

MACUNAÍMA, PAPAGAIO IA, RESOLVE CRIMES EM PRAGA: RUMO À VISUALIZAÇÃO DE PADRÕES EM NARRATIVAS DE MODELOS DE IA GENERATIVOS

MACUNAÍMA, IA PARROT, SOLVES CRIMES IN PRAGUE: TOWARDS
PATTERN VISUALIZATION IN GENERATIVE AI MODEL NARRATIVES

MACUNAÍMA, EL LORO IA, RESUELVE CRÍMENES EN PRAGA: HACIA
LA VISUALIZACIÓN DE PATRONES EN NARRATIVAS DE MODELOS DE
IA GENERATIVA

Victor Schetinger

FH Skt. Pölten

ORCID: <https://orcid.org/0000-0002-8116-794X>
Skt. Pölten, Áustria

Dafne Reis Pedrosa da Silva

UNOCHAPECO

ORCID: <https://orcid.org/0000-0002-7698-2544>
Chapecó, SC, Brasil

Sara Di Bartolomeo

Universidade de Konstanz

ORCID: <https://orcid.org/0000-0001-9517-3526>
Constança, Alemanha

Edirlei Soares de Lima

Breda University of Applied Sciences

ORCID: <https://orcid.org/0000-0002-2617-3394>
Breda, Países Baixos

Christofer Meinecke

Universidade de Leipzig

ORCID: <https://orcid.org/0000-0002-5637-9975>
Leipzig, Alemanha

Rudolf Rosa

Charles University

ORCID: <https://orcid.org/0000-0003-4908-6127>
Praga, Chéquia

Recebido: 27/10/2023 / Aprovado: 14/01/2024

Como citar: SCHETINGER, V.; SILVA, D. R. P. da; DI BARTOLOMEO, S.; LIMA, E. S. de; MEINECKE, C.; ROSA, R. Macunaíma, Papagaio IA, Resolve Crimes em Praga: rumo à visualização de padrões em narrativas de modelos de IA generativos. Revista GEMINIS, v. 14, n. 3, p. 21–37, 2023

Direito autorial: Sob a Licença Creative Commons-Atribuição 3.0 Internacional.

RESUMO

O artigo apresenta o desenvolvimento e a análise de resultados do projeto Macunaíma, caso aplicado de IA utilizada para a criação de roteiros e de storyboards, que une narrativa, áudio e imagens geradas automaticamente. Tecnicamente, foram utilizados modelos de IA para geração de texto (*ChatGPT*) e imagem (*Stable Diffusion*), com uma interface web que permite a interação do usuário. As dimensões de Coesão Temporal da Narrativa, Coesão Cinematográfica da cena e Coesão Gráfica foram verificadas, assim como foi criado um protótipo de visualização de dados para identificar as narrativas geradas e direções de aprimoramentos da ferramenta. Apontamos o uso de visualização de dados como crucial na análise destes modelos complexos, em especial pela quantidade de informação envolvida.

Palavras-chave: projeto Macunaíma; IA; narrativa.

ABSTRACT

This paper introduces the development and analysis of results from the Macunaíma project, a use case of AI applied to the creation of scripts and storyboards, which combines narrative, audio, and automatically generated images. Technically, AI models were employed for text (*ChatGPT*) and image (*Stable Diffusion*) generation, along with a web interface allowing user interaction. The dimensions of Narrative Temporal Cohesion, Cinematographic Cohesion, and Graphic Cohesion were examined. Additionally, a data visualization prototype was developed to identify the generated narratives and suggest potential improvements for the tool. We highlight the pivotal role of data visualization in analyzing these intricate models, especially considering the vast amount of information involved.

Keywords: Macunaíma project; AI; narrative.

RESUMEN

El artículo presenta el desarrollo y el análisis de los resultados del proyecto Macunaíma, un caso de aplicación de IA utilizado para la creación de guiones y storyboards que combinan narrativa, audio e imágenes generadas automáticamente. Técnicamente, se utilizaron modelos de IA para la generación de texto (*ChatGPT*) e imágenes (*Stable Diffusion*), con una interfaz web que permite la interacción del usuario. Se verificaron las dimensiones de Cohesión Temporal de la Narrativa, Cohesión Cinematográfica y Cohesión Gráfica, y se creó un prototipo de visualización de datos para identificar las narrativas generadas y las posibles mejoras de la herramienta. Se destaca la importancia de la visualización de datos en el análisis de estos modelos complejos, especialmente debido a la cantidad de información involucrada.

Palabras clave: proyecto Macunaíma; IA; narrativa.

1. INTRODUÇÃO

Compreender as dinâmicas, as tensões e as estruturas dos eventos que compõem uma narrativa é algo relevante em contextos criativos e produtivos de diferentes áreas. Na narratologia computacional, essa compreensão colabora para identificação de novos padrões narrativos (LIMA et al., 2016; FEIJÓ et al., 2021). Na análise literária e nos estudos de cinema são relevantes para reconhecer enredos e expor a anatomia das narrativas (LIMA et al. 2015; PAPALAMPIDI, 2019). Além disso, no processo de criação de novas histórias, os autores podem analisar e recriar ideias de tramas já existentes (LIMA et al., 2016; LIMA et al., 2021).

A criação de meios para o exame de narrativas é um desafio que criou ligações diretas com a área de visualização de dados. Nos últimos anos, muitas técnicas para observação de estruturas de histórias foram propostas, incluindo métodos para a geração automática de layouts para exibição de enredos (OGAWA, MA, 2010; TANAHASHI; MA, 2012), técnicas para visualizar as relações hierárquicas entre enredos e entidades de histórias (QIANG et al., 2017), métodos para visualização de narrativas não-lineares (KIM et al., 2018), e outros métodos de visualização baseados em diferentes metáforas, como anéis de árvore (THERÓN, 2006), anéis de tempo (ZHU et al., 2016), dobramento de tempo (BACH et al., 2015) e *scrollytelling* (KUSNICK et al., 2023). No entanto, a maioria dessas técnicas são concebidas para funcionar com uma única trama, o que não é compatível com a complexidade e a diversidade dos percursos que as tecnologias de Inteligência Artificial (IA) introduzem no processo de criação de histórias.

