Proceedings of the Pilot4HWID Workshop on Pilot Implementation for Testing Human-Work Interaction Designs

Online and Bari, Italy, August 31, 2021

In conjunction with **INTERACT2021**

Edited and organized by

Morten Hertzum
Torkil Clemmensen
Barbara Rita Barricelli
Pedro F. Campos
Frederica Gonçalves
José Abdelnour Nocera
Ganesh Bhutkar
Arminda Guerra Lopes

Welcome

Welcome to the Pilot4HWID workshop on pilot implementation for testing human-work interaction designs. These workshop proceedings contain the eleven papers presented at the workshop. The copyright to the papers remains with their authors.

Pilot implementations are field tests of properly engineered, yet unfinished, systems. In contrast to lab tests, the users in a pilot implementation use the system for performing real work. In contrast to full-scale implementations, the objective of a pilot implementation is to learn. The Pilot4HWID workshop aims to (a) help mature this technique for evaluating human-work interaction designs during the process of their development and implementation, (b) collect case studies that analyze experiences with conducting and learning from pilot implementations, and (c) formulate a research agenda for future work on pilot implementations – addressing their strengths, limitations, conduct, and impact. Further information about the workshop is available in the workshop description (Hertzum et al., 2021), which is part of the INTERACT2021 proceedings.

The workshop is organized by Working Group 6 – Human Work Interaction Design – under IFIP TC13 (https://hwid.unibs.it/). The organizers are

Morten Hertzum
University of Copenhagen, Denmark

Torkil Clemmensen
Copenhagen Business School, Denmark

Barbara Rita Barricelli
Università degli Studi di Brescia, Italy

Pedro F. Campos
ITI/LARSyS and University of Madeira, Portugal

Frederica Gonçalves
ITI/LARSyS and University of Madeira, Portugal

José Abdelnour Nocera
University of West London, UK, ITI/LARSyS and University of Madeira, Portugal

Ganesh Bhutkar
Vishwakarma Institute of Technology, India

Arminda Guerra Lopes
Polytechnic Institute of Castelo Branco, ITI/LARSyS, and University of Madeira, Portugal

Reference

Hertzum, M., Clemmensen, T., Barricelli, B.R., Campos, P.F., Gonçalves, F., Nocera, J.A., Bhutkar, G., & Lopes, A.G. (2021). Pilot implementation: Testing human-work interaction designs. In *Proceedings of the INTERACT2021 Conference on Human-Computer Interaction*. Switzerland: Springer Nature.

Table of Content

Dropping a bomb or providing a gentle loving touch? Towards a relation artefact theory of pilot implementation 4
T. Clemmensen

AttnGAN: Realistic text-to-image synthesis with attentional generative adversarial networks 10
S. Mathesul, G. Bhutkar, A. Rambhad

Prototype design of a multi-modal AI-based web application for hateful content detection in social media posts 16
T. Pradhan, B. Padamwar, G. Bhutkar, A. Pangaonkar, S. Pareek, K. Partani

Driver-vehicle interaction analysis: Pilot implementation for driver behavior classification using smartphone sensor data 23
P. Wawage, Y. Deshpande

University digital engagement of students 32
M. C. Pereira, J. C. Ferreira, S. Moro, F. Gonçalves

Pilot implementation in virtual reality: Pitfalls, tensions and organizational issues 36
C. França, M. Sousa, F. Gonçalves, É. R. Gouveia, B. Gouveia, P. Campos

Socio-technical and collaborative design of pilot implementations of digital twins 41
B. R. Barricelli

Co-designing workshop as pilot implementation for complex workplaces 46
P. Saadati, J. Abdelnour-Nocera, T. Clemmensen

Mental happiness - A pilot implementation for elderly people 53
A. G. Lopes, I. Cerejo

Usability evaluation of mobile app to support road safety 58
S. Salve, G. Bhutkar

Pilot implementation: Organizational alignment when implementing an IT-system 63
L. K..Herbæk, C. E. D. Hansen, T. Clemmensen

Dropping a bomb or providing a gentle loving touch? Towards a relation artefact theory of pilot implementation

Torkil Clemmensen[1][0000-0002-0934-2336]

[1] Copenhagen Business School, Howitzvej 60, DK-2000, Frederiksberg, Denmark
tc.digi@cbs.dk

Abstract. This position paper is about socio-technical interventions in pilot implementation contexts. It argues that human work interaction design provides massive push towards such interventions. It does so through theorizing the continuous relation-building between empirical work analysis and interaction design activities that creates new local solutions for the stakeholders involved. The question is how hard of soft that this push should be.

Keywords: Sociotechnical user experiences, digital work legacy, interaction design interoperability, human work interaction design

1 Introduction

This paper is about pilot implementation understood as socio-technical interventions in organizational and wider contexts. Pilot implementation is a notion that stems from IT health research [2] and thus allude to medical science epistemology of 'effect'-driven intervention [9]. Today it is however a broader concept that captures the moment in design when the wider organizational and beyond context is involved in the design [7,8,16,20] and has recently been established as (part of) 'organizational implementation' that aims at organizational change [6].

