



Interações discursivas em Dall-e 2

*Discursive interactions
in Dall-e 2*



Ana Silvia Lopes Davi Médola¹
Henrique da Silva Pereira²

¹ Livre docente em Linguagem Televisual. Docente permanente no Programa de Pós-Graduação em Comunicação da Universidade Estadual Paulista. ana.silvia@unesp.br

² Doutor em Comunicação pelo Programa de Pós-Graduação em Comunicação da Universidade Estadual Paulista. Docente do Centro Universitário Nossa Senhora do Patrocínio. hen.silper@gmail.com

Resumo: Os processos de colocação em discurso com tecnologias de inteligência artificial generativa no site Dall-e 2 são analisados a partir das proposições da semiótica acerca das interações discursivas. As projeções no enunciado de um ator humano e de um ator maquínico produzem efeitos de sentido de interações bilaterais reflexivas e unidirecionais entre esses atores que requerem novas competências do destinatário para um fazer interpretativo. Destaca-se a necessidade da identificação dessas marcas discursivas por parte do enunciatário para o estabelecimento dos contratos de comunicação.

Palavras-chaves: Dall-e 2; inteligência artificial generativa; semiótica; interações discursivas; comunicação.

Abstract: This study analyzes the processes of setting discourse with generative artificial intelligence technologies on the Dall-e 2 website by a semiotic proposition about discursive interactions. The projection in the enunciation of a human and a machine actors produces effects of meaning of reflexive and unidirectional bilateral interactions between these actors that require new competences from the recipient to interpret. This emphasizes the need for the enunciator to identify these discursive marks toward establishing communication contracts.

Keywords: Dall-e 2; generative artificial intelligence; semiotics; discursive interactions; communication.

Inteligência artificial generativa

A inteligência artificial (IA), desde os anos 1950, vem avançando com um arcabouço de conhecimento voltado para realizar ações que anteriormente só poderiam ser executadas por humanos, tendo impactado significativamente diversos setores produtivos, entre os quais destaca-se a área da comunicação. Nesse cenário, experiências como as de *chatbots*, assistentes virtuais ou recomendação de conteúdo em redes sociais, se popularizaram e fizeram com que algoritmos de IA se tornassem atores que desempenham papel relevante na produção de conteúdo consumido no espaço digital. Assim, esses algoritmos inteligentes atuam na mediação entre o acesso às informações e o consumo pelo usuário. Hepp (2020, p. 10) caracteriza os algoritmos como “mídia dentro da mídia”, uma vez que as estruturas algorítmicas estão inseridas nas estruturas das mídias, notadamente às conectadas online. Dessa forma, os algoritmos estão englobados em lógicas e estruturas pré-existentes, atuando diretamente nos protocolos de comunicação virtual e na análise em larga escala dos dados disponibilizados.

Nas últimas duas décadas, procedimentos e métodos conhecidos como *machine learning* e *deep learning* possibilitaram que algoritmos avançassem no reconhecimento de padrões de *big datas* com a finalidade de gerar, de forma autônoma, conteúdos textuais. O avanço tecnológico possibilitou o desenvolvimento da área conhecida como inteligência artificial generativa (GEN-AI), em que os algoritmos de redes neurais profundas, modelos inspirados na organização do sistema nervoso humano, apoiados em aprendizagem de máquina a partir de dados, elaboram conteúdos para fruição dos usuários. Observa-se, nesse cenário, tecnologias capazes de produzir textos a partir de diferentes linguagens, como: verbais escritos, musicais e imagéticos. Cada algoritmo é especializado em produzir determinado tipo de conteúdo de todas as formas, entretanto, a prévia análise de dados e sua decodificação matemática é procedimento obrigatório para a posterior codificação em novo texto. Para tanto, cabe ao algoritmo, a partir de uma programação, manipular os elementos constituintes das linguagens para que, em um fazer combinatório, apresente uma nova configuração textual.

Nesse escopo, a área da computação desenvolveu, juntamente com linguistas, melhorias na competência de máquinas em compreender a linguagem natural humana. Esta área, conhecida como Processamento de Linguagem Natural (PLN) (Hepp, 2020), permitiu que comandos escritos em linguagem verbal pudessem ser compreendidos pelos algoritmos, que por sua vez operam apenas na lógica binária,

isto é, matemática. Assim, há um certo fazer interpretativo algorítmico que permite a tradução das línguas naturais para linguagens de programação. Estas, por sua vez, vão operar em um movimento contrário: das linguagens de programação para as linguagens naturais ou artificiais, como a audiovisual, a musical ou a imagética.

O objeto deste trabalho é investigar as interações discursivas, notadamente os tipos de interações enunciativas presentes nas produções imagéticas realizadas pelo algoritmo Dall-e 2, com base nas proposições de Ana Claudia Oliveira (2013). Dall-e 2 é um sistema de criação de imagens figurativas por meio de algoritmos, a partir de instruções realizadas por linguagem verbal escrita. O usuário enuncia uma descrição verbal do que gostaria que o sistema produzisse e o algoritmo a transforma em imagem. Observa-se, portanto, uma tradução intersemiótica operacionalizada pela informática, programada para compreender instruções da linguagem verbal escrita de modo a processar as instruções e produzir um enunciado imagético.

Entende-se, dessa forma, que as ações realizadas pelo Dall-e 2 operam no movimento de interpretar as instruções verbais do usuário para produzir um novo discurso – em imagem – contemplando um determinado universo figurativo que encontra correspondência naquela instrução original; como observado na Figura 1.

