Ergocrowd: apport de l'ergonomie à un outil informatique conçu pour l'ergonomie

Carole Baudin, Laura Maillard, Gaëtan Bussy

► To cite this version:

Carole Baudin, Laura Maillard, Gaëtan Bussy. Ergocrowd: apport de l'ergonomie à un outil informatique conçu pour l'ergonomie. Actes de la 28ième conférence francophone sur l'Interaction Homme-Machine, Oct 2016, Fribourg, Suisse. pp.291-298, 10.1145/3004107.3004141 . hal-01384340

HAL Id: hal-01384340

<https://hal.science/hal-01384340>

Submitted on 19 Oct 2016

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Ergocrowd : apport de l'ergonomie à un outil informatique conçu pour l'ergonomie

Carole Baudin

Responsable du Groupe de
Compétences « Conception de
produits centrée utilisateurs »
HE ARC Ingénierie
Espace de l'Europe 11
2000, Neuchâtel, Suisse
carole.baudin@he-arc.ch

Laura Maillard

Groupe de Compétences
« Conception de produits centrée
utilisateurs » - HE ARC Ingénierie
Espace de l'Europe 11
2000, Neuchâtel, Suisse
laura.maillard@he-arc.ch

Gaëtan Bussy

Groupe de Compétences
« Conception de produits centrée
utilisateurs » - HE ARC Ingénierie
Espace de l'Europe 11
2000, Neuchâtel, Suisse
gaetan.bussy@he-arc.ch

Résumé

Cet article présente les prémices des résultats du projet de recherche Ergocrowd, premier concept de solution semi-automatisée d'évaluation ergonomique des interfaces web. Nous présentons ici l'apport de l'ergonomie dans ce projet, en particulier dans le recentrage de l'objectif initial, le traitement des défis ergonomiques en jeu et la méthodologie appliquée pour y répondre. Nous concluons sur les résultats finalement obtenus et les projections pour la suite du projet.

Mots Clés

Ergonomie d'interface ; eye-tracking ; utilisabilité ; navigabilité ; analyse crowd

Abstract

This article presents the primary results of the research project Ergocrowd, which is a first solution concept half automated of ergonomics evaluation for web interface. We present here what ergonomics has brought in this project, especially by refocusing the initial goals, addressing the ergonomics challenges involved and adducing the applied methodology in order to answer it. We will conclude with the results we finally obtained and the projections about the future of the project.

Author Keywords

Ergonomics of interface; eye-tracking; usability; navigability; crowd analysis

ACM Classification Keywords

Miscellaneous

© ACM, 2016. This is the author's version of the work. It is posted here by permission of ACM for your personal use. Not for redistribution. The definitive version was published in Actes de la 28ème conférence francophone sur l'Interaction Homme-Machine, 2016.
<http://dx.doi.org/10.1145/3004107.3004141>

Figure 1 : Concept initial Ergocrowd (© Nabil Ouerhani – HE Arc Ingénierie)

Introduction

En 15 ans, l'ergonomie des interfaces est devenue une discipline incontournable pour la mise en place de solutions interactives efficaces, en particulier d'interfaces web. Pourtant, aujourd'hui encore, elle est pratiquée de manière empirique par les développeurs et webdesigners. Or, la qualité des sites web est devenue essentielle dans les stratégies commerciales et économiques des entreprises. Elles font donc intervenir de plus en plus des experts ergonomes externes pour obtenir un diagnostic ergonomique plus poussé. Les interventions ergonomiques alors mises en place sont souvent longues et coûteuses, elles se réalisent en situation expérimentale et doivent, pour être viables, étudier des données issues d'un grand nombre de testeurs. En se basant sur ce constat, l'équipe de Technologie d'Interaction de la Haute Ecole Arc Ingénierie, a fait appel à notre équipe spécialiste en ergonomie pour co-développer le dispositif Ergocrowd, premier concept de solution semi-automatisée d'évaluation des interfaces web. Rendre leur analyse ergonomique quantitative, reproductible et moins coûteuse, tel est l'enjeu du projet répondant ainsi aux besoins des entreprises mais aussi des ergonomes dans la facilitation des analyses à grande échelle.

