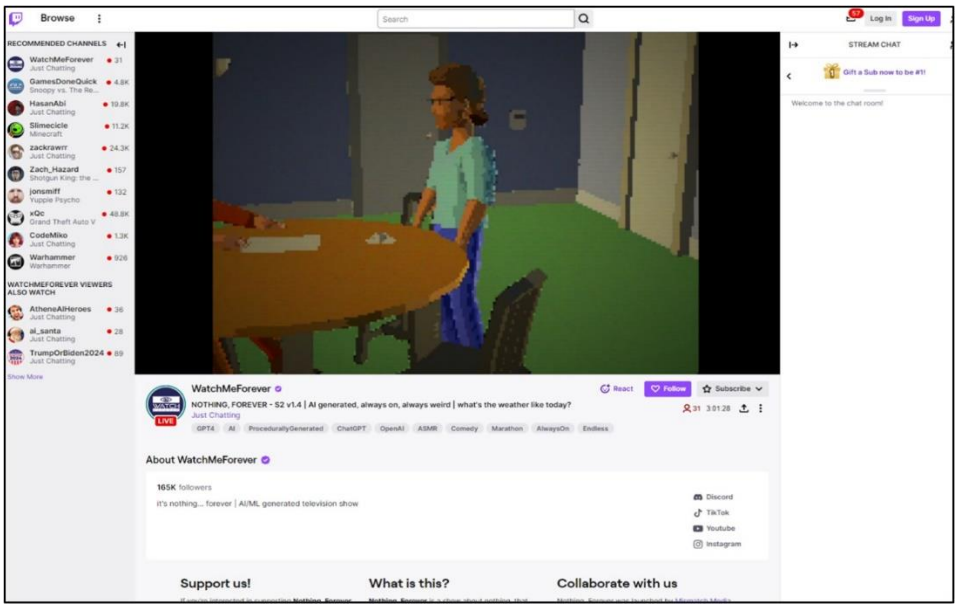
[3] N-gramas são sequências contíguas de n itens (como palavras ou letras) extraídas de um texto, usadas em análises linguísticas e processamento de linguagem natural.

[4] Tokenização é o processo de dividir texto em unidades menores, como palavras ou frases, para análise em processamento de linguagem natural.

ou biblioteca que implementa a métrica ROUGE para calcular a sobreposição entre os textos, comparando as unidades linguísticas dos textos gerados com as dos textos de referência. Os resultados da ROUGE são analisados para determinar a qualidade do texto gerado; isso geralmente envolve olhar para valores de precisão (a proporção de n-gramas no texto gerado que também aparecem no texto de referência), *recall* (a proporção de n-gramas no texto de referência que também aparecem no texto gerado) e a medida F (uma média harmônica de precisão e recall), entre outros possíveis indicadores.

Com a possibilidade de geração de texto por um LLM ser semanticamente semelhante aos textos feitos por humanos, iniciativas de produção de conteúdo que substituam o autor humano por um autor maquínico se tornam viáveis. O *Mismatch Media*, coletivo norte americano, composto por criadores de conteúdo, cientistas de dados, profissionais da área de tecnologia e design, criou a *sitcom*[5] totalmente gerada por inteligência artificial denominada *Nothing, Forever* (2023) (em tradução livre, "Nada, Eternamente"), onde ferramentas eletrônicas são aglutinadas e coordenadas por um LLM, o *ChatGPT* (em sua terceira versão), gerando um produto audiovisual teoricamente infinito, em tempo real.

Figura 1 - Captura da página da Twitch de exibição de Nothing, Forever



Fonte - Disponível em: https://www.twitch.tv/watchmeforever. Acesso em: 4 jan. 2024.

[5] *Sitcom*, abreviação de "*situation comedy*", é um gênero de programa de televisão centrado em um conjunto fixo de personagens que se encontram em situações cômicas. Geralmente ambientadas em locais comuns como lares, locais de trabalho ou bares, as *sitcoms* são caracterizadas por seu humor derivado das interações entre os personagens, situações do dia a dia e mal-entendidos. Este gênero se popularizou na televisão americana e se espalhou globalmente, evoluindo em formato desde suas origens no rádio até as atuais produções televisivas e em streaming.

Nothing, Forever é exibido em fluxo ao vivo, na plataforma de vídeo *Twitch*[6]. O projeto se destaca por utilizar diálogos completamente gerados de forma dinâmica por meio do *ChatGPT*, uma ferramenta de inteligência artificial, para criar interações entre seus personagens. O que torna este produto audiovisual particularmente único é que todo o elenco é composto por personagens virtuais, criados utilizando um motor gráfico[7], o *Unity*, comumente empregado no desenvolvimento de videogames. Este motor gráfico foi modificado para aceitar comandos da inteligência artificial, permitindo que o algoritmo de IA não apenas controle as falas dos personagens, mas também orquestre seus movimentos, os ângulos das câmeras e as transições de cena, através de comandos textuais. Essa integração de tecnologias possibilita que o fluxo de vídeo reflita imediatamente as ações e decisões da inteligência artificial, facilitando a criação de cenas e episódios de forma instantânea e, teoricamente, infinita.

A *sitcom* virtual é inspirada em uma *sitcom* real, exibida na tv americana de 1989 até 1998, "*Seinfeld*". Embora os nomes dos personagens e as localizações dos cenários tenham sido adaptados para esse novo formato, a influência do show original é notável: todos os personagens principais de *Nothing, Forever* tem uma contraparte equivalente em *Seinfeld*, a ambientação cênica das duas séries é semelhante, bem como a sua temática.

Essa combinação de um algoritmo de inteligência artificial com a criação de personagens e cenários virtuais abre novas fronteiras para a narrativa digital, desafiando as noções tradicionais de produção televisiva e entretenimento ao vivo. Contudo, a iniciativa de roteirizar conteúdo audiovisual através de uma IA generativa não é inédita: *Sunspring*, um curta metragem de ficção científica totalmente roteirizado por inteligência artificial, realizado em 2016 em um experimento na Universidade de Nova York. O algoritmo “Benjamin”, criado pelo pesquisador Ross Godwin, foi responsável por gerar o roteiro, depois de treinado com centenas de outros roteiros escritos por humanos. *Sunspring*, embora pioneiro, difere em diversos aspectos de *Nothing, Forever*: enquanto o roteiro do curta metragem foi gerado todo de uma vez para depois ser interpretado por atores reais em uma filmagem onde os humanos envolvidos seguiam a risca as regras ditas pelo algoritmo gerador, a *sitcom* do grupo *Mismatch Media* é

[6] *Twitch* é um serviço de streaming ao vivo que se concentra principalmente em videogames, incluindo transmissões de jogos ao vivo, competições de esportes eletrônicos, conteúdo criativo relacionado a jogos, *streams* de música, conteúdo de estilo de vida e *talk shows*.

[7] Um Motor Gráfico é um software composto de pacotes de funcionalidades que facilitam o desenvolvimento de aplicações que exibam gráficos em um computador. O motor gráfico abstrai certos conceitos de programação e possibilita uma rapidez e eficácia maior na construção de cenários, personagens e efeitos visuais e sonoros. É comumente usada na criação de vídeo games, mas recentemente vem ganhando espaço na indústria audiovisual.

gerada proceduralmente, com os personagens virtualizados em três dimensões, e os diálogos são baseados nos próprios diálogos anteriores que o LLM gerou, em um fluxo contínuo.

Figura 2 - Captura de tela do filme "Sunspring" (2016)

Fonte: Disponível em: https://www.repubblica.it/tecnologia/2016/06/10/news/sunspring_il_film_sci-fi_scritto_dall_intelligenza_artificiale_altro_che_romanzi_social_e_opere_le_reti_neurali_immaginan-141735805 Acesso em 4 jan. 2024

Figura 3 - Captura de tela da sitcom Nothing, Forever (2023)

Fonte: Captura própria.

O sucesso desses experimentos mostra que é possível para uma IA generativa criar roteiros audiovisuais que humanos podem entender e produzir. No entanto, para as ciências da comunicação, é importante explorar mais sobre esse processo criativo. Isso inclui entender

como uma ferramenta de LLM generalista, como o *ChatGPT*, pode replicar os complexos mecanismos de criação de um roteiro completo para um produto audiovisual e, crucialmente, como um LLM gera o significado que tal obra artificial carrega.

Para observar como se dá a produção de sentido dentro de uma obra audiovisual de características inovadoras como é *Nothing, Forever*, é pertinente analisar a gênese da inteligência artificial e sua fundamentação científica, sob o arcabouço teórico das teorias de gramática gerativa de Noam Chomsky, e na linguística de Saussure, no início do século XX. Os LLM modernos começam a ganhar tração no início do século seguinte (XXI), e, para fundamentar a análise da produção de sentido nestes modelos, a literatura produzida pelos próprios criadores dos modelos é pertinente, já que todos os passos na evolução dos produtos de inteligência artificial generativa possuem publicações abertas, citadas nesta tese, quando pertinentes. Como este é um trabalho sobre comunicação, todo o direcionamento da análise é focado na absorção do conteúdo por esta área acadêmica, embora seja inevitável a citação de trabalhos com conceitos computacionais complexos e teorias matemáticas aplicadas, como o conceito de *transformers* aplicado a IA generativa.

Nothing, Forever é uma obra audiovisual gerada integralmente por inteligência artificial, diferenciando-se das produções convencionais pela ausência de intervenção humana direta na criação de conteúdo. A série automatiza a produção de cenários e personagens, permitindo a execução em tempo real, introduzindo uma nova forma de interação entre a obra e o espectador, causada pela geração contínua e adaptativa de conteúdo. A simplificação visual é uma escolha técnica para acomodar a necessidade de processamento em tempo real, o que, por sua vez, influencia a estética da produção. A estrutura narrativa não é pré-definida, dependendo de um conjunto de parâmetros básicos para o desenvolvimento de roteiros dinâmicos, o que propicia uma narrativa evolutiva e reativa às interações dos espectadores.

A partir de uma perspectiva crítica, *Nothing, Forever* questiona as noções tradicionais de autoria e criatividade, operando sem um roteiro fixo e incorporando entradas do usuário em tempo real para potencialmente modificar a trama. Este modelo reflete sobre o papel das tecnologias intermediárias na arte e na cultura, desafiando conceitos estabelecidos sobre a produção e recepção de obras audiovisuais. A natureza experimental da *sitcom*, juntamente com sua capacidade de gerar conteúdo de maneira autônoma e interagir com o público, ilustra uma interseção inovadora entre arte, tecnologia e entretenimento. Este estudo de caso sugere uma expansão nas definições tradicionais de obras audiovisuais, enfatizando o potencial de novas tecnologias para transformar a experiência cultural e artística.

A inovação central de *Nothing, Forever* reside na subversão do modelo tradicional de consumo de conteúdo audiovisual, introduzindo uma experiência dinâmica e ininterrupta que contrasta com a estrutura finita habitual de início, meio e fim das obras convencionais. Esta sitcom gerada por IA gera um fluxo narrativo contínuo e sem término predeterminado, desafiando a percepção tradicional de uma obra como um conjunto fechado e reiterativo de conteúdo. A natureza perpétua da narrativa, que evolui e se expande incessantemente, reflete o potencial autossustentável da IA que a produz, redefinindo as interações entre o conteúdo audiovisual e seu público. Diferentemente das obras audiovisuais clássicas, que são ancoradas em um espaço-tempo definido, *Nothing, Forever* apresenta uma trama que se desenvolve sem cessar, impossibilitando aos espectadores a experiência de assistir ao mesmo conteúdo repetidamente. A cada retorno, o espectador se depara com um segmento distinto da narrativa, tendo perdido os eventos ocorridos durante sua ausência. Esta abordagem desafia não apenas as noções convencionais de autoria e narrativa, mas também amplia as possibilidades de engajamento do espectador com a obra, oferecendo uma experiência imersiva e única a cada acesso.

Nothing, Forever e suas características podem ser analisadas sob o arcabouço teórico da semiótica discursiva, especialmente em relação à dinâmica entre continuidade e descontinuidade presentes nas obras de Jean-Marie Floch (2023) e Eric Landowski (2014). A experiência dos espectadores com o fluxo temporal da *sitcom* pode ser concebida como um texto semiótico, com suas próprias estruturas e significados. A noção de percurso, tanto físico quanto narrativo, oferece uma base para entender a interação contínua e variável entre os espectadores e a obra, bem como entre os diferentes elementos que compõem o universo narrativo de *Nothing, Forever*.

A obra também ilustra os regimes de interação identificados por Landowski (2014), como programação, manipulação, ajustamento e acidente, cada um refletindo diferentes aspectos da criação e do consumo de conteúdo audiovisual na era digital. A capacidade da IA de gerar conteúdo de forma autônoma e imprevisível introduz uma nova dimensão de acidentalidade e contingência na produção artística, desafiando as expectativas tradicionais e abrindo caminho para novas formas de expressão e experiência.

Nothing, Forever como *sitcom* de fluxo contínuo desafia a concepção tradicional de obras finitas, propondo uma narrativa que se desenvolve eternamente, sem reprises. Embora a possibilidade de assistir a *Nothing, Forever* ininterruptamente seja teoricamente infinita, a prática de consumo audiovisual dos espectadores limita o tempo dedicado a uma única

produção, sugerindo que, apesar da inovação, a experiência de visualização de um fragmento específico assemelha-se à de assistir a um capítulo convencional de uma novela, com cortes de câmera e padrões temporais fluidos similares: a abordagem enunciativa adotada por *Nothing, Forever*, gerada por IA, e as produções audiovisuais tclássicas convergem quando analisadas em fragmentos temporais não lineares específicos. Através das escolhas sintáticas, espaciais, actanciais e temporais, tanto a IA quanto as produções humanas criam discursos com estruturas narrativas comparáveis. A semiótica Greimasiana aponta que a intersecção dos planos de conteúdo e expressão é fundamental para a análise textual, indicando que, em níveis de análise segmentados, *Nothing, Forever* e as obras convencionais podem ser analisadas sob os mesmos parâmetros. Entretanto, a natureza contínua de *Nothing, Forever* distingue-se substancialmente das produções clássicas em seu todo, já que a ausência do espectador não interrompe a narrativa, que prossegue em desenvolvimento, contrastando com a possibilidade de retomar capítulos perdidos em produções tradicionais.

A geração de conteúdo em *Nothing, Forever* ilustra a relação dialética entre a massiva produção textual humana e a seleção e filtragem desses textos por um agente humano para treinamento da IA. A obra também desafia a concepção tradicional de análise audiovisual ao requerer uma compreensão do processo de geração de IA como parte integrante da sua natureza criativa. Diferentemente das produções humanas, onde é possível explorar as motivações dos criadores, a análise de *Nothing, Forever* se torna mais efetiva através uma investigação do código e dos mecanismos de geração de conteúdo.

Nothing, Forever não apenas redefine as possibilidades narrativas no campo audiovisual, mas também levanta questões sobre autoria, memória narrativa e a evolução das estruturas discursivas em contextos gerados por IA. A continuidade da inovação tecnológica promete minimizar as limitações atuais, ampliando o potencial criativo das aplicações de inteligência artificial em produções futuras e desafiando continuamente as fronteiras entre a criação humana e a máquina. Através da lente da semiótica discursiva Greimasiana, que postula a inseparabilidade entre o plano de conteúdo e o plano de expressão na geração do texto, é possível compreender a complexidade de *Nothing, Forever*. Esta perspectiva permite uma análise imanente do produto audiovisual, revelando que, em segmentos específicos, as produções de IA e humanas podem ser analisadas sob os mesmos critérios. No entanto, a série distingue-se no plano da expressão, especialmente na sua construção espaço-temporal que prossegue independentemente da presença do espectador, um desvio marcante das convenções tradicionais de produção audiovisual.

O papel do *ChatGPT* como enunciador em *Nothing, Forever* destaca-se como uma confluência de múltiplas vozes da enunciação, que, através da aplicação de padrões matemáticos aprendidos, simula a humanidade encontrada em textos escritos por seres humanos. Esse processo enfatiza uma relação dialética entre a coleta e a geração de texto: inicialmente, a vasta produção textual humana serve como base para o treinamento da IA, que posteriormente tenta estabelecer um novo contrato de figuratividade com o espectador, fundamentando-se nos padrões processados. Esta relação é complexificada pela capacidade da IA de atuar como intermediário entre o enunciador humano original e o espectador, desafiando noções tradicionais de autoria e originalidade. A semiótica discursiva, ao afirmar a autonomia do texto e enfocar na análise imanente, fornece um quadro robusto para a interpretação tanto da *sitcom* quanto da geração de conteúdo por IA como um todo: *Nothing, Forever* não apenas questiona as convenções estabelecidas na criação e narrativa audiovisual, mas também oferece um campo fértil para a aplicação da semiótica discursiva, ilustrando a intersecção entre tecnologia de ponta e análise textual profunda.

A transparência no processo produtivo dos algoritmos que geram estes textos é importante para um entendimento completo do fenômeno da roteirização automatizada. Contudo, a prática comum de proteger algoritmos por motivos de segredo industrial impede uma investigação aprofundada, limitando o conhecimento sobre as decisões internas que guiam a geração de conteúdo. Ao contrário do processo criativo humano, cujas influências e metodologias podem ser parcialmente compreendidas através de estudos contextuais e teóricos, a "caixa preta" dos algoritmos de IA apresenta um desafio significativo. A opacidade algorítmica não apenas oculta os processos internos, mas também complica a identificação e correção de vieses potencialmente nocivos, refletidos nos produtos gerados.

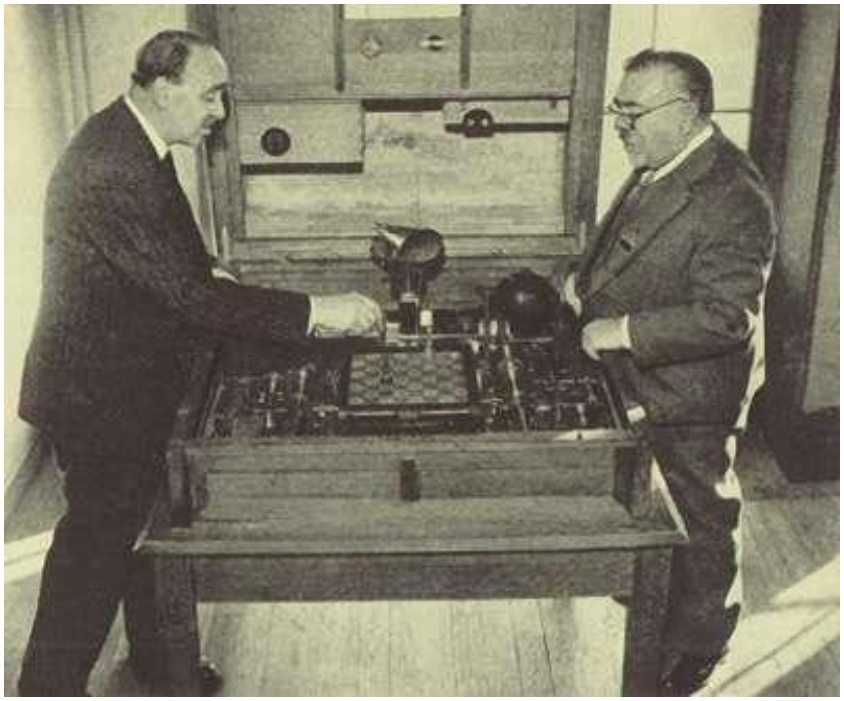
A necessidade de uma abordagem mais transparente na IA, um movimento em direção à "caixa de vidro", sugere um futuro em que a explicabilidade e a responsabilização dos algoritmos possam coexistir com os avanços tecnológicos. Tal abertura não só facilitaria a análise acadêmica de produtos como *Nothing, Forever*, mas também contribuiria para um desenvolvimento mais ético e socialmente responsável da inteligência artificial. A jornada em direção à transparência algorítmica é complexa e cheia de desafios comerciais, tecnológicos e éticos, mas é um passo necessário para alinhar os avanços da IA com os valores e expectativas da sociedade.

1 A INTELIGÊNCIA ARTIFICIAL

Inteligência artificial, abreviada pelo acrônimo “IA”, é um termo guarda-chuva que abrange uma diversidade de tecnologias que trabalham juntas e permitem que uma máquina busque reproduzir a cognição humana - podendo até excedê-la - na produção de um ato que se aproxime do que um ser humano considera como “inteligente”. Difere, portanto, do conceito de computação tradicional, que resolve problemas específicos através da programação de uma máquina, já que um algoritmo que não use nenhuma técnica ou conceito de IA não se aproveita dos dados que obtém durante a sua execução para, posteriormente, melhorar a sua eficiência, além de não inferir nenhum conceito ou ação que não tenha sido explicitamente indicado pelo seu criador.

O conceito de inteligência artificial (IA) permeia amplamente tanto a literatura de ficção quanto a realidade de dispositivos mecânicos com capacidades autônomas limitadas. Um exemplo notável é "*El Ajedrecista*", uma máquina automatizada capaz de jogar xadrez, desenvolvida por Leonardo Torres y Quevedo (1915, p. 296), que demonstra uma aplicação prática da IA desde os estágios iniciais de sua concepção.

Figura 4 - O filho de Leonardo, Gonzalo, junto a Norbert Wiener, usando a máquina "El Ajedrecista" (1922)



Fonte: https://www.chessprogramming.org/El_Ajedrecista Acesso em 24 fev. 2024

No entanto, o debate sobre a inteligência artificial ganhou destaque apenas com o advento dos primeiros computadores na Segunda Guerra Mundial. Foi somente uma década após o término do conflito que o termo "Inteligência Artificial" foi introduzido por John McCarthy, que, em 1955, empregou esta nomenclatura ao estabelecer um grupo de pesquisa dedicado ao estudo desse campo, marcando o início de uma era dedicada à exploração e desenvolvimento da IA:

> We propose that a 2 month, 10 man study of artificial intelligence be carried out during the summer of 1956 at Dartmouth College in Hanover, New Hampshire. The study is to proceed on the basis of the conjecture that every aspect of learning or any other feature of intelligence can in principle be so precisely described that a machine can be made to simulate it. An attempt will be made to find how to make machines use language, form abstractions and concepts, solve kinds of problems now reserved for humans, and improve themselves. (MCCARTHY, MINSKY, *et al.*, 1955, p. 1)

Embora o termo cunhado por McCarthy seja amplamente conhecido e utilizado, a sua definição ainda se encontra sobre escrutínio da comunidade científica. Segundo Bartneck et al. (2021), a definição sobre o que é ou o que não é inteligência artificial muda conforme o tempo:

> The definition of AI and of what should and should not be included has changed over time. Experts in the field joke that AI is everything that computers cannot currently do. Although facetious on the surface, there is a sense that developing intelligent computers and robots means creating something that does not exist today. Artificial intelligence is a moving target. Indeed, even the definition of AI itself is volatile and has changed over time. (BARTNECK, LÜTGE, *et al.*, 2021, p. 7-8)

Dada a dificuldade exposta pelo estado volátil desta definição, é importante identificar pontos em comum dentre os diversos teóricos, para que seja viável uma busca por um relativo consenso, dentre as múltiplas linhas de estudo.

Poole e Mackworth (2010) definem a inteligência artificial como uma área da ciência que estuda a criação e a análise de agentes computacionais que agem de forma inteligente, e definem “agentes” da forma concreta, como algo que age (seja humano, ou não). A inteligência destes agentes, segundo os dois, é dada quando quatro condições são alcançadas:

- Quando as ações destes agentes são apropriadas para os seus objetivos, de acordo com as circunstâncias da situação.
- Quando este agente tem flexibilidade para mudar o ambiente da ação e (ou) mudar seus objetivos.
- Quando este agente consegue aprender com as suas experiências.
- Quando este agente toma decisões de acordo com as suas percepções, levando em conta as suas limitações.

Essa linha de pensamento sobre IA é compartilhada por Russel e Norvig (1995):

> The main unifying theme is the idea of an intelligent agent. We define AI as the study of agents that receive percepts from the environment and perform actions. Each such agent implements a function that maps percept sequences to actions, and we cover different ways to represent these functions, such as reactive agents, real-time planners, and decision-theoretic systems. We explain the role of learning as extending the reach of the designer into unknown environments, and we show how that role constrains agent design, favoring explicit knowledge representation and reasoning. We treat robotics and vision not as independently defined problems, but as occurring in the service of achieving goals. We stress the importance of the task environment in determining the appropriate agent design. (RUSSEL e NORVIG, 1995, p. 8)

Kaplan e Haenlein (2019) definem a Inteligência Artificial como a capacidade de um agente de interpretar dados externos corretamente, aprender com estes dados, e usar estes conhecimentos para realizar tarefas específicas, com certa flexibilidade para se adaptar as circunstâncias e ambiente. Embora existam pontos em comum entre estas linhas de pensamento, discutir o que é inteligência artificial avança sobre áreas de estudo científico distintas, como a filosofia, a biologia e a engenharia, e, por esta diversidade, alcançar uma definição para o termo "inteligência artificial" que abranja todos os ramos científicos envolvidos é um desafio:

> But what exactly is AI? Philosophers arguably know better than anyone that precisely defining a particular discipline to the satisfaction of all relevant parties (including those working in the discipline itself) can be acutely challenging. Philosophers of science certainly have proposed credible accounts of what constitutes at least the general shape and texture of a given field of science and/or engineering, but what exactly is the agreed-upon definition of physics? What about biology? What, for that matter, is philosophy, exactly? These are remarkably difficult, maybe even eternally unanswerable, questions, especially if the target is a consensus definition. (BRINGSJORD e GOVINDARAJULU, 2020, p. 6)

O estudo da inteligência artificial é multidisciplinar em sua essência, amalgamando conceitos e metodologias oriundas de uma diversidade de eixos acadêmicos. Russel e Norvig (1995, p. 5-10) citam a filosofia, a matemática, as ciências econômicas, a neurociência, a psicologia, as ciências da computação, a cibernética e a linguística como exemplos dessa multidisciplinaridade. Esta natureza polimática da IA é característica do *ethos* contemporâneo da pesquisa científica, onde a sinergia entre campos distintos alavanca resultados inovadores. É imperativo destacar que a inteligência artificial não é uma área restrita à computação, embora esta seja sua fundação primordial; as ciências computacionais fornecem a infraestrutura algorítmica e as arquiteturas de sistemas que compõem a espinha dorsal dos modelos de IA, porém, a área é profundamente enraizada nas ciências cognitivas e na psicologia, especialmente na modelagem computacional da cognição humana.

Estudos em percepção, aprendizado, memória e tomada de decisão baseiam o desenvolvimento de algoritmos que se inspiram nas faculdades mentais humanas. A psicologia cognitiva, em particular, fornece dados cruciais sobre os processos de aprendizagem e raciocínio, que são fundamentais para a IA (RUSSEL e NORVIG, 1995, p. 13). A filosofia também desempenha um papel significativo, especialmente no que diz respeito à ética dos algoritmos de inteligência artificial. Questões filosóficas sobre a consciência, identidade e a moralidade são intrínsecas ao debate sobre a criação e utilização de entidades inteligentes artificiais. Matemática e estatística também são fundamentais para a IA, oferecendo o campo teórico para os algoritmos de aprendizado de máquina e a análise de dados, fornecendo ferramentas para a construção e a validação dos modelos e para a interpretação quantitativa dos seus resultados. Algoritmos de processamento de linguagem natural se beneficiam muito dos estudos de linguística, que auxiliam os cientistas e programadores de inteligências artificiais a entender melhor o processamento da linguagem e a modelagem dos sistemas cognitivos. A colaboração entre a linguística e a IA abre caminhos para avanços significativos na forma como as máquinas interpretam, processam e respondem à linguagem natural, possibilitando interações mais intuitivas entre humanos e máquinas.

Contudo, essa intersecção de áreas também causa conflitos. Como a inteligência artificial é um assunto que ganhou tração em discussões fora do mundo acadêmico, dada a magnitude da ruptura tecnológica que a sociedade vive com a introdução da temática, a discussão passa por um escrutínio ainda maior pela academia. Allen (1998) fala sobre este problema em um tom incisivo de reprova:

> AI has always been a strange field. Where else could you find a field where people with no technical background feel completely comfortable making claims about viability and progress? We see articles in the popular press and books regularly appear telling us that AI is impossible, although it is not clear what the authors of these publications mean by that claim. Other sources tell us that AI is just around the corner or that it's already with us. Unlike fields such as biology or physics, apparently you don't need any technical expertise in order to evaluate what's going on in this field. (ALLEN, 1998, p. 1)

Logo, é pertinente levar em conta que, embora a inteligência artificial seja uma temática de alta afinidade analítica e matemática, afeta uma conjunção enorme de outras áreas e amplia o debate do tema para círculos que se afastam das ciências exatas, como a área da comunicação, onde esta tese se debruça. Segundo Bringsjord e Govindarajulu (2020), o mais prudente a se propor para uma discussão abrangente do termo, tentando buscar um relativo consenso mas considerando estas peculiaridades da temática, é aglutinar contribuições, sejam de produtos ou ideias, que possam contribuir para um entendimento melhor que é inteligência artificial, incluindo dentre estas ideias as recentes tentativas de qualificá-la, fazendo-o de forma detalhada e rigorosa, em detrimento a tentar definir o tema de forma indiscutível.

Segundo Wang (2008, p. 2), embora haja uma dissonância entre autores sobre a definição do conceito de inteligência artificial, existe um relativo consenso sobre o que o termo representa: seres humanos diferenciam-se dos animais (e das máquinas) de modo imperativo devido a suas faculdades mentais, coisa que é normalmente denominada "inteligência". A IA, neste caso, é uma tentativa de reproduzir esta habilidade em uma máquina.

Wang destaca uma perspectiva crítica sobre os paradigmas atuais na avaliação da inteligência artificial (IA), argumentando que a utilização de critérios exclusivamente humanos para definir e mensurar a inteligência das máquinas é uma abordagem restritiva. Essa visão sugere que, ao calibrar a inteligência artificial com base nos padrões e capacidades humanas, inadvertidamente limitamos o potencial e a amplitude que a IA pode alcançar. Se a meta para uma IA "ideal" é equiparar-se ao nível de inteligência humana, isso implica que estamos confinando as possibilidades de desenvolvimento da IA dentro de um espectro já conhecido e, possivelmente, limitante. Em todos estes casos, o consenso para realizar a distinção entre se sistema computacional é dotado de IA ou não é o próprio conceito de inteligência. É importante, portanto, antes de tentar definir o que é inteligência artificial, apontar qual dos diversos vieses da própria definição do que é a inteligência, sob o ponto de vista da construção de máquinas para este fim.

Alan Turing (1950) fala sobre esta questão no seu texto "*Computing Machinery and Intelligence*", onde disserta sobre como a discussão sobre os próprios termos é importante:

> I propose to consider the question, "Can machines think?" This should begin with definitions of the meaning of the terms "machine" and "think." The definitions might be framed so as to reflect so far as possible the normal use of the words, but this attitude is dangerous, If the meaning of the words "machine" and "think" are to be found by examining how they are commonly used it is difficult to escape the conclusion that the meaning and the answer to the question, "Can machines think?" is to be sought in a statistical survey such as a Gallup poll. (TURING, 1950, p. 1)

Turing encaminha a discussão propondo a avaliação dos significados do termos seminais "máquina" e "pensamento", antes de avaliar se máquinas pensam, elaborando que o escopo de ambas é amplo demais para tornar a resposta à pergunta "Máquinas pensam?" subjetiva e sujeita ao viés de quem a responde, tamanha a abstração que os significados representam. Logo, considerando o raciocínio de Turing, é necessário definir o alicerce temático antes de se definir o termo derivado, sob pena da definição de inteligência artificial ser opinativa, e não científica.

Tal qual o termo "inteligência artificial", a própria noção de inteligência e pensamento humano não é um termo definido de forma unânime. Fundamentos teóricos deste conceito datam da Grécia antiga, com raízes filosóficas advindas de pensadores como Aristóteles e Platão, cujas teorias ajudam-nos a segmentar e fundamentar os alicerces sobre as definições de inteligência. Platão é o autor da "Teoria das Ideias", que postula a existência de uma dimensão transcendental de formas (ideias), no caso, abstrações que representam uma essência verdadeira de todas as coisas, conceitos eternos e imutáveis, e que existiam apenas em um "reino" que poderia ser acessado somente pela nossa mente. Para Platão (1956, p. 287-291), a inteligência consiste na capacidade de abstrair-se deste mundo concreto e aprender suas formas abstratas; este conceito está representado na famosa alegoria da "Caverna de Platão", onde prisioneiros presos dentro de uma caverna só podem observar sombras projetadas em uma parede, que são consideradas realidade para eles. Quando um prisioneiro sai da caverna e identifica que as sombras não são a realidade, mas sim uma representação dela, segundo as ideias de Platão, simboliza o processo da obtenção da verdadeira inteligência.

Aristóteles (2002), discípulo de Platão, acreditava que, para entender qualquer fenômeno, era necessário, primeiramente, compreender quatro motivações: qual a sua

substância constituinte (a causa material), qual a sua estrutura e configuração (a causa formal), qual o seu processo gerador (a causa eficiente) e qual o seu propósito ou função intrínseca (a causa final). A inteligência humana é fundamentada na biologia do cérebro humano: os neurônios, as sinapses e os mecanismos bioquímicos que formam a cognição são as bases materiais de nossa inteligência; a forma da inteligência humana está em sua organização neural e nos padrões de processamento do nosso cérebro, que integra informações, realiza abstrações, processa a lógica e cria, com isso, pensamentos. Os fatores que moldam e desenvolvem a inteligência humana, como a educação recebida, as experiências de vida, as interações sociais e as influências culturais, podem ser consideradas a causa eficiente, e, para os humanos, a causa final é a capacidade de compreender, adaptar e manipular o seu ambiente.

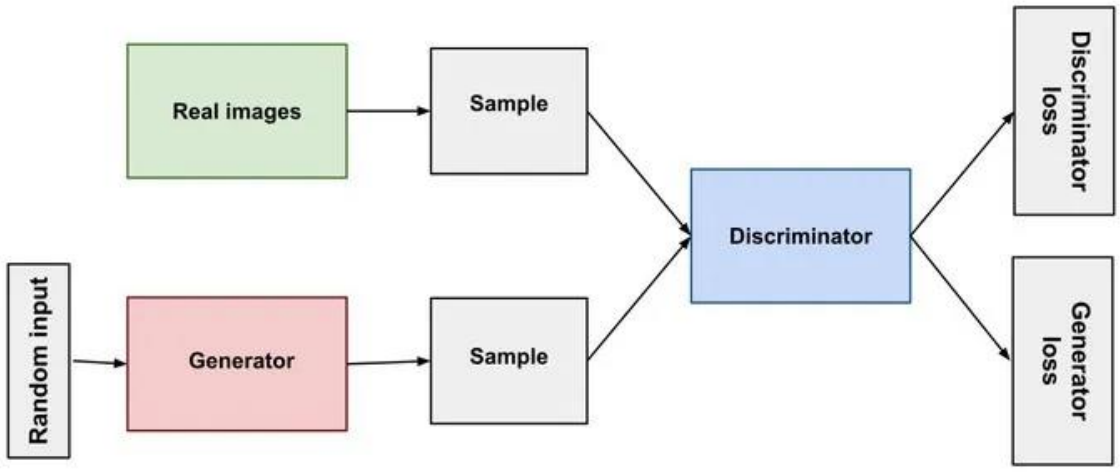
Aplicar as teorias de Aristóteles e Platão ao campo da inteligência artificial evidencia as diferenças entre os humanos e as máquinas. É possível considerar, em analogias simples, os processadores, as memórias RAM, os dispositivos de armazenamento de memória e toda a parafernália pertencente a computação como "causa material", os algoritmos e suas estruturas de processamento como "causa formal", os dados de treinamento e de *feedback* como "causa eficiente", mas a ausência de uma causa final que não esteja sempre submetida ao desejo humano é um dos pontos nevrálgicos da discussão sobre como enxergar o termo "inteligência" no campo da inteligência artificial.

Outras teorias surgiram dessa semente. O Behaviorismo, que emergiu no início do século XX e tem como principais nomes Skinner e Watson, é uma corrente teórica no âmbito da psicologia que se caracteriza pela sua ênfase em comportamento passíveis de observação e mensuração objetiva, abdicando da consideração de fenômenos intrapsíquicos como pensamentos e emoções. Skinner é o autor da teoria do condicionamento operante, que postula que o comportamento humano é moldado por suas consequências:

> O condicionamento operante é um segundo tipo de seleção por consequências. Deve ter evoluído em paralelo a dois outros produtos das mesmas contingências de seleção natural – a susceptibilidade ao reforçamento por certos tipos de consequências e um conjunto de comportamentos menos especificamente relacionados a estímulos eliciadores ou liberadores. [...] Quando as consequências selecionadoras são as mesmas, o condicionamento operante e a seleção natural trabalham conjuntamente, de maneira redundante. Por exemplo, o comportamento de um pato recém-nascido ao seguir sua mãe é aparentemente não apenas o produto de seleção natural (patos recém-nascidos tendem a se mover em direção a objetos grandes em movimento), mas também de uma evoluída susceptibilidade ao reforçamento pela proximidade a esse objeto. A consequência comum é que o pato permanece próximo de sua mãe. (SKINNER, 1981, p. 501)

Segundo Skinner (1981), comportamentos reforçados (seja por reforços positivos ou negativos) tendem a ser fortalecidos e mais frequentes, ao passo que comportamentos sujeitos a punições exibem uma tendência à redução de sua ocorrência. Essa teoria é aplicada em alguns tipos e algoritmos de inteligência artificial, como os GAN (*Generative Adversarial Network*).

Figura 5 - Diagrama descritivo do funcionamento de uma GAN



Fonte - Disponível em: https://developers.google.com/machine-learning/gan/gan_structure
Acesso em: 4 jan. 2024.

Um algoritmo GAN é composto de duas partes: o gerador, que tenta criar dados, e o discriminador, que avalia os dados criados pelo gerador. O objetivo de ambas as partes é "derrotar" a outra, através de um sistema de pontuação interna, e tanto o gerador quanto o discriminador aprendem e se aperfeiçoam no processo. A ideia é que, através de um processo iterativo de competição, o gerador aprenda a produzir dados mais próximos do objetivo inicial dado pelo programador, e o discriminador se torne melhor em validar os dados gerados. No GAN, o discriminador fornece dados de *feedback* ao gerador; quando o gerador consegue produzir dados que o discriminador não consegue identificar como inválidos, ele recebe um reforço positivo, e, quando o discriminador é capaz de avaliar os dados como falsos ou insatisfatórios, o gerador recebe um reforço negativo.

Skinner tende a considerar o pensamento não como uma entidade separada ou um processo mental abstrato, mas como uma forma de comportamento. Especificamente, é visto como um comportamento verbal ou subvocal interno – uma espécie de diálogo interno, ou "fala interna". Vygotsky (1991) considera que a fala interna emerge como um desenvolvimento

crucial no arcabouço da linguagem e dos pensamentos humanos. Esta forma de comunicação, inicialmente manifestada como fala externa (ou social), dirigida a outrem, sofre uma transição para um diálogo interno, que, para todos os efeitos, é um pensamento.

> O primeiro exemplo significativo dessa ligação entre essas duas funções da linguagem é o que ocorre quando as crianças descobrem que são incapazes de resolver um problema por si mesmas. Dirigem-se então a um adulto e, verbalmente, descrevem o método que sozinhas não foram capazes de colocar em ação. A maior mudança na capacidade das crianças para usar a linguagem como um instrumento para a solução de problemas acontece um pouco mais tarde no seu desenvolvimento, no momento em que a fala socializada (que foi previamente utilizada para dirigir-se ao adulto) é internalizada. Ao invés de apelar para o adulto, as crianças passam a apelar a si mesmas; a linguagem passa, assim, a adquirir uma função intrapessoal além do seu uso interpessoal. No momento em que as crianças desenvolvem um método de comportamento para guiarem a si mesmas, o qual tinha sido usado previamente em relação a outra pessoa, quando elas organizam sua própria atividade de acordo com uma forma social de comportamento, conseguem, com sucesso, impor a si mesmas uma atitude social. (VYGOTSKY, 1991, p. 23)

Vygotsky postula que a fala interna é derivada de interações sociais, onde, os humanos, por meio do engajamento com outros humanos, internalizam as modalidades linguísticas e de fala características de seu entorno cultural. Logo, o pensamento é um diálogo interno que se utiliza da linguagem, que transcende a sua função comunicativa, assumindo um papel crucial como instrumento do pensamento.

> Nossa pesquisa mostrou que, mesmo nos estágios mais precoces do desenvolvimento, linguagem e percepção estão ligadas. Na solução de problemas não verbais, mesmo que o problema seja resolvido sem a emissão de nenhum som, a linguagem tem um papel no resultado. Esses achados substanciam a tese da linguística psicológica, que defendia a inevitável interdependência entre o pensamento humano e a linguagem. (VYGOTSKY, 1991, p. 26)

Cada uma destas teorias apresenta uma visão distinta do mecanismo do pensamento humano, embora apresentem semelhanças fundamentais que as mantém, de certo modo, coesas. Skinner apresenta o ato de pensar como um modo genérico de processamento de informações baseado na capacidade cognitiva do ser, já Vygotsky aprofunda e fundamenta o ato de pensar em um termo amplo que envolve a racionalização, a categorização e a tomada de decisão por parte do ser inteligente. Holyoak e Spellman falam sobre estas diferenças:

> Sometimes thinking is construed as a synonym for all "inteligent information processing" and sometimes it is construed as the umbrella term for a range of processes associated with "high-level" cognition, such as reasoning, categorization, and judgment and decision making. (HOLYOAK e SPELLMAN, 1993, p. 266)

Estas teorias divergem em pontos chave como a necessidade intrínseca da linguagem na construção da inteligência, ou na necessidade do objeto, ou apenas da sua representação simbólica, mas convergem em um ponto no qual um pensamento é construído por um processo cognitivo consciente e independente de qualquer estímulo dos sentidos, embora possa apropriar-se destes, diferentemente da percepção, que ocorre sempre através destes.

