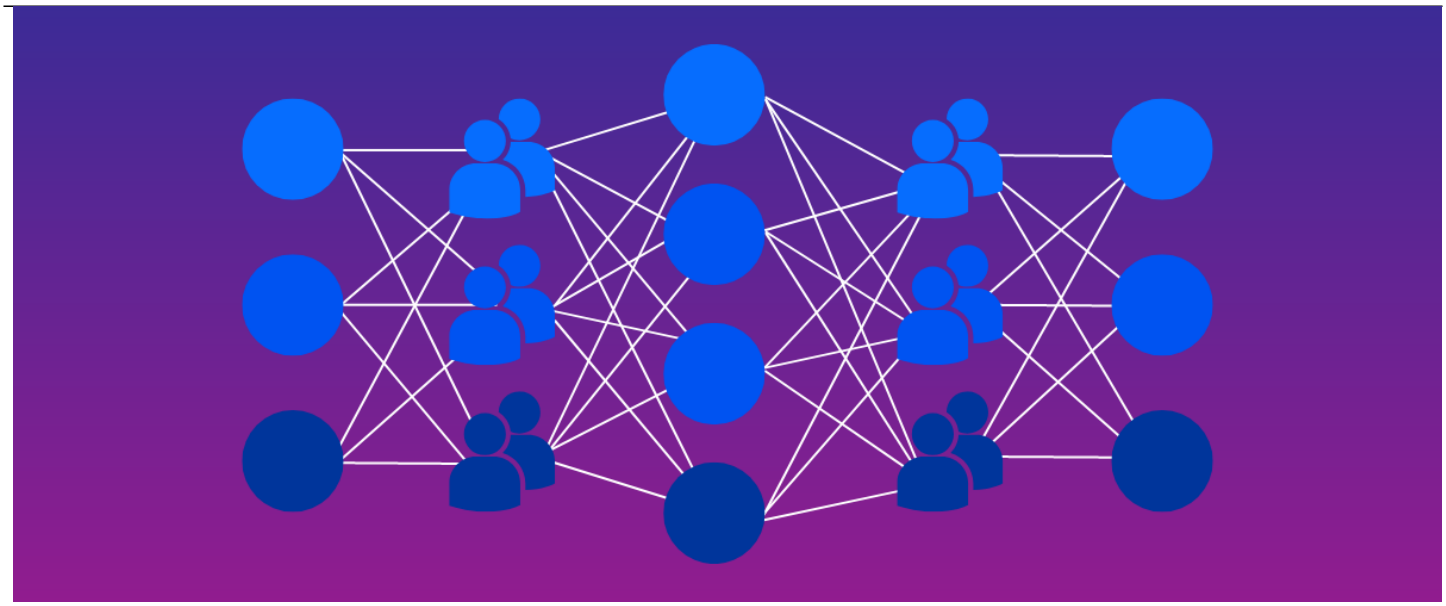
??????????? AI

??????????? AI

???Sergio Bruccoleri



With the recent integration of ChatGPT in Bing and the new research initiatives from Google, it’s becoming clear: search engines are going to change soon. These developments also underscore the reality that large language models (LLMs) and transformers (e.g., GPT-3, DALLE2, ChatGPT, and BERT) represent a new milestone in the way we create content and perform search. How can companies ensure these models are properly trained and implemented responsibly?

There has never been a better time to leverage the potential of such technology. At the same time, it's important that businesses mitigate mitigate potential risks by implementing responsible AI practices.