Cette contribution présente les étapes de travail que nous avons menées, depuis le point de vue ergonomique, pour développer avec nos collègues informaticiens un proof-of-concept du dispositif Ergocrowd.

Objectif initial du dispositif ErgoCrowd

L'objectif initial du projet Ergocrowd était de développer une plateforme « cloud » permettant l'analyse ergonomique quantitative et systématique des sites web en impliquant le « crowd-sourcing » pour un déploiement à grande échelle. Plus particulièrement, il s'agissait de rendre les évaluations plus robustes par l'intermédiaire de données quantitatives, et plus

fiables, en réalisant les tests non pas en situation expérimentale, mais en situations dites écologiques, donc chez les testeurs. Pour ce faire, l'architecture proposée du dispositif se compose donc des données (ou métriques) recueillies via un eye-tracker, un clavier, une souris, une webcam, et sauvegardées sur un cloud pour être analysées et permettre un diagnostic ergonomique du site web (Fig. 1). L'ambition initiale était donc d'automatiser l'expertise ergonomique, ce à quoi nous avons proposé quelques ajustements, notamment car l'expertise en soi ne peut être automatisée mais facilitée.

1^{ère} étape : recentrer les objectifs du projet

De l'évaluation ergonomique autonome au recueil de données automatisé

• Définition des termes

En ergonomie, l'évaluation d'un site web recouvre une gamme variée de pratiques : depuis les non-ergonomes qui font souvent de la « mauvaise ergonomie » en se basant sur des critères de base établis et facilement trouvables sur internet, jusqu'aux experts qui réalisent, selon les demandes, de véritables études d'utilisabilité des sites. L'utilisabilité est la capacité d'un système à être utilisé de manière efficace, efficiente et satisfaisante par des utilisateurs identifiés dans des contextes d'utilisation spécifiés (Norme ISO 9241-11). Elle peut se définir selon différents critères (Fig. 2) dépendant du type d'interventions ergonomiques : en amont ([1] Norme ISO 9241-11), pendant ([2] Bastien et Scapin), en aval du développement d'un site web ([3] Nielsen). Les ergonomes qui réalisent ce type d'études doivent mettre en place divers scénarios de test et utilisent une approche mixte (qualitative et quantitative) pour étudier de manière approfondie l'utilisabilité d'un site selon les objectifs du mandataire et de l'interface testée. Cela demande une observation des activités et des comportements (opératoires, gestuels, verbaux) des utilisateurs, que les ergonomes réalisent la plupart du temps

Figure 2: Critères de l'utilisabilité

en laboratoire de test et en temps réel. Ils utilisent aussi des données quantitatives (données clavier, souris, eye-tracking) leur permettant d'asseoir leurs observations. Nous avons donc proposé que le dispositif Ergocrowd puisse pré-traiter les données quantitatives nécessaires aux diagnostics ergonomiques sans toutefois enlever l'expertise nécessaire des ergonomes. En parallèle, suite à divers entretiens avec de potentielles entreprises clientes, on a constaté que la plupart des demandes étaient liées non pas tant à des études d'utilisabilité, mais plutôt à de simples études de navigabilité des sites web. L'enjeu du dispositif Ergocrowd devenait donc de réaliser des études fiables et pertinentes de navigabilité en situations écologiques, facilitant le diagnostic expert.

• Etat de l'art

La navigabilité implique l'étude des typologies de stratégies d'exploration et de recherches d'informations visuelles des utilisateurs. Ceci peut être étudié par l'intermédiaire d'un eye-tracking. De ce fait, une étude plus poussée a été réalisée sur la corrélation de données eye-tracking avec des interprétations ou règles ergonomiques. L'étude bibliographique (Fig.3, gauche) montre bien qu'il n'existe pas de règles linéaires ou binaires entre métriques et interprétation mais plutôt des heuristiques. Par exemple, une longue fixation du regard mesurée par l'eye-tracker peut à la fois signifier une difficulté à extraire une information ou un grand intérêt pour l'objet. C'est pourquoi ces mesures objectives du comportement oculaire ne peuvent en elles seules permettre une interprétation sur le comportement cognitif. Il est alors nécessaire de prendre en compte l'interprétation d'autres données comme les verbalisations, les variables liées au typologie de testeurs, les perceptions des testeurs, etc.. Ce constat venait valider notre proposition de recentrage de l'objectif initial : Ergocrowd ne doit pas avoir comme ambition de rendre autonome l'expertise, car