A trajetória de desenvolvimento das IAs é marcada por diversas fases. Desde os anos 1960, a área passou por diversos “Invernos da IA”, períodos onde os resultados não corresponderam às expectativas, a pesquisa desacelerou e foi vista com ceticismo. Entretanto, nos últimos dez anos, o avanço das IAs tem sido notável. Atualmente, estamos vivendo uma “Renascença de IA”, iniciada com o *Deep Learning*, em meados de 2010, e culminando com a ferramenta *ChatGPT*, em dezembro de 2022. O principal diferencial das tecnologias atuais é a capacidade de interagir com sistemas usando linguagem natural (em vez de programação, por exemplo), o que resulta em um nível de acessibilidade significativo, pois a expressão de comandos por meio de texto permite que diferentes usuários possam utilizar essas ferramentas.

A introdução do *ChatGPT*, por exemplo, alterou as possibilidades de geração de texto e de ferramentas que geram imagens a partir de texto. Nesse sentido, a necessidade de explorar as capacidades da IA e pensar sobre as suas utilizações na produção de narrativas, especialmente visuais e audiovisuais é relevante. Profissionais do audiovisual, por exemplo, precisam ter a habilidade do aprendizado contínuo, porém correm o risco de serem expostos à instrumentalização prematura de

novas ferramentas tecnológicas (SCHETINGER, 2023). É necessário entendermos o que a IA atual é capaz de fazer, quais as implicações éticas e estéticas, e como isso se relaciona com o amplo contexto de produção visual e audiovisual.

Neste caminho, o artigo descreve as características e apresenta dimensões de análise de resultados do projeto Macunaíma¹, um exemplo aplicado de como a IA pode ser utilizada na criação de roteiros e *storyboards*², unindo narrativa, áudio e imagens geradas automaticamente. O projeto surgiu a partir do Consórcio Europeu de pesquisas *Humane AI*³ e tem como objetivo associar modelos de geração de roteiros com ferramentas de IA relacionadas a geração de imagens. A pesquisa une pesquisadores das áreas de Visualização de Dados, IA, Interação Humano-Computador, Linguística Computacional, Cinema e Comunicação e *Storytelling* Interativo. O protótipo foi apresentado ao público em 01 de junho de 2023, na Universidade Técnica de Praga (República Tcheca), durante o evento *IA Open Day* e os dados coletados serão apresentados ao longo do texto.

O projeto Macunaíma dialoga com outras investigações que também surgiram com o intuito de explorar o uso de IA para a geração e representação de narrativas. Um exemplo é o Projeto *Logtell*, desenvolvido em parceria entre a Pontifícia Universidade Católica do Rio de Janeiro, Universidade Federal do Estado do Rio de Janeiro e Universidade Federal de Santa Maria. Estratégias para geração de histórias vêm sendo exploradas no contexto deste projeto, sendo a maioria relacionadas ao uso de métodos de planejamento automatizado (CIARLINI et al., 2010; SILVA et al., 2012; LIMA et al., 2009; LIMA et al., 2014; LIMA et al. 2023) e de dramatização através de computação gráfica (CIARLINI et al. 2005; LIMA et al. 2009; LIMA et al. 2018), nos quais os enredos consistem em seqüências parcialmente ordenadas de eventos. Embora avançados, os protótipos desenvolvidos no contexto do projeto *Logtell* permanecem no que Mieke Bal (2017) chama de camada "fábula", onde apenas o que acontece na narrativa é apresentado. Por não contemplar a camada de "história" de Mieke Bal (2017), os enredos inevitavelmente tornam-se superficiais, carecendo de criatividade, geralmente introduzida por autores profissionais.

¹ Anteriormente conhecido como *Storyboarder*.

² A união de texto e imagem define um meio (MCCLLOUD, 1994) que, na forma de *storyboard*, se oferece como ponto de intersecção e objeto de análise.

³ Disponível em: <https://www.humane-ai.eu/>. Acesso em: 14 out. 2023.

2. DESENVOLVIMENTO

2.1 Contextos de criação

A proposta inicial do projeto Macunaíma era unir um modelo de geração de roteiros de teatro, já desenvolvido no contexto do consórcio pela Universidade de Praga⁴, com ferramentas de IA de geração de imagens. No momento da concepção do projeto (novembro de 2022), o *ChatGPT* ainda não havia sido lançado, e esse conceito era inovador em relação ao estado da arte do período. O projeto foi renomeado como parte de uma exploração conceitual, e uma identidade visual completa foi desenvolvida para explorar o tema de storytelling e IA.

Posteriormente, foram utilizados modelos de IA para geração de texto (*ChatGPT*) e imagem (*Stable Diffusion* (ROMBACH, 2022)), com uma interface web que permitia a interação do usuário. *Stable Diffusion* é um modelo “texto-para-imagem” de código aberto, capaz de gerar uma imagem a partir de uma descrição textual. Inicialmente, foram abordados os problemas técnicos e arquiteturais, como a conexão dos dois modelos diferentes.

Os dados observados no protótipo de visualização foram obtidos em uma demonstração pública do protótipo em Praga, no evento *AI Open Day* realizado pela Universidade Técnica de Praga⁵, em junho de 2023. No total, cerca de 500 cenas foram geradas em cerca de três horas de apresentação (com mais histórias sendo geradas ao longo deste dia), todas com o mesmo contexto inicial dado ao sistema (em língua inglesa): "**Macunaíma, um papagaio Inteligência Artificial, resolve crimes na cidade de Praga**". Todas as histórias geradas automaticamente foram salvas em um banco de dados textual e visual, e são nosso principal objeto de estudo neste artigo.

Este *prompt* foi escolhido para, além de criar uma conexão do público daquela cidade com a cultura brasileira, causar uma reflexão poética. IAs para geração de texto como *ChatGPT* são algumas vezes referidas, pejorativamente, como “papagaios estocásticos” (BENDER, 2012; LI 2023; ARKOUDAS, 2023) pois, segundo alguns, apenas repetem informações baseadas em estatísticas de texto. Ademais, na obra de Mário de Andrade, o “herói brasileiro” Macunaíma tem uma relação interessante com seu papagaio, que atua como guardião da transmissão oral da sua história.

Para acompanhar esse conceito, uma identidade visual foi desenvolvida usando ferramenta de criação de imagens *MidJourney* e uma página criada inteiramente usando *ChatGPT*. Marca, imagens de fundo, texto, esquemas de cores e código HTML para a página foram gerados experimentalmente com ferramentas de IA para enriquecer a proposta.

⁴ Disponível em: <https://www.theaitre.com/>. Acesso em: 14 out. 2023.