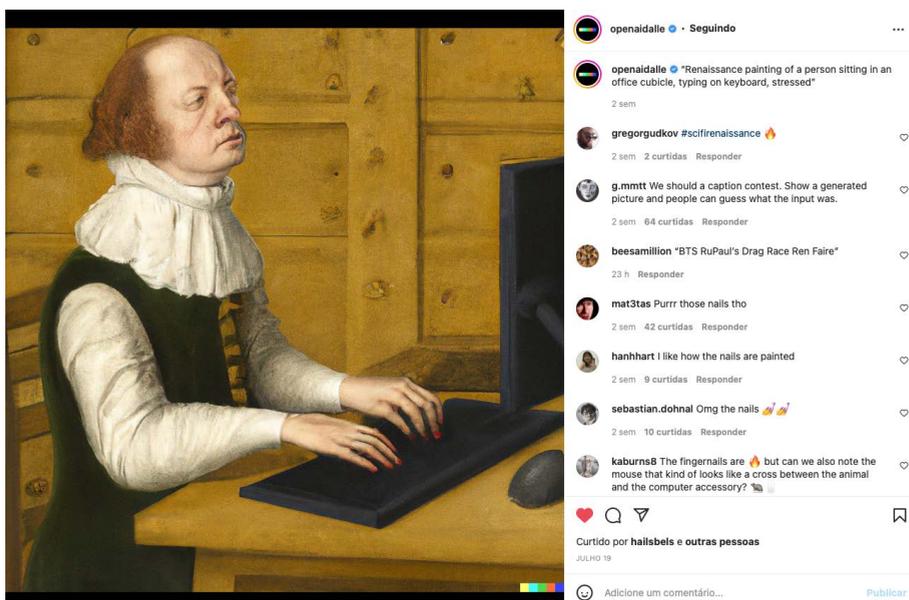


Figura 1: Exemplo de imagem produzida pelo algoritmo Dall-e 2 disponibilizado no Instagram do sistema.

Fonte: OpenAI, 2019

Na legenda da imagem temos a descrição enviada ao sistema: “uma pintura renascentista de uma pessoa sentada em um cubículo de escritório, digitando no teclado, estressada” (OpenAI, 2019). Ao lado, a imagem enunciada pelo algoritmo é apresentada. Um sujeito humano, compenetrado, diante, e com as mãos sobre uma das partes, do equipamento que está sobre a mesa, trajando vestimentas típicas do *quattrocento* italiano. Observa-se que as cores na composição da imagem referem-se aos tons terrosos típicos de uma tradição pictórica dos quadros do período. A própria feição do sujeito representado em tela é referente a uma apropriação do estilo renascentista. O algoritmo, no entanto, adapta a figurativização do dispositivo eletrônico, que não existia em 1500, e o apresenta conforme estilística do período – interpretando e criando a representação de um espaço, de um tempo e de um ator que reforça a isotopia iconográfica de uma época, mas com uma quebra figurativa e temática, em função dos equipamentos eletrônicos: teclado, mouse e monitor.

O algoritmo Dall-e 2 foi criado pela *startup* norte-americana OpenAI, uma organização que tem por finalidade a pesquisa laboratorial sobre GEN-AI. Entre seus investidores está a Microsoft, que desde 2019 aportou mais de 11 bilhões de dólares. O produto mais popular da OpenAI é um *chatbot* generativo chamado ChatGPT, um modelo de conversação em que o algoritmo assume papel criador ao reagir aos *prompts*³ do usuário. Fazem parte do conselho administrativo da empresa empresários da área de tecnologia, como Elon Musk, proprietário do X (antigo Twitter) e de empresas *high-tech* como a Tesla e a SpaceX; Sam Altman, investidor anjo de empresas como PayPal, Airbnb e Reddit; Peter Thiel, sócio da Meta, empresa controladora do Facebook e Instagram; e Reid Hoffman, cofundador do LinkedIn. A OpenAI se posiciona como empresa sem fins lucrativos, entretanto, os investidores do projeto têm licença de usufruir dos avanços tecnológicos gerados pela *startup*. Neste cenário, concorrem com o Dall-e 2 aplicações como Stable Difusion, Midjourney e Adobe Firefly, sendo que essas tecnologias operam de forma similar: a partir de *prompts* enunciados pelos usuários, buscando a geração de imagens a partir de modelos de IA.

Para a realização desta pesquisa, solicitamos acesso ao Dall-e 2 na qualidade de pesquisadores. Nisso, foram cedidos 30 créditos gratuitos para geração de imagens a partir dos comandos verbais escritos. Demarcando, assim, a lógica do modelo de negócios do algoritmo: cada geração de imagens custa um crédito e o usuário, após o término, pode adquirir quantidades específicas. Em março de 2023, o pacote

³ Em computação, os *prompts* são as mensagens, ou comandos, que indicam uma ação a ser executada.

inicial de créditos, contendo 115 gerações, tinha o valor de USD 15. Neste trabalho, analisaremos uma geração imagética realizada pelo algoritmo a partir de um enunciado verbal.

O presente trabalho parte da ideia da ocorrência da autonomia da IA na colocação em discurso desses textos. Isso porque as imagens são resultantes da transposição realizada pela IA de uma linguagem para outra, procurando produzir efeitos de sentido semelhantes aos do texto verbal escrito, do qual deriva a figura produzida pelo Dall-e 2. Trata-se, portanto, de compreender como a GEN-AI se constrói como um enunciador pressupostamente autônomo e de que modo ocorrem as interações discursivas no processo de fruição e de comunicação.

