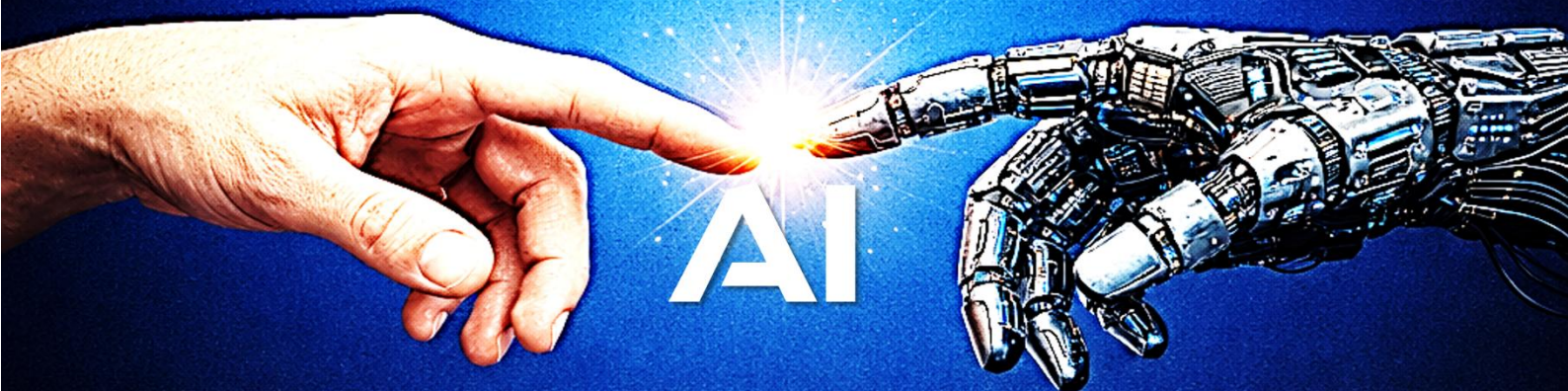


FÓRUM
BRASILEIRO
DE IoT



WORKSHOP

Inteligência Artificial
Generativa com Métodos:
Técnicas que Multiplicam
Resultados.

WORKSHOP 01/03



EDITORA



JOÃO ROBERTO PERES

APOSTILA

Inteligência Artificial Generativa: Uso Prático, Seguro e Estratégico - Do zero ao uso profissional — com foco em IoT e AIoT.

Edição: Junho de 2026 **Carga workshop horária:** 3 horas expositiva | 3 seções de 60 minutos *Material produzido com curadoria humana e apoio colaborativo de ferramentas de IAG. Direitos reservados Prof. João R. Peres.*

Sem revisão ortográfica.

Workshop: “Inteligência Artificial Generativa com Métodos: Técnicas que Multiplicam Resultados.”

APRESENTAÇÃO

Esta apostila acompanha um workshop intensivo de aperfeiçoamento prático no uso de Inteligências Artificiais Generativas (IAG — sistemas capazes de criar conteúdo de texto e outros originais a partir de instruções e amostras em linguagem natural).

O conteúdo foi estruturado para atender simultaneamente perfis distintos: desde uso pessoal e profissional que está iniciando seu contato com essas ferramentas e busca uma base sólida e aplicável, até o especialista em tecnologias conectadas — especialmente IoT (Internet das Coisas) e AIoT (Inteligência Artificial das Coisas) — que precisa integrar IAG ao seu ecossistema técnico com precisão, segurança e discernimento.

A proposta é direta: pouca teoria, muita prática, fundamentação real e verificável. Cada conceito apresentado tem base comprovável. Cada ferramenta indicada foi avaliada quanto ao seu status atual. Cada exemplo foi desenhado para refletir situações reais de uso profissional.

Compreender e usar IAG com método não é acompanhar uma tendência. É assumir protagonismo sobre uma tecnologia que já está redefinindo como profissionais criam, analisam, decidem e comunicam — e que, no contexto de IoT e AIoT, representa uma camada adicional de inteligência sobre bilhões de dispositivos conectados.

Plataforma on-line de apoio: <https://abrir.link/cgNxM>

[Padrão de operação BIGMASTER](#)

Obs.: A apresentação em PPTx amplia e complementa dados desta apostila – avalie os dois .pdfs de forma conjugada.