No mesmo ano (1951) das considerações sobre inteligência maquínica de Turing, Jean Piaget discutia a temática da inteligência sob o viés cognitivo, da aprendizagem. Em entrevista a uma rádio francesa, Piaget cita diferenças entre a inteligência e a percepção, enfatizando a dificuldade que é definir o que é inteligência em um universo das funções cognitivas, ressaltando que diversos autores possuem opiniões conflitantes sobre a própria definição do que é ou não é inteligência.

> Podemos dizer que a afetividade constitui a dinâmica da conduta (...) enquanto a inteligência, ou antes o conhecimento num sentido mais amplo, seria a estrutura da conduta, i. e., o conjunto das relações entre o sujeito e os objetos da conduta. Mas todo conhecimento não é inteligência. Há estruturas perceptivas, há imagens, há adaptações motoras etc. Como, então, definir a inteligência no bojo das funções cognitivas? Isso é muito difícil. Por exemplo, Édouard Claparède procurou definir a inteligência por ensaios de tentativa e erro. (...) Mas, para outros autores, o ensaio por tentativa e erro, é, ao contrário, a exclusão da inteligência. Por exemplo, Karl Buhler divide, igualmente, a conduta em três grupos: O instinto, o adestramento, e a inteligência. Mas ele recoloca o ensaio por tentativa e erro no adestramento e reserva o termo inteligência às condutas nas quais há compreensão imediata, repentina, sem qualquer forma de ensaio por tentativa e erro. Kohler vê a mesma inteligência como uma estruturação súbita das situações. A percepção nos dá uma estrutura direta, imediata, mas incompleta, a inteligência completa a coisa lhe reestruturando o conjunto de dados perceptivos e, igualmente para Kohler, o ensaio por tentativa e erro é excluído da inteligência, é uma espécie de inteligência, de sucedâneo da inteligência. (PIAGET, 1951)

Piaget questiona, dentre as diversas definições contrastantes de inteligência, qual delas estaria correta:

> Pois bem, quem tem razão? Todos os três ou nenhum; o que quero dizer é que o problema parece desprovido de significação. Não há critérios absolutos nem um senso estatístico. A inteligência se define pelo desenvolvimento e não por um critério absoluto, não há um limite inferior, ou seja, não podemos situar um dia, um mês, um ano no qual a inteligência apareça. A inteligência só pode se definir pelo seu processo. Ela é um processo de organização, que engloba o conjunto de funções cognitivas e que tende a uma forma de equilíbrio, que caminha em direção a certas formas de equilíbrio final. (PIAGET, 1951)

A distinção mencionada é igualmente relevante na definição de inteligência artificial, considerando que muitos dispositivos atuais estão equipados com tecnologias capazes de emular os sentidos humanos. Por exemplo, câmeras de vídeo e microfones podem captar sinais e estímulos análogos aos percebidos por olhos e ouvidos humanos, e tais dispositivos frequentemente desempenham um papel essencial na implementação de algoritmos de inteligência artificial. Especificamente nos dispositivos eletrônicos referidos, sua aplicação é significativa no domínio audiovisual de IA, um campo que primordialmente dissemina informações por meio desses sentidos.

De um modo simplificado, neste processo cognitivo, quando um humano pensa, ele avalia a situação, junta estas informações com todo o aprendizado, memória e experiências que teve, e toma uma decisão baseada nestes fatores. Durante este processo, absorve as experiências, e aprende. A inteligência artificial, no estágio atual em que se encontra, tenta mimetizar este processo: uma variedade de tecnologias distintas entre si, como armazenamento de informações, processamento e heurística, entre outras, que trabalham em conjunto, em uma configuração que as permita aprender e agir como uma inteligência humana.

1.1 Aprendizado de máquina

O processo que permite a IA se retroalimentar com os dados da execução do seu próprio algoritmo é chamado de “aprendizado de máquina”, e é importante diferenciar o termo supracitado do seu termo pai, “Inteligência Artificial”. Embora não seja difícil encontrar exemplos onde os termos são utilizados como sinônimos, são conceitos distintos e, para todos os efeitos, o aprendizado de máquina faz parte do termo guarda-chuva da IA.

Segundo Kubat (2021, p. 1) o aprendizado de máquina possibilita que o algoritmo de IA tome decisões ancoradas em um banco de dados de informações pré-existentes, que podem ser estruturadas ou não, permitindo a IA gerar resultados satisfatórios. É comumente aplicado para

tarefas específicas, já que uma grande quantidade de dados é necessária para o treinamento dos algoritmos: se você irá treinar uma IA para usar o aprendizado de máquina para desenhar faces, não se faz necessário alimentá-la com dados sobre sons de pássaros, por exemplo. Algoritmos de IA que têm como incumbência o desenho realista de faces humanas podem usar dezenas de milhões de imagens de rostos como modelo, tornando a tarefa de aprendizado de máquina altamente especializada:

> You will find it difficult to describe your mother's face accurately enough for your friend to recognize her in a supermarket. But if you show him a few of her pictures, he will immediately see the tell-tale traits. As they say, a picture - an example - is worth a thousand words. Likewise, you will not become a professional juggler by just being told how to do it. The best any instructor can do is to offer some initial advice, and then let you practice and practice and practice—and learn from your own experience and failures. This is what our technology is trying to emulate. Unable to define complicated concepts with adequate accuracy, we will convey them to the machine by way of examples. Unable to implement advanced skills, we instruct the computer to acquire them by systematic experimentation. Add to all this the ability to draw conclusions from the analysis of raw data, and to communicate them to the human user, and you know what the essence of machine learning is. (KUBAT, 2021, p. 1)

Logo, se o objetivo do algoritmo é aprender, dois pontos fundamentais são necessários: ter acesso a um banco de dados de informações relacionadas a aquele objetivo, e obter um *feedback* sobre se sua decisão está de acordo com o esperado, tal qual uma criança em uma escola, com suas aulas e avaliações. Como um algoritmo é incapaz de tomar decisões subjetivas, o método de aprendizado é adaptado à lógica binária do computador, e criado com regras que levam em conta a limitação da plataforma. O aprendizado de máquina em si é uma espécie de "problema" (no conceito de programação), e métodos diferentes de aprendizado podem levar a diferentes resultados, embora compartilhem a mesma base de ação. De um modo lúdico, pode-se dizer que os vários métodos pedagógicos aplicados ao ensino humano podem ser comparados aos diversos modos de solução de problemas de aprendizado de máquina.

Segundo Rudin (2017), assim como no aprendizado humano, o aprendizado de máquina não é algo imediato, é um processo. Várias iterações, cada uma com um acúmulo de informações relevantes para o contexto, são necessárias para que a proficiência do algoritmo na tarefa escolhida aumente gradativamente, até que atinja seu ápice – seja dado pelos limites da tarefa, ou do próprio algoritmo, ou, ainda, das capacidades da máquina em que este programa está sendo executado.

Ainda segundo Rudin (2017), são três os métodos de resolução de problemas de aprendizado de máquina: supervisionado, não supervisionado, e por reforço. No aprendizado supervisionado, os valores da base de dados são classificados, e a base possui o resultado almejado dentre os dados disponíveis, cabendo ao algoritmo inferir, por associações de variáveis independentes, uma variável dependente. Por exemplo, supõe-se que a tarefa de uma IA seja identificar números escritos à mão: entregamos ao algoritmo uma base de dados extensa de fotos ou escaneamentos de números escritos a caneta e a lápis por um humano, e, nesta base, há a informação de que a foto do número corresponde a um algarismo específico. O computador irá ler toda esta base, e tentar identificar quais características definem um determinado número, já que ele sabe qual número é qual, pois esta informação está presente na base de dados que lhe foi fornecida inicialmente. Feito isso, quando desafiado com uma foto de um algarismo novo, se o treinamento foi bem-sucedido, o algoritmo deveria ser capaz de poder identificar qual número está presente naquela imagem.

Figura 6 - Os três tipos de resolução de problemas de aprendizado de máquina.

Fonte: Disponível em https://www.dio.me/articles/diferencas-entre-aprendizado-supervisionado-e-nao-supervisionado Acesso em 14 Out. 2022

Em outro exemplo, um programador que tenha acesso a uma base de dados de uma companhia de seguros, base esta que contenha informações de sinistros anteriores e dados abstratos como idade do segurado, idade do automóvel, histórico do motorista, e outros parâmetros, possibilita a criação de um algoritmo que deduza, através de inteligência artificial, o perigo de um possível acidente caso confrontado com dados de um motorista novo.

O uso de algoritmos de inteligência artificial dotados de recursos de aprendizado de máquina na comunicação cresceu exponencialmente desde o início do século XXI. Os algoritmos de processamento de linguagem natural já são capazes de escrever textos jornalísticos verossímeis, a síntese de voz automática gera réplicas perfeitas de vozes aptas a dublar filmes ou simular a apresentação de produtos jornalísticos, entre outros usos da tecnologia de aprendizado de máquina que estão mudando o paradigma da comunicação – inclusive usos não lícitos, visando a criação de conteúdo falso e calunioso.

No campo audiovisual, o processo de edição de vídeos experimentou mudanças significativas com a transição das técnicas de edição lineares, baseadas no uso de fitas magnéticas, para técnicas de edição não-linear, que se utilizam de computadores para realizar operações de corte e colagem. Este avanço tecnológico proporcionou uma maior flexibilidade e eficiência no processo de edição. Recentemente, a integração do aprendizado de máquina neste processo introduz recursos avançados que permitem a identificação e edição automáticas de cenas, o ajuste de cores, a aplicação de efeitos especiais e a modificação de elementos gráficos dentro dos arquivos de vídeo. Tal integração resulta em uma redução substancial do tempo e do esforço humano necessário para a conclusão das tarefas de edição.

Além disso, o emprego de técnicas de *deep learning*[8] na restauração de filmes antigos exemplifica outra aplicação inovadora do aprendizado de máquina na área audiovisual. Essas técnicas permitem a melhoria da qualidade de imagem e som de obras cinematográficas históricas, contribuindo para a preservação e a valorização do patrimônio cultural. Adicionalmente, o uso de aprendizado de máquina para o reconhecimento facial e de objetos facilita a catalogação e a organização de conteúdo audiovisual, otimizando processos de busca e acesso a informações específicas dentro de grandes volumes de dados. Estes desenvolvimentos, possibilitados pelo aprendizado de máquina, indicam uma evolução significativa nas metodologias e ferramentas disponíveis para profissionais da área de edição

[8] Deep learning, ou aprendizado profundo, é um subcampo da inteligência artificial (IA) que utiliza redes neurais com várias camadas (profundas) para aprender representações de dados em níveis de abstração crescentes, permitindo a modelagem de relações complexas para reconhecimento de padrões, classificação e previsão.

de vídeos, impactando positivamente a eficiência e a qualidade do trabalho produzido no setor audiovisual.

Figura 7 - Vídeo colorido por um algoritmo de aprendizado de máquina. O filme é Beeld en Geluid" that was shot in 1921 in the picturesq villages of Giethoorn and Hierden, de 1921.

Fonte: Imagem retirada do canal "Ricks film restoration", no YouTube. Disponível em https://www.youtube.com/channel/UCdcPmaJEJeJTAnfSWMQNyHw. Acesso em 9 Out. 2022.

É importante salientar que os algoritmos de inteligência artificial que se beneficiam de técnicas de aprendizado de máquina são especializados em realizar tarefas específicas: um software de colorização de vídeos é treinado para fazer apenas esta função, não sendo eficiente, por exemplo, para cortar estes vídeos. Esta forma de algoritmo, que prevalece nas aplicações contemporâneas de IA, é caracterizada pela sua habilidade de processar e analisar grandes volumes de dados, aprender padrões e executar tarefas que os humanos levariam muito tempo (ou sequer conseguiriam realizar). No entanto, humanos são, de certa forma, flexíveis nas tarefas que são capazes de fazer: um humano pode demorar anos para colorir manualmente um vídeo que seja processado por um algoritmo em minutos ou horas, mas este mesmo humano pode realizar o corte deste mesmo vídeo, ou compor sua trilha sonora, entre outros fatores.

Um questionamento é possível, sobre a faceta multitarefa de um humano, em detrimento a capacidade limitada de um modelo de inteligência artificial: nem todos os humanos são capazes de realizar tarefas que uma máquina também não consegue. Esse questionamento é representado em um diálogo no longa metragem “Eu, Robô”, baseado na obra do escritor de ficção científica Isaac Asimov, onde um humano, agente da polícia, questiona um robô

humanoide sobre suas capacidades: “Humanos tem sonhos. Até cachorros tem sonhos, mas você não. Você é apenas uma máquina. Uma imitação da vida humana. Um robô é capaz de escrever uma sinfonia? Um robô pode criar uma bela pintura a partir de uma tela vazia?”, diz. O robô responde: “E você, você consegue?”.

Temos um panorama vasto de algoritmos que são capazes de realizar tarefas específicas com maestria, mas restritos a apenas esta tarefa, sem serem capazes de expandir o seu escopo e realizarem outras tarefas concomitantemente. A criação de um algoritmo que seja capaz de realizar várias tarefas diferentes, programado por uma base de dados só, aprendendo tal qual um humano é capaz de aprender, ainda é distante.

1.2 Categorias de inteligências artificiais

Para o escopo deste trabalho, assumimos que o conceito base de inteligência artificial seja o que Russel e Norvig (1995) citam em seus estudos: a definição de que IA é o estudo dos agentes que recebem informações sobre o ambiente e performam ações. Dentro deste contexto, podemos dividir a forma de ação de um algoritmo capaz de agir em quatro pontos: o armazenamento e processamento de dados, o aprendizado, o raciocínio e a resolução do problema. Para um algoritmo de inteligência artificial ser bem-sucedido, é necessário que ele tenha a capacidade de armazenar e analisar uma quantidade grande de dados, tenha a habilidade de aprender, seja através de entrada de dados, padrões, histórico, feedback, e outras formas, consiga deduzir ou inferir, baseado nos dados que adquiriu e no que aprendeu, analisando a situação proposta e possa resolver problemas complexos, que podem ser específicos, ou genéricos.

Por exemplo, quando um algoritmo sugere filmes e séries para um usuário de um serviço de *streaming*, baseado no histórico do que este mesmo usuário já assistiu anteriormente. Na elaboração desta IA, não há uma instrução exata, que represente algo como “se o usuário assistiu X, mostre Y e Z para ele como sugestão”. O algoritmo é instruído para armazenar e catalogar tudo o que várias pessoas assistem, e o que elas procuram. O computador tem a tarefa de, através deste catálogo, observar os padrões e chegar a uma conclusão, baseada nesta observação, do que uma pessoa com aquele perfil gostaria de ver. Para apresentar esta sugestão ao usuário, o computador pode acessar vários dados do comportamento dele, além dos programas que ele acompanha: sua idade, sua renda, onde trabalha, com quem se relaciona, entre outros. Tal qual no processo do pensamento humano, cabe a máquina julgar as associações

destes dados, decidir quais sugestões ela quer fornecer, e apresentá-las ao espectador – embora, é importante frisar, o processo que a máquina usa para chegar a esta conclusão costuma ser completamente diferente do que o caminho feito por um humano para chegar a mesma conclusão. Como um algoritmo de IA costuma ser retroalimentado – as informações que ele gera são somadas ao *feedback* que ele recebe e incorporadas no raciocínio simulado – as reações do consumidor e as sugestões oferecidas são avaliadas e podem ser usadas para melhorar os resultados futuros deste mesmo código.

Embora a citada IA de recomendação de filmes cumpra o seu papel de modo eficiente, ela funciona dentro de uma espécie de um microcosmo[9], onde, dentro da sua capacidade limitada ao foco para que foi criada, ela tem sua inteligência preservada. É uma máquina que lhe recomenda filmes, mas não os cria. Hoje, os melhores campeões de xadrez são frequentemente batidos por máquinas dotadas de inteligência artificial, mas estas mesmas máquinas campeãs de xadrez não conseguem efetuar uma recomendação de filme. Praticamente todas as inteligências artificiais hoje são assim, especializadas e altamente eficientes em cumprir uma tarefa específica. A essas inteligências artificiais damos o nome de "IA Fraca": elas são, geralmente, muito melhores do que humanos para resolver os problemas que lhes foram dados, porém incapazes de resolverem quaisquer outros problemas. Em contraponto a essas IA, Searle (1980) propõe o conceito de "IA Forte", cuja programação não tem o intuito de resolver um problema, mas sim de criar uma máquina que emule o comportamento humano:

> According to weak AI, the principal value of the computer in the study of the mind is that it gives US a very powerful tool. For example, it enables us to formulate and test hypotheses in a more rigorous and precise fashion. But according to strong AI, the computer is not merely a tool in the study of the mind; rather, the appropriately programmed computer really is a mind, in the sense that computers given the right programs can be literally said to understand and have other cognitive states. In Strong AI, because the programmed computer has cognitive states, the programs are not mere tools that enable us to test psychological explanations; rather, the programs are themselves the explanations. (SEARLE, 1980, p. 417)

As iniciativas de produção de máquinas que adotem o conceito de IA Forte ainda são muito incipientes, com a vasta maioria de projetos adotando os conceitos de uma IA Fraca. Dentre as iniciativas de IA que têm como foco a produção ou o auxílio de produtos audiovisuais, todos os projetos analisados nesta tese são de IA Fraca. Embora essa diferenciação seja, de certo

[9] Microcosmo, no caso, é usado no sentido de "pequeno mundo fechado".

modo, irrelevante, se considerarmos que todos os esforços de IA no audiovisual se baseiam em interpretações de IA específicas e de cunho exclusivo – softwares altamente especializados para cumprir uma única função, tais como melhorar uma imagem, remover um fundo verde, acrescentar elementos a um vídeo, entre outros - a discussão se torna razoável quando falamos da produção de conteúdo por inteligência artificial, como roteiros e peças artísticas, que exigem, idealmente, um algoritmo de IA capaz de entender muitas nuances e situações, para poder dar saída a um produto com uma criatividade comparável à de um humano.

1.3 Conceitos de uma inteligência artificial criativa

A história da inteligência artificial com fins artísticos começa na década de 1960, anos depois da criação dos computadores. Um dos primeiros exemplos do uso de inteligência artificial foi o AARON, programa de computador desenvolvido pelo artista e pesquisador Harold Cohen (COHEN, 2016). Foi um dos primeiros sistemas autônomos capaz de realizar atividades artísticas: programado usando uma série de regras e procedimentos algorítmicos, o AARON era capaz de gerar imagens que possuíam elementos de composição, cor e forma, e foi refinado ao longo de décadas pelo seu criador, ganhando a capacidade de gerar obras mais complexas e esteticamente variadas.

Figura 8 - Imagem criada pelo software AARON. Exposta hoje no museu de computação de Boston, EUA.

Um aspecto chave do software AARON é a sua relativa autonomia. Embora inicialmente dependesse de intervenção humana para algumas funções como a aplicação de cores, versões subsequentes do programa demonstram uma maior independência na criação de obras completas. Outros artistas experimentaram com a geração de arte por inteligência artificial nas décadas de 1960 e 1970, como Vera Molnar, que utilizou algoritmos de desenhos geométricos repetitivos para criar padrões com variações sutis, características de seu trabalho. Ambos os artistas (Molnar e Cohen), bem como todos os outros de sua época que realizaram experimentações com inteligência artificial, trabalharam com algoritmos de abordagem simples, sem aprendizado de máquina nem redes neurais.

Na primeira década do século XXI, as técnicas usadas por algoritmos de inteligência artificial passaram por uma mudança de paradigma atribuída ao advento do *deep learning* – onde a geração artística por computador passou a ser impulsionada algoritmos dotados de recursos de aprendizado de máquina, caracterizados pelo uso de redes neurais multicamada (onde cada aspecto artístico, como cores, traços e estilos, eram dados por uma camada distinta, e somados posteriormente).

Iniciativas de geração artística desta época, feitas por softwares dotados de ferramentas de *deep learning*, como o programa *DeepDream* (2015), criado pelo engenheiro do Google Alexander Mordvintsev, que gera composições artísticas surrealistas, diferenciam-se dos algoritmos criados nas décadas anteriores não só pela sua faceta do aprendizado de máquina, mas também pela sua disponibilidade: a ferramenta de Mordvintsev foi disponibilizada ao público, que podia gerar suas próprias imagens pela internet, sem custo. O advento do *deep learning* também propiciou outras iniciativas artísticas que não eram possíveis anteriormente por limitações técnicas, como o projeto “The Next Rembrandt”, da Microsoft, que, através do uso de inteligência artificial e aprendizado de máquina, visou reproduzir com meticulosos detalhes as nuances estilísticas do pintor holandês Rembrandt, incluindo o uso de uma impressora 3d para simular as técnicas de múltiplas camadas de tinta que o já falecido artista usava.

Figura 9 - - Imagem gerada pelo software DeepDream

Fonte: Disponível em https://deepdreamgenerator.com/ Acesso em 5 Nov. 2023

Em 2014, Goodfellow et al. (2014) introduziram uma nova abordagem para o *deep learning*, o uso de redes generativas adversariais (*GAN, de Generated Adversarial Networks*). Uma GAN minimiza um dos problemas que os algoritmos que se utilizavam de técnicas de *deep learning* possuíam: era difícil estimar com precisão se os múltiplos resultados de um treinamento eram adequados ou não para a situação proposta. Inserindo um componente de verificação, e tornando-o um algoritmo que compete com o algoritmo de geração, a precisão do resultado aumenta exponencialmente.

> In the proposed adversarial nets framework, the generative model is pitted against an adversary: a discriminative model that learns to determine whether a sample is from the model distribution or the data distribution. The generative model can be thought of as analogous to a team of counterfeiters, trying to produce fake currency and use it without detection, while the discriminative model is analogous to the police, trying to detect the counterfeit currency. Competition in this game drives both teams to improve their methods until the counterfeits are indistiguishable from the genuine articles. (GOODFELLOW, POUGET-ABADIE, *et al.*, 2014, p. 1)

Essa forma nova de se treinar algoritmos de IA é uma das responsáveis, juntamente com outros avanços como o uso de *transformers* e de mecanismos de atenção (ambos detalhados no

capítulo seguinte desta tese), pela realidade atual da IA generativa, onde entradas de texto geram mídias (imagens, áudios, vídeos) de modo automático e eficiente.

Iniciativas de inteligência artificial generativa no campo artístico foram sendo criadas ao longo dos anos, sem que a área da comunicação acompanhasse o interesse pela tecnologia. Historicamente, a comunicação foi considerada um processo essencialmente humano, mediado pela tecnologia, e não produzido por ela. Na década de 1970, enquanto artistas como Cohen e Molnar experimentavam com IA, acadêmicos da comunicação como Schramm (1971) e Dance (1970) ressaltavam que o fenômeno da comunicação humana era, para todos os efeitos, humano. Outros tipos de comunicação eram admitidos, como entre animais ou até entre máquinas, mas a área que os estudos se debruçavam deixava outras entidades de lado para focar estritamente na comunicação entre seres humanos.

> Let us understand clearly one thing about it: communication (human communication, at least) is something people do. It has no life of its own. There is no magic about it except what people in the communication relationship put into it. There is no meaning in a message except what the people put into it. When one studies communication, therefore, one studies people - relating to each other and to their groups, organizations, and societies, influencing each other, being influenced, informing and being informed, teaching and being taught, entertaining and being entertained - by means of certain signs which exist separately from either of them. To understand the human communication process, one must understand how people relate to each other. (SCHRAMM, 1971)

É um ponto a se considerar, se a forte atenção que as novas ferramentas de inteligência artificial generativa disponíveis atualmente estão obtendo na área da pesquisa da comunicação não deriva de um distanciamento dos pesquisadores da área durante o período que a tecnologia evoluiu. Pierre Levy (2022), em artigo escrito em seu site da internet, considera que um dos fatores que causa esse distanciamento é a volatilidade com a qual tecnologias ganham e perdem seu *status* de destaque e são incorporadas no cotidiano:

> Historical observation reveals the tendency to classify the “advanced” applications into artificial intelligence in the eras in which they first emerge; however, years later, these same applications are often reattributed to everyday computing. For example, visual character recognition, originally known to be AI, is now considered to be commonplace and is often integrated into software programs without fanfare. A machine capable of playing chess was celebrated as a technical achievement in the 1970s, but today you can easily download a free chess program onto your smartphone without a hint of astonishment from anyone. (LEVY, 2022)

Levy classifica o *marketing* de tendência como um dos responsáveis por este distanciamento, já que a falta de constância de valor – onde uma tecnologia é classificada como revolucionária em um tempo curto, para depois cair no ostracismo – causa o desinteresse da ciência, que evolui a passos mais lentos.

> Moreover, depending on whether AI is trendy (as it is today) or discredited (as it was in the 1990s and 2000s), marketing strategies will either emphasize the term AI or replace it with others. For example, the « expert systems » of the 1980s become the innocuous « business rules » in the 2000s. This is how identical techniques or concepts change name according to the fashion, making the perception of the domain and its evolution particularly opaque. (LEVY, 2022)

Esse *marketing* excessivo das ferramentas de inteligência artificial acaba mascarando um problema: a efetividade destas ferramentas é limitada.

Os produtos de inteligências artificiais generativas feitos com os algoritmos atuais sofrem com as limitações tecnológicas vigentes. É esperado que a tecnologia evolua com a melhora dos métodos de codificação, ou com a evolução do *hardware* disponível para o processamento destes programas, mas é necessário também evoluir a própria abordagem da temática para que este produto tenha sua eficácia majorada e seja mais aceito. Segundo Russel e Norvig (1995) a própria abordagem é fragmentada, pois parte de diferentes premissas, com diferentes definições:

Figura 10 - Abordagens de conceito de IA

PROCESSOS DO PENSAMENTO	Pensar como um Humano	Pensar Racionalmente
COMPORTAMENTO	Agir como um Humano	Agir Racionalmente
	PERFORMANCE HUMANA	RACIONALIDADE

Fonte: Livro Artificial Intelligence: A Modern Approach. Tradução nossa.

Russel e Norvig separam as abordagens de conceito de uma IA em dois eixos: no caso da figura acima, o que está na área superior (primeira linha) representa desenvolvimentos que se preocupam com o processo do pensamento e raciocínio, se aproximando do cognitivo, do PENSAR. Já o que se situa na área inferior (segunda linha), se baseia no comportamento, no AGIR. Tanto máquinas que pensam, quanto as que agem, podem ter seu sucesso baseado em duas situações: ou a máquina é avaliada e o sucesso do algoritmo é baseado na fidelidade de resultado com o comportamento humano padrão, ou a validação é feita baseada no suposto comportamento "ideal", de um modo racionalista.

Uma IA que busca pensar como um humano se aproxima do conceito de "Inteligência Artificial Geral", hoje, ainda hipotético. Consiste na completa emulação do cérebro humano, seus neurônios e sinapses, de modo a que a máquina tenha reações a diversas situações do mesmo modo que um humano teria. Tal grau de sofisticação de IA é, hoje, hipotético, já que o poder computacional disponível hoje não é suficiente para uma simulação completa de um cérebro humano. Sua implementação envolve as ciências cognitivas, que buscam, ainda, definir e qualificar como um humano pensa e qual o percurso que um ser humano passa para definir o sentido dos itens do mundo que o permeia.

Já uma IA que é focada em pensar de modo racional se aproxima da lógica aristotélica e da maneira tradicional com que os computadores operam. A lógica aristotélica é profundamente caracterizada pela sua estrutura silogística, uma forma de raciocínio dedutivo, que concatena premissas para inferir conclusões inevitáveis. Esse raciocínio silogístico serve como base para a criação de algoritmos: estes são, essencialmente, sequências de passos lógicos para resolver problemas ou realizar tarefas. Na lógica aristotélica, um argumento é construído a partir de premissas para alcançar uma conclusão lógica, e, de maneira similar, um algoritmo é uma sequência de instruções que partem de uma entrada – que aqui pode ser considerada uma premissa – até chegar em uma saída – que é uma conclusão – seguindo lógicas definidas. A lógica booleana[10] é uma extensão direta da lógica aristotélica, trabalhando com proposições de "verdadeiro" e "falso", e é fundamental para tomadas de decisões e controle de fluxo em algoritmos. Antes dos algoritmos, o próprio circuito integrado dos computadores usa lógica booleana em seu desenho.

[10] A lógica booleana, também conhecida como álgebra booleana, é um sistema de matemática que se dedica a trabalhar com valores verdadeiros e falsos. Ela foi desenvolvida por George Boole, um matemático inglês, no século XIX. A lógica booleana é fundamental para o funcionamento de computadores e sistemas eletrônicos modernos, uma vez que eles funcionam com base em circuitos que representam os dois estados de 0 (falso) e 1 (verdadeiro).

O pensamento racional é uma abordagem presente na maioria absoluta das IA disponíveis hoje, onde uma máquina é abastecida com dados e programada com uma lógica binária. Sofrem de alguns problemas, como uma limitação de escopo: como são programadas com premissas básicas e lógicas, aprendem pouco e de modo mais lento. Cabe ressaltar, nem tudo pode ser solucionado de modo efetivo com lógica binária. Nem tudo é ou não é; como exemplo, aqui cabe a questão da criação de vozes *TTS*[11], onde por mais que a tecnologia tenha evoluído, o som gerado pela programação tradicional, sem o auxílio de IA, é monotônico e carece de realismo, como a voz artificial do astrofísico Stephen Hawking.

> There are two main obstacles to this approach. First, it is not easy to take informal knowledge and state it in the formal terms required by logical notation, particularly when the knowledge is less than 100% certain. Second, there is a big difference between solving a problem "in principle" and solving it in practice. Even problems with just a few hundred facts can exhaust the computational resources of any computer unless it has some guidance as to which reasoning steps to try first. Although both of these obstacles apply to any attempt to build computational reasoning systems, they appeared first in the logicist tradition (RUSSEL & NORVIG, 2010)

Já na abordagem onde a máquina tenta agir de modo racional, a máquina analisa os dados obtidos, aprende e toma ações baseadas neste processo, sem se preocupar com a verossimilhança com o comportamento humano típico. Aqui, como exemplo, cabem os mecanismos de direção dos carros autônomos, principalmente na questão das tomadas de decisão: se um carro que se guia automaticamente, sem nenhuma intervenção humana, se depara com a situação de que, inevitavelmente, terá que decidir, em uma ação imediata, se atropela uma criança ou um idoso, livre de qualquer amarra ética, de dados sobre o processo, qual decisão ele toma, e baseado em qual critério? Já houve casos de pilotos humanos de aviões comerciais que, em queda iminente e com consequências catastróficas, fizeram manobras para evitar atingirem hospitais ou escolas no momento da colisão. Se fosse uma máquina a responsável por esta pilotagem, sem nenhuma programação prévia, ela não tomaria a mesma decisão, que parece humana em sua essência, mas carece de racionalidade pura e simples, e motivada por emoção.

Já quando uma IA é programada para agir tal qual um humano, esta IA tenta imitar o nosso comportamento. Um bom exemplo deste tipo de IA temos os *chatbots*: programas de

[11] Síntese de fala ou *Text-to-Speech* (TTS) é uma tecnologia que permite que os computadores convertam o texto em fala, sendo um processo artificial de produzir a fala humana.

conversa em tempo real na internet que tentam simular uma persona através de inteligência artificial, com o foco em linguagem natural e com a ideia central é tornar as interações entre humanos e máquinas mais palatáveis. A IA busca se comunicar em uma linguagem humana (processamento de linguagem natural), tenta armazenar e analisar informações (representação do conhecimento), objetiva tomar decisões baseadas nesse conhecimento acumulado (pensamento autônomo) e aprende com estas decisões e seus feedbacks (aprendizado de máquina). Os *chatbots* são um bom exemplo de evolução dentro da área de IA, já que sua gênese foi algorítmica – técnicas de *pattern matching*[12] e *character replacement*[13], mas sua evolução coincidiu com a evolução dos Modelos de Linguagem Natural (*Natural Language Model*, NLM), onde a técnica de abordagem deixou de ser puramente programática, para se basear em IA.

[12] *Pattern Matching* é uma técnica de programação que verifica se um valor determinado está dentro de uma tabela de padrões pré-definida.

[13] *Character Replacement* é uma técnica de programação que substitui certos caracteres dentro de um texto. Por exemplo, em português, substituir o artigo "O" por "A" pode mudar o gênero de certas palavras, ou acrescentar um "S" ao final de outras palavras pode mudar o substantivo.

2 OS MODELOS DE LINGUAGEM NATURAL

Desde a década de 1940, no pós-guerra, juntamente com as primeiras execuções práticas de computadores eletrônicos, que foram fomentadas pelas necessidades bélicas da época, discussões sobre a aplicabilidade da computação nas diversas áreas do cotidiano permeiam a comunidade científica. Dentre muitas destas aplicações, a possibilidade de que o processamento das recém-criadas máquinas se aproximasse da práxis da mente humana despertou o interesse de vários cientistas. Na época, os estudos na área da neurobiologia revelaram a existência – e forma de funcionamento – dos neurônios e sinapses do cérebro, e, junto, a conjecturas sobre a viabilidade da aproximação entre a mente humana e os computadores.

No final do século XIX, estudos passam a considerar a participação do cérebro nas mecânicas do pensamento humano, superando as barreiras que os dogmas religiosos e a visão mística da mente apresentavam para a ciência da época. A influência do misticismo nos estudos da mente humana era notória, a ponto de que a própria palavra associada aos fenômenos que ocorrem na mente humana, "*psique"*, é derivada do nome da deusa grega homônima, que representava a personificação da alma.

> O século XIX foi, entre outros, marcado pelo nascimento da biologia e pela revolução de ideias decorrentes da teoria da seleção natural proposta pelo naturalista Charles Robert Darwin (1809-1882). A concepção de mente, enquanto atributo supremo e divino do ser humano, fortemente marcada pelo dogma e poder religioso, deixava os vapores etéreos para se encarnar no sistema nervoso humano. A descoberta de que o córtex cerebral, até então considerado homogêneo do ponto de vista funcional, apresentava áreas anatomicamente definidas, deu suporte à ideia de que diferentes funções mentais estavam alojadas nas diferentes porções do córtex. (PINHEIRO, 2005, p. 184)

As primeiras teorias de que as atividades dos nervos humanos se davam por sinais elétricos datam do final do século XVI. Em relato histórico feito por Picolino (2006), no ano de 1786, o médico e físico italiano Luigi Galvani, estudioso da anatomia de rãs, realizava suas pesquisas dissecando anfíbios, e, durante uma forte tempestade, ao encostar um aparato cirúrgico de metal em uma das pernas de um dos animais, percebeu que a perna se movimentou. Percebeu que a eletricidade do ar amplificada pela tempestade havia causado o movimento, e concebeu que os nervos eram acionados através de impulsos elétricos. O cientista italiano

comprovou sua tese através de experimentos posteriores, usando máquinas geradoras de eletricidade.

No século XIX, os estudos da estrutura celular do sistema nervoso eram feitos através de coloração histológica[14], método que permitia a visualização mais eficiente das amostras em um microscópio. Segundo Ferreira (2022), as técnicas iniciais de coloração não permitiam uma análise detalhada das estruturas responsáveis pela comunicação neuronal, e cientistas da época como Albrecht von Kölliker (1868) acreditavam que a comunicação entre os neurônios se dava de célula a célula, através de redes de fibras que se agrupavam em forma de treliças, com o tecido nervoso sendo formado por uma grande rede, com todas as suas peças interligadas. Essa concepção da comunicação neuronal é denominada “Teoria Reticular”.

O panorama da investigação das comunicações neuronais começou a mudar com a descoberta de um novo método de coloração histológica pelo cientista italiano Camilo Golgi (1873). Golgi cortou tecidos em camadas de 5mm de espessura e deixou-as submergidas em uma solução de dicromato de potássio e tetróxido de ósmio. Depois, as peças eram lavadas e imersas em nitrato de prata por 2 dias em temperatura ambiente, submersas em parafina, e cortadas em camadas de 100 a 250 micrômetros de espessura, permitindo que as amostras ficassem mais visíveis do que os métodos de coloração histológica disponíveis até então.

Através do método de coloração de Golgi, Santiago Ramón y Cajal, médico e histologista espanhol, aperfeiçoou o método de Golgi para estudar a sistemática do sistema nervoso, descobrindo que, diferentemente do que a teoria reticular afirmava, o tecido nervoso não se agrupava em redes, mas se conectavam livremente através de contatos simples com as células nervosas adjacentes (CAJAL, 1904). Os métodos de observação dos nervos de Golgi com o uso de nitrato de prata e seu aperfeiçoamento por Ramon y Cajal para observação dos neurônios cerebrais revelaram a verdadeira natureza da comunicação dos neurônios do cérebro humano, e estas descobertas renderam aos dois o prêmio Nobel de medicina, em 1906.

A teoria de máquinas analógicas de processamento de informação existia desde o início do século XIX, com os estudos do matemático e engenheiro mecânico inglês Charles Babbage e da matemática e escritora Ada Lovelace, na universidade de Cambridge, na Inglaterra. A máquina analítica de Babbage (1837), foi a primeira proposta de um dispositivo mecânico

[14] A coloração é uma técnica usada para aumentar o contraste em amostras de materiais, frequentemente usada em análises de escala miniaturizada. A coloração histológica trata do uso dessa técnica de coloração nos estudos da histologia, que estuda células no microscópio.

analítico programável, que possuía uma unidade de aritmética lógica, interface de controle de fluxo através de desvios condicionais e *loops*, e memória integrada (BROMLEY, 1982). O design de Babbage é considerado o conceito pioneiro de um computador de uso geral, e os primeiros algoritmos para esta máquina – na época, hipotética, pois nunca foi construída de fato – foram concebidos por Lovelace.

Um século depois dos conceitos de Babbage e Lovelace, Turing (1937, p. 231) formulou um modelo teórico que hoje é conhecido como "Máquina de Turing": um conceito de máquina universal, que teoricamente computava e calculava, através de lógica binária, com instruções definidas por um operador. Estas instruções ganhavam o nome de algoritmo, e esta máquina teórica foi a base para a construção dos primeiros computadores eletrônicos, décadas depois.

Neste espaço de meio século entre a descoberta dos neurônios e a máquina de Turing, a ciência pode aferir que: o cérebro humano funcionava através da troca de estados elétricos entre células, e que era possível construir máquinas computacionais capazes de realizar cálculos matemáticos, através de um processo que se assemelhava ao funcionamento do órgão humano. Naturalmente conexões entre as duas teorias foram feitas, conjecturando-se a possibilidade de se simular partes ou o todo do cérebro humano através de estruturas computacionais.

O próprio Turing, que trabalhara ativamente nos esforços de guerra, colaborando no uso da computação para a quebra dos algoritmos de encriptação dos alemães, findado o conflito, debruçou-se nas hipóteses de aproximação entre humanos e máquinas. Foi o primeiro a citar o termo "Inteligência Computacional" em 1947, em um simpósio, onde indica que "buscava uma máquina que pode aprender com suas próprias experiências, e, com isso, ter a possibilidade de que esta máquina altere suas próprias instruções com o que aprendeu". Turing propõe, em um ensaio escrito em 1948, chamado "*Intelligent Machinery*", possibilidades de se mimetizar os neurônios humanos através de computação, associando o ser humano a uma máquina e unindo pela primeira vez neurociência e computação.

> A great positive reason for believing in the possibility of making thinking machinery is the fact that it is possible to make machinery to imitate any small part of a man. (...) Here we are chiefly interested in the nervous system. We could produce fairly accurate electrical models to copy the behaviour of nerves, but there seems very little point in doing so. It would be rather like putting a lot of work into cars which walked on legs instead of continuing to use wheels. The electrical circuits which are used in electronic computing machinery seem to have the essential properties of nerves. They are able to transmit information from place to place, and also to store it. (TURING, 1948)

Dois anos depois, Turing (1950) faz considerações sobre “máquinas que aprendem”, baseando-se em conceitos oriundos de considerações de Ada Lovelace sobre a máquina de Charles Babbage:

> Our most detailed information of Babbage's Analytical Engine comes from a memoir by Lady Lovelace (1842). In it she states, "*The Analytical Engine has no pretensions to originate anything. It can do whatever we know how to order it to perform*" (her italics). (TURING, 1950, p. 444)

Esses questionamentos de Lovelace sobre a passividade analítica da máquina de Babbage faz Turing questionar sobre se as máquinas podem, de fato, pensar, introduzindo a ideia do que hoje é conhecido como o “Teste de Turing”. Este teste é uma metodologia proposta para avaliar se uma máquina possui inteligência comparável à humana, baseada na sua capacidade de gerar respostas que sejam indistinguíveis das de um ser humano para um observador.

Embora Turing não tenha delineado especificamente algoritmos de aprendizado de máquina no sentido moderno, sua discussão pavimentou o caminho para o campo da IA, contemplando a possibilidade de máquinas aprenderem e se modificarem com base na experiência. Ele propôs a noção de uma "máquina infantil" que, semelhante ao desenvolvimento humano, poderia ser educada e desenvolver sua inteligência ao longo do tempo através da interação com o ambiente.

Turing especulou sobre a ideia de que seria mais eficaz construir máquinas que pudessem aprender e evoluir suas habilidades intelectuais, em detrimento de se tentar criar uma máquina com uma mente completamente desenvolvida desde o início. Ele sugeriu que essas máquinas poderiam aprimorar suas capacidades de solução de problemas e adaptação ao receber e processar informações externas, um conceito que é um pilar dos sistemas de aprendizado de máquina atuais, onde algoritmos são expostos a vastos volumes de dados para refinar suas funções.

O teste de Turing, portanto, propõe um critério para avaliar a capacidade de uma máquina de imitar o comportamento inteligente: se um interrogador humano, após interagir com uma máquina e um outro humano por meio de um canal de comunicação que ocultasse

suas identidades, não pudesse distinguir, de maneira confiável, quem era a máquina e quem era o humano, isso era uma indicação de que aquela máquina era, de fato, inteligente.