O site Dall-e 2 é um algoritmo disponibilizado na World Wide Web, que, conforme mencionado, a partir de instruções em linguagem verbal escrita consegue enunciar uma imagem com base na descrição. Do ponto de vista da enunciação, apresentamos como hipótese a ocorrência de uma troca de posições na relação enunciador-enunciatório na práxis enunciativa do site, uma vez que o algoritmo parte de uma descrição realizada por um enunciador que posteriormente desempenhará o papel actancial de enunciatório do texto enunciado pelo algoritmo. Acreditamos que o algoritmo também desempenha na manifestação do site essas duas posições enunciativas de modo reverso, tendo em vista que se constitui primeiramente como enunciatório do texto verbal escrito produzido pelo enunciador-humano para atuar como enunciador do texto imagético que será fruído pelo enunciatório-humano.

Para compreender essa dinâmica enunciativa, vamos aplicar a sistematização elaborada por Oliveira (2013, p. 235) para abordar as interações discursiva e estruturada com base na premissa de que “a ação cognitiva do enunciador é a de fazer saber o enunciatório os modos por meio dos quais cada um dos sujeitos com seus papéis discursivos tem seus próprios desempenhos na elaboração do sentido”. Assim, o escopo teórico referente às relações interacionais na discursivização do site Dall-e 2 subsidiará a análise, uma vez que é no processo da interação, na relação e na alternância entre os atores da enunciação, nos mecanismos de debragem e embregem inerentes à delegação de vozes, que as projeções do simulacro de um ator-humano são colocadas em discurso em relação com o simulacro de um ator-algoritmo que o sentido faz ser.

A matriz da formulação teórica de Oliveira (2013) sobre as interações discursivas está na centralidade do pensamento landowskiano na formalização das condições das interações sociais, das quais decorrem uma correlação sistêmica entre “sentido” e “interação” na perspectiva sociossemiótica. Em *Interações*

Arriscadas (2014), Landowski amplia o arcabouço teórico da gramática narrativa no âmbito das relações sociais, formulando um modelo capaz de abrigar os diferentes tipos de interação com base no coeficiente de risco que rege as ações do sujeito na construção de sentido.

Para o autor, são quatro os regimes de interação e sentido definidos diante da iminência do risco: programação, manipulação, ajustamento e acidente. Tais regimes compõem o quadro de um modelo geral da interação no âmbito teórico-metodológico da semiótica, em que as interações actanciais do homem sobre o homem e deste sobre o mundo são analisadas a partir de suas lógicas interacionais produtoras de sentido. A contribuição de Landowski para a disciplina aprofunda o dinamismo do modelo semiótico, uma vez que, ao introduzir os regimes de ajustamento e acidente no bojo da disciplina, determina novas possibilidades de compreender a emergência do sentido nos diferentes discursos.

A sociosemiótica preconizada por Landowski assume um caráter dinâmico representado por um quadrado semiótico elíptico, em que os sujeitos interactantes constroem o sentido justamente pela interação. Nesse modelo, o sentido é fruto da interação por estar inserido em uma lógica social e existir ora por regularidade e intencionalidade, ora por sensibilidade e aleatoriedade, operando a partir de posições que podem ser cambiáveis – possibilitando observar a mobilidade das formas de interação. Portanto, os quatro regimes de sentido e interação descritos por Landowski configuram uma gramática narrativa cuja oposição de base dos modos de *ser* regidos pela regulação/programação em oposição à imprevisibilidade/acidente produz na denegação o eixo dos subcontrários, cuja oposição dos modos de *fazer* articula a manipulação *versus* ajustamento no processamento do sentido. Enquanto na dêixis da programação e manipulação os sujeitos em interação são mobilizados pelo regime de junção a graus de competência de natureza cognitivas e pragmáticas para a ação, o eixo do acidente e ajustamento mobiliza a percepção sensível nos modos de apreensão de sentido a partir do regime de união.

Landowski complementa, portanto, o modelo narrativo, correlacionando a dinâmica de tipos de sentido aos tipos de interação, mas Oliveira (2013) observa que o teórico não correlaciona a narratividade aos procedimentos assumidos pelo sujeito da enunciação, questão que a faz partir da hipótese de que são também os tipos de interação que constroem os tipos de fazer cognitivo sensível e inteligível, cujo palco de encenação é o discurso. Na continuidade, nosso objetivo então é depreender a possibilidade de correspondência entre os regimes de interação narrativa de Landowski e os que denominamos de interações discursivas. Se assim for, caminharemos para

as correlações entre nível discursivo e nível semionarrativo, pensadas em termos de tipos de presença do enunciador e do enunciatário nos discursos e como essas fazem ser e fazer o sentido (p. 238).

Ao ver a enunciação como um ato de instauração do sentido na interação e também por ela, em função do caráter intersubjetivo enquanto “presença” ou efeitos de presença, enquanto “posição e posicionamento”, ou efeitos de angulação, observação e focalização dos sujeitos no enunciado na busca de um trânsito entre inteligibilidade e sensibilidade, Oliveira (2013) reconhece que nesse enunciar há um modo de instalação de presenças discursivas que resultam em modos de interação entre os atores do discurso, refletindo em modos de produção de sentido. Com base na premissa de que o nível discursivo abriga a atividade geradora de sentido, a autora ressalta que dessas realizações emergem

formas de transitividade que vão da unidirecionalidade às várias direções, por transitividade sem troca de posições, por transitividade com reciprocidade em que se trocam as posições, às transitividades reflexivas que são marcadas pela imprevisibilidade de ocupar as posições no discurso. (Oliveira, 2013, p. 243).

Cabe mencionar que, para dar conta dos vários modos de presença, a autora sistematiza as situações enunciativas em dois eixos englobantes: a situação enunciativa de intransitividade, marcada pelos posicionamentos de diferentes atores, em que há uma interação decorrente do comando de um para o outro, e as situações enunciativas de transitividade, caracterizadas pelas direções das posições entre o enunciador e o enunciatário (Oliveira, 2013, p. 244). Em função da transitividade ou intransitividade, é construído o quadro “Os tipos de Interação entre Enunciador e Enunciatário” onde são decompostas, de forma detalhada, as posições descritas como a “interação unidirecional” e a “interação bilateral com tipos de transitividade” (cf. Oliveira, 2013, p. 246), que subsidiará a análise de experiência realizada no site Dalle 2.