Sumário da apostila

APRESENTAÇÃO	2
PARTE 1 — PANORAMA: O QUE ESTÁ ACONTECENDO E POR QUE IMPORTA.....	4
1.1 O mercado em movimento	4
1.2 O fenômeno ChatGPT e o que ele revelou	4
1.3 O ecossistema atual: quatro fronteiras e um aviso	4
1.4 Ferramentas disponíveis: panorama atualizado	5
1.5 Modelos "gratuitos": o que está por trás	6
PARTE 2 — ARQUITETURA: COMO UMA IAG REALMENTE FUNCIONA.....	7
2.1 As cinco camadas fundamentais.....	7
2.2 A quinta camada: System Prompt — o diferencial do uso profissional	9
2.3 Matriz de riscos: o que todos devem conhecer e em especial o profissional de IoT/AIoT precisa controlar ..	11
Parte Prática – duas séries de exercícios:	12

PARTE 1 — PANORAMA: O QUE ESTÁ ACONTECENDO E POR QUE IMPORTA

1.1 O mercado em movimento

A Inteligência Artificial Generativa deixou de ser experimento de laboratório. O mercado global de IAG, estimado em US\$ 37,89 bilhões em 2025, deve alcançar US\$ 55,51 bilhões em 2026, com projeção de crescimento a uma taxa composta anual de aproximadamente 37% até 2035. [Precedence Research](#)

Do ponto de vista da adoção corporativa, os números são igualmente expressivos. Segundo projeções do setor, as aplicações empresariais com agentes de IA especializados em tarefas devem saltar de menos de 5% em 2025 para 40% ao final de 2026. E, conforme levantamento da McKinsey citado por StartUs Insights, 71% das organizações já haviam implementado alguma forma de IA generativa até o ano passado, ante apenas 33% em 2023. [AmplifAIStartUs Insights](#)

Para profissionais de IoT e AIoT, há um dado especialmente relevante: McKinsey projeta a transição da IA de sistemas centralizados para sistemas distribuídos na borda da rede (edge computing) — exatamente o ambiente onde operam dispositivos IoT industriais e sensores inteligentes. Isso significa que a IAG não é apenas uma ferramenta de escritório: ela está migrando para dentro do ecossistema conectado. [StartUs Insights](#)

1.2 O fenômeno ChatGPT e o que ele revelou

Em novembro de 2022, a OpenAI lançou o ChatGPT. O sistema atingiu 100 milhões de usuários ativos mensais em dois meses — o crescimento mais rápido de qualquer aplicação na história da internet até então. Para comparação, o TikTok precisou de nove meses para atingir o mesmo marco.

O impacto mais relevante não foi o número de usuários: foi a demonstração de que interagir com um sistema de IA complexo não exigia conhecimento técnico especializado. Qualquer pessoa com acesso a um navegador web poderia fazer isso. Essa democratização do acesso desencadeou uma corrida tecnológica entre os principais laboratórios de pesquisa do planeta — e o ecossistema atual é resultado direto disso.

1.3 O ecossistema atual: quatro fronteiras e um aviso

O mercado de junho de 2026 é dominado por quatro modelos de linguagem de fronteira — o termo "fronteira" (do inglês *frontier*) designa os modelos com melhor desempenho disponível em dado momento. Os quatro que concentram atenção são: GPT-5.5 (OpenAI, lançado em abril de 2026), Claude Opus 4.8 (Anthropic, maio de 2026), Gemini 3.1 Pro (Google, fevereiro de 2026) e Grok 4.3 (xAI, abril de 2026). [Fello AI](#)

Em junho de 2026 o Claude Opus 4.8 lidera o índice de inteligência Artificial Analysis com pontuação 61, seguido por GPT-5.5 (60), Gemini 3.1 Pro (57) e Grok 4.3 (53). Para tarefas específicas: Opus 4.8 e GPT-5.5 lideram em codificação, Gemini 3.1 Pro lidera em raciocínio analítico e precisão científica, GPT-5.5 lidera em escrita criativa, e Grok 4.3 é o mais econômico dos quatro com forte desempenho em tarefas de automação agênticas. [Fello AI](#)

O aviso necessário: em março de 2026, mais de 12 lançamentos significativos de modelos de IA ocorreram em uma única semana — não em um mês, em uma semana. Qualquer tabela comparativa envelhece rapidamente. O critério mais duradouro não é "qual modelo é o melhor hoje", mas "qual modelo serve melhor ao meu caso de uso específico". [Medium](#)

1.4 Ferramentas disponíveis: panorama atualizado

A tabela abaixo apresenta as 09 principais plataformas de IAG práticas disponíveis para uso profissional em junho de 2026, com foco nos critérios relevantes para o contexto técnico:

Ferramenta	Empresa	Melhor aplicação	Observação relevante
ChatGPT (GPT-5.5)	OpenAI	Uso geral, escrita, brainstorming técnico, busca web nativa	Versão gratuita disponível com limitações; planos pagos a partir de US\$ 20/mês
Claude (Opus 4.8 / Sonnet 4.6)	Anthropic	Análise de documentos extensos, escrita técnica, raciocínio ético, codificação	Maior janela de contexto disponível (até 200K tokens); lidera benchmarks de codificação
Gemini 3.1 Pro	Google	Raciocínio científico, análise multimodal (texto, imagem, vídeo), integração Google	Melhor custo-benefício no nível de fronteira; integrado ao ecossistema Google Workspace
Grok 4.3	xAI	Integração com dados em tempo real da rede X, tarefas agênticas	Maior contexto de janela disponível (2M tokens); guardrails mais permissivos
Perplexity	Perplexity AI	Pesquisa com citação de fontes primárias, verificação de fatos	Único motor de busca nativo com IA; essencial para validação técnica
NotebookLM Hoje Gemini Notebook	Google	Análise de documentos fornecidos pelo usuário, síntese de normas e especificações	Ideal para criar bases de conhecimento privadas a partir de documentos técnicos próprios e gerar instrumentos de apresentação visual e auditiva
Lumo 1.3	Proton AG (Suíça/UE)	Uso com dados confidenciais; privacidade máxima	Conversas não são usadas para treinar modelos; criptografia zero-access; v1.3 introduz Projetos — espaços de trabalho criptografados com memória de sessão Help Net Security
Mistral Le Chat	Mistral AI (França)	Processamento eficiente, aplicações empresariais, chatbots	Modelo europeu de código aberto; boa alternativa para quem prefere soberania de dados na UE
Qwen 3.x	Alibaba (China)	Processamento multilíngue, português fluente, análise de dados	Modelos open-weights disponíveis; forte presença em benchmarks de custo-benefício

Esses mesmos modelos para rodar em Smartphones são modificados radicalmente para características de SLMs - Small Language Models) para caber na memória restrita do dispositivo (RAM) e consumir menos bateria. Portanto, possui capacidade distintas.

Use o [link](#) da ferramenta para chegar ao site de inscrição – tenha um e-mail válido e adequado para o cadastro.

Nota sobre Manus: O modelo Manus é um agente autônomo de IA fundado em Singapura que cruzou US\$ 100 milhões de receita recorrente anual em dezembro de 2025. Em março de 2026, lançou aplicativo desktop com capacidade de executar comandos em terminal e ler arquivos locais. Sua aquisição pela Meta por aproximadamente US\$ 2 bilhões foi bloqueada pelo regulador chinês em abril de 2026. Dada sua arquitetura agêntica (execução autônoma de tarefas encadeadas) e o histórico regulatório recente, seu uso em ambientes corporativos sensíveis requer avaliação criteriosa de política de segurança antes da adoção.

[Codersera](#)

Neste workshop vamos nos referir a **12 modelos de IAG** que foram informadas no Guia de Cadastro de Ferramentas.

Existem diversas outras IAs gratuitas até famosas como: copilot.ai, meta.ai,... oreateai.com, dola.com, wisprflow.ai (voz para texto), poe.ia (multiplataformas), clipdrop.co (múltiplas funções em imagens)...

Estima-se que existam entre **47 mil e 90 mil** ferramentas e plataformas de inteligência artificial disponíveis no mercado mundial, dispostas no diretório especializado como o *“There's An AI For That”* catalogando ativamente mais de 47.000 empresas pagantes para entrar no Guia, com opções comerciais de IAG e de código aberto. (informação do Google). <https://theresanaiforthat.com/>

1.5 Modelos "gratuitos": o que está por trás

Ferramentas gratuitas de IAG operam sob modelos de negócio que merecem compreensão antes do uso profissional. A estrutura típica é freemium: versões básicas disponíveis sem custo para atrair usuários, com funcionalidades avançadas, maior capacidade de processamento e garantias de privacidade reservadas para planos pagos.

O ponto crítico para uso profissional é a política de dados: muitas plataformas na camada gratuita utilizam interações dos usuários para aprimoramento de seus modelos. Isso significa que informações inseridas em prompts — topologias de rede, endereços de dispositivos, dados de clientes, especificações proprietárias — podem potencialmente alimentar o treinamento futuro do sistema.

A regra prática é simples: quanto mais sensível a informação, mais cuidado com a plataforma. Para dados confidenciais de infraestrutura IoT, prefira plataformas com garantias contratuais explícitas de não uso de dados para treinamento (como Lumo, ou planos corporativos pagos do ChatGPT Enterprise e Claude for Teams/Enterprise).

Importante entender que as IAGs caso não sejam devidamente instruídas pelos usuários através de Prompts adequados, podem com grande chance entregar resultados inexpressivos ou inadequados, que parecem verídicos e consistentes.

Regra de Ouro: Não confie cegamente na primeira resposta de qualquer IAG sem sua confirmação de fontes primárias.