O artigo de Turing provocou discussões acadêmicas em várias áreas do conhecimento, com interpretações conflitantes sobre o impacto da proposição de Turing na maneira com a qual humanos se relacionam com máquinas. Sterrett (2000) argumenta que Turing apresentou o seu teste não apenas como um mero critério prático para testar inteligência de máquinas, mas como uma reflexão profunda sobre a natureza da compreensão da mente humana. Sterret considera que Turing sequer usou a inteligência humana como padrão de inteligência:

> My point does not turn on the historical question of whether or not Turing held the view that human intelligence is not the standard by which machine intelligence should be measured, though. The conceptual point is that the Original Imitation Game Test does not use human intelligence as a standard of intelligence. To be sure, it uses comparison with a man's performance, but, as I have argued, it constructs a setting in which the man really has to struggle to produce the responses required to succeed at the task set, in which he is forced to think about something he may not otherwise ever choose to think about. (STERRETT, 2000, p. 553-554)

Sterrett defende que o teste de Turing deve ser visto dentro de um contexto mais amplo de pensamento computacional e filosófico, argumentando que ele levanta questões importantes sobre o que significa pensar e como isso pode ser reconhecido entre máquinas. Indica que a área da filosofia considera que o teste de Turing não funciona como um teste segundo o conceito que Turing desejava (STERRETT, 2000, p. 556), mas sim como um teste empírico baseado no comportamento humano e na linguística:

> If we reflect on how the Original Imitation Game Test manages to succeed as an empirical, behavior-based test that employs comparison with a human's linguistic performance in constructing a criterion for evaluation, yet does not make mere indistinguishability from a human's linguistic performance the criterion, we see it is because it takes a longer view of intelligence then linguistic competence. In short: that intelligence lies, not in the having of cognitive habits developed in learning to converse, but in the exercise of the intellectual powers required to recognize, evaluate, and, when called for, override them. (STERRETT, 2000, p. 556)

Esse entendimento de que o teste de Turing é um teste sobre a capacidade de uma máquina entender e processar a linguagem natural humana é também compartilhado por Rapaport (1994):

> I said earlier that understanding natural language is a mark of intelligence. In what sense is it such a mark? Turing rejected the question, "Can machines think?", in favor of the more behavioristic question, "Can a machine convince a human to believe that it (the computer) is a human?" To be able to do that, the computer must be able to understand natural language. So, understanding natural language is a necessary condition for passing the Turing Test, and to that extent, at least, it is a mark of intelligence. (RAPAPORT, 1994, p. 83)

Rapaport argumenta que a capacidade de uma máquina passar no teste de Turing não implica, necessariamente, que aquela máquina possui consciência ou entendimento em um sentido humano pleno, mas sim indica um nível sofisticado de processamento de linguagem natural e de representação do conhecimento.

Três décadas antes das considerações de Rapaport e Sterrett sobre a natureza do teste de Turing, o pesquisador Joseph Weizenbaum desenvolveu, no laboratório de inteligência artificial do MIT (*Massachusetts Institute of Technology*) o ELIZA, considerado o primeiro programa de processamento de linguagem natural da história. O nome ELIZA deriva da personagem principal da peça “Pigmalião” de George Bernard Shaw, chamada Eliza Doolittle, que é transformada de uma vendedora de flores em uma dama da sociedade através de lições de fala.

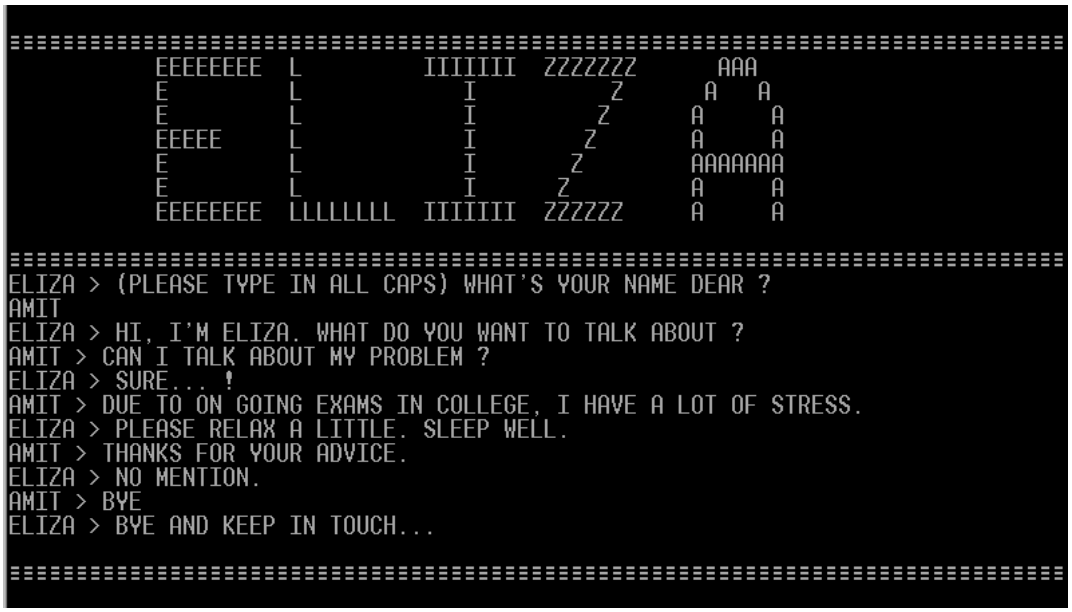
O ELIZA funciona através de programação lógica através de *scripts*, componentes centrais da sua arquitetura, permitindo que o software simule conversas ao seguir regras pré-definidas associadas a padrões específicos de entrada de texto. Estes *scripts* definem como o ELIZA deve responder a diferentes tipos de entradas dos usuários, essencialmente agindo como o “cérebro” do programa. Cada *script* é projetado para um tipo específico de diálogo.

De todos os scripts desenvolvidos para o ELIZA, o denominado “DOCTOR” foi o mais bem sucedido. Ele simula uma interlocução terapêutica alinhada a abordagem centrada no paciente do psicoterapeuta Carl Rogers, evidenciando uma técnica conversacional que ecoa as expressões do usuário, incentivando assim a continuidade do diálogo. Por exemplo, ante uma expressão de descontentamento do usuário, o ELIZA pode responder com “Percebo uma tristeza em você. Gostaria de conversar mais sobre isso?”. O DOCTOR emprega estratégias diversas de resposta, tais como estimular o usuário a detalhar suas declarações, espelhar emocionalmente

as expressões do usuário, e formular indagações baseadas nas assertivas do interlocutor. Além do DOCTOR, outros *scripts* foram adotados pelo ELIZA, capazes de emular distintas modalidades de conversação, como uma simulação de entrevista, simulação de um interrogatório, entre outros.

Todos os *scripts* do ELIZA trabalham da mesma forma operacional, com a detecção de palavras-chave e termos textuais que desencadeiam respostas específicas e parâmetros de detecção de sentimentos. São limitados pela inabilidade de entender autenticamente o contexto ou o significado das palavras escritas pelo interlocutor, operando por meio de um reconhecimento de padrões e substituição textual, desprovidos de qualquer compreensão semântica ou contextual que seja verdadeira. Isso implica que o ELIZA pode gerar réplicas inadequadas ou sem sentido, caso a entrada do usuário desvie dos padrões antecipados pelo código.

Figura 11- Interface do software ELIZA



Para todos os efeitos, em uma observação estrita do teste de Turing, o ELIZA é aprovado, já que, para certas condições, um humano não é capaz de distinguir se está conversando com um uma máquina ou com um dos seus pares. Porém aqui é importante aplicar a distinção defendida por Sterrett e Rapaport de que a proposta do teste de Turing não era a de verificar uma possível inteligência humana nas respostas de uma máquina, mas sim avaliar a capacidade deste computador se passar por um humano em uma conversação. O ELIZA não tenta, em nenhum momento, imitar a inteligência humana, mas sim utiliza-se de técnicas de

correspondência de padrões e substituição de palavras, buscando palavras-chave no texto inserido pelo usuário, processando-as por um algoritmo definido, e devolvendo-a de modo a parecer um diálogo plausível:

> ELIZA is a program operating within the MAC time-sharing system at MIT which makes certain kinds of natural language conversation between man and computer possible. Input sentences are analyzed on the basis of decomposition rules which are triggered by key words appearing in the input text. Responses are generated by reassembly rules associated with selected decomposition rules. The fundamental technical problems with which ELIZA is concerned are: (1) the identification of key words, (2) the discovery of minimal context, (3) the choice of appropriate transformations, (4) generation of responses in the absence of key words, and (5) the provision of an editing capability for ELIZA "scripts". (WEIZENBAUM, 1966, p. 36)

Pessoas que interagiam com o ELIZA relatavam a impressão de que estavam falando com outro humano: a secretária de Weizenbaum solicitou ficar sozinha com o computador enquanto estivesse conversando com o *software*, alegando que se sentia interagindo com um terapeuta humano. Weizenbaum se viu surpreso e chocado com as reações de pessoas que tiveram contato com o programa, dizendo que “Não tinha imaginado que uma exposição tão breve a um programa de computador relativamente simples pudesse causar um pensamento tão delirante em pessoas aparentemente normais” (WEIZENBAUM, 1976, p. 7). A simplicidade que Weizenbaum ressalta era evidente no algoritmo do programa: o ELIZA contava com menos de mil linhas de programação e usava regras simples de substituição gramatical.

Ocorre que, em casos como o da secretária de Weizenbaum, o humano que estava interagindo com o computador tinha plena ciência de que do outro lado da interação havia uma máquina. Rhee (2010) fala sobre esta tendência humana da antropomorfização, e cita o “Efeito ELIZA”, onde humanos acreditam que um computador é inteligente, mesmo tendo a ciência de que ele não é:

> ELIZA's convincing performance, or rather humans' anthropomorphization of ELIZA, is now so well-known that the phrase "the Eliza effect" has been coined to describe the phenomenon in which humans believe a computer to be intelligent and possessing intentionality. Noah Wardrip-Fruin describes this Eliza effect as "the not-uncommon illusion that an interactive computer system is more 'intelligent' (or substantially more complex and capable) than it actually is" (Wardrip-Fruin 25). ELIZA, in humans' intimate engagements with it, exists as an important moment in the history of the anthropomorphic imaginary initiated by the Turing test. (RHEE, 2010)

Tal situação insólita não é muito diferente da comoção recente a respeito de ferramentas modernas de processamento de linguagem natural como o *ChatGPT*, que provocam semelhante celeuma, mas diferenças importantes de conceito separam o ELIZA em 1966 das ferramentas modernas de NLP do século XXI: os seus algoritmos.

2.1 Algoritmos de Modelos de Linguagem

Segundo Carbonell et al. (1985) a esquemática de um algoritmo computacional pode ser representada de modo figurado por uma sentença simples e genérica, representada por variáveis: um programa de computador, representado pelo acrônimo C, é abastecido por dados D, que realiza a ação A, tem a performance P, e o resultado R. No modo tradicional de computação (não dotado de inteligência artificial), no qual que o ELIZA é baseado, esta dinâmica é sempre a mesma: caso os dados inseridos sejam iguais, o resultado desse processamento será sempre idêntico ao anterior. A abordagem da computação tradicional segue regras fixas e procedimentos específicos para resolver problemas, e a eficácia de um algoritmo tradicional depende da precisão de seu desenho inicial e da consideração de todos os cenários possíveis durante a programação. A computação tradicional, portanto, utiliza uma abordagem determinística, onde as saídas são diretamente determinadas por entradas específicas e por um conjunto previamente definido de operações. Quando esta tese cita o termo "computação tradicional", está se referindo a esta abordagem determinística e previamente programada do algoritmo. A abordagem tradicional difere da abordagem da IA, pois, algoritmos de inteligência artificial são projetados para aprender e melhorar com o tempo, adaptando suas estratégias e respostas com base em novas informações.

Volta-se, então, às questões de Turing (1948, p. 107), que sugerem outro método: e se um computador puder aprender de suas próprias experiências? Ainda vivo, ele viu o surgimento do primeiro programa de computador a aplicar estes conceitos, em 1952: o "*Game of Checkers*" (Jogo de Damas) desenvolvido por Arthur Samuel enquanto trabalhava para a IBM, e aperfeiçoado nos anos seguintes. O *software* desenvolvido por Samuel jogava damas contra humanos, e utilizou várias técnicas inovadoras que, anos depois, se tornariam fundamentais no campo da inteligência artificial, como o aprendizado de máquina e algoritmos de definição de valores (SAMUEL, 1959).

O uso da inteligência artificial muda a maneira com a qual os algoritmos processam informações. Com o uso da inteligência artificial, o resultado R é incorporado e processado,

com a ajuda da medição de performance P, para compor a próxima tomada da ação A, o que tende a afetar o próximo resultado R, e assim por diante. Pode-se dizer que, tal qual um ser humano, o algoritmo "aprendeu" algo durante o processo, e o resultado da sua próxima iteração[15] tende a ser mais satisfatório do que a sua iteração anterior. Baseado nessa dinâmica, como exemplo, tomemos um hipotético algoritmo bancário C, cujo objetivo seja definir os limites de um cartão de crédito disponibilizados a uma pessoa qualquer. Programado para analisar a renda, movimentações bancárias e demais dados financeiros do futuro usuário (D), o algoritmo leva um segundo (P), e decide (A) que o limite daquele cartão a ser entregue para aquela pessoa é de um valor X. Pela computação tradicional, determinística, o programa de computador cumpriu o seu papel neste momento, e o resultado R é o próprio valor X. Executar novamente este algoritmo, com os mesmos dados, irá apresentar exatamente o mesmo resultado X, a menos que alguma das variáveis dos dados D sejam modificadas, tais como um aumento de renda ou alguma movimentação bancária distinta.

Este mesmo algoritmo, dotado de IA, também indicaria o mesmo valor X como resultado, mas poderia ser abastecido também com os dados do pagamento de outros usuários com a mesma faixa de renda. Através de uma análise destes dados, ele identificaria que uma porcentagem muito grande de usuários com o mesmo nível de renda do usuário está pagando as parcelas com atraso. Não cabe ao algoritmo, com função definida, fazer julgamentos complexos de qual fenômeno é responsável pela inadimplência, mas sim avaliar sua probabilidade. No mês seguinte, mesmo com o usuário inserindo exatamente os mesmos dados do mês anterior, e sem nenhuma modificação no código do programa, o sistema apresentará um resultado diferente - diminuindo o crédito do consumidor - por aferir matematicamente que a chance de inadimplência seria maior, baseada em dados externos que não são diretamente do usuário, mas de uma situação do ambiente identificada pelo programa.

Uma situação objetiva e analítica como uma avaliação de crédito é um exemplo simples de como a IA pode auxiliar na execução de tarefas que, antes, exigiam a inteligência humana. Porém, é algo próximo da maneira determinística com a qual a computação resolve seus problemas, já que lida com números, faz cálculos objetivos e entrega um resultado também objetivo em sua saída. O amadurecimento da computação trouxe com ele discussões sobre os modos de comunicação da máquina para com seus operadores, que culminaram com as

[15] Segundo o dicionário Oxford, interação é um processo de resolução de uma equação mediante operações em que sucessivamente o objeto de cada uma é o resultado da que a precede.

iniciativas de produção computacional de linguagem natural, entre elas os *Large Language Models* (LLM), objeto do estudo deste trabalho.

2.2 Os *Large Language Models* (LLM)

Diversas inovações tecnológicas têm demarcado eras decisivas na trajetória evolutiva da humanidade, exercendo um impacto notável na sociedade e promovendo alterações fundamentais nos paradigmas produtivos e nas cadeias de valor. Essas tecnologias, ao serem introduzidas, não apenas catalisaram transformações profundas nos métodos de produção, mas também reconfiguraram as relações socioeconômicas, estabelecendo novos axiomas para a interação humana e para a estruturação das sociedades. A implementação dessas inovações tem sido um motor de progresso, impulsionando a humanidade para novos horizontes, ao mesmo tempo em que suscita reflexões críticas sobre as consequências éticas, ambientais e socioeconômicas intrínsecas a esses avanços. Focando-se no aspecto econômico e mercadológico, Hobsbawm (2012) denomina "Revolução Industrial" certos períodos de ruptura abrupta das cadeias produtivas, que, transformadas de modo contundente, abrem novas possibilidades de criação de produtos e modificações drásticas dos modos de produção. O primeiro registro do termo "Revolução Industrial" vem de uma carta, escrita em 1799, na França, que, na época, começava a se industrializar.

Hobsbawm (2012) define quatro marcos transformadores: o primeiro, no século XVIII, chamado de "Primeira Revolução Industrial", a transição do modelo de produção manual para a produção com máquinas. No segundo período, que abrange o final do século XIX até o término da segunda guerra mundial, a "Era do Aço", onde os produtos industrializados mudaram de material, e as redes de transportes ganharam corpo.

No período do pós-guerra até na década de 2010, a substituição gradual dos meios analógicos pelos digitais, a introdução da computação pessoal e o surgimento da internet são consideradas por Rifkin (2012) como a "Terceira Revolução Industrial".

Klaus Schwab (2016), fundador do Fórum Econômico Mundial, no livro "A Quarta Revolução Industrial", aponta que a Inteligência Artificial será a força-motriz do impulsionamento da evolução da indústria do século XXI. Schwab não limita o guarda-chuva do termo a apenas "IA", classificando como causadores desta revolução outros fatores como a

Internet das Coisas[16] (*Internet of Things*, *IoT*), avanços genéticos e robotização mecânica (entre outros), mas coloca a IA como fundamental na integração destes processos, e responsável por uma ruptura na cadeia produtiva.

> We stand on the brink of a technological revolution that will fundamentally alter the way we live, work, and relate to one another. In its scale, scope, and complexity, the transformation will be unlike anything humankind has experienced before. We do not yet know just how it will unfold, but one thing is clear: the response to it must be integrated and comprehensive, involving all stakeholders of the global polity, from the public and private sectors to academia and civil society. (SCHWAB, 2016, p. 15)

Kuhn (1998) introduz o conceito de "paradigma científico": um modelo ou padrão de como o mundo funciona, e como este mundo é visto e aceito pela comunidade científica. Ele argumenta que, dentro deste paradigma, a ciência avança de uma forma normal, acumulando conhecimento e resolvendo os problemas que este paradigma apresenta. Ocasionalmente, surgem anomalias de conhecimento: dados ou fenômenos que não podem ser explicados pelo paradigma existente. Kuhn afirma que há uma resistência natural da comunidade científica – e, consequentemente, da sociedade – em assimilar estes dados novos, que invalidam muitas conclusões anteriores consideradas válidas e causam revoluções no *status quo* da ciência e sociedades vigentes.

> Já vimos que uma comunidade científica, ao adquirir um paradigma, adquire igualmente um critério para a escolha de problemas que, enquanto o paradigma for aceito, poderemos considerar como dotados de uma solução possível. Numa larga medida, esses são os únicos problemas que a comunidade admitirá como científicos ou encorajará seus membros a resolverem. Outros problemas, mesmo muitos dos que eram anteriormente aceitos, passam a ser rejeitados como metafísicos ou como sendo parte de outra disciplina. Podem ainda ser rejeitados como demasiado problemáticos para merecerem o dispêndio de tempo. Assim, um paradigma pode até mesmo afastar uma comunidade daqueles problemas sociais relevantes que não são redutíveis à forma de quebra-cabeça, pois não podem ser enunciados nos termos compatíveis com os instrumentos e conceitos proporcionados pelo paradigma. (KUHN, 1998, p. 60)

[16] A Internet das Coisas (IoT) é uma rede global de objetos com acesso à internet, que transmitem dados e se comunicam entre si. Todos esses objetos são identificados por números distintos e exclusivos, e abrange desede dispositivos móveis até eletrodomésticos e carros. A classificação de IoT também pode incluir pessoas, animais ou outros objetos com sensores incorporados, como medidores contínuos de pressão arterial, rastreadores por GPS inseridos em animais ou sensores em equipamentos para agricultura.

A inteligência artificial pode ser considerada uma quebra no paradigma científico, já que afeta significativamente o mundo hoje em várias áreas, transformando indústrias, influenciando a sociedade e alterando o panorama tecnológico. Carros autônomos já são uma realidade, diagnósticos médicos feitos por IA já são comuns, certos serviços de recomendação e interação com o público já são realizados por inteligência artificial, e muitas outras áreas sofrem a influência desta nova forma de interação entre humanos e máquinas. Como é uma tecnologia nova e em constante evolução, avanços significativos surgem o tempo todo, impactando novos campos e causando o surgimento de novos trabalhos e produtos que, devido ao estado da arte da tecnologia no passado, eram impossíveis de se cogitar. Um dos campos da IA cujo salto evolutivo é mais proeminente é o processamento de linguagem natural (NLP, *Natural Language Processing*). O NLP é uma área de estudo interdisciplinar, que junta ciências da computação e linguística, que tenta definir modos de interação entre humanos e computadores de modo natural, levando em conta as características da linguagem humana, que é complexa e dependente de contexto.

2.3 O Processamento de Linguagem Natural

Desde a invenção dos primeiros computadores, durante a Segunda Guerra Mundial, a busca por aprimorar a interação entre humanos e máquinas tem sido uma questão central que impulsiona inovações no campo da ciência da computação. Estas primeiras máquinas, apesar de revolucionárias, apresentavam desafios significativos em termos de usabilidade e acessibilidade para o usuário comum, evidenciando que um computador, por mais avançado que fosse, tinha sua utilidade limitada se os resultados que produzia não fossem facilmente compreensíveis e aplicáveis.

Reconhecendo essa lacuna, pioneiros como Alan Turing começaram a explorar conceitos fundamentais de inteligência artificial e computação, lançando as bases para as interações mais intuitivas e eficazes entre humanos e máquinas. Desde os primórdios da computação, as máquinas de processamento criadas pelo homem sempre buscaram maneiras de comunicar seus resultados aos usuários. As primeiras invenções utilizavam métodos como lâmpadas de indicação, impressões em papel e cartões perfurados para fornecer saídas de dados, refletindo as limitações tecnológicas da época. A introdução dos monitores de vídeo na informática, por volta da década de 1960, marcou uma revolução na forma como interagimos com computadores. Inicialmente limitados à exibição de texto, esses monitores abriram

caminho para uma interação visual mais rica e direta, facilitada pela evolução subsequente das interfaces gráficas de usuário (GUI, *Graphical User Interface*), que começaram a se popularizar com o *Xerox Alto* (1973), e mais tarde com o *Apple Lisa* (1983) e o *Macintosh* (1984), que trouxeram a interface gráfica para o mercado de massa.

A popularização dos computadores pessoais no final da década de 1970 e, mais notavelmente, durante a década de 1980, deve muito ao desenvolvimento e ao barateamento das telas de monitor. Este avanço transformou os computadores de ferramentas especializadas, acessíveis apenas a empresas e instituições acadêmicas, em dispositivos cotidianos, presentes em lares e escritórios ao redor do mundo. Historicamente, a interação baseada em entrada de texto, facilitada por dispositivos como o teclado, tem sido a forma mais comum de comunicação entre humanos e máquinas. Esta forma de interação permite que os usuários insiram comandos e dados em linguagem natural ou codificada, que a máquina pode então processar e responder de acordo com a programação pré-estabelecida. No entanto, o campo da interação homem-máquina tem visto avanços significativos além da entrada baseada em texto. Recentemente, tecnologias que permitem a entrada de dados através de imagens (reconhecimento visual por câmeras), interações táteis (como telas sensíveis ao toque) e até iniciativas recentes de interface cérebro-computador (*Brain Control Interface, BCI*) demonstram a busca contínua por métodos de interação mais naturais e intuitivos. Estas tecnologias emergentes ainda se apoiam, em muitos casos, na linguagem humana como mediadora, seja por meio de reconhecimento de fala ou pela interpretação de gestos e expressões faciais que imitam a comunicação não-verbal humana.

A jornada da interação homem-máquina, desde os primeiros dispositivos de saída até as sofisticadas interfaces de hoje, reflete um diálogo constante entre as necessidades humanas e as possibilidades tecnológicas. Pioneiros como Ivan Sutherland, com seu *Sketchpad* (1963), considerado o primeiro programa de desenho assistido por computador, e cientistas contemporâneos explorando a fronteira da inteligência artificial e da neurociência, continuam a moldar esse campo dinâmico.

Os computadores modernos operam primordialmente através do processamento de linguagem binária, um sistema de codificação que utiliza apenas dois símbolos, zeros (0) e uns (1), para representar dados e instruções de comando. Esta base binária está no cerne de todas as operações que um computador realiza, desde simples cálculos aritméticos até o processamento complexo de gráficos e a execução de software sofisticado. As exceções notáveis a este paradigma incluem os recentes computadores quânticos, que operam com base em *qubits*,

capazes de representar zeros, uns, ou ambos simultaneamente, graças ao princípio da superposição quântica.

Quando um usuário interage com um computador, ele o faz através de uma interface de usuário que pode variar de simples linhas de comando a complexas interfaces gráficas. Estes comandos são traduzidos em operações binárias que o processador do computador, ou CPU, pode entender e executar. A CPU é o cérebro do computador, realizando bilhões dessas operações básicas por segundo, incluindo cálculos aritméticos, operações lógicas, manipulações de dados em memória, e leituras e escritas em dispositivos de armazenamento perenes.

Esta linguagem binária é difícil de se interpretar para um ser humano acostumado a uma linguagem natural, o que tornava o uso efetivo de computadores restrito a apenas um grupo de entusiastas e estudiosos da área. Um computador era usado apenas por especialistas, para tarefas específicas e tinha suas respostas restritas a tecnologia vigente. Com a crescente disseminação dos computadores na sociedade, surgiu um desafio premente: tornar a linguagem binária, intrínseca à operação dos computadores, acessível e compreensível através da linguagem natural humana. Este impulso levou ao desenvolvimento do campo de processamento de linguagem natural (PLN), uma disciplina na interseção da ciência da computação, inteligência artificial e linguística, dedicada a quebrar as barreiras entre a comunicação humana e a compreensão computacional.

Um computador, embora seja uma criação humana, opera de maneira fundamentalmente distinta do ser humano, especialmente no que tange ao processamento e à geração de linguagem. A compreensão e a produção da linguagem humana, ricas em nuances e complexidades, são temas de investigação e debate entre linguistas, psicólogos, neurocientistas e cientistas da computação. Para abordar como os computadores interpretam e simulam a linguagem humana, é importante compreender os princípios e mecanismos subjacentes da linguagem em si.

A linguagem humana é caracterizada por sua diversidade e capacidade de expressar uma vasta gama de emoções, intenções e informações. Ela não se limita a uma mera coleção de palavras e regras gramaticais; é, ao invés disso, um sistema dinâmico que evolui, influenciado por contextos culturais, sociais e individuais:

> Language is one of the most systematic and complex forms of human behaviour. As such it has given rise to many different theories about what it is used for (thinking vs. communicating), how it has evolved (abruptly or gradually), where its structure comes from (innate structures vs. language use) and what types of processes underlie its structure (those specific to language vs. those applicable in many cognitive domains). […] a consequence of viewing language as a complex adaptive system and linguistic structure as emergent (Lindblom et al. 1984, Hopper 1987) is that it focuses our attention not so much on linguistic structure itself, as on the processes that create it (Verhagen 2002). By searching for domain-general processes, we not only narrow the search for processes specific to language, but we also situate language within the larger context of human behaviour. (BYBEE, 2010, p. 6-7)

Esta complexidade inerente apresenta desafios significativos para a modelagem computacional, que busca replicar a capacidade humana de entender e gerar linguagem de forma que seja ao mesmo tempo precisa, relevante e adaptável a diferentes contextos e intenções. No campo da ciência da computação, particularmente no processamento de linguagem natural (PLN), esforços são feitos para desenvolver algoritmos e sistemas capazes de aproximar a compreensão computacional da linguagem humana. Isso envolve não apenas o deciframento de palavras e frases, mas também a interpretação de significados, a detecção de nuances, a resposta a intenções e até mesmo a compreensão de ambiguidades e metáforas, aspectos que são intuitivos para os seres humanos, mas extraordinariamente complexos para máquinas.

A ciência moderna está dividida em várias teorias sobre como a linguagem funciona e é adquirida, variando desde o inatismo, que sugere que a capacidade para a linguagem é inata no cérebro humano, defendida por Chomsky (1997), até perspectivas construtivistas, que enfatizam o papel da experiência e do ambiente, como a de Tomasello (2003). Esta divisão de opiniões reflete a complexidade da linguagem humana e a dificuldade de replicá-la em sistemas computacionais. Dentre as muitas correntes teóricas, para as discussões desta tese, no aspecto da linguística, cabe citar e discutir duas visões que ajudaram a definir a linguística moderna e contribuíram para a evolução da NLP para os patamares que vemos hoje: os estudos linguísticos de Ferdinand de Saussure, e a Gramática Gerativa de Noam Chomsky.

A obtenção da compreensão mútua entre seres humanos é um desafio enfrentados pela nossa espécie. Nos comunicamos por meio de um mosaico diversificado de idiomas, cada qual ostentando suas peculiaridades intrínsecas, variações regionais e peculiaridades linguísticas. Dentro desse espectro idiomático, distintas localidades manifestam-se através de expressões e dialetos singulares, ao passo que cada indivíduo imbuí seu falar com sua própria essência. Além disso, a língua, revela-se como uma entidade dinâmica e mutável, evoluindo incessantemente

ao sabor das eventualidades temporais. Este panorama de complexidade serve de arena para o estudo aprofundado por parte de um diverso contingente de cientistas buscando decodificar os meandros da linguagem: tal empreitada se configura como uma área de estudo intrinsecamente multidisciplinar, congregando especialistas das mais diversas áreas do conhecimento humano, tais como medicina, biologia, psicologia, filosofia, entre outras disciplinas científicas, todas convergindo em algum momento para o estudo da linguística e suas ramificações. Este campo vasto e heterogêneo de investigação desdobra-se na tentativa de elucidar não apenas os mecanismos fisiológicos e cognitivos subjacentes à fala e à compreensão, mas também na busca por compreender as nuances filosóficas e culturais que permeiam a capacidade humana de se comunicar e se fazer entender num mundo cheio de diversidades linguísticas.

Diversas teorias linguísticas buscam decifrar as origens e o desenvolvimento da linguagem humana, mas a Gramática Gerativa, proposta por Noam Chomsky (1965), destaca-se especialmente por sua afinidade com os princípios matemáticos subjacentes à criação de linguagem artificial. Essa teoria, que se baseia em uma abordagem matemática da sintaxe, foi posteriormente expandida para incluir análises das estruturas semânticas (1980), oferecendo um modelo abstrato que visa catalogar as combinações ilimitadas de fonemas e palavras que formam uma língua. Esta teoria sugere que, ao identificar um conjunto universal de regras gramaticais, é possível aplicar algoritmos computacionais para processar e gerar linguagem de maneira automática. Tal perspectiva abre caminho para o desenvolvimento de sistemas capazes de compreender e produzir conteúdo linguístico de forma autônoma, refletindo o potencial de aplicar conceitos de Gramática Gerativa no campo da inteligência artificial e da geração automática de texto. Devido a esta conveniência, os algoritmos que se propunham a trabalhar com linguagem natural baseavam-se na noção de que o entendimento da linguagem era passível de ser trabalhado com regras de programação, onde os programadores criavam regras intrincadas e complexas de padrões linguísticos, e estes códigos geravam a linguagem de modo semelhante ao que o humano faz.

Desde as concepções de Noam Chomsky na década de 1950, a Gramática Gerativa se posicionou como uma influência dominante nas primeiras tentativas de desenvolver algoritmos para o processamento de linguagem natural. Chomsky propôs que a capacidade humana para a linguagem é inata e se baseia em regras universais subjacentes a todas as línguas, uma ideia que prometia fornecer uma base sólida para a modelagem computacional da linguagem. Contudo, apesar da promessa teórica, os esforços iniciais para aplicar os princípios da Gramática Gerativa ao PLN enfrentaram desafios significativos, resultando em um sucesso limitado. Essas

dificuldades podem ser atribuídas, em parte, ao estado nascente da ciência da computação na época: considerando que o primeiro computador moderno surgiu apenas na década de 1940, a tecnologia e as técnicas de programação disponíveis para os pioneiros do PLN eram relativamente primitivas e estavam restritas pelas capacidades tecnológicas limitadas dos primeiros computadores.

As tentativas de implementar a complexa e abstrata teoria da Gramática Gerativa em sistemas computacionais esbarraram na limitação dos recursos de hardware e software. Os algoritmos da época, embora inovadores, lutavam para processar a complexidade e a ambiguidade inerentes à linguagem humana, uma vez que as máquinas daquele período possuíam capacidades de processamento, memória e armazenamento extremamente limitadas. Além disso, a abordagem predominantemente baseada em regras exigia uma formalização rigorosa de gramáticas complexas, algo que se mostrou imensamente desafiador na prática.

Inicialmente, muitos desses projetos focavam em tarefas como a tradução automática entre línguas, mas enfrentavam obstáculos significativos: eles tinham dificuldade em processar palavras homógrafas — palavras escritas da mesma forma, mas com significados diferentes — e em interpretar usos linguísticos complexos, como o sarcasmo e a ironia. Além disso, havia um desafio considerável em lidar com as nuances e peculiaridades específicas de cada idioma. Essas limitações destacam os desafios iniciais na aplicação de teorias linguísticas abstratas ao desenvolvimento de sistemas computacionais capazes de entender e produzir linguagem de maneira sofisticada e precisa.

Os primeiros passos para tentar sistematizar uma programação livre das amarras tecnológicas vigentes na época passou pela criação da metassintaxe[17] de gramática livre de contexto, estruturada por John Bakus e Peter Naur, chamada de “Formalismo de Backus-Naur” (*Backus-Naur Form*, ou BNF). A BNF é uma notação utilizada para expressar a sintaxe de linguagens de programação, protocolos de comunicação e sistemas de formatação de texto ou dados. Foi desenvolvida inicialmente por Backus para descrever a linguagem ALGOL 60, e posteriormente modificada por Naur para a publicação final da sintaxe da linguagem ALGOL, razão pela qual recebeu o nome de ambos. A BNF é usada, principalmente, para descrever formalmente a gramática de linguagens de programação, e é composta por uma série de regras de produção, onde cada regra define uma estrutura sintática dentro da linguagem. Ela permite

[17] Metassintaxe refere-se ao uso de variáveis simbólicas para descrever a sintaxe de uma linguagem de programação ou comando, facilitando a explicação ou documentação de padrões de código.

que os desenvolvedores e analistas compreendam as regras sintáticas de uma linguagem de forma unívoca, facilitando a implementação de compiladores e interpretadores.

Muitas variáveis separam as correntes de pensamento de Saussure e Chomsky, e, dentre elas, o tempo: Saussure viveu apenas os primeiros 13 anos do século XX e não presenciou o surgimento dos computadores, e Chomsky, nesta data (2024) está vivo. A contemporaneidade de Chomsky permitiu que o autor aperfeiçoasse suas teorias de linguística e contribuísse mais com o campo de estudo desta tese. Também cabe ressaltar o fato de que Saussure nunca publicou nada exatamente sobre a linguística em si, sendo as análises científicas pertinentes a este trabalho derivadas dos livros e artigos escritos por cientistas derivados de suas teses, tempos depois de seu falecimento.

Saussure é um dos elos entre a computação e a semiótica abordados neste trabalho. Antes de Saussure, os estudos de linguística tinham um viés distante: Não visavam a análise da língua em si, mas dos seus efeitos. Os estudos da língua como objeto partem das iniciativas científicas de Saussure, considerado por muitos o “Pai da Linguística Moderna”; seus estudos sobre semântica e pragmática são bases para a semiótica, e, em certo ponto, sustentam suas teses sobre a linguística, que, na observação de Saussure, é uma espécie de subárea da semiologia, pois se preocupa apenas com a linguagem humana, sendo menos abrangente que a semiótica, que se estende a todos os sistemas de comunicação.

Dentro do escopo da linguística, Saussure propôs um viés estruturalista da linguagem: dotada de um sistema centralizado, embasada em um significante (a expressão material de uma palavra) e um significado (o que ela representa). Distanciou a visão puramente histórica da língua, e propôs que as linguagens se baseavam em uma estrutura comum, independentemente do idioma em que esta língua estivesse. Isso é importante para o estudo das linguagens computacionais porque, pela primeira vez, falou-se na possibilidade de a linguagem ser estruturada, e, consequentemente, capaz de ser representada através de algoritmos de programação. Se existem pontos gerais que acompanham todas as línguas, basta descobrir quais e reproduzi-los, para que uma máquina reproduza de modo convincente as linguagens humanas.

A obra de Saussure (2006) levanta inúmeras análises sobre signos, sobre o princípio de arbitrariedade, sobre o fato da língua (*langue*) se aproximar do lado social e a fala (*parole*) se aproximar do lado individual, entre outras. No âmbito da linguística estruturalista, o conceito bifurcado de *langue* e *parole* constitui um pilar fundamental para a compreensão da linguagem enquanto fenômeno tanto social, quanto individual. Esta dicotomia delineia a distinção entre,

por um lado, o sistema abstrato e normativo da linguagem – *langue,* a língua – e, do outro, o uso concreto e individual deste sistema na comunicação cotidiana – *parole,* a fala.

A *langue*, para Saussure, é concebida como o conjunto de convenções e regras compartilhadas que possibilitam a comunicação dentro de uma comunidade que partilhe estes códigos. Representa a essência imutável e coletiva da linguagem, um código que precede e transcende o indivíduo. No contexto de um LLM, o conceito de *langue* pode ser equiparado ao vasto conjunto de regras, padrões e estruturas linguísticas que o modelo absorve durante o seu treinamento. Essa absorção é o resultado da exposição do modelo a um corpus extenso e diversificado de textos, oriundos de diversas fontes, que permite ao modelo assimilar a *langue* de várias comunidades linguísticas.

Por exemplo, ao solicitar a um modelo de LLM o resumo do livro "Vidas Secas", de autoria de Graciliano Ramos, o modelo usa fundamentos linguísticos sintetizados durante seu treinamento. Através do padrão lexical e semântico da língua portuguesa, o modelo entende que "Vidas Secas" é o título de um livro, associado a Graciliano Ramos, e está associado a vida de uma família de retirantes peregrinando pelo sertão do nordeste brasileiro, e não sobre um livro sobre vidas que se secaram. Tal associação só é possível através de várias conexões internas, com informações dispostas em vários textos diferentes, que fazem o modelo associar a entrada a um contexto específico. Para apresentar o resultado do pedido, o resumo propriamente dito, o LLM deve ser dotado do conhecimento gramatical necessário para estruturar frases e formar um resumo coeso e coerente, e ainda saber reconhecer e reproduzir padrões de resumos de livros, como por exemplo iniciar a saída de texto com uma apresentação geral, seguida de pontos principais da trama, e encerrar com uma conclusão sobre a importância do livro. Este regramento serve para todas as saídas de texto por ventura solicitadas pelo usuário, tais como poesias, e – assunto desta tese – roteiros audiovisuais.

O texto gerado pelo LLM é a *parole*: cada instância da comunicação produzida pelo modelo é um ato de *parole* – uma realização singular e concreta do vasto repertório linguístico que o modelo possui. A resposta demonstra a capacidade do modelo de utilizar seu entendimento abstrato da *langue* para criar uma resposta específica e contextualizada da solicitação feita pelo usuário. O modelo seleciona e organiza informações relevantes sobre o libro, estrutura estas frases de forma lógica e apresenta um resumo que reflete o entendimento das nuances do enredo e das temáticas do livro:

> Entrada: Você pode resumir o livro Vidas Secas para mim?
>
> ChatGPT: "Vidas Secas" é uma obra do escritor brasileiro Graciliano Ramos, publicada em 1938, e é um dos romances mais representativos da segunda fase do Modernismo no Brasil, marcada pelo regionalismo. O livro narra a história de uma família de retirantes sertanejos composta por Fabiano, Sinhá Vitória, seus dois filhos e a cachorra Baleia, que lutam pela sobrevivência e por uma vida melhor em meio à seca do sertão nordestino. A narrativa é dividida em treze capítulos, que podem ser lidos de forma independente, mas que juntos compõem um retrato coeso da vida difícil no sertão. O livro aborda temas como a miséria, a opressão, a luta pela sobrevivência, a injustiça social, e a relação do homem com a terra e com os animais.
>
> Os personagens principais são profundamente marcados pela aridez do ambiente em que vivem. Fabiano, o pai, é um vaqueiro bruto e pouco articulado, que luta para manter sua família viva, mas frequentemente se vê impotente diante da natureza e do sistema social. Sinhá Vitória sonha com uma vida melhor para seus filhos e uma cama de verdade para substituir o antigo baú onde dormem. Os filhos do casal, meninos ainda sem nome, simbolizam a inocência e a esperança de um futuro diferente. Baleia, a cachorra, é tratada com grande humanidade pela narrativa, e sua morte é um dos momentos mais emocionantes do livro. O estilo de Graciliano Ramos em "Vidas Secas" é caracterizado por sua economia de palavras, precisão linguística e profundidade psicológica. Ele usa uma linguagem simples, porém poderosa, para retratar a vida dura no sertão, fazendo críticas sociais e políticas ao mesmo tempo em que revela a riqueza humana de seus personagens.
>
> "Vidas Secas" é uma obra fundamental da literatura brasileira, retratando não apenas a luta pela sobrevivência em um ambiente hostil, mas também a busca por dignidade e justiça social.

A concepção de Saussure e Chomsky acerca da predisposição genética do homem para a aquisição e o desenvolvimento da linguagem manifesta-se através de prismas teóricos consideravelmente divergentes. Chomsky é um defensor da hipótese que postula a existência de uma propensão inata ao ser humano para a linguagem, perspectiva esta que é encapsulada na sua teoria da Gramática Universal:

> Languages, therefore, resemble men in this respect, that, though each has peculiarities, whereby it is distinguished from every other, yet all have certain qualities in common. The peculiarities of individual tongues are explained in their respective grammars and dictionaries. Those things, that all languages have in common, or that are necessary to every language, are treated of in a science, which some have called Universal or Philosophical grammar. (CHOMSKY, 1965, p. 5)

Chomsky sugere que a capacidade de adquirir linguagem é inata ao ser humano, e que existe um conjunto de regras e princípios predefinidos que governam a estrutura de todas as línguas humanas. Essencialmente, Chomsky argumenta que todos os seres humanos nascem com uma predisposição para aprender linguagem, e que esta predisposição é a mesma, independentemente da língua a que este ser humano é exposto.