This paper argues that human work interaction design (HWID) approach provides massive push towards such interventions, through the continuous relations-building between empirical work analysis and interaction design activities that creates new local solutions for the stakeholders involved. How hard should this push be: should pilot implementation be as dropping a bomb or providing a gentle loving touch?

2 Pilot implementation as a movement from the technical towards the social

Socio-technical HCI design approaches can be used for different directions of socio-technical interventions. Sometimes technology is assumed to lead to social changes, (e.g., [12]), sometimes social changes are assumed to change technology, (e.g., [19].

For example, the practice based computing approach [22] is mostly interested in the empirical effects of new technology such as ongoing work interaction design for learning and development, while the experience design approach [5] build on theoretical knowledge about human psychology to come up with novel interaction designs. The HWID platform (e.g., this workshop) support any of the directions for socio-technical interventions. However, the HWID platform is very focused on the socio-technical relations themselves.

A socio-technical approach to human-automation collaboration would take participatory and co-design approaches seriously across HCI, CSCW, IS, UI engineering, and technical psychology, and all the way through the lifecycle of an artefact. A requirement to begin doing that is to have an open mind about how to link the social and the technical also when it comes to pilot implementation.

3 HWID Relation Artefacts

HWID relation artefacts are simply IT artefacts that relates empirical work analysis and interaction design (Clemmensen, in preparation). They are different from 'relational artefacts', that is, social robots [21] and other anthropomorphic interfaces, which are things that people relate to. HWID relation artefacts relate the work analysis with the interaction design. Furthermore, HWID relation artefacts may impose a certain order in time and space of the design. In this paper, the relation artefacts are: interaction interoperability checkups, digital legacy interventions, and organizational strategy alignments, **Fig. 1**. They present a move from the technical to the social.

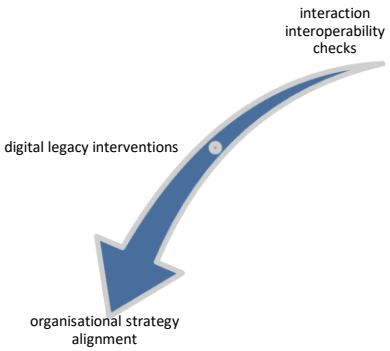


Fig. 1. Relation artefacts for socio-technical interventions

The following proposed relation artefacts for interventions begins with the technical interaction designs and moving towards organizational and social interventions in workplaces.

Interaction interoperability checkups. Interaction interoperability checkups are relation artefacts that aim to increase the UX related to interoperability of interaction designs. The interoperability of interaction designs has been studied as continuity in

multi-device interactions where interactions move, or transition, from one device to another. A first checkup on the interoperability of novel interaction designs may focus on UX of sequential multi-device use. [18]. Second, interoperability has been studied as socio-technical interoperability of HCI in work domains with multiple workers and multiple devices collaborating. Kwon et al. [14] suggested that socio-technical interoperability of HCI in work domains concerns (1) sharedness, (2) readiness, (3) awareness, (4) adaptiveness, and (5) coupledness of the multiple workers and multiple devices collaborating. To assess these dimensions would be a second checkup. Third, cross-validations of the proposed solution's interoperability could be done by evaluating it in the lab, the field, and the gallery [13], which all may contribute to increases in novel knowledge and confidence in the interoperability checks of the novel interaction designs. When doing interaction interoperability checks and other design interventions not any single of the three evaluation approaches are the correct one, all should probably be used in a triangulation manner.

Digital legacy interventions. Digital legacy interventions are socio-technical interventions that in some way change (improve, hopefully) the proposed interaction designs relation to the organization's legacy systems, for example by increasing their UX. To do this requires reconceptualizing what we mean by organizational legacy systems and what are interaction designs for digital legacy. It raises questions such as how employees' experience their organization's business-critical but obsolete systems, and what can employees do about their own digital legacy in their organization. The socio-technical view of UX of legacy systems is then that they are socio-technical systems that are technically and/or socially obsolete, old, and need lots of maintenance, but solve problems for organizations and meet individual employees' needs [11]. Creating new relations between work/organization analysis and interaction design is what solves legacy issues, not simply software modernization. Furthermore, employee's digital legacy is from a HWID perspective central for digital legacy interventions. Users' digital legacy can be defined as "the meaningful and complex way in which information, values, and possessions are passed on to others" [4]. Employees could benefit from organizational owned add-ons, plugins, and data scrapers that could support doing legitimate extracting from organizational storage and other places where employees generate and collect personal digital data and support them in transferring their digital legacy to private storage or to their next employers' storages.

Organizational strategy alignments. A factor for successful organizations is the close linkage of the IT strategy and business strategy [1]. Strategy as practice [10] tells that it is important how management practices are used to put strategy into practice. Pilot implementation thus should entail relation artefacts that are about morphing interaction design for human work morphed into organizational strategies, for example by aligning the organizational UX culture with the business and organizational goals. The activities that IT managers engage in to ensure they are in the room when important business decisions about product direction and business strategy are made have been identified by UX leaders from industry and by researchers [3,15]. The ac-

tivities include broader questions of developing and managing a UX culture in the organization. UX leaders see UX strategy at the corporate level as being about that the UX team are aligned with the overall goals and objectives of the business. They aim to shape the strategic plans, operational needs, and interdependencies between their own organization and the rest of the company, to and to increase UX team's effectiveness and synergies with other business functions. They see UX strategy at the level of a business unit as being about plans for delivering products, systems, or services that offers a high value to customers, and differentiates the company's brand. However, this requires multiple parts of the organizations to be involved [15]. Thus, UX strategy alignment has to be done within a UX organizational culture that can support the strategy and make it realistic and ensure it has an impact on company outcomes [3].