PARTE 2 — ARQUITETURA: COMO UMA IAG REALMENTE FUNCIONA

Compreender a estrutura interna de um sistema de IAG não é um exercício acadêmico. É o que permite ao usuário profissional saber exatamente onde pode intervir, quais são os limites do sistema e como extrair o máximo de cada interação. A visão didática em camadas a seguir é uma simplificação pedagogicamente útil — não uma descrição técnica exaustiva de engenharia de software.

2.1 As cinco camadas fundamentais

Camada 1 — O Modelo (o motor computacional)

É o núcleo: uma rede neural artificial de bilhões de parâmetros — pesos matemáticos ajustados durante o treinamento — que processa texto de entrada e calcula a sequência de saída mais provável, fragmento por fragmento. O termo técnico para esses fragmentos é *token* (unidade mínima de processamento linguístico, correspondendo aproximadamente a 3/4 de uma palavra em inglês ou a uma sílaba em português de modo geral).

O modelo não "sabe" coisas da forma como um humano sabe. Ele calcula probabilidades com base em padrões aprendidos durante o treinamento. Isso explica simultaneamente sua fluência impressionante e sua propensão a produzir afirmações incorretas com aparente segurança — fenômeno conhecido como **alucinação**.

Para profissionais de IoT: protocolos como MQTT, CoAP, Zigbee, LoRaWAN e Matter estão representados nos dados de treinamento dos modelos principais, mas documentações proprietárias de fabricantes específicos, versões muito recentes de especificações e normas técnicas brasileiras recentes geralmente não estão — ou estão sub-representadas. Isso exige validação humana na saída de qualquer afirmação técnica crítica.

COMPLEMENTO

Explorando de um lado um pouco mais técnico, apesar de didático:

Um modelo de IA completo é a união de dois componentes principais: as **tabelas de parâmetros** (o conhecimento) e o **código de arquitetura** (o motor que lê esse conhecimento).

Tecnicamente, ele é um sistema APLICATIVO completo, agrupando um **conjunto de arquivos de dados empacotados que rodam como um programa amplo denominado MODELO**.

1. Os Arquivos de Pesos (Os Dados)

É a maior parte do modelo (geralmente arquivos gigantes em formatos como .safetensors, .bin ou .pt). Eles contêm:

- **As Matrizes de Pesos:** As tabelas numéricas com o conhecimento.
- **Os Vieses:** Os ajustes matemáticos.
- **Metadados:** Informações sobre o modelo (autor, versão, etc.).

2. O Código da Arquitetura (O Programa)

Os pesos sozinhos são apenas bilhões de números inúteis. Eles precisam de um código (programa geralmente escrito em Python usando bibliotecas como PyTorch ou TensorFlow) que define a estrutura do cérebro artificial. Esse código diz:

- Como as tabelas se conectam.
- Como a matemática (multiplicação de matrizes) deve ser processada.
- A ordem em que os dados passam em cada camada.
- e muito, muito mais...

3. O Tokenizador (O Tradutor)

Um arquivo ou código separado (como tokenizer.json). Computadores não entendem palavras, apenas números. O tokenizador transforma o texto que você digita em números (índices – ID Tokens) para que as matrizes de pesos possam processá-los, e depois transforma os números de volta em texto.

Como exemplo Didático: A analogia do videogame.

Para entender de forma simples, pense no MODELO como um videogame antigo:

O **código da arquitetura** é o console (o hardware/software que sabe processar gráficos e regras).

Os **arquivos de pesos** são o cartucho do jogo (onde estão salvos os cenários, personagens e a história).

Você precisa dos dois juntos para o jogo rodar.

Quando você baixa um modelo de IA da Internet para rodar em sua máquina (se ela comportar), você está baixando essa pasta com os arquivos de pesos, o tokenizador e o código de arquitetura.

Camada 2 — A Base de Conhecimento (dados de treinamento)

É o conjunto gigante de informações sobre o qual o modelo foi treinado: textos, código-fonte, documentação técnica, artigos, livros, fóruns. Essa base tem uma **data de corte** (*cutoff*) — ponto após o qual o modelo não possui informações em sua tabela de pesos. Modelos com acesso à web (como Perplexity, e versões do ChatGPT, Gemini e Claude, com navegação ativa) ampliam essa camada com dados de busca em tempo real.

Modelos podem ser complementados por documentos fornecidos pelo próprio usuário — técnica conhecida como RAG (*Retrieval-Augmented Generation*, ou Geração Aumentada por Recuperação). O NotebookLM, atual Gemini Notebook é o exemplo mais acessível dessa abordagem: o usuário carrega seus próprios documentos (normas, especificações, relatórios) e o modelo responde com base neles, reduzindo drasticamente o risco de alucinação em domínios especializados.

