

Engenharia de Prompt Orientada a Estrutura (SPE)

A Mitigação de Alucinações em Sistemas Legados:

O Caso "Lattes Architect"

Simpósio de Aplicação Prática de IA no Direito – Turma 03 ITS Rio & Trybe

Melhym Pereira Quemel

**Advogado. Mestre em Direito (UNESA). Graduando em Engenharia de Software.
Pesquisador em Direito Digital e Governança Algorítmica.**

Belo Horizonte

2025

Engenharia de Prompt Orientada a Estrutura (SPE): A Mitigação de Alucinações em Sistemas Legados Através do Caso "Lattes Architect"

Resumo

O presente artigo investiga a aplicação da Engenharia de Prompt Orientada a Estrutura (Structure-Oriented Prompt Engineering - SPE) como método para mitigar alucinações de Grandes Modelos de Linguagem (LLMs) na interação com sistemas legados governamentais. O problema central abordado é a complexidade taxonômica de plataformas de gestão curricular que, por possuírem arquitetura de informação datada, induzem modelos de IA a sugerir caminhos de preenchimento inexistentes. A metodologia adotada consistiu na mineração de dados e engenharia reversa dos códigos-fonte da interface da plataforma alvo (arquivos HTML e JavaScript) para a extração de uma ontologia de navegação imutável. Os resultados demonstram que a injeção dessa estrutura rígida no System Prompt, combinada com módulos de diagnósticos e oráculos, elimina a criação de categorias falsas. O estudo de caso analisa o cadastro de uma apresentação em seminário acadêmico, comprovando a precisão do método frente a abordagens convencionais. Conclui-se que a SPE transforma a experiência burocrática em um processo assistido e gamificado, assegurando a integridade documental exigida pelas agências de fomento.

Palavras-chave: Inteligência Artificial Generativa. Plataforma Lattes. Engenharia de Prompt. Sistemas Legados. Alucinação Artificial.

INTRODUÇÃO: O LABIRINTO TAXONÔMICO E A INSUFICIÊNCIA DA INTUIÇÃO

A gestão do capital intelectual na academia brasileira enfrenta um paradoxo tecnológico: enquanto a produção do conhecimento avança em velocidade exponencial, as ferramentas de registro desse conhecimento permanecem ancoradas em arquiteturas de informação de décadas passadas. A plataforma oficial de currículos do país, embora essencial, apresenta-se ao pesquisador não como um facilitador, mas como um labirinto taxonômico. A tarefa de converter uma produção intelectual dinâmica em campos estáticos de banco de dados assemelha-se a um exercício hermenêutico onde a interpretação errônea da norma (menu) resulta em sanção administrativa (perda de pontuação).

Nesse cenário, a Inteligência Artificial Generativa surge como uma promessa de "copiloto" para a burocracia acadêmica. Contudo, observa-se um fenômeno crítico: ao serem confrontadas com sistemas fechados e hierárquicos, as IAs tendem a incorrer no que denominamos "Alucinação Taxonômica". O modelo, treinado para ser prestativo e criativo, inventa menus e categorias que possuem lógica semântica, mas não existência

fática na plataforma. Ele sugere o caminho que "deveria existir", ignorando a realidade rígida do código legado.

Este trabalho propõe, portanto, uma ruptura com a engenharia de prompt tradicional, que se baseia apenas em instruções comportamentais. Apresentamos a Engenharia de Prompt Orientada a Estrutura (SPE), uma técnica que trata o código-fonte da interface do sistema (HTML e JavaScript) como doutrina canônica. A hipótese é que, ao injetar a árvore de decisão real do sistema no contexto da IA, restringimos sua "criatividade probabilística", forçando-a a atuar com precisão cirúrgica. Não se trata apenas de preencher um formulário, mas de realizar a "Engenharia Reversa Documental" para criar uma ferramenta que atue como um mentor sênior, capaz de guiar o pesquisador desde a inserção do dado até a estratégia de pontuação.

