Ning Zheng

Cover Letter for Discussion Leader

CSE-60687 (Research Methods)

Paper: Guidelines for Human-AI Interaction

General Topic: Identifying general design guidelines for infusing AI technologies in human-facing applications.

Specific Behavior or Activity Studied: As automated inferences are typically performed under uncertainty, often producing false positives and false negatives, demonstrating unpredictable behaviors that can be disruptive, confusing, offensive, and even dangerous, AI-infused systems can violate established usability guidelines of traditional user interface design. Therefore, new guidelines should be investigated.

Specific Research Questions: What are the reusable guidelines at the current stage to help with the design and evaluation of AI-infused systems that people can understand, trust, and can engage with effectively?

Paradigm: The paper first synthesizes a unified set of design guidelines from a variety of communities and sources, and then systematically examine those guidelines in a variety of AI-infused systems to validate their applicability and relevance. To address the problem of lack of rigorous validations of proposed design heuristics in specific domains, the guidelines in this paper are developed using a four-stage process:

- Phase 1: Consolidating guidelines
- Phase 2: Heuristic evaluation from participants
- Phase 3: User study with AI-driven features of products
- Phase 4: Expert evaluation of revisions

Importance: The guidelines serve as a source for designers working with AI and can help create intuitive and effective AI-infused systems that people can understand, trust, and engage with effectively.

Claim: The guidelines will result in better, more human-centric AI-infused systems and facilitate future research.

State of Knowledge: The *"ways of knowing"* of this paper draws upon established theories and approaches in heuristic evaluation[1], qualitative analysis (e.g., affinity diagramming) etc. However, the methodology inevitably suffers validity threats, such as individual bias, where data analysts' personal preferences might influence data annotation; subjective bias, which can occur if participants improve in subsequent tests due to familiarity with the testing procedure or content; and sampling bias: the method includes a step in inspecting existing AI-infused products or features, the generalizability of those products are likely to be biased.

Evidence: Evaluations are conducted following Phase 3 to determine: 1) the relevance of the guidelines, and 2) the clarity of the guidelines. These evaluations are based on participants' responses to usability questions in a questionnaire. To gain deeper insights, particularly regarding negative feedback, the authors engage in discussions with participants to understand their reasons for giving negative ratings. During the expert evaluation in Phase 4, findings are presented on experts' preferences between the original and revised guidelines through bar charts.

Story Structure: The paper begins by exploring the significance of investigating design guidelines for AI-infused systems within the HCI community. It then presents a detailed description of the methodology, presenting each step of the research process alongside relevant theories and established approaches to lay a theoretical groundwork. Finally, the paper summarized the design guidelines and discusses their limitations.

---

[1] Jakob Nielsen and Rolf Molich. 1990. Heuristic evaluation of user interfaces. In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '90). Association for Computing Machinery, New York, NY, USA, 249–256. https://doi-org.proxy.library.nd.edu/10.1145/97243.97281