4 Dropping a bomb or providing a gentle loving touch?

First-stage intervention = {relation artefact type intervention a}
IF evaluation = {nonresponse}
THEN second-stage intervention = {intensify relation artefact type intervention a}
ELSE IF evaluation = {response}
THEN at second stage = {continue with relation artefact type intervention b}

Fig. 2. Decision rules in adaptive interventions in socio-technical design. The sequence is continued with relation artefacts type intervention c, and repeated as long as it takes to reach closure with type intervention c. Adapted from [17].

In HWID, like in most HCI design, evaluations are most of the time formative evaluation with the purpose of improving the design. This implies that there is a systematic and sequential overlap between construction and evaluation/ intervention activities. In HWID pilot implementation the sequence of artefact designs/actions and evaluations thereof can perhaps be conceptualized as an 'adaptive intervention'. Adaptive intervention is a method proposed in psychology to allow greater individualization and adaptation of intervention options (i.e., intervention type and/or strength) over time [17]. Adaptive interventions in socio-technical design are thus a string of different relation artefacts that are evaluated to adapt to workers' and organizations' characteristics and changing needs over time, with the general aim to optimize the long-term effectiveness of the overall socio-technical intervention. **Fig. 2** shows how decision rules can be used to operationalize adaptive interventions with relation artefacts.

In conclusion, this paper's answer to the title question is to do pilot implementation with adaptive interventions, which is perhaps closer to the gentle loving touch than to dropping a bomb. Secondly, the paper proposes the following: (1) HWID relation artefacts, (2) a movement from the technical to the social, and (3) the notion of adaptive interventions, as possible elements towards a theory of pilot implementation as relation artefacts used for adaptive sociotechnical HCI design interventions.

5 References

[1] Walter Baets. 1992. Aligning information systems with business strategy. *J. Strateg. Inf. Syst.* 1, 4 (1992), 205–213.

[2] Jørgen P Bansler and Erling Havn. 2010. Pilot implementation of health information systems: Issues and challenges. *Int. J. Med. Inform.* 79, 9 (2010), 637–648.

[3] Liam Friedland. 2019. Culture eats UX strategy for breakfast. *interactions* 26, 5 (2019), 78–81.

[4] Rebecca Gulotta, William Odom, Jodi Forlizzi, and Haakon Faste. 2013. Digital Artifacts as Legacy: Exploring the Lifespan and Value of Digital Data. In *CHI '13*. DOI:https://doi.org/10.1145/2470654.2466240

[5] Marc Hassenzahl. 2010. Experience Design: Technology for All the Right Reasons. *Synth. Lect. Human-Centered Informatics* 3, 1 (2010), 1–95.

[6] Morten Hertzum. 2021. Organizational Implementation: The Design in Use of Information Systems. *Synth. Lect. Human-Centered Informatics* 14, 2 (2021), i–109.

[7] Morten Hertzum, Jørgen P Bansler, Erling C Havn, and Jesper Simonsen. 2012. Pilot implementation: Learning from field tests in IS development. *Commun. Assoc. Inf. Syst.* 30, 1 (2012), 20.

[8] Morten Hertzum, Maria Ie Manikas, and Arnvør á Torkilsheyggi. 2019. Grappling with the future: The messiness of pilot implementation in information systems design. *Health Informatics J.* 25, 2 (2019), 372–388.

[9] Morten Hertzum and Jesper Simonsen. 2010. Effects-driven IT development: an instrument for supporting sustained participatory design. In *Proceedings of the 11th Biennial Participatory Design Conference*, 61–70.

[10] Paula Jarzabkowski. 2004. Strategy as practice: recursiveness, adaptation, and practices-in-use. *Organ. Stud.* 25, 4 (2004), 529–560.

[11] Ravi Khadka, Belfrit V. Batlajery, Amir M. Saeidi, Slinger Jansen, and Jurriaan Hage. 2014. How do professionals perceive legacy systems and software modernization? In *Proceedings of the 36th International Conference on Software Engineering - ICSE 2014*, 36–47. DOI:https://doi.org/10.1145/2568225.2568318

[12] Rob Kling. 2000. Learning about information technologies and social change: The contribution of social informatics. *Inf. Soc.* 16, 3 (2000), 217–232.

[13] Ilpo Koskinen, Finland Thomas Binder, and Johan Redström. 2008. Lab, field, gallery, and beyond. *Artifact J. Des. Pract.* 2, 1 (2008), 46–57.

[14] Gyu Hyun Kwon, Tonya L Smith-Jackson, and Charles W Bostian. 2011. Socio-cognitive aspects of interoperability: Understanding communication task environments among different organizations. *ACM Trans. Comput. Interact.* 18, 4 (2011), 1–21.