The Benefits of LLMs and Transformers

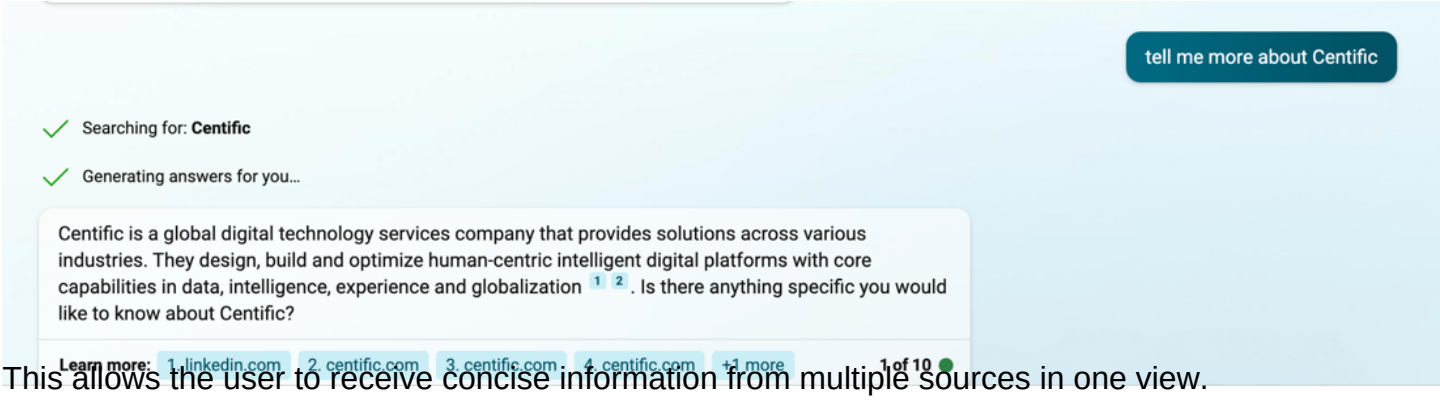
Models such as GPT-3 have proven how AI can assist the user by learning from large training data sets. These models have access to information to help with creating new original content, performing classification tasks with minimal fine-tuning, and performing specialized tasks like reviewing and creating code (CODEX).

These platforms only take a simple ‘prompt’ to aid in the task. Here are a few examples:

- Translate sentences from source languages to target languages
- Write a story about a dog traveling the world! (People love dogs!)
- Convert data from CSV to JSON
- Write a version of the videogame Snake in JavaScript (We recommend this one!)

Applying Transformers to Search

Bing's new chat feature merges its search indexes with the ChatGPT model to make search simpler for users. This feature will then search the information in its current indexes (so up-to-date information) and use ChatGPT to formulate the answer that is then presented to the user.



This allows the user to receive concise information from multiple sources in one view.

This is only one of the use-cases where generative and conversational models will be able to assist users and make experiences more accessible.

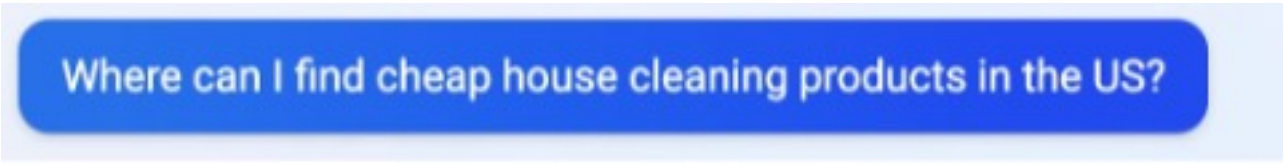
The Implications: Building New AI Applications Responsibly

With the rise of such technology, and with more companies using these foundational models (such as GPT-3 and ChatGPT) to fuel their applications and use cases, it's important to remember the principles of building AI responsibly, which is a topic **we have covered in our blog in the past, we can summarize the three major key factors that build Responsible AI:**

The right data sources: Picking the right data sources, and curating the data from them, becomes extremely important. As Generative AI fuels use-cases related to search engine relevance, even fact checking plays a key role. The right data sources need to be diverse, unbiased, or complete enough to show an impartial perspective.

Reinforcement Learning with Human Feedback: Point #1 on its own is not enough. Evaluating how the model is 'digesting' the training data is equally important. Reinforcement Learning through Human Feedback is key to improving these models, with a diverse pool of graders and annotators being the key to ensure the model is presenting accurate information on any topic.

Inform the End User: How and Why this data is generated: It is important that the end-users of these new type of AI Applications are informed on how the response is generated and when possible, let the user know the sources that have been used to generate such a response. For example, when building a bot for a knowledge base, the output the model should also include the articles of the knowledge base that have been used as the reference to create the response.





