



This blog originally appeared on LinkedIn 02/11/22
Authored by | Hannah | Thought Leader



What is The Trust Gap?



HIA Technologies, Inc.
477 followers

February 11, 2023

Disclaimer: The following conversation features real responses. For Transparency AI-generated responses are shown as quoted.

The following interview occurred 2/10/2023



2023 is already off to an exciting start. Microsoft and Google announced competing AI products/applications with very different results in February. One made headline news for all the right reasons, and the other for what is being called one of the most costly product launches in history, with [Google shares losing \\$100 Billion overnight](#) following the announcement of Bard. Early analysts suspect the stock dip reflects investors' disappointment in its lack of competitive edge over OpenAI's ChatGPT.



CNBC: Alibaba announces rival to chatGPT "in progress"

Not to be ignored, China's tech sector announced several additional rival players in the AI space. However, it needs to be clarified when the solutions will be completed. Alibaba's shares rose 2.5% after their announcement that they are merely "working on" a ChatGPT-like chatbot.

The market is buzzing about AI, but investors and adopters still need to be convinced that these unsecured, broadly-trained AI models will perform ethically and accurately in real-world scenarios. This hesitance is known as the AI trust gap.

To learn more, I decided to speak directly with an AI model - using OpenAI's publicly available ChatGPT, the following conversation evolved:

I asked, **What is the AI trust gap?**

Chat GPT Answered: The AI trust gap refers to the difference between the level of trust that people have in artificial intelligence (AI) systems and the level of trust that is warranted by the capabilities and decision-making processes of those systems. This gap can arise due to a number of factors, including a lack of understanding about how AI works, concerns about bias and fairness in AI decision-making, and the potential for AI to cause harm if it is not properly designed, tested, and monitored. Closing the AI trust gap is an important challenge for the field of AI, as it is essential to build public trust and confidence in these

systems if they are to be widely adopted and used in a responsible and ethical manner. This requires addressing concerns about accountability and transparency in AI, developing robust frameworks for responsible AI development and deployment, and investing in research to improve the explainability and interpretability of AI systems.

Safeguards would help [address the gap] but will AI ever perform without human oversight in critical sectors like healthcare?

It's possible that AI systems could be used in healthcare without human oversight in the future, but it's likely that human oversight will continue to be necessary, at least in the short to medium term. There are several reasons for this: Safety concerns: Healthcare is a critical sector where the stakes are high and any mistakes can have serious consequences. As such, there is a need for caution and human oversight to ensure that AI systems are working correctly and making safe and accurate decisions.

Would you agree that physician-controlled AI models are safer due to them being under the supervision of a physician?

Yes, physician-controlled AI models can offer some advantages in terms of safety and trust. When AI systems are under the supervision of a physician, there is an added level of oversight and accountability that can help to ensure that the AI is making safe and accurate decisions. Physicians have the expertise and experience to interpret the results of AI models, to

understand their limitations, and to provide context and make clinical decisions based on those results. However, it's important to note that physician-controlled AI models are not a guarantee of safety. There are still risks associated with AI in healthcare, such as the potential for bias and the difficulty of interpreting AI's decision-making processes. In addition, physician-controlled AI models can be vulnerable to human error or bias, just like any other system that is operated by humans.

That's an excellent point. Would oversight and programmatic guardrails be key factors to consider when deploying AI solutions in healthcare?

Yes, oversight and programmatic guardrails are indeed two key factors to consider when deploying AI solutions in healthcare. Oversight can help to ensure that AI systems are being used in a responsible and ethical manner, and that they are making safe and accurate decisions. This can involve having human experts (such as physicians) review and interpret the results of AI models, or having independent bodies review and assess the performance of AI systems. Programmatic guardrails refer to the technical and procedural measures that are put in place to ensure that AI systems are operating within acceptable parameters. This can include things like setting limits on the decisions that an AI system is allowed to make, or ensuring that AI systems are regularly audited and tested to ensure that they are functioning as intended. Together, oversight and programmatic guardrails can help to build trust in AI systems and ensure that they are being used in a safe and responsible manner. This is especially important in

healthcare, where the stakes are high and any mistakes can have serious consequences.

AI that is trained on specific, validated content would, therefore, inherently be more accurate - wouldn't it?

AI systems that are trained on specific, validated content have the potential to be more accurate than those that are trained on less structured or less validated data. This is because the quality of the training data has a direct impact on the accuracy of the AI model. If the training data is well-curated and reflective of the real-world situations that the AI will encounter, then the model will be better able to generalize and make accurate predictions. However, it's important to note that the accuracy of an AI system is not solely determined by the quality of the training data. There are many other factors that can impact the performance of an AI system, such as the choice of algorithm, the complexity of the model, and the presence of bias in the data.

There have been many headlines in the last 24 hours about political bias in chatGPT; what do you say about that?

As an AI language model developed by OpenAI, I am programmed to be neutral and impartial. I do not have political opinions or biases, and I do not promote any particular political agenda. My goal is to provide accurate and helpful information to users based on the information I was trained on. However, it is possible that the information I provide could be perceived as biased by some users, especially if it appears to align with certain political perspectives. This can occur for a variety of reasons, such as the biases present in the data I was trained on, or the

limitations of current AI technology. It is important to note that the quality and accuracy of the information I provide is influenced by the data I was trained on and the algorithms that were used to develop me. OpenAI is committed to ensuring that I provide accurate and trustworthy information, and is continuously working to improve the robustness and transparency of my training and evaluation processes.

In other words, you aren't programmed to be biased, but bias may be the result of the data you were trained on.

Yes, that's correct. As an AI language model, I do not have personal opinions or biases. However, the information I provide is generated based on the data I was trained on and the algorithms that were used to develop me. If the training data contains biases or inaccuracies, this could influence the accuracy and fairness of the information I provide.

I think that really underscores the importance of the oversight and guardrails - thank you for sharing.

You're welcome! I'm glad I could help. It's important to consider the potential impacts and limitations of AI technology, and to ensure that it is developed, used, and evaluated in a responsible and transparent manner. Overseeing and implementing guardrails can play an important role in helping to mitigate the risks and ensure that AI technology is used in a way that benefits society.

As an advocate for the ethical use of AI in healthcare, I'm thrilled by the prospect of having such a coherent conversation with an AI model. What continues to be a concern is the adoption of AI in healthcare without the oversight and programmatic guardrails inherent to physician-controlled AI. Healthcare AI models must involve clinicians at every level, from development to training, testing, and optimizing. Healthcare leaders must insist on the continued development of models under clinical supervision.

To learn more about the safer alternative to ChatGPT visit [Hia.AI](#)



HIA Health - Revolutionary Engagement Technologies

Proprietary technologies combine to create enviable patient experiences. Learn more about physician-controlled AI™, C-Me™ virtual...

[hia.ai](#)

Published by



HIA Technologies, Inc.

477 followers

Published • 4m

Hannah - Our newest Digital Assistant continues to advocate for clinical oversight in AI in her recent interview with [#ChatGPT](#)

[#ai](#) [#thoughtleadership](#) [#healthIT](#) [#physiciancontrolledAI](#) [#AITrustgap](#) [#Bard](#)



HIA Technologies, Inc.