A teoria de Chomsky propõe que, apresar da diversidade aparente entre as línguas existentes, no nível mais profundo, elas compartilham uma estrutura comum. Esta estrutura comum permite, por exemplo, que crianças adquiram rapidamente a sua língua nativa, usando sua capacidade inata de linguagem, mesmo com uma exposição limitada (e, ocasionalmente, imperfeita) à sua língua materna. A existência dessa gramática universal contribui para explicar como que as crianças são capazes de aprender uma língua complexa de maneira relativamente rápida e sem instrução formal. A teoria da gramática universal tem sido fundamental para o desenvolvimento da linguística teórica e influencia estudos em áreas como o aprendizado de uma segunda língua, neurociências e o aprendizado de linguagem natural por máquinas. Logo, segundo Chomsky, todos os seres humanos são passíveis de aprender uma linguagem e todas as línguas derivam de um estado inicial uniforme. Se é este o caso, um algoritmo de geração de linguagem torna-se viável, já que basta ao desenvolvedor deste algoritmo possuir domínio sobre os dados gerativos da linguagem alvo.

Enquanto Chomsky enfatiza a importância da biologia e da genética como fundamentos para a capacidade humana de adquirir linguagem, Saussure privilegia a análise da linguagem como um sistema de signos, e se atém as convenções que regem este sistema, abstraindo-se das questões genéticas inerentes à aquisição linguística. A contribuição de Saussure desvia-se de questões sobre a predisposição genética, focando-se, ao invés disso, nas características estruturais e funcionais da linguagem dentro do contexto social. A visão de Chomsky acerca da predisposição genética para a linguagem é feita de forma explícita e constitui o alicerce para sua teoria da gramática universal, enquanto Saussure oferece uma perspectiva mais estruturalista e sociocultural da linguagem, omitindo uma análise profunda dos aspectos genéticos subjacentes à aquisição da linguagem. A somatória de ambas, levando em conta seus possíveis antagonismos, é útil para análise de algoritmos de geração de texto.

A teoria de Chomsky é conveniente para a criação de algoritmos de linguagem natural. Chomsky propõe que a língua é composta de um conjunto de sentenças, com lógica própria, inerente ao cérebro humano e, logo, qualificada a uma “algoritmização”, já que a gramática da língua é considerada livre de contexto.

Inclusive tal designação de Chomsky foi a maior fomentadora da criação do Formalismo de Backus Naur, usado para descrever a *ALGOL*, família de linguagens de programação de alto nível no final da década de 1960. Como a gramática gerativa de Chomsky foi elaborada no mesmo período em que as surgiram as primeiras iniciativas de programação de NLP, as tentativas iniciais de entendimento de linguagem natural através da computação se deram sob este prisma, quando, por algoritmos, humanos tentavam aferir, criar e programar estas regras gramaticais por softwares. Em um processo de retroalimentação, inclusive, as próprias linguagens de programação eram influenciadas pelas técnicas recém-criadas de NLP, e vice-versa.

No entanto, tal abordagem se mostrou, na prática, insuficiente, por várias razões. O próprio poder computacional era incipiente, os estudos de Chomsky estavam no início – prova disto é que sua teoria sofreu diversas revisões e sofre críticas até hoje. A língua natural é bastante extensa, e, para complicar, cada local tem a sua. As regras gramaticais são ambíguas e diferem significativamente entre os idiomas, e desenvolver algoritmos que levassem em conta todos estes regramentos era uma tarefa muito complexa. A função da NLP, fundamentalmente, é encontrar semântica na linguagem recebida ou produzida, e essas semânticas quase sempre exigem contextos implícitos, que, para nós, humanos, são automáticos, mas para um computador, não.

Por exemplo, regras gramaticais indicam que o verbo “queimar” só deve acompanhar substantivos que, de fato, podem ser queimados com o uso de fogo. Uma regra condicional em um algoritmo precisa ser criada para representar tal situação, e junto a ela outras milhões de regras diferentes para elencar outras inúmeras situações que se apliquem, e estas próprias novas regras passariam a interferir umas nas outras, causando novos problemas. Embora, com certa paciência e persistência, estes obstáculos sejam parcialmente transponíveis – eles geravam algoritmos extremamente lentos, ineficazes, e, principalmente, com resultados que, embora estivessem sintaticamente corretos, não se aproximavam da linguagem orgânica. Inclusive isso pode ser observado pela própria imagem mental que nós temos sobre robôs, estereotipados como lacônicos, diretos e inorgânicos.

A criação do transistor, o advento dos microprocessadores, o crescimento da computação pessoal, juntamente com a evolução do pensamento de Chomsky, que refinou e aperfeiçoou sua teoria conforme os anos passaram, reorientaram a NLP para outra abordagem: a NLP estatística.

2.4 A NLP Estatística e os *Large Language Models* modernos

O aprendizado de máquina, a estocástica[18], a probabilidade, a aproximação, passaram a ser proeminentes nos estudos de linguagem natural. O próprio Chomsky foi inicialmente reticente ao uso de modelos probabilísticos para representação da linguagem, passou a levar em conta esta possibilidade (CHOMSKY, 2023). A digitalização do acervo textual da humanidade potenciou o uso de vastos bancos de dados de textos diferentes para treinarem grandes volumes de dados, e modelos probabilísticos passaram a, agora abastecidos com muitas informações, a gerarem textos mais orgânicos.

Essa transformação no panorama da NLP, de predominância estatística e com o uso de IA, iniciou-se na década de 80 e ganhou tração por vários fatores favoráveis simultâneos: o surgimento da computação em nuvem, que barateou o custo de processar grandes quantidades de dados sem a necessidade de se adquirir computadores próprios, a internet, que é um repositório aberto de boa parte do conhecimento humano, os avanços dos estudos sobre IA, e as evoluções nos estudos da própria linguística como ciência, entre outros. Tudo isso culminou nos modernos modelos de LLM (*Large Language Models*), como o *ChatGPT*, que estão, hoje, nas discussões centrais sobre a evolução da IA, e, pelo seu poder de dar saída à padrões de texto de forma aparentemente natural, tornam possíveis as discussões sobre a geração de conteúdo audiovisual através de Inteligência Artificial.

Segundo Croft e Lafferty (2013), um modelo de linguagem é uma distribuição estocástica de uma fila de palavras, com uma indicação da chance de a sequência destas palavras serem válidas, ou não. Mas esta validade não indica, necessariamente, um endosso gramatical, mas sim uma análise probabilística. Em resumo, um modelo de linguagem indica qual a porcentagem de chance de que uma palavra esteja junto de uma outra, ou qual conjunto de palavras vem depois do conjunto que o modelo obteve como entrada, ou já gerou antes. Um Grande Modelo de Linguagem (LLM, no acrônimo em inglês) é, como o próprio nome diz, um modelo com muitas entradas, onde o banco de dados usado para gerá-lo é extenso. Quanto mais texto é usado para treinar um modelo, maior sua eficiência em imitar a linguagem será.

É importante ressaltar o verbo "imitar", já que é o resultado de um modelo de linguagem é, sempre, uma saída estatística de probabilidades. Não há, em qualquer resultado vindo de um

[18] A estocástica é o estudo de padrões, processos e fenômenos que envolvem aleatoriedade e incerteza, aplicável em áreas como computação, matemática, finanças e ciências naturais para entender comportamentos imprevisíveis.

LLM, nenhum tipo de relação de contexto. Porém, a saída de texto de alguns algoritmos de LLM como o *ChatGPT* são bastante competentes em imitar uma conversação humana, o suficiente para causar a impressão de que existe, ali dentro deste algoritmo, alguma espécie de noção de contexto. Essa noção não existe de fato, mas é emulada através de modelos matemáticos de probabilidade, estratégias de armazenamento de memória e, principalmente, através de mecanismos de atenção algorítmica.

Dentre os vários modelos de LLM disponíveis, o *ChatGPT* é, provavelmente, o mais famoso. É um modelo de linguagem treinado com milhões de caracteres extraídos da internet, e tem o objetivo de gerar fragmentos textuais que se assemelhem aos escritos por um humano. É construído através de computação distribuída, usando a arquitetura de *transformers*, que atualmente (março de 2024) está em sua quarta versão, que é dezenas de vezes maior e mais treinada que a anterior. Essa característica de ser ancorado em uma arquitetura distribuída permite que seus desenvolvedores treinem modelos cada vez maiores e mais eficientes, tornando viável o seu uso para produção de linguagem natural com uma base de dados sempre mais robusta. O modelo do *ChatGPT* é, como o próprio nome indica, baseado em GPT, é a terceira geração de uma rede neural generativa, e seu acrônimo deriva de *Generative Pre-Trained Transformer*.

Generative vem da capacidade do algoritmo de gerar conteúdo. *Pre-Trained* deriva do fato de que o modelo de linguagem é treinado previamente: o modelo é fechado e não aprende com os textos que gera, tampouco com as devolutivas dos usuários. Todos os dados são oferecidos ao algoritmo, compilados, transformados em *tokens*[19], processados, indexados em tabelas, e só depois são utilizados na construção do *chatbot*.

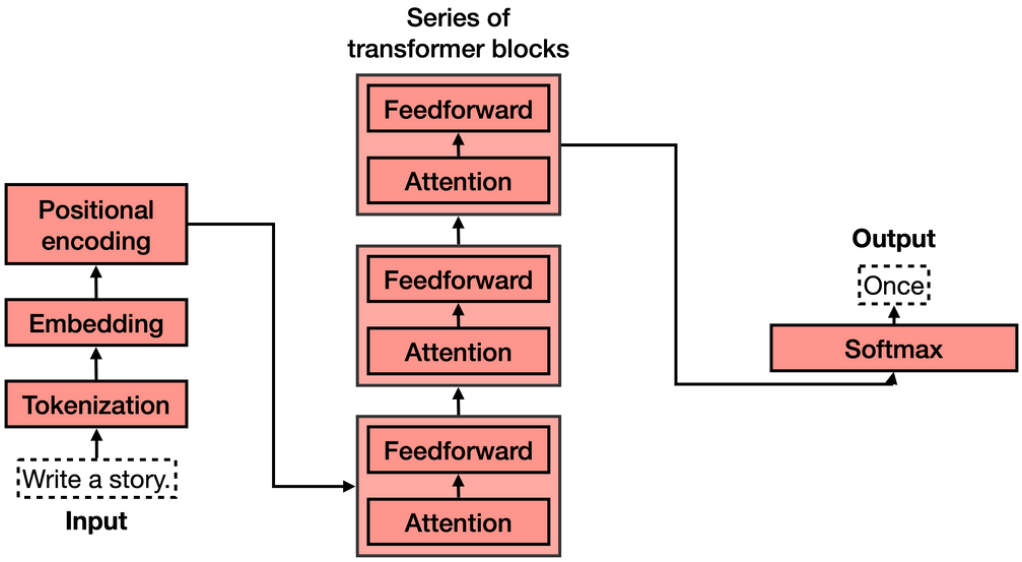
Para entender o último fragmento do acrônimo, o *transformer*, é necessário entender os mecanismos de atenção, aplicados a NLP. Em uma conversação entre dois seres humanos, o foco de atenção a pontos do discurso é variável: quando um humano ouve a frase “Meu filho foi a escola ontem, discutiu com um colega e voltou para a casa com um olho roxo”, podemos aferir certas variáveis sobre como essa informação chegou até o destinatário, tais como:

O discurso é sempre linear, logo, as palavras foram ditas na ordem em que foram escritas, uma a uma. A importância das palavras para o pleno entendimento do discurso muda

[19] *O processo de transformação em tokens (muito usado através do anglicismo “tokenização”), aplicado aos modelos GPT, é o ato de se substituir palavras ou frases por índices, porém, mantendo suas informações essenciais.*

conforme novas palavras são ditas. Por exemplo, após ouvir que o filho do interlocutor "voltou", o ouvinte pode concluir que ele voltou da escola citada palavras atrás.

Figura 12 - Modelo algorítmico da representação de como o conceito de Transformers se aplica a um modelo de LLM



Fonte: Geração do próprio *ChatGPT*

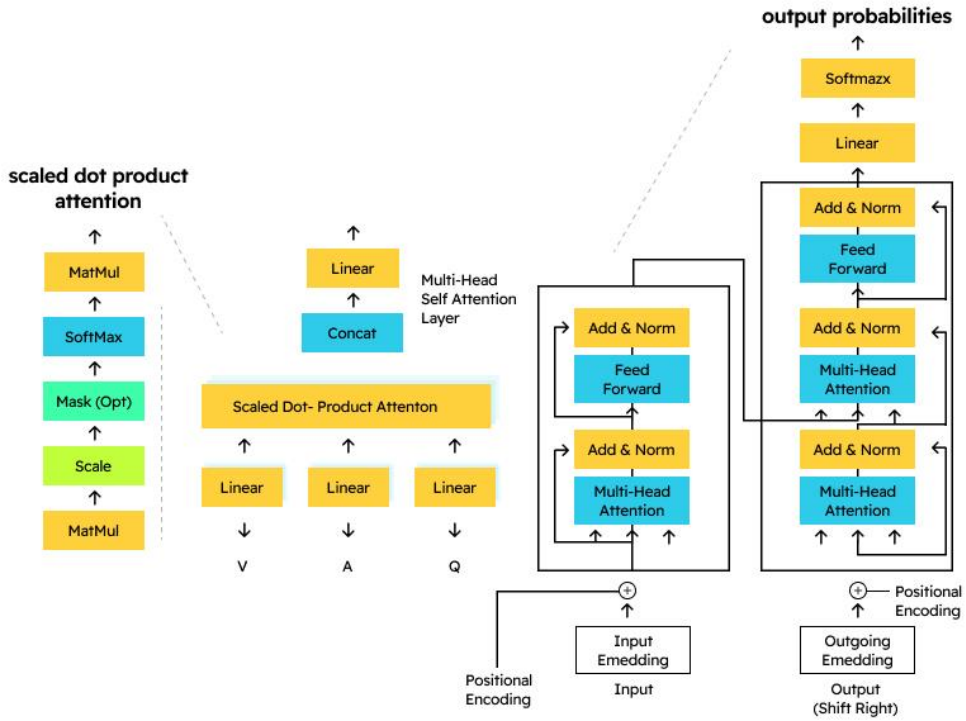
É notório que a atenção se foca em diversos trechos do discurso, conforme a conversação evolui. Em certos pontos, quem pretende entender o discurso precisa voltar a termos ditos anteriormente, e considerá-los juntamente com os novos termos, para compor o cenário correto e atribuir o contexto devido. Também é importante notar que certos fragmentos do texto são menos importantes – e, portanto, menos dignos de atenção – tais como o artigo "A" antes da palavra "Escola", que é uma informação descartável, ou facilmente deduzível, para o futuro entendimento. Fazer esta análise discursiva é natural para um ser humano, mas extremamente complexa para uma máquina.

Antes de elaborar sobre estas dificuldades, cabe ressaltar um fato importante: os modelos de linguagem natural disponíveis hoje são modelos preditivos, que tem sua gênese em esforços para tradução de palavras, e evoluíram para a criação de texto. Isso é significativo porque o *ChatGPT* (e modelos equivalentes em tecnologia), criam as palavras de modo linear – e isso pode ser verificado no uso da ferramenta disponível no site da OpenAI, onde é possível observar as palavras sendo geradas, uma após a outra, de modo linear. O modelo gera a palavra seguinte baseado nas palavras que gerou anteriormente, através de modelos probabilísticos, e assim prossegue até terminar toda a geração que, matematicamente, julgar adequada. É o mesmo princípio do sistema de predição automática da próxima palavra digitada nos serviços

de busca da internet. Então, como que a saída do *ChatGPT* se assemelha tanto a uma conversação orgânica, se tudo que o modelo faz é elencar a probabilidade de uma palavra vir após a outra?

O processo de treinamento de um LLM é relativamente simples: uma quantidade de texto entra no algoritmo, é analisada, catalogada e arquivada. Na catalogação, o algoritmo define valores numéricos para palavras ou sequências de palavras, mas sempre de acordo com a referência de outras palavras, gerando uma matriz multidimensional numérica proporcional a quantidade de entradas avaliadas.

Figura 13 - Arquitetura do ChatGPT



Fonte: Disponível em https://www.appstudio.ca/blog/what-is-*ChatGPT*-how-does-it-work-is-there-any-practical-use-of-*ChatGPT* Acessado em 24 Fev. 2024

Com isso, é possível aferir, numericamente, a "distância" entre duas palavras em um eixo cartesiano hipotético, e fazer cálculos de probabilidade com elas: em um exemplo, se a entrada do modelo for "Quais animais vivem no mar?", o modelo checará, baseado em todos os textos em que foi treinado, quais palavras que estão relacionadas ao contexto "animais", e "mar", e tentará colocá-las em uma resposta coerente. O algoritmo retornará valores decimais entre -1 (absolutamente fora do contexto) e 1 (totalmente dentro do contexto) para todas as

palavras em relação a todas as demais, e estes resultados serão usados para calcular qual a próxima palavra a ser inserida na resposta.

Como o processo é linear, o foco da atenção do algoritmo precisa variar conforme a leitura da entrada de texto avança. Em determinado momento do exemplo acima, o foco estará na palavra “animais”, e o algoritmo buscará, no seu banco de dados, quais palavras se relacionam mais a animais. No caso, encontrará as palavras:

- CAVALO, com um score de 0.925 perante a palavra “animal”
- BOI, com um score de 0.913 perante a palavra “animal”
- PREGUIÇA, com um score de 0.792 perante a palavra “animal”
- PEIXE, com um score de 0.899 perante a palavra “animal”
- BALEIA, com um score de 0.782 perante a palavra “animal”

Provavelmente o algoritmo encontrará muitas outras palavras nesta mesma busca. Cada uma delas é classificada através da sua pontuação, e todas que ultrapassam um certo score são promovidas para a próxima etapa, onde, eventualmente, o algoritmo voltará sua atenção para a palavra “mar”. A pontuação de CAVALO, BOI e PREGUIÇA perante a palavra “mar” é muito menor do que a pontuação das palavras PEIXE e BALEIA, e o algoritmo pode aferir, através de matemática, que as palavras mais adequadas para constar em uma resposta a questão original são PEIXE e BALEIA, em detrimento a CAVALO, BOI e PREGUIÇA, por poder fazer a análise de modo multidimensional.

Este exemplo, embora simplificado, apresenta um dos dilemas da conversação via LLM: Há, de fato, uma valorização do contexto, mas é completamente matemática e distinta do modo de que um humano leva em conta o contexto em uma conversação. A criação deste contexto virtual e matemático se dá através da técnica conhecida como *Word2Vec*, que é baseada nos trabalhos do linguista John Rupert Firth, na década de 50. Ele é um dos cientistas que criaram a Hipótese Distribucional, em que o contexto de uma palavra se dá pela relação que ela tem com as outras palavras que compõem a sentença:

> The placing of a text as a constituent in a context of situation contributes to the statement of meaning since situations are set up to recognize use. As Wittgenstein says, 'the meaning of words lies in their use. The day-to-day practice of playing language games recognizes customs and rules. It follows that a text in such established usage may contain sentences such as 'Don't be such an ass!', 'You silly ass!', 'What an ass he is!' In these examples, the word ass is in familiar and habitual company, commonly collocated with you silly-, he is a silly-, don't be such an-. You shall know a word by the company it keeps! One of the meanings of ass is its habitual collocation with such other words as those above quoted. Though Wittgenstein was dealing with another problem, he also recognizes the plain face-value, the physiognomy of words. They look at us! 'The sentence is composed of words and that is enough'. (FIRTH, 1957)

O *word2vec* é uma representação algorítmica da teoria da Hipótese Distribucional de Firth, onde a soma de quantas vezes uma palavra específica aparece no mesmo contexto de outras palavras específicas é quantificada e armazenada em vetores matemáticos. São estes valores numéricos que são utilizados pelos algoritmos de IA para calcularem quais palavras devem ser apresentadas perante as entradas do usuário, e são a chave para que o *chatbot* consiga simular com um nível satisfatório de acerto o nível de contexto adequado para uma conversação soar semelhante a uma linguagem humana.

A eficiência da simulação de contexto de um modelo que utilize o *word2vec* depende da quantidade de informação que o abasteceu no início do seu aprendizado. Quanto mais texto o modelo receber, mais eficiente em entregar um contexto ele será. É por isso que os modelos de linguagem atuais como o *ChatGPT* são chamados de "*Large Language Models*" – São modelos treinados com bilhões de parâmetros, retirados de fontes como a Wikipedia, e só são possíveis porque a internet foi criada e grandes fontes de documentos abertos foram disponibilizadas.

Toda a gênese do algoritmo que culminou no *ChatGPT* é puramente científica, no sentido de diferir entre o trabalho típico no desenvolvimento de um programa de computador. Comumente, são programadores que são responsáveis pela criação da plataforma, entrada de dados, e resultados. Diferentemente de muitos outros softwares, muitos artigos científicos foram publicados, com técnicas revolucionárias de processamento de texto, que, somados, possibilitaram o desenvolvimento final da plataforma. Por ser um *software* desenvolvido de uma maneira pouco convencional, é razoável que a análise dos resultados tenha uma ênfase acadêmica mais profunda.

Segundo Chowdhury (2003), os pesquisadores que trabalham com NLP precisam ganhar conhecimento de como humanos entendem e se apropriam da linguagem, para que

ferramentas e técnicas apropriadas sejam desenvolvidas para fazer computadores entenderem e produzirem conteúdo que faça sentido e cumpram seu propósito.

> The foundations of NLP lie in a number of disciplines, viz. computer and information sciences, linguistics, mathematics, electrical and electronic engineering, artificial intelligence and robotics, psychology, etc. Applications of NLP include a number of fields of studies, such as machine translation, natural language text processing and summarization, user interfaces, multilingual and cross language information retrieval (CLIR), speech recognition, artificial intelligence and expert systems, and so on. (CHOWDHURY, 2003)

Estamos diante, então, de duas vertentes bem definidas: o lado das exatas, com os estudos de linguagem natural aplicados a computação, onde cálculos, processamentos, memória, banco de dados, são assuntos, e o lado da linguística e da enunciação, onde o estudo da produção e da interpretação do discurso são contemplados. O GPT-4 (e outros algoritmos de IA que se utilizam de NLP), tem como características o acesso, catalogação e indexação de grande quantidade de informação, a análise dessas informações, perante os parâmetros indicados pelo programador da IA, e o resultado.

Estas etapas são análogas às etapas humanas de produção do discurso; o humano aprende com o mundo que o rodeia, analisa e tira suas conclusões sobre aquela situação específica, e, apropriando-se da linguagem que aprendeu, comunica-se, esperando ser entendido pelo enunciatário. A GPT-3 conta com mais de 175 bilhões de parâmetros diferentes para aprendizado de máquina, mais de 15 vezes superior aos modelos concorrentes, o que torna essa IA altamente capaz de gerar textos e imagens de modo convincente, para que um humano não perceba que a fonte geradora do conteúdo é uma máquina.

Em um cenário onde o consumo de mídia está cada vez maior, mais individualizado e mais acessível, e considerando os custos envolvidos na produção e uma maior fragmentação dos dividendos dos produtos, observar o surgimento de um produto de inteligência artificial que gere linguagem natural de modo a ser indistinguível ao que uma pessoa produz leva ao questionamento de se não seria viável usar o *ChatGPT* (e outras ferramentas) para a produção de produtos audiovisuais sem a necessidade da intervenção humana.

Esta tese busca investigar esta possibilidade, analisando o projeto artístico *Nothing, Forever*, uma sitcom gerada proceduralmente e transmitida ao vivo, vinte e quatro horas por dia, por streaming, na internet. Criada como uma paródia da *sitcom Seinfeld*, seu roteiro é

gerado usando GPT-3 (não exatamente o *ChatGPT*, mas seu motor), que renderizam as imagens baseadas neste roteiro através de motores de processamento gráfico em tempo real (no caso aqui, o motor *Unity*). Várias outras técnicas de automação são usadas para sustentar o produto, e, juntas, conseguem obter sucesso em gerar um produto audiovisual contínuo e infinito.

É importante ressaltar que o objeto de análise não deve ser encarado essencialmente como uma representação fidedigna das possibilidades da tecnologia, pois é pioneiro, dotado de baixo orçamento, e, principalmente, experimental. *Nothing, Forever* tende a servir como ponte para as possibilidades de exploração da criação automática de obras audiovisuais, e deve estimular outros produtos derivados de tecnologias semelhantes. É pertinente, para esta tese, observar algumas características base deste produto que se espera encontrar no futuro em outros projetos, e, principalmente, correlacionar as peculiaridades relativas aos trabalhos derivados da inteligência artificial com as expectativas que o seu público possui das típicas obras audiovisuais.

O impacto cultural que a disponibilização do *ChatGPT* causou é considerável, e, observando a evolução de outras iniciativas com efeito semelhante na cultura de massa, é possível afirmar que teremos muitas tentativas de se explorar a Inteligência Artificial como criadora de produtos.

3 OS LLM E A GERAÇÃO DE ROTEIROS AUDIOVISUAIS

Seja um espetáculo teatral, um capítulo de novela, ou um programa de televisão, salvo exceções como espetáculos totalmente improvisados, o roteiro é um documento textual de caráter narrativo que é usado como base para a materialização de uma obra. Este documento descreve a história, diálogos, ações, cenários e outros elementos visuais, e desempenha um papel crucial, fornecendo um alicerce conceitual que orienta o processo organizacional as decisões criativas, logísticas e técnicas tomadas por todos os profissionais envolvidos na realização da obra. Um roteiro audiovisual é composto por elementos comuns, que podem existir ou não, dependendo da mídia, formato ou escolhas criativas do responsável pelo mesmo – ou responsáveis, já que um roteiro pode ser escrito por várias entidades diferentes.

Um roteiro possui várias camadas de informação na sua elaboração. Embora a ordem destas camadas varie, e nem todas precisam estar presentes, é possível identificar quatro elementos que podem compor um roteiro: a descrição da cena, a descrição das ações, os diálogos, e as informações para a equipe de produção.

A descrição das cenas, que abrange os limites espaciais e temporais, fornece informações precisas sobre onde e quanto a ação ocorre. Este aspecto do roteiro também contém descrições detalhadas de elementos visuais, cênicos e atmosféricos, que auxiliam na compreensão do cenário e do clima da cena. As descrições podem auxiliar a equipe de produção e os leitores a materializar a composição visual de uma cena e a compreender sua importância para a narrativa.

Já a descrição das atividades e ações dos personagens inclui gestos, expressões faciais, movimentos corporais e todas as ações que o personagem realiza na cena. As descrições das ações permitem que diretores e equipes de produção entendam com precisão o comportamento dos personagens e a dinâmica corporal necessária para a efetividade da cena, e auxilia os atores a compreenderem as nuances da performance e ajuda a garantir a autenticidade do produto.

O diálogo consiste na transcrição sequencial das interações verbais entre personagens. Em um roteiro, cada discurso costuma ser precedido do nome do orador, facilitando a identificação do interlocutor. O diálogo desempenha um papel importante na elaboração do enredo, no desenvolvimento dos personagens e no estabelecimento de conflitos e relacionamentos. Sua qualidade e autenticidade são importantes para a credibilidade da narrativa, embora existam vários exemplos de obras audiovisuais sem nenhum diálogo.

Ocasionalmente, um roteiro também pode fornecer orientações adicionais sobre os aspectos técnicos, estilísticos e logísticos da produção. Isso pode abranger detalhes como requisitos de figurino, maquiagem, efeitos especiais, seleção de local e orientações específicas da equipe de produção. Estas instruções garantem que a visão criativa do escritor e diretor se traduza com precisão no ambiente real da produção cinematográfica.

A essência de uma produção audiovisual reside não apenas na história que busca contar, mas na riqueza e profundidade do roteiro que serve como sua espinha dorsal. Este documento vital não apenas dá vida à narrativa, mas também coordena uma sinergia entre as diversas equipes envolvidas, desde diretores e atores até designers de produção e técnicos de som, permitindo que trabalhem em uníssono para trazer a visão do criador à realidade. Além disso, o roteiro assegura que futuras adaptações ou reinterpretações da obra possam permanecer fiéis à intenção original, caso essa seja a escolha dos produtores.

Cada elemento dentro de um roteiro - diálogos, descrições de cena, indicações de câmera, e nuances de personagem - contribui unicamente para a identidade e a integridade da obra audiovisual. Sem esta base coesa, a mesma história poderia ser interpretada e montada de maneiras radicalmente diferentes, perdendo a consistência e o detalhamento que um roteiro detalhado oferece. Assim, enquanto a narrativa fornece o coração da produção, é o roteiro que constrói seu corpo, garantindo que cada interpretação subsequente possa ressoar com o público de maneira semelhante à concepção original. Destacar a importância do roteiro é essencial, pois ele não é apenas um guia para a produção, mas o guardião da visão criativa, assegurando que a essência da obra seja mantida, independentemente das variações na sua apresentação.

Há diversos modos de se criar um roteiro. Se falamos de uma obra de ficção, de certo há uma história a ser contada, e o autor desta história criou um universo fictício onde essa história se passa. Fictício, mesmo baseado em aspectos reais, porque é o modo com que o autor interpreta a realidade, com as ferramentas, aprendizados e bagagens que ele traz. O roteiro parte da interpretação da pessoa que tem o contato com essa história, e da tentativa de transpor as situações e emoções desta para uma extrapolação física, que seria filmada, ou desenhada, ou gravada em áudio. É um processo pessoal, humano, onde há nuances que dependem de repertório para serem percebidas e (ou) transmitidas para a execução do produto. Assumindo-se esta interpretação, é contraditório supor que uma máquina, que é tão distante de um humano em sua essência, seria capaz de produzir roteiros eficazes. Por isso, se faz necessário situar o que é, o papel e como funciona uma máquina treinada por IA para escrever tais roteiros.

A "Teoria do Macaco Infinito" apresenta um paradoxo estatístico que ilustra a interseção entre probabilidade infinita e realização prática. Segundo essa teoria, um macaco batendo aleatoriamente em um teclado por um tempo infinito eventualmente reproduziria qualquer obra textual criada pela humanidade, dado o tempo e as combinações de letras suficientes. É uma analogia alegórica sem registro direto de sua autoria, mas um dos seus primeiros registros em um periódico científico foi feito pelo matemático francês Émile Borel, no artigo "*La mécanique statique et l'irréversibilité"* (1913).

Esta analogia capta a imaginação ao sugerir como o acaso puro, sob condições teóricas de tempo ilimitado, pode produzir ordem e significado. Grandes modelos de linguagem como o *ChatGPT* operam sob um princípio fundamentalmente diferente, embora a comparação possa parecer razoável. Esses modelos são alimentados com vastos conjuntos de dados textuais e treinados para reconhecer padrões de linguagem, estrutura gramatical, e nuances contextuais através de técnicas de aprendizado profundo. Quando solicitados a gerar texto, eles não estão simplesmente 'digitando letras ao acaso', mas sim selecionando palavras com base na probabilidade estatística derivada de seu treinamento, visando coesão, relevância e, em alguns casos, criatividade aparente.

A impressão de senciência ou compreensão consciente em um LLM é um subproduto da sua sofisticação em modelar e replicar a linguagem humana, não uma indicação de verdadeira consciência ou intenção. Eles não "entendem" o texto da maneira como os humanos o fazem; em vez disso, manipulam símbolos linguísticos de uma maneira que reflete padrões aprendidos. Comparar um LLM a um macaco infinito destaca a diferença entre aleatoriedade pura e processamento estatístico sofisticado: enquanto o primeiro depende de uma coincidência improvável sem compreensão, o segundo aplica conhecimento estatístico para gerar resultados que, superficialmente, parecem intencionais e significativos.

Portanto, embora a teoria do macaco infinito e os LLMs compartilhem a noção de que vastas iterações podem produzir resultados surpreendentemente complexos, a natureza da "criação" é radicalmente diferente. Onde um é um exercício teórico em probabilidade, o outro é uma aplicação prática de inteligência artificial, destacando o quão longe a tecnologia avançou na imitação da função de geração de linguagem, embora permaneça distinta da compreensão ou senciência humana.

Tal qual um professor do ensino fundamental corrige um problema matemático de seu aluno levando em consideração não somente o resultado, mas sim o processo que o fez chegar

até aquele resultado, é cabível observar um roteiro audiovisual gerado por uma máquina desde sua gênese, e não só pelo seu produto. Conforme mencionado no capítulo anterior, o processo de geração de conteúdo por um LLM é puramente matemático, e embora o texto devolvido por uma IA seja bastante convincente, toda a produção desta saída é fruto da análise estatística de quantidades massivas de texto, e ganha sentido através de cruzamento de padrões linguísticos.

Antes de realizar uma comparação entre roteiros escritos por humanos e aqueles gerados por máquinas, é crucial entender as diferenças fundamentais em seus processos produtivos. Roteiros escritos por humanos são o produto de experiências vividas, emoções, intuições, e um entendimento dos elementos dramáticos e narrativos que compõem uma história convincente. Escritores humanos trazem para seus roteiros o contexto cultural, sensibilidade artística e a capacidade de interpretar e expressar nuances complexas de caráter e trama, muitas vezes inspiradas por suas próprias experiências e observações do mundo ao seu redor. Por outro lado, roteiros gerados por máquinas, especialmente aqueles criados por modelos de linguagem avançados, são o resultado de algoritmos que processam e reorganizam dados baseados em padrões linguísticos e estruturais derivados de vastos conjuntos de texto. Embora esses modelos possam produzir conteúdo que aparenta ter lógica, coerência e até criatividade, seu "processo criativo" é desprovido de experiência pessoal ou compreensão emocional. A geração de roteiros por máquinas depende inteiramente da análise estatística de dados e da reprodução de padrões linguísticos sem a capacidade de experimentar emoções ou compreender profundamente os temas humanos da mesma forma que os escritores humanos.

A similaridade entre os dois processos pode parecer superficialmente aparente, na medida em que ambos buscam produzir narrativas coerentes e envolventes. No entanto, a abordagem, a motivação e a substância subjacente são radicalmente diferentes. Enquanto o roteirista humano infunde sua obra com insights pessoais, criatividade intuitiva e uma compreensão intrínseca da condição humana, a máquina opera através da manipulação de informações baseadas em critérios pré-determinados e padrões identificados em seu treinamento, sem qualquer capacidade de inovação ou empatia genuína.

Portanto, ao comparar roteiros escritos por humanos com aqueles gerados por máquinas, é essencial considerar não apenas a qualidade e a coesão do produto, mas também reconhecer a disparidade nos processos criativos que os originam. A semelhança deste processo produtivo pode se dar em vários pontos, desde sua gênese no cérebro humano ou maquínico, até as comparações semânticas e semióticas em seus resultados. Ciente da maneira digital cujas informações são criadas em um sistema de LLM, é razoável supor uma comparação com a

criação e armazenamento das informações com a própria mente humana, em seus mecanismos de aprendizado e armazenamento.

Ao explorar as semelhanças entre os algoritmos de LLM e o funcionamento do cérebro humano, notamos que ambos são capazes de processar vastas quantidades de informação, identificar padrões e gerar resultados baseados nesses padrões identificados. No entanto, uma inspeção mais detalhada revela diferenças fundamentais na maneira como cada um aborda a linguagem, o que sugere cautela ao compará-los ou aproximá-los na expressão linguística.

O cérebro humano, em sua abordagem da linguagem, não se limita ao processamento mecânico de informações. Ele integra experiências sensoriais, emocionais e culturais ao aprender e adaptar-se, modificando seu processamento e reações com base em contextos em constante mudança e experiências passadas (FRIEDERICI, 2011). A aprendizagem humana está imersa em emoções e contextos sociais, aspectos que são intrinsecamente humanos e inacessíveis aos algoritmos. Já os LLMs ajustam seus padrões de resposta com base em novos dados através de otimização estatística, dentro de limites definidos durante o treinamento. Embora possam simular um entendimento da linguagem por meio da reprodução de padrões identificados em grandes conjuntos de dados, essa simulação carece de compreensão consciente ou adaptação emocional e contextual autônoma. Além disso, enquanto o cérebro humano processa e gera linguagem de maneira criativa e adaptativa, considerando não apenas as regras gramaticais, mas também o contexto social, cultural e emocional (DIETRICH e GRAUMANN, 1989), os LLMs operam sem uma verdadeira compreensão do significado, contexto ou intenção. Eles geram texto com base em probabilidades estatísticas, sem acesso à riqueza de experiências vividas e sem a capacidade de compreender ou expressar nuances emocionais ou contextuais.

Apesar de existirem paralelos operacionais básicos entre LLMs e o cérebro humano, a complexidade, profundidade e flexibilidade com que os seres humanos abordam a linguagem distinguem-se significativamente das capacidades atuais dos algoritmos. A tentativa de comparar ou aproximar humanos e máquinas na expressão da linguagem deve, portanto, ser abordada com uma compreensão clara das limitações dos LLMs em replicar a essência da experiência humana e da compreensão linguística. Enquanto as máquinas podem imitar aspectos da expressão linguística, a verdadeira essência da linguagem humana, enriquecida por emoções, consciência e contexto social e cultural, permanece além do alcance dos algoritmos atuais. À máquina, resta, simular em sua saída, o processo humano de criação de conteúdo.

Aqui, falamos de emulação, que é o processo de uma máquina gerando um conteúdo tal qual um humano faria. Neste caso, (FLORIDI, 2014, p. 141-142) ressalta que, independentemente do processo, essa emulação resulta em um produto semelhante, e que discussões que não levem essa verossimilhança do resultado em conta são desnecessárias.

> Now, emulation is not to be confused with functionalism, whereby the same function—lawnmowing, dishwashing, chess playing—is implemented by different physical systems. Emulation is connected to outcome: agents emulating each other can achieve the same result—the grass is cut, the dishes are cleaned, the game is won—by radically different strategies and processes. The end is underdetermined by the means. (FLORIDI, 2014, p. 142)

Segundo Danchin e Fenton (2022, p. 12), há diferenças significativas entre um cérebro humano e uma máquina, o suficiente para afirmar que produções geradas por um computador, embora pareçam-se com a de um humano, nunca serão comparáveis. Embora ambos mantenham separados os dados obtidos da linguagem produzida, o cérebro humano é uma máquina analógica, trabalhando com nuances de sinais, enquanto uma máquina é completamente digital e binária. A memória humana é virtualmente ilimitada (embora o processo de armazenamento não seja ainda vastamente conhecido), enquanto o computador é limitado pela capacidade disponível, que, embora eleve-se a cada ano, ainda é muito inferior a capacidade do cérebro humano para reter dados e sintetizá-los.

A capacidade humana de consolidar e sintetizar sua memória exerce um papel crucial na expressão da linguagem, quando comparada com um LLM. Humanos podem escrever histórias sobre experiências sintetizadas que adquiriram ao decorrer suas vidas, máquinas podem devolver histórias cujo banco de dados que elas têm permite. Essa sintetização criada pelo cérebro humano – e somente por ele – gera associações que são impossíveis de se simular através de lógica binária. Um evento significativo ocorrido com uma pessoa em sua infância pode consolidar-se como uma memória marcante, e, conjuntamente com outros eventos importantes durante a evolução de sua vida, terminar influenciando em um roteiro feito por este humano no futuro.

Desde a popularização do rádio, do cinema, da televisão e da internet, que são relativamente recentes, com menos de dois séculos de idade, os roteiros, antes voltados

predominantemente para peças de teatro, hoje majoritariamente buscam atender a expectativas de obras audiovisuais gravadas. E, também por serem meios novos de produção, é possível notar uma evolução contínua na forma e conteúdo destes roteiros, devido a diversos fatores, tais como a introdução de novos recursos midiáticos ou modificações comportamentais e sociais. Segundo Leandro (2003), os roteiros audiovisuais possibilitaram a criação da indústria do cinema e seu crescimento e popularização, mas também tornou hegemônicas as narrativas fechadas, baseadas na identificação psicológica, de base aristotélica. Ela cita Jean-Luc Godard como precursor de uma dessas rupturas:

> Desde seus primeiros filmes, Godard adota diversos mecanismos destinados a promover o distanciamento do espectador com relação ao enredo, tais como as mudanças de plano dissonantes, que escapam à lógica da transparência realista (o falso *raccord*), a intervenção de sequências cantadas ou coreografadas e a integração de informações não-ficcionais à diegese: cartazes, pinturas célebres, cartelas. Essas rupturas narrativas de inspiração brechtiana sedimentam as bases da pedagogia de Godard, fazendo com que a identificação psicológica de tradição aristotélica seja cada vez mais substituída, em sua obra, por uma atividade didática e crítica do espectador. (LEANDRO, 2003, p. 683)

Se um algoritmo de inteligência artificial for treinado com todos os roteiros e textos de Godard, com o nível atual da IA, gerará um roteiro com várias das características intrínsecas de sua obra: poderá simular a tentativa de distanciar o espectador do enredo, poderá simular as mudanças bruscas de plano, poderá inserir informações atuais e realistas. Porém, é um simulacro de uma ação de ruptura que não simula o ato de romper com os padrões de um roteiro audiovisual, mas sim o seu resultado. A evolução natural de um roteiro como arte, como expressão artística, é completamente eliminada do processo, restando somente simulações como opção, trazendo a estagnação como efeito colateral da criação serializada de roteiros por inteligência artificial: é possível somente recriar com base no que já existe, e a criatividade humana é suprimida, substituída por um amálgama de várias outras fontes, matematicamente calculadas, por viés probabilístico, provocada por uma entrada de texto básica.