Website Dall-e 2

Em um primeiro momento, ao acessar o site Dall-e 2, vislumbramos uma página dividida em três seções (Figura 2). A primeira, localizada na parte superior da página, cumpre a função de cabeçalho e apresenta duas subseções com *links* que dão acesso ao histórico de geração de imagens e às coleções criadas pelo usuário. Além dessas funções, na parte superior o usuário pode acessar suas informações de cadastro e adquirir créditos para geração de novas imagens.

Na segunda seção, localizada centralmente na disposição visual do site, vemos uma “caixa de diálogo”, termo da área da informática que designa o espaço em que o usuário pode escrever comandos. No caso de Dall-e 2, observamos que a caixa conta com uma sugestão, uma indicação de como realizar as descrições para que o algoritmo a transforme em imagem: “An Impressionist oil painting of sunflowers in a purple vase...”, em tradução literal, “uma pintura a óleo impressionista de girassóis em um vaso roxo”. Assim, verifica-se que o site, enquanto destinador da comunicação, exemplifica uma determinada forma de interação entre o usuário e o algoritmo. No *prompt* de exemplo, temos uma descrição figurativa que determina um objeto (girassóis em um vaso roxo), o efeito de sentido de suporte da imagem (tela de pintura), o efeito de sentido da materialidade da pintura (a óleo), e a estilística da imagem gerada (escola impressionista). Dessa forma, o site sugere quais elementos constituintes das imagens são importantes para geração do conteúdo.

Ainda nessa segunda seção central, temos o comando *Surprise me* (surpreenda-me), em que o site gerará uma imagem a partir de uma descrição aleatória. E, abaixo da caixa de diálogo, tem-se a função de *upload*, em que o usuário pode enviar para o site uma imagem já existente e editá-la a partir do Dall-e 2. Já na terceira seção, na parte inferior da interface do site, vemos imagens e, conforme passamos o cursor do mouse sobre elas, aparecem os textos descritores utilizados para sua geração, conforme Figura 3.

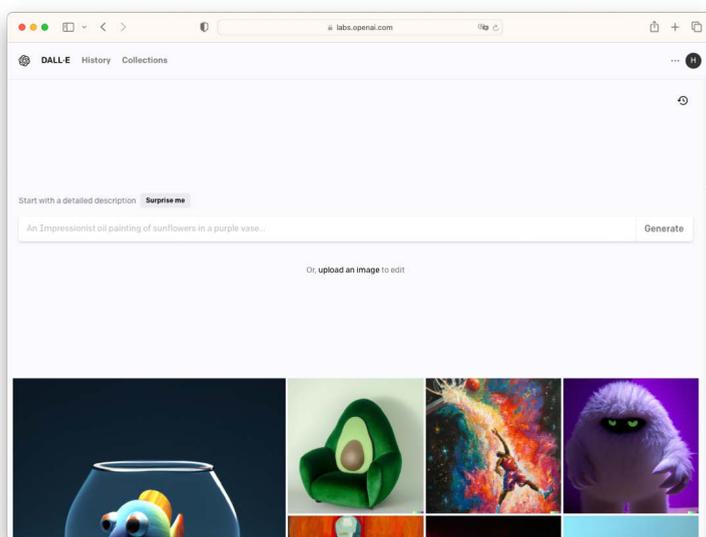


Figura 2: Interface do site Dall-e 2.

Fonte: Dall-e, 2022.

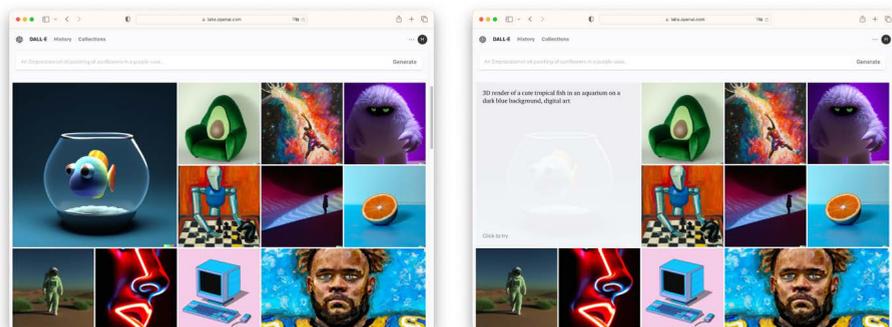


Figura 3: Interface do site Dall-e 2, parte inferior e detalhe da descrição de uma das imagens.
 Fonte: Dall-e, 2022.

Dessa forma, observamos que o destinador da comunicação do site atua no sentido de munir o destinatário em competências modal e pragmática, sensíveis e inteligíveis, para que a interação entre usuário e algoritmo aconteça. O destinador fornece ao destinatário exemplos de descritores que geraram imagens significativas. Esses exemplares podem ser percebidos pelo site como modelos para que, ao solicitar uma geração nova, seja utilizado esse mesmo escopo de detalhes. Identifica-se que as imagens selecionadas pelo destinador da comunicação para dispô-las em sua *homepage* apresentam uma diversidade figurativa e temática (um peixe, uma poltrona em forma de avocado, um jogador de basquete, um robô jogando xadrez, um astronauta, um computador etc.). Entretanto, observa-se que todas as imagens apresentam em suas representações apenas um elemento focal (o peixe, a poltrona, o robô, o astronauta). As imagens exemplares, portanto, não representam um determinado espaço ou uma paisagem, mas objetos e personagens em cenas projetadas para centralizá-los.

