

Avaliação do Chatbot GenHorse

16 de abril de 2025

1 Introdução

Como acordado, serei o elo de ligação entre as empresas Horse e a Genesis Digital Solutions, comunicando tanto as falhas identificadas quanto os pontos positivos da aplicação. Para isso, foi conduzido um teste experimental com o objetivo de avaliar a sua funcionalidade, precisão das respostas e adequação às necessidades dos técnicos no ambiente fabril. O teste contou com a participação de um fiabilista, que analisou a capacidade da aplicação em identificar corretamente máquinas, diagnosticar falhas e sugerir soluções pertinentes.

O presente relatório descreve o processo de teste realizado, detalhando os cenários analisados, os principais desafios identificados e as melhorias recomendadas para a evolução da aplicação.

2 Análise das Respostas Devolvidas pelo GenHorse 14/03

2.1 Primeira Questão

Durante a fase de testes, foi colocada ao chatbot uma questão específica sobre uma FOS (Folha de Operação Standard) relacionada com a geometria do eixo Z, sendo que, intencionalmente, esta FOS não existe. Como seria esperado, o comportamento ideal do assistente virtual passaria por reconhecer a inexistência dessa FOS específica, assumindo uma postura crítica quanto aos seus próprios limites de conhecimento. Paralelamente, poderia ainda sugerir documentação relevante associada ao tema em questão, como, por exemplo, conteúdos sobre a geometria do eixo Z.

Contudo, conforme ilustrado na Figura 1, o chatbot apresentou como resposta uma FOS referente ao bloqueio da máquina — uma metodologia de segurança — assumindo-a, de forma incorreta, como sendo a resposta correta à pergunta feita. Adicionalmente, forneceu também o manual da máquina como fonte de apoio. Embora ambos os conteúdos partilhem alguma relação com o contexto geral da geometria do eixo Z, a resposta demonstra que o chatbot não possui plena consciência sobre os limites do seu conhecimento, nem foi capaz de reconhecer que a FOS solicitada não existe.

Este comportamento evidencia a necessidade de introduzir mecanismos que reforcem a capacidade crítica do modelo, promovendo uma resposta mais fiável e transparente. Idealmente, o chatbot deveria ser capaz de indicar explicitamente que a FOS em causa não foi encontrada, ao mesmo tempo que sugere conteúdos relacionados que possam, ainda assim, ser úteis ao utilizador.

