Trust by Design: Building Credibility in Al with User-Centered Principles

Zahra Langford

zahrala@gmail.com

Abstract

The current state of AI products for decision-making in the market is disheartening. As the hype fades, realization sets in that these tools are unreliable, and more likely to create extra work in the short term than reduce it. These products often fail in complex real-life scenarios, are difficult to use, and untrustworthy. As technology creators at an inflection point in our industry, we have an opportunity to slow down and take time to reevaluate the processes used to build AI products so that we can move forward more thoughtfully while earning users' respect and trust.

The goal of this paper is to contribute to the practical understanding of how to build better AI products for decision-making from a user experience (UX) design perspective. It details design principles,

- Center Users Core Principle: Prioritize user needs at every stage of AI product development.
- Articulate Trustworthiness Ensure transparency in AI product algorithms and operation.
- Exhibit Confidence Demonstrate the reliability of AI product outputs.
- Illuminate Logic Explain the reasoning behind AI product decisions.
- Facilitate Action Enable users to act effectively on AI-generated output.

It includes visual examples and practical strategies for applying these principles to the AI products you may be working on.

Together, we can impact the usability and trustworthiness of AI products industry wide. We can increase the user-centricity of our products even if we only focus on one or two of these design principles to champion within our organizations. These principles support the creation of products that delight users and make their lives easier, while rebuilding the trust that is so critical to AI technology moving forward.

Biography

Zahra Langford is a user experience design leader with 15+ years championing end-to-end design of artificial intelligence (AI), healthcare, finance, and ecommerce products for SaaS B2B and B2C companies. She designed her first machine learning app in 2012 as Principal Interaction Designer at the Xerox Innovation Group. She is known for user empathy, curiosity, and measurable results. She currently works as the UX Program Manager and Designer at Rogue Credit Union. Zahra is also a small business owner; she is always experimenting with AI products to augment her workflows to understand customers and create nature-inspired fashion and home décor designs.

Copyright Zahra Langford 2024.

1 Introduction

The recent explosion of AI products aimed at improving decision-making, combined with the rush to get technology to market has led to unusable products, broken promises, and a deepening lack of trust in AI companies, further compounding trust issues inherent in the field (Chakravorti 2024).

The *AI Gone Wrong* page maintained by tech.com, a news and reviews resource, documents a growing list of AI product errors, mistakes, and failures by prominent technology companies (Drapkin, 2024). Only 15% of product leaders in North America and Europe report that their users are embracing AI features (Pratt, 2024). As we descend into the "trough of disillusionment" of the Gartner Hype Cycle[™] for AI (Gartner Research, 2024). There is an opportunity to step back and reassess our process before we alienate users entirely with public missteps, half-truths, and misunderstandings of their needs and wants.

The design, development, and delivery of these products is often missing real connection to users and their goals. Few are asking, "Do users want or need these products and features?" "Are they solving real user problems?" "How can we demonstrate accuracy and confidence?" Let us stop leading with technology and start leading with user-centered principles.

Our primary challenge is to build usable and trustworthy products that meet users' needs while balancing technological advancement and competitive advantage.

2 Purpose

This paper aims to support the delivery of AI products for decision-making by providing design principles to ensure these products are desirable, useful, usable, and trustworthy for target-users. It documents the use of design principles to create better AI products and demonstrates how to apply these to real-life products. By conclusion of this paper, you should understand industry challenges, have design principles to address them, and practical ideas for transforming this information into actionable steps to build the next generation of transformational AI products.

3 Definitions

3.1 Artificial Intelligence (AI)

- AI Systems capable of performing tasks that normally require human intelligence, consisting of algorithms and models that enable these systems to learn from data, recognize patterns, and make decisions (GeeksforGeeks 2024).
- All Als referenced in this paper are Narrow AI (designed and trained for specific tasks or domains) expert systems and do not include physical/mechanical systems (GeeksforGeeks 2024).
- The Als in this context can consist of one or more sub-systems, including Computational Linguistics (CL), Large Language Models (LLM), Machine Learning (ML), and Natural Language Processing (NLP).