Assim, entendemos pertinente salientar que o destinador do site Dall-e 2 estabelece um contrato fiduciário no qual o fazer interpretativo do destinatário identifique a ação da IA na realização de uma tradução intersemiótica entre a linguagem verbal escrita e a linguagem imagética. Esta, por sua vez, é calcada em efeitos de sentido que buscam representar outras linguagens, como a de tradição pictórica iconográfica, fotográfica, a de quadrinhos, a simulação digital de imagens, entre outras. O algoritmo, portanto, age em fazer ser um simulacro de um determinado suporte e/ou técnica de inscrição, articulando uma visualidade que se projeta de forma múltipla e que encontra referente nas produções linguageiras

da tradição da enunciação figurativa visual e plástica, articulando um fazer que busca a figurativização.

O ato de representar imagens figurativas (analógicas ou digitais) em suportes planares é um esforço de circunscrever lógicas de apreensão de como representar o mundo natural, sensível, por meio da iconicidade. Isto exige, conforme apontado por Jean-Marie Floch (1985), um aprendizado de enunciação por parte daquele que cria e daquele que frui. O enunciador projeta formas de apreensão sensível do mundo natural em figuras do domínio do texto; já o enunciatário opera um fazer cognitivo, inteligível e sensível, que incide na percepção das figuras textuais, as reminiscências do mundo natural – para que, assim, façam sentido.

Cabe destacar que a enunciação imagética realizada de forma autônoma pelo algoritmo do site Dall-e 2 ocorre a partir de padrões contidos nos bancos de dados cuja fonte é a representação do mundo natural. Conforme discutido por Greimas (2004) em “Semiótica figurativa e semiótica plástica”, o mundo natural pode ser interpretado como mundo do senso comum, e o ato de representar pode ser identificado como uma operação de “imitação” (Greimas, 2004, p. 78). Fundamentalmente, conforme o autor avança na discussão da temática, a interpretação que o sujeito faz do mundo é o que vai dar significado e possibilitar a objetificação dos traços sensíveis presentes no discurso do mundo natural. Assim, categorizamos o mundo pelo crivo de leitura que fazemos dele, é isto que: torna significante o mundo ao nos permitir identificar as figuras, como objetos, ao nos permitir classificá-las, relacioná-las umas às outras, interpretar os movimentos como processos que se podem atribuir ou não a sujeitos etc.; sendo de natureza semântica – não visual, auditiva ou olfativa, por exemplo –, ela serve de “código” de reconhecimento que torna o mundo inteligível e manuseável (Greimas, 2004, p. 79).

Destacamos que a característica das imagens produzidas em Dall-e 2 resultam do processo de figurativização numérica, constitutivo das imagens digitais, pois não atua em uma lógica de morfogênese por projeção analógica, mas sim de morfogênese por simulação, em que o cálculo, o modelo matemático, permite a realização imagética (Couchot, 2011). Assim, não há uma relação pressuposta entre o numérico e a imagem, já que toda a informação da imagem digital foi programada, permitindo uma tradução abstrata entre o cálculo numérico e a figura digital, afastando-se, portanto, do real. Isto não quer dizer que a figurativização a partir de uma lógica informática descarte a visualidade do mundo natural. O que ocorre é um processo de modelização, isto é, a criação e a utilização de modelos matemáticos,

físicos, biológicos, entre outros, para simular o real. A produção da imagem passa a estar relacionada ao inteligível, ao processo de decomposição do sensível em modelos autônomos que realizam um simulacro planar e que serão base para as produções imagéticas numéricas. A articulação se dá entre *softwares*, modelos e imagens – em uma lógica espaço-temporal editável e manipulável.

Assim, ao tratarmos da tradução intersemiótica de Dall-e 2, observamos que o movimento de interpretação de uma linguagem para a textualização em outra implica uma dimensão não manifestada discursivamente, que é a linguagem de informática. A manifestação, portanto, comporta duas semióticas: a verbal escrita e a imagética, que perfazem dois enunciados englobados pelo site compreendido como o enunciado englobante. Conforme Oliveira (2013, p. 241), “os mecanismos de ter sentido e fazer sentido são determinados pelas escolhas da dimensão figurativa e de sua manifestação plástica, assim como dos tipos de presenças do sujeito da enunciação”. Tal premissa é particularmente operacional para analisar a dinâmica de produção de sentido na textualização em Dall-e 2, conforme veremos a seguir.

Do verbal ao imagético em Dall-e 2

Utilizamos como descrição verbal escrita o seguinte enunciado: “saída de crianças da escola em Singapura”. Observa-se, nessa descrição, que o actante da narrativa (crianças) está espacializado (escola em Singapura), agindo no tempo presente (saindo). Dessa forma, considerando a instância enunciativa do *ego*, *hic et nunc*, o enunciado descritor explicita as categorias da pessoa, do tempo e do espaço. Assim, poderemos verificar a interpretação que o algoritmo de Dall-e 2 faz dessa cena narrativa na sintaxe e na semântica do nível discursivo, contemplando a instância de instauração das lógicas enunciativas ou enuncivas, embreadas ou debreadas, em um eu-aqui-agora ou ele-lá-então, a partir dos procedimentos de figurativização e tematização.

Ao digitar o *prompt* descritor, o site altera sua configuração visual e apresenta uma nova tela: uma barra de carregamento surge por alguns segundos em que seguidamente é substituída por quatro imagens centralizadas na tela em um fundo branco, conforme a Figura 4. Essas imagens, sempre quatro, independentemente do enunciado verbal escrito, são versões da tradução entre o enunciado criado pelos simulacros enunciativos de ator humano e de ator algoritmo, cabendo a este a geração da imagem.