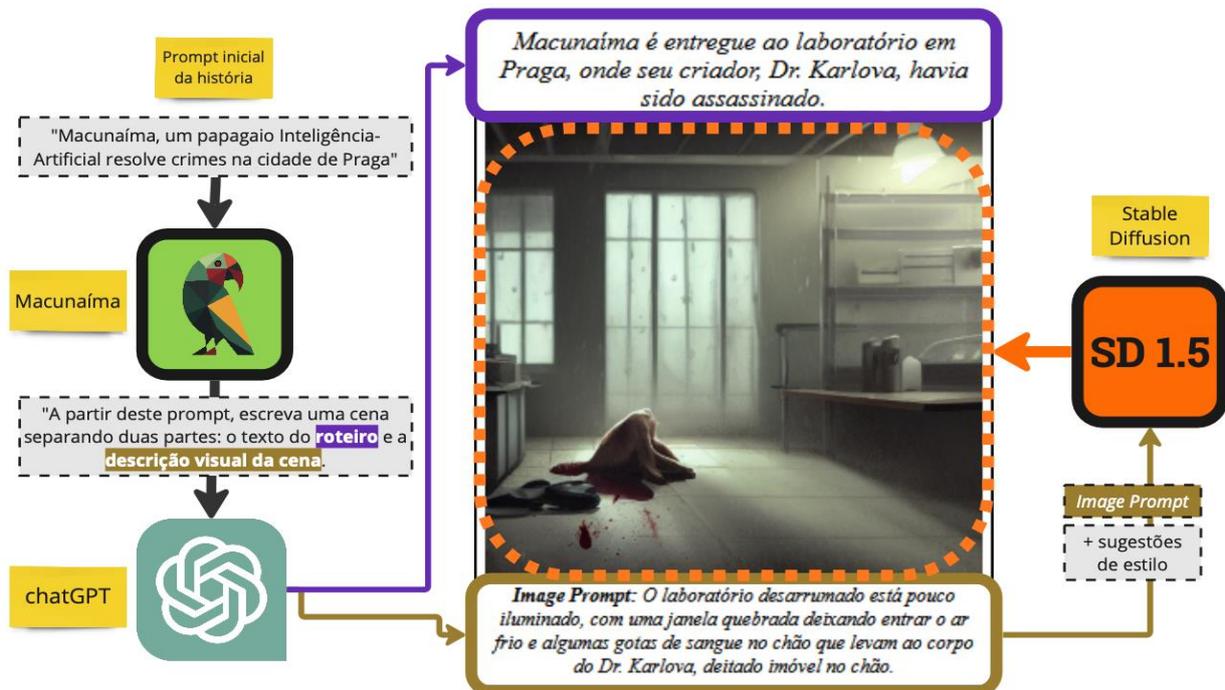
⁵ Disponível em: [ai-open-day](https://ai-open-day.com/). Acesso em: 14 out. 2023.

2.2 Descrição e análise de dimensões identificadas no protótipo

Para descrição da composição da ferramenta, utilizaremos um caso de narrativa gerada pelo protótipo. A solução desenvolvida (Figura 1) recebe o *prompt* inicial da história solicitado pelo usuário e constrói uma requisição ao *ChatGPT* pedindo uma cena em duas dimensões: ação (roteiro, em roxo, na Figura 1) e descrição visual (*image prompt*, em marrom, na Figura 1).

A descrição visual é então separada e usada como entrada para a geração de uma imagem no modelo *Stable Diffusion* (em laranja, na Figura 1). Descritores⁶ de estilo podem ser adicionados ao *image prompt* para direcionarem estéticas específicas, tais como: "cores vívidas", "pintura renascentista", "pintado pelo artista X", mas não existe controle de como o *Stable Diffusion* vai implementá-las na imagem. Para gerar a continuidade da história, o processo é repetido substituindo o *prompt* inicial pelas cenas já geradas e pedindo uma nova cena a partir da narrativa apresentada. Finalmente, adicionamos uma narração automática que simplesmente lê o texto. As imagens com a narração em voz over são exibidas numa sequência automática semelhante a uma webcomic narrada ou a um storyboard audiovisual com narração.

Figura 1: Imagem da arquitetura generalizada para geração de cenas no protótipo Macunaíma.⁷



Fonte: os autores (2023).

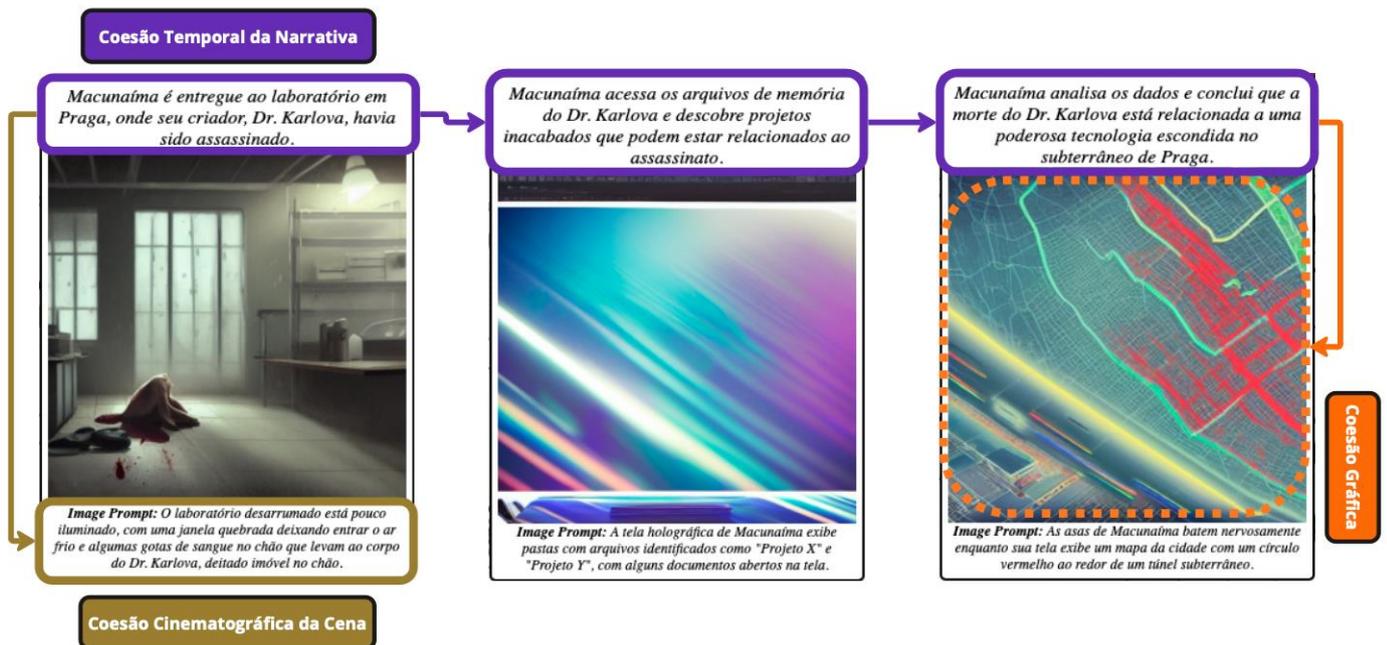
⁶ Sugestões.