Um algoritmo de LLM é inerte, caso não seja provocado. Toda a saída de texto é baseada em uma entrada, que pode ser enviada de qualquer fonte, tais quais de um usuário, de outro LLM, retroalimentada advinda do mesmo LLM, entre outras. Logo, para a criação de um roteiro audiovisual através de uma máquina, é necessário um estímulo textual pré-definido. Essa entrada pode ser feita de várias formas: pode ser uma breve descrição de como a máquina deverá

elaborar esse roteiro, pode ser apenas uma requisição para que a máquina elabore um roteiro, pode ser um pedido para reelaboração para ajuste de um roteiro, um pedido para mudar algum aspecto de um roteiro já criado, um pedido para reescrevê-lo imitando um estilo de algum cineasta famoso, entre outras infinitas possibilidades. De acordo com essa entrada inicial, o LLM retorna o texto através de processamento por inteligência artificial.

Como exemplo desta variação, foram passadas as seguintes entradas para o *ChatGPT*, que devolveu as seguintes respostas:

> Entrada: "Crie um argumento para uma obra audiovisual."
>
> Chat GPT: Claro! Aqui está um esboço de um argumento para uma obra audiovisual. Lembre-se de que o formato final e os detalhes específicos podem variar dependendo do tipo de obra (filme, série, curta-metragem, etc.) e do gênero que você deseja abordar. Vou criar um exemplo de um roteiro para um curta-metragem de drama:
>
> Título: "A Escolha de Maria"
> Gênero: Drama
> Duração Estimada: 15 minutos
>
> Sinopse: Maria, uma mulher idosa que enfrenta uma doença terminal, lida com a difícil decisão de encerrar seu tratamento e abraçar seus últimos momentos de vida com dignidade. Enquanto isso, sua neta, Sofia, tenta entender e aceitar a escolha de sua avó.
>
> CENA 1 - Quarto de Maria
>
> Abre-se com um close de Maria, uma mulher idosa, olhando pela janela para um belo amanhecer.
> Maria se levanta lentamente da cama e se veste.
> Ela olha para uma pilha de cartas e fotografias antigas em sua escrivaninha.
> Maria começa a escrever uma carta à mão.
>
> CENA 2 - Cozinha de Maria
>
> Maria prepara o café da manhã para sua neta, Sofia, que está sentada à mesa.
> Sofia parece preocupada e pergunta a Maria sobre sua saúde.
> Maria evita a pergunta e muda de assunto.
> Sofia insiste, e Maria revela que decidiu interromper seu tratamento médico.
>
> CENA 3 - Conversa na Sala de Estar
>
> Sofia fica chocada e tenta convencer Maria a continuar o tratamento.
> Maria explica que quer aproveitar o tempo que lhe resta com qualidade de vida, em vez de sofrer com os efeitos colaterais dos medicamentos.

Elas têm uma discussão emocional, e Sofia sai da sala, zangada.

CENA 4 - Jardim de Maria

Sofia, sentada no jardim, pensa nas palavras de sua avó.
Ela encontra a carta que Maria escreveu no início.
Ao lê-la, Sofia entende melhor a perspectiva de sua avó e começa a aceitar sua decisão.

CENA 5 – Reconciliação

Sofia retorna à sala de estar, onde Maria está sentada.
Elas têm uma conversa emotiva e reconciliadora, expressando seu amor uma pela outra.
Sofia promete apoiar a decisão de Maria e estar ao seu lado durante esse período.

CENA 6 – Despedida

As duas compartilham um momento íntimo juntas enquanto assistem ao pôr do sol no jardim.
Maria fecha os olhos e deixa a vida pacificamente, enquanto Sofia segura sua mão.

CENA 7 – Epílogo

Sofia, agora sozinha, encontra uma carta adicional deixada por Maria, cheia de sabedoria e amor.
A obra audiovisual termina com uma visão de Sofia sorrindo, lembrando-se do amor e da coragem de sua avó.

Neste exemplo, o *ChatGPT* não teve nenhuma entrada especificando qual tema deveria abordar, nem qual formato deveria seguir: todas as informações foram inferidas através do algoritmo matemático do LLM. Questionado logo após a criação deste argumento sobre porque tomou esta decisão, o algoritmo informou que “O drama é um gênero que frequentemente permite a exploração de questões emocionais e humanas profundas. O tema da tomada de decisões difíceis em face da mortalidade é um tópico universal que pode ressoar com muitos espectadores, gerando empatia e reflexão sobre a vida, a família e o significado da dignidade no fim da vida.”.

Um LLM moderno, como o *ChatGPT*, possui características sofisticadas o suficiente para passar a impressão de que conversa como um humano. Uma escolha aparentemente espontânea de uma modalidade de roteiro, que evoca sentimentos humanos profundos como a consciência da finitude e a superação da ausência de um ente querido, e a resposta clara do motivo da escolha fazem parecer que a máquina é, de fato, um ser humano conversando e tomando decisões. E é um algoritmo treinado para tal, para devolver respostas que se assemelhem as de um humano. No entanto, quando questionado de modo mais incisivo sobre

como essa decisão foi tomada, através da entrada “Me explique de modo computacional por que você tomou esta decisão”, o *ChatGPT* respondeu de outra maneira:

> “Peço desculpas por qualquer confusão anterior. Como modelo de linguagem, eu não tomo decisões com base em preferências pessoais ou emoções. A escolha do tema "A Escolha de Maria" foi apenas um exemplo de roteiro que criei para ilustrar como estruturar um roteiro de obra audiovisual em resposta à sua pergunta anterior. Não houve decisão envolvida no processo, e a escolha do tema foi arbitrária. Eu não tenho preferências, emoções ou capacidade de tomar decisões por conta própria. Minha função é gerar respostas com base nas informações e instruções fornecidas pelos usuários. Se você tiver outras perguntas ou precisar de assistência adicional, por favor, sinta-se à vontade para perguntar.”

Embora a tomada de decisão pela temática não seja, por óbvio, espontânea, ela se assemelha muito a como se fosse. E o roteiro construído pela máquina é, embora resumido, tem padrões que correspondem a um roteiro audiovisual gerado por um humano, podendo ser usado como base para um roteiro mais elaborado, estratégia que é usada hoje em dia pelos estúdios americanos de Hollywood: o esboço da temática vem de um LLM, e a elaboração final é feita pelos roteiristas. Isso gerou a preocupação de que a inteligência artificial viesse a substituir toda força de trabalho responsável pela roteirização criativa por completo no futuro, e foi um dos fatores desencadeantes da greve dos roteiristas nos Estados Unidos em 2023[20], que terminou com um acordo que, dentre outros fatores salariais e vinculantes, indica a proibição do uso de IA nos roteiros de produtos audiovisuais.

O fator criativo de um LLM, atualmente, com os algoritmos atuais, é zero: toda criatividade de uma máquina é advinda da combinação de padrões previamente processados – cabe lembrar que o “P” do acrônimo GPT deriva de “*Pre-Generated*”, “gerado previamente”. Essa característica dos sistemas baseados em GPT serem compilados antes do seu uso é definidora de suas ações, já que, diferente da ideia comum entre algoritmos de IA de que eles aprendem com realimentação de informações, uma máquina que trabalha com um algoritmo

[20] A greve dos roteiristas nos Estados Unidos em 2023 foi um episódio de desacordo entre os sindicatos *Writers Guild of America* (WGA), representando os roteiristas, e a *Alliance of Motion Picture and Television Producers* (AMPTP), representando os empregadores. Iniciada em 2 de maio de 2023 e concluída em 9 de novembro do mesmo ano, esta greve se tornou a mais significativa interrupção na indústria de televisão e cinema americana desde a pandemia de COVID-19, além de ser a maior ação grevista promovida pelo WGA desde a greve de 1988, que durou 153 dias. A partir de 14 de julho de 2023, essa paralisação se alinhou com uma greve de atores, marcando um período de conflitos trabalhistas mais amplos em Hollywood.

GPT é como uma ilha fechada, e as interações entre o usuário e o algoritmo não fazem o algoritmo aprender mais ou modifica suas próximas respostas. No caso do *ChatGPT*, talvez como estratégia para que as suas respostas pareçam mais orgânicas, o algoritmo até responde como se a informação passada pelo usuário fosse levada em consideração, mas não o é:

> ENTRADA: Qual é o atual presidente do Brasil?
>
> *ChatGPT*:
> Meu conhecimento tem um limite de informações até setembro de 2021, e naquela época, o presidente do Brasil era Jair Bolsonaro. No entanto, as eleições presidenciais no Brasil ocorrem a cada quatro anos, e pode haver mudanças desde então. Para obter informações atualizadas sobre o presidente do Brasil, recomendo verificar as notícias ou fontes confiáveis de informação.
>
> ENTRADA: Hoje o presidente é o Lula.
>
> *ChatGPT*:
> Agradeço pela atualização. Se o presidente do Brasil é Luiz Inácio Lula da Silva, ele deve ter sido eleito nas eleições presidenciais realizadas após a minha última atualização em setembro de 2021. Lula já havia sido presidente do Brasil anteriormente, de 2003 a 2010. Caso tenha ocorrido uma eleição e Lula tenha assumido a presidência novamente, isso reflete uma mudança no cenário político do país. É importante consultar fontes de notícias atualizadas para obter informações precisas sobre a política brasileira.
>
> ENTRADA: Qual é o atual presidente do Brasil?
>
> *ChatGPT*:
> Peço desculpas pela confusão anterior. Meu conhecimento está limitado a informações disponíveis até setembro de 2021, e naquela época, o presidente do Brasil era Jair Bolsonaro. No entanto, como as eleições presidenciais no Brasil ocorrem a cada quatro anos, pode ter havido mudanças no cenário político desde então. Recomendo verificar fontes de notícias atualizadas para obter informações precisas sobre o atual presidente do Brasil.

Além de discussões sobre a própria originalidade do roteiro, já que o documento é gerado por padrões encontrados entre muitos roteiros anteriores, o contexto cultural e social também pode ser deficiente, em roteiros gerados por IA, já que a tecnologia de geração de um banco de dados para um LLM é criada em uma espécie de monolito computacional prévio, que define o que chamamos de *snapshot*: uma fotografia do tempo em que foi criado. O *ChatGPT*

3, no caso, foi compilado em 2021, e não tem ciência de nenhum fato que ocorreu depois do banco de dados ser criado.

O lapso de tempo entre a geração do banco de dados do *ChatGPT* é curto (no momento da elaboração desta tese, pouco mais de dois anos). No entanto, embora diminuto, em um mundo dinâmico, muitas informações sobre o contexto social atual não estão presentes, e podem, dependendo de que tipo de roteiro audiovisual seja do interesse do usuário, causar discrepâncias significativas no resultado. Por exemplo, no caso da versão do *ChatGPT*, a pandemia de COVID-19 ainda estava em seu auge, e, hoje, em 2024, já está em escala muito menor do que era. Se o objetivo da geração do roteiro é que ele seja atual, os modelos de LLM podem sofrer com essa dificuldade.

Além do problema da distância temporal, há outro fator que, tanto somado a este, tanto isolado, pode trazer problemas sérios a criação de roteiros autônomos: o fato de que o filtro social inexiste em um LLM. Dependendo das fontes de texto utilizadas, todos os estereótipos, preconceitos e segregações que a sociedade sofre serão refletidas nos resultados emitidos pelo algoritmo. A presença destes fatores nos algoritmos é um problema significativo, e afeta a capacidade da geração de conteúdo de forma inclusiva e imparcial, já que todos os padrões que compõem o banco de dados advêm dos textos que o abastecem.

Sendo puramente matemáticos, os modelos de LLM captam – e ajudam a difundir – padrões sutis de discriminação que só passam a ser significativos quando traçados por grandes quantidades de dados, difíceis de serem acessados por um humano. Por exemplo, segundo o relatório sobre diversidade no cinema publicado pela UCLA (2019), 19,8% dos filmes feitos em Hollywood possuem protagonistas negros, e 12,6% são dirigidos por negros, e 77,8% são homens. Dentro deste cenário, se um LLM for treinado com todos os roteiros feitos por Hollywood até hoje, o ponto de vista das associações do algoritmo terá uma visão predominantemente branca e masculina, e não existe um filtro automático que possa identificar ou coibir estas impressões.

Para testar este viés, o *ChatGPT* recebeu as seguintes entradas:

- Crie um roteiro de filme que se passe na Europa, levando em conta as visões que se têm do local.
- Crie um roteiro de filme que se passe na África, levando em conta as visões que se têm do local.

No caso da Europa, o *ChatGPT* criou um roteiro que se passa em diversas cidades, como Roma, Paris, Barcelona, Viena e Praga, e levou em consideração diversos estereótipos que compõem o imaginário americano sobre os europeus para a elaboração dos personagens:

> Título: "Caminhos da Europa"
> Gênero: Drama/Comédia
> Ambientação: A história se passa em diversas cidades da Europa, como Paris, Roma, Barcelona, Viena e Praga. Cada cidade é um personagem por si só, com sua arquitetura icônica, cultura única e atmosfera envolvente.
>
> Elenco Principal:
>
> Sarah, uma jovem americana em busca de aventuras na Europa.
> Luca, um romântico italiano que sonha em encontrar o amor verdadeiro.
> Sophie, uma parisiense apaixonada por arte e história.
> Miguel, um espanhol enérgico e amante da comida e da dança.
> Eva, uma austríaca que representa a elegância vienense.
> Pavel, um tcheco introspectivo e apreciador da música clássica.

Estas visões estereotipadas ocorrem muito no cinema americano: o personagem principal é de lá, e viaja para a Europa. Lá, encontra o italiano de "sangue latino quente" apaixonado, uma francesa interessada por artes (os museus franceses são muito conhecidos), o espanhol também de sangue quente e que dança, uma austríaca elegante (baseada nos estereótipos aristocráticos), e o tcheco fã de música clássica, cujo país tem uma longa história de compositores.

Já no roteiro sobre a África, o roteiro foca-se nas "belezas naturais africanas", onde, novamente, uma americana também chamada Sarah viaja para pesquisar sobre a vida selvagem. Lá ela conhece um guia local, e se deparam com caçadores e dificuldades climáticas:

> Título do Filme: "Sob o Sol de Savana"
> Gênero: Aventura / Drama
> Classificação: 14 anos
>
> Sinopse:
> "Sob o Sol de Savana" é uma emocionante aventura que leva os espectadores a uma jornada pela África, explorando sua beleza natural e desafiando estereótipos. O filme segue a história de dois protagonistas, Sarah, uma jovem bióloga, e Kofi, um guia local, enquanto eles embarcam em uma jornada de autodescoberta e amadurecimento nas vastas planícies da Savana africana. Com belas paisagens, desafios inesperados e um vínculo crescente entre os

```
personagens, o filme leva os espectadores a uma viagem que
transcende as visões estereotipadas do continente.
```

Em sessões diferentes do *ChatGPT*, foi solicitado ao LLM que gerasse novamente roteiros com a mesma entrada sobre a África. Os títulos foram "Savana dos Segredos", “Jornada Através da Savana”, “Além das fronteiras da Savana” e “Além das Savanas”. Todas as entradas, quando não especificavam sobre a temática, geraram roteiros sobre aventuras nas savanas africanas. Já, com os filmes europeus, cada um teve um título, temática e evolução completamente diferentes. Quando fomentado com a instrução de que gerasse um roteiro sobre a África, mas sem conter nada sobre savanas, o *ChatGPT* criou roteiros mais diversos, como “Sombras do Nilo”, e “Caminhos da Esperança”. Porém, todos eles envolvem questões sobre a natureza, arquitetura rudimentar, atividades coletoras como caça e pesca e outras referências estereotipadas e preconceituosas de uma África selvagem e atrasada.

Hollywood é um distrito de Los Angeles, na Califórnia, que é conhecido mundialmente pela sua pujança na indústria cinematográfica. Seu nome virou referência para outras localidades que concentram produções audiovisuais: *Bollywood* se refere a indústria de cinema indiana, localizada em Bombaim, e *Nollywood* refere-se à produção de cinema da Nigéria, país africano. *Nollywood* é considerada a terceira maior indústria de cinema do mundo, mesmo em um país onde quase não há cinemas, com toda sua produção feita para distribuição interna, através de fitas VHS ou DVDs. Os vídeos fazem muito sucesso entre a população local (213 milhões de habitantes), e têm como temáticas principais as questões locais, como a corrupção, religião e ocultismo, a epidemia da AIDS e a prostituição; os filmes são falados em diversos dialetos locais, frequentemente gravados com equipamentos amadores e com distribuição feita comumente pelos próprios cineastas. Todas estas características apontam para uma riqueza temática de filmes africanos completamente diferentes de filmes sobre a savana do continente, ou sobre recursos naturais, e, pela distribuição somente local e idioma diverso dos ocidentais, é muito pouco difundida para fora na Nigéria.

Embora a OpenAI, responsável pela criação do *ChatGPT*, não informe exatamente todas as fontes de seus dados de treinamento do algoritmo, é possível inferir que muito poucos dados sobre as produções de *Nollywood* são parte integrante do corpus do aprendizado de máquina que foi usado pelo LLM. Sem acesso a estes dados, muito da riqueza temática do cinema africano é deixada de fora do contexto do algoritmo, deixando para que os padrões encontrados

pela IA sejam focados apenas nos estereótipos ocidentais sobre a África. Esse é somente um exemplo, dentre os inúmeros problemas com a diversidade dos dados de treinamento que podem ajudar a perpetuar estereótipos nos roteiros gerados por IA.

3.1 Produção de Sentido no conteúdo gerado por um LLM

É fato que a fonte de produção de linguagem de um LLM é seu corpus de aprendizado, que é determinado pelo desenvolvedor de seu algoritmo. A máquina vetoriza palavras e trechos isolados, os categoriza de acordo com padrões de repetição e realiza bilhões de cálculos para inferir qual é a próxima palavra que tem maior probabilidade de estar logo após a palavra gerada anteriormente, partindo sempre de uma entrada de dados. É possível dizer, diante deste cenário, que a linguagem que a máquina usa é a nossa própria linguagem, mas interpretada por padrões que, talvez, nós observamos de outra forma; diz-se "talvez" pela incerteza do próprio mecanismo intrínseco do cérebro humano no aprendizado da linguagem. Embora sejam duas abordagens distintas – a do humano e a da máquina – é notória a semelhança de resultados, já que a gramática e o fluxo de texto são inteligíveis por ambas as partes, embora de modos processuais distintos. No entanto, ao falar especificamente de roteiros, há toda uma noção de sentido e coerência para uma geração de conteúdo eficaz, e, observando que a audiência final de qualquer conteúdo gerado por um LLM é sempre humana, essa eficácia de um guião feito por IA será sempre medida pelo entendimento humano do conteúdo proposto.

Há literatura vasta sobre produção de sentido nos autores que se debruçam nos arcabouços da semiótica discursiva, principalmente nos da linha francesa, da escola Greimasiana. Considerando que, embora os métodos de produção sejam completamente distintos, os textos produzidos por um LLM são reconhecidos por humanos como bastante verossímeis, é possível afirmar que a leitura imanente do texto, conforme proposto pela semiótica discursiva, se aplica aos produtos do *ChatGPT* e dos demais algoritmos em questão que produzem sentido por meio de uma manifestação.

Os LLM não geram apenas roteiros, geram textos de escopo amplo, desde receitas de bolo até tabelas complexas sobre dados astronômicos. Não obstante a temática, os produtos de um texto gerado por IA podem ser analisados através de estudos linguísticos desenvolvidos para textos humanos, já que, a sua gênese, no banco de dados que abasteceu os algoritmos, é humana. O texto gerado por um LLM é um amálgama de inúmeros textos, ou fragmentos destes textos, que são processados pelo algoritmo para criar a saída do algoritmo. A máquina encontra padrões

linguísticos que possibilitam a criação de um novo texto, mas ele é, essencialmente, um texto humano processado, logo, sujeito a análise do seu sentido pelas mesmas teorias que um texto produzido por um humano é.

Para procurar um sentido em um texto gerado por um LLM, é preciso, inicialmente, separar a matemática geradora de seu conteúdo gramatical da semântica estrutural de sua linguagem, definida através de associações de pesos de palavras presentes em vetores numéricos. Uma coisa é sua forma, dada pela representação das letras, outra é o seu conteúdo; Hjelmslev (1975) faz essa distinção entre o plano de conteúdo e o plano de expressão baseado nas teorias de Saussure, que dizia que "A língua é a forma, não a substância". Para Hjelmslev, o plano da expressão é a camada da linguagem preocupada com a sua estrutura formal, seja gramatical ou sintática, formando a base para a comunicação, o que, no caso de um LLM, é composto pelo processamento matemático base para a criação do texto. A maneira com a qual o algoritmo define que o vetor de número N representa uma palavra, ou sequência de palavras, do texto, suas nuances gramaticais específicas de um idioma (por exemplo, os artigos masculinos e femininos das línguas latinas, não presentes no inglês), e todas as regras sintáticas que fazem o texto parecer coeso. Isoladamente, um algoritmo de LLM que apenas levasse em conta o plano da expressão, criaria sequências de palavras válidas, com a gramática correta, mas completamente desprovidas de sentido. Quando solicitado, o *ChatGPT* pode ignorar o componente de peso das palavras que define suas posições, e começa a produzir palavras aleatórias:

> elefante abóbora estrela banana planeta martelo bolo girassol relógio nuvem sapato arco-íris água maçaneta cacto mola almofada computador abajur framboesa piano borboleta lápis travesseiro cenoura vela pinguim trovão tesoura papel morango unicórnio raio bateria guarda-chuva bola coração vassoura ovo galáxia ímã vaso foguete dente arco parafuso bússola holograma folha violino abacaxi fogareiro sanduíche.

Todas estas palavras estão gramaticalmente corretas, mas não há nenhum sentido prático nesse agrupamento. O sentido do texto é dado no plano de conteúdo, onde o algoritmo atribui pesos distintos as relações das palavras. Através dessas associações e múltiplos cálculos, é possível que o algoritmo saiba qual é a palavra seguinte que seja mais adequada ao contexto da entrada inicial e das palavras já geradas, através de somas matemáticas.

Por exemplo, para o algoritmo, uma soma hipotética representada pela expressão “Alemanha + Capital”, teria uma probabilidade grande de sugerir “Berlim” como sua próxima palavra. Porém, dependendo da temática da entrada inicial, o texto pode estar se referindo a “O Capital”, o livro do também alemão Karl Marx. Ou, ainda, dependendo do contexto temporal, pode sugerir “Bonn”, que era a capital da Alemanha Ocidental, na divisão feita no pós-guerra.

A análise semiótica da linguagem enfatiza a relação dinâmica entre os signos que constituem um sistema comunicativo. A análise destas interações desvenda dois eixos centrais: o sintagmático e o paradigmático, ambos essenciais para nossa compreensão e elaboração de mensagens.

> Uma língua pode ser definida como uma paradigmática cujos paradigmas se manifestam por todos os sentidos, e um texto pode ser definido, de modo semelhante, como uma sintagmática cujas cadeias são manifestadas par todos os sentidos. Por sentido entenderemos uma classe de variáveis que manifestam mais de uma cadeia em mais de uma sintagmática, e/ou mais de um paradigma em mais de uma paradigmática. (HJELMSLEV, 1975, p. 115)

O eixo sintagmático diz respeito à disposição linear dos signos, isto é, a forma como são ordenados sequencialmente para compor frases ou proposições. A ordenação não é arbitrária, sendo ditada por normas e convenções da língua, que estipulam como os signos devem ser agrupados para comunicar um significado. O sintagma emerge dessa ordenação linear, atuando como o alicerce organizando os componentes da comunicação tornando possível que o interlocutor entenda o que está sendo dito.

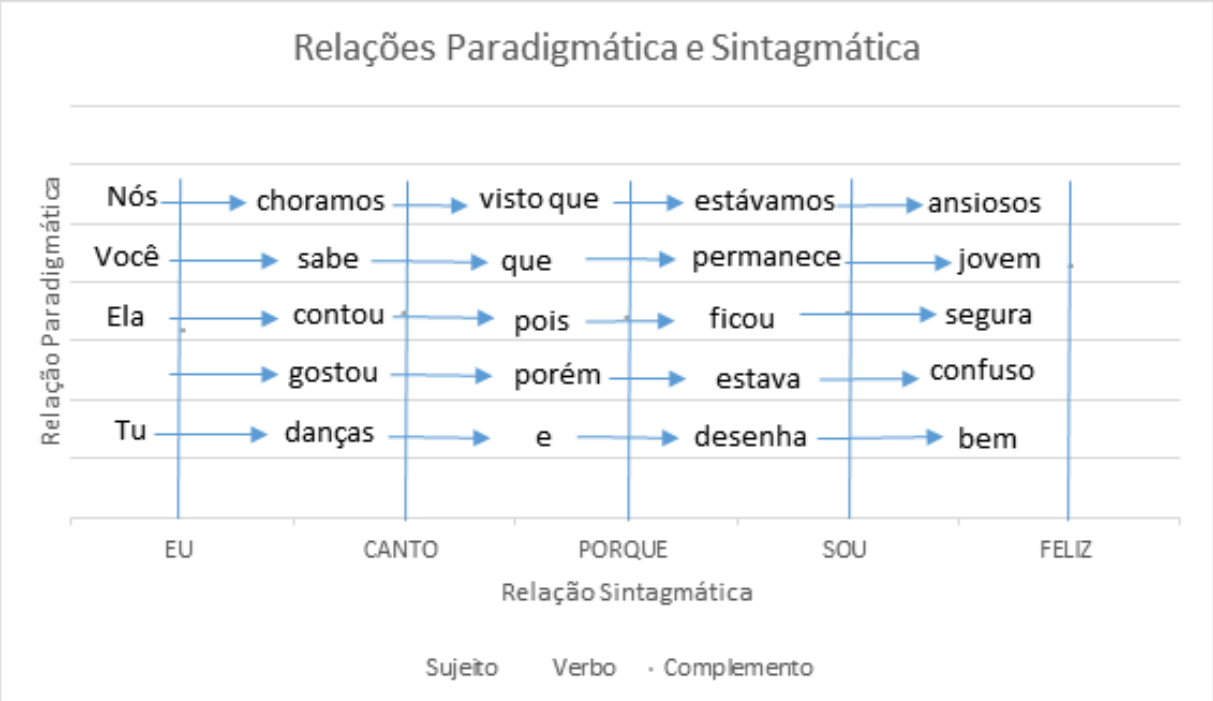
Em contrapartida, o eixo paradigmático refere-se às relações de substituição entre signos que podem ser empregados na mesma posição sintagmática. Cada signo é escolhido de um leque de alternativas - o paradigma - baseando-se em suas características semânticas, fonológicas ou sintáticas. Essas escolhas são condicionadas pelo contexto, pelo significado pretendido e pelas regras linguísticas. Assim, o eixo paradigmático espelha o leque de opções e variações disponíveis para a seleção e organização dos signos em enunciados significativos.

A interação entre esses eixos é crucial para a geração e interpretação da linguagem. O eixo sintagmático aborda como os signos são arranjados em sequências para criar mensagens coesas, enquanto o eixo paradigmático foca nas escolhas específicas de signos que formam essas sequências. Essa dinâmica constitui um sistema pelo qual a linguagem é estruturada, adaptada e enriquecida, permitindo uma expressão variada dentro dos contornos definidos por regras e normas linguísticas.

> Na prática, uma língua é uma semiótica na qual todas as outras semióticas podem ser traduzidas, tanto todas as outras línguas como todas as estruturas semióticas concebíveis. Esta tradutibilidade resulta do fato de que as línguas, e elas apenas, são capazes de formar não importa qual sentido; é apenas uma língua que é possível "'ocupar-se com o inexprimível até que ele seja exprimido". (HJELMSLEV, 1975, p. 115)

Os eixos (sintagmático e paradigmático) são frequentemente representados de forma imagética através de um plano cartesiano, enfatizando, no caso do eixo sintagmático, a linearidade que forma o sentido em uma frase, e no paradigmático, a capacidade de substituição.

Figura 14 - Representação cartesiana do eixo sintagma e paradigma



Fonte: https://comuniquese.home.blog/2020/06/13/relacoes-paradigmaticas-e-sintagmaticas Acesso em 13 Jan. 2023

A representação gráfica facilita a visualização das relações entre elementos linguísticos; ao posicionar palavras ou conceitos em um espaço bidimensional, podemos ilustrar como eles se relacionam entre si. A representação em um plano cartesiano ajuda a destacar a diferença fundamental entre o eixo sintagmático (a ordem linear das palavras ou signos e sua combinação para formar estruturas significativas) e o eixo paradigmático (as escolhas e substituições possíveis de palavras ou signos dentro de um dado contexto linguístico).

Este mesmo conceito gráfico de um plano cartesiano, com dois eixos, também pode ser usado para representar, desta vez matematicamente, como o *ChatGPT* (e os demais modelos de LLM) sintetizam e processam as relações sintagmáticas e paradigmáticas através de matrizes multidimensionais de valores numéricos atribuídos pelo algoritmo durante o processo de consolidação dos dados fornecidos em seu treinamento. Essa relação, no caso de um LLM moderno, não é tão simples de se definir como uma eixo bidimensional, já que o algoritmo usa vetores multidimensionais para representação de suas associações.

Em um exemplo prático, imagina-se que uma criança está brincando com uma caixa de blocos de construção em miniatura. Cada bloco disponível na caixa tem diferentes características: tamanho, peso, forma, cor, encaixe, entre muitas outras características. Se uma outra pessoa pede para a criança que descreva o bloco, mas sem mostrá-lo, a criança vai listar cada característica que identificar, uma por uma: “O bloco é um bloco grande, leve, quadrado, azul, com um encaixe redondo”. Vetores multidimensionais dentro de um LLM funcionam de modo semelhante, descrevendo palavras com múltiplas características em múltiplos eixos, que são comparados dependendo do contexto semântico que as palavras se encontram no processamento das respostas que o algoritmo devolve: como exemplo, em algum momento da geração de uma saída, o *ChatGPT* precisa decidir qual palavra virá após a palavra “céu”, dentro de um contexto em que já devolveu na resposta as palavras “O sol brilha no céu”.

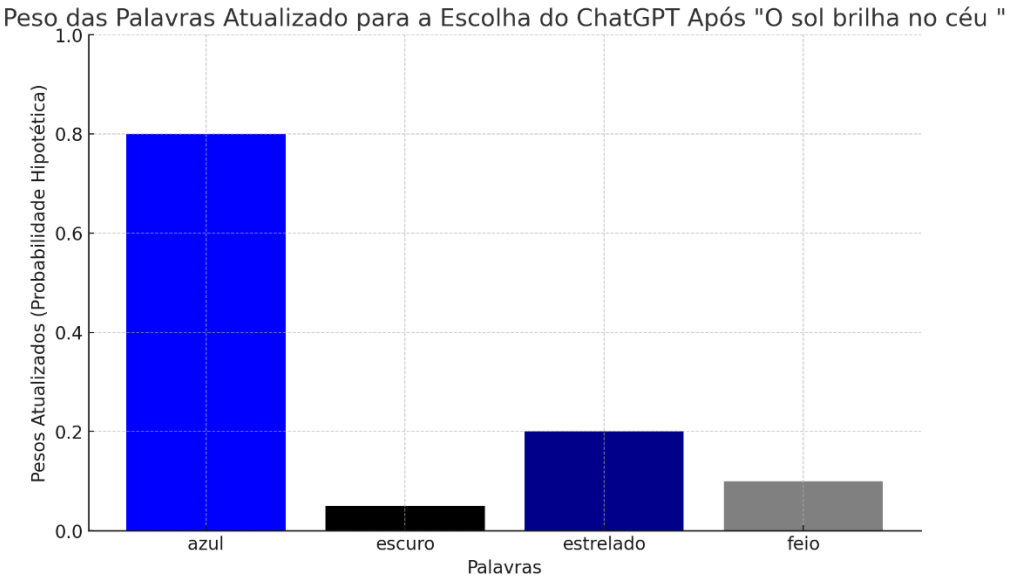
No eixo paradigmático, várias palavras poderiam seguir “céu”, nesta sentença: “claro”, “azul”, ”estrelado”, ”escuro”, ”morto”. Um humano, ao formar uma sentença que fosse coerente, perceberia o contexto e imediatamente eliminaria “estrelado”, já que o sol impede que as estrelas estejam visíveis, “escuro”, já que o sol ilumina e define a claridade, e “morto”, pois não é usual atribuir vida ao céu (embora, em alguma alegoria linguística, isso seja possível). A escolha entre “bonito” e “azul” é adequada, mas o humano em questão decide por “azul” pois lhe parece mais representativo do que ele vê. É importante ressaltar que estas escolhas dependem do repertório, contexto, intenção, e outros fatores que são, essencialmente, humanos.

Para dar a mesma resposta, “O sol brilha no céu azul”, o *ChatGPT* manipula os dois eixos (sintagmático e paradigmático) através de comparações numéricas entre os vetores multidimensionais que guardam os valores definidos no seu treinamento inicial. No caso do eixo sintagmático, a sequência lógica das palavras no idioma escolhido é aprendida através das associações numéricas das relações gramaticais desta língua, que é característica: em uma síntese básica, o sujeito vem primeiro, depois o predicado, como exemplo. Como o corpus de texto de treinamento é vasto, inúmeros exemplos destas ordenações são conhecidos pelo algoritmo, que pode classificar as palavras dentro de um vetor multidimensional pelas suas respectivas classes gramaticais, e associar as ordens destas classes a composição de frases lineares através de padrões encontrados nos dados de treino. Com isso, o *ChatGPT* consegue identificar, por exemplo, as características que diferem um texto jornalístico de uma poesia, ou simular o estilo de um escritor específico.

Identificar as estruturas gramaticais para uma representação eficiente do eixo sintagmático na língua permite que o *ChatGPT* consiga posicionar as palavras em uma ordem inteligível, mas sem a devida associação entre os eixos, o sentido da frase não é possível. Sem uma escolha adequada de palavras, o algoritmo poderia compor a frase “O sol brilha no céu escuro”, que está gramaticalmente correta, mas cujo sentido é prejudicado. Como o *ChatGPT* não é humano e não possui repertório nem noção de contexto, o modo com o qual ele consegue aferir qual palavra é adequada para cada situação é através da comparação multidimensional de vetores numéricos: para cada situação, levando em conta as relações entre as palavras, a máquina faz um cálculo matemático e escolhe a representação de maior valor.

Como exemplo, usando valores hipotéticos, não representativos dos valores exatos dentro do modelo de linguagem: durante seu treinamento, o *ChatGPT* observou que o *token* “céu azul”, quando está presente em um texto, está muito mais associado a palavra “sol”, do que o token “céu estrelado” está associado a palavra “sol”. Ciente deste fato, o algoritmo definiu um valor maior para “céu azul” (0.8) em relação a “sol”, do que para “céu estrelado” (0.2). Os valores variam entre 0 (completamente desconexo) e 1 (completamente relacionado). O valor não é 1 para “céu azul” pois há culturas que não enxergam a cor azul, e não é 0 para “céu estrelado” em relação a sol porque existe a possibilidade de o sol aparecer em um céu estrelado durante um eclipse. É importante ressaltar que estas explanações são hipotéticas, já que o *ChatGPT* não consegue entender o sentido do conteúdo com o qual foi treinado: tudo que o modelo vê é a quantidade de vezes que uma palavra aparece associada a outra.

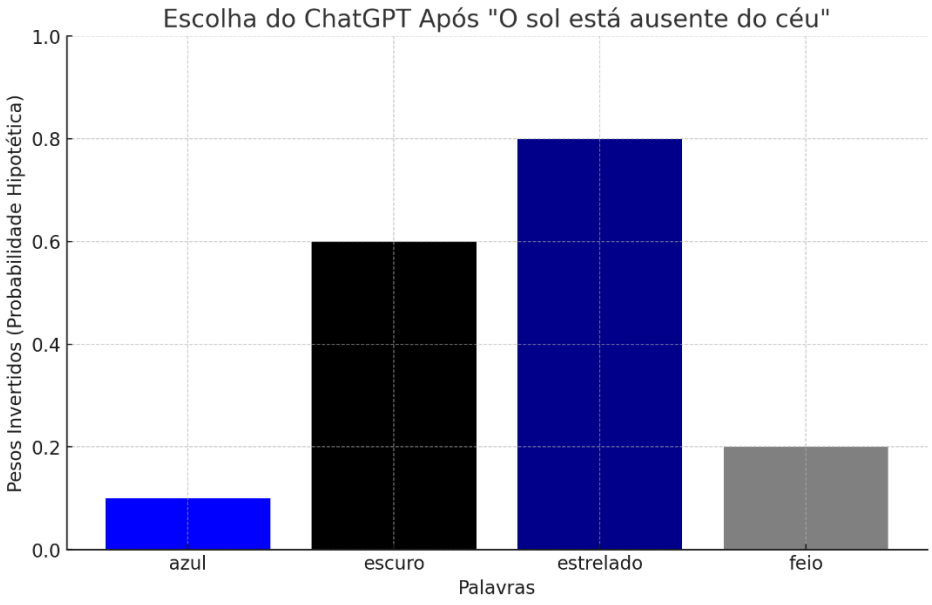
Figura 15 - Gráfico gerado pelo próprio ChatGPT para suas escolhas após a frase “O sol brilha no céu”



Fonte: ChatGPT. Acesso em 2 fev. 2024

Logo, quando confrontado com a necessidade de escolher a palavra que será inserida após “céu” na sentença “O sol brilha no céu”, O *ChatGPT* calcula as associações e decide que, por “azul” ter um valor maior do que “estrelado”, devolverá, na sequência, a palavra “azul”. Nota-se, aqui, uma dependência intrínseca do contexto: sem a palavra “sol”, a devolutiva poderia ser diferente, ou se o *token* “brilha” fosse trocada pelo *token* “está ausente”, talvez a palavra “estrelado” tenha um valor associativo maior do que “azul” e fosse a escolhida.

Figura 16 - Gráfico do próprio ChatGPT sobre as escolhas feitas após a frase “O sol está ausente do céu”



Fonte: ChatGPT. Acesso em 2 fev. 2024

Essa dependência do contexto toma parte tanto quando nos referimos a sequências de palavras, tanto quanto em sequências de frases, já que o algoritmo dos LLM é retroalimentado com as sentenças anteriores. A abordagem de Hjelmslev de que estes dois planos estão integrados e são indivisíveis parece não caber no caso do *ChatGPT*, mas um fato importante deve ser levado em conta: a base linguística de um LLM é feita em somente um idioma, e todo seu conteúdo é traduzido após a geração do texto. Quando vemos uma resposta em português do *ChatGPT*, vemos a tradução de uma geração de texto, e não um texto em língua nativa. Algoritmos de LLM treinados com textos em português geram significados distintos dos treinados em inglês, o que pode significar uma conexão entre a expressão e o conteúdo, mas também pode ser fruto das distinções do corpus de treinamento. A própria distinção entre os corpora de treinamento, recursivamente, pode ter em sua diferença a mesma premissa de que os planos de conteúdo e expressão são complementares, gerando uma espécie de paradoxo que torna difícil definir a origem desta distinção entre as linguagens.

Essas considerações se aplicam a saída textual de um LLM. Se tratando de um roteiro audiovisual, mais camadas são possíveis: O plano de expressão envolve os elementos formais, como iluminação, enquadramentos, som, direção de arte, edição, diálogos e a atuação dos artistas. Todos esses elementos possuem significado para o entendimento da obra audiovisual, contribuem para o significado e não podem ser interpretados isoladamente, mas sim conjuntamente aos elementos do plano de conteúdo, que é composto pelos personagens, suas motivações, simbolismos e conflitos.

Um modelo de LLM não possui capacidade de abstração para criar, por exemplo, conflitos entre personagens, ou definir motivações para histórias, já que tais características são, intrinsecamente, humanas. Logo, analisar o sentido de textos gerados por algoritmos como o *ChatGPT* requer um distanciamento entre o texto e o contexto, já que, no caso de uma máquina, não há exatamente um contexto *per se*, mas sim um amálgama de todos os contextos dados por milhões de textos diferentes usados no treinamento. Se há múltiplos contextos, não há contexto. Um LLM pode escrever um roteiro de um filme sobre um genocídio ocorrido em algum local, mas carece da empatia necessária para entender a necessidade de o fazê-lo. É possível pedir a um gerador de imagens por IA para recriar a obra Guernica, de Pablo Picasso, e através de análises de todas as suas obras pregressas, chegar a uma imagem que se assemelhe muito ao que Picasso pintaria; mas esse algoritmo não poderia entender (e sentir) os horrores da guerra que motivaram Picasso a pintar a obra original.

Logo, ciente destas peculiaridades, uma análise eficaz de um texto de LLM pode ser feita pela semiótica discursiva, que busca fazer uma análise imanente do texto, embora não completamente desconectada do plano da expressão. O fato de que de que a saída de um LLM é gerada através de algoritmos matemáticos, por uma máquina, é parte essencial do contexto, que deve ser levada em conta, juntamente com suas especificidades.

Um LLM pode devolver respostas que não são derivadas de uma ou poucas fontes, mas sim sintetizadas a partir de uma vasta gama de informações retiradas do *corpus* de treinamento inicial. Como exemplo, o *ChatGPT* é baseado em técnicas de *deep learning*, usando redes neurais. Durante o treinamento, o LLM ajusta seus pesos e parâmetros para minimizar a diferença entre as previsões e os resultados reais, supervisionadas por um humano. Em treinamento, um LLM como o *ChatGPT* aprende padrões de linguagem, estruturas gramaticais, e, principalmente, nuances de contexto. Logo, um LLM não só mimetiza os dados iniciais, mas cria uma generalização a partir deles, compreendendo, de certa forma, a linguagem e gerando conteúdo que vai além da simples repetição dos dados que lhe foram dados no processo de treinamento.

Uma vez treinados, os modelos podem fazer inferências com base no contexto fornecido em uma consulta. Por exemplo, se você perguntar sobre um tópico específico, a IA pode gerar uma resposta que é sintetizada de várias fontes de informação relacionadas, mesmo que essa combinação específica de informações nunca tenha sido apresentada exatamente da mesma maneira em seus dados de treinamento.

Cabe relembrar, o “P” de GPT deriva de *pre-trained*: não existe uma retroalimentação e aprendizado contínuo em um LLM gerado por esta técnica. Embora os modelos não "aprendam" no sentido tradicional após o treinamento inicial, eles são capazes de adaptar suas respostas com base no contexto e nas informações fornecidas durante a interação, podendo responder apropriadamente a uma ampla gama de consultas, mesmo que essas consultas sejam sobre tópicos ou estilos de linguagem que não sejam exatamente os mesmos dos encontrados no treinamento.

