From Usability Testing and Text Analysis to User–Response Criticism

Florentina Armaselu
florentina.armaselu@uni.lu
University of Luxembourg, Luxembourg

Catherine Emma Jones
catherine.jones@uni.lu
University of Luxembourg, Luxembourg

Introduction

Studies on user experience in the digital medium are often related to Human-Computer Interaction (HCI) and the construction of user models or the performance of usability tests in order to support design and evaluation of digital artefacts. User modelling research has mainly focused on the construction of "usable" and "useful" tools providing the users with "experiences fitting their specific background knowledge and objectives" (Fischer, 2001: 65). A variety of characteristics have been used to inform such models, from demographic information (age, gender, native language) or relevant experience (novice, advanced, expert), to interests, goals and plans (general interest categories, task-related objectives/sequences of actions) or contextual information (location, time, physical environment) (Sosnovsky and Dicheva, 2010: 33-34). Many of the approaches merge cognitive science and artificial intelligence (Webb et al., 2001; Biswas and Robinson, 2010; Mohamad and Kouroupetroglou, 2013), whilst usability testing, as a technique from user-centred design, often involves the iterative refinement of a prototype based on user's feedback (Massanari, 2010). Usability studies also evaluate how a tool is actually used (Brown and Hocutt, 2015) exploring constructs such as ease of use and learnability. Other researches, from the fields of philosophy of technology or digital hermeneutics, go beyond the usefulness and usability aspects of the technology, trying to address questions related to the "human, social, cultural, ethical, and political implications of those technologies" (Fallman, 2007: 296) or to the "self-interpretation of human beings" (Capurro, 2010: 10) in the light of the code. Further directions of study propose a re-orientation of the HCI as "an aesthetic field" (Bertelsen and Pold, 2004: 23) or a cultural perspective on the "reflexive relationship between user and medium" as a "remediation" of the self (Bolter and Grusin, 2000: 230), considered as "humanistic HCI" (Bardzell and Bardzell, 2016).

This article tries to bridge the fields of HCI and Digital Humanities (DH), where HCI techniques are used to evaluate tools developed in DH projects and the results of this evaluation are analysed via DH methods, with the intention of potential development inspired by the literary theory of aesthetic response (Iser, 1980). The paper elaborates on previous work (Armaselu and Jones, 2016) and presents two case studies of usability tests conducted within the framework of interface and game design for digital historical editions and digital cultural heritage (Section 2). Section 3 describes the type of analysis applied to users' responses, whereby we propose a typology of users and interpretation of users' experience, followed by conclusions and future work (Section 4).

Two case studies

The first case refers to the design and implementation of an XML-TEI-based platform (Transviewer) allowing the exploration of digital editions of historical documents through features for page-by-page navigation, side-by-side view (facsimile/transcription), free-text and named entities search. The usability tests inspired by previous studies (Nielsen, 2000; Jones and Weber, 2012) involved a user-group of 8 researchers in history, political science and linguistics, 4 males and 4 females, aged 25-64. They had to complete 17 tasks using the prototype and to fill-in a USE-based questionnaire (Lund, 2001). During the experiments, the users were asked to think-aloud and the audio and screen interactions were recorded. The common language was English, although none of the participants was a native speaker.

The second case uses data collected during three sessions of gameplay conducted as part of a requirements gathering and co-design process for Pilot 4 of the H2020 Crosscult project. Players were asked to play a board game and contribute reflections as they encountered historical objects pinned to various locations in the city (the Board was derived from a map of Luxembourg City). The first session contained 6 players (5 females, 1 male), the second 5 players (all female) and the third 5 players (4 males, 1 female). All players, aged 25 to 50, worked in a research environment, and none of the participants used English as

his/her mother tongue. In the first session, players had 10 roles of the dice to score as many points as possible in successive rounds, the game was shortened so players had to score the most points and reach the end of a score board first.

Analysing user response

For both cases, the users' responses from the questionnaires were transcribed, when not already in electronic form. Partial transcription of the think-aloud audio recordings for the first case was performed (reflections on the experience, improvement suggestions, expressions of disorientation or frustration); the transcription of the second case video-audio recordings is not yet completed, therefore not included in the study. The transcribed snapshots were pre-processed (TXT, XML, R) according to the formats required by the analysis phase. Three types of software were used: Textexture – a tool for representing the text as a network (Paranyushkin, 2011); TXM – a statistical tool for corpus analysis; TheySay – a sentiment analysis package.