[15] Florian Lachner, Philipp Naegelein, Robert Kowalski, Martin Spann, and Andreas Butz. 2016. Quantified UX: Towards a Common Organizational Understanding of User Experience. *Proc. 9th Nord. Conf. Human-Computer Interact. - Nord. '16* (2016), 56:1--56:10.

DOI:https://doi.org/10.1145/2971485.2971501

[16] Troels Mønsted, Morten Hertzum, and Jens Søndergaard. 2020. A Socio-Temporal Perspective on Pilot Implementation: Bootstrapping Preventive Care. *Comput. Support. Coop. Work* 29, 4 (2020), 419–449.

[17] Inbal Nahum-Shani, Min Qian, Daniel Almirall, William E Pelham, Beth Gnagy, Gregory A Fabiano, James G Waxmonsky, Jihnhee Yu, and Susan A Murphy. 2012. Experimental design and primary data analysis methods for comparing adaptive interventions. *Psychol. Methods* 17, 4 (2012), 457.

[18] Dimitrios Raptis, Jesper Kjeldskov, and Mikael B. Skov. 2016. Continuity in Multi-Device Interaction. In *Proceedings of the 9th Nordic Conference on Human-Computer Interaction - NordiCHI '16*. DOI:https://doi.org/10.1145/2971485.2971533

[19] Nick Seaver. 2018. What should an anthropology of algorithms do? *Cult. Anthropol.* 33, 3 (2018), 375–385.

[20] Arnvør á Torkilsheyggi and Morten Hertzum. 2014. User participation in pilot implementation: Porters and nurses coordinating patient transports. In *Proceedings of the 26th Australian Computer-Human Interaction Conference on Designing Futures: the Future of Design*, 290–299.

[21] Sherry Turkle, Will Taggart, Cory D Kidd, and Olivia Dasté. 2006. Relational artifacts with children and elders: the complexities of cybercompanionship. *Conn. Sci.* 18, 4 (2006), 347–361.

[22] Volker Wulf, Claudia Müller, Pipek Volkmar, David Randall, Markus Rohde, and Gunnar Stevens. 2015. Practice-Based Computing: Empirically Grounded Conceptualizations Derived from Design Case Studies. In *Designing Socially Embedded Technologies in the Real-World*, Volker Wulf, Kjeld Schmidt and David Randall (eds.). Springer London, London, 111–150. DOI:https://doi.org/10.1007/978-1-4471-6720-4_7

AttnGAN: Realistic Text-to-Image Synthesis with Attentional Generative Adversarial Networks

Shubham Mathesul, Ganesh Bhutkar, Ayush Rambhad
Vishwakarma Institute of Technology, Pune, India
shubham.mathesul19@vit.edu

Abstract. In this paper, we propose a prototype design for manifold refinement to fine grained text-to-image generation by using Attentional Generative Adversarial Networks. We concentrate on creating realistic images from text descriptions. We have used a collection of Attentional Generative Adversarial Network layers that are able to correctly select the modal meaning at the word-level and sentence-level. Generative Adversarial Networks (GANs) prove to be fundamental structure for many design applications from Game design, Art, Science and Modelling applications. Our prototype is easy to implement and practical; choosing the most relevant word vectors and using those vectors to generate related image sub-regions. With due consideration to findings of usability testing, the develpment team in future iterations of the application, hopes to improve the generated image resolution. They plan to provide a choice for created variety of images with further improvements to the image generation algorithm.

Keywords: GAN, Text-to-Image Synthesis, Artificial Intelligence, Artificial Neural Networks, DAMSM, Attentional Generative Adversarial Networks.

1 Introduction

The objective of the research work was to use Generative Adversial Networks (GAN) to create photorealistic images based on text descriptions provided by the user. Understanding the relationship between visual information and natural languages is a critical first step toward Artificial Intelligence. Although GAN has had impressive results over the years, it still lacks the ability to intelligently select the modal meaning at the word-level and sentence-level[3]. To solve this issue, we propose the use of Attentional Generative Adversarial Networks(AttnGAN) which provides a manifold refinement to fine grained text-to-image generation process.

Automatic generation of image designs based on natural language descriptions is a primary task in many domains, including art generation, design industry and computer-aided design. It also propels research in multimodal learning and inference through vision and language, which has been one of the most active research fields in recent years[3]. The use of AttnGAN for automatic generation of

images will help image designers to instantly create highly compelling imagedesigns as per their preference ranging from musicalbum covers, room interior designs, game posters and will solve their distress of designing an image by hand or using computer-aided designing (CAD) software, which usually consumes quite a few productive hours out of their day and many times the users are not always fully satisfied with their image design.

2 Literature Review

In the literature review, following interesting papers were studied to get useful insights into the doamin.

In the first paper, Xu et al. built a attentional generative network for the AttnGAN to use in a multi-stage cycle to produce high-quality images[3]. Their AttnGAN outperformed many previous GAN models, increasing the highest observed inception performance on the Caltech-University of California San Diego(UCSD) Birds dataset by 14.14 percent. Their attention-driven image-text matching score helped them to assess the similarity of an image-sentence pair based on an attention model between the image and the text.