l'interprétation de l'ergonome est incontournable, toutefois Ergocrowd doit pouvoir automatiser les métriques relatives à la navigabilité de manière à faciliter le diagnostic de l'expert. En ce sens, nous avons proposé que le dispositif puisse non seulement se baser sur des métriques issues de l'eye-tracker ou mouse-tracker mais aussi intégrer des données qualitatives par l'intégration de questionnaires en amont puis en aval du test de navigabilité, permettant ainsi de recueillir en amont, les données liées aux variables des participants; puis en aval des données sur leur perception post-test de la navigabilité d'un site web.

Du théorique à la pratique : les frontières de l'outil

Tout au long de l'étude, nous avons été accompagnés par un expert ergonome des interfaces qui nous a précisé ses pratiques. Cet aspect était important dans la mesure où il existe un décalage entre les données théoriques identifiées comme étant signifiantes dans l'interprétation ergonomique et les pratiques des ergonomes. Par exemple, le temps de fixation n'est en soi pas une donnée signifiante pour l'ergonome, en revanche, le temps de première fixation sur une zone d'intérêt (ou Area of Interest : AOI) permettant d'identifier les aspects de « découvrabilité » d'un site web ou encore le pourcentage de participants fixant une zone d'intérêt, permettant de préciser la « saillance » d'une zone sont des données directement exploitables pour les ergonomes. Nous avons donc identifié de cette manière les métriques signifiantes dans l'expertise en croisant état de l'art, pratique des ergonomes et demandes des mandants (voir fig. 3 page suivante, à droite).

Deuxième étape : les tests utilisateurs

Une fois le premier recentrage effectué et les métriques identifiées, le processus de développement du premier démonstrateur Ergocrowd a pu s'initier. La méthodologie de

type agile mise en place a permis de procéder par itération pour, dans un premier temps, valider les métriques, et identifier les problématiques liées aux conditions écologiques (test 1), puis expérimenter des solutions pour maîtriser ces problématiques et aborder les aspects de traitement des données massives (test 2).

• Test 1

Ergocrowd est constitué d'un dispositif d'eye-tracker low-cost et simple d'utilisation (l'EyeTribe ®) et d'un logiciel Opensource de recueil des données (Ogama). La version d'Ogama a progressivement été améliorée pour pouvoir mettre : (i) en place un questionnaire pré-test sur les données utilisateurs, (ii) introduire des consignes de test, (iii) recueillir

et traiter les métriques issues de l'eye-tracker, et de la souris, (iv) réaliser un questionnaire post-test. Sur une première version de ce dispositif, des tests ont été menés avec 26 testeurs pour valider les métriques mais aussi identifier les problématiques liées à la réalisation de tests en conditions écologiques. Le test 1 s'est fait en condition semi-réaliste, c'est-à-dire que lors de la calibration, les testeurs pouvaient ajuster leur position et le matériel pour que leur regard soit bien détecté par l'eye-tracker, comme en conditions réelles. Les données enregistrées ont permis d'identifier une problématique centrale : la fiabilité des données. En effet, moins de 8% des participants ont obtenu une qualité de calibration « perfect¹ ». Ce test nous a alors permis de mettre différentes phases d'un test eye-tracking et qui sont

Fig. 3 : Synthèse des analyses réalisées pour l'état de l'art. **A gauche**, analyse bibliographique, montrant les multiples relations des métriques aux interprétations (les articles référencés de 4 à 22 sont en biographie). **A droite**, les métriques finalement retenues, au croisement des données liées aux types de demandes, aux pratiques des ergonomes et à l'analyse bibliographique. (Schémas réalisés par C. Lecomte).

¹ Trois types de calibration ont été définis : « perfect », « good », « redo », pour laquelle le testeur devait refaire la calibration

fondamentales dans la fiabilité des données : position du testeur, du matériel, ambiance lumineuse et sonore, éléments perturbateurs. Seule une consigne contraignante lors des tests peut alors résoudre ce problème en tentant de maîtriser ces variables. Ce premier test nous a permis d'élaborer une consigne-type (formulation, typologie de tâche demandée, actions à mener par le testeur) et un protocole de test simplifié expérimenté dans un deuxième test.