3.2 Al Products

- Al-driven products where the product and the Al that powers its features and functionality are built at the same time and the product is not functional without Al.
- Al-assisted products products that already existed, to which Al-driven features are integrated. Example: Microsoft 365 products integrate Copilot as an Al assistant.

This paper will use the term "AI products" to refer to both AI-driven and AI-assisted products. The focus will be on products that analyze large chunks of structured and unstructured information to generate results such as recommendations, predictions, identification, and summarization that support strategic decision-making.

3.3 Design Principles

Product design principles are value statements that apply to the presentation layer of products, outlining the key objectives that products should deliver for their users. They frame design decisions, typically focusing on UX design fundamentals such as "Provide the ability to undo and redo actions" or "Be flexible," but also those that align with a specific mission and vision like "Simplicity" and "Inclusion."

Every organization can benefit from a set of core design principles. In this case, the focus is on a set of principles tailored to the unique challenges of designing and developing AI products.

The principles in this paper are written to integrate into a company's core design principles. If an organization does not have existing principles, let the principles documented in the following pages inspire you to help create them and incorporate these with an AI perspective. Feel free to customize them, every company has a unique way of articulating principles to match their mission, vision, target users and brand.

4 Challenges

The rapid growth of AI in technology comes with inherent internal challenges to the creation of usercentered and trustworthy products.

- Companies face pressure to make statements about AI capabilities and release products before they are fully developed and tested. Announcing exciting new AI capabilities can boost financials, but failing to deliver as promised can result in decreased stock prices, loss of valuation, lawsuits, and penalties (Donachie 2024).
- Companies often lack incentive to fully disclose what their AI products can and cannot do, clearly explained, stored, and anonymized, and identify the level of accuracy and confidence in results.
- The highly competitive nature of the industry fosters secrecy to protect competitive advantages, limiting transparency about how the technology works.
- There is often a drive to acquire more users to improve product algorithms, with less consideration given to retention and repeat use.
- It may be difficult to explain how the AI product generated a specific output due to its independent learning capabilities.

In addition, pushing technological boundaries while maintaining user focus is difficult. Users often cannot articulate their needs and may not recognize the benefits of certain technologies until they can experiment with them.

"What people say, what people do, and what they say they do are entirely different things." - Margaret Mead

These challenges can limit our efforts. However, the design principles presented in this paper will aid in overcoming them.

5 Design Principles

This collection of design principles for AI decision support products has evolved over the course of 15+ years in collaboration with esteemed data science, design, and engineering colleagues.

5.1 Core Principle: Center Users

This key principle should be at the heart of everything. Users come first. We need to understand the target users deeply before applying AI to solve their challenges. Users should guide technology, not the other way around. Simply staying committed to users, their needs and wants, will organically increase the quality of AI products.

Think of users as more than sources of training data, consider the advantages of building long-term relationships with repeat usage that will drive product stability and success. A user-centered foundation can inform the integration of AI functionality that learns about users as they interact with the product, which can automatically evolve a more personalized experience for them over time.

Make centering users a priority with these straightforward strategies to improve user focus and product relevance from inception.

5.1.1 Everyone is not Your User

Remember that "everyone" is not a legitimate target-user population. Identify exactly who you are designing for and what problems you are solving for them before completing major work. This is your **North Star.** Start with something easy and clear, it can evolve as you learn more. Keep it documented and visible as the project progresses. Remember that there may be additional users beyond your primary ones such as admins and specialists. Here are ideas to help you get clear on your target user (or users).

Try to answer these simple questions, use relevant qualitative and quantitative data if available.

- Who are the users? Define characteristics such as demographics, profession, income, device, screen size, and OS.
- What do they want? Understand their challenges, needs, and desires.
- What problems are we solving for them? Explain how the product will improve their lives.
- What does user success look like and how is it measured?

Stakeholders and team members should be able to explain the product to people on the street with limited technical backgrounds. For example: Users X and Y, who [key characteristics], struggle with A and B due to C and D. This product addresses these challenges by [how it improves their lives], enabling them to achieve E and F more effectively, with success measured by G and H.

Build user data collection into products so you can continue to learn more about them and make relevant improvements over time. The user focus does not end with the first deployment, this data will contribute to effective product improvement road mapping.