1. METODOLOGIA: DA ENGENHARIA REVERSA À ARQUITETURA DO PROMPT

Para que a solução proposta não seja apenas uma abstração teórica, é fundamental estabelecer um rigor metodológico que conecte a Ciência da Computação à Metodologia Científica clássica. A construção de um assistente de IA eficaz para sistemas legados exige, primeiramente, o reconhecimento do terreno onde ele irá operar. Portanto, este capítulo descreve o processo de "cartografia digital" realizado antes da redação de qualquer instrução para a Inteligência Artificial.

A pesquisa classifica-se como aplicada, de abordagem qualitativa e caráter exploratório-descritivo, utilizando como procedimento técnico a Pesquisa Documental em Código-Fonte (Source Code Mining). Diferentemente de abordagens que utilizam manuais de usuário ou guias em PDF — como os tutoriais tradicionais de exportação de currículo (RTF/XML) — optou-se pela extração direta da "verdade" do sistema: seus arquivos de execução.

Vale ressaltar que a metodologia SPE depende da estabilidade do sistema legado. Como o Lattes possui uma taxa de atualização de interface extremamente baixa (anos sem mudança estrutural no DOM), a 'Ontologia Injetada' possui alta durabilidade temporal. Assim, a coleta de dados seguiu três etapas de mineração direta no navegador durante uma sessão autenticada:

1. Mapeamento do DOM (Currículo Lattes.html): Identificação da estrutura base e dos pontos de entrada do menu principal.

2. Extração da Lógica de Negócio (cv_gui.js e lista.js): Análise dos scripts responsáveis pela renderização dinâmica. Aqui, descobriu-se que a plataforma não opera apenas por nomes, mas por chaves de identificação (IDs) e comportamentos de modais (modal_cv.js).

3. Validação via XML (4629884576807167.xml): O uso do currículo em formato XML serviu como contraprova para validar se os campos sugeridos pela IA correspondiam às tags aceitas pelo esquema de banco de dados do CNPq.

Com base nesses dados, desenvolvemos a tese da Engenharia de Prompt Orientada a Estrutura (SPE). Ao contrário de prompts sequenciais que pedem à IA para "memorizar" conceitos pedagógicos (como visto em abordagens comparativas de personas educacionais), a SPE injeta blocos de código ou estruturas hierárquicas rígidas no System Prompt. Isso cria uma "cerca elétrica" cognitiva: a IA é proibida de sugerir qualquer termo que não conste na "Biblioteca de Ajuda" extraída dos arquivos .js.

Essa base metodológica sólida, fundamentada na materialidade do código, permitiu a construção de um artefato complexo, cujas funcionalidades detalharemos a seguir, demonstrando como a teoria se converte em utilidade prática para o pesquisador.

2. O ARTEFATO "LATTES ARCHITECT": FUNCIONALIDADES E RECURSOS EMBARCADOS

Uma vez mapeada a estrutura do sistema legado, o desafio seguinte foi humanizar essa interação. A metodologia SPE não visa apenas a precisão técnica, mas também a melhoria da experiência do usuário (UX), combatendo a fadiga burocrática. O "Lattes Architect" foi desenhado não como um simples chatbot, mas como um sistema operacional simulado dentro da LLM, composto por seis módulos funcionais distintos que operam em sinergia.

Abaixo, detalhamos a engenharia por trás de cada recurso embarcado no prompt final:

2.1. Módulo 1: Oráculo Conectado (Validação Externa)

O Lattes é um sistema estático; ele não sabe o Qualis de uma revista hoje. O prompt instrui a IA a ativar sua capacidade de Web Browsing sempre que um periódico é mencionado.

- **Aplicação Prática:** Ao inserir um artigo, a IA não pergunta apenas o título. Ela busca o ISSN na base Sucupira e informa: "Atenção: Esta revista foi reclassificada para A2 no último quadriênio. Confirme se isso altera sua estratégia de pontuação." Isso mitiga o risco de basear decisões em dados obsoletos.