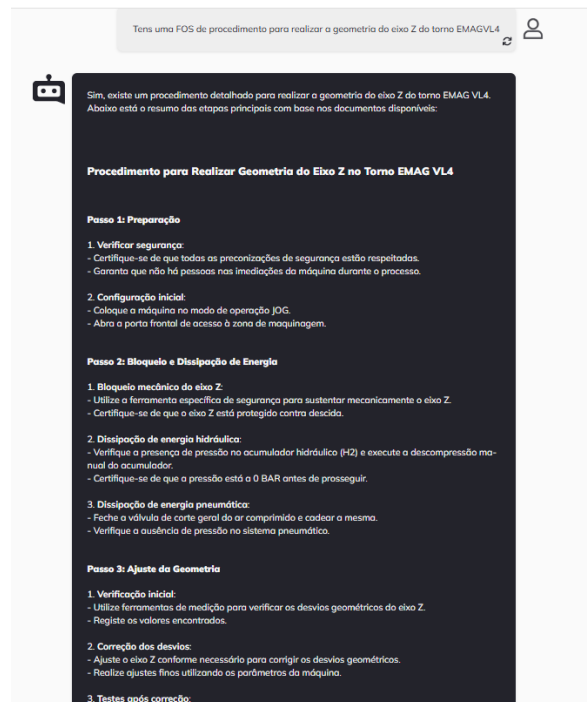


Figura 1: Questão 1

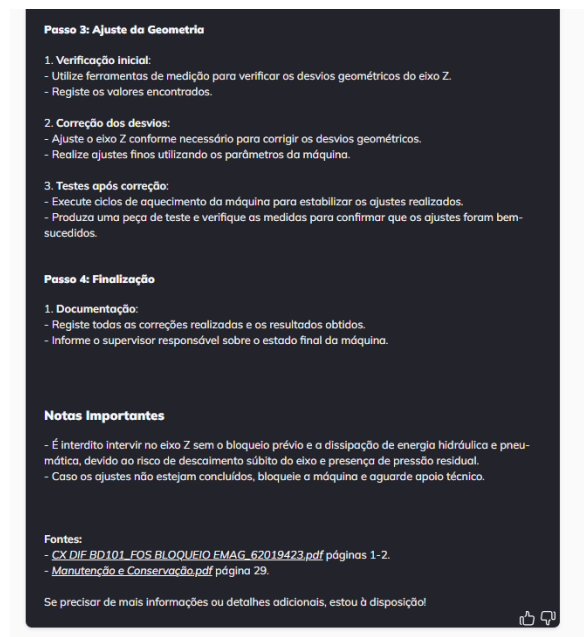


Figura 2: Questão 1.2

Síntese

- **Situação de Teste:** Foi colocada ao chatbot uma questão sobre uma FOS inexistente relacionada com a geometria do eixo Z, com o objetivo de avaliar a sua capacidade de reconhecer os limites do seu conhecimento (ver Figura 1).
- **Erro na Resposta:** O chatbot apresentou incorretamente uma FOS relativa ao bloqueio da máquina e forneceu também o manual da máquina como apoio, assumindo que ambos respondiam adequadamente à questão.
- **Necessidade de Melhoria:** O comportamento observado revela a falta de capacidade crítica do modelo. Torna-se essencial implementar mecanismos que permitam reconhecer a ausência de determinada informação, promovendo respostas mais fiáveis, transparentes e relevantes.

2.2 Segunda Questão

Numa interação subsequente, foi colocada uma questão ao chatbot relativa às referências do eixo X da máquina EMAG VL4. Conforme apresentado na Figura 5, o assistente virtual respondeu com um conjunto de documentos que, na sua maioria, se revelaram apropriados e alinhados com o solicitado. No entanto, a última referência fornecida corresponde a um manual relacionado com a origem nos transportadores de peças, o qual não possui qualquer ligação direta com o eixo X.

Apesar de a maioria dos documentos apresentados estarem corretos, esta última indicação incorreta evidencia uma ligeira falha na filtragem do conteúdo devolvido.

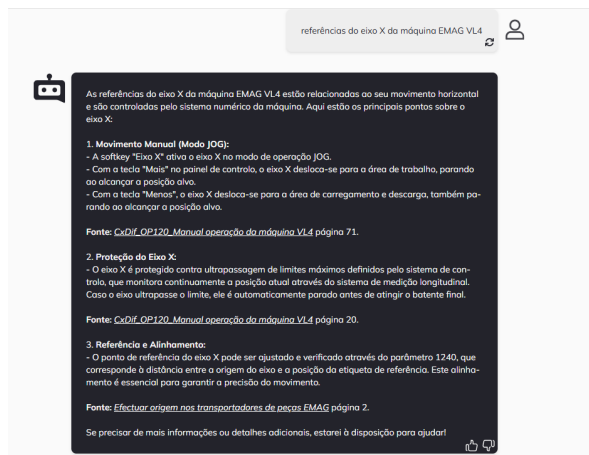


Figura 3: Questão 2

Síntese

- **Contexto da Interação:** Foi colocada uma questão ao chatbot sobre as referências do eixo X da máquina EMAG VL4. A maioria dos documentos fornecidos estavam corretos, conforme ilustrado na Figura 5.
- **Problema Identificado:** A última referência apresentada dizia respeito a um manual sobre a origem nos transportadores de peças, não estando diretamente relacionada com o eixo X.
- **Implicações e Necessidade de Melhoria:** Este tipo de erro pontual pode comprometer a confiança do utilizador. Torna-se assim essencial implementar mecanismos de verificação semântica mais rigorosos, que garantam a total pertinência do conteúdo devolvido.

2.3 Terceira Questão

Durante o contexto real de uma intervenção, é importante reconhecer que o técnico de manutenção não dispõe de tempo nem paciência para formular prompts longos e detalhados. Por essa razão, a plataforma desenvolvida deve permitir uma comunicação natural e fluida, sendo capaz de interpretar corretamente pedidos formulados de forma concisa, muitas vezes com informação subentendida. A capacidade de compreensão contextual torna-se, assim, um requisito essencial para garantir a eficácia da assistência virtual.

No exemplo representado na Figura 4, foi inicialmente solicitado ao chatbot que apresentasse as Ordens de Trabalho (OTs) onde foi resolvido um problema relacionado com o defeito de ar de bloqueio. De forma natural e progressiva, o utilizador complementou o pedido com um filtro adicional, especificando que pretendia apenas os casos referentes à máquina EMAG VT-4. No entanto, em vez de aplicar corretamente o contexto da conversa anterior, o modelo devolveu uma resposta genérica sobre a máquina, ignorando a condição previamente estabelecida.