5.1.2 Find Real or Representative Users

Access to users is critical to understanding, helpful for answering questions, and validating direction. Product team members are not the users of the product they are bringing to life and should not attempt to be. If there is no access to real users, find people in the company who can accurately represent users such as customer support representatives or business analysts. Ask people who are close enough, in demographics and characteristics, to the target user and who are easily accessible. I have found that internal personnel can be immensely helpful because you do not need an NDA, and interactions can be more informal.

5.1.3 Address Bias

This could be a topic for an entirely separate conference, so this paper does not go into detail. The most important thing to keep in mind is that it is never safe to assume that AI algorithms do not contain bias. Approach them as if they have bias built in and the capacity to produce biased output when trained on existing content. Try to do everything you can to reduce this bias, it should be a continuous process (Timbo 2023).

5.1.4 Keep it Simple

Keep it simple stupid (KISS). Refer to your north star when deciding what features and capabilities to include and how to present them in the interface. Make sure the most important things are the easiest to identify and understand. Secondary functionality can be less prominent or a click or two deeper. One of the most common issues with most products is that they try to do too much and crowd everything into a single view.

5.1.5 Consider Accessibility

Consider accessibility before starting the project. Identify Web Content Accessibility Guidelines (WCAG) 2.2 Level A (minimum) requirements and Level AA requirements (widely considered legally acceptable) to integrate into product planning (W3C Web Accessibility Initiative 2024).

5.2 Articulate Trustworthiness

Users should not have to search for the information they need to make decisions about placing trust in Al products. A summary of trust indicators should appear prominently on the product's entry pages and in any descriptions of the product and its capabilities. Include confirmation of trustworthiness at key points of interaction during the product's use (see Design Principles 5.3-5).

Al products on the market can be very vague about what they do. Companies force users to create accounts and experiment with the capabilities to understand if it will help them with their workflows. If users discover it will not, they have already sunk a significant amount of time and money in subscription or access fees. This obfuscation can alienate users, making it unlikely that they will return, even after product teams implement improvements.

When users first encounter the product, they immediately seek indicators that it is trustworthy and can meet their needs. Begin with a clear explanation of your **North Star**: who the product is for, what problems it solves, and what makes it special. Then, present the company's high-level values, articulating their commitment to transparency and ethics, including specific details in clear and easy to understand language. (Deodhar, Borokini and Weber 2024).

- Model definition, use, and behavior.
- Model training data content and source.
- Established industry frameworks incorporated to arrive at correct and usable output.
- Data privacy policy with the ability to opt out of data tracking. Describe how the system collects and utilizes user data.
- Security information describing how the product keeps users' data secure.
- Privacy or security certifications obtained.

5.3 Exhibit Confidence

Assure users of the accuracy of the AI product's output. Here are strategies for demonstrating reliability and confidence.

- Reference supporting sources or industry data and include reputable sources used to generate and validate results.
- Use an established confidence scoring methodology or design a validated one that is clear to non-technical users.
- Clearly label outputs that are speculative and need further validation.
- Provide ways to make independent validation easy.
- Ask users to give feedback on the quality or usefulness of the output.

While researching this paper using Microsoft's Copilot in Bing, the tool gave me data that seemed inaccurate. I prompted it to give me a confidence score for that data. The response: "Certainly! When generating responses, I do not directly provide a confidence score. However, I strive to be accurate and helpful." This was extremely unhelpful and diminished my trust in the output.

Here is an example from my work at HealthPointe solutions as the Director of Product Design (2019-23). This is a page from the Enhanced Medical Chart product that helps medical providers get a useful and searchable summary of their patients' health. It surfaces insights about potential conditions from unstructured notes in the patient records that would otherwise go untreated.

In this scenario, a provider has searched for 'HCC' (Hierarchical Condition Categories: sets of medical codes linked to specific clinical diagnoses) to find any undiagnosed conditions. The product surfaced three suspected HCCs and it gives a confidence level for each. Selecting one of the HCCs links the provider directly to the places in the unstructured medical record where the evidence is located, so the provider can validate for themselves the decide whether to document an official diagnosis and start a treatment plan for that condition or disregard it.