In another paper, Zhu et al. proposed an approach a dynamic memory module to optimise the fuzzy image contents, where the initial images are not well generated [4]. Their technique involved choosing the important text information based on the initial image material allowed their application to produce images accurately from the text definition.

In the next paper, Lee et al. proposed a framework to solve two fundamental concerns in GANs - catastrophic forgetting of the discriminator and mode collapse of the generator[5]. The team achieved this by using GANs for contrastive learning and information maximization approach to understand the source for improvements. Using this framework, the team was able to improve the stability while training the GAN and improving GAN performance over the test dataset.

3 Research Methodlogy

3.1 Application Design

Our proposed implementation methodology is to train a convolutional neural network(CNN) trained on text features encoded by a text-encoder which provides sentence-level and individual word-level features. Both the generator network G and the discriminator network D perform feed-forward inference training on the text

features provided by the text-encoder and output a photo-realistic image as output of the network. The proposed Attentional Generative Adversarial Network (AttnGAN) constitues of two neoteric units: the attentional generative network and the deep attentional multimodal similarity model. To understand the details of the text-to-image generation process, the application architecture of AttnGAN is shown in the Fig.1 below.

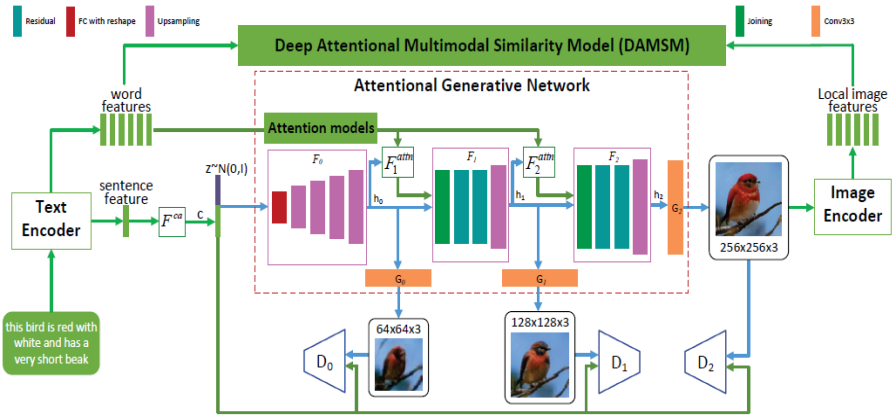


Fig. 1. AttnGAN application architecture [3]

3.2 Attentional Generative Network

Current GAN-based models for text-to-image generation usually encode the entire sentence-text definition into a single vector as the condition for image generation, but lack fine-grained word-level detail[3]. We propose to draw all the different image sub-regions, based on keywords that are most important to each suchsub-regions. With reference to Fig.1. above, the proposed attentional generative network has *m* generators*(G0,G1, ...,Gm−1)*, which take the hidden states*(h0, h1, ..., hm−1)* as input and generate images of small-to-large resolutions*(ˆx0, ˆx1, ..., ˆxm−1)*. Here, *z* is often a noise vector taken from a normal standard distribution. $\bar{e}$ is a vector of a global sentence, *e* is the module of vectors of words. The attention model is given the word features and image features as an input from the preceding hidden layer and then we compute a word-context vector for each image sub-region depending on its hidden features. To get the final realistic image design after multiple iterations, the final function of attentional generative network generates a balance of two expressions: Firstly, GAN loss where the unconditional loss decides whether the image is realistic or fake and the conditional loss determines whether or not the image and the text description fit. Second, the Deep Attentional Multimodal Similarity Model(DAMSM) computed image-text matching loss, which will be detailed in the following section.

3.3 Deep Attentional Multimodal Similarity Model

The DAMSM employs artificial neural networks that map image sub-regions with word features in the sentence to a shared space, which helps in measuring text-image similarity at the individual word-level to compute the loss for image generation. The DAMSM loss matches the complete image with a sequence of words forming the text description.

3.4 Evaluation Metrics

Since the inception score cannot on its own indicate whether the produced picture is well dependent on the provided text description, we suggest that we employ R-precision as a supplemental evaluation metric for text-to-image synthesization. We find the first *R* relevant images related to the query, then we test the top results of the relevant images and then we find that *r* images are actually relevant and hence we calculate the R-precision as *r/R*. To extract the global vectors from the produced pictures and the provided text descriptions, first the image and text encoders learnt in our DAMSM are used. Then, similarities between global vectors of images and global text vectors are calculated. Finally, for each picture we classify candidate text descriptions with decreasing similarity and find the top r descriptions of R-precision calculations.

4 Usability Testing

Usability testing is performed to capture the user experience of the AttnGAN-the developed application prototype with its users and to get the suggestions for effective application design of next version. Due to pandemic situation in India, the development team created a questionnaire to collect the user's quantitative responses using the Likert scale [8]. Different types of users including game designers, art designers, computer aided design(CAD) designers, albumcover designers and interior designers were provided with the online application and were asked their response/feedback about the quality, quantity of the generated image, their opinion on using artificial intelligence for image designing along with their suggestions for further improvement in the designed proptotype. About 11 users including 8 male and 3 female users were involved in this pilot testing [7].