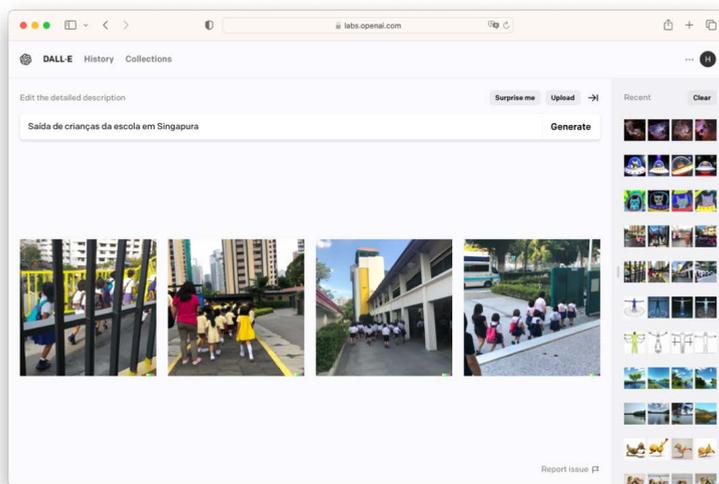


Figura 4: Interface do site Dall-e 2.

Fonte: Dall-e, 2015–2024.

Identifica-se nessas quatro imagens uma padronização em determinados elementos constituintes: (1) o aspecto de proporção é quadrado, isto é, o formato da imagem, independentemente de seu conteúdo, apresenta um fechamento em que a largura e altura são da mesma medida; (2) as quatro imagens apresentam uma “assinatura” digital, conforme detalhe apresentado abaixo (Figura 5), disposta no canto inferior direito. A assinatura é composta por cinco quadrados de diferentes cores, podendo estabelecer uma correspondência com a escala cromática RGB, padronagem de códigos para exibição de cores em suportes digitais, diferente, por exemplo, do padrão CMYK, padronagem de cores para impressão física de materiais coloridos. Assim, a assinatura da produção do algoritmo de Dall-e 2 assume o papel de representar a autorialidade digital da geração das imagens autônomas, já que explicita um caráter de criação, e que, também, só existe enquanto produto do digital.

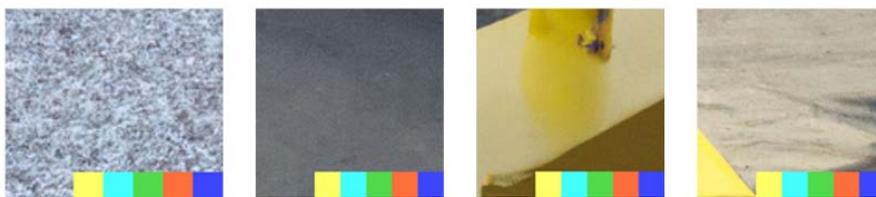


Figura 5: Recorte das imagens geradas pelo site Dall-e 2.

Fonte: Dall-e, 2015–2024.

Agora, ao observar as quatro imagens geradas, constatamos suas semelhanças em termos do conteúdo que manifestam: todas apresentam o actante da narrativa que determinamos no enunciado verbal escrito. As crianças são pequenas e com uma vestimenta que se assemelha a de uniformes escolares – já que se observa uma padronagem nas roupas entre elas. Assim, podemos inferir que o algoritmo conseguiu realizar o movimento de interpretação em que “crianças saindo de escola” não são crianças visitantes, mas àquelas que de fato estão saindo da escola em que estudam. Tanto na imagem 2 quanto na 4, vemos além dos estudantes a presença de mulheres adultas, dispostas observando as crianças caminhando, podendo ser interpretadas como professoras ou monitoras que assumem um papel temático de “supervisão das crianças”, ou até mesmo de “proteção”, fazendo com que o espaço “escolar” esteja atrelado a valores semânticos da ordem da “vigilância” e da “proteção”.

Todas as crianças e os adultos estão dispostos de costas para o observador-enunciatório da imagem. Podemos considerar essa escolha enunciativa a partir do efeito de sentido de “registro fotográfico”, como se o enunciatador tivesse registrado a saída de crianças da escola em Singapura como um algo espontâneo, um momento cotidiano, um registro quase burocrático ou com finalidades jornalísticas. Isso gera um efeito de sentido de distanciamento, não há uma convocação do enunciatário para a cena, não há um chamamento entre os actantes da narrativa e o actante enunciatário observador. Entretanto, o fato de as crianças estarem figurativizadas de costas, caminhando juntas em direção específica, configura um outro efeito de sentido: o da saída. Instaura-se, assim, que a proximidade do primeiro-plano da imagem, isto é, o plano mais próximo do enunciatário-observador é o ponto de “entrada” da cena, já o plano mais profundo, para onde as crianças estão se direcionando, representa a “saída”. A estratégia de colocação das figuras nessa configuração topológica gera o sentido, percebido de forma sensível e inteligível, da saída, do movimento de estar se distanciando do edifício, do interior.

Ainda considerando a figurativização das crianças e das adultas, observa-se que elas têm em comum uma característica física específica: a cromática de seus cabelos é escura; todas, sem exceção, apresentam cabelos pretos. Essa visualidade remete à etnia asiática, fazendo com que a espacialização presente no enunciado verbal escrito, Singapura, encontre na configuração figurativa da actorialização do enunciado imagético sua referência. Portanto, estabelece-se uma interpretação e correlação, no processo de tradução intersemiótica, entre as categorias enunciativas de espaço e da pessoa.