Search results from HealthPointe's Enhanced Medical Chart showing undiagnosed conditions by strength of evidence. No real patient data was used to generate these screens. Proprietary. © HealthPointe Solutions. All Rights Reserved.

5.4 Illuminate Logic

This principle is an early vision of what is now evolving into "Explainable AI." Trust builds on understanding. Products should reveal how they work by showing the source data and, when applicable, a decision tree used in the final output. This can be difficult as companies must safeguard proprietary methods unless committed to an open-source model. However, there are ways to do this that meet the users' need for understanding while protecting competitive advantage (Majewski, 2022).

- Link to original source material so the user can evaluate its trustworthiness.
- Offer a high-level explanation of key factors used to reach a conclusion. For example, explain that the model predicted X due to factors Y and Z.

Most Al-driven search engines now include links to references in the top-level results summary. However, this does not seem to have made its way into more complex decision support products.

Here is another example from HealthPointe Solutions' Enhanced Medical Chart, linking a code prediction to specific appearances in the patient's medical record. It shows the exact information used to make the recommendation.

HEALTHPOINTE"	Enhanced Medical Chart					
Page No: 1		Julia	ana Moore (05/05/1976)			
Contents & Search			\sim			
Legend: Conditions and Symptoms Medications	s 🔽 Procedures 🗸 La	bs ✔Allergies ✔Life Styles ✔Dates ✔Other				
ED PIT note - ED Notes (Inpatient Encounter) PIT Course (ED PIT note) 83 yo male brought in by EMS complaining of polisible tonight. He reportedly passed out 1-2 times. He way is transfusion. On exam patient appears slightly pale. Contusions note Workup initiated in triage I initiated the medical screening exam of to begin wo **** A *** PA 05/16/20 2154 Discharge Summary - Encounter Notes (Inpatient Er 5/22/2020 3:35 PM. Hospitalist Discharge Summary. Patient: Date: (5/22/2020) yo. Gender: male Admission Date: (5/17/20) Length of S PATIENT FOLLOW UP INFORMATION (Discharge Sur Major/New hospital diagnosis: hypovolemic shock) (G Significant Laboratory results needed for follow up: CB Significant Pathology results needed for follow up. NA Anticoagulation and follow up needs: D/C coumadin Medication changes and reason: PPI BID metoprotor New DME: N/A Consultant recommendations: follow up with heme/onc at Discharge Disposition: Home or Self Care Immunization History (Discharge Summary)	Gi bleed . Patient reports Information for Con MS DRG 791 Clinical Codes Rank ICD Code 1 K92.2 Appearances: 5 Section H&P - Encounter Notes ED PIT note - ED Note requiced and GI outpatient	is he noticed frank blood in the toilet dition or Symptom	05/16/2020			
Previous: Vaccinations	Gc	To Page: 1	Next: Page 2 >			

Detail of patient record from HealthPointe's Enhanced Medical Chart showing evidence to support its automated condition coding. No real patient data was used to generate these screens. Proprietary. © HealthPointe Solutions. All Rights Reserved.

5.5 Facilitate Action

The promise of AI extends beyond accurate and trustworthy output. Empower users to make informed decisions and act on information, driving change, and delivering tangible results within the same ecosystem. AI products often address only one part of a user's workflow. Ideally, users should not have to move information in and out of other systems to complete those workflows. This can be difficult to accomplish and adds more effort, but therein lies the true value of AI.

Financial institutions employ AI products to make automated decisions on loan and credit card applications, reducing effort on the straightforward applications while sending more complex cases to a human for processing. Healthcare coding systems can make automated decisions for billing purposes, using code prediction like the example above, sending invoices to insurance companies with a high rate of success.

Actions can be as simple as sending an email with a report to ask another person to complete the necessary work, such as making a phone call, submitting a referral, or scheduling appointments. If you cannot integrate actions into the product, explain how users can act on the information provided in diverse ways. Include the ability to export data and reports out of the product in a way that is consumable by other systems.