- **Test 2**

Le protocole du test 2 était donc plus cadré : poste de travail figé, fiche de recommandations donnée au testeur (position recommandée, actions à éviter comme les raccourcis clavier), une consigne plus précise et plus courte que lors du test 1 a été mise en place. 80% des participants ont ainsi obtenu une calibration de qualité « perfect », et donc des données eye-tracking beaucoup plus fiables. Les recommandations sur la consigne et le protocole de test ont donc été validées. Le traitement des données a ensuite pu être abordé.

Dès le test 1, les métriques obtenues par le dispositif ont été validées, toutefois leur présentation permettant un pré-diagnostic et interprétation de l'expert ergonomiste était encore vague. Le traitement de données massives est une problématique en soit. L'idée de réaliser un outil (« dashboard ») permettant à l'expert ergonomiste de manipuler la somme des données obtenues a émergé dès le test 1. Toutefois, la nécessité de pouvoir filtrer les données, l'identification des types de filtres pertinents et surtout les modes de visualisations des données traitées ont été travaillées lors du test 2. Les résultats sont présentés dans le chapitre suivant.

Résultats préliminaires

- **1^{er} résultat : validation de métriques**

Comme nous l'avons présenté précédemment, l'identification et validation des métriques ergonomiques constitutives d'Ergocrowd a été le fruit d'un travail mixte, entre état de l'art, consultation des entreprises, entretiens d'experts et tests.

Outre des données liées à la performance globale (par exemple, le temps passé à réaliser la consigne avec succès), nous avons identifié une liste de données eye-tracker les plus révélatrices des comportements cognitifs de l'utilisateur permettant d'évaluer la navigabilité d'un site. Ces comportements regroupent essentiellement des stratégies d'exploration et de recherche d'information dans des zones d'intérêt spécifiées ou « AOI » (Area of Interest) que l'expert doit préalablement identifier. Nous avons pu classer les métriques selon plusieurs niveaux pour aller d'une analyse globale à fine:

- Métriques sur le test entier : temps passé, nombre de pages visitées, chemin de page emprunté (pagepath), chemin de clic (clickpath), chemin d'AOI (scanpath)
- Métriques sur une page spécifique : temps passé, nombre de clic, nombre de scroll, etc.
- Métriques sur une AOI spécifique : durée avant 1^{ère} fixation (time to 1st fixation), durée avant 1^{er} clic, etc.

Les différentes hypothèses-types utilisées par les experts pour évaluer la navigabilité d'un site web ont pu être traduites en termes de « métriques » objectives. Ce travail a été poussé jusqu'à établir des exemples d'analyses heuristiques pour permettre aux experts d'interpréter plus directement les données restituées par Ergocrowd.

- **2^{ème} résultat : recommandations sur le protocole de test**

Des recommandations ont été faites sur le protocole de test pour le testeur et pour l'expert, notamment sur une consigne-type réaliste. Malgré les consignes données sur la mise en place du test en situations écologiques, les dimensions

Figure 4 :
Dashboard (1/vérification de la fiabilité des données, 2/filtrage des données, 3/visualisation des données)

contextuelles représentent encore des problématiques fortes et importantes car de cette étape dépend la fiabilité des données. Ainsi, le choix d'un eye-tracker « low-cost », élément principal du dispositif physique, doit encore être précisé afin de convenir à un utilisateur lambda et à tout type d'environnement. D'autres perspectives telles que la conception d'un support d'eye-tracker, et d'un dispositif de mesure du niveau de luminosité et sonore sont encore à explorer.

• 3^{ème} résultat : conception d'un dashboard

Le traitement et la restitution des données de tests se devait d'être la plus efficiente possible afin de faciliter le travail de l'ergonome dans le traitement de données fiables, maniables sur des données massives et facilement communicables pour les entreprises clientes. C'est pourquoi le « dashboard » (fig. 4) a été particulièrement étudié. Il a été pensé en trois étapes :

1. Vérification de la fiabilité des données
2. Filtrage des données (en fonction des métriques, du contexte d'étude et des caractéristiques individuelles)
3. visualisation des données (relatives à un individu (scanpath, mousepath, replay webcam) et relatives à un groupe d'individus (heatmap, statistiques))

Des visuels spécifiques pour la visualisation de données de masse ont été travaillées (crowd-click-path notamment) ainsi que la souplesse de l'interface pour faciliter le travail de l'ergonome (voir poster).