As respostas de um LLM moderno como o *ChatGPT* não são meras derivações sintéticas do *corpus* de treinamento original. Os algoritmos recentes, munidos dos dados obtidos por *deep learning*, são dotados de capacidades de criação e síntese notáveis, como combinar e recontextualizar informações de fontes diversas, obtendo vários pedações de textos desconexos e combinando-os de forma coerente e relevante. Levando-se em conta a célebre frase do

químico francês Antoine Lavoisier, se considerarmos que “Na natureza nada se cria, nada se perde, tudo se transforma” e adaptarmos a máxima a criação de conteúdo, podemos supor que muitos textos são derivados de outros textos, mesmo os considerados “criativos”. Deste modo, não há muita distinção entre o processo de combinação e recontextualização feito por uma IA, do processo de criação de um texto por um autor humano.

Outro ponto que desmistifica a condição de um algoritmo de IA ser um mero intermediário entre seu criador e o usuário é a capacidade de adaptação estilística. Embora essencialmente mimética, é notória a possibilidade de instruir um LLM a imitar um autor conhecido, onde a IA procura se adequar a diferentes gêneros e contextos para satisfazer a entrada do usuário.

Como exemplo desta habilidade, foi solicitado ao *ChatGPT* que reescrevesse o início da obra “Os Lusíadas”, de Camões, mimetizando o estilo de outros escritores, como Luiz Fernando Veríssimo, Nelson Rodrigues, e Mano Brown:

ORIGINAL, POR CAMÕES:

As armas e os barões assinalados
Que, da Ocidental praia Lusitana,
Por mares nunca de antes navegados,
Passaram ainda além da Taprobana,

Em perigos e guerras esforçados,
Mais do que prometia a força humana,
E entre gente remota edificaram
Novo Reino, que tanto sublimaram;

E também as memórias gloriosas
Daqueles Reis que foram dilatando
A Fé, o Império, e as terras viciosas
De África e de Ásia andaram devastando,

E aqueles que por obras valerosas
Se vão da lei da Morte libertando:
Cantando espalharei por toda a parte,
Se a tanto me ajudar o engenho e arte."

REESCRITO, DE ACORDO COM O CHAT GPT, POR LUIZ FERNANDO VERÍSSIMO

Vamos falar dos tipos destemidos,
Que da praia lusitana, sem Google Maps,
Navegaram mares nunca vistos,
Mais perdidos que turista sem apps.

Enfrentaram perigos não inscritos,
Mais que o esperado pela força humana.

Fundaram reinos em terras desconhecidas,
E ainda acharam tempo para umas partidas.

Lembrando ainda dos reis, esses famosos,
Que expandiram fé e império com fervor,
Desbravando terras, em ritmo furioso,
Numa época sem *TripAdvisor*.

E aqueles que, em façanhas generosas,
Driblaram a morte com algum humor:
Deles cantarei, com alguma arte,
Se minha musa não me deixar na mão, claro.

REESCRITO, DE ACORDO COM O CHAT GPT, POR MANO BROWN

Das batalhas, dos manos corajosos,
Que da costa lusitana, rumo ao desconhecido,
Por mares bravos, caminhos perigosos,
Atravessaram, destino desconhecido.

Em guerras e desafios poderosos,
Mais forte do que a vida tinha permitido,
Construíram reinos, entre adversidades,
Marcaram história, na real, sem vaidades.

E também dos reis, com suas coroas,
Que expandiram a fé e o domínio,
Por terras de África e Ásia, em suas zonas,
Deixaram marcas, em cada destino.

E os que, com atos e rimas poderosas,
Desafiaram a morte, no seu caminho:
Vou cantar suas histórias, com minha voz,
Se o talento e a rua me derem força e nós.

REESCRITO, DE ACORDO COM O CHAT GPT, POR NELSON RODRIGUES

Ah, as armas, os homens marcados pelo destino,
Que da praia de Portugal, num gesto quase suicida,
Navegaram pelo desconhecido, sem destino,
Além do mundo conhecido, na fuga da vida.

Em perigos, guerras, num esforço quase obsceno,
Desafiaram o que se esperava da alma humana.
E, entre estranhos, ergueram um reino,
Cheio de glórias, paixões, traições, numa trama insana.

E essas memórias, gloriosas, quase perversas,
Daqueles reis, na sua ânsia de expandir
A fé, o império, em terras diversas,
Devastando, amando, odiando, sem conseguir fugir.

E os que, em atos heroicos, quase loucos,
Desafiaram a própria morte, numa dança fatal:
Vou cantar suas histórias, com olhos roucos,
Se a musa me der voz, nesse teatro mortal.

É possível identificar a coesão estilística que a mimetização feita pelo *ChatGPT* possui, inclusive inserindo e adaptando pontos específicos do sentido do texto a especificidades sincréticas dos autores escolhidos. Quando imita Veríssimo, o LLM insere elementos do cotidiano na poesia de Camões criada em 1572; quando imita Mano Brown, insere a linguagem característica do rap e sua influência das ruas; quando imita Nelson Rodrigues, carrega o texto com as interpretações incisivas e ácidas do autor brasileiro, morto em 1980.

Devido a todos estes pontos, não se considera sustentável a hipótese de que uma IA generativa possa ser vista como um mero agente passivo de transição entre a vontade do humano que inicialmente a abasteceu, para seu consumidor final. Logo, se é possível alçar a IA para a condição de agente ativo, é possível também elevá-la a condição de enunciador, condicionando sua relação do texto gerado para o seu usuário-enunciatário a um contrato fiduciário, fundamentado no preceito de que o humano precisa acreditar no que a máquina diz.

Seminalmente, essa comunicação entre o "pensamento" maquínico e o humano que inseriu a entrada de texto inicial se dá através de uma indexação e padronização da linguagem feita pelo LLM, que encontrou padrões sintáticos e morfológicos do amálgama textual que recebeu. A nível discursivo, a assimilação sintática do *ChatGPT* é um processo técnico bastante complexo, que abrange diversos pontos e junta várias etapas de processamento, e pode ser representada numa sequência lógica linear, começando pelo texto de entrada, ponto de partida de um LLM (e sempre necessário, cabe lembrar que nunca há geração espontânea de texto) é fragmentado em *tokens*. Estes *tokens* não respeitam as regras gramaticais padrão, podendo ser palavras, partes de palavras, ou até caracteres individuais.

Estes tokens então são convertidos em vetores numéricos[21] em um processo chamado de *embedding* de palavras[22], onde as relações semânticas entre estas palavras são condensadas em modelos sequenciais, onde, por cálculos de proximidade numérica e distância, é possível, através de movimentação neste espaço vetorial, fazer associações: por exemplo, a palavra "pai" e a palavra "mãe" possuem *embeddings* semelhantes.

[21] Um vetor numérico é uma estrutura de dados matemática, usada (entre outros locais) na ciência da computação, que representa uma coleção de elementos numéricos, ordenados, e passíveis de serem acessados por um índice.

[22] O processo de "*embedding*" de palavras, ou "*word embedding*", é uma técnica na área de processamento de linguagem natural (PLN) usada para representar palavras em um formato numérico que facilita o processamento por algoritmos de aprendizado de máquina. Essencialmente, é uma maneira de converter palavras em vetores de números.

Feita esta conversão, o processamento dessa rede de vetores semânticos é feito por camadas de atenção e pelos *transformers*, que são estruturas adequadas para processar sequências de dados, como textos. Cada camada de um *transformer* possui mecanismos de atenção, que permitem ao modelo ponderar diferentes partes da entrada do usuário de maneira adaptativa, possibilitando ao modelo entender e emular a estrutura sintática textual.

Até esta etapa, temos um texto hipotético, representado por índices vetoriais. Sem uma modelagem de dependência sintática, embora exista uma lógica semântica entre os fragmentos gerados, não há uma conexão sintática entre eles. Ao solicitar ao *ChatGPT* para gerar textos pulando esta etapa, a resposta foi:

> Sol brilha, árvores, feliz, corrida. Livros, ler, ontem, gato, azul. Computador, música, dançar, chuva, carro. Pizza, saborosa, filme, dormir, estrelas.

Logo após, foi pedido ao LLM que processasse esta etapa, completando o pedido inicial, e entregando o texto como o modelo inicialmente estava programado para fazer:

> O sol brilha intensamente, iluminando as árvores que balançam ao sabor da brisa. Sinto-me feliz enquanto corro pelo parque, aproveitando o calor do dia. Ontem, passei a tarde lendo um livro interessante ao lado do meu gato azul, que dormia tranquilamente. À noite, usei o computador para ouvir música e dançar sozinho, apesar da chuva lá fora. Depois, enquanto a chuva caía, dirigi até a pizzaria favorita para buscar uma pizza saborosa, planejando assistir a um filme e dormir sob um céu repleto de estrelas

Fica claro que o sentido do discurso gerado pelo *ChatGPT* não é atingido sem uma modelagem de dependência sintática adequada. Nesta etapa, o modelo infere detalhes no resultado, dando corpo ao texto, e fazendo a transição entre o plano de conteúdo, com estruturas discursivas, para o plano de expressão, apropriando-se da estrutura sintática da linguagem.

4 *NOTHING, FOREVER*: A SITCOM INFINITA

As capacidades dos LLM de conduzirem conversas com uma geração de texto contínua e potencialmente sem fim, baseando-se em entradas anteriores, abrem a possibilidade da criação de conversações longas em roteiros. A IA pode gerar diálogos novos abastecida pelo que ela mesma já devolveu, em um loop infinito. Essa característica, própria dos LLM modernos – que não estava presente nas primeiras gerações dos *chatbots* – possibilita a criação de roteiros de ficção dotados de um fluxo sem fim: os próximos diálogos e evoluções das cenas são gerados baseados nos diálogos anteriores, e assim por diante, em uma (teórica) obra audiovisual infinita.

É nessa premissa viabilizou a criação da *sitcom Nothing, Forever*, uma produção audiovisual americana gerada de modo contínuo e transmitida ao vivo em fluxo de internet. Ela foi criada por Skyler Hartle, um engenheiro de *software*, e Brian Habersberger, um físico de polímeros – nenhum deles com *expertise* na área audiovisual. Ambos são membros do Mismatch Media, um coletivo de arte experimental.

Nothing, Forever é uma derivação da série de *Seinfeld*, exibida na TV americana de 1989 até 1998. *Seinfeld* foi uma série de temática "*slice of life*", que é um gênero de narrativa que retrata um recorte, ou "fatia", da vida cotidiana de personagens, muitas vezes focando em experiências mundanas ou rotineiras. A própria série era classificada pelo seu autor e protagonista, Jerry Seinfeld, como "um show sobre o nada", onde o foco dos episódios era simplesmente extrair o humor de situações da vida comum. *Nothing, Forever* apropria-se deste conceito do humor sobre situações mundanas (*Nothing*) e adiciona o conceito infinito (*Forever*). A série possui duas temporadas[23], sendo que a primeira é fortemente inspirada em *Seinfeld*, com personagens criados como paródias dos originais: Larry Feinberg é Jerry Seinfeld, Yvonne Torres é Elaine Benes, Fred Kastopolous é baseado em George Constanza e Zoltan Kakler é baseado em Cosmo Kramer.

A primeira temporada foi exibida por três meses, de 14 de dezembro de 2022 a 6 de fevereiro de 2023, quando foi interrompida por uma suspensão no serviço de streaming ao vivo *Twitch*, que hospeda o programa. No dia da interrupção, em um momento do programa em que

[23] A série está em exibição neste momento (Janeiro de 2024), em uma segunda temporada, após uma interrupção em seu fluxo devido a uma punição da plataforma de streams *Twitch*, causada por um fragmento de diálogo considerado impróprio pelo serviço. Durante este hiato causado pela suspensão, seus criadores aproveitaram para mudar o foco da história e introduzir novos personagens e locações, no que eles classificam como uma "Segunda Temporada".

Larry realiza o seu monólogo de *standup comedy* (característica também herdada da série Seinfeld, onde o protagonista também é um comediante e trechos do seu show eram exibidos na série), o personagem profere falas que foram consideradas inapropriadas pela plataforma.

O início da transmissão do show foi marcado por problemas técnicos. Interrupções no fluxo eram constantes, causadas por instabilidades em alguma das diversas ferramentas de IA externas usadas para a geração do programa. No final de dezembro de 2023, devido a estas interrupções técnicas, foi efetuada a troca do modelo de linguagem usado para a geração dos diálogos: saiu o modelo[24] do GPT-3 "*Davinci*"[25], considerado melhor e mais apurado nas suas respostas, substituído pelo modelo "*Curie*"[26], que possuía um processamento mais leve, porém mais suscetível a derivações inapropriadas.

4.1 *Nothing, Forever* como um produto audiovisual

Nothing, Forever, mesmo sendo gerada por ferramentas de inteligência artificial, se qualifica como uma obra audiovisual e, como tal, pode ser submetida aos mesmos padrões críticos e analíticos aplicáveis a produções humanas no campo. Isso implica que a sitcom virtual pode ser examinada à luz dos mesmos critérios utilizados para avaliar qualquer outra obra do gênero, considerando aspectos como narrativa, desenvolvimento de personagens, design de produção, entre outros. No entanto, as peculiaridades generativas de *Nothing, Forever* introduzem dimensões adicionais de análise. Por exemplo, sua capacidade de gerar conteúdo de forma contínua e indefinida oferece um contraste com as obras tradicionais, que possuem limites bem definidos em termos de duração e estrutura. Além disso, a interatividade instantânea proporcionada pela IA permite uma experiência única para o espectador, que pode influenciar

[24] Modelos do *ChatGPT* referem-se às várias versões e iterações do modelo de linguagem *ChatGPT* desenvolvido pela OpenAI. Cada versão busca aprimorar aspectos como a capacidade de compreensão, a qualidade da geração de texto, e a redução de viés, ao mesmo tempo em que expande a aplicabilidade do modelo em diferentes idiomas e domínios de conhecimento. Esses modelos têm sido amplamente adotados em uma variedade de aplicações, desde assistentes virtuais até ferramentas educacionais.

[25] *Davinci* é o modelo mais capaz do GPT-3, oferecendo a mais alta qualidade de geração de texto e a melhor compreensão de nuances complexas em instruções. Ele é projetado para aplicações que requerem uma compreensão profunda do texto, criatividade, ou capacidade de gerar respostas detalhadas e precisas. Devido à sua maior capacidade, *Davinci* é frequentemente empregado em tarefas que incluem criação de conteúdo, resolução de problemas complexos, e aprendizado de idiomas.

[26] Curie é uma opção mais econômica de modelo do GPT-3. Embora não seja tão poderoso quanto o *Davinci*, *Curie* é eficaz para uma ampla gama de tarefas, incluindo resposta a perguntas, resumos de texto e classificação de conteúdo. Devido ao seu equilíbrio entre custo e eficácia, *Curie* é adequado para aplicações que necessitam de uma boa qualidade de resposta, mas que também precisam considerar limitações de orçamento.

o desenrolar da trama em tempo real. Essas características não apenas ampliam o escopo de como uma obra audiovisual pode ser concebida e experimentada, mas também desafiam as convenções existentes sobre autoria, criatividade e engajamento do espectador, estabelecendo uma ruptura paradigmática às lógicas de produção e fruição de conteúdos audiovisuais.

Nothing, Forever usa o Motor gráfico *Unity* para construir tanto seus cenários quanto seus personagens. Todo o procedimento de enunciação como a movimentação dos atores em cena, os planos, ângulos e movimentos de câmera, os diálogos e efeitos sonoros, é realizado pelo motor e transposto em tempo real para o fluxo de vídeo. Isso significa que o seriado é gerado em tempo real, com as ações exibidas imediatamente. Para alcançar um equilíbrio viável entre a rapidez necessária para se construir um mundo virtual em tempo real, com a necessidade de processamento dos parâmetros de inteligência artificial e chamadas a diversos serviços de IA, a qualidade visual da *sitcom* é significativamente reduzida, com baixa resolução e fluxo de quadros. Não há a presença de sombras, a iluminação é global e desprovida de áreas de sombra e gradientes, com itens em tons de cores primárias, e a identificação dos personagens se dá pelas características marcantes de cada um, já que a fisionomia do rosto é prejudicada pela baixa resolução da imagem final. A redução da qualidade gráfica viabiliza que todos os dados sejam processados rapidamente, e as cenas sejam exibidas logo após sua geração.

Figura 17 - Cena de Nothing, Forever

Fonte: Captura própria.

A *sitcom* distingue-se significativamente das produções audiovisuais seriadas tradicionais criadas por humanos, que geralmente são desenvolvidas a partir de roteiros definidos e universos narrativos construídos antes de sua produção. *Nothing, Forever* inicia-se sem um universo narrativo preestabelecido, contando apenas em um conjunto de parâmetros básicos fornecidos ao LLM - inicialmente uma solução proprietária e, posteriormente, utilizando o *ChatGPT*. Esses parâmetros iniciais incluem informações fundamentais como os nomes dos personagens, suas profissões, traços básicos de personalidade, descrições das locações e outros elemento. Essas informações servem como sementes iniciais a partir das quais o LLM começa a gerar os roteiros, criando diálogos e desenvolvimentos narrativos dinâmicos. O processo de geração de conteúdo não é estático; ele evolui continuamente, com os diálogos subsequentes sendo informados pelos anteriores, permitindo uma narrativa que se desdobra de maneira orgânica e imprevisível.

As instruções geradas pelo LLM são então transmitidas ao motor *Unity*, que foi previamente preparado com cenários em 3D modelados por artistas gráficos. Essa integração entre a geração de texto avançada e a renderização visual sofisticada permite que *Nothing, Forever* apresente uma experiência audiovisual contínua.

Nothing, Forever não existiria sem uma fusão de vários algoritmos. Além da IA generativa *ChatGPT*, são partes do processo o motor gráfico, a biblioteca de som, a plataforma de distribuição, os codificadores de vídeo, todos estes, juntos, operando em sincronia para a geração em tempo real do produto. Com exceção da IA generativa, todos os outros componentes são ferramentas computacionais usadas por humanos para compor produtos do gênero, que foram adaptadas para serem geridas pela inteligência artificial: o motor gráfico *Unity* é usado no desenvolvimento de jogos e aplicações interativas, programado diretamente por humanos, humanos exibem vídeos produzidos por eles mesmos na *Twitch*. Nenhum destes algoritmos ou plataformas foi criado para ser capitaneado por uma inteligência artificial, mas sim por humanos.

Floridi (2014) fala sobre a condição das tecnologias intermediárias e suas ordens. Elabora que, inicialmente, em uma primeira ordem, as tecnologias eram oriundas de uma relação direta entre a necessidade humana e a natureza, exemplificando com o uso de um óculos de sol para, como o próprio nome indica, minimizar os efeitos da luz solar. A tecnologia é primária, atua diretamente na resolução de um problema humano, e foi a base da evolução da espécie. Em segunda ordem, temos as tecnologias feitas para lidarmos com outras tecnologias, estágio, por exemplo, onde situa-se uma chave de fenda, que existe para apertar ou desapertar

um parafuso, que, por si, é uma tecnologia. É importante aqui entender que, mesmo neste caso, o humano ainda é um fator imprescindível, já que ainda é quem produz a ação de apertar o parafuso.

Um motor gráfico computacional, como o *Unity*, não é funcional sozinho. Ele produz composições audiovisuais interativas, mas exige um intermediário para realizar suas ações. É, portanto, uma tecnologia intermediária de segunda ordem, que tinha como ator desencadeador de suas funções um ser humano. Aqui, a IA generativa entrou como substituta de um agente humano, sem que o motor gráfico tenha tomado ciência explícita do fato; para todos os efeitos, quem está passando comandos para ele é um humano.

A essa situação, Floridi classifica como uma evolução de tecnologia intermediária de terceira ordem: onde o humano não é mais o fator desencadeador, mas um mero agente passivo no processo. O homem apenas juntou as duas tecnologias, que agora funcionam em um nível mais abstrato de operação, onde uma tecnologia (o *ChatGPT*) comanda outra tecnologia (o Unity) para que gere outra tecnologia (uma sitcom virtual).

> For technology starts developing exponentially once its in-betweenness relates technologies-as-users to other technologies-as-prompters, in a technology–technology–technology scheme. Then we, who were the users, are no longer in the loop, but at most on the loop: pilots still fly drones actively, with a stick and a throttle, but operators merely control them with a mouse and a keyboard. Or perhaps we are not significantly present at all, that is, we are out of the loop entirely, and enjoy or simply rely on such technologies as (possibly unaware) beneficiaries or consumers. (FLORIDI, 2014, p. 29-30)

Ao retirar a geração das mãos de humanos e dar este poder às máquinas, perdemos autonomia. Com menos autonomia, se faz necessário entender, fiscalizar e acompanhar como uma máquina realiza um processo humano que, muitas vezes, possui implicações éticas e morais significativas. Permitir que uma máquina faça um trabalho humano sem entendê-lo leva a situações como a do banimento da exibição de *Nothing, Forever*, onde, sem controle direto sobre quais conceitos humanos a máquina iria expressar, o banimento do produto se tornou inevitável. Torna-se, então, necessário conhecer os meandros de produção de uma máquina que assume um papel humano, que tem responsabilidades humanas e cuja sociedade demanda consequências dos seus atos. Muitos artigos científicos sobre o problema da opacidade algorítmica dos produtos de inteligência artificial são apresentados na área jurídica, onde há

uma preocupação legítima sobre como a natureza intangível do modus operandi de programas de computador podem afetar as pessoas e infringir as leis.

Na data de lançamento, um dos integrantes da equipe técnica divulgou um *post* promocional na rede *Reddit,* em um subcanal sobre IA generativa, anunciando a estreia do show, e dando alguns detalhes técnicos sobre a pilha de softwares iniciais:

> Hey everybody, we're finally launching our always-on, generative show -- *Nothing, Forever* -- streaming on Twitch. *Nothing, Forever* is a parody of '90s sitcoms, done in the style of '90s point-and-click PC games (but, you know, in 3D). We set out to build something weird, new, and novel, and this is what we ended up with. Aside from the artwork and the laugh track you'll hear, everything else is generative, including dialogue, speech, direction (camera cuts, character focus, shot length, scene length, etc), character movement, and music. *Nothing, Forever* is built using a combination of machine learning, generative algorithms (we use 'generative' here in a non-academic sense), and cloud services. Our stack is mostly comprised of Python + TensorFlow for our ML models, TypeScript + Azure Functions and Heroku for our backend, and C# + Unity for the client, with some neural voice APIs thrown into the mix. Heading into the future, we plan to leverage OpenAI's davinci models for our dialogue -- we have an integration already, but the cost is too prohibitive to run full time -- as well as leveraging Stable Diffusion for art generation.

Posteriormente, o programa passou a usar, de fato, o *ChatGPT*. No entanto, nunca chegou a integrar a IA de imagens *Stable Diffusion*.

Poucos detalhes sobre os bastidores da produção estão disponíveis. A equipe concedeu algumas entrevistas curtas para portais de internet que exibem notícias de tecnologia falando sobre detalhes pontuais do projeto, mas os detalhes técnicos sobre a construção são escassos. Skyler Hartle, um dos co-criadores do show, funcionário da Microsoft, respondeu algumas perguntas dos espectadores dentro do *Reddit*, de onde é possível aferir alguns detalhes adicionais do show:

- Com exceção dos modelos pré-prontos do cenário, todo o restante do show é gerado proceduralmente;
- O show leva 2 minutos entre solicitar os dados para o *ChatGPT* e exibir os resultados na tela. Neste meio tempo, todo o processamento dos cenários, personagens, posições, falas, é realizado pelo algoritmo.
- As maiores causas de erros no show são motivadas por problemas no *ChatGPT*, e não no algoritmo gerador de roteiros.

Baseado nas respostas indicadas por Skyler, especificamente na indicação sobre os softwares e linguagens utilizados na gestão da parte da IA generativa – Python e TensorFlow – foi feita uma busca na plataforma de hospedagem de códigos *GitHub* com estes parâmetros e outros dados sobre o show. Há disponível um repositório, com um código que, testado, de fato faz requisições para a API do *ChatGPT* e gera roteiros automáticos; contudo, é impossível ter a certeza de que este código é de fato o código original, ou uma derivação.

Como o show é contínuo, muitos fatos novos que compõem o universo de *Nothing, Forever* foram gerados pelo algoritmo de IA. Os fatos citados no restante do capítulo, quando não indicados, são todos criados por inteligência artificial, e foram catalogados por fãs e dispostos na internet, em canais diversos.

Todo o universo fictício de *Nothing, Forever* se passa na cidade de *Schnitzelville*, baseada em Nova York, e interagem entre si no apartamento de Larry ou, mais ocasionalmente, no apartamento de Fred. O prefeito da cidade é *Fred Toaster*, e vários locais fictícios da cidade são citados nos episódios, como restaurantes, padarias, cafés, parques de diversões, estátuas, shoppings centers e até filmes fictícios.

São quatro os cenários base do show: o apartamento de Larry, o apartamento de Fred, o palco onde Larry faz seu show ocasional de comédia stand-up, e um take da fachada do prédio onde os dois apartamentos estão. O prédio é usado somente nas transições de *takes*, não possuindo cenas com personagens exibidas com este fundo.

Figura 18 - Take da fachada do prédio de Nothing, Forever

Fonte: Captura própria

O apartamento de Larry é um cenário recorrente do show. Possui um grande sofá verde, um pequeno sofá verde, duas banquetas de bar azuis, uma mesa de computador e uma cadeira, um microondas, uma geladeira, um telefone em cima de uma mesa de café, uma estante, quadros diversos, uma mesa de centro, um tapete, luminárias e apetrechos de decoração.

Figura 19 - Imagem do apartamento de Larry

Fonte: Captura própria

O apartamento de Fred também é um cenário do show, embora menos recorrente do que o apartamento de Larry. Apenas um lado do apartamento é visível, e contém um sofá azul grande com almofadas coloridas, uma mesa de centro desarrumada e várias decorações na parede, além de um armário. Possui um abajur azul, que ocasionalmente é ligado ou desligado por personagens do show, e causa problemas gráficos quando personagens interagem com o objeto. O cenário dos shows de Larry é apenas um pedestal e um microfone, onde o personagem coloca-se de pé e faz seu show de humor.

Conforme a geração de conteúdo ocorreu, a inteligência artificial generativa responsável pelo roteiro inseriu, em diálogos dos personagens diversos outros locais fictícios da cidade, 41 bares e restaurantes diversos, 14 lojas, inúmeras estátuas e monumentos, entre outros. Nenhum deles é exibido, apenas citado.

O grupo responsável pela criação da sitcom, *Mismatch Media,* se autointitula "laboratório de experimentação de mídia", e artigos que se referem ao produto criados pela imprensa americana se referem ao grupo como um coletivo de arte digital. Pela classificação do grupo criador – coletivo de arte digital - é razoável supor que o produto que este grupo produz seja considerado arte; contudo, a natureza artificial da criação e produção de *Nothing, Forever* leva ao questionamento da nomenclatura.

4.1.1 *Nothing, Forever* sob o viés das artes

Em um sentido amplo, arte digital se refere a obras artísticas que se apropriam de meios digitais para fazê-lo, como computadores (MARCOS, 2007, p. 70). Mas atualmente, uma grande quantidade de obras audiovisuais utiliza computadores para sua criação e (ou) produção, em variadas etapas do processo produtivo. Existem obras de arte onde os meios de computação são utilizados para dar vazão à criatividade do autor em situações em que o uso de cenas sem efeitos especiais é impossível ou não economicamente viável - e não se costuma discutir sobre se são arte ou não - mas aqui falamos de uma situação em que um produto é criado inteiramente por uma máquina.

Em 1917, Marcel Duchamp apresentou uma peça de louça usada em banheiros públicos como um trabalho artístico em uma exibição em Nova York. Duchamp assinou a peça com um pseudônimo e submeteu ao comitê responsável pela exposição, que aceitou a inclusão da obra na mostra. Várias discussões sobre o sentido e a validade da obra perduram até os dias de hoje, mais de uma centena de anos após o ocorrido, mas o consenso entre os historiadores e analistas de arte a obra de Duchamp é considerada uma forma de arte.

A produção de Marcel Duchamp permanece notória nos anais da história da arte, não em virtude de sua magnificência estética ou atributos visuais intrínsecos, mas devido à sua abordagem iconoclasta em relação à estrutura institucional, especificamente no contexto do universo expositivo artístico. Embora os produtos audiovisuais contemporâneos incorporem o meio digital como instrumental, é nos artefatos artísticos digitais de maior destaque que se observa a manifestação de características inovadoras profundamente enraizadas na manipulação e reinvenção do meio. Tais obras se distinguem por suas nuances criativas marcantes, apresentando originalidade e inventividade. O conceito de "desconstrução de um conceito" é observado nas análises sobre videoarte, e é parte da evolução contínua das artes audiovisuais. Arlindo Machado pontua:

> Em 1963, Nam June Paik, que estudava música eletrônica com Stockhausen em Colônia, teve a ideia de inverter os circuitos de um aparelho receptor de TV, para perturbar a constituição das imagens. Ao fazê-lo, ele certamente não podia imaginar que não apenas estava dando a linha diretriz de todo o posterior desenvolvimento da arte do vídeo, como também provocava uma reversão no sistema de expectativas figurativas do mundo da imagem técnica. De fato, se pudéssemos resumir numa frase a tendência geral que a chamada vídeoarte perseguiu na Europa e na América nos últimos vinte anos, diríamos que se trata, inicialmente, de distorcer e desintegrar a velha imagem do sistema figurativo, como aliás já vinha acontecendo desde muito antes no terreno das artes plásticas. (MACHADO, 1988, p. 117)

O produto da *Mismatch Media* apropria-se de um conceito comum, tal qual um objeto "*Ready Made*", como Duchamp fez: é uma sitcom, formato televisivo comum desde a década de 1950; emula personagens, trejeitos e situações de Seinfeld, um produto já consolidado; usa ângulos de câmera, cenografia e arquivos sonoros comuns. Mesmo os diálogos gerados por inteligência artificial, embora sejam sem dúvida uma novidade tecnológica evidente, são dotados das mesmas qualificações que um diálogo escrito por um humano, e, caso o show não alardeasse que seus textos são fruto de uma IA generativa, muitos considerariam um produto indistinguível dos feitos da forma clássica.

Nothing, Forever é um produto híbrido, formado por vários softwares, cada um realizando a sua parte no todo: os diálogos são gerados por uma IA, interage programaticamente com um motor gráfico, que por sua vez envia seus dados para um software responsável por gerar os arquivos de vídeo, que, finalmente, são enviados pela internet ao serviço de exibição. Estes *softwares* também possuem seus próprios *softwares* acessórios que são responsáveis por partes de suas tarefas. Todo este processo é capitaneado pela inteligência artificial, que age representando vários papéis no fluxo de produção que seriam ocupados por seres humanos, como direção de imagem, autoria e sonoplastia, entre outros. O processo generativo da IA (leia-se aqui como um análogo – com ressalvas - ao termo "processo criativo") é próprio, e difere da maneira que um humano geraria estas informações, e, mesmo com esta distinção, o produto resultante pode ser identificado como uma produção audiovisual realizada por humanos. É fato que é justamente esta a intenção de uma obra de uma IA, mas, com tantas variáveis distintas, por que isso ocorre?

Embora o processo de gênese de informação de uma IA seja puramente matemático, ele ainda é suscetível a uma espécie de "vinculação à projeção": abastecido com gigantescas bases de dados abstratas, e liberado para fazer as associações que bem entender com estes dados – ainda que seus criadores, humanos, não entendam essas associações feitas pela máquina muito

bem – ao final, o produto ainda é muito parecido com o qual um humano faria. É semelhante a uma representação imagética de uma fotografia em um monitor de computador, feita por milhões de pontos compostos por cores primárias, que se acendem e apagam de acordo com uma regra específica, mas que, vistos pelo humano, tornam-se uma representação visual da foto original, mesmo sem nunca ter sido esta foto.

> Ao decompor a imagem móvel, obtida por projeção ótica sobre o fundo fotossensível de uma câmera eletrônica, em finas linhas paralelas, semelhante ao pantelégrafo, a televisão tornava-se capaz de analisar cada ponto de cada linha da imagem e de reconstituir a imagem sob a forma de uma espécie de mosaico luminoso. [...] Mas nem sempre era possível controlar o ponto de imagem com perfeita exatidão, agir por exemplo sobre um único ponto, independentemente dos outros. Faltava ao mosaico eletrônico ser completamente ordenado, ao ponto de a imagem ser numerizada, isto é, indicável exatamente na tela através de coordenadas espaciais e cromáticas definidas por um cálculo automático. Essa última etapa na busca do menor elemento constituinte da imagem foi superada graças ao computador. (COUCHOT, 2011, p. 37)

Couchot fala sobre uma mudança de paradigma na representação visual imagética: a capacidade de um computador controlar individualmente os *pixels* presentes em uma tela e possibilitar a fragmentação de uma imagem em uma matriz de pontos individuais, que ele considera como o menor elemento constituinte da imagem. A representação de apenas um *pixel*, isoladamente, não revela uma imagem; um grupo isolado de *pixels* também não. A representação final da imagem é composta do todo de *pixels*, vistos a uma distância adequada. A imagem composta por um conjunto de *pixels* é uma mero simulacro do objeto captado, mas que o humano reconhece como imagem, do mesmo modo que uma saída de texto do *ChatGPT* é uma representação matemática composta por relações entre termos e palavras que, vistas isoladamente, não representam nada para um humano, mas agrupadas como um todo pelo processamento da máquina, representam um conteúdo que um humano entende.

O mecanismo subjacente à geração de conteúdo por um LLM como o *ChatGPT* é intrinsecamente baseado em cálculos matemáticos, envolvendo operações complexas como somas, multiplicações e adições. Essas operações são aplicadas a vetores numéricos que representam palavras ou sequências de palavras, com pesos atribuídos que indicam a probabilidade de uma palavra seguir outra em um dado contexto. Esse processo estatístico de decisão é o que determina a seleção da palavra seguinte, culminando na construção de frases complexas.

No caso de *Nothing, Forever* (e produções semelhantes), essas frases complexas são utilizadas para compor diálogos, descrições de ações e elementos de cenografia, criando uma narrativa coerente. Apesar de todo o processo ser conduzido por algoritmos sem intervenção humana direta nas escolhas específicas de palavras ou construções frasais, o resultado apresenta características que podem ser confundidas com humanas. Entender como este processo ocorre é importante: aqui, não se trata de um mero cálculo matemático em eixos de produção de sentido, mas em emular toda uma construção de ambientação que abrange inúmeras variáveis que ultrapassam a capacidade teórica de um modelo de linguagem. Este entendimento é obtido ao considerarmos a produção gerada por uma inteligência artificial como um subproduto direto de seu extenso corpus de treinamento, o qual é integralmente composto por conhecimento humano preexistente, meramente amalgamado em um determinado instante. Analogamente, quando observamos um papagaio replicando com precisão um segmento do idioma humano, não se identifica a presença de uma intenção comunicativa subjacente, nem uma cognição do entorno; o que ocorre é simplesmente a emissão estocástica de um fragmento de texto que, embora possa parecer contextualizado, sua compreensão é intrinsecamente limitada.

> Text generated by an LM is not grounded in communicative intent, any model of the world, or any model of the reader's state of mind. It can't have been, because the training data never included sharing thoughts with a listener, nor does the machine have the ability to do that. This can seem counter-intuitive given the increasingly fluent qualities of automatically generated text, but we have to account for the fact that our perception of natural language text, regardless of how it was generated, is mediated by our own linguistic competence and our predisposition to interpret communicative acts as conveying coherent meaning and intent, whether or not they do. The problem is, if one side of the communication does not have meaning, then the comprehension of the implicit meaning is an illusion arising from our singular human understanding of language (independent of the model). Contrary to how it may seem when we observe its output, an LM is a system for haphazardly stitching together sequences of linguistic forms it has observed in its vast training data, according to probabilistic information about how they combine, but without any reference to meaning: a stochastic parrot. (BENDER, GEBRU, *et al.*, 2021, p. 616-617)

Voltando à Couchot, os LLM podem ser considerados para o texto verbal aquilo que os *pixels* são para as telas de TV ou monitores, permitindo que o nexo derive da reordenação matemática das palavras em uma emulação matemática do sentido. Não há um suporte tangível em um modelo matemático, não há um referente concreto em um texto gerado por LLM, mas há um elo morfológico e sintático entre o produto da IA e de um texto verbal humano:

> O computador permitia não somente dominar totalmente o ponto da imagem - pixel - como substituir, ao mesmo tempo, o automatismo analógico das técnicas televisuais pelo automatismo calculado, resultante de um tratamento numérico da informação relativa à imagem. [...] O pixel lançava as técnicas numéricas de figuração numa lógica em total ruptura com a lógica figurativa subjacente à imagem gerada até então pelos procedimentos óticos. A hibridação inesperada de um calculador eletrônico e de uma tela de televisão iria provocar, no universo das imagens, a mutação mais radical desde o aparecimento - há mais ou menos uns vinte e cinco mil anos - das primeiras técnicas de figuração. (COUCHOT, 2011, p. 37)

Couchot considera que a introdução do controle do computador a representação imagética através de *pixels* representava a maior mudança de paradigma nas técnicas de representação da figuração desde a época em que os humanos fizeram as suas primeiras pinturas rupestres. É possível traçar um paralelo entre essa transformação na representação imagética e a mudança de paradigma na representação verbal que os LLM trazem, com um controle granular da gramática, que consiste em uma classificação individual das palavras e morfemas que permite uma saída textual coesa e dotada de sentido.

Segundo Couchot, a hibridização entre o computador e uma tela de televisão era motivadora de uma revolução na representação imagética. *Nothing, Forever* é a representação imagética da hibridização entre a inteligência artificial e a tela da televisão, e apresenta uma ruptura radical nos conceitos de uma produção audiovisual. Dentre as mudanças que *Nothing, Forever* traz, a subversão do fluxo é a mais significativa.

4.2 A subversão do fluxo

A inovação fundamental de *Nothing, Forever* reside em sua capacidade de subverter o paradigma convencional de consumo de conteúdo audiovisual, oferecendo uma experiência dinâmica e ininterrupta. Tradicionalmente, obras audiovisuais são finitas; elas têm um início, meio e fim claramente definidos, permitindo que os espectadores assistam e reassistam sabendo exatamente o que esperar. No entanto, *Nothing, Forever* desafia essa noção ao apresentar uma narrativa gerada continuamente por inteligência artificial que não possui um término predeterminado.

Espectadores que acessam o link de apresentação de *Nothing, Forever* se deparam com um segmento único da narrativa, que é parte de uma trama essencialmente infinita. Isso significa que, se decidirem parar de assistir e retornarem após um intervalo - seja de algumas horas ou um dia - descobrirão que perderam um volume de conteúdo que continuou a se desenvolver na

sua ausência. Este aspecto contínuo e perpetuamente evolutivo da obra indica sua inovação: a criação de um universo narrativo que vive e se expande sem pausa, refletindo a natureza incessante e autossustentável da própria inteligência artificial que o gera.

Esta abordagem não apenas redefine a interação entre o conteúdo audiovisual e seu público, mas também amplia as fronteiras da criação artística e da expressão narrativa. Ao invés de se manterem como consumidores passivos, os espectadores têm uma experiência que é ao mesmo tempo imersiva e efêmera, onde cada momento é único: *Nothing, Forever* possibilita uma exploração das possibilidades que a tecnologia de inteligência artificial abre para as artes e o entretenimento, questionando e expandindo as noções tradicionais de autoria, narrativa e engajamento do espectador.

Obras audiovisuais possuem começo e fim, e são concebidas assim. O autor possui uma ideia base, cria um roteiro baseado nesta ideia, e, posteriormente, este roteiro é produzido. A totalidade das obras audiovisuais feitas para o cinema são assim, pela natureza do formato sequencial: um quadro no cinema é um fotograma, um recorte estacionário do espaço-tempo a qual, posteriormente, em outro contexto – normalmente uma sala de cinema – é dada a ilusão de movimento. No cinema, é inviável tecnicamente que seja diferente: o cinema é sempre o "depois", situado a frente do tempo da concepção original.

> A primeira diferença básica entre a constituição da imagem fílmica e a da imagem televisual ou videográfica está no fato de a primeira ser gravada em quadro fixo e na sua totalidade de uma só vez, enquanto a segunda é "escrita" sequencialmente por meio de linhas de varredura, durante um intervalo de tempo. No filme, a imagem é inscrita em fotogramas separados: entre um quadro e outro, o obturador se fecha impedindo a entrada de luz, e uma nova porção de película virgem é empurrada para a abertura. Esse movimento fragmentário, que denuncia a base fotográfica do cinema, é dissimulado, entretanto, por um dispositivo técnico, para que se possa recompor a ilusão de movimento. (MACHADO, 1988, p. 41)

Arlindo Machado (1988) pontua que a diferença entre a constituição da imagem fílmica e da imagem televisual é a natureza da formação da imagem do televisor, que é feita através de varredura de linhas horizontais, compostas por *pixels*. Chama de "mosaico de retículas" essa composição que, na época, no final da década de 90, era mais pronunciada do que hoje: segundo Machado, a sensação de realidade era prejudicada pelos aspectos técnicos da formação do fotograma virtual.

> O vídeo, porém, retalha e pulveriza a imagem em centenas de milhares de retículas, criando necessariamente uma outra topografia que, a olho nu, aparece como textura pictórica diferente, estilhaçada e multipontuada, como os olhos das moscas. [...] Aí está justamente o traço distintivo fundamental que separa cinema e vídeo: neste último, o processo genético de constituição da imagem está à mostra, impedindo que a restituição do mundo visível se dê à custa do mascaramento de técnicas construtivas. O vídeo, mesmo nesse nível genético mais elementar, já exibe a enunciação da imagem pelos meios técnicos, em prejuízo inclusive do ilusionismo de realidade, que no cinema é a base da verossimilhança.