Here is an example from HealthPointe Solutions' Digital Care product that assists medical providers in managing their patient population and improving outcomes for high-risk patients. A provider can select a patient or patient group with similar characteristics such as conditions, risk factors, or appointment

recency. Then the provider can act on the insights or circumstances. Actions in this context are: Offer transportation assistance, assign to social work or nurse to call patient.

@	Utilization Gap for Juliana Moore Lab tests are irregular and sporadic									
Summa	ary									
In	Progress	Medium	Evidence • Tests see	emed to be irregular.						
Status		Priority	Prior can	cellations						
10/18/2	2021	Yesterday								
Date Op	ened	Date of Latest Action								
Recommended Actions Offer transportation assistance 1 gap Assign to social work 2 gaps Action History										
Status	Action			Date Assigned	Assigned by	Assigned to	Date Completed	Outcome		
-	Schedule tests for the next 6 months			10:30 am	You	Broadway Health Scheduling				

Detail of patient records from HealthPointe's Digital Care product showing recommended actions that the interface can initiate. No real patient data was used to generate these screens. Proprietary. © HealthPointe Solutions. All Rights Reserved.

6 Integrating Principles in Real Life

Integrating these design principles into products and product development cycles can present an entirely different challenge. The culture of some work environments may not be supportive, particularly if they have achieved success with a technology-first approach. The principles described in this paper represent an ideal and unattainable state. However, championing even one or two of these ideas will increase the quality of the AI products we produce. Here are recommendations for navigating internal roadblocks to increase involvement and influence.

6.1 Start with Domain Expertise

Software and software quality engineers can champion Principles *Exhibit Confidence* and *Illuminate Logic*. Designers and product managers can explore *Facilitate Action*. Marketing can execute *Articulate Trustworthiness*, but the details about model behavior, data, data collection, data use, and established frameworks that algorithms incorporate should come from engineering.

Eventually everyone involved in the planning, design, and delivery of AI products can benefit from a shared understanding of all principles, particularly core principle: *Center Users*. Without the framework they provide, products can go off the rails quickly without any real awareness of the missed opportunities.

6.2 Choose Strategically

Target products where things are more flexible and open; where the roles of project team members are fluid, and contribution from multiple disciplines is welcome (Czernek 2018).

- Start with smaller projects rather than large ones. Quick wins can boost confidence and trust in the principles for use on future important products.
- Select from scratch products in the early planning stages or those that have newly formed teams. These can be easier to influence and introduce innovative ideas.

6.3 Plan Testing and Validation

Get involved in the planning process as early as possible. Consider users in project planning by asking the user definition questions from the *Center Users* section (5.1.1). If direct involvement is not possible, ask your manager to pose these questions. Write user stories correctly to facilitate easy testing.

6.4 Connect with Design

If your team has product or UX designers, it makes sense to build a relationship with them. It is mutually beneficial and will increase the quality of the products you deliver together. Ask designers if you can sit in on their early sessions and invite them to relevant engineering meetings. Maintain frequent connection to share status, ask/answer questions, discuss technical challenges, and co-create solutions.

7 Conclusion

There is increasing recognition that UX design principles have a critical role to play in the creation of AI products that delight users and make their lives easier while rebuilding trust (Agidi 2024). Along with an awareness that continuing to build technology-first products is unsustainable and ill advised.

Conversations about "Explainable AI" are gaining more traction and companies like IBM, Salesforce and OpenAI have added trust indicators and usage policies. However, interfaces demonstrating even one of the design principles described in this paper are rare.

There are so many opportunities, and it is the perfect time to advocate for a better way. The application of these design principles will usher in a new era of AI products. Pick a principle and start today.

8 Special Thanks

I would like to thank my colleague and friend <u>Nathan Gnanasambandam</u>, Chief Technology Officer at HealthPointe Solutions, for his inspiration and support on this topic, as well as his friendship. We have worked together to build successful, user-centered AI products for over a decade.

9 Disclosure and Acknowledgement

This section discloses and acknowledges relevant third-party work from which the current work is derived; especially content generated with the use of AI and AI-assisted technologies.

Use of AI in the writing of this paper.