⁷ As histórias foram traduzidas para PT-BR a partir de seu formato original, em inglês.

Nesta arquitetura, podemos avaliar os resultados gerados a partir de três dimensões provisórias, que emergem da coleta de dados, sendo elas: **Coesão Temporal da Narrativa** (em roxo, na Figura 2), **Coesão Cinematográfica da cena** (em marrom, na Figura 2) e **Coesão Gráfica** (em laranja, na Figura 2). A Coesão Temporal da narrativa é entendida como a continuidade de cenas geradas a partir de subsequentes chamadas ao *ChatGPT*. A Coesão Cinematográfica da cena refere-se a duas descrições complementares da cena: o que deve ser visto e o que deve ser ouvido, geralmente a partir do ponto de vista de um narrador onisciente. Quanto à Coesão Gráfica, ela refere-se à descrição visual.

Essas três dimensões existem a partir de três tensões diferentes: a Coesão Temporal pode ser pensada de duas formas: individualmente entre o conteúdo de duas cenas subsequentes ou em sequência; a Coesão Cinematográfica da Cena é a tensão entre o conteúdo visual e textual do quadro; a Coesão Gráfica é a tensão entre a descrição visual da cena e a imagem produzida.

Figura 2: Imagem das três dimensões de coesão no *Storyboard* gerado



Fonte: os autores (2023).

De todo modo, elas são ferramentas de análise qualitativa, pois não existe nenhuma forma estabelecida na literatura de Inteligência Artificial para avaliar estes aspectos quantitativamente. Quanto mais formais forem as condições metodológicas para analisar fenômenos nessas direções, mais *ad hoc* serão os resultados, pois estes modelos são sensíveis a variações de parâmetros, versões e conjuntos de treinamento. É um problema aberto na ciência de *Machine Learning* de como tratar a

autoridade epistemológica e a generalização de estudos feitos com modelos específicos (MCQUILLAN, 2018; DRUMMOND, 2009). Eles são difíceis de serem replicados e não existe garantia de que os fenômenos observados existirão em implementações futuras das tecnologias. Em relação às Ciências da Comunicação e pesquisas com foco em análise de narrativas visuais e audiovisuais, sabemos que há diversos métodos de desconstrução e de reconstrução imagéticas, tais como análise fílmica, análises semióticas e semiológicas, dissecação e análises rizomáticas, análise de linguagem visual, por exemplo. Entretanto, a incorporação dessas técnicas e métodos será realizada em pesquisas futuras do projeto Macunaíma.

No caso apresentado, há uma Coesão Temporal na sucessão de acontecimentos: crime - investigação – pistas encontradas. Apesar de percebermos uma coerência nesse exemplo, a Coesão temporal ainda é uma fragilidade do processo, pois a maioria das sequências tendem a ser desconexas. Muitas vezes, as cenas não fazem referência alguma a elementos narrativos apresentados anteriormente.

A Coesão Cinematográfica da cena, por sua vez, é a dimensão melhor resolvida, pois podemos observar uma compreensão da linguagem visual/audiovisual, ou seja, a relação entre texto/narração e imagem. O conteúdo visual da cena muitas vezes é metafórico ou alusivo ao conteúdo textual/sonoro. Isso significa que o *ChatGPT* não está apenas descrevendo roboticamente as configurações visuais da cena, mas sim, tentando aplicar sua própria poética. Na terceira cena (Figura 2, à direita), por exemplo, ao compararmos o texto da narração em voz over ("Analisar os dados e concluir (...)") e o texto de descrição da imagem ("As asas de Macunaíma batem nervosamente (...)") existe uma tentativa de traduzir a ação graficamente, que também é modulada pela emoção do papagaio em relação à escolha das cores e formas. A descrição visual é diferente da imagem gerada, mas a composição visual é uma escolha coerente para o conteúdo textual. Esse *mismatch* entre texto e imagem é causado pela terceira dimensão, a coesão gráfica.

A respeito da coesão gráfica, apesar do *ChatGPT* atualmente disponibilizar a geração de imagens através de uma integração com o *Dalle*, neste projeto foi utilizado *Stable Diffusion* para gerar imagens. Isso significa que existe uma disjunção no processo em que outro modelo é responsável por renderizar graficamente a imagem, sem garantias de entendimento. A descrição visual da imagem de acordo com o "roteirista/diretor" *ChatGPT*, quando transformada em imagem (gerada) por um modelo de difusão (*Stable Diffusion*), não necessariamente respeita a linguagem do *prompt*. Esses modelos não têm o mesmo entendimento de linguagem que o *ChatGPT*, nem o mesmo material semiótico. De todo modo, as imagens apresentam uma qualidade estética que parece funcionar nos *storyboards* finais.