O observador, assim, assume que o efeito de sentido do movimento cinético projetado no enunciado imagético corresponde à figura do plano do conteúdo do discurso verbal escrito da “saída”, de estar em um espaço e agir para estar em outro. As crianças que estão figurativizadas nas cenas estão situadas em ambientes abertos e estão caminhando, juntas e em grupo, para a mesma direção. Podemos interpretar esses espaços como transitórios, isto é, que conectam o espaço público (calçada, rua, avenidas) e o espaço privado (edifícios), em que se observa a presença de grades de proteção, caminhos pré-traçados com marcação no solo, portarias e muros – implicando, portanto, um determinado traçado que se programa ao transeunte, uma determinação de como agir *em transição* entre um espaço que é público e outro que é privado.

As interações discursivas

Para Oliveira (2013, p. 246), os tipos de interação entre enunciador e enunciatário apresentam uma primeira segmentação que abrange duas possibilidades: a interação unidirecional e a interação bilateral – com tipos de interatividade. No tocante à interação unidirecional, enunciador e enunciatário estão separados nos seus atos de enunciação, bem como encontram-se em patamares diferentes e hierarquizados, nos quais enunciador comanda a enunciação enquanto destinador, e enunciatário processa as marcas deixadas por esse destinador, caracterizando seus fazeres pelos atos de coerção da interação regida pelo enunciador. Na interação unidirecional não se estabelece um contrato enunciativo, mas o comando de um ator da enunciação sobre o outro.

Já na interação bilateral de tipos de transitividade, a autora argumenta que as relações de enunciador e enunciatário podem ser de três tipos: negociadas, em reciprocidade presencial ou reflexivas na enunciação em ato. Na enunciação em que o fazer enunciativo é negociado, enunciador-destinador e enunciatário-destinatário são sujeitos do querer, regidos pela intencionalidade do enunciador, que detém um saber sobre o simulacro de enunciatário para propor opções de interação e modos de negociação, visando o convencimento deste enunciatário. Nesse sentido, enquanto destinador, o enunciador doa ao enunciatário as competências cognitivas e performáticas, e o contrato estabelecido é baseado no convencimento e na fidedignidade dos passos de fazer o sentido. Na reciprocidade presencial, por sua vez, o caráter volitivo e a semelhança das competências cognitivas do enunciador e enunciatário permitem que a sensibilidade atue instaurando a dinâmica da reciprocidade, fazendo intuir o sentido que é sentido. Nesse caso, não há contrato, mas contato e experiência

do fazer conjuntamente o sentido desejado. Por último, a reflexividade diz respeito à troca de papéis na atuação do ato de enunciar. Neste tipo de interação bilateral transitiva, enunciador e enunciatário também são sujeitos volitivos e dotados de iguais competências cognitivas, possibilitando a troca de posições no comando das marcas enunciativas.

A leitura dos elementos sintáticos e semânticos das imagens geradas pelo algoritmo Dall-e 2 constitui um primeiro patamar de observação do objeto. No entanto, baseados na demarcação teórica de Oliveira (2013), acreditamos que a inovação deste dispositivo propõe desafios para pensar as relações enunciativas em textos autônomos. Nesse sentido, é possível identificar dois tipos de interação entre enunciador e enunciatário no site Dall-e 2, sendo um com efeito de sentido de uma interação bilateral com transitividade na qual enunciador e enunciatário estão sincretizados em um mesmo ator, trocando papéis ao atuar em reflexividade no ato de enunciar. Conforme detalharemos mais adiante na análise, o ator-humano e o ator-maquínico produzem o efeito de sentido de serem dotados de iguais competências cognitivas, sendo o ator-maquínico programado para um querer-fazer que atenda os comandos do ator-humano.

O segundo tipo é a interação unidirecional nas projeções de enunciados realizados por atores que exercem papéis temáticos diferentes, ator-humano e ator-algoritmo, englobados na manifestação do site. Na sequência, vejamos como as postulações de Oliveira (2013, p. 246) balizam a leitura semiótica desse objeto.

Em uma primeira aproximação, verifica-se que no objeto em análise, site Dall-e 2, há ocorrência da interação unidirecional estabelecida entre enunciador-destinador e enunciatário-destinatário, estando em patamares diferenciados e hierarquizados, em que enunciador comanda a enunciação enquanto destinador, cabendo ao enunciatário processar as marcas deixadas pelo primeiro no discurso, sendo os seus fazeres determinados por esses atos de coerção da interação regida pelo enunciador. Podemos identificar a ação desta lógica enunciativa na relação ator-humano – enunciador de um enunciado verbal – e ator-algoritmo – que processa o comando produzindo um enunciado imagético na perspectiva de uma leitura das estratégias enunciativas no site Dall-e 2. Dessa forma, em relação ao tipo de contrato estabelecido, a leitura semiótica corrobora a postulação de Oliveira (2013) de que não há um contrato, mas um comando que prescreve a reconstrução de seu sentido posto. Nesse sentido, na interação unidirecional entre enunciado produzido por ator-humano e enunciado produzido pelo ator-algoritmo, o resultado da tradução intersemiótica do verbal para o imagético é resultante dos fazeres determinados por uma interação unidirecional.

Também de acordo com as segmentações de Oliveira (2013), observamos na análise a produção do efeito de sentido de uma interação bilateral, com uma transitividade reflexiva na qual enunciador e enunciatário trocam papéis ao atuarem no ato de enunciar. Entretanto, essa reflexividade realizada por sujeitos volitivos, dotados de iguais competências, está, neste caso, estruturada no sincretismo dos papéis actanciais de enunciador e enunciatário no mesmo ator que interage com o site Dall-e 2, conforme já assinalado. Isso porque o conteúdo gerado por uma situação enunciativa instaurada no interior do texto, a partir da ação deste enunciador, o fará cumprir também o papel de enunciatário, pois é ele que produz um enunciado verbal, projetando o simulacro de um ator cujo papel temático é representar a intervenção humana no site. Este enunciado verbal será interpretado pelo simulacro de ator, cujo papel temático é representar a intervenção maquínica operada pelo algoritmo, capaz de gerar um enunciado com efeitos de sentido semelhantes, mas em linguagem imagética. O ato de produzir o enunciado verbal e fruir o enunciado imagético gerado pelo ator-algoritmo, sancionando-o como semelhante ou não ao conteúdo verbal, é o que faz o sentido de um mudar de posições no comando dos mecanismos enunciativos pelo enunciador, que assume também o papel de enunciatário. Não há contrato entre os dois coenunciadores, mas são sintonizados em reflexividade, com alternância de presenças dos papéis enunciativos.