- No use of any AI or AI-assisted tools to create figures, images, or artwork.
- No images violate state or federal laws.
- Use of two AI or AI-assisted tools to help complete the research for this paper: Microsoft Copilot in the Bing browser and ChatGPT-4.
- Use of AI for specific use cases:
 - Editing proposal After writing original, unformatted first and second drafts, ChatGPT-4 was used to organize and edit that content to fit the proposal format.
 - Improving paper clarity and impact ChatGPT-4 was used to identify opportunities to reduce the use of the passive voice in the final draft.

References

- Agidi, R. (2024, April 13). *Designing for trust: a designer's guide to building responsible AI products.* Retrieved August 11, 2024, from Medium: <u>https://medium.com/@agidirukevwe/designing-for-trust-a-designers-guide-to-building-responsible-ai-products-9d5d0d07f85d</u>
- Chakravorti, B. (2024, May 03). *Al's Trust Problem: Twelve persistent risks of Al that are driving skepticism.* Retrieved August 11, 2024, from <u>https://hbr.org/2024/05/ais-trust-problem</u>
- Czernek, K. (2018, November 13). Why Software Engineers Should Engage in Product Development. Retrieved August 11, 2024, from Medium: <u>https://medium.com/hackernoon/why-software-engineers-should-engage-in-product-development-627d4d581795</u>
- Deodhar, S., Borokini, F., & Weber, B. (2024, August 5). *How Companies Can Take a Global Approach to Al Ethics*. Retrieved August 11, 2024, from Harvard Business Review: <u>https://hbr.org/2024/08/how-companies-can-take-a-global-approach-to-ai-ethics</u>
- Donachie, P. (2024, March 18). SEC Accuses Two Firms of Lying About AI. Retrieved August 11, 2024, from WealthManagement.com: <u>https://www.wealthmanagement.com/regulation-compliance/sec-accuses-two-firms-lying-about-ai</u>
- Drapkin, A. (2024, August 16). *AI Gone Wrong: An Updated List of AI Errors, Mistakes and Failures.* Retrieved August 20, 2024, from tech.co: <u>https://tech.co/news/list-ai-failures-mistakes-errors</u>
- Gartner Research. (2024, June 17). *Hype Cycle for Artificial Intelligence, 2024.* Retrieved August 15, 2024, from Gartner: <u>https://www.gartner.com/en/documents/5505695</u>
- GeeksforGeeks. (2024, August 8). What is Artificial Intelligence AI Definition, Types, Working -GeeksforGeeks. Retrieved August 11, 2024, from GeeksforGeeks: https://www.geeksforgeeks.org/artificial-intelligence-an-introduction/
- Heikkila, M. (2024, July 22). Al companies promised to self-regulate one year ago. What's changed? Retrieved August 11, 2024, from MIT Technology Review: <u>https://www.technologyreview.com/2024/07/22/1095193/ai-companies-promised-the-white-house-to-self-regulate-one-year-ago-whats-changed/</u>
- Majewski, S. (2022, August 16). *Explainable AI: A Way to Explain How Your AI Model Works*. Retrieved August 18, 2024, from Medium: <u>https://medium.com/dlabs-ai/explainable-ai-a-way-to-explain-how-your-ai-model-works-5766b1f11deb</u>
- Pratt, Louron. (2024, July 17). *Product survey findings: Only 15% of users are embracing AI features*. Retrieved August 20, 2024, from Mind the Product: <u>https://www.mindtheproduct.com/product-survey-findings-only-15-of-users-are-embracing-ai-features/</u>
- Timbo, R. (2023, September 28). AI Bias: What It Is & How to Prevent It? Retrieved August 11, 2024, from Revelo: <u>https://www.revelo.com/blog/ai-bias</u>
- Tonye, G. (2021, August 25). *Machine Learning Confidence Scores All You Need to Know as a Conversation Designer*. Retrieved August 11, 2024, from Medium: <u>https://medium.com/voice-tech-global/machine-learning-confidence-scores-all-you-need-to-know-as-a-conversation-designer-8babd39caae7</u>
- W3C Web Accessibility Initiative. (2024, August 6). *How to Meet WCAG (Quick Reference).* Retrieved August 11, 2024, from W3C Web Accessibility Initiative: https://www.w3.org/WAI/WCAG22/quickref/