The first experiment with Textexture drew attention to noteworthy connections between different clusters of meaning related to users' experience as expressed in their responses. Figure 1 presents two examples: the first highlights how the notion of *trust* is related to the side-by-side view feature of the interface, as allowing the users to compare the transcription with the scanned original and make sure it can be trusted (left); the second illustrates the linking of the sub-networks for *player (reflection, discussion, exchange, opinion)*, *place (location, malta, luxembourg)* and *story (card, map, point)*, which reveals the relations, at a conceptual level, between the significant features and interactions of the game.
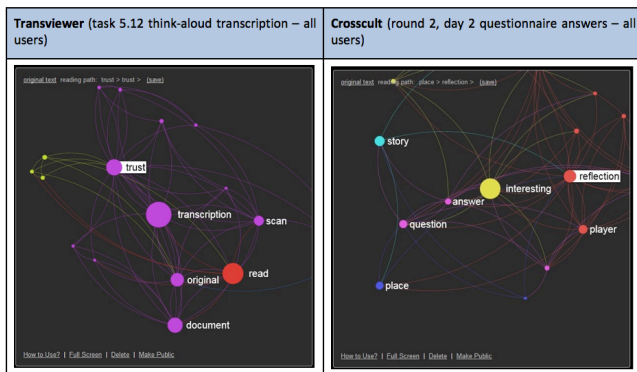

Figure 1.Textexture

TXM allowed contrasting the specificities scores (Lafon, 1980) corresponding to each user, in terms of overuse/deficit of words usage, as compared to the rest of the corpus. Table 1 shows the positive/negative specificities diagrams based on these measures for three groups of linguistic features. The scores above/under a banality threshold (+/-2.0) indicate highest specificity for responses from particular types of respondents, which allowed us to make hypotheses about a potential "typology" of users that can be described within both case studies.

Table 1. TXM: User-response specificities

For instance, some users are characterised by an overuse of *I, my* or *you, your*, others by an alternation of them, which can create the impression of an "immersive", "distant" or "versatile" point of view: "Which *I* found strange. Yes, *I* have not yet used the big arrow buttons", "if *you* scroll, *you* have to scroll both" (Transviewer); "prefer to elaborate *my* own answer, without influence", "*I* think it triggers *your* own thinking process" (Crosscult). Similarly, the use of conditionals, negations and uncertainty adverbs are suggestive of a "sceptical" user, in contrast to experiences described with appreciative adjectives and superlatives indicative of an "enthusiastic" standpoint.

After exploring the results in TXM and identifying possible types of users, we analysed the responses via TheySay (Table 2).

	Transviewer (think aloud transcription)	Crosscult (questionnaire answers)
Immersed	Positive (**0.446**, 0.129, 0.425, 2163)	Positive (**0.533**, 0.082, 0.385, 253)
Sceptical	Positive (**0.451**, 0.163, 0.386, 2591)	Positive (**0.646**, 0.097, 0.256, 206)
Enthusiastic	Positive (**0.596**, 0.119, 0.286, 733)	Positive (**0.721**, 0.121, 0.158, 261)

Table 2.TheySay: overall and polarity scores (positive, neutral, negative, word count)

The results enabled us to explore differences in sentiment between the types of users. For example, the "enthusiastic" user from both experiments scores highly with respect to the measure of positive polarity, whilst the sceptical user scores are a bit lower but, interestingly enough, higher than the immersed user's.

It was also observed that sometimes, irrespective the type of user, sentences with high score for humour may actually point to interaction-related aspects like disorientation, confusion, contrariety: "I was … where was I?", "I clicked on people but I don't know what happened" (scores 0.996 and 1, Transviewer); "I've never been in the flow because I can't focus on other gamers", "didn't use any, but I don't think I would" (scores 0.996 and 1, Crosscult).

Conclusion and future work

The paper describes two case studies in interface and game design dealing with the application of textual analysis to user-response via three systems, for visualisation of the text as a network (Textexture), corpus analysis (TXM), and sentiment analysis (TheySay). The research is still in progress and more experiments with new cases are expected to further support, test and validate the proposed user typologies and interpretation modalities, which might in the future inform humanistic interface design and approaching of user models. In addition, we expect to explore the theoretical matters, assuming that this kind of analysis, beyond its usability-oriented value, may inspire new paths of reflection on user's self-projection in the digital space, at the intersection of digital hermeneutics, digital aesthetics, and the theory of literary response.

Bibliography

Armaselu, F., Jones, C.E. (2016). "Towards a Digital Hermeneutics? Interpreting the User's Response to a Visualisation Platform for Historical Documents." DHBenelux 2016: Conference Abstracts, Belval, Luxembourg. http://www.dhbenelux.org/wp-content/uploads/2016/05/106_Armaselu-Jones_FinalAbstract_DHBenelux_long.pdf.