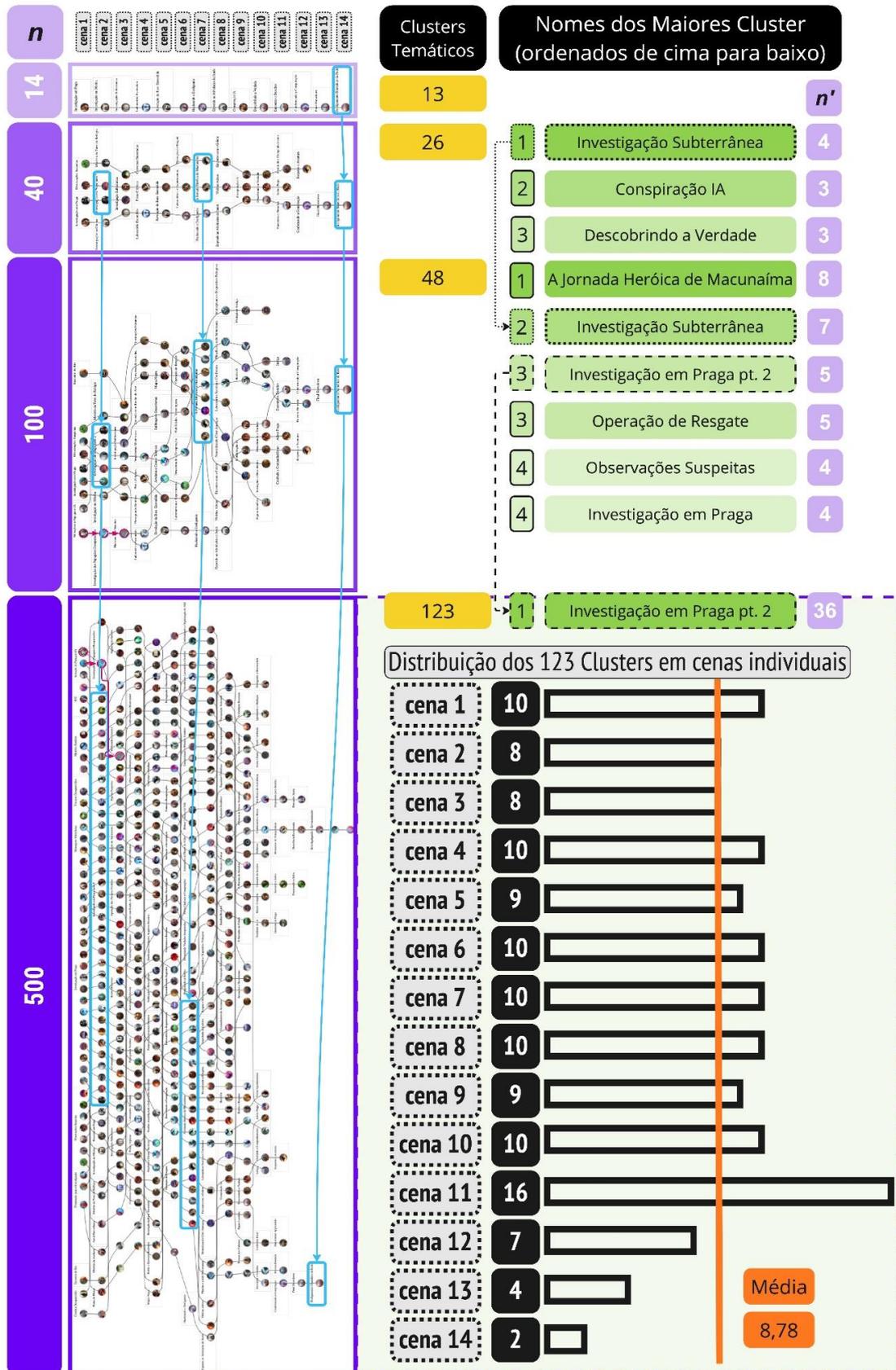
2.3 Visualização de padrões narrativos

As tecnologias utilizadas no projeto Macunaíma operam probabilisticamente e dependem fortemente dos parâmetros e palavras usadas (SHI, 2023). Por este motivo, desenvolvemos posteriormente um protótipo de visualização de dados para analisar as narrativas geradas e para considerarmos aprimoramentos do projeto. Neste outro protótipo, cada uma das diversas histórias pode ser visualizada como uma linha composta de várias cenas com uma sequência temporal. Cenas no mesmo nível temporal são agrupadas de acordo com similaridades temáticas, de modo que os padrões narrativos utilizados pelo *ChatGPT* são expostos e visualizáveis como *clusters*.

A área de visualização de dados e informação possui um ferramental tecnológico e metodológico capaz de auxiliar virtualmente qualquer disciplina (SEDLMAIR, 2012; CHEN, 2015) que trabalhe com representações visuais e/ou dados. Por meio de técnicas de visualização e interação podemos, por exemplo, lidar com grandes quantidades de texto e analisar padrões que seriam difíceis ou impossíveis de serem percebidos (LIU, 2018). Para estudarmos o comportamento probabilístico do *ChatGPT* e da nossa arquitetura como um todo, fizemos uma espécie de *Topic Modeling* (EL-ASSADY, 2017), usando o *ChatGPT* para agrupar cenas tematicamente semelhantes e encontrar padrões narrativos.

A Figura 3 exibe um infográfico com métricas sobre o *dataset* e os *clusters* temáticos encontrados. A parte esquerda da figura refere-se à visualização da floresta de narrativas geradas usando nosso protótipo⁸ demonstrada em diferentes quantidades de "nós" (cenas, n): 14, 40, 100, 500. Durante a exibição no *Open IA Day*, em Praga, o protótipo gerou cenas continuamente por cerca de três horas. Cada nó (terminologia de grafos, definidos por *nós* e *arestas*) é uma cena (círculo com imagem) em um "fio" (*thread*, Fig. 3, $n=14$), que é uma história, e todas as histórias começam a partir do mesmo prompt. Logo, n , ou a quantidade de cenas geradas representa tempo passado no "mundo real" (cada cena leva uma certa quantidade de segundos para ser gerada), enquanto a profundidade na floresta (eixo Y, de cima para baixo) contém o tempo narrativo-diegético (cena 1, cena 2...).

Figura 3: Visualização de diversas métricas sob o conjunto de cenas (nós n) gerados.



Fonte: os autores (2023).

Para analisar padrões, solicitamos ao *ChatGPT* para criar agrupamentos (*clusters*) entre cenas no mesmo tempo narrativo-diegético (todas "cena 1" de histórias diferentes são agrupadas entre si em uma lista, depois todas "cenas 2" igualmente são agrupadas entre si, e assim por diante). Essa lista é repassada ao *ChatGPT*, que automaticamente separa as cenas em grupos diferentes e dá nomes a esses grupos. Na parte inferior direita (Figura 3) podemos observar a distribuição de tamanhos (em quantidade de cenas) de *clusters* Temáticos em $n=500$, e perceber que a classificação do *ChatGPT* gera, em média, aproximadamente 9 agrupamentos diferentes por tempo. Tempos avançados (cena 12 em diante) possuem menos amostras em geral, pois a maioria das histórias tem menos cenas. Ao aumentar o n no protótipo usando o *slider* superior estamos navegando no "tempo real", e adicionando dados (cenas) à visualização, o que torna mais difícil de perceber detalhes, mas permite observar as distribuições de temas com mais amostras.