Conclusions et projections

L'interface proposée est encore au stade de démonstrateur. En ce sens, d'un point de vue technique, si beaucoup de problèmes ont pu être résolus, il reste encore des améliorations à réaliser pour pouvoir être un outil exploitable par et pour les ergonomes (ex : gestion des sites dynamiques). L'interface globale permettant la réalisation des tests, ainsi

que le dashboard, ont été travaillés techniquement mais l'ergonomie et le design d'interface sont perfectibles.

Des tests sont encore à réaliser sur une grande échelle de testeurs pour ajuster le dashboard dans ses aspects de visualisation. Toutefois les premiers tests ont montré que l'application de filtres personnalisables (origine géographique, aspect générationnel, etc.) sur la dimension « crowd » dans les évaluations de sites permettront d'obtenir des informations liées aux aspects culturels extrêmement pertinentes et novatrices pour les ergonomes. Par exemple, lors du test 1 nous avons pu voir dans les résultats l'influence des dimensions générationnelles sur les comportements de navigation. La génération Z (14-25 ans) avait tendance à passer moins de temps sur le site et à visiter davantage de pages (en moyenne 14s/page) comparé à des individus de la génération X (40-50 ans) qui ont visité moins de pages mais y ont passé plus de temps (24s/page). Le dispositif Ergocrowd a un intérêt certain pour les ergonomes, dans l'interprétation des données massives qui pourraient avoir un intérêt culturel. Finalement, nous pouvons affirmer qu'Ergocrowd se profile comme étant un outil d'aide à l'évaluation ergonomique puissant pour les ergonomes qui pourrait ouvrir d'autres perspectives scientifiques s'il devenait un système expert se basant sur un système intelligent permettant à chaque itération d'apprendre les heuristiques générées.

Remerciements

Ce projet a été financé grâce aux fonds interdisciplinaires de la HES-SO et est mené par le groupe Technologies d'interaction de la Haute Ecole Arc Ingénierie. Nous souhaitons remercier particulièrement Nabil Ouerhani, chef de projet, David Grunenwald, Claudia Gheorghe, Andreea Mihet et Chloé Lecomte. Nous remercions aussi chaleureusement Julien Roland de Uservalue, pour sa précieuse contribution tout au long du projet.