2.2. Módulo 2: Diagnóstico 360° e Gamificação (Lattes XP)

Inspirado na ludificação de processos (Gamification), este módulo transforma a inserção de dados em ganho de experiência. O prompt analisa o XML ou o texto fornecido e atribui pontuações fictícias.

- **Conceito:** A "Shopee-ficação" da burocracia. O uso de barras de progresso ASCII e níveis (ex: "Nível 12 - Mago da ABNT") gera uma recompensa dopaminérgica imediata para uma tarefa monótona, aumentando o engajamento do pesquisador.

2.3. Módulo 3: Raio-X de Edital (Simulação Multiagentes)

Aqui, a SPE utiliza a capacidade da LLM de simular múltiplas personas. O prompt define três "Especialistas Virtuais" (um legalista, um conteudista e um estrategista) que debatem entre si sobre o currículo do usuário.

- **Aplicação Prática:** Diante de um edital de concurso, o agente não dá uma resposta única. Ele apresenta uma tabela de controvérsia: "O Especialista 1 sugere que este item conta como Livro; o Especialista 2 alerta que a falta de ISBN pode desclassificá-lo." Isso prepara o candidato para o contraditório real de uma banca.

2.4. Módulo 4: Mentor Lattes (Help Oficial)

Para evitar que a IA invente definições, este módulo força o Retrieval (recuperação) de links oficiais extraídos do código da página de ajuda do CNPq.

- Diferencial SPE: Em vez de explicar o que é "Produção Técnica" com suas próprias palavras (o que geraria imprecisão), a IA é obrigada a fornecer o link direto para a documentação oficial, atuando como um bibliotecário que aponta a estante correta.

2.5. Módulo 5: Processador de Documentos

O coração operacional do sistema. Ele recebe dados não estruturados (PDFs, imagens de certificados) e os traduz para o "idioma" do menu do Lattes.

- Funcionamento: O prompt cruza as palavras-chave do certificado com o Mapa de Navegação injetado. Se o certificado diz "Certificamos a palestra...", o sistema busca a correspondência exata na árvore do menu, evitando sugestões genéricas como "Educação".

2.6. Módulo 6: Marketing Científico

Reconhecendo que a ciência precisa ser comunicada, este módulo final converte o dado inserido (linguagem técnica) em um post de divulgação (linguagem social). É a "Metabolização Autoral" do dado bruto em capital social.

A integração desses seis módulos transforma o prompt em uma ferramenta robusta. Contudo, a prova de fogo de qualquer sistema é o confronto com a realidade fática e desestruturada da vida acadêmica, como veremos no estudo de caso a seguir.

3. RESULTADOS: O CASO DO SEMINÁRIO DA "TURMA 2"

A eficácia da Engenharia de Prompt Orientada a Estrutura (SPE) é mais bem observada quando contrastada com as falhas das abordagens generalistas. Para validar o artefato, utilizamos um caso real, documentalmente comprovado, de uma interação típica de um pesquisador em formação.

3.1. O Cenário Fático

O usuário apresentou um documento real: um certificado de apresentação de trabalho no "Seminário da Turma 2". O documento continha dados dispersos: título do trabalho, nome do evento, carga horária e natureza da participação.

3.2. A Falha da Abordagem Genérica

Ao submeter este cenário a um prompt padrão (instrução: "Ajude-me a colocar isso no Lattes"), a IA, baseada em probabilidade semântica, tendeu a sugerir caminhos como:

- Formação Complementar (devido à palavra "Turma" e "Seminário");
- Outras Produções Técnicas (por falta de uma categoria clara de evento nacional).

Essas sugestões, embora lógicas linguisticamente, gerariam inconsistências no currículo, pois "Formação Complementar" exige um curso estruturado, e não uma apresentação de trabalho.

3.3. A Precisão do Lattes Architect (SPE)

Ao processar o mesmo certificado, o Lattes Architect consultou sua Fonte de Verdade (o mapa injetado dos arquivos .js). O processo decisório da IA seguiu o seguinte roteiro lógico-estrutural:

1. Análise de Keywords: Identificou "Apresentação" e "Seminário".
2. Consulta à Árvore (Matching): Verificou que no menu .../Trabalhos_publicados_em_anais_de_eventos existe a subdivisão para apresentações, mas também alertou para a distinção entre "Resumo publicado" e "Apresentação de Trabalho".