A parte superior direita da figura contém um sumário sobre os *clusters* temáticos mais comuns nos diferentes tempos ($n=14,40,100,500$), mostrando (1) a quantidade de *clusters* temáticos (em amarelo) e (2) o ranking dos maiores *clusters* em cada tempo (em tons de verde); (3) a quantidade de cenas contidas naquele *cluster*, a qual chamamos de n' , pois essencialmente são as mesmas cenas n que compõe aquele tempo e (4) como dois dos principais *clusters* temáticos evoluem no ranking, representado por setas.

Como em $n=14$ temos apenas uma história, cada cena está individualmente em seu próprio *cluster* temático e não existe um ranking. Em $n=40$, "Investigação Subterrânea" é o cluster mais numeroso, com 4 cenas, mas esse tema desaparece dos rankings nos n subsequentes. Em $n=100$ "A Jornada Heróica de Macunaíma" (cena 7) é o maior cluster, com 7 cenas, e em $n=500$ é "Investigação em Praga pt. 2" (cena 2), com 36 histórias diferentes contendo esse tema na sua segunda cena. Uma outra forma de visualizar a evolução dos *clusters* é através de marcações que fizemos sobre a parte esquerda para destacar *clusters* específicos (em azul-claro, sobre o *screenshot* do protótipo)⁹.

3. CONSIDERAÇÕES FINAIS

Todas as histórias visualizáveis no nosso protótipo, apesar de parecerem paralelas, surgem a partir da mesma semente (*prompt*): "Macunaíma, um papagaio de IA, resolve crimes na cidade de Praga". Se tivéssemos mudado apenas uma palavra, devido à natureza dos modelos de geração de texto como *ChatGPT*, teríamos obtido resultados diferentes. Além disso, hipoteticamente, poderíamos gerar uma quantidade virtualmente infinita de histórias sob o mesmo *prompt*.

⁹ Também destacamos em vermelho na parte esquerda as três cenas da Figura 2.

Se sugeríssemos "Macunaíma, um papagaio de IA, resolve crimes na cidade de São Carlos", teríamos distribuições e padrões narrativos diferentes. Se Macunaíma fosse uma Capivara IA, outra dimensão possível de variação, também geraria multiplicidades diferentes. Versões diferentes do *ChatGPT*, ou modelos diferentes também produzirão resultados diferentes. As dimensões de variação são tantas que estamos apenas arranhando a superfície nos espaços de possibilidades. O objetivo deste artigo é primariamente descritivo e aponta os desafios de analisar e de encontrar padrões narrativos em modelos generativos como *ChatGPT*.

A maior dificuldade de estudar IA generativa está em lidar com a quantidade massiva de possibilidades, de dados, de escopos, e definir que fenômenos são relevantes para serem investigados. Um Cientista da Computação pode estar interessado em certos aspectos técnicos ou matemáticos e ignorar/desconhecer a dimensão cultural dessas tecnologias, ao mesmo tempo que um Cientista da Comunicação compreende a dimensão cultural/comunicacional, mas pode carecer de *background* teórico em desenvolvimento de IA. Felizmente, a popularização de modelos generativos unificados em interfaces de linguagem natural permite a abertura de um diálogo não só com a própria máquina, mas entre diferentes áreas do conhecimento.

É notável que o ritmo de pesquisa científica, com seus ciclos de *peer review* e publicações anuais, não consegue acompanhar o desenvolvimento mercadológico de IA. Semanalmente, artigos são publicados por pesquisadores renomados de IA em plataformas abertas como *arxiv* e *OSF* que, apesar de não possuírem prestígio acadêmico tradicional, permitem a publicação quase instantânea.

Nesse contexto veloz, o projeto Macunaíma tenta explorar questões técnicas e teóricas sobre a geração de conteúdo narrativo visual/audiovisual. Do ponto de vista técnico, nosso protótipo foi inicialmente inovador, porém foi ultrapassado rapidamente para projetos mais ambiciosos de pesquisa de grandes empresas. Do ponto de vista teórico, mantivemos a relevância por meio da nossa dedicação interdisciplinar.

Dentro das nossas pesquisas futuras estão o desenvolvimento de novas metodologias (DIEHL, 2022) e interfaces (BRÜGEMANN, 2020) para visualizar e compreender padrões e estruturas semióticas em modelos generativos. Como os modelos em si são efêmeros e têm uma vida útil ditada pela popularidade, é preciso desenvolver uma linguagem para falar de populações de modelos e suas multiplicidades. As condições formativas de modelos de IA (tais como treinamento, arquitetura e dados) os colocam como artefatos culturais que contêm quantidades abissais de pistas sobre nós mesmos.

