

# Review of: "[Perspective] AI Is All About Typing the Right Phrase"

Michail Ploumis

**Potential competing interests:** No potential competing interests to declare.

The post touches upon some interesting challenges related to the current space of AI application, which need to be tackled during the phase of requirements specification, designing as well as engineering of an AI system. Cases in point being data privacy, bias, transparency, explainability, accountability and fairness, as correctly referred and discussed by the author.

**Data privacy.** An AI model is trained on structured (tabular) or unstructured (graphs) data for a given downstream task, while the case is often that the data are personalized. The model should be able to learn patterns from a centralized or decentralized data pool, while avoiding jeopardizing the identity of the individuals. The task at hand is complex, as according to the method employed, there are challenges imposed on the developer(s). Should the author need to have a more comprehensive view on data privacy preservation techniques, they are advised to consult research work done on the fields of federated learning, multi-party computation and differential privacy.

**Bias.** Depending on the AI application, the dataset an AI model is trained on must be balanced, i.e., consist of examples describing a population thoroughly at the desired granularity. In case the problem is such that only an imbalanced dataset is available, the developers should consider appropriate pre-training statistical approaches, as well as suitable AI models, such as one-class learning models. Otherwise, the well-known problem of “garbage in – garbage out” is manifested. Last, more often than not, there are untended cognitive biases imposed by the developers, such as selective perception and confirmation bias, which affect the algorithms, leading to lack of fairness.

**Explainability & Accountability.** An AI model is heavily depended on math and algorithmics, which are complex even to the developers, let alone non-expert people, who often perceive them as “black boxes”. This fact effectively leads to the fair question of “can we explain why the model made such decision?”. This question is crucial in AI applications. For example, in a fraud detection model, a transaction can be classified as suspicious, which further means that the account owner is considered a fraud suspect and will be subject to further investigations and potential judicial

sanctions. What if the model made a wrong prediction and the individual is falsely accused? The list of examples is exhaustive and applies to most AI applications fields.

The perspective of the author regarding the challenges referred is a fair judgement, however he proceeds to the section of introduction with a referral of Dall-e, an AI image generator, in his attempt to discuss the leap in evolution technology has made. The two main flaws of this section are, (1) it is loosely connected to the following section and (2) it doesn't convey the objective of an introductory section. The introductory section is supposed to ease the reader in the main focus of the study, while the focus the post is not clear.

The following section, "AI in earth science", is an attempt to examine the current status, use cases, challenges and opportunities of machine intelligence in field of earth science. The fundamental flaws of this section is that (1) he refers to this reading as a scientific research paper, he has already stated it's a post, (2) the purpose of such reading is best suited to a systematic literature review and (3) it is loosely connected to the preceding and following sections.

Last, the author switches the focus again and turns his attention to how AI can potentially negatively impact creativity and suggests a way to overcome it.

In conclusion, while the challenges, i.e., privacy, explainability, bias, the potential negative impact of generative AI are valid points, they are conveyed in a somewhat superficial and one-sided manner. The author should make clear whether this reading is a post or a scientific research paper, while he should also narrow down is focus. A suggestion to start with is to discuss how one or multiple challenges specifically affect AI in earth science. Another suggestion would be to turn this post into a systematic literature review of AI in earth science.