Bardzell, J., Bardzell, S. (2016). Humanistic HCI. interactions 23, 2 (February 2016), 20-29. DOI=http://dx.doi.org/10.1145/2888576. http://interactions.acm.org/archive/view/march-april-2016/humanistic-hci.

Bertelsen, O.W., Pold, S. (2004). "Criticism as an approach to interface aesthetics." In NordiCHI '04, October 23-27, 2004 Tampere, Finland, ACM 1-58113-857-1/04/10, pp. 23-32. http://www.interactiondesign.us/courses/taught/2010_AD590/pdfs/Bertelsen_2004.pdf.

Biswas, P., Robinson, P. (2010). "A brief survey on user modelling in HCI." In Proceedings of the International Conference on Intelligent Human Computer Interaction (IHCI) 2010. http://www.cl.cam.ac.uk/~pr10/publications/ihci10.pdf.

Bolter, J.D., Grusin, R. (2000). Remediation: Understanding the New Media, The MIT Press, 1999, paperback edition 2000.

Brown, M.E., Hocutt, D.L. (2015). "Learning to Use, Useful for Learning: A Usability Study of Google Apps for Education." JUS, Journal of Usability Studies, Vol. 10, Issue 4, August 2015, pp. 160-181. http://uxpajournal.org/usability-study-google-apps-education/.

Capurro, R. (2010). "Digital Hermeneutics: An Outline." In AI & Society, 2010, 35 (1), 35-42. http://www.capurro.de/digitalhermeneutics.html.

Fallman, D. (2007). "Persuade Into What? Why Human-Computer Interaction Needs a Philosophy of Technology." In Persuasive 2007. Yvonne de Kort et al. (Eds.). Heidelberg, Springer, pp. 295-306.

Fischer, G. (2001). "User Modeling in Human–Computer Interaction." In User Modeling and User-Adapted Interaction, 11: 65.doi:10.1023/A:1011145532042, pp 65–86. http://link.springer.com/article/10.1023/A:1011145532042.

Iser, W. (1980). The Act of Reading: A Theory of Aesthetic Response, The John Hopkins University Press, 1978, paperbacks edition 1980.

Jones, C., Weber, P. (2012). "Towards Usability Engineering for online Editors of Volunteered Geographic Information: A perspective on learnability." Transactions in GIS 16(4).

Lafon P. (1980). "Sur la variabilité de la fréquence des formes dans un corpus." Mots N°1, pp. 127-165. http://www.persee.fr/doc/mots_0243-6450_1980_num_1_1_1008.

Nielsen, J. (2000). "Why You Only Need to Test with 5 Users." NN/g, Nielsen Norman Group, Evidence-Based User Experience Research, Training, and Consulting. https://www.nngroup.com/articles/why-you-only-need-to-test-with-5-users.

Massanari, A. L. (2010). "Designing for imaginary friends: information architecture, personas and the politics of

user-centered design." In *New Media & Society*, 12(3) 401–416. DOI: 10.1177/1461444809346722, SAGE. http://nms.sagepub.com/content/12/3/401.full.pdf.

Mohamad, Y., Kouroupetroglou, C. (2013). "User modeling", Research and Development Working Group Wiki, last modified on 10 May 2013, at 05:07. https://www.w3.org/WAI/RD/wiki/User_modeling.

Paranyushkin, D. (2011). "Identifying the Pathways for Meaning Circulation using Text Network Analysis", Nodus Labs. Published October 2011, Berlin. http://noduslabs.com/research/pathways-meaning-circulation-text-network-analysis/.

Sosnovsky, S., Dicheva, D. (2010). "Ontological technologies for user modelling", Int. J. *Metadata, Semantics and Ontologies,* Vol. 5, No. 1, pp.32–71. http://www.dfki.de/~sosnovsky/papers/IJMSO_5_1_Paper_03.pdf.

USE Questionnaire: Usefulness, Satisfaction, and Ease of use. Based on: **Lund, A.M.** (2001). "Measuring Usability with the USE Questionnaire", STC Usability SIG Newsletter, 8:2.
http://garyperlman.com/quest/quest.cgi?form=USE.

Webb, G.I., Pazzani, M.J., Billsus, D. (2001). "Machine Learning for User Modeling." In User Modeling and User-Adapted Interaction, 11:19-29, 2001. Kluwer Academic Publishers, Printed in the Netherlands. http://www.umuai.org/anniversary/webb-umuai-2001.pdf.