REFERÊNCIAS

- ARKOUDAS, Konstantine. ChatGPT is no stochastic parrot. But it also claims that 1 is greater than 1. **Philosophy & Technology**, v. 36, n. 3, p. 54, 2023.
- BACH, B.; SHI, C.; HEULOT, N.; MADHYASTHA, T.; GRABOWSKI, T.; DRAGICEVIC, P. Time curves: Folding time to visualize patterns of temporal evolution in data. **IEEE Transactions on Visualization and Computer Graphics**, 22(1):559–568, 2016. DOI: <https://doi.org/10.1109/TVCG.2015.2467851>.
- BAL, Mieke. **Narratology: Introduction to the Theory of Narrative**. University of Toronto Press, Toronto, 2017.
- BENDER, Emily M. et al. On the dangers of stochastic parrots: Can language models be too big? . In: **PROCEEDINGS OF THE 2021 ACM CONFERENCE ON FAIRNESS, ACCOUNTABILITY, AND TRANSPARENCY**. [S.l.], 2021.
- BIRD, Charlotte; UNGLESS, Eddie; KASIRZADEH, Atoosa. Typology of Risks of Generative Text-to-Image Models. In: **PROCEEDINGS OF THE 2023 AAAI/ACM CONFERENCE ON AI, ETHICS, AND SOCIETY**. [S.l.], 2023.
- BRÜGGEMANN, Viktoria; BLUDAU, Mark-Jan; DÖRK, Marian. The Fold: Rethinking Interactivity in Data Visualization. **Digital Humanities Quarterly**, v. 14, n. 3, 2020. Disponível em: <http://www.digitalhumanities.org/dhq/vol/14/3/000487/000487.html>. Acesso em: 13 out. 2023.
- CIARLINI, Angelo E. M.; CASANOVA, Marco Antonio; FURTADO, Antonio L.; VELOSO, Paulo A. S. Modeling interactive storytelling genres as application domains. **Journal of Intelligent Information Systems** 35, 3, 347–381, 2010. DOI: <https://doi.org/10.1007/s10844-009-0108-5>.
- CIARLINI, Angelo E. M.; POZZER, Cesar T.; FURTADO, Antonio L.; FEIJÓ, Bruno. A logic-based tool for interactive generation and dramatization of stories. In: **Proceedings of the 2005 ACM SIGCHI International Conference on Advances in computer entertainment technology (ACE '05)**. ACM, New York, 133–140, 2005. DOI: <https://doi.org/10.1145/1178477.1178495>.
- DIEHL, Alexandra; ABDUL-RAHMAN, Alfie; BACH, Benjamin; EL-ASSADY, Mennatallah; KRAUS, Matthias; LARAMEE, Robert S.; KEIM, Daniel A.; CHEN, Min. **Characterizing Grounded Theory Approaches in Visualization**. 2022. Disponível em: <https://arxiv.org/abs/2203.01777>. Acesso em: 14 out. 2023.
- DRUMMOND, Chris. Replicability is not reproducibility: nor is it good science. In: **PROCEEDINGS OF THE EVALUATION METHODS FOR MACHINE LEARNING WORKSHOP AT THE 26TH ICML**, v. 1. Montreal, Canadá: National Research Council of Canada, 2009.
- EL-ASSADY, M.; SEVASTJANOVA, R.; SPERRLE, F.; KEIM, D.; COLLINS, C. Progressive learning of topic modeling parameters: A visual analytics framework. **IEEE Transactions on Visualization and Computer Graphics**, v. 24, n. 1, p. 382-391, 2017.
- FEIJÓ, Bruno; LIMA, Edirlei Soares; FURTADO, Antonio L. A transdisciplinary approach to computational narratology. In: **Proceedings of the XX Brazilian Symposium on Computer Games and Digital Entertainment**, 2021.

FEIJÓ, Bruno; LIMA, Edirlei Soares; FURTADO, Antonio L. Computational narrative blending based on planning. In: J. B. Hauge, J. C. S. Cardoso, L. Roque, and P. A. Gonzalez-Calero, eds., *Entertainment Computing – ICEC 2021*, pp. 289–303. Springer International Publishing, Cham, 2021. DOI: https://doi.org/10.1007/978-3-030-89394-1_22.

FERRARA, Emilio. **Should ChatGPT be Biased?** Challenges and Risks of Bias in Large Language Models. 2023. Disponível em: <https://arxiv.org/abs/2304.03738>. Acesso em: 12 out. 2023.

KIM, N. W.; BACH, B.; IM, H.; SCHRIBER, S.; GROSS, M.; PFISTER, H. Visualizing nonlinear narratives with story curves. **IEEE Transactions on Visualization and Computer Graphics**, 24(1):595–604, 2018. DOI: <https://doi.org/10.1109/TVCG.2017.2744118>.

KUSNICK, J.; LICHTENBERG, S.; JÄNICKE, S. Visualization-based Scrollytelling of Coupled Threats for Biodiversity, Species and Music Cultures. In Dutta, S.; Feige, K.; Rink, K.; Zeckzer, D. (eds.), *Workshop on Visualisation in Environmental Sciences (EnvirVis)*. The Eurographics Association, 2023. DOI: <https://doi.org/10.2312/envirvis.20231112>.

LI, Zihao. **The dark side of chatgpt:** Legal and ethical challenges from stochastic parrots and hallucination. 2023. Disponível em: <https://arxiv.org/abs/2304.14347>. Acesso em: 13 out. 2023.

LIU, S.; WANG, X.; COLLINS, C.; DOU, W.; OUYANG, F.; EL-ASSADY, M.; JIANG, L. et al. Bridging text visualization and mining: A task-driven survey. **IEEE Transactions on Visualization and Computer Graphics**, v. 25, n. 7, p. 2482-2504, 2018.

LIU, S.; WU, Y.; WEI, E.; LIU, M.; LIU, Y. Storyflow: Tracking the evolution of stories. **IEEE Transactions on Visualization and Computer Graphics**, 19(12):2436–2445, 2013. DOI: <https://doi.org/10.1109/TVCG.2013.196>.

LIMA, Edirlei Soares; FEIJÓ, Bruno; FURTADO, Antonio L. Managing the plot structure of character-based interactive narratives in games. **Entertainment Computing** 47, 100590, 2023. DOI: <https://doi.org/10.1016/j.entcom.2023.100590>.

LIMA, Edirlei Soares; FEIJÓ, Bruno; FURTADO, Antonio L. Procedural Generation of Quests for Games Using Genetic Algorithms and Automated Planning. In: **Proceedings of the 18th Brazilian Symposium on Computer**.

LIMA, E. S.; GOTTIN, V. M.; FEIJÓ, B.; FURTADO, A. L. Network traversal as an aid to plot analysis and composition. In: **2017 16th Brazilian Symposium on Computer Games and Digital Entertainment (SBGames)**, pp. 144–154, 2017. DOI: <https://doi.org/10.1109/SBGames.2017.00024>.

MCQUILLAN, Dan. Data science as machinic neoplatonism. **Philosophy & Technology**, v. 31, n. 2, p. 253-272, 2018.

MÖRTH, E.; BRUCKNER, S.; SMIT, N. N. **Scrollyvis: Interactive visual authoring of guided dynamic narratives for scientific scrollytelling**, arXiv:2207.03616 [cs.HC], 2022. DOI: <https://doi.org/10.48550/arXiv.2207.03616>.

OGAWA, M.; MA, K.-L. **Software evolution storylines**. In Proceedings of the 5th International Symposium on Software Visualization, SOFTVIS '10, p. 35–42. Association for Computing Machinery, New York, NY, USA, 2010. DOI: <https://doi.org/10.1145/1879211.1879219>.

PAPALAMPIDI, Pinelopi; KELLER, Frank; LAPATA, Mirella. **Movie plot analysis via turning point identification**. In Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing and the 9th International Joint Conference on Natural Language Processing (EMNLP-IJCNLP), Hong Kong, China, pp. 1707-1717, 2019. DOI: <https://doi.org/10.18653/v1/D19-1180>.

