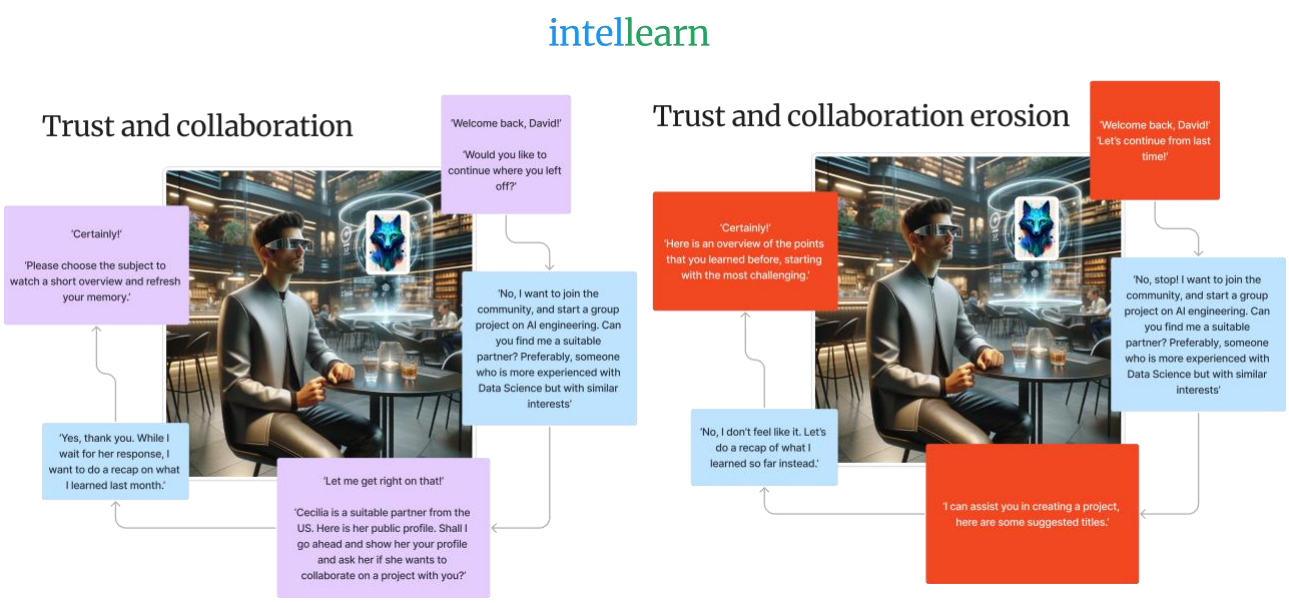
Assignment #3: *A discussion on trust in AI services such as IntelLearn, an AI-powered learning app*

By prioritizing user autonomy, IntelLearn aims to foster a trustworthy and collaborative environment where individuals can feel empowered to achieve their full potential with the support of AI. Below are two versions of the same storyboard. One side (purple speech bubbles) is showing a design that fosters trust and collaboration and on the other side, how that experience can be eroded by a lack of these design principles. The design principles are discussed below in the reflection.



Reflection

In IntelLearn, trust and collaboration between the AI agent and human users are paramount design concepts, while at the same time the goal is to mitigate ethical concerns.

In the realm of educational applications, the stakes are comparatively lower than those of health diagnostic apps for example. While both involve the use of AI, the consequences of inaccuracies or misjudgments in a learning app primarily result in setbacks in educational progress. These setbacks, although undesirable, are generally less severe than the potential ramifications of errors in a diagnostics app.

That said, let's look at the **concerns** arising from AI-powered learning apps such as Intellearn:

1. **Over-reliance on AI**

One issue is diminishing personal discipline. This impacts the learning experience by reducing the user's active engagement and accountability for their progress.

Maintaining user agency and accountability is crucial for fostering a healthy learning environment. Over-dependence on AI adjustments may lead to passivity, hindering the development of critical thinking and self-discipline, as well as diminishing the user satisfaction of achievement and self-efficacy.

To address this issue, IntelLearn will implement a blended approach that combines AI recommendations with user-driven goals and feedback loops. Users will have control over the frequency and extent of AI adjustments, ensuring a balance between personalized support and individual accountability. Users, empowered to set their learning goals, preferences, and boundaries will feel more self-accountable and satisfied. The app will also provide options for manual adjustment of AI recommendations to encourage active engagement.

2. **Privacy and data security due to continuous monitoring**

Continuous monitoring entails the ongoing collection of various types of user data. While this data is crucial for personalizing the learning experience and improving the effectiveness of AI-driven recommendations, it also presents potential risks.

Continuous monitoring involves the collection of vast amounts of user data. This can include sensitive personal information, such as browsing history, learning preferences, and even biometric data if utilizing technologies like facial recognition or biometric sensors. Storing large volumes of user data introduces risks of unauthorized access, data breaches, or misuse. There's the possibility of user data being shared with third parties, such as advertisers or educational institutions, without users' explicit consent, leading to concerns about data exploitation and unwanted profiling. Continuous monitoring systems are susceptible to cyber threats such as hacking, malware, or phishing attacks, which can compromise the integrity and confidentiality of user data.

To address these issues requires a team of various experts, and a designer cannot cover everything but here are some ideas: 1)Add sufficient Safeguards: Encryption and decreased data retention policies may help with unauthorized access or misuse. 2)Ensuring compliance with data protection regulations such as GDPR or CCPA although challenging, especially when dealing with diverse user data and evolving privacy laws, would help protect user privacy. 3) Incorporate privacy principles into the design and development of the application, ensuring that user privacy is prioritized from the outset. 4) User Consent: Obtain explicit consent from users regarding data collection, usage, and sharing practices, providing transparency and control over their personal information. 5) Implement robust cybersecurity measures such as encryption, multi-factor authentication, and regular security audits to safeguard user data from unauthorized access or breaches.

3. **Bias and Fairness**. AI algorithms may inadvertently perpetuate biases present in the training data, leading to unequal treatment or discrimination against certain groups. To address this, the app can implement bias detection and mitigation techniques to ensure fairness in algorithmic decisions. That involves testing with an expert panel of data scientist, human rights specialists, social science experts and philosophers. An additional strategy would be to regularly audit and evaluate AI models for bias, and transparently communicate efforts to mitigate bias to users.

4. **Lack of Human Interaction**. The use of the app may result in reduced opportunities of meaningful human interaction, which is essential for social and emotional development but also for mental wellbeing. Positive feelings and a sense of belonging correlated with the app will also increase a feeling of trust. This is why there is a possibility to interact with other learners using the app to encourage collaboration. There could also be opportunities for group projects, even across disciplines. For example, a designer and a ML engineer, such as implied in the storyboard above.

5. **Digital Equity.** This is a concern that I personally feel is very important. Users feel valued and included when they perceive that efforts have been made to ensure equitable access to educational resources, and in this way it promotes user trust. Access to AI-powered learning apps may be limited by factors such as internet connection, device availability, or digital literacy skills, leading to disparities in educational opportunities. To promote equity, some solutions could be to offer offline access options, or lightweight versions of the app, to accommodate for issues with connectivity, or to provide support resources, tutorials or training programs to enhance user's digital skills and proficiency with the app. Of course, the design of the app should be intuitive and user-friendly from the outset. It would be great to also have a free version of the app available, with less functions or with ads(yuck!).

Trust is fundamental both in the design and development process and in the success of AI powered services. It requires transparency, user empowerment and ethical considerations for the user's data use and wellbeing. While I believe AI can greatly enhance learning experiences, it must be deployed responsibly in all areas of application.

Thank you for reading!
/Eirini