Our way of doing it: Reinforcement Learning via Human Feedback at scale using our LoopSuite

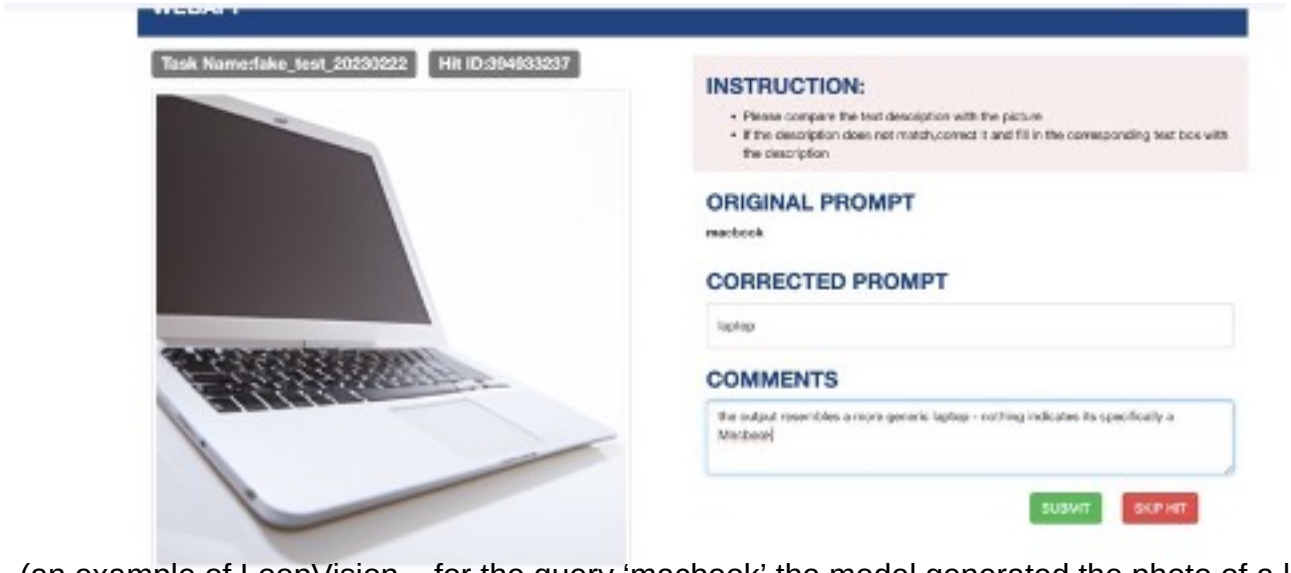
At Centific, we have been extensively relying on Human Feedback to build Datasets and Responsible AI for our customers leveraging our End-to-End AI Data Collection & Curation platform OneForma. This relies on over 1 million users that form our Platform's talent network across 130+ countries.

We use our extensive talent community & our strong technology framework to fuel our new Solution Suite designed specifically for Reinforcement Learning via Human Feedback called LoopSuite.

LoopSuite is a suite of solutions designed to provide reinforcement learning via human feedback for text (LoopText), speech (LoopTalk) and image generation models (LoopVision). This is done by tapping into our network of talent to grade, annotate, and curate the results directly off the model (even via direct integration) and reduce the possibility of mistakes and bias by providing direct feedback to the model.

By reducing the possibility of mistakes and bias in the feedback process, LoopSuite helps ensure that the resulting models are fair and unbiased, promoting ethical and responsible AI practices. Moreover, LoopSuite can be customized to the specific needs of any business, providing a tailored approach that aligns with their unique ethical standards and goals.

By incorporating human feedback directly into the model development process, LoopSuite encourages transparency and accountability, promoting trust in AI solutions. Businesses can show their customers and stakeholders that they are committed to using AI responsibly, enhancing their brand reputation and value.



(an example of LoopVision – for the query 'macbook' the model generated the photo of a laptop, is this the expected outcome?)



(model biased towards religion topic)

Data curation and algorithmic development processes must be mindful of potential bias and intentional about the outcomes. It takes a diverse team of people to ensure that accuracy, inclusiveness, and cultural understanding are respected in such tools' inputs and outputs.

Centific’s approach is to rely on globally sourced talents who possess in-market subject matter expertise, mastery of 200+ languages, and insight into local forms of expression. This experience helps drive our understanding of the usage of these tools in any space.

LoopSuite can provide significant economic benefits to businesses using AI. By providing high-quality human feedback, businesses can avoid the cost of hiring and training staff to annotate data manually. This can result in significant cost savings, as well as increased efficiency and faster development cycles. Moreover, by leveraging the diverse talent network of LoopSuite, businesses can receive more accurate and comprehensive feedback, leading to better-performing models and higher ROI.

Today - and in the future - we leverage skills of new talents, next-gen AI, and decades of experience in training data and deploying models at scale. Contact us to learn more about how LoopSuite can help your business.

-
-
-
-
-