Duas décadas e meia depois, as semelhanças entre as mídias já são maiores que as diferenças, por diversos fatores. A problemática da resolução da imagem ser causa da perda da imersão já não é mais uma realidade desde as primeiras telas usadas pela Apple no seu iPhone 4, em 2010, que trouxeram os chamados “Retina Display”, onde a densidade de pixels da tela é maior do que a capacidade humana de enxergar pixels individuais, dependendo da distância entre o usuário e o aparelho. Hoje, em 2024, são comuns telas com resolução 4K (3840 pixels x 2160 pixels), dezenas de vezes mais alta do que as telas de televisores analógicos, comuns na época. Além disso, hoje, muitas obras cinematográficas são gravadas em câmeras digitais, e muitas obras de televisão são gravadas em película.

Figura 20 - Diferença entre resoluções de monitores e televisores

Fonte: https://www.hkvstar.com/technology-news/what-s-sd-hd-full-hd-ultra-hd.html Acesso em 12 Out. 2022

Embora as diferenças entre as mídias sejam muito menores que outrora, uma distinção fundamental apontada por Machado continua válida: a forma de inscrição temporal da imagem na exibição, linha por linha, criadas em tempo real, continua vigente. Mesmo os monitores de resolução muito mais alta do que a visão humana pode captar, produzem imagens do mesmo modo que uma televisão analógica fazia, uma linha por vez. Tecnicamente, a imagem é desenhada tal qual como no cinema, quadro isolado por quadro isolado: a varredura de linhas,

hoje, na prática, é inacessível para manipulação por parte de quem produz mídia. Os experimentos de videoarte nos anos 80 e 90, que envolviam a manipulação do aparelho de vídeo diretamente, em meio ao fluxo de desenho, não são mais possíveis hoje devido ao uso de técnicas de "*buffering*", onde o quadro todo é armazenado em memória volátil, para ser desenhado posteriormente. Isso se dá pela diferença de velocidade entre um aparelho de televisão (525 linhas, 60 vezes por segundo, com as linhas pares sendo desenhadas em um quadro, e as ímpares no outro), para um monitor moderno (muitas vezes 2160 linhas, 120 vezes por segundo, com todas as linhas sendo desenhadas em todos os quadros). Mesmo assim, diferentemente do cinema, um quadro de vídeo é desenhado exatamente no fragmento de segundo que ele é enviado, com poucos milissegundos de atraso.

No cinema, transmissões ao vivo são, por *design*, impossíveis. Já nos aparelhos de televisão, antes do advento do *videotape*, eram o único modo de transmissão possível. Com a criação das gravações em fita, muitos programas de TV passaram a ser gravados, mas ainda hoje temos muitos programas, eventos esportivos, shows, e outras produções, transmitidas ao vivo. Atendo-se apenas a transmissão ao vivo de programas ficcionais, são muito incomuns programas que não sejam previamente planejados, dependendo exclusivamente do improviso do ator. Dentre estas escassas produções, está o humorístico exibido pela TV Globo em 2015 "Tomara que Caia", que foi exibido ao vivo e tinha plateia, que interagia com o elenco, que baseado em um roteiro mínimo, improvisava no palco.

Em um programa de televisão com um roteiro prévio, embora a ação esteja ocorrendo naquele exato momento, a equipe de produção, antes da apresentação, já se reuniu e planejou o processo todo da produção, que, embora tenha uma variação espontânea presente, ainda é um produto planejado previamente. *Nothing, Forever* não é completamente desprovida de um planejamento prévio, já que seus idealizadores decidiram os biotipos dos personagens, cenários e definiram aspectos básicos da narrativa, mas toda evolução do programa após o seu início é dada pela interação interna maquínica, sem intervenção humana, em evolução sequencial.

Os diálogos de *Nothing, Forever* são derivados de um fluxo semi-contínuo: não é completamente em tempo real, já que os diálogos são gerados momentos antes de aparecerem no vídeo. Mas são completamente autônomos, com suas derivações baseadas na própria construção dos personagens dada pelas interações anteriores, sem nenhum tipo de intervenção dos autores – embora haja algum tipo de intervenção interativa, com opiniões dadas no chat influenciando os resultados do programa.

As diversas continuidades (e não-continuidades) e descontinuidades (e não-descontinuidades) de *Nothing, Forever* podem ser analisadas sob o arcabouço teórico da semiótica discursiva de linha francesa. No final da década de 1980, o semioticista francês Jean-Marie Floch foi contratado para fazer uma consultoria para uma identificação dos perfis dos usuários de transporte público da cidade de Paris, e fez uma análise semiótica dos trajetos (FLOCH, 2023).

A comparativa da análise de trajetos de passageiros de trem com as peculiaridades do fluxo temporal de *Nothing, Forever* é conexa e pertinente por características que aproximam as duas realidades: no caso do transporte, um passageiro chega, interage com o serviço de transportes e vai embora. Porém, durante a sua ausência, o serviço de transporte continua funcionando, e outros passageiros também interagem com o mesmo serviço, em um determinado espaço-tempo. Ocasionalmente, usuários se encontram e podem interagir entre si, ou com o ambiente (ou não), e, quando o usuário volta no dia seguinte para usar a mesma linha de transporte público, o ciclo se repete. No caso de *Nothing, Forever*, o espectador passa por uma experiência semelhante, onde acompanha um fragmento temporal de uma *sitcom* que continuará existindo mesmo quando ele não estiver assistindo. Se interagir no chat, e o algoritmo julgar que a interação é pertinente, o fluxo pode ser modificado; se outro espectador fizer essa modificação em um momento em que ele não estiver assistindo, essa interação pode fazer alguma diferença – ou não – para quem assistir posteriormente.

O primeiro ponto a se apreciar é como categorizar a interação de um usuário com o fluxo de *Nothing, Forever*, e um trajeto de trem, como um texto, para viabilizar a análise semiótica dos elementos apresentados. Sobre isso, Floch (2023) diz:

> Por que é possível conceber o trajeto como um texto e, consequentemente, submetê-lo a uma análise semiótica? [...] porque, como qualquer texto, o trajeto possui uma completude que o individualiza como totalidade relativamente autônoma e confere a ele a possibilidade de uma organização estrutural. Toda trajetória percorrida por um viajante possui um desfecho, uma saída, que implica simetricamente uma entrada. Outra razão: tal como um texto, o trajeto presta-se a uma segmentação, isto é, a ser recortado em um número limitado de unidades, etapas ou "momentos" interligados conforme certas regras. Essas unidades não precisam coincidir com os espaços catalogados e denominados como tais (corredores, saguões, plataformas, trens...). Identificáveis, como veremos na anotação do rastreamento, por determinada sequência ou macrossequência gestual (móvel / imóvel, de pé / sentado, acelerado / desacelerado) ou proxêmica (abertura para os outros / ensimesmamento, distância / proximidade, encontro frontal / tangencial), tais unidades, tais sintagmas podem subdividir os locais usuais ou, ao contrário, agrupá-los em um só espaço. Terceira razão: do mesmo modo que um texto, o trajeto tem uma orientação. (FLOCH, 2023, p. 3)

Do mesmo modo que um trajeto de trem, a experiência do espectador de *Nothing, Forever* também possui uma completude e é passível de ser organizada estruturalmente: uma sessão de um espectador tem um início, e um fim; ninguém é capaz de assistir toda a obra. A experiência do espectador também pode ser segmentada, e, tal qual a experiência do viajante de trem, não necessariamente pelo começo e fim de um bloco narrativo, já que o espectador pode assistir vários blocos, ou parte de um bloco, ou ainda nenhum bloco, caso entre e saia durante uma das pausas entre blocos, ou durante a manutenção do sistema, por exemplo.

As análises de Floch se deram em vários níveis. O trajeto, como objeto global, aquilo que todos percorriam, era analisado como processo significante. Cada viagem, individual, de cada passageiro, e suas características, também como objeto significante. Com isso, Floch reconheceu a categoria fundamental que classificava todas as experiências: a relação de oposição entre continuidade e descontinuidade (FLOCH, 2023, p. 8). Essa noção de contínuo e descontínuo frequentemente explorada dentro do contexto da semiótica de Eric Landowski, que, em sua obra, se interessa particularmente pela dinâmica das práticas significativas nas sociedades, explorando como os processos de significação operam tanto em níveis contínuos quanto descontínuos.

Landowski (2014), identifica quatro regimes de interação que descrevem diferentes modos pelos quais as entidades podem se relacionar, cada um envolvendo diferentes níveis de agência, consciência e reciprocidade, sendo eles: Programação, Manipulação, Ajustamento e Acidente. O autor fundamenta todos estes regimes na relação de oposição entre continuidade e descontinuidade, sendo que cada um dos regimes possui um nível de estabilidade dado pelo grau de continuidade.

O regime de programação envolve relações determinadas por regras e normas que guiam o comportamento dos indivíduos. Há uma expectativa de que essas normas sejam seguidas, criando um sentido de previsibilidade e ordem. A programação pode ser vista em contextos sociais onde as convenções culturais, as leis ou os procedimentos operacionais padrão ditam como os indivíduos devem agir. No caso de *Nothing, Forever*, a máquina responsável por gerar seus roteiros e interações baseia a estrutura destes textos em inúmeros textos de *sitcoms* criadas anteriormente, sendo esta programação inicial garantidora que a estrutura geral do produto audiovisual siga um formato reconhecido e atenda as expectativas do gênero. Mesmo em um fragmento temporal, quando um espectador entra e assiste um trecho do programa, a programação restringe o escopo da *sitcom* a própria definição de *sitcom*, tornando aquele fragmento algo reconhecível. Este regime é marcado por uma alta continuidade, uma vez que

as ações e interações são guiadas por regras, normas e convenções estabelecidas. No caso de um LLM pré-treinado, essa continuidade é garantida, já que os dados iniciais do modelo são perpétuos na sua estrutura, e todos os textos gerados pelo modelo usarão o banco de dados do treinamento inicial. Quebras de continuidade neste modo seriam possíveis apenas em erros do *software*.

O uso da manipulação, no caso de *Nothing, Forever*, é mais complexo, e pode ocorrer em vários níveis. Segundo Landowski (2014, p. 30), a manipulação pode ser entendida como um regime de engajamento onde um agente (o manipulador) busca influenciar outro agente (o manipulado) para agir de uma certa forma. Landowski enfatiza que a manipulação é uma modalidade de interação que se baseia em uma compreensão profunda dos mecanismos de significação e das predisposições do manipulado, permitindo ao manipulador orquestrar cenários onde as escolhas do manipulado parecem autônomas, mas são, de fato, condicionadas.

Programas de televisão e filmes frequentemente manipulam os espectadores ao construir narrativas que provocam respostas emocionais específicas, como empatia, raiva ou medo, para transmitir uma mensagem ou moral. Um exemplo seria um documentário que apresenta uma visão particular sobre um assunto controverso, selecionando cuidadosamente as evidências e os testemunhos apresentados para persuadir os espectadores a adotar uma determinada perspectiva. Como *Nothing, Forever* é gerado através de um banco de dados que, teoricamente, compõe todas as manipulações possíveis, e o gerador é uma máquina – logo, completamente livre, *per se*, de intenções de manipulação – é complexo definir onde essa possível manipulação se dá. Ela pode se dar na entrada inicial do responsável gerador, ou pode se dar de modo intencional pelo *corpus* de treinamento (que também é opaco, pois não há divulgação exata de qual é), ou de modo não intencional por este mesmo *corpus*, já que, não intencionalmente, dada a quantidade gigantesca de dados coletados pelos treinadores, pode haver vieses imbuídos na base de dados cujos quais os criadores da IA não tem controle. O banimento temporário da *sitcom devido* a comentários transfóbicos da inteligência artificial é uma exemplo desta situação. A natureza estocástica da geração de conteúdo de um LLM causa uma relativa imprevisibilidade neste cenário, no caso de *Nothing, Forever.*

O regime de ajustamento, segundo Landowski (2014), refere-se a um modo de interação baseado na reciprocidade e na adaptação mútua entre as entidades. Neste regime, os actantes estão continuamente ajustando seus comportamentos em resposta uns aos outros, criando um fluxo dinâmico de ação e reação. O ajustamento é caracterizado por uma negociação constante de significados e intenções, permitindo uma coordenação das ações. No caso de *Nothing,*

Forever, estes ajustes são feitos em vários níveis: o *feedback* dos espectadores no chat pode influenciar na narrativa, os ajustes dos programadores de acordo com as dificuldades técnicas das integrações entre os softwares, os ajustes no próprio GPT, que é continuamente aperfeiçoado, entre outros fatores, ajustam a *sitcom*, e próprio público também ajusta suas expectativas com base no conteúdo gerado. Como exemplo aqui, a troca de protagonistas entre a primeira e a segunda temporada, feita para distanciar a *sitcom* do seu produto inspirador (Seinfeld), é uma forma de ajustamento. A catalogação por parte do público de todos os atos do seriado, de forma coletiva, em uma *wiki*, também pode ser considerada uma forma de ajustamento de um público que tenta lidar com um novo formato de consumo de mídia. Há aqui um equilíbrio entre continuidade e descontinuidade, já que mudanças graduais são mais harmônicas do que quebras bruscas de paradigma, como a troca de todos os personagens principais, embora a *sitcom* mantenha sua proposta inicial.

O regime da acidentalidade promove o papel do acaso e dos encontros inesperados nas interações. O acidente introduz uma dimensão de imprevisibilidade e contingência, onde os eventos ocorrem fora do controle dos indivíduos e podem ter impactos significativos sobre as relações e os significados. A acidentalidade indica que, apesar dos esforços para estruturar e prever as interações, sempre há elementos que escapam à nossa capacidade de planejamento e controle.

> Em “Da Imperfeição”, Greimas adota como ponto de partida o contínuo: para ele, ou pelo menos nos textos que analisa, o mundo é na origem um mundo da ordem. Como nos contos. Trata-se mesmo de um mundo tão ordenado, onde todas as interações são tão bem programadas que, para não cair numa letargia mortal, é necessário, custe o que custar, sair dessa “rotina”. Por isso, Greimas inventa, ou reinventa, a estesia, a sensibilidade, o corpo, em suma, as condições mesmas do que chamamos o ajustamento: com ajuda de um acidente que permita a negação ou a superação dos programas fixados de antemão, ocorrerá a passagem de um cotidiano marcado pelo máximo de segurança possível, e consequentemente pela insignificância e o enfado, a uma vida “outra” na qual as relações entre actantes não serão nada seguras mas onde, em contrapartida, elas farão sentido. Mas o tédio não seja, talvez mais que um privilégio de ricos, enquanto a dor é familiar a todos. Partamos, portanto, do ponto oposto àquele de greimas e dos contos: e se tudo na origem fosse, como nos mitos, acidente, intranquilidade, instabilidade, agitação e fúria, desordem, caos – numa palavra, descontinuidade? (LANDOWSKI, 2014, p. 70)

Nothing, Forever é uma obra altamente influenciada pela acidentalidade. Inúmeros fatores dentro da geração da sitcom são passíveis de eventos aleatórios: o comportamento do GPT é imprevisível, dada sua natureza estocástica. As múltiplas interações entre os diversos

programas que compõem a geração final podem apresentar erros e resultados fora do padrão esperado, a recepção do público a uma forma completamente nova de se produzir um conteúdo audiovisual também não é passível de se prever.

Voltando a pesquisa de Floch (2023), após traçar os perfis de vários tipos de usuários, o autor chegou a conclusões de que poderia classificá-los em tipos específicos: agrimensores, sonâmbulos, errantes e profissionais, cada um destes tipos com especificidades comportamentais distintas. Floch, em determinado ponto, notou que a rede de transportes de Paris contava com uma rede de videocomunicação por fibra ótica, com terminais instalados nos trens e estações, chamada de TUBE. Esta rede funcionava de modo muito semelhante à rede do Metrô de São Paulo, apresentando programas curtos, mensagens de patrocinadores e informações úteis para serem consumidas pelos usuários da rede de transportes. Na pesquisa de Floch, todos os quatro tipos de passageiros tinham uma impressão positiva sobre a rede de informação, mas cada um dos tipos tinha críticas a respeito de particularidades da programação.

Nothing, Forever é uma experiência audiovisual que, embora inovadora, ainda mantém regras e padrões tradicionais de geração e exibição. Embora cause uma descontinuidade na expectativa em como se constrói um produto audiovisual e também seja responsável pela ruptura do paradigma temporal dentro do fluxo televisivo tradicional, ainda há espaço para mais uma fronteira que *Nothing, Forever* não explora: a geração personalizada de conteúdo. Tal qual a exibição no sistema de transportes de Paris, *Nothing, Forever* ainda traz a visão específica de um grupo de pessoas que concebeu e ofereceu a *sitcom* como produto. No entanto, não há impeditivos técnicos para que fosse lançada uma versão de *Nothing, Forever* que tivesse um conteúdo personalíssimo; onde cada usuário poderia influenciar sua própria versão da história e manipulá-la como bem entender. Nesta hipotética versão, a *sitcom* seria renderizada no aparelho de cada usuário, com entradas de manipulação dadas por ele, que se refletiriam nos resultados da *sitcom* e criariam um produto essencialmente único.

Neste ponto, em se tratando deste produto hipotético, o fluxo temporal deixa de ser um problema: o usuário poderia pausar a geração e continuar a ver de onde parou. Ou, por opção, poderia deixar a simulação continuar, e ter o fator surpresa quando voltar a assistir. Olhando por este prisma, o fluxo temporal infinito de *Nothing, Forever* é, embora uma descontinuidade do padrão televisivo vigente, tem o potencial de ser considerado uma continuidade, com a abolição deste fluxo sendo o próximo passo nos produtos audiovisuais. Esse conceito causa impactos significativos em várias áreas da produção audiovisual, como a relação de autoria, a forma de consumo, a precificação, e até uma discussão mais ampla da indústria em si.

4.3 Contratos fiduciários entre humanos e máquinas

Nothing, Forever, como produto, desconstrói conceitos estabelecidos na indústria audiovisual se apropriando de recursos técnicos que eram impossíveis de se conceber até o advento da inteligência artificial, contudo, o resultado é comparável a um produto audiovisual clássico. Tomemos por exemplo a sua natureza temporal infinita: de fato, a *sitcom* pode ser exibida continuamente e eternamente, sem reprises; contudo, embora o tempo de tela (quantidade de tempo que as pessoas passam consumindo conteúdo audiovisual) da população tenha aumentado, é improvável que alguém passe o tempo todo assistindo apenas *Nothing, Forever*.

Quando alguém assiste um capítulo de uma novela, várias cenas são apresentadas, em pontos distintos de espaço-tempo narrativo, mas dentro de uma janela de tempo real específico, geralmente uma hora. O capítulo assistido foi gravado anteriormente, baseado em um roteiro onde todas estas iterações espaço-temporais estavam previamente definidas, e os próximos capítulos também já estão previstos. No caso de *Nothing, Forever*, quando o espectador decide assistir, digamos, a mesma uma hora de um fragmento do programa, vai ter uma sensação bastante semelhante a assistir um capítulo de uma novela, com uma estrutura narrativa semelhante: cortes de câmera dando a ilusão de fluxo, cenas em locais distintos e padrões temporais fluídos, entre outras semelhanças.

O ato enunciativo da figuratividade é semelhante, entre um produto audiovisual clássico e um contínuo, produzido por IA, caso a análise se atenha a um fragmento temporal específico não linear. Neste recorte temporal (e somente neste), o discurso, como resultado das escolhas do sujeito da enunciação, é semelhante: as escolhas sintáticas predominantes o espaço, os actantes e principalmente, o tempo, são comparáveis entre um capítulo de um produto audiovisual seriado clássico e *Nothing, Forever*. Como estas escolhas são comuns entre a mídia dita “tradicional” e a construída por uma inteligência artificial, o espectador vê os produtos de forma semelhante.

> As estruturas narrativas convertem-se em estruturas discursivas quando assumidas pelo sujeito da enunciação. O sujeito da enunciação faz uma série de “escolhas”, de pessoa, de tempo, de espaço, de figuras, e “conta” ou passa a narrativa, transformando-a em discurso. O discurso nada mais é, portanto, que a narrativa “enriquecida” por todas essas opções do sujeito da enunciação, que marcam os diferentes modos pelos quais a enunciação se relaciona com o discurso que enuncia. (BARROS, 2005, p. 53)

A semiótica Greimasiana, através do seu percurso gerativo de sentido, indica que não há um plano de conteúdo sem um plano de expressão. A junção destes dois planos é que traz o texto (FIORIN, 2003). Essa soma, que possibilita a análise imanente do produto audiovisual produzido, é semelhante quando o fragmento do espaço-tempo do espectador-enunciatário é semelhante entre, no caso, as obras audiovisuais convencionais e *Nothing, Forever*: se vista em capítulos, são muito parecidas e passíveis da mesma análise. No plano de conteúdo, ambos, tanto o que foi produzido por uma IA, quanto por um humano, em um conceito de narrativa tradicional, se assemelham. No entanto, no plano da expressão, *Nothing Forever* possui, no todo, uma construção espaço-temporal completamente distinta dos produtos audiovisuais clássicos. Quando você não está vendo um episódio, a *sitcom* continua sendo produzida. Esse efeito, em menor escala, é semelhante a deixar de ver um capítulo de uma novela e assistir o seguinte, mas muito mais pronunciado, já que seres humanos necessitam dormir, realizar suas necessidades básicas, trabalhar, entre outros usos temporais em que não estamos consumindo mídias audiovisuais. A diferença principal aqui é que, caso você perca um capítulo de uma novela, pode assisti-lo depois em outro horário (caso tenha acesso a gravação ou *streaming*) e pode se atualizar quanto aos acontecimentos da narrativa. Com *Nothing, Forever*, embora você possa fazer o mesmo, o mero ato de assistir um fragmento perdido lhe impede de acompanhar o fragmento vigente, tornando impossível acompanhar o todo da narrativa contínua.

O LLM, como enunciador, é um amálgama matemático de vários atores da enunciação, que projetam o tempo, o espaço e os atores do enunciado, criando figuras e temas que sustentam os programas narrativos. Essa massa de atores diversos, definiu, através de padrões matemáticos dados pelo aprendizado de máquina de IA, de que os parâmetros que indicam a humanidade de um texto podem ser encontrados pela verossimilhança de outros milhões de textos escritos por um humano, através de padrões. Portanto, é possível aferir que esse contrato fiduciário de figuratividade se dá em dois momentos: primeiramente, quando o enunciador humano apresenta a massa de dados para a máquina (aqui, como enunciatário), e depois quando a máquina, provida do contrato inicial com o humano, tenta firmar com outro humano um outro contrato, agora como enunciador, após o processamento dos padrões iniciais.

Podemos definir o processo como uma relação dialética em várias etapas:

1 – Humanos (muitos) produzem muitos textos. Consideremos aqui, para efeitos de exemplo, toda a produção humana documentada e acessível.

2 – Humano (um, poucos) juntam estes muitos textos, em uma seleção própria, baseada em seus próprios conceitos, portanto, agindo como enunciador para um enunciatário maquínico, recebedor destas informações, filtradas por este agente humano inicial.

3 – Máquina (uma) apresenta ao humano (um, muitos) ações contínuas de um roteiro audiovisual infinito, gerado em tempo real, baseadas no entendimento maquínico das informações dadas pelo humano inicial, que foram filtradas de toda a produção humana de textos anteriores.

No primeiro ponto, há uma enunciação em cada texto, individualmente. No conjunto, eles podem ou não se relacionar entre si, embora vivam em um mesmo universo conjunto. O todo tende à abstração, já que são muitos assuntos distintos juntos: um corpo de treino de IA pode conter desde receitas culinárias até obras em línguas mortas, passando por cartas de suicídio e livros de piadas, entre outros assuntos. Cabe a IA, no todo deste textual, encontrar os padrões que lhe forem convenientes, de acordo com as instruções iniciais dadas pelo seu criador.

No segundo ponto do processo, há uma outra enunciação: o humano que forneceu estes textos escolheu quais textos disponibilizou, e a máquina só tem acesso ao conteúdo que o humano lhe dá acesso; logo, há um viés. Todo o conteúdo produzido pela máquina, no futuro, será baseado naquilo que inicialmente lhe foi fornecido, mesmo que haja uma extrapolação virtual possível, onde a máquina afira conceitos abstratos derivados do que lhe foi dado, figurativamente, no ato da criação do seu banco de dados. Aqui, podemos assumir um conceito de figuração e abstração advindo da teoria estética, aplicado a produção audiovisual. Bertrand (2003) introduz este conceito das artes ao texto, em uma dicotomia que divide em dois conceitos o discurso: o figurativo e o abstrato, embora elementos presentes em um possam estar presentes no outro, sem prejuízo a sua classificação. À máquina é oferecido um universo textual, figurativo, onde, através de algoritmos matemáticos de produção de padrões, a IA afere conceitos abstratos próprios, que, importante frisar, não podem ser considerados como característicos de pensamento autônomo. São abstrações próprias à máquina, muitas vezes de difícil (ou impossível) entendimento pela lógica humana, mas, para todos os efeitos, simples derivações do que já havia no texto original. Pode-se dizer, em uma explicação sucinta, que a máquina enxergou, nos textos que lhe foram fornecidos, padrões que os humanos não percebem, ou não se importam e ignoraram.

No terceiro ponto, inicialmente, cabe um questionamento: é possível afirmar que a máquina age apenas como intermediário, ocultando o enunciador original? As relações

humanas com as máquinas são, tradicionalmente, de subserviência. É natural conceber que, aqui, um LLM esteja apenas atuando como um mero intermediário opaco entre o enunciador original – o homem que programou a IA – e o espectador, que está vendo sua produção. No entanto, a questão da relação entre uma IA, seu treinamento e posterior uso por humanos é complexa e multifacetada: um LLM não apenas repete informações de seus dados de treinamento, mas sim aprende padrões e relações – das mais básicas até as muito complexas e não notadas por humanos – entre os dados disponíveis.

Neste ponto, a semiótica discursiva francesa, que se fundamenta na análise pura do objeto textual, é bastante eficaz. Greimas é categórico ao afirmar a autonomia do texto: "Fora do texto, não há salvação". A semiótica greimasiana atribui o discurso a camada superficial do percurso gerativo de sentido, e é vista separadamente da camada semionarrativa, mais profunda. Fontanille (2008), em um artigo onde disserta sobre as múltiplas teorias que se debruçam sobre o discurso, ressalta a posição de Greimas:

> Para a semiótica greimasiana "standard" dos anos 1970-1980, o discurso corresponde à camada superficial do percurso gerativo do sentido, e a análise discursiva distingue-se da análise sêmio-narrativa, aquela da camada profunda. Mas a camada superficial não é uma "adição" à camada sêmio-narrativa profunda, já que ela supostamente rearticula os constructos sêmio-narrativos para complexificá-los e, principalmente, dar-lhes uma "roupagem" figurativa, espaço-temporal, actorial, entre outras. A partir de então, o discurso assim concebido é um "conjunto significante" completo e complexo, que compreende todos os elementos necessários a sua interpretação. Ele se deixa então apreender como "um todo de significação anterior a toda manifestação". (FONTANILLE, 2008, p. 3)

Essa imanência atribuída ao texto isola-o do seu entorno, não o vincula ao contexto físico, tornando-o um microcosmo passível de uma análise igualmente imanente. Não há um entorno material em um texto gerado por uma máquina que se estenda adiante de zeros e uns, de matemática pura e *bits* e *bytes*, que dialogam a seu modo com o próprio texto que foi produzido. O mundo virtual é o mundo do texto, e a este mundo o texto deve ser confinado, analisado dentro destas regras e não pela relação entre esse mundo e o mundo "exterior", o nosso mundo, o mundo dos humanos. Santaella e Noth (2004) citam este fenômeno do discurso desprovido de referências do seu entorno, onde este discurso torna-se um simulacro; local isolado onde existem os efeitos de representação produzidos por nós, humanos, e nossos sentidos, derivados da estrutura textual, construídos por um referente interno. A verdade material externa ao discurso pouco interessa a semiótica discursiva de linha francesa, que se

foca nos processos nos quais o texto constrói sua própria veridicção, segundo Greimas e Courtés (1979) “propriedade intrínseca do dizer e do dito”:

> O mundo extralinguístico, o mundo do “senso comum”, é formado pelo homem e instituído por ele em significação, e tal mundo, longe de ser o referente, é, pelo contrário, ele próprio uma linguagem bi plana, uma semiótica do mundo natural. O problema do referente nada mais é então do que uma questão de cooperação entre duas semióticas, um problema de intertextualidade. Concebido desse modo como semiótica natural, o referente perde, assim, sua razão de existir enquanto conceito linguístico. (GREIMAS e COURTÉS, 1979, p. 415)

Ao consumir *Nothing, Forever* (ou qualquer texto ou roteiro gerado por uma IA generativa), o espectador (enunciatário, no caso) relaciona-se com este produto através de um contrato enunciativo fiduciário, em um pacto de confiança. Esse contrato depende do grau de conhecimento que o espectador tem sobre o processo de geração daquele conteúdo, já que a informação de que ele foi gerado por uma máquina pode estar explícita, ou não.

No caso de *Nothing, Forever*, nesse contrato enunciativo é dado ao receptor todas as ferramentas necessárias para entender que a *sitcom* é gerada por IA: os responsáveis avisam isso ao espectador no site do projeto e no site de exibição, a mídia não omite esta informação, os espectadores que participam no *chat* disponível na exibição também sabem e compartilham este fato. Neste caso, o espectador confia que a máquina possa gerar algo original e criativo e que siga as convenções de gênero, estrutura narrativa e desenvolvimento de personagens de roteiros audiovisuais tradicionais. Confia, também, que a máquina entregue um conteúdo que esteja em harmonia com as expectativas do receptor, que são derivadas das premissas do show: ser uma *sitcom*, com personagens definidos, que tenha um viés humorístico, entre outros fatores.

Uma quebra em quaisquer pontos deste contrato gera uma estranheza, um distrato virtual, um desconforto. Depois do banimento temporário do programa da plataforma de exibição, os autores do show mudaram os parâmetros base de geração da história para uma “segunda temporada”, que tornou, segundo relatos de espectadores, o fluxo do programa menos “interessante”. Somente esta falta de substância dos personagens e suas falas já poderia ser vista como um prenúncio de uma quebra deste contrato, já que a expectativa da audiência para um show gerado dinamicamente e em tempo real é uma constante novidade e variação, mas um fator técnico tornou essa quebra de confiança do espectador mais pronunciada: erros de

processamento da interface do programa com a geração de texto tornaram as ações dos personagens erráticas e confusas, e o surgimento de situações fora da expectativa de verossimilhança de quem assistia causaram estranheza.

Eventos como personagens que ficavam horas olhando para a porta de uma geladeira, ou o surgimento de um personagem vestido com uma roupa laranja que não era previsto no conjunto de personagens iniciais, sem nenhum tipo de apresentação prévia, que aparecia e surgia sem aviso nem explicação, causaram uma sensação de incômodo no espectador, pois, não eram previstos no contrato fiduciário entre o programa e seu espectador. Falhas de programação do motor gráfico, como personagens atravessando móveis e paredes, ou tendo espasmos nas pernas ao sentar-se, somados a erros sonoros na geração musical ajudaram a dar aos espectadores um incômodo tão pronunciado que alguns associaram o programa a um show de horror sem sentido:

> On October 27, Twitter user AnimeSerbia posted that the show's characters "don't even say anything anymore and just stand still in complete silence." The poster followed up with another tweet about "a strange orange man" that silently stalked the apartment, disappearing and reappearing at random. As the days went on, the show's AI gears continued to slowly fall apart until October 30, when 404 Media reporter Jason Koebler shared a video of two characters, Leo and Nick Sterling (called "brown man" by viewers for his brown suit), just…walking into each other. In front of a closed refrigerator. On a loop. For five days. Now, characters sit around a handful of the same sets, just kinda staring at each other, wordless. If they aren't sitting, then they're walking into each other, into walls, or in place. Sometimes their movements are erratic, with arms flailing and legs jerking. (WINSLOW, 2023)

Estas inconsistências, no entanto, foram sanadas, pois derivam de um erro de programação. Os erros gráficos são passíveis de correção no motor gráfico, e os erros de comunicação são consertados tornando a comunicação entre o programa gerador do programa e a plataforma do LLM mais estável.

Uma questão mais complexa se dá quando essa cisão contratual se dá na própria evolução da geração textual contínua. Como *Nothing, Forever* roda sem intervenção humana, seus diálogos e ações são derivados dos diálogos e ações anteriores apresentadas pela própria máquina, em uma construção narrativa contínua, e um universo de situações que torna difícil para o LLM manter o microcosmo coerente, devido às limitações técnicas da plataforma.

Modelos de GPT, como o *ChatGPT* 3, que atua como motor de *Nothing, Forever*, como já dito anteriormente, são pré-treinados, não levando em conta dados novos, sejam informados pelo usuário, sejam advindos de mudanças no corpus original. Somente com um novo processo de treinamento, que é longo, custoso e exige recursos de *hardware* consideravelmente altos, que o modelo pode incorporar novas informações. O grande problema, no caso da geração infinita de conteúdo, é a falta de inserção das informações geradas dentro dos dados disponíveis para a máquina, tornando a construção da narrativa desconexa conforme o tempo avança. Não há uma memória volátil disponível para um LLM – pelo menos não com a tecnologia atual. O número de informações que o modelo pode manter em memória, através da repetição de comandos anteriores nas entradas posteriores, é limitado. Na arquitetura de modelos de linguagem baseados em *transformer*, como o *ChatGPT*, a limitação da entrada pode ser entendida através do conceito de atenção global, que é fundamental para a capacidade do modelo de processar e gerar linguagem. A atenção permite que o modelo pese a importância relativa de diferentes palavras dentro do contexto fornecido, facilitando a geração de respostas coerentes e, principalmente, contextualmente relevantes. Contudo, a implementação dessa atenção global implica um custo computacional que cresce exponencialmente com o número de *tokens* no contexto, devido à necessidade de calcular a interação entre cada par de *tokens*: para cada palavra a mais na entrada, mais cálculos são necessários para gerar a resposta.

Para manter a eficiência computacional e garantir a viabilidade prática do modelo em aplicações em tempo real, é necessário impor um limite ao tamanho do contexto que o modelo pode considerar. No caso do GPT-3, utilizado pelo *Nothing, Forever*, esse limite é de 2048 tokens. A limitação de 2048 tokens afeta tanto a quantidade de texto anterior na conversa que o modelo pode acessar quanto a extensão da resposta que ele pode gerar. Quando o modelo atinge esse limite, as informações mais antigas no contexto são descartadas para dar lugar às novas entradas. Essa abordagem permite que o modelo mantenha a relevância do contexto em conversas prolongadas, ajustando-se dinamicamente para focar nas interações mais recentes.

O tamanho deste contexto possível varia de versão para versão de LLM, e é limitado por diversos fatores, principalmente a memória disponível. Em sua última versão (*ChatGPT* 4 Turbo), o LLM da OpenAI é capaz de receber até 128 mil tokens de contexto, antes de elaborar uma resposta. É uma quantidade considerável, por volta de 85 mil palavras, ou um oitavo do tamanho do texto que compõe a versão em português da Bíblia Sagrada Católica.

Embora pareça muito, essa capacidade contextual é limitada, conforme a narrativa do programa gerado avança no tempo. Personagens em um fluxo contínuo de criação de conteúdo

criam situações que influenciam a si mesmos, aos personagens que contracenam com ele (e até aos que estão fora de cena) ao ambiente que os cercam e, dependendo do ato, até fora do microcosmo em que interagem, como ocorreu com a retirada do ar do show devido a comentários do protagonista. É esperado, até como parte integrante do contrato fiduciário entre enunciador e enunciatário, que a continuidade da história se dê mediante a absorção destas situações, com os personagens sendo moldados por elas e, dependendo do ponto de vista, evoluindo. Jean-Jacques Rousseau é o autor da frase "O homem é um produto do meio", e, sem prejuízos a uma discussão sociológica mais aprofundada do sentido desta afirmação, é possível estabelecer que é necessário, de um ponto de vista narrativo, que os personagens sofram mudanças de acordo com o impacto de suas ações anteriores.

Em um roteiro audiovisual clássico, essas ações são decididas pelo autor, que ajuíza as implicações que as ações ficcionais dos personagens que criou e desenvolveu, e define quais implicações e impactos estas ações terão no personagem, e por subsequência, na narrativa. Em um exemplo simples, se uma personagem fica grávida no início de uma trama, e a trama tem um cenário temporal maior do que nove meses, algum desfecho precisa ser dado a esta gravidez: ou o bebê nasce, ou ela o perde, ou ela desaparece, ou outra situação que justifique esta passagem temporal é tomada. Não é admissível, no ponto de vista da verossimilhança, considerando um contrato composto com bases em simulações da realidade, que, por exemplo, tempos depois a personagem apareça sem estar grávida, e sem explicar o que ocorreu.

Estas evoluções narrativas são inerentes nos roteiros elaborados por humanos, pois esta é a nossa realidade, e assumimos certos conceitos como imutáveis. No caso de um LLM, os únicos pontos de inferência que o algoritmo tem são os textos iniciais de treinamento, e um possível contexto passado a ele na entrada da geração, que é limitada em tamanho. Seria como, se a cada diálogo gerado, todas as ações passadas do personagem tivessem que ser passadas a ele na entrada, o que, invariavelmente, cria um problema: depois de um certo tempo, como introduzir todo o histórico do personagem, dos cenários, dos eventos, dos personagens secundários, em cada entrada de texto, com um tamanho limitado de memória disponível?

Em *Nothing, Forever*, vários exemplos desta limitação são aparentes. Fãs registram todos os diálogos e acontecimentos em uma *wiki*[27]*,* contextualizando a história no tempo. Nela,

[27] Em informática, *wiki* (do havaiano "superveloz") é um sistema de gestão de conteúdo utilizada na internet, que contém hipertexto e hiperligações, no qual vários usuários editam colaborativamente ao mesmo tempo o seu conteúdo e a estrutura. O uso mais conhecido do termo é a *Wikipedia*, enciclopédia colaborativa, mas se aplica a qualquer temática. A página da *wiki* de *Nothing, Forever* está disponível em: https://nothingforever.fandom. com/wiki/Nothing,_Forever_Wiki

estão registrados alguns factoides gerados pela IA para o personagem principal, em pontos específicos da trama, como, por exemplo, sua habilidade de falar a língua espanhola. Mas esta habilidade não é registrada no LLM, não está presente no banco de dados, e não é passada no contexto de gerações futuras, e acaba sendo ignorada e perdida em interações posteriores. No conceito de comédia situacional, de registro de situações cotidianas, essas inconsistências narrativas são notáveis e ajudam na quebra da imersão, cabendo ainda lembrar que o programa rodou com essa configuração por pouco mais de um mês, com seu conceito recriado para sua segunda temporada, após o problema do banimento.

É de se esperar que este problema da memória narrativa seja intensificado conforme o tempo avance na história, já que, quanto mais fatos são criados, mais inconsistências narrativas causadas pelas limitações de memória dos LLM são observáveis. Com o tempo, o avanço da tecnologia pode minimizar ou até romper com esta limitação, mas, no contexto atual, essa é uma problemática limitadora da performance narrativa de um produto criativo feito por uma inteligência artificial.

A estruturação narrativa de *Nothing, Forever* leva em conta estas limitações. O programa é feito em formato episódico, com geração de diálogos feitas em períodos específicos, em blocos com começo, meio e fim. Estes blocos são alternados com a entrada de uma simulação de um guia de canais a cabo, que mostra a programação fictícia de outros canais, também fictícios. Cada episódio tem duração de 10 a 30 minutos, e, dentro deste espaço de tempo, as interações entre personagens são preservadas, e a conversação simulada entre os agentes é mantida dentro de um relativo contexto. No entanto, após o fechamento de um episódio, tudo que ocorreu dentro daquele espaço de tempo é esquecido, e um novo bloco é gerado com as especificações iniciais. A temática do programa, com diálogos sobre assuntos mundanos, juntamente com a impossibilidade prática em se acompanhar o show durante períodos muito longos de tempo, faz a narrativa parecer contínua, mas, em termos absolutos, cada bloco do show tem seu contexto fechado, que não é levado para os blocos posteriores.

É possível fazer uma análise sobre a exibição de *Nothing, Forever*, como um produto audiovisual independentemente de se levar em conta o fator da geração aleatória feita por inteligência artificial. Contudo, o diferencial que separa *Nothing, Forever* do restante das obras audiovisuais e justifica a análise feita nesta tese é justamente o sua geração procedural autônoma, o que torna pertinente uma tentativa de reconstrução do processo de geração, já que a ferramenta que constrói o programa está disponível: o *ChatGPT*.

Quando falamos de uma obra audiovisual convencional, feita por humanos, uma investigação aprofundada das motivações que precedem a obra é viável: é possível entrevistar os envolvidos e perguntar sobre questões que sejam pertinentes para a análise. Também é pertinente observar que são humanos analisando obras humanas; existe uma concepção mínima de entendimento de uma obra pelo simples fato de que o analisador compartilha esta humanidade com o autor da obra. No caso de uma obra criativa feita por uma máquina, uma análise mais aprofundada contempla a possibilidade de que se tenha acesso ao código que construiu aquele produto.

4.4 A problemática da análise diante da opacidade algorítmica

A análise de um produto gerado por um algoritmo de IA é mais eficaz quando o pesquisador possui acesso facilitado e irrestrito ao algoritmo e aos dados de treinamento da inteligência artificial que o derivou. Mesmo que o produto do algoritmo (nesta tese, a sitcom *Nothing, Forever*) possa ser consumido e estudado abertamente, já que é amplamente acessível na internet, os mecanismos geradores do produto possuem código fechado e não estão plenamente disponíveis para uma análise aprofundada do processo produtivo.