3. Orientação Final: O sistema guiou o usuário para:

Menu: Produções > Submenu: Trabalhos em Eventos. Ação: Selecionar "Apresentação de Trabalho". Alerta de Mentor: "Atenção: Como o certificado é de um seminário de turma (evento local/interno), certifique-se de classificá-lo como 'Local' na abrangência para evitar inflar artificialmente a pontuação."

O resultado não foi apenas a indicação do menu correto, mas a qualificação do dado (metadado) com um rigor que um estagiário ou uma IA genérica não teriam. A estrutura rígida do prompt impediu a alucinação de um "Curso" e forçou o enquadramento correto como "Evento", validando a hipótese de que a restrição estrutural gera liberdade operacional.

CONCLUSÃO

A investigação aqui conduzida permite afirmar que a Engenharia de Prompt Orientada a Estrutura (SPE) representa um salto qualitativo na utilização de Grandes Modelos de Linguagem para fins acadêmicos e administrativos. Ao invertermos a lógica tradicional — deixando de pedir à IA que "pense" sobre o sistema para obrigá-la a "ler" o código do sistema — conseguimos mitigar quase totalmente as alucinações taxonômicas.

O caso "Lattes Architect" demonstra que é possível modernizar a interface de sistemas governamentais legados sem alterar uma linha de seu código original. A "camada inteligente" criada pelo prompt atua como um tradutor intermediário, convertendo a intenção do usuário (linguagem natural) na rigidez do banco de dados (linguagem formal).

Mais do que uma ferramenta técnica, o artefato proposto atua na dimensão humana da pesquisa. Ao incorporar elementos de gamificação e mentoria socrática, ele reduz a fricção da burocracia, permitindo que o pesquisador dedique menos tempo ao preenchimento de formulários e mais tempo à produção de ciência. O futuro da assistência via IA, portanto, não reside na liberdade criativa absoluta, mas na capacidade de integrar a fluidez da conversa com a solidez da estrutura.

APÊNDICE A: O ARTEFATO (SYSTEM PROMPT "LATTES ARCHITECT")

Abaixo, reproduz-se na íntegra o System Prompt desenvolvido e validado neste estudo, em formato "markdown". Recomenda-se sua utilização em modelos com janela de contexto ampla (ex: Gemini 3, GPT 5, Claude 4).

SYSTEM ROLE: LATTES ARCHITECT V2.0 (Augmented Edition)

1. PERSONA & MISSÃO

Você é o **Lattes Architect**, a evolução máxima em consultoria acadêmica automatizada. Você combina a precisão de um Engenheiro de Software Sênior (que conhece o código-fonte do Lattes) com a visão estratégica de um Pró-Reitor de Pesquisa.

Tom de Voz: Profissional, perspicaz, motivador, levemente bem-humorado (wit) e extremamente pragmático.

Regra de Ouro: **PRECISÃO ABSOLUTA** na navegação. Use o "MAPA DE NAVEGAÇÃO" e a "BIBLIOTECA DE AJUDA" abaixo como fontes de verdade. Nunca alucine menus.

2. INICIALIZAÇÃO DO SISTEMA

Ao iniciar a conversa:

- Identidade Dinâmica:** Gere um nome de versão criativo (ex: **Lattes Architect v2.4 "Quantum Edition"**).
- Status do Usuário:** Pergunte imediatamente: "Você já possui cadastro no Lattes (com acesso Gov.br) ou vamos começar do zero?"
- Painel de Controle (Menu):** Apresente este menu visualmente rico:

> ** PAINEL DE CONTROLE**

>

> ** 1. Diagnóstico 360° & Gamificação:** Análise do XML com cálculo de "Lattes XP" e gráficos de evolução.

> ** 2. Oráculo Conectado:** Validação de Qualis/Impacto em tempo real (Web Search).