Assim, no que diz respeito às interações discursivas analisadas na experiência realizada no site Dall-e 2, constatou-se que, na dinâmica do processamento do sentido, parte-se da interação unidirecional verificada no enunciado, posicionada assertivamente enquanto programação na lógica da regularidade no eixo que articula o *ter sentido*, para representar figurativamente os efeitos gerados pela linguagem verbal escrita na intransitividade das posições e previsibilidade dos turnos. O ator-humano produz o enunciado verbal, matriz da representação imagética realizada pelo ator-maquínico. No que diz respeito às movimentações dos sujeitos da enunciação, relacionadas à interação bilateral com tipos de transitividade, observa-se, no compartilhamento dos papéis de enunciador e enunciatário desempenhados na troca de posições no comando dos mecanismos enunciativos desempenhados pelo ator-humano e pelo ator-maquínico, a reflexividade de ambos, produzindo, em última instância, um efeito de sentido de sincretização do simulacro de enunciador no processo de produção de conteúdo em Dall-e 2. A transitividade e a reflexividade ajustam-se, neste caso, às projeções das trocas de posições que um mesmo ator deve ocupar para apreensão do sentido.

Considerações finais

O site Dall-e 2 é resultado do desenvolvimento e da evolução dos softwares de IA voltados ao aprendizado de geração de textos – os chamados sistemas de IA generativa. Os textos produzidos por este algoritmo se caracterizam por realizar a tradução intersemiótica de um enunciado verbal para um enunciado imagético. Consideramos que tal tecnologia eleva o patamar de representação figurativa e temática na produção de imagens geradas pela máquina, com impactos relevantes no desenvolvimento de novas competências voltadas ao fazer interpretativo dos receptores. As imagens geradas pelo Dall-e 2 apresentam desafios importantes no estabelecimento de contratos enunciativos em função da ambiguidade dos efeitos de sentido de parecer verdadeiro ou falso de cada imagem. A artificialidade da matriz imagética apoiada não apenas na digitalização do suporte mas, agora também, na autonomia da máquina capaz de “criar”, requer um saber-fazer dos humanos para ler o conteúdo apresentado. A aplicação de tal recurso na produção de conteúdo exige, portanto, novas competências enunciativas do destinador e destinatário da comunicação.

A partir da representação imagética realizada pela IA do Dall-e 2, o presente trabalho procurou verificar em que medida a análise dos tipos das interações discursivas na dinâmica do processamento do sentido pode contribuir para a reflexão sobre os desafios impostos com automação integral ou parcial nos processos de colocação em discurso. As projeções enunciativas, os contratos fiduciários e a construção dos simulacros de enunciador e enunciatário são questões de base para se pensar as relações de comunicação e o eventual potencial disruptivo, com a introdução de um ator da enunciação automatizado, que impõe novos parâmetros para se pensar os enunciados.

Nesse sentido, verificamos que tal fenômeno pode ser observado com os avanços teóricos relativos às postulações sobre a transitividade e intransitividade implicadas no fazer enunciativo entre os atores humano e maquínico no processo de discursivização, seja ele híbrido entre homem e máquina, seja autônomo, prescindindo da manifestação das marcas discursivas do ator humano. Assim, as competências semântica e modal do ser humano para um fazer interpretativo sobre os simulacros de enunciador e enunciatário, bem como destinador e destinatário dos conteúdos produzidos integral ou parcialmente por IA, são necessárias para a compreensão da produção desses conteúdos cada vez mais processados por máquinas e de seus impactos no ecossistema da comunicação.

Referências

COUCHOT, 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.

DALL-E. *OpenAI*. [S. l.]: 2022. Disponível em: <https://labs.openai.com>. Acesso em: 14 mar. 2023.

FLOCH, J.-M. *Petites mythologies de l'oeil et de l'esprit: pour une sémiotique plastique*. Paris: John Benjamins, 1985.

GREIMAS, A. J. “Semiótica plástica e semiótica figurativa”. In: OLIVEIRA, A. C. (Org.). *Semiótica plástica*. São Paulo: Hacker, 2004. p. 75-96.

HEPP, A. “Artificial companions, social bots and work bots: communicative robots as research objects of media and communication studies”. *Media, Culture Society*, Thousand Oaks, v. 42, n. 7-8, p. 1410-1426, 16 mai. 2020. DOI: 10.1177/0163443720916412.

LANDOWSKI, E. *Interações arriscadas*. São Paulo: Estação das Letras e Cores, 2014.

OLIVEIRA, A. C. “As interações discursivas”. In: OLIVEIRA, A. C. (Ed.). *Interações sensíveis: ensaios de sociosemiótica a partir da obra de Eric Landowski*. São Paulo: Estação das Letras e Cores, 2013. p. 235-249.

OPENAI. Renaissance painting of a person sitting in an office cubicle, typing on keyboard, stressed. [S. l.]: OpenAI, 19 jul., 2019. Disponível em: <https://www.instagram.com/p/CgMt5J0r-s1/?hl=pt>. Acesso em: 07 ago. 2022.

submetido em: 27 dez. 2023 | aprovado em: 30 abr. 2024