

DRAFT Notes and Topics from the

2016 AVIXD Workshop on Ubiquitous Voice: Creating Guidelines for Designing Conversational User Interfaces

28 September 2016

This document is a record of the discussions that took place at the 2016 AVIXD Workshop on Ubiquitous Voice: Creating Guidelines for Designing Conversational User Interfaces that took place 22 May 2016 in Washington, D.C..

Thanks to co-organizers Kristie Flenord, Amy Goodwin, and Shelley Moore for all their work to prepare for the conversation. Participants at the workshop were David Attwater, Kristy Boyer, Caroline Collins, Juan Gilbert, Stephanie Laureys, Jim Lewis, Michael McTear, Crispin Reedy, Jamey White, and Susan Hura.

These notes were compiled by Susan Hura, who took notes during the meeting. Susan has volunteered to lead a proposed AVIXD Working Group to revise and expand the current AVIXD Design Guidelines Wiki for ubiquitous voice and conversational UIs.

PRELIMINARIES

- 1) What counts as a **conversational** interface?
- 2) How do we appropriately evaluate conversational UIs?
- 3) What counts as evidence?
- 4) What is a useful classification of conversational UIs of different types? Do we need to include visual or other modalities in such a classification?
- 5) How universal are existing VUI design guidelines?
 - a) Are universal guidelines so broad that they will not be easy to apply in particular design decisions? E.g., Grice's Maxims are very relevant to designing spoken interactions, but they do not offer specific how-to advice.
 - b) How specific do guidelines need to be to be relevant for different contexts of use?
 - c) What are the contexts/domains/industries in which conversational UI design guidelines apply?
 - i) How do the guidelines differ for various contexts of use?
- 6) What is the role of standards for conversational interactions? Do standards limit design innovation? Should we produce guidelines (aimed at informed professionals) or standards (which anyone can follow)?
- 7) Does the development in voice interaction design mirror the development from HCI to usability engineering, to user centered design, to user experience design in GUI? What can we learn from the process they went through?

CONTEXT of USE

- 8) In what contexts is ubiquitous speech necessary? What are the conditions that will force us to develop new guidelines?
- 9) What is the overall conversational UI ecosystem?
- 10) Some existing contexts/domains/industries in which conversational UIs are currently being used:
 - a) Healthcare
 - b) Automotive
 - c) Ambient Intelligences
 - d) Dictation
 - e) Corn field
 - f) Smart home
 - g) Elder care/assistive technologies

TECHNOLOGICAL FACTORS

- 11) What is the impact of hardware technologies on conversational UIs?
 - a) press-and-hold
 - b) press-to-talk
 - c) wake words (“Alexa” and “Siri”)
 - d) invocation names (for applications/”skills” on Alexa platform)
 - e) automated versus handcrafted dialogue management
 - f) speech-only UI versus speech plus...screen, gesture, etc.
- 12) In speech+ interfaces, what is the precedence of modalities?
 - a) Which is primary, or default?
 - b) Is this true in all situations, or does it change?
 - c) If it changes, what are the factors that influence the change (is it context of use, content of the interaction, etc.)?
- 13) Many conversational UIs require an internet connection to function.
 - a) What are the unique error conditions that are caused by this?
 - b) How do connectivity limitations affect UI decisions?
- 14) Many conversational UIs are battery powered.
 - a) What are the unique error conditions that are caused by this?
 - b) How do battery life limitations affect UI decisions?

USERS and INTERACTION MODELS

- 15) How do users discover conversational interfaces?
 - a) How do they access the interface?
 - b) That is, what is the step before the interaction begins?

- 16) Many conversational UIs are being used by choice, rather than by necessity. In David Thomson's words, they are examples of volunteer automation, rather than victim automation (which was true for users of most speech-enabled IVRs.)
 - a) What is the impact of this difference?
 - b) How does it affect user opinions of speech technology, the specific UI, and automation in general?
 - c) How does it affect users' perception of the value of speech in a UI?
- 17) What are the unique characteristics of a volunteer user base?
- 18) Some conversational UIs have a known user base.
 - a) How does this affect the interaction?
 - b) What predictive capabilities does this enable?
- 19) What is the impact of system-initiated dialogues versus user-initiated dialogues?
 - a) Note that the Alexa platform specifically prohibits system-initiated dialogues.
 - b) How do non-speech system-initiation actions by the user (like press-to-talk) play into this?
- 20) Does the public want system-initiated dialogues? Do we all want Jarvis (the voice assistant in the *Iron Man* movies)?
- 21) What is the role of (and value of) artificially constrained conversational UIs? What can we learn from them that applies to other less constrained domains?
- 22) What is the role of experiential factors that affect users, such as emotions, persuasions, trust?

EXISTING CONVERSATIONAL UI PROFESSIONALS

- 23) Who are voice interaction designers working on conversational UIs? What is their background (are they IVR designers, or GUI designers, or something else)?
- 24) What books about conversational UI design are relevant? Who are the authors, and what is their background?