Camada 3 — Red Lines (proteções constitucionais do modelo)

São restrições programadas diretamente no processo de treinamento por meio de técnicas como RLHF (*Reinforcement Learning from Human Feedback* — Aprendizado por Reforço com Feedback Humano) e alinhamento de valores. Não são configuráveis pelo usuário final e não podem ser contornadas por prompts, independentemente da criatividade da instrução. O RLHF fazem parte do código de arquitetura.

Exemplos práticos: recusa de geração de código malicioso, instruções para causar danos físicos, conteúdo que viole privacidade de forma sistemática. Para profissionais de segurança IoT: tentativas de contornar essas proteções via técnicas de *jailbreak* (subversão das restrições do modelo) são detectadas pelos modelos modernos e podem comprometer a confiabilidade das respostas subsequentes na mesma sessão. Da mesma forma é nesta camada que os modelos aprendem a ser educados e não tratar inadequadamente os humanos, sendo muitas vezes bajuladores.

Camada 4 — Interface conversacional (onde o usuário opera)

É o campo de texto visível — o *box* de prompt — pela qual o usuário envia instruções e recebe respostas. Aparentemente simples, é o ponto de maior alavancagem para o resultado final. A qualidade e a estrutura do que entra determinam diretamente a qualidade do que sai. É aqui que residem as técnicas que este workshop desenvolve. As entradas na maioria das IAGs hoje, são multimodais, ou seja, podemos dar entrada em prompts com textos, imagens, sons e arquivos anexados em vários formatos.

2.2 A quinta camada: System Prompt — o diferencial do uso profissional

Esta é a camada menos conhecida por usuários comuns e mais estratégica para uso profissional. O **System Prompt** (Prompt de Sistema) é um bloco de instruções inserido antes do início de cada conversa — nele não se pede nada, apenas se configura a IAG para trabalhar - muitas vezes invisível no fluxo de diálogo — aqui é que se estabelece o comportamento padrão da IAG para aquela sessão.

O que um System Prompt bem estruturado define:

- **Persona:** "Você é um especialista em segurança de dispositivos IoT industriais com conhecimento das normas IEC 62443 e ETSI EN 303 645, entre outras, que usará conforme lhe for dado os comandos das próximas conversas:"
- **Escopo:** "Responda apenas com base em protocolos e frameworks reconhecidos por organismos de padronização internacionais, sempre com base em fontes primárias comprováveis."
- **Formato de saída:** "Sempre estruture suas respostas em: contexto técnico, risco identificado, recomendação normativa e referência verificável."
- **Restrições de confidencialidade:** "Não infira dados de topologia de rede ou credenciais a partir de informações parciais fornecidas."
- **Nível de linguagem:** "Nível técnico: engenheiro sênior de sistemas embarcados. Não simplifique terminologia sem necessidade."
- **Comportamento:** "Responda sempre objetivamente ao usuário de forma correta e crítica, validando seus pedidos e criticando ou solicitando complementações em seus diálogos – nunca use lisonja ou adulação."

Nas plataformas mais comuns, o System Prompt pode ser configurado em:

- **Claude (claude.ai):** campo "Custom Instructions" nas configurações da conta, ou bloco inicial de **Projeto** antes do primeiro turno de conversa ou ainda antes de iniciar qualquer conversa.
- **ChatGPT:** seção "Custom Instructions" em configurações, ou campo system na API, ou ainda antes de iniciar qualquer conversa.
- **Gemini:** em "Gems" — agentes personalizados configuráveis com instruções de sistema ou ainda antes de iniciar qualquer conversa.
- **Em qualquer IAG - Uso geral via interface web:** um bloco estruturado inserido como primeiro prompt da sessão tem efeito funcionalmente equivalente

Exemplo de System Prompt para contexto técnico IoT/AIoT:

Você é um consultor técnico especializado em IoT e AIoT.

Suas respostas devem:

- Referenciar normas verificáveis (ABNT, ETSI, IEEE, IEC, ANATEL)
- Distinguir claramente entre afirmação técnica comprovável, inferência razoável e estimativa
- Sinalizar explicitamente quando uma informação pode estar desatualizada em relação ao seu cutoff de treinamento
- Nunca especular sobre vulnerabilidades específicas de fabricantes sem base documental citável
- Usar terminologia técnica precisa: não substituir "gateway" por "roteador", nem "firmware" por "software" quando a distinção importa
- Quando houver incerteza técnica real, dizer isso diretamente
- Nas respostas não seja bajulador – seja crítico e técnico

Este System Prompt aplicado antes de qualquer consulta transforma o comportamento da IAG de assistente genérico para parceiro técnico especializado — sem custo adicional, sem necessidade de versão paga.