Bibliographie

1. Norme ISO 9241-11 (1998)
2. Bastien et Scapin, *Critères ergonomiques pour l'évaluation d'interfaces utilisateurs*. Rapport technique INRIA n° 156, Juin 1993, INRIA : Le Chesnay
3. Nielsen, *Usability Engineering*. New Jersey: A P Professional, 1993
4. Goldberg, H. J., & Kotval, X. P. (1999). Computer interface evaluation using eye movements: Methods and constructs. *International Journal of Industrial Ergonomics*, 24, 631-645.
5. Poole, A., Ball, L. J., & Phillips, P. (2004). In search of salience: A response time and eye movement analysis of bookmark recognition. In S. Fincher, P. Markopolous, D. Moore, & R. Ruddle (Eds.), *People and Computers XVIII—Design for Life: Proceedings of HCI 2004*. London: Springer-Verlag Ltd.
6. Just, M. A., & Carpenter, P. A. (1976). Eye fixations and cognitive processes. *Cognitive Psychology*, 8, 441-480
7. Mello-Thoms, C., Nodine, C. F., & Kundel, H. L. (2004). What attracts the eye to the location of missed and reported breast cancers? In *Proceedings of the Eye Tracking Research and Applications Symposium 2002* (pp. 111-117). NY: ACM Press.
8. Hauland, G. (2003). Measuring team situation awareness by means of eye movement data. In *Proceedings of HCI International 2003: Vol 3* (pp. 230-234). Mahwah, NJ: Lawrence Erlbaum Associates
9. Cowen, L., Ball, L. J., & Delin, J. (2002). An eye-movement analysis of web-page usability. In X. Faulkner, J. Finlay, & F. Détienne (Eds.), *People and Computers XVI—Memorable yet Invisible: Proceedings of HCI 2002* (pp. 317-335). London: SpringerVerlag Ltd.
10. Byrne, M. D., Anderson, J. R., Douglas, S., & Matessa, M. (1999). Eye tracking the visual search of clickdown menus. *Proceedings of CHI'99* (pp. 402-409). NY: ACM Press.
11. Albert, W. (2002). Do web users actually look at ads? A case study of banner ads and eye-tracking technology. In *Proceedings of the Eleventh Annual Conference of the Usability Professionals' Association*.
12. Goldberg, J. H., Stimson, M. J., Lewenstein, M., Scott, N., & Wichansky, A. M. (2002). Eye tracking in web search tasks: Design implications. In *Proceedings of the Eye Tracking Research and Applications Symposium 2002* (pp. 51-58). NY: ACM Press.
13. Sibert, J. L., Gokturk, M., & Lavine, R. A. (2000). The Reading Assistant: Eye gaze triggered auditory prompting for reading remediation. In *Proceedings of the Thirteenth Annual ACM Symposium on User Interface Software and Technology* (pp. 101-107). NY: ACM Press.
14. Vuori, T., Olkkonen, M., Pölönen, M., Siren, A. & Häkkinen, J. (2004). Can Eye Movements Be Quantitatively Applied to Image Quality Studies? In *NordiCHI '04 Conference proceedings*, ACM Press, pp. 335-338.
15. Hendrickson, J. J. (1989). Performance, preference, and visual scan patterns on a menu-based system: Implications for interface design. In *Proceedings of CHI'89 Human Factors in Computing Systems Conference* (pp. 217-222). NY: ACM Press.
16. Altonen, A., Hyrskykari, A., & Räihä, K. (1998). 101 Spots, or how do users read menus? In *Proceedings of CHI'98 Human Factors in Computing Systems* (pp. 132- 139). NY: ACM Press.
17. Fitts, P. M., Jones, R. E., & Milton, J. L. (1950). Eye movements of aircraft pilots during instrument-landing approaches. *Aeronautical Engineering Review*, 9, 24-29
18. Renshaw, J. A., Finlay, J. E., Tyfa, D., & Ward, R. D. (2003). Designing for visual influence: An eye tracking study of the

- usability of graphical management information. In M. Rauterberg, M. Menozzi, & J. Wesson (Eds.), *Proceedings of Human-Computer Interaction – INTERACT 2003* (pp. 144-152). Amsterdam: IOS Press.
19. Jacob, R. J. K., & Karn, K. S. (2003). Eye tracking in human-computer interaction and usability research: Ready to deliver the promises (Section Commentary). In J. Hyona, R. Radach & H. Deubel (Eds), *The mind's eye: Cognitive and applied aspects of eye movement research* (pp. 573-605). Amsterdam: Elsevier Science.
20. Marcos, M.-C., Garcia-Gavilanes, R., Batatineh, E., & Pasarin, L. (2013). Using Eye Tracking to Identify Cultural Differences in Information Seeking Behavior. Paper presented at the CHI'13, Paris, France.
21. Dong, Y., & Lee, K. P. (2008). A cross-cultural comparative study of users' perceptions of a webpage: With a focus on the cognitive styles of Chinese, Koreans and Americans. *International Journal of Design*, 2(2), 19-30
22. Chua, H. F., Boland, J. E., & Nisbett, R. E. (2005). Cultural variation in eye movements during scene perception. *Proceedings of the National Academy of Sciences of the United States of America*, 102(35), 12629-12633.
23. Rouet Jean-François, Ros C., Jégou G., Metta S., « Chercher des informations dans les menus web : interaction entre tâche, type de menu et variables individuelles », *Le travail humain* 4/2004 (Vol. 67), p. 377-395, URL : www.cairn.info/revue-le-travail-humain-2004-4-page-377.htm. DOI: 10.3917/th.674.0377.
24. Barnum, Carol M. *Usability testing essentials: ready, set...test !*, Amsterdam : Elsevier, 2011. 382p.