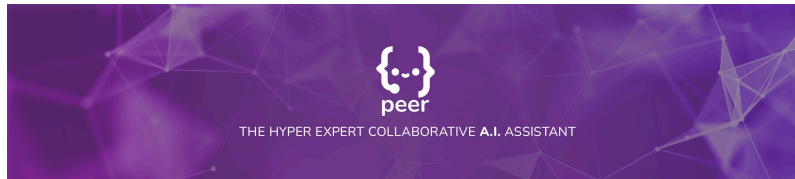




## PEER Newsletter - Issue 3

*This project has received funding from the European Union's Horizon Europe research and innovation programme under grant agreement No 101120406.*



### Trust VS Trustworthiness

**In the field of human-centric AI, distinguishing between trust and trustworthiness is not only a semantic exercise**, but it is the foundation for designing systems that are truly ready for "real-world" deployment.

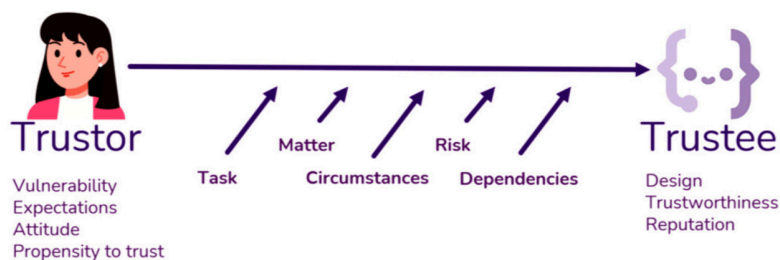
The PEER project directly addresses this distinction by operationalising the concept of trustworthiness to bridge the gap between user expectations and AI capabilities.

#### What is the difference between trust and trustworthiness?

##### TRUST

**Trust** happens between two parties – the trustor and the trustee – in a situation where the trustee has to do something, typically to achieve the trustor's goal. In this process, there are several factors:

- For the trustor, key factors include their vulnerability, expectations, attitude toward technology, and their overall propensity to trust.
- The circumstances also matter, such as the type of task at hand or the level of risk involved, and this is important to understand that trust is not static – it's **dynamic and evolves** based on outcomes and experiences over the use of the AI system.
- The trustee – which is the AI assistant in the PEER project – must exhibit certain qualities or attributes, notably the attribute of trustworthiness, which is one of the keys focused on in the PEER project.



## TRUSTWORTHINESS

While trust can fluctuate, **trustworthiness** refers to the inherent **qualities** or **attributes** that are **perceived by the users** and that make the system deserving of trust in the first place. Note that different users might interpret the same system's qualities in very different ways based on their personal experiences, expectations, and context.

### How PEER Operationalises Trustworthiness?

The PEER project is fundamentally designed to enhance the trustworthiness of AI assistants in sequential decision-making scenarios (such as route planning or manufacturing processes). By focusing on the trustee's attributes, PEER aims to create a solid foundation upon which user trust can be built and sustained.

**Here is how PEER aligns with the key components of trustworthiness:**

- **Human-Centric Design as a Trustworthiness Enabler:** Unlike traditional "black box" systems, PEER embeds the user at the centre of the entire AI lifecycle. By facilitating a bidirectional communication flow, the PEER AI assistant does not just output decisions; it engages in a dialogue, explaining its reasoning and adapting to user feedback. This transparency is a critical attribute of trustworthiness, directly addressing the user's need for explainability.
- **Competence through "Hyper Expert" Collaboration:** A trustworthy system must be competent. PEER's "Hyper Expert" collaborative assistant is designed to support complex, sequential decision-making by combining data-driven insights with human expertise. This hybrid intelligence ensures that the system acts as a capable partner, reinforcing the attribute of ability essential for trustworthiness.
- **Responsiveness and Adaptation:** Trustworthiness is also about responsiveness. PEER's continuous feedback loops allow the system to learn from user preferences and adapt to evolving contexts (e.g., changing obstacles in a route planning use case). This adaptability demonstrates that the system is acting in the user's best interest, a core component of perceived trustworthiness.
- **Operational Frameworks for Trust:** PEER contributes to a shared European framework for operationalising trustworthy AI. By defining qualitative and quantitative metrics for interactivity, acceptance, and explainability, PEER transforms abstract ethical principles into measurable system attributes.

 [Discover more about PEER](#) 

[PEER official website](#)

[PEER YouTube](#)

You are receiving this email because you signed up for our newsletter.

**FOLLOW us on social media!**



If you would like to unsubscribe, please [click here](#).