The responses received in this pilot testing showed that the respondent users were satisfied with the image generated from the AttnGAN project. About 80% of the respondents were keen on using the application in future to assist their domains and would also recommend it to their peers. Although many of the respondents were happy with the quality of the generated image, they implored that an improvement in the quality/resolution of images would help them even more with their designs. There were mixed opinions with the variety of images that the application provided as each image provided a distinct design which some users preferred while the

others deferred. They appealed for the option to choose between the generated variety of images. As the prototype aims to cover varied domains like CAD image generation, musicalbum cover generation, the users felt the necessity for the prototype to be explored more with other domains other than bird images to satisfy their need for this prototype. Almost everyone agreed on the fact that artificially generated images acted as a satisfactory base for designers to further build their design upon.

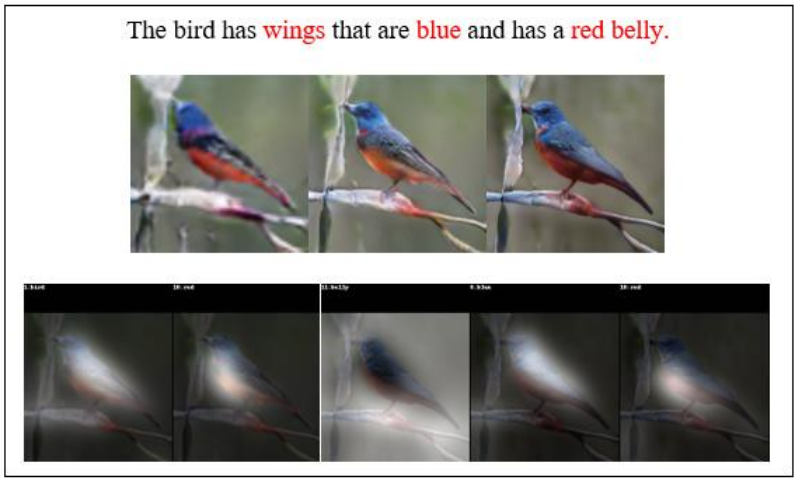


Fig. 2. Results of trained AttnGAN model

5 Results and Conclusion

We have compared the prototype AttnGAN's findings with its previous generated images to help explain what it has learned. The model reported an inception score of 12.80 and a R-precision score of 72.88% on the Caltech-University of California San Diego Birds(CUB) dataset which measured the quality and the variance in the generated image design. The initial phase of prototype emphasizes on the shape and the color of the object and then allocates the attention to the key features of the text descriptions. With the inputs from pilot testing learned that the users were content with the generated image and stated they would use it as a base to their designs if given the option to choose between the generated images and if the prototype is further improved on other image domains.

In this paper, we proposed the prototype AttnGAN for photo realistic text-to-image synthesis. We introduced a deep attentional multimodal similarity model to compute the text-image matching loss for training the generator of the AttnGAN. Experimental results clearly depicts the importance of the proposed model for text-to-image generation in ever-growing fields of science, art and design. During the

experiment, we observed that the size of training group with standard increase (flip, rotation, scale) increased, training outcomes improved up to a certain amount with the help of increased data which significantly improves the image generation algorithm. However, due to CPU and GPU limitations the current system consisted of only 3 hidden layers and trained on around 1204 images from the dataset and the resolution set to 256x256 in resolution. In future, the developer team plans to work on the learning from the user experience of pilot testing to improve resolution of generated image from the prototype in order to assist the designers with higher quality images and to scale the images to a higher resolution. They also plan to add an option for the users to choose between the varied image designs generated from the application over multiple domains.

Acknowledgement

We thank all the users / participants of usability testing of AttnGAN, in Pune, India, for their cooperation and support.

6 References

1. Campo, M.: “Architecture, Language and AI - Language, Attentional Generative Adversarial Networks (AttnGAN) and Architecture Design”. Proceedings of the 26th CAADRIA Conference, Vol 1, Hong Kong, 211-220 (2021).
2. Hashemiet, A., Mozaffari, S.: “Secure Deep Neural Networks using Adversarial Image Generation and Training with Noise-GAN”. Computers & Security, Vol 86, 372-387 (2019).
3. Tao, X., Pengchuan, Z., Qiuyuan, H., Han, Z., Zhe, G., Xiaolei, H., Xiaodong, H.: "AttnGAN: Fine-Grained Text to Image Generation with Attentional Generative Adversarial Networks". IEEE / CVF Conference on Computer Vision and Pattern Recognition, 1316-1324 (2018).
4. Minfeng, Z., Pingbo, P., Wei, C., Yi, Y.: “DM-GAN: Dynamic Memory Generative Adversarial Networks for Text-To-Image Synthesis”. IEEE / CVF Conference on Computer Vision and Pattern Recognition (CVPR), 5795-5803 (2019).
5. Hajar, E., Majid, M., Ming, D., Chinnam, R.: “SPA-GAN: Spatial Attention GAN for Image-to-Image Translation”. IEEE Transactions on Multimedia, 391-401 (2020).
6. Lee, K., Ngoc-Trung, T., Ngai-Man, C.: “InfoMax-GAN: Improved Adversarial Image Generation via Information Maximization and Contrastive Learning”. NeurIPS Workshop on Information Theory and Machine Learning, 3941-3951 (2021).
7. Hertzum, M., Bansler, J.P., Havn, E., Simonsen, J.: “Pilot Implementation: Learning from Field Tests in IS Development”. Communications of the Association for Information Systems, Vol 30 (1), 313-328 (2012).
8. Bhutkar, G., Mehetre, A., Sagale, U., Kendre, A., Bhole, D., Jathar, N.: “Heuristic Evaluation of an Eye-Free Android Application”. Proceedings of the 6th International Conference on M4D Mobile Communication Technology for Development, Kampala, Uganda, Karlstad University Studies, 206-215 (2019).
9. California Institute of Technology. “Caltech-UCSD Birds 200”. CNS-TR-2010-001. (2010). http://www.vision.caltech.edu/visipedia/CUB-200.html , accessed dated July 29, 2021.