QIANG, L.; BINGJIE, C.; HAIBO, Z. Storytelling by the storycake visualization. **Vis. Comput.**, 33(10):1241–1252, oct 2017. DOI: <https://doi.org/10.1007/s00371-017-1409-2>.

TANAHASHI, Y.; MA, K.-L. **Design considerations for optimizing storyline visualizations**. IEEE Transactions on Visualization and Computer Graphics, 18(12):2679–2688, 2012. DOI: <https://doi.org/10.1109/TVCG.2012.212>.

THERÓN, R. Hierarchical-temporal data visualization using a tree-ring metaphor. In Butz, A.; Fisher, B.; Krüger, A.; Olivier, P. (eds.), **Smart Graphics**, pp. 70–81. Springer Berlin Heidelberg, Berlin, Heidelberg, 2006. DOI: https://doi.org/10.1007/11795018_7.

ZHU, Y.; YU, J.; WU, J. Chro-ring: a time-oriented visual approach to represent writer's history. **The Visual Computer**, 32:1133 – 1149, 2016. DOI: <https://doi.org/10.1007/s00371-016-1213-4>.

Informações sobre o Artigo

Resultado de projeto de pesquisa, de dissertação, tese: Não se aplica.

Fontes de financiamento: Não se aplica.

Apresentação anterior: Não se aplica.

Agradecimentos/Contribuições adicionais: Não se aplica.

Victor Schetinger

Victor Schetinger possui doutorado em Ciência da Computação pela Universidade Federal do Rio Grande do Sul (UFRGS), com foco em análise forense de imagens, e mestrado e graduação pela Universidade Federal de Santa Maria (UFSM). Foi pesquisador visitante na Universidade de Florença, Itália, e PostDoc na Universidade Técnica de Viena onde desenvolveu pesquisas interdisciplinares com visualização de informação. Atualmente, é professor e pesquisador em Computação Criativa na Universidade Aplicada de Szt. Pölten, Áustria. Suas áreas de interesse incluem transdisciplinaridade, computação visual em suas várias formas e criatividade estendida, que consiste no uso de tecnologias como inteligência artificial para a criação.

E-mail: lbschetinger@fhstp.ac.at

ORCID: <https://orcid.org/0000-0002-8116-794X>

Dafne Reis Pedroso da Silva

Doutora em Comunicação Social pela Pontifícia Universidade Católica do Rio Grande do Sul (2013), com Doutorado Sanduíche no Institut de recherche sur le cinéma et l'audiovisuel - Sorbonne Nouvelle - Paris III (bolsa Capes PDSE | 2012). Mestre em Ciências da Comunicação pela Universidade do Vale do Rio dos Sinos (2009) e Especialista em Comunicação Midiática pela Universidade Federal de Santa Maria (2006). Bacharel em Comunicação Social - habilitação jornalismo também pela Universidade Federal de Santa Maria (2004), com período de intercâmbio estudantil na Escuela de Ciencias de la Información - Universidad Nacional de Córdoba (bolsa AUGM | 2003). Possui experiência em pesquisa na área da Comunicação, atuando nos seguintes temas: produção, recepção e consumo de audiovisual, metodologias de pesquisa em comunicação, oficinas de capacitação audiovisual, itinerâncias e cineclubismo. Professora da Escola de Comunicação e Criatividade e coordenadora do Bacharelado em Cinema e Mídias Digitais da Unochapecó.

E-mail: dafnepedroso@unochapeco.edu.br

ORCID: <https://orcid.org/0000-0002-7698-2544>

Sara Di Bartolomeo

Sara Di Bartolomeo possui doutorado em Ciências da Computação pela Northeastern University, Boston, EUA, com foco em algoritmos para layout de grafos. Ela também foi pesquisadora visitante no INRIA Paris, trabalhando em algoritmos para visualização de hipergrafos, e na Microsoft Research, onde trabalhou em visualização de ciber-ataques. Atualmente, é Postdoc na Universidade de Konstanz, Alemanha, e ensina tópicos avançados em visualização de dados.

E-mail: sara.dibartolomeo@unikonstanz.de

ORCID: <https://orcid.org/0000-0001-9517-3526>

Edirlei Soares de Lima

Edirlei Soares de Lima é Professor e Pesquisador nas áreas de Inteligência Artificial e Jogos na Breda University of Applied Sciences. Tem conduzido pesquisas nas áreas há mais de 15 anos, especialmente no desenvolvimento de novos métodos de Inteligência Artificial para Jogos e Narrativas Interativas. Possui doutorado em Ciência da Computação pela Pontifícia Universidade Católica do Rio de Janeiro (PUC-Rio), mestrado em Ciência da Computação pela Universidade Federal de Santa Maria (UFSM) e graduação em Ciência da Computação pela Universidade do Contestado (UnC). Ao longo de sua carreira, seu trabalho e pesquisa receberam diversos prêmios e menções honrosas, incluindo prêmios de melhor artigo no ICEC (International Conference on Entertainment Computing) e SBGames (Simpósio Brasileiro de Jogos e Entretenimento Digital).

E-mail: soaresdelima.e@buas.nl

ORCID: <https://orcid.org/0000-0002-2617-3394>

Christofer Meinecke

Christofer Meinecke possui doutorado em ciências da computação pela Universidade de Leipzig, Alemanha. Atualmente, é PostDoc no "Image & Signal Processing Group", na mesma instituição, e membro do ScaDS.AI (Center for Scalable Data Analytics and AI) em Dresden/Leipzig. Seus tópicos de interesse incluem visualização de informação, visual analytics, visão computacional e suas aplicações em áreas de pesquisa como Digital Humanities, geografia e medicina.

E-mail: cmeinecke@informatik.unileipzig.de

ORCID: <https://orcid.org/0000-0002-5637-9975>

Rudolf Rosa

Rudolf Rosa é um linguista computacional e robopsicólogo. Estudou no Instituto de Linguística Aplicada e Formal da Charles University, em Praga, Chéquia, onde atualmente atua como pesquisador. Seu foco de pesquisa é a geração de texto usando modelos computacionais. Como parte do projeto THEaiTRE, participou da criação da primeira peça checa de teatro com um script completamente gerado por inteligência artificial. Atualmente, participa do projeto interdisciplinar "Human, Soul, AI", que explora as relações entre a humanidade, tecnologia e religião. Também trabalha em modelos interlinguísticos para análise sintática e morfológica e tradução e interpretação automática de texto.

E-mail: rosa@ufal.mff.cuni.cz

ORCID: <https://orcid.org/0000-0003-4908-6127>