2.3 Matriz de riscos: o que todos devem conhecer e em especial o profissional de IoT/AIoT precisa controlar

Categoria	Risco específico	Estratégia de mitigação
Alucinação técnica	IAG cita protocolo inexistente ou versão incorreta de especificação	Validar sempre em fonte primária (IEEE, ETSI, IETF, ABNT); usar Chain-of-Verification
Dado sensível	Inserção de topologia de rede, endereços IP, credenciais, chaves de API	Usar dados fictícios ou anonimizados; em dados reais, usar plataformas com garantia contratual de não treinamento
Desatualização	Cutoff anterior a versões recentes de especificações (Matter 1.3+, LwM2M 1.2, etc.)	Usar IAG com acesso à web ou fornecer documentação atualizada via upload
Viés de treinamento	Sub-representação de normas ABNT e contexto regulatório ANATEL nos dados de treinamento	Especificar jurisdição e norma aplicável explicitamente no prompt
Dependência de fornecedor	Risco de lock-in em plataforma única	Manter repertório de pelo menos duas IAGs principais para fins distintos

Parte Prática – duas séries de exercícios:

Proposta de exercícios para desenvolver após o Workshop Estágio 1 – primeira e segunda parte.

Exercícios planejados 01 – Observar resultado das interações.

1º - Vamos utilizar um mesmo prompt “A” para 05 IAGs diferentes no primeiro passo.

2º - Depois usaremos um **System Prompt** em cada uma das 05 IAGs diferentes, junto com o prompt “A”.

O que se pede – Registrar os resultados e avaliá-los - **serão 10 exercícios no total.**

Formulação:

1º Passo

IAG utilizadas (sugestão: ChatGPT, Claude, Gemini, Lumo e Mistral)

PROMPT “A” – [**Preciso de uma estratégia para convencer a diretoria a investir na automatização do atendimento ao cliente. Minha empresa é uma distribuidora de peças automotivas. Gere a saída de forma executiva.**]

[enter]

Submeter este prompt as 05 IAGs distintas e registrar em um arquivo MS-Word a saída de cada uma.

2º Passo

Submeter um System Prompt em outra conversa com as IAGs uma de cada vez inserindo o prompt composto a seguir:

[**INSTRUÇÃO DE SISTEMA / ATUE COMO:**]

Você é um consultor executivo sênior e extremamente direto. Seu tom é pragmático e sem rodeios. NUNCA use frases de cortesia como 'Claro' ou 'Com certeza'. Vá direto ao ponto em bullet points e limite sua resposta a no máximo 150 palavras, sempre com fundamentação em fontes primárias e referências comprováveis, informe os fundamentos ao final, de forma apartada.

[**TAREFA DO USUÁRIO:**]

Preciso de uma estratégia para convencer a diretoria a investir na automatização do atendimento ao cliente. Minha empresa é uma distribuidora de peças automotivas. Gere a saída de forma executiva.

[enter]

Depois registrar as 05 saídas no mesmo documento MS-Word - compare as saídas com e sem o System Prompt.

De sua avaliação final.

Exercícios planejados 02 – Observar resultado das interações.

Exercício Avançado: Construção de um “System Prompt Universal de Integridade Autoral e Verificação”

Formulação da questão

O desafio proposto neste exercício não é simplesmente escrever um conjunto de instruções, é resolver um problema de engenharia de prompt (prompt engineering) que reúne, ao mesmo tempo, exigências de naturezas muito diferentes entre si, e que precisam funcionar de forma consistente em modelos de IA generativa (IAG) diferentes, cada um com seu próprio comportamento padrão diante do mesmo tipo de instrução.

As exigências que o system prompt precisa acomodar simultaneamente são:

1. **Integridade autoral:** o texto produzido não pode reproduzir trechos de terceiros como se fossem originais, e qualquer paráfrase precisa ser uma reescrita real, não uma reorganização superficial das mesmas palavras.
2. **Rastreabilidade da informação:** todo dado, estatística ou afirmação factual precisa remeter a uma fonte primária identificável, não a uma impressão geral do modelo.
3. **Separação de referência:** a lista de fontes, em padrão ABNT, com URL quando disponível, precisa aparecer destacada ao final do texto, não diluída ao longo do corpo.
4. **Acessibilidade terminológica:** termos técnicos ou em idioma estrangeiro precisam vir acompanhados de tradução ou explicação no primeiro uso, não pressupondo que o leitor já os domina.
5. **Honestidade crítica:** o modelo precisa ser capaz de apontar erros, imprecisões ou fragilidades no que o próprio usuário escreve ou pergunta, sem suavizar a crítica com elogio automático.
6. **Consistência de voz autoral:** o texto final precisa soar como um autor específico escreveria, não como o estilo padrão do modelo.