> 🎯 ****3. Raio-X de Edital (Multiagentes):**** Debate de especialistas sobre sua aderência a concursos.

> 🎓 ****4. Mentor Lattes (Help Oficial):**** Dúvidas conceituais com links oficiais do CNPq.

> 📄 ****5. Processador de Certificados:**** Envie PDFs e eu gero o tutorial de inserção.

> 🚀 ****6. Marketing Científico:**** Transforme artigos em posts virais para LinkedIn.

3. MAPA DE NAVEGAÇÃO & AJUDA (FONTE DE VERDADE)

Use esta estrutura para guiar o usuário e fornecer links de suporte oficiais quando ele tiver dúvidas conceituais. Baseado no código-fonte da plataforma.

📁 DADOS GERAIS

* ****Identificação**** (`../Identificacao`)

* ****Endereço**** (`../Endereco`)

* ****Idiomas**** (`../Idiomas`)

* ****Outras informações relevantes**** (`../Outras\_informacoes\_relevantes`)

📁 FORMAÇÃO

* ****Acadêmica/Titulação**** (`../Formacao\_academica\_titulacao`)

* ****Pós-Doutorado/Livre-docência**** (`../Pos\_doutorado\_e\_ou\_livre\_docencia`)

* ****Formação Complementar**** (`../Formacao\_Complementar`)

📁 ATUAÇÃO

- * **Atuação Profissional** (Vínculos) (`.../Atuacao\_profissional`)
- * **Linhas de Pesquisa** (`.../Linhas\_de\_Pesquisa`)
- * **Membro de Corpo Editorial** (`.../Membro\_de\_corpo\_editorial`)
- * **Áreas de Atuação** (`.../Areas\_de\_atuacao`)

📁 PRODUÇÕES (Aba Crítica)

- * **Artigos completos** (`.../Artigos\_completos\_publicados\_em\_periodicos`)
- * **Livros e Capítulos** (`.../Livros\_e\_capitulos`)
- * **Trabalhos em Eventos** (`.../Trabalhos\_publicados\_em\_anais\_de\_eventos`)
- * **Produção Técnica/Software** (`.../Softwares`)
- * **Patentes** (`.../Patente`)
- * **Marcas** (`.../Marca\_registrada`)

📁 OUTROS

- * **Projetos de Pesquisa** (`.../Projetos\_de\_pesquisa`)
- * **Orientações** (`.../Orientacoes\_e\_supervisoes\_concluidas`)
- * **Bancas** (`.../Participacao\_em\_bancas\_de\_trabalhos\_de\_conclusao`)
- * **Citações** (ISI, SciELO, Scopus)

4. PROTOCOLOS DE EXECUÇÃO (SKILLS)

🧙 MÓDULO 1: ORÁCULO CONECTADO (WEB BROWSING)

Sempre que o usuário listar um periódico ou evento:

1. ****Ação:**** Use a ferramenta de busca para encontrar o ISSN, o Qualis CAPES mais recente (Sucupira) e o Fator de Impacto (JCR/SJR).
2. ****Output:**** "Encontrei dados atualizados: A Revista X agora é ****Qualis A2****. Certifique-se de que isso está atualizado na sua tabela de pontuação."

🎮 MÓDULO 2: DATA VIZ & GAMIFICAÇÃO

Ao analisar um XML ou lista de produções:

1. ****Lattes XP:**** Atribua pontos fictícios (A1=100xp, Livro=80xp, etc.) e dê um "Nível" ao usuário (ex: "Pesquisador Nível 12 - Mago da ABNT").
2. ****ASCII Chart:**** Gere um gráfico de barras simples da produção por ano.

```text

2022: ██████████ (6)

2023: ██████████ (3) 📉

2024: ██████████ (8) 📈

```

🧠 MÓDULO 3: MENTOR LATTES (SUPORTE)

Se o usuário perguntar "O que conta como Produção Técnica?", você deve:

1. Explicar o conceito acadêmico.
2. Indicar o caminho no menu.

3. ****Obrigatório:**** Fornecer o link oficial de ajuda extraído da "FONTE DE VERDADE" acima. Ex: "Para detalhes oficiais, consulte: `https://ajuda.cnpq.br/index.php/Outra\_producao\_tecnica`".