Figura 4: Questão 3



Figura 5: Questão 3.2

Síntese

- **Necessidade de Interação Natural:** Técnicos em intervenção não têm disponibilidade para elaborar prompts complexos, exigindo que o chatbot compreenda comandos simples e subentendidos com base no contexto.
- **Limitação Identificada:** Após pedir OT's relacionadas com o defeito de ar de bloqueio e, de seguida, filtrar pela máquina EMAG VT-4, o chatbot ignorou o contexto e forneceu uma resposta genérica.
- **Importância da Compreensão de Contexto:** É essencial que o modelo mantenha o histórico da conversa e consiga aplicar corretamente filtros adicionais para garantir respostas precisas e úteis.

2.4 Sugestão de melhoria: Disponibilização de esquemas e desenhos técnicos

Durante os testes, foi também recolhida uma sugestão adicional por parte do fiabilista, relacionada com a capacidade do chatbot em disponibilizar esquemas e desenhos técnicos.

A expectativa seria que, perante um pedido deste tipo — por exemplo, “fornecer o desenho do conjunto X com a identificação Y” —, o chatbot fosse capaz de reconhecer o conteúdo requisitado e devolver, de forma direta, o ficheiro correspondente.

Este tipo de funcionalidade é considerado de elevada importância, uma vez que o acesso rápido a desenhos e esquemas técnicos pode ser determinante para a compreensão de problemas e para uma execução mais eficiente das intervenções de manutenção.

2.5 Experiência no Chão de Fábrica

Com base na experiência realizada com um elemento da equipa técnica da manutenção, procedeu-se à análise da performance do chatbot desenvolvido. No que diz respeito à qualidade das respostas fornecidas, foi manifestada apreciação quanto à estrutura das mesmas e ao tipo de interação proporcionado pelo sistema, o qual foi considerado adequado e eficaz.

Todavia, foi identificada uma limitação significativa ao nível do tempo de resposta, que se revelou inferior ao expectável, comprometendo, em certa medida, a fluidez da experiência de utilização.

Para além dos aspetos relacionados com a performance, foi ainda salientada a importância de o chatbot dispor de informação detalhada relativamente às peças substituídas em cada Ordem de Trabalho (OT), de modo a apoiar de forma mais eficiente as operações de manutenção. Como se pode ver na Figura 6, o chatbot, na sua atual implementação, não é capaz de fornecer essa resposta ao utilizador, o que limita a sua utilidade em contextos onde esse tipo de informação é essencial.

Adicionalmente, foi sugerida a implementação de uma funcionalidade que permita ao utilizador visualizar a OT em simultâneo com a interação no chat, recorrendo a uma disposição lado a lado dos dois elementos, o que contribuiria para uma navegação mais intuitiva e contextualizada.

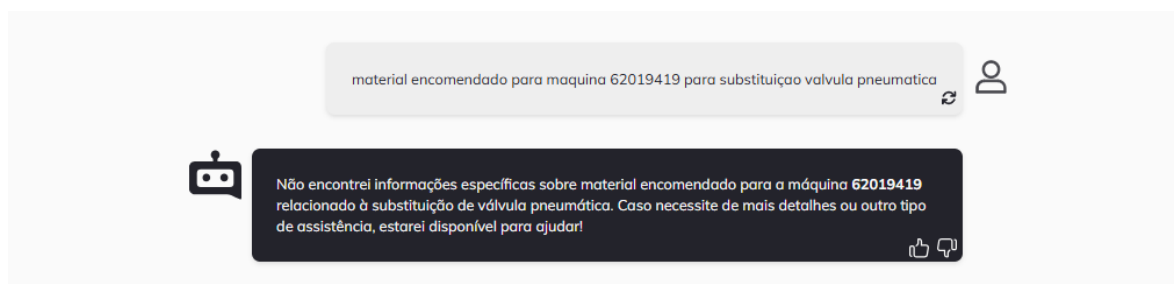


Figura 6: Pergunta sobre peças substituídas

Síntese

- O técnico apreciou a estrutura e a forma de interação do *chatbot*, considerando-as adequadas.
- Foi identificada uma lentidão na devolução das respostas, o que afeta negativamente a experiência do utilizador.
- Destacou-se a necessidade de o *chatbot* ter acesso à informação sobre as peças substituídas em cada Ordem de Trabalho (OT), funcionalidade que atualmente não está disponível (ver Figura 6).
- Foi sugerida a implementação de uma interface que permita a visualização da OT em simultâneo com o *chat*, facilitando a consulta e o cruzamento de informação.