Prototype Design of a Multi-Modal AI-Based Web Application for Hateful Content Detection in Social Media Posts

Tejas Pradhan[1], Balaji Padamwar[2], Ganesh Bhutkar[3], Aditya Pangaonkar[4], Shriram Pareek[5] and Kishan Partani[6]

Department of Computer Engineering, Vishwakarma Institute of Technology, Pune, India

tejas.pradhan18@vit.edu

Abstract. Hate Speech and hateful content is a major problem on digital platforms. The ever increasing amount of content posted daily on social media has led to an excessive amount of digital hate being spread in the form of posts, images and comments. The proposed system is developed in order to act as a tool for determining if a particular social media post is hateful and is aimed to aid any benign social media user who has been affected by hate speech and wants to report it. The proposed system uses a multimodal artificial intelligence based approach by classifying different formats of posts, i.e., images and comments or captions separately. Ensemble convolutional neural network architecture is used for this classification, thus, proving to be a strong tool for finding evidence of any prevalent hates speech. This system is tested using the Likert scale for its user interface, accuracy and utility. Based on the result this paper proposes a prototype design of a web application, which can be used for hateful content detection.

Keywords: Hate Speech, classification, convolutional neural networks, ensemble, multimodal, and interface.

1 Introduction

Social media is an extremely fast mechanism to spread information. With the increasing amount of social media usage, millions of users are sharing photos, posts and reviews everyday. While social media is a really effective tool, it has also become a channel for quick and aggressive spread of hatred among different communities. Mitigation of hate speech using Artificial Intelligence (AI) complimenting the laws of a country like India can help bolster safer social media platforms [1]. A majority of the young population is affected by offensive content online, which mainly includes teens [2]. Potential threats to teenagers and adolescents include exposure to inappropriate content, online abuse and cyber bullying [3]. Managing and curbing the spread of hateful content through social media is the need of the hour in this digital world. However, detecting hateful content is not deterministic in nature. The main reason behind this is that hateful content is extremely subjective. It is dependent on the context and community being addressed at hand.

The subjective nature of offensive content on social media has made it very difficult for autonomous intelligent systems to detect it [4]. Many existing content filtering algorithms and systems use post tagging and sentiment analysis based approaches to detect hate speech. Hateful content is not just in the form of text, but also images and text embedded in images (memes). The proposed system aims to incorporate all of these aspects into the web application and develop a suitable methodology to detect hateful content and possibly report it. It uses a hateful content detection engine which is trained on multiple hate speech tagged comments and other textual format using data from Twitter and Facebook. The engine also includes a model trained on violent images to detect any image, which is potentially offensive or inappropriate.

The focus of the proposed web application is to provide utility to users who want to report hateful content and need evidence for the same. The target user group for this system includes:

1. Victims of social media hatred,
2. Respondents, i.e., users who have been provoked by or responded aggressively to hateful content,
3. Blog or small social media forum owners.

The victims of social media hatred often find it difficult to report hateful content. Social media aggression due to online hate is another important issue at hand, which needs to be restrained. Blog owners, too, need to rely on third party content filtering mechanisms. The application will be tested for its accuracy and utility value taking the above user group into consideration. Thus, the robustness of the system will also be tested. Since the target user group involves a wide variety of users, feedback will be taken on the design using the Likert Scale [9] and suggested improvements will be incorporated into the final prototype.

2 Related Work

When dealing with the detection and classification of hate speech a lot of problems would arise, that is if the speech is universally accepted hateful or not handling bias, or dealing with unbalanced data [1]. To classify hate speech, language and words are necessary. Researchers have deployed a wide range of more sophisticated feature representations, including word n-grams, syntactic features, and distributional semantics [4]. This article within the journal considers the uncertainty; lastly that the most effective understanding of Waldron’s argument in his book is that hate speech tends to cause damage - a weak sort of the consequentialist case for its proscription. His argument isn't advanced by his apparent reliance on speech-act theory [5].

In the paper, HaterNet, an intelligent system that is presently getting used by the Spanish National workplace Against Hate Crimes of the Spanish State Secretariat for Security is conferred that identifies and monitors the evolution of hate speech in Twitter. The contributions of this analysis include many interesting aspects. It introduces the primary intelligent system that monitors and visualizes, mistreatment social network analysis techniques, hate speech in Social Media and conjointly it introduces a completely unique public dataset on hate speech in Spanish consisting of 6000 expert-labeled tweets [6]. This article within the journal aims to critique and augment posts associated with cyber hate. This would be achieved by emphasizing a particular feature of the content in contrast to alternative options, like simple access, size of audience, and namelessness, which are usually overlooked [7].

