

DESIGN AND EVALUATION CRITERIA FOR CONVERSATIONAL USER INTERFACES

Kyoko Sugisaki - UZH | Andreas Bleiker - yasashii | <https://usablebots.tech>
sugisaki@ds.uzh.ch info@yasashii.ch

Visibility of system status

- 1 Is it always clear, whose turn in the conversation it is?
- 2 Does the CUI indicate clearly when it is processing (in particular: the case of response delay)?
- 3 Is it always clear that the CUI has acknowledged, understood and accepted/rejected the human partner's goal or intention?
- 4 Can the human partner clearly recognize who said what in the conversation history?
- 5 Is the CUI remembering information relevant to the current context of the user? (i.e., it does not forget answers to previous questions as long as they are potentially relevant for the conversation).
- 6 Can the human partner request details about any previous arrangement (e.g., scheduled task, alarm settings)?

Match Between System and Real World

- 1 Is the level of proactivity appropriate (e.g., offering tips and explanations)?
- 2 Are the CUI's utterances and actions truthful (i.e., providing correct information)?
- 3 Does the human-likeness, personality and identity of the CUI fit to its ability and its role? (i.e., does it set the right expectations about its actual abilities?)
- 4 Does the CUI use terminology appropriate to the knowledge of the human partner?
- 5 Is the CUI able to learn about the preferences of the human partner (e.g., terminology) and use it at later stage?
- 6 Is the CUI cooperative with the human partner for the solution of his problem and the achievement of his goal?

Error Prevention

- 1 Does the CUI master common typos, misspellings, grammatical challenges (e.g., incorrect punctuation, declinations, word order, code-switching)?
- 2 Is the CUI able to establish consensus (common ground) in case of ambiguous wording and unclear intention/goal of users?
- 3 Does the CUI ask for confirmation before executing potentially irreversible actions (e.g., permanent deletion of data)?

Help Users Recognise, Diagnose, and Recover from Errors

- 1 In the case of misunderstandings, does the CUI express clearly what type of clarification is needed by the human partner?
- 2 Does the CUI offer hints on how the human partner can help the chatbot to understand him/her (e.g., offer alternatives, ask simple clarifying questions)?
- 3 Can the human partner correct the misunderstandings of the CUI?
- 4 Can the human partner easily reverse or stop the execution of his/her command?

Recognition Rather than Recall

- 1 Does the conversation flow (the timing and rhythm of turn-taking) feel natural and engaging? e.g. appropriate waiting times, feel of immersion into the conversation?
- 2 If the chatbot expects a certain format of input: Does it offer an example of the type of answer it expects (e.g., time and date formats)?
- 3 If the human partner is supposed to make a selection: Does the CUI offer a list of options or provide examples that are helpful to make the decision?
- 4 If the human partner is asked for a final consequential decision (e.g., payment): does the CUI offer a summary of the agreement?

Flexibility and Efficiency of Use

- 1 Can advanced users use shortcuts (e.g., F1-key for help) and shorthands (e.g., '?' instead of 'what do you mean')?
- 2 Are autocomplete and autocorrect used effectively to speed up entry and reduce errors of understanding?
- 3 Are buttons and other GUI elements used in alternation with text entry for selections and choices (e.g., on mobile devices)?
- 4 Does the CUI understand synonyms and variations within its domain?
- 5 Does the CUI understand various styles of language?
- 6 Does the CUI understand the response of the human partner and handle all the provided information correctly, if he/she provides too much or little information on its original question?
- 7 Can the human partner request, at any time, that the chatbot conversation is escalated to a human conversation partner?
- 8 Can the human partner request that the conversation log can be forwarded or referred to for the purpose of continuing the conversation in another communication channel (e.g., email) or at a later stage?
- 9 Does the CUI use knowledge about the human partner (e.g., individual preferences and routines) from previous conversations with the chatbot and its environment (e.g., birth date from the platform)?

User Control and Freedom

- 1 Can the human partner decide when a conversation starts? (e.g., the chatbot does not unsolicitedly re-establish conversation to spam with "helpful" news and updates).
- 2 Can the human partner decide when a conversation ends?
- 3 Can the human partner determine the pace of the conversation? (e.g., take a break and continue at a later stage without the CUI forgetting relevant parts of the conversation and without being inappropriately timed out?)
- 4 Can the human partner steer the conversation topic and its focus (i.e., to initiate, change, continue or close a topic)?
- 5 Can the human partner access, change and delete or overrule any (explicitly or implicitly stored) preferences that the chatbot uses to personalise the conversation? (privacy and transparency)

Help and Documentation

- 1 Is help context-sensitive and on request?
- 2 Does the CUI provide the appropriate amount of information about itself (e.g., its identity, abilities, competencies and responsibilities)?
- 3 Does the CUI appropriately provide information about how to interact more efficiently with it (e.g., shortcuts, abbreviations)?
- 4 Does the CUI handle clarifying questions (e.g., terminology, its capabilities) in the middle of the task/topic?
- 5 Once clarification is completed, can the conversation be continued seamlessly?

Aesthetic and Minimalist Design

- 1 Is the length of the utterances of the CUI appropriate?
- 2 Are the utterances of the CUI relevant for the context and the user's goal?
- 3 Does the CUI handle generic and off-topic/off-domain requests (e.g., small talk) appropriately? (i.e., does the way it handles small talk fit to its personality and role?)
- 4 Does the CUI communicate in an appropriate, pleasant and polite manner?
- 5 Does the CUI respond gracefully, if the human partner are not polite (e.g., insults and sexual abuse)?
- 6 Does the CUI follow a minimal structure of conversation (e.g., greeting at the opening and closing of conversation, self identification) without forcing the human partner to follow this structure (e.g., requiring to greet the chatbot before being able to start a topic)?

Consistency and Standards

- 1 Are the utterances of the CUI orthographically and grammatically correct and consistent?
- 2 Does the CUI use domain-specific terminology consistently and correctly?
- 3 If the CUI uses a specific style of language (e.g., the use of emojis, mixed language, socio-demographically specific language, dialects, local variety, humour, second language): Does the CUI not only produce this style in its utterances, but does it also understand it in the utterances of the user?
- 4 Does the CUI follow conventions, guidelines and best practices of the environment in which it is integrated (e.g., Facebook Messenger, WeChat, Alexa, Google Assistant etc.)?
- 5 Does the CUI follow established conversational conventions of other chatbots (e.g., to get help or clarifications? Are common shortcuts used (incl. emojis)?)

Something missing?