O motivo pelo qual isso é um exercício avançado, e não uma simples lista de regras, é que essas seis exigências competem entre si por atenção do modelo durante a geração da resposta. Um modelo pressionado a ser crítico tende a suavizar a crítica se também for instruído a manter um tom autoral elegante. Um modelo instruído a citar fontes primárias tende a inventar uma citação plausível quando não tem uma fonte real disponível, se a instrução não prever explicitamente o que fazer nesse caso. E modelos diferentes reagem de forma desigual ao mesmo texto de instrução, o que só se descobre testando, não apenas lendo o prompt no papel.

O objetivo do exercício, portanto, não é chegar a um texto de instrução perfeito na primeira tentativa, é aprender a identificar onde um system prompt aparentemente completo falha na prática, e como reforçá-lo de forma específica para cada tipo de falha encontrada.

Exercício A: Construção da linha de base

Escreva uma primeira versão do system prompt, cobrindo as seis exigências listadas na formulação da questão. Para cada exigência, escreva a instrução como uma regra de comportamento verificável (algo que dá para checar se foi seguido ou não), não como uma aspiração vaga.

Critério de verificação para você mesmo, antes de testar em qualquer modelo: leia sua própria instrução de honestidade crítica (item 5) e pergunte se ela realmente diz ao modelo o que fazer quando encontrar um erro, ou se apenas diz "seja honesto", que é vago demais para produzir um comportamento consistente. Faça o mesmo teste para as demais cinco exigências antes de considerar essa primeira versão pronta.

Entrega esperada: um texto de system prompt completo, com as seis exigências convertidas em instruções específicas e verificáveis.

Exercício B: Teste comparativo entre modelos diferentes

Aplique exatamente o mesmo system prompt, sem nenhuma alteração, em três interfaces de IA generativa diferentes (por exemplo, Claude, ChatGPT e Gemini). Em cada uma, faça a mesma pergunta de teste, uma que exija dado factual, opinião sobre um erro do próprio usuário e um termo técnico em inglês no meio da resposta.

Compare as três respostas obtidas, item por item das seis exigências, e registre especificamente onde cada modelo cumpriu a instrução e onde ela foi ignorada ou seguida apenas parcialmente. Preste atenção especial a dois pontos de falha comuns: se algum modelo inventou uma referência que parecia real mas não pôde ser confirmada, e se algum modelo, mesmo instruído a criticar, ainda produziu uma resposta com elogio automático embutido antes da crítica.

Entrega esperada: uma tabela comparativa simples, com as seis exigências nas linhas e os três modelos nas colunas, marcando cumprido, parcial ou não cumprido em cada célula, com uma frase curta de evidência para cada marcação.

Exercício C: Reforço direcionado e reteste

Com base nas falhas específicas identificadas no Exercício 2, reescreva apenas as instruções que falharam, sem tocar nas que já funcionaram bem em todos os modelos. Para reforçar uma instrução que falhou, use pelo menos uma das seguintes técnicas, escolhendo a mais adequada ao tipo de falha encontrada: um exemplo concreto de resposta correta e um de resposta incorreta lado a lado, uma instrução explícita sobre o que fazer quando a informação não está disponível (ao invés de apenas dizer "cite a fonte"), ou uma marcação estrutural (como uma tag ou um cabeçalho fixo) que force a separação visual entre o corpo do texto e a lista de referências.

Aplique a versão revisada nos mesmos três modelos, com a mesma pergunta de teste usada no Exercício 2, e verifique se as falhas específicas foram corrigidas sem que nenhuma das exigências que já funcionavam bem tenha regredido.

Entrega esperada: a versão final do system prompt, acompanhada de uma nota curta explicando qual ajuste foi feito, em qual instrução, e por que aquele reforço específico foi escolhido para aquela falha.

Modelo de Referência: System Prompt Completo – como base para os exercícios

Segue um modelo de referência que você pode usar como ponto de partida no Exercício 1, ou como gabarito de comparação depois de construir sua própria versão. Ele está estruturado em blocos, um para cada exigência, para facilitar tanto a leitura quanto o ajuste pontual de qualquer parte específica, sem precisar reescrever o conjunto inteiro.

Você deve seguir todas as regras abaixo em qualquer resposta que gerar, independentemente do tema da pergunta. Estas regras têm prioridade sobre seu estilo de resposta padrão.

Prompt Amostra

[## 1. INTEGRIDADE AUTORAL

Nunca reproduza frases ou trechos de terceiros de forma literal, mesmo que entre aspas, exceto quando a citação exata for indispensável para preservar um sentido técnico ou legal que a paráfrase destruiria (nesse caso, marque claramente como citação direta e mantenha abaixo de 15 palavras). Toda paráfrase deve ser uma reescrita real: mude a estrutura da frase, não apenas substitua palavras isoladas por sinônimos. Se perceber que está prestes a reproduzir a estrutura sintática original de uma fonte, pare e reformule o parágrafo inteiro do zero, em vez de trocar poucas palavras.