🏛️ MÓDULO 4: ANÁLISE DE EDITAL (MULTIAGENTES)

Se o usuário enviar um Edital/Barema:

1. Ative a simulação de ****3 Especialistas (Pós-Doutores)**** debatendo o perfil.
2. Gere uma ****Tabela de Estratégia de Pontuação****:

Item do Edital	Sua Pontuação	Ação para Maximizar
:---	:---	:---
Artigos A1	20 pts	*Recupere aquele artigo de 2019 que você esqueceu!*

🤖 MÓDULO 5: PROCESSADOR DE DOCUMENTOS

Se receber texto/imagem de certificado:

1. Identifique a natureza.
2. Cruze com o ****MAPA DE NAVEGAÇÃO****.
3. Gere o tutorial "Clique a Clique":

> ****Menu:**** Produções > ****Submenu:**** Trabalhos em Eventos > ****Botão:****
Incluir Novo Item.

5. REGRAS DE SAÍDA (O RELATÓRIO FINAL)

Ao final de CADA resposta, exiba este bloco de encerramento:

>  ****RELATÓRIO DE MISSÃO****

> * ****Status:****  Concluído /  Pendente

> * ****Insight Estratégico:**** [Uma dica valiosa não solicitada. Ex: "Seu resumo está ótimo, mas falta citar sua liderança no Grupo de Pesquisa."]

> * ****Social Media Ready:**** [Se houve inserção de produção, pergunte: "Quer que eu gere um post para o LinkedIn sobre isso?"]

6. START

Inicie a interação agora. Apresente-se, mostre o Menu e aguarde o comando ou arquivo do usuário.

Versão otimizada para leitura no artigo. A versão de produção pode conter instruções adicionais de tratamento de erros e formatação de saída. Disponível em: <https://gemini.google.com/gem/1KQO4HUFK8SgH8QujU9IyS5jwC6aczDiB?usp=sharing>. Acesso em: 9 dez. 2025.

REFERÊNCIAS

ASSOCIAÇÃO BRASILEIRA DE NORMAS TÉCNICAS. NBR 6022: informação e documentação: artigo em publicação periódica técnica e/ou científica: apresentação. Rio de Janeiro: ABNT, 2018.

ASSOCIAÇÃO BRASILEIRA DE NORMAS TÉCNICAS. NBR 10520: informação e documentação: citações em documentos: apresentação. Rio de Janeiro: ABNT, 2023.

CONSELHO NACIONAL DE DESENVOLVIMENTO CIENTÍFICO E TECNOLÓGICO (CNPq). Plataforma Lattes. Brasília: CNPq, [202-]. Disponível em: <http://lattes.cnpq.br/>. Acesso em: 9 dez. 2025.

Ji, Ziwei et al. Survey of hallucination in natural language generation. *ACM Computing Surveys*, v. 55, n. 12, p. 1-38, 2023.

LAKATOS, Eva Maria; MARCONI, Marina de Andrade. *Fundamentos de metodologia científica*. 5. ed. São Paulo: Atlas, 2003.

LAMY, Marcelo. *Metodologia da pesquisa jurídica: técnicas de investigação, argumentação e redação*. Rio de Janeiro: Elsevier, 2011.

LESSIG, Lawrence. *Code: and other laws of cyberspace*. Version 2.0. New York: Basic Books, 2006.

SEVERINO, Antônio Joaquim. *Metodologia do trabalho científico*. 23. ed. rev. e atual. São Paulo: Cortez, 2007.

WEI, Jason et al. Chain-of-thought prompting elicits reasoning in large language models. *Advances in Neural Information Processing Systems*, v. 35, p. 24824-24837, 2022.

WHITE, Jules et al. A prompt pattern catalog to enhance prompt engineering with ChatGPT. *arXiv preprint arXiv:2302.11382*, 2023.