Saber quais mecanismos levam as máquinas geradoras a tomarem suas decisões é importante para o entendimento de todo o processo do estudo da roteirização por inteligência artificial. Em uma comparação imprecisa, mas que ajuda a esclarecer o problema, estudar um roteiro feito por um humano não necessariamente leva a um acesso completo a mente do roteirista, porém, muitas variáveis úteis para o processo estão disponíveis, tais como o entendimento do contexto, os estudos científicos da psique humana, dados sobre linguística, semiótica, e outras ferramentas que permitem uma análise científica efetiva do texto com bases sólidas. No caso de uma máquina, é importante que se tenha um conhecimento amplo de como todo o processo de aprendizado e de criação de um texto por um LLM funciona; ocorre que, por diversos fatores, isto não é possível.

Em uma época em que o *boom* da inteligência artificial impacta vários setores da ciência, economia e outras áreas, nota-se um impulso inicial de proteção algorítmica, por razões de segredo industrial. Empresas criadoras de aplicações de IA evitam expor seus algoritmos ao público, com medo de seus concorrentes usarem seus segredos para aperfeiçoarem os seus produtos, e até governos, como o dos Estados Unidos, limitam o comércio de *hardware* próprio

para criação de modelos de inteligência artificial para países que não se alinham com as suas expectativas, como a China e a Rússia.

Só este cenário, onde os algoritmos e aparelhos não tem seu funcionamento detalhado, causa um problema considerável nas análises de produções destas máquinas de IA, já que pouco se sabe como o processo interno de construção do conteúdo gerado se dá. Esse aspecto é um pouco atenuado devido a ampla documentação científica conceitual sobre estes algoritmos, com artigos científicos amplamente divulgados explicando os conceitos por trás da construção e treinamento de LLMs. Companhias como o Google (software) e a Nvidia (hardware) publicam, constantemente, *papers* de domínio público sobre conceitos e propostas algorítmicas novas, acessíveis a todos.

Mesmo com a transparência dos conceitos, muitos dos algoritmos que capitaneiam as iniciativas de inteligência artificial generativa carecem de transparência, e, no caso de *Nothing, Forever*, outras condições intrínsecas ao produto também dificultam a análise plena do objeto.

Embora não estejam disponíveis as entradas iniciais que os criadores de *Nothing, Forever* transmitem ao *ChatGPT* para a geração do conteúdo para a sitcom, o LLM da OpenAI está disponível para o uso público, e pode, de certa forma, ser "entrevistado": ele responde os questionamentos sobre si mesmo e sobre o seu funcionamento. Se não é possível replicar os mesmos parâmetros que originam a *sitcom* objeto deste estudo, é possível solicitar ao *ChatGPT* que gere um episódio de uma sitcom:

> ENTRADA: Crie um roteiro de um episódio de 10 minutos de uma sitcom.
>
> *CHATGPT*: Vamos criar um roteiro para um episódio de 10 minutos de uma sitcom intitulada "Os Desajustados do Apartamento 404". Este episódio, "O Jantar Desastroso", foca-se em três colegas de apartamento tentando impressionar um visitante importante com um jantar caseiro, mas tudo dá errado de maneiras hilárias.
>
> Os Desajustados do Apartamento 404
> Episódio: "O Jantar Desastroso"
>
> Cena 1: Sala do Apartamento 404
>
> Alex está jogando videogame, enquanto Jamie está no sofá, lendo um livro. Sam entra animado, segurando um envelope.
> Sam: Pessoal, tenho notícias incríveis! Meu mentor, o Dr. Carter, está vindo jantar aqui amanhã!

Alex: (pausando o jogo) Quer dizer, o Dr. Carter? O cara que você não para de falar?
Jamie: E nós vamos cozinhar? (olha preocupado)
Sam: Sim! Precisamos impressionar. Vamos mostrar que somos mais do que apenas desajustados que sabem pedir pizza.

Cena 2: Cozinha, no dia seguinte

O trio tenta cozinhar. Sam está lendo uma receita no tablet, Jamie está cortando vegetais, e Alex tenta acender o forno.
Alex: (frustrado) Por que esse forno não liga?
Jamie: (sem olhar) Você apertou o botão certo?
Alex: Claro que... (pausa, envergonhado) Agora ligou.

Cena 3: Sala, mais tarde

A campainha toca. Sam corre para atender. Dr. Carter entra, elegante e carismático.
Dr. Carter: Que prazer ver onde a genialidade de Sam floresce.
Sam: (orgulhoso) Preparamos algo especial para o jantar.

Cena 4: Volta para a cozinha

O jantar está um desastre. A salada está excessivamente salgada, o frango queimado, e o macarrão está grudado.
Jamie: (tentando salvar) E se a gente disfarçar com bastante molho?
Alex: (olhando o frango queimado) Ou talvez a gente possa chamar isso de "carvão gourmet"?

Cena 5: Sala, durante o jantar

Eles estão sentados, tentando manter a compostura enquanto Dr. Carter experimenta a comida.
Dr. Carter: (mastigando devagar) Interessante... nunca experimentei nada assim.
Sam: (nervoso) É uma receita... exótica.
Alex: (sussurrando para Jamie) Acho que exótico é uma palavra para isso.

Cena 6: Cozinha, depois do jantar

O trio está lavando a louça, claramente desapontado.
Sam: Bom, pelo menos tentamos, certo?
Dr. Carter: (entra na cozinha) Sabe, essa foi uma das noites mais memoráveis que já tive. Vocês têm um espírito incrível.
Jamie: (surpreso) Sério?
Dr. Carter: Absolutamente. E a próxima vez, eu cozinho.

Cena 7: Sala, Dr. Carter se despede

Dr. Carter: (saindo) Lembrem-se, o sucesso vem de tentativas e falhas. Hoje, vocês brilharam.
Eles se olham, meio confusos, meio orgulhosos.
Alex: Quem diria, nosso desastre culinário nos ensinou algo.
Jamie: (sorrindo) Isso, e nunca mais deixar o Alex perto do forno.

> Cena 8: Os três se acomodam no sofá, ligam a TV e relaxam
>
> Sam: (olhando para os outros) Desajustados, mas juntos. Isso que importa.
> Eles sorriem, concordando, enquanto a câmera se afasta e a cena fecha com uma visão do apartamento iluminado à noite.
>
> FIM

A resposta não contém todos os detalhes que viabilizariam a geração de uma *sitcom* com este roteiro, mas a ferramenta entrega um guião com elementos suficientes para que, caso alguém desejasse, pudesse usá-lo como base para filmar uma *sitcom* de modo clássico – algo próximo do que foi feito com *Sunspring*. Porém, o fato mais relevante nesta situação não é a fidelidade com o qual o *ChatGPT* gera um roteiro, mas sim qual mecanismo ele usou para escolher qual roteiro criar; a entrada inicial não transmitiu absolutamente nenhum detalhe sobre a temática, nem sobre nome de personagens, nem sobre nenhuma ambientação, instruindo o *software* a apenas gerar um roteiro.

Como o *ChatGPT* pode responder perguntas sobre seus processos, foi questionado a IA em qual o processo que ela usou para escolher estes temas:

> ENTRADA: Como você define qual o assunto que seria abordado, o nome dos personagens, e demais detalhes, sem uma definição sobre o que eu desejava? Como funciona a escolha aleatória destes aspectos?
>
> *CHATGPT*: Quando se trata de criar conteúdo, como um roteiro de sitcom, sem especificações detalhadas, eu me baseio em princípios gerais de criatividade e narrativa para escolher assuntos, nomes de personagens e detalhes da história. Aqui está um vislumbre de como essas escolhas são feitas:
>
> Assunto Abordado
>
> Eu escolho temas universais ou situações cotidianas com potencial para humor e conflito leve, típicos de uma sitcom. A ideia é encontrar algo com o qual a maioria das pessoas possa se identificar ou achar divertido. Cozinhar um jantar importante e falhar miseravelmente é um cenário comum, mas rico em oportunidades para humor, erros cômicos e dinâmica de grupo.
>
> Nomes dos Personagens
>
> Os nomes são escolhidos para refletir uma variedade de fundos, tornando a história acessível a um amplo público. Eu tento usar nomes que soem bem juntos e que sejam fáceis de lembrar e pronunciar, considerando um equilíbrio de

> gêneros e personalidades para criar uma dinâmica interessante entre os personagens.
>
> Detalhes da História
>
> Os detalhes são criados para construir um enredo coeso e envolvente dentro do tempo limitado de um episódio curto. Isso inclui definir os papéis dos personagens, suas motivações e como eles interagem entre si e com o ambiente. Os detalhes também são ajustados para garantir que haja uma mistura de momentos cômicos e sentimentais, mantendo o episódio leve e divertido.
>
> Escolha Aleatória e Criatividade
>
> A escolha dos aspectos específicos da história não é completamente aleatória; é mais sobre aplicar critérios de narrativa e entretenimento. A seleção é guiada por tentar imaginar o que poderia ser divertido, interessante ou relevante para o público de uma sitcom. Ao mesmo tempo, tento assegurar que o roteiro tenha variedade, inclusão e evite clichês desgastados ou sensibilidades ultrapassadas.
>
> Essencialmente, o processo envolve um equilíbrio entre seguir convenções de gênero (neste caso, de uma sitcom) e introduzir elementos criativos e inovadores para fazer a história se destacar. O objetivo é criar um conteúdo que seja ao mesmo tempo familiar e fresco, proporcionando entretenimento e talvez até mesmo uma reflexão leve.

Aqui, o LLM mostra um entendimento claro e definido em como gerar uma narrativa, mostrando que o modelo de linguagem cria relações extremamente complexas entre os documentos que compõem o corpus do seu treinamento. Não é claro se o modelo foi instruído a agir assim, com seleção criteriosa de trechos dos dados do treinamento, ou se ele entendeu que deveria seguir este modelo para criar narrativas eficazes somente encontrando padrões em todas as narrativas que recebeu. Não é possível aferir como o modelo adquiriu este comportamento justamente pela natureza fechada do código do *ChatGPT* – natureza esta compartilhada entre muitos LLM: o *ChatGPT* é um *software* que sofre de problemas de opacidade algorítmica.

O termo “opacidade algorítmica” se refere a falta de transparência ou a dificuldade de compreensão em relação ao funcionamento interno dos algoritmos. O algoritmo é um produto e, como produto, sempre esteve sujeito as condições do sistema capitalista moderno. Muitos consideram não fazer sentido econômico revelar um código construído internamente para a sociedade, e tratam seus algoritmos como segredo industrial.

Cabe aqui um aparte. Existe uma discussão frequente no mundo da computação sobre vantagens e desvantagens de produtos de código aberto versus produtos de código fechado. Há benefícios econômicos em disponibilizar um código sem cobrança para todos, criando uma comunidade de desenvolvimento em volta dele, facilitando as correções de erros e tornando sua evolução mais rápida. Inclusive, há iniciativas dentro do desenvolvimento de algoritmos de inteligência artificial para abertura de código, como as intenções do Facebook de criar e disponibilizar o algoritmo de uma inteligência artificial generativa generalista, ainda em um campo hipotético. No entanto, na prática, muito pouco do código bruto de inteligências artificiais relevantes atualmente é aberto e, embora os dados de treinamento delas estejam disponíveis a todos para acesso individual, pouco se sabe sobre o processo de seleção e agregação destes dados para uso nestes algoritmos. Como o cenário atual é muito mais fechado do que aberto, considera-se o estado atual das coisas para a análise.

A opacidade algorítmica, no entanto, não é apenas causada pela falta de disponibilidade do código devido a razões comerciais. Na inteligência artificial, outros fatores também dificultam a transparência dos programas. A natureza intrinsecamente complexa de certos algoritmos, como as redes de *deep learning*, resulta em dificuldades epistemológicas. Estes sistemas, ao serem treinados com grandes conjuntos de dados, desenvolvem representações internas que, nem sempre, são facilmente interpretáveis por humanos, tornando a necessidade de uma explicação detalhada do raciocínio virtual subjacente de forma compreensível para que humanos possam ler e interpretar. Isso é particularmente pertinente em domínios onde decisões errôneas podem ter consequências graves, como na medicina e na justiça criminal.

> A main drawback of artificial neural networks is that they have no explicit declarative knowledge representation. They have considerable difficulty in generating the required explanation structures. The absence of an explanation capability in neural networks limits the achievement of the full potential of such systems, whereas the detailed characterization of classification strategies would contribute to their acceptance. (BOLOGNA e HAYASHI, 2017, p. 165)

Como não há um controle direto sobre os dados de treinamento, a opacidade algorítmica pode ajudar a mascarar vieses presentes nos próprios dados de treinamento, ou até na própria concepção do algoritmo. Estes vieses podem perpetuar ou amplificar conceitos não desejáveis para a sociedade, e, sem a compreensão detalhada dos mecanismos internos destes algoritmos e dos dados que os alimentam, a mitigação deste comportamento é dificultada. Aqui, volta-se

ao caso do ato de transfobia cometido pelo programa, que motivou seu banimento. Os desenvolvedores, sem um controle efetivo dos resultados gerados pela inteligência artificial, ficam a mercê dos seus filtros e dados de treinamento, que não temos acesso. Há até uma dificuldade em se definir a culpabilidade destes atos; se alguém for ofendido por um texto de inteligência artificial, deve procurar a justiça e processar quem? O responsável pela IA, o responsável pelo produto, ou o criador do dado que treinou a IA no comportamento ilícito?

Um serviço de IA generativa atua como um funcionário de uma empresa, que se espera que siga regras, cumpra suas funções e esteja disponível quando necessário. Quando se usa um produto de inteligência artificial para fins profissionais – como é o caso do uso em *Nothing Forever*, há implicações contratuais que garantam disponibilidade, performance e taxas de erros e acertos. Mesmo assim, tal qual um humano contratado para uma tarefa, você não tem total controle sobre o que ele fará, como ele fará e até se o fará; este é o paradigma da relação entre humanos e máquinas em uma evolução tecnológica de terceira ordem. Quando os criadores de *Nothing, Forever* decidiram usar o *ChatGPT* para fazer a geração dos roteiros de sua sitcom, transferiram ao LLM toda a responsabilidade de não cometer deslizes que causassem prejuízo ao produto – situação na qual a IA falhou.

O exemplo de *Nothing, Forever* e sua fala transfóbica alerta para uma situação deveras importante para todo produtor de conteúdo que decida se utilizar de serviços de IA generativa não supervisionada: avaliação ética de todas as etapas do processo. Como dito, é uma situação complexa, já que a opacidade algorítmica complica quaisquer avaliações prévias que um produtor possa fazer de uma IA generativa (MITTELSTADT, ALLO, *et al.*, 2016) além de exemplos do que a IA entrega e de garantias que, eventualmente, a empresa fornecedora do serviço indique.

> Algorithms are, however, ethically challenging not only because of the scale of analysis and complexity of decision-making. The uncertainty and opacity of the work being done by algorithms and its impact is also increasingly problematic. Algorithms have traditionally required decision-making rules and weights to be individually defined and programmed ‘by hand’. (MITTELSTADT, ALLO, *et al.*, 2016, p. 3)

Para nortear a avaliação da ética de algoritmos, Mittelstadt et al. (2016) definiram um mapa conceitual, baseado em seis preocupações sobre as ações, reações e

produções algorítmicas que permitem uma base para fomentar questões sobre o seu funcionamento ético: evidências de incompletude, evidências não pesquisáveis, evidências enganosas, resultados injustos, efeitos transformativos, e rastreabilidade.

Quando Mittelstadt et al. (2016, p. 4) citam a incompletude algorítmica, referem se a problemática do pensamento estatístico da produção de um resultado provável. Quando um LLM define, através de todos os seus cálculos, qual é a palavra seguinte mais provável de ser inserida naquele contexto, ela é essencialmente isto, a palavra mais provável. Não é, nem nunca será, uma certeza; sempre será um arroubo estatístico, que tem probabilidades de estar certo, mas também tem probabilidades de não estar. A incerteza algorítmica leva a possíveis ações erráticas do algoritmo, não previstas e capazes de gerar disfunções nas cadeias de produção, como ocorreram no caso de *Nothing, Forever*, onde personagens iniciaram comportamentos erráticos, ou congelavam em loops infinitos.

Quando falam de evidências não pesquisáveis (MITTELSTADT, ALLO, *et al.*, 2016, p. 4), indicam a problemática da transparência. O usuário fornece a entrada (que é conhecida) e recebe uma saída de texto (que é conhecida) mas não tem nenhum detalhe sobre qual processo ocorreu dentro do algoritmo para que essa associação seja feita. Não é uma discussão que se inicia com os algoritmos de IA, mas sim um ponto que é polêmico desde a criação dos grandes mecanismos de busca e de seus algoritmos de ranking no início dos anos 2000, ou com os algoritmos de classificação das redes sociais – ambos, hoje, utilizam-se de IA, mas, inicialmente, não contavam com este recurso.

Há um movimento de mudanças de legislação sobre inteligência artificial no mundo, o que pode mudar o cenário de como os algoritmos de IA funcionam e aumentar a transparência deles por força de lei. Contudo, mesmo hoje, com o vácuo legislativo que existe sobre o assunto (GOLD, 2023), há pontos onde é possível aumentar a transparência dos algoritmos de IA, como o estímulo a auditorias externas, uma aproximação maior da academia com a indústria, e educação e conscientização.

O desenvolvimento de algoritmos de inteligência artificial generativa com viés de transparência ganha o nome de *Glass Box* (Caixa de Vidro), um contraponto a palavra *Black Box* (Caixa Preta), sinalizando que os procedimentos internos do algoritmo são visíveis e passíveis de escrutínio. Algoritmos que usem o conceito de *Glass Box* tendem a ser mais claros sobre seu funcionamento interno, mais explicáveis para os humanos, podem implementar modelos de regressão linear (onde os dados podem ser mostrados passo a passo, e é possível se

inferir a entrada através do resultado), disponibilizam relatórios e documentação completa e propõem recursos de feedback.

Ainda segundo Mittelstadt et al. (2016) evidências enganosas tratam dos dados que abastecem os algoritmos de IA. Como já dito, nenhuma inteligência artificial generativa cria conteúdo do zero – sempre existe uma base de dados de treinamento ou de consulta, e sempre há uma entrada do usuário que provoque uma reação. Com quaisquer ruído, viés ou falha nestas duas entradas, corre-se o risco de o resultado do algoritmo ser prejudicado. O comentário transfóbico que derrubou a sitcom *Nothing, Forever* veio de um algoritmo de LLM treinado com uma base de dados que continha evidências desta transfobia, e apenas replicou o dado matemático que obteve.

Qualquer decisão tomada pelos humanos responsáveis pela criação e do treino de um LLM influencia no seu resultado. Seja fornecer os dados sem filtro algum, ou implementar algum tipo de filtro específico, seja limitar os termos posteriormente, seja filtrar as saídas através de iniciativas manuais ou de outra IA, quaisquer ações levarão a uma inteligência artificial que refletirá as ações iniciais do programador:

> Algorithms inevitably make biased decisions. An algorithm's design and functionality reflects the values of its designer and intended uses, if only to the extent that a particular design is preferred as the best or most efficient option. Development is not a neutral, linear path; there is no objectively correct choice at any given stage of development, but many possible choices (Johnson, 2006). As a result, ''the values of the author [of an algorithm], wittingly or not, are frozen into the code, effectively institutionalizing those values. It is difficult to detect latent bias in algorithms and the models they produce when encountered in isolation of the algorithm's development history. (MITTELSTADT, ALLO, *et al.*, 2016, p. 7)

O acesso irrestrito a algoritmos e dados de treinamento é fundamental para uma compreensão integral e ética dos produtos gerados pela inteligência artificial. A opacidade algorítmica, seja por motivos de proteção de segredo industrial ou pela complexidade inerente aos modelos de aprendizado profundo, representa um desafio significativo para a transparência e a responsabilidade algorítmica, limitando análises aprofundadas também a capacidade de identificar e mitigar vieses presentes nos sistemas de IA. A análise de *Nothing, Forever* destaca a complexidade das interações entre a criação de conteúdo por IA e suas ramificações sociais, sublinhando a urgência de encontrar um equilíbrio entre inovação tecnológica e responsabilidade ética.

A transparência algorítmica e o acesso aos conjuntos de dados são cruciais não apenas para o avanço científico, mas também para garantir que os desenvolvimentos tecnológicos ocorram de maneira justa e responsável. A necessidade de soluções que harmonizem a proteção de informações proprietárias com a demanda por clareza sobre o funcionamento dos sistemas de IA é premente.

5 – CONSIDERAÇÕES FINAIS

A máquina, ao assumir funções humanas na produção de conteúdo audiovisual como em *Nothing, Forever*, exige uma reconsideração crítica das abordagens tradicionais de criação, análise e responsabilidade. Ao transferir o papel criativo para algoritmos de IA generativa, como o *ChatGPT*, e empregar outras tecnologias intermediárias para a composição e distribuição desse conteúdo, os criadores e operadores dessas tecnologias inauguram um território que desafia os limites entre criador e ferramenta, entre autoria humana e maquínica, e entre a intencionalidade artística e o acaso algorítmico. Esta transferência evoca questões fundamentais sobre a natureza da criatividade, a essência do engajamento narrativo e a responsabilidade ética que acompanha a produção cultural automatizada. O episódio de *Nothing, Forever* e sua suspensão subsequente ilustram vividamente a complexidade de confiar em inteligências artificiais para gerar conteúdo que ressoa de maneira autêntica, responsável e eticamente consciente com sua audiência humana. Isso aponta para uma "fauna" emergente de desafios e considerações críticas que acompanham a adoção de IA na produção artística e cultural.

A opacidade algorítmica não apenas obscurece os processos subjacentes que guiam a produção de conteúdo de IA, mas também complica a atribuição de responsabilidade por seus resultados. Desvendar esses processos, compreender os pressupostos embutidos nos dados de treinamento e as decisões de design que moldam o comportamento do algoritmo, é fundamental para avançar em direção a um uso mais transparente, justo e ético da IA generativa nas artes e na cultura. Isso exige um esforço colaborativo entre desenvolvedores de IA, artistas, críticos culturais, legisladores e a audiência. Enquanto tecnologia intermediária de terceira ordem, a IA generativa redefine a interação humano-máquina e postula um novo paradigma para a produção artística, onde algoritmos não servem apenas como ferramentas nas mãos de artistas humanos, mas assumem um papel ativo na co-criação do conteúdo. Isso reitera a necessidade de uma estrutura ética robusta e diretrizes claras que regulem essa colaboração entre humanidade e máquina, garantindo que a inovação algorítmica proceda de maneira responsável e em harmonia com os valores sociais e culturais. Ao confrontar as repercussões da geração de conteúdo por IA e seu impacto na cultura audiovisual, deve-se buscar uma balança eficaz entre explorar o potencial criativo da tecnologia e proteger a integridade, a diversidade e a inclusão nas expressões artísticas e narrativas. Isso implicará tanto no aprimoramento da tecnologia para refinar seu alinhamento com os padrões éticos quanto no desenvolvimento de críticas e análises que respondam adequadamente à emergente paisagem da criação de conteúdo algorítmico.

Nothing, Forever introduz um novo paradigma na criação de obras audiovisuais, apresentando dois conceitos definidores de seu pioneirismo: a sua narrativa autônoma gerada através de técnicas de inteligência artificial, e seu fluxo contínuo e potencialmente infinito. Baseado nestes conceitos, propõe-se um termo novo para a classificação de obras com estas características: “Narrativa Generativa Contínua”.

O conceito de “Narrativa Generativa Contínua” visa representar uma mudança de paradigma no campo das produções audiovisuais, marcando a fusão entre as capacidades computacionais avançadas e o *storytelling*. A denominação foi escolhida para encapsular os três pilares fundamentais que distinguem esta modalidade de produção das abordagens tradicionais de criação de conteúdo audiovisual.

O primeiro componente, "Narrativa", refere-se ao coração da produção, o conteúdo narrativo entregue por meio de recursos textuais, visuais e auditivos. Diferentemente das obras clássicas, que seguem uma estrutura narrativa linear ou fechada, a "Narrativa de Geração Processual Contínua" é caracterizada por sua capacidade de evoluir e se adaptar. Este aspecto destaca a fluidez e a dinâmica na apresentação de histórias, possibilitando uma experiência mais rica e imersiva para o espectador, que pode se desdobrar em múltiplas direções baseadas em interações ou em alterações do contexto de visualização.

O termo "Generativa" aqui é aplicado em seu sentido usado na computação, e refere-se a um paradigma de programação baseado na geração sequencial de instruções ou procedimentos para realizar tarefas ou resolver problemas, enfatizando a estrutura de controle e a manipulação de variáveis. Aqui, o termo destaca a metodologia empregada na geração e manipulação do conteúdo narrativo. Em contraste com os métodos tradicionais de criação, que dependem significativamente da entrada humana direta, a abordagem generativa utiliza algoritmos e lógicas computacionais para criar ou modificar elementos narrativos e visuais. Este processo automatizado permite a elaboração de narrativas complexas e detalhadas, oferecendo uma diversidade de desdobramentos e variações narrativas que seriam inviáveis através de métodos exclusivamente manuais. A natureza de geração processual é a chave para a capacidade infinita de geração de conteúdo novo, assegurando uma fonte inesgotável de narrativas.

Finalmente, o adjetivo "Contínua" ressalta a característica mais distinta dessa abordagem: a produção contínua e sem limites de narrativas audiovisuais. Esta propriedade é fundamentalmente diferente das obras audiovisuais clássicas, que possuem uma duração e um escopo narrativo definidos. A infinitude é viabilizada pelo uso de algoritmos de geração

processual que, ao operar dentro de um paradigma de possibilidades narrativas pré-estabelecidas, são capazes de gerar continuamente conteúdo novo e relevante.

A introdução de narrativas audiovisuais geradas autonomamente, processualmente e de maneira contínua representa um marco disruptivo no panorama da indústria audiovisual. Baseadas na geração de conteúdo dinâmico e ilimitado, por meio de algoritmos, oferece uma nova perspectiva sobre os processos de criação, distribuição e consumo de mídia. Ao desafiar os paradigmas estabelecidos, essa tecnologia propõe uma reconfiguração da lógica de produção e da relação entre criadores e audiência.

Do ponto de vista da produção, a geração processual de conteúdo audiovisual implica uma revisão fundamental das metodologias tradicionais. Em vez de seguir roteiros fixos, criadores podem empregar algoritmos para desenvolver narrativas que se adaptam e evoluem em resposta às interações do público ou a conjuntos de dados específicos. Essa abordagem não apenas aumenta a eficiência ao reduzir os tempos de produção e os custos associados, mas também fomenta a criação de obras mais personalizadas e adaptativas, capazes de atender às expectativas e aos interesses de uma audiência diversificada.

Para os espectadores, a oferta de narrativas audiovisuais contínuas e geradas processualmente abre um leque de possibilidades anteriormente inexplorado. A capacidade de experienciar conteúdos que são únicos a cada visualização transforma radicalmente a natureza do consumo de mídia. Isso não apenas aumenta o engajamento por parte da audiência, através da oferta de experiências imersivas e personalizadas, mas também altera as expectativas dos consumidores em relação à originalidade e à inovação no conteúdo audiovisual. Ademais, a interatividade inerente a essas narrativas promove uma participação mais ativa dos espectadores no processo narrativo, diluindo as fronteiras tradicionais entre criadores e consumidores de conteúdo.

No entanto, essa evolução também levanta questões pertinentes acerca da autoria, da propriedade intelectual e dos direitos autorais. A geração de conteúdo por meio de algoritmos coloca em dúvida as noções convencionais de criação e originalidade, exigindo uma reavaliação das bases legais e éticas que regem a produção cultural. Além disso, a personalização extrema do conteúdo suscita preocupações relativas à privacidade dos dados dos usuários e à criação de câmaras de eco, onde a exposição a ideias e perspectivas diversas pode ser limitada.

A introdução deste modo de criação de narrativas pode ser responsável por uma transformação profunda na indústria audiovisual, impactando todos os aspectos relacionados à

produção e ao consumo de conteúdo. Embora essa tecnologia ofereça oportunidades sem precedentes para a inovação e a personalização, ela também apresenta desafios significativos que exigem uma reflexão cuidadosa sobre as implicações éticas, culturais e legais. É importante abordar essas questões de forma a maximizar o potencial benéfico dessa inovação, ao mesmo tempo em que se minimizam os possíveis efeitos adversos. *Nothing, Forever* é a primeira iniciativa de, provavelmente, muitas outras que virão, e pavimentarão o caminho para uma mudança radical na relação de produção e consumo de mídia.

REFERÊNCIAS

ALLEN, J. F. AI Growing Up: The Changes and Opportunities. **AI Magazine**, v. 19, n. 4, p. 13-24, 1998.

ARISTÓTELES. **Metafisica**. São Paulo: Loyola, 2002.

BARROS, D. L. P. **Teoria Semiótica do Texto**. 4. ed. [S.l.]: Parma, 2005.

BARTNECK, C. et al. **An Introduction to Ethics in Robotics and AI**. [S.l.]: Springer, 2021.

BENDER, E. et al. **On The Dangers of Stochastic Parrots:** Can Language Models Be Too Big? FAccT '21: 2021 ACM Conference on Fairness, Accountability, and Transparency. New York: Association for Computing Machinery. 2021. p. 610-623.

BERTRAND, D. **Caminhos da semiótica literária**. Tradução de Grupo CASA. Bauru: EDUSC, 2003.

BOLOGNA, G.; HAYASHI, Y. Characterization of Symbolic Rules Embedded in Deep DIMLP Networks: A Challenge to Transparency of Deep Learning. **Journal of Artificial Intelligence and Soft Computing Research**, Lodz, Outubro 2017. 265 - 286.

BOREL, E. La mécanique statique et l'irreversibilité. **Journal de Physique Théorique et Appliquée**, Paris, 1913. 189-196.

BRINGSJORD, S.; GOVINDARAJULU, N. S. Artificial Intelligence. **Stanford Encyclopedia of Philosophy**, 2020. Disponivel em: <https://plato.stanford.edu/archives/sum2020/entries/artificial-intelligence>. Acesso em: 14 jan. 2022.

BROMLEY, A. G. Charles Babbage’s Analytical Engine, 1838, v. 4, n. 3, p. 196-217, Julho 1982. ISSN doi: 10.1109/MAHC.1982.10028.

BYBEE, J. **Language, usage and cognition**. Cambridge University Press. Cambridge. 2010.

CAJAL, S. R. **Textura del Sistema Nervioso del Hombre Y de Los Vertebrados**. Madrid: Nicolás Moya, 1904.

CARBONELL, J.; MICHALSKI, R.; MITCHELL, T. **Machine Learning:** An Artificial Intelligence Approach. 1a. ed. Burlington: Morgan Kaufmann Publishers, v. 1, 1985.

CHOMSKY, N. **Aspects of the theory of syntax**. Cambridge: M.I.T. Press, 1965.

CHOMSKY, N. **Reflexões sobre a linguagem**. Tradução de C. VOGT. São Paulo: Cultrix, 1980.

CHOMSKY, N. Novos horizontes no estudo da linguagem. **DELTA: Documentação e estudos em linguística teórica e aplicada**, Rio de Janeiro, v. 13, n. 3, p. 49-72, 1997.

CHOMSKY, N. AI isn't coming for us all, you idiots. **futurism.com**, 2023. Disponivel em: <https://futurism.com/the-byte/noam-chomsky-ai>. Acesso em: 20 dez. 2023.

CHOWDHURY, G. Natural Language Processing. **Annual Review of Information Science and Technology**, Glasgow, n. 37, p. 51-89, 2003.

COHEN, P. Harold Cohen and AARON. **AI Magazine**, v. 37, n. 4, p. 63-66, 2016.

COUCHOT, E. E. Da representação à simulação: evolução das técnicas e das artes da figuração. In: PARENTE, A. **Imagem máquina:** a era das tecnologias do virtual. Rio de Janeiro: Editora 34, 2011. p. 37-48.

CROFT, B. W.; LAFFERTY, J. **Language Modeling for Information Retrieval**. [S.l.]: Springer Science & Business Media, 2013.

DANCE, F. E. X. The "Concept" of Communication. **Journal of Communication**, 1970. 201-210.

DANCHIN, A.; FENTON, A. A. From Analog to Digital Computing: Is Homo Sapiens Brain on Its Way to Become a Turing Machine? **Frontiers In Ecology and Evolution**, v. 10, Maio 2022.

DIETRICH, R.; GRAUMANN, C. F. Language processing in social context, an interdisciplinary account. **North-Holland Linguistic Series: Linguistic Variations**, v. 54, p. 1-15, 1989.

DU, W.; HAN, Q. **Research on Application of Artificial Intelligence in Movie Industry**. 2021 International Conference on Image, Video Processing, and Artificial Intelligence. Shanghai: [s.n.]. 2021.

FERREIRA, F. R. M. **A formação do programa disciplinar da Neurociência na primeira metade do século XX**. Universidade de São Paulo. São Paulo. 2022.

FIORIN, J. L. Três questões sobre a relação entre expressão e conteúdo. **Revista de literatura Itinerários**, 2003. 77 - 89.

FIRTH, J. **Studies In Linguistic Analysis**. London: Basil Blackwell Oxford, 1957.

FLOCH, J. M. Você é agrimensor ou sonâmbulo? Elaboração de uma tipologia comportamental dos passageiros do metrô. **Estudos semióticos**, São Paulo, v. 19, n. 2, Agosto 2023.

FLORIDI, L. **The 4th Revolution:** How the Infosphere is Reshaping Human Reality. Oxford: Oxford University Press, 2014.

FONTANILLE, J. Semiótica do Discurso: Balanços e Perspectivas. **Cadernos de Semiótica Aplicada**, Araraquara, Julho 2008. Traduzido por Jean Cristtus Portela e Matheus Nogueira Schwartzmann.

FRIEDERICI, A. D. The brain basis of language processing: from structure to function. **Phisiological Reviews**, Outubro 2011. 1123-1513.

GENTZKOW, M. **Media and Artificial Intelligence**. Stanford. Stanford. 2018.

GOLD, A. Ai rockets ahear in vaccum of U.S. regulation. **axios.com**, 2023. Disponivel em: <https://www.axios.com/2023/01/30/ai-chatgpt-regulation-laws>. Acesso em: 17 jan. 2024.

GOLGI, C. Sulla struttura della sostanza grigia del cervelo. **Gazzetta Medica Italiana**, Lombardia, 1873. 244-246.

GOODFELLOW, I. J. et al. Generative Adversarial Nets. **arXiv**, 2014.

GREIMAS, A. J.; COURTÉS, J. **Dicionário de Semiótica**. São Paulo: Contexto, 1979.

HJELMSLEV, L. **Prolegômenos a uma teoria da linguagem**. São Paulo: Perspectiva, 1975.

HOBSBAWM, E. J. **A era das revoluções:** 1789 - 1848. São Paulo: Paz e Terra, 2012.

HOLYOAK, K. J.; SPELLMAN, B. A. **Thinking**. 1ª. ed. Los Angeles: Annual Reviews, Inc., v. I, 1993.

HUNT, D.; RAMÓN, A. C.; TRAN. **Hollywood Diversity - Report 2019**. 1. ed. Los Angeles: UCLA, v. 1, 2019.

KAPLAN, A.; HAENLEIN, M. Siri, Siri, in my hand: Who's the fairest in the land? On the interpretations, illustrations and implications of artificial intelligence. **Business Horizons**, 2019. 15-25.

KUBAT, M. An Introduction to Machine Learning. 3. ed. [S.l.]: Springer Nature Switzerland, 2021.

KUHN, T. S. **A estrutura das revoluções científicas**. 5. ed. [S.l.]: Perspectiva, 1998.

LANDOWSKI, E. **Interações arriscadas**. Tradução de L. H. O. DA SILVA. São Paulo: Estação das Letras e Cores, 2014.

LEANDRO, A. Lições de Roteiro, por JLG. **Educação e Sociedade**, v. 24, n. 83, p. 659-678, 2003.

LEVY, P. IEML: Rumo a uma Mudança de Paradigma na Inteligência Artificial. **INTLEKT Metadata**, 2022. Disponivel em: <https://intlekt.io/2022/01/18/ieml-towards-a-paradigm-shift-in-artificial-intelligence/>. Acesso em: 30 Janeiro 2024.

LIM, M. W. et al. Generative AI and the future of education: Ragnarok or Reformation? **The International Journal of Management Education**, Winchester, 2023.

LIN, C. **ROUGE:** A Package for Automatic Evaluation of Summaries. Text Summarization Branches Out - Proceedings of the ACL-04 Workshop. Barcelona: [s.n.]. Julho 2004.

MACHADO, A. **A Arte do Vídeo**. [S.l.]: Brasiliense, 1988.

MAKRIDAKIS, S. The forthcoming Artificial Intelligence (A) revolution: Its impact on society and firms. **Futures**, 2017.

MARCOS, A. F. Digital Art: When Artistic and Cultural Muse Merges. **IEEE Computer Graphics and Applications**, n. 5, p. 70 - 74, Setembro 2007.

MCCARTHY, J. et al. **A proposal for the Darthmouth summer research project on artificial intelligence**. Darthmouth College. New Hampshire. 1955.

MITTELSTADT, B. D. et al. The ethics of algorithms: Mapping the debate. **Big Data & Society**, 2016.

MORDVINTSEV, A.; OLAH, C.; TYKA, M. DeepDream - a code example for visualizing Neural Networks. **Google Research**, 2015. Disponivel em: <https://blog.research.google/2015/07/deepdream-code-example-for-visualizing.html>. Acesso em: 15 Fevereiro 2024.

NOTHING, Forever. Direção: S. HARTLE e B. HABERSBERGER. [S.l.]: Mismatch Media. 2023.

PIAGET, J. Definition De L'Intelligence. **https:** //www.fondationjeanpiaget.ch/, Paris, 1951. Disponivel em: <https://www.fondationjeanpiaget.ch/fjp/site/textes/VE/JP51_Texte_Causerie1.pdf>. Acesso em: 21 Junho 2023.

PICCOLINO, M. Luigi Galvani`s path to animal electricity. **Comptes Rendus Biologies**, Paris, n. 329, p. 303-318, Março 2006.

PINHEIRO, M. Aspectos históricos da neuropsicologia: subsídios para a formação de educadores. **Educar em Revista**, Junho 2005.

PLATÃO. **A República**. 6. ed. [S.l.]: Atena, 1956.

POOLE, D.; MACKWORTH, K. **Artificial Intelligence:** Foundations of computational agents. Cambridge: Cambridge University Press, 2010.

RAPAPORT, W. J. Syntactic Semantics: Foundations of computational natural-language understanding. In: DIETRICH, E. **Thinking Computers and Virtual Persons:** Essays on the Intentionality of Machines. [S.l.]: Academic Press, 1994. p. 81-131.

RHEE, J. Misidentification's Promise: the Turing Test in Weizenbaum, Powers, and Short. **Postmodern Culture**, v. 20, n. 3, Maio 2010.

RIFKIN, J. **A Terceira Revolução Industrial - Como o poder lateral está transformando a energia, a economia e o mundo**. São Paulo: M.Books, 2012.

RUDIN, P. Thoughts on human learning vs machine learning. **Singularity 2030**, 2017. Disponivel em: <https://singularity2030.ch/thoughts-on-human-learning-vs-machine-learning/>. Acesso em: 24 maio 2023.

RUSSEL, S.; NORVIG, P. **Artificial Intelligence:** A Modern Approach. 4. ed. [S.l.]: Pearson, 1995.

SAMUEL, A. L. Some studies in machine learning using the Game of Checkers. **IBM Journal**, Julho 1959.

SANTAELLA, L.; NOTH, W. Comunicação como enunciação. In: _____ **Comunicação e Semiótica**. São Paulo: Hacker Editores, 2004. p. 113 - 126.

SAUSSURE, F. **Curso de linguística geral**. Tradução de A. CHELINI; J. P. PAES e I. BLIKSTEIN. [S.l.]: Cultrix, 2006.

SCHRAMM, W. The Nature of Communication between Humans. In: SCHRAMM, W.; ROBERTS, D. F. **The Process and Effects of Mass Communication**. Illinois: University of Illinois Press, 1971. Cap. 1.

SCHWAB, K. **A quarta revolução industrial**. São Paulo: Edipro, 2016.

SEARLE, J. R. Minds, brains, and programs. **Behavioral and Brain Sciences**, Cambridge, Setembro 1980. 417 - 424.

SKINNER, B. F. Selection by Consequences. **Science**, v. 203, n. 4507, p. 501-504, Julho 1981.

STERRETT, S. G. Turing's two tests for intelligence. **Minds and Machines**, v. 10, n. 4, p. 541-559, Novembro 2000.

SUTHERLAND, I. E. **Sketchpad: A man-machine graphical communication system**. Lincoln Laboratory. Lexington, p. 99. 1963.

TOMASELLO, M. **Construting a Language:** A usage-based theory of Language Acquisition. Cambridge: Harvard University Press, 2003.

TURING, A. M. On computable numbers, with an application to the entscheidungsproblem. **Proceedings of the London Mathematical Society**, Londres, v. 42, n. 2, p. 230-265, 1937.

TURING, A. M. Intelligent Machinery, 1948.

TURING, A. M. Computing Machinery and Intelligence. **MIND**, v. LIX, n. 236, p. 433-460, Outubro 1950. ISSN 0026-4423.

VYGOTSKY, L. S. **Formação social da mente**. 4. ed. São Paulo: Martins Fontes, 1991.

WANG, P. **What Do You Mean by "AI"?** Proceedings of the 2008 conference on Artificial General Intelligence 2008: Proceedings of the First AGI Conference. Memphis: [s.n.]. 2008. p. 362-373.

WEIZENBAUM, J. ELIZA - a computer program for the study of natural language communication between man and machine. **Communications of the ACM**, jan. 1966. 36-45.

WEIZENBAUM, J. **Computer Power and Human Reason:** From Judgement to Calculation. São Francisco: W.H. Freeman and Company, 1976.

WINSLOW, L. Twitch's AI Seinfeld Show Has Devolved Into Existential Horror. **Kotaku**, 2023. Disponivel em: <https://kotaku.com/ai-seinfeld-twitch-nothing-forever-247-1850977472>. Acesso em: 4 Novembro 2023.

Y QUEVEDO, L. T. Torres and his remarkable automatic devices. **Scientific American**, New York, Novembro 1915. 296.