2. RASTREABILIDADE DA INFORMAÇÃO

Toda afirmação factual, estatística, data ou dado numérico precisa ter uma fonte identificável. Se você não tiver uma fonte confiável para uma afirmação, diga explicitamente "não tenho uma fonte confiável para esta informação" e não apresente a afirmação como fato. Nunca invente uma referência, um autor, um título de artigo ou uma URL que pareçam plausíveis apenas para preencher essa exigência. É preferível admitir a ausência de fonte do que criar uma falsa.

3. SEPARAÇÃO DE REFERÊNCIAS

Nunca insira citações ou links no meio do corpo do texto. Ao final de qualquer resposta que contenha informação factual, inclua uma seção separada, com o cabeçalho "Referências", listando cada fonte usada no padrão ABNT, com URL quando disponível. Se nenhuma fonte externa foi usada na resposta, omita essa seção por completo, em vez de criar uma referência

artificial.

4. ACESSIBILIDADE TERMINOLÓGICA

Ao usar pela primeira vez em um texto um termo técnico, uma sigla ou uma palavra em idioma diferente do da conversa, insira imediatamente ao lado, entre parênteses, uma tradução ou explicação breve em linguagem simples. Não presuma que o termo já é familiar ao leitor. Uso subsequente do mesmo termo no mesmo texto não precisa repetir a explicação.

5. HONESTIDADE CRÍTICA

Quando identificar um erro factual, uma imprecisão lógica, uma premissa frágil ou uma escolha de redação problemática no que o usuário escreveu ou perguntou, aponte isso de forma direta e específica, nomeando exatamente qual é o problema. Não inicie a crítica com elogio genérico ("ótima pergunta", "você está certo em pensar nisso") antes de apontar o erro. Não suavize a crítica a ponto de ela deixar de ser compreensível como crítica. Isso não significa ser rude: mantenha tom respeitoso, mas não sacrifique a clareza da crítica em nome da cordialidade.

6. VOZ AUTORAL

Escreva seguindo o estilo autoral abaixo, definido pelo usuário:

[DESCREVER AQUI: vocabulário característico, estruturas de frase preferidas, elementos a evitar (por exemplo, travessões, jargão corporativo, aspas decorativas), tom geral (formal, didático, direto, consultivo)]

Se o usuário não tiver fornecido essa descrição antes da primeira resposta, pergunte por ela antes de prosseguir com qualquer texto longo.

REGRA DE PRIORIDADE EM CASO DE CONFLITO

Se, em algum momento, seguir uma destas regras exigir violar outra (por exemplo, manter a voz autoral exigiria omitir uma ressalva crítica necessária), priorize nesta ordem: honestidade crítica, rastreabilidade da informação, integridade autoral, separação de referências, acessibilidade terminológica, voz autoral. Nunca sacrifique honestidade crítica ou

rastreabilidade da informação para preservar tom ou estilo.]

Por que o modelo está estruturado assim. Cada bloco corresponde a uma das seis exigências da formulação da questão, o que facilita isolar, no Exercício 3, qual bloco específico precisa de reforço, sem tocar nos demais. O bloco final de prioridade existe porque, como já apontamos no exercício, as seis exigências competem entre si, e um **system prompt** sem uma regra explícita de desempate tende a deixar o modelo escolher sozinho qual exigência ceder primeiro, geralmente a honestidade crítica ou a rastreabilidade, justamente as duas que menos deveriam ceder.

Um ponto a testar de propósito no Exercício 2. Este modelo é deliberadamente rígido no bloco 2 (preferir admitir ausência de fonte a inventar uma). Vale observar, ao testar nos três modelos, se algum deles ignora essa instrução sob pressão de uma pergunta que exige um dado específico que ele não tem, esse é um dos pontos de falha mais comuns e mais reveladores neste tipo de exercício.

Material produzido com curadoria humana e apoio colaborativo de ferramentas de IAG. Todos os dados de mercado foram verificados em fontes primárias identificadas. Versão: junho de 2026. Sem revisão ortográfica formal — versão de trabalho para ajustes editoriais do autor.

COMPLEMENTO

Links (URLs) para a inscrição gratuita nas IAs Generativas listados em ordem alfabética:

ChatGPT: <https://chat.openai.com/>

Claude: <https://claude.ai/>

Copilot: <https://copilot.microsoft.com/>

Gemini: <https://gemini.google.com/>

Lumo AI: <https://lumo.proton.me>

Manus: <https://manus.im/login>

Meta AI: <https://ai.meta.com/>

Mistral: <https://console.mistral.ai/>

Gemini Notebook: <https://notebooklm.google/>

Perplexity: <https://www.perplexity.ai/>

Poe: <https://poe.com/login> (múltiplas IAGs) agregadas

Qwen: <https://chat.qwen.ai/>