3 System Architecture

The hateful content detection system is deployed as a web application and is intended to be a detection tool for checking if a particular image or comment falls into the category of hateful content. This will be specifically useful for people who have been affected by hate speech and want evidence to report it. The input is taken in the form of text or image. This input is forwarded to the hate speech detection engine wherein we have used 1-D and 2-D convolutional neural networks for image and text classification and the cumulative means of results is used for a final classification – Hateful or Safe.

3.1 Inputs To the System

Since the nature of the proposed system is multimodal, input to the application can be given in three ways:

1. Text (Comments or Posts on Social Media)
2. Image (Photos from Social Media)
3. Image with Embedded Text (Memes)

Based on the type of given input the system either uses pure image classification mechanisms, pure text classification mechanisms or both, i.e., multimodal classification mechanisms. Prior to classification, text needs to be extracted from the image. This is done using optical character recognition. The extracted text is cleaned using multiple natural language processing techniques before testing it for hateful content. Extraction of text from images is done using Optical Character Recognition (OCR).

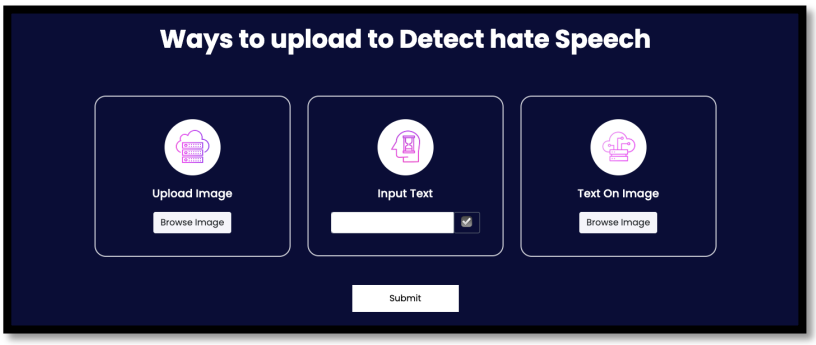


Fig. 1. Input Design of the Web Application

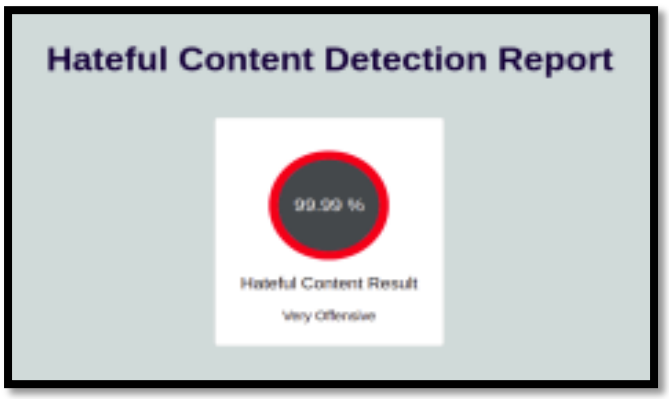


Fig. 2. Report Design of the Web Application

3.2 Detecting Hateful Content using Classification Algorithms

The classification technique follows a multi-modal / ensemble approach. In the first stage, images and text are classified separately. We have used deep neural networks for classifying text and images since they have shown to perform better than most statistical machine learning approaches for hate speech detection [5]. For image classification, the system uses a 2-D Convolutional Neural Network with 2 convolutional layers followed by max pooling layers respectively. This convolutional neural network has been trained on 1000 images of potentially hateful and non-hateful images.

For detecting hate speech from text, each comment is first cleaned, preprocessed and tokenized to make it suitable for classification. Social media data contains a lot of unnecessary characters and emoticons which are not important when it comes to classification. Text preprocessing includes removing these characters, removing stop-words (a, an, the etc.) and preserving only meaningful words. This text is then converted into word vectors using the count vectorisation approach

before finally sending it to the hate speech classifier. Text classification uses a 1-D convolutional Neural Network with each word represented as a 1-dimensional vector [8].

3.3 Output of the System

A mean of the individual classification probabilities is calculated and used for detection of hateful content. This probability is then converted into percentage and the output is given as a report to the user. The report contains the percentage of the given input to be hate-speech along with the category. The current implementation of our application can categorize violence.

4 Usability Testing

The proposed system is tested in two phases. In the first phase, users testing the system are asked to give different forms of input to the system and test its robustness. A diverse group of users - including victims or witnesses of hateful content and cyber-bullying have been asked to test the system by feeding in different images, posts and sentences of hate speech which they had previously encountered. This group includes 8 males and 4 females. The average age range is 17-22 with 2 users above the age of 45 years. They are asked to fill a detailed questionnaire, which highlights different aspects of the system including design, accuracy of classification and overall user experience. The feedback of a preliminary group of 12 users is recorded using 'Likert scale' [10].

Current data from the testing suggests that the users found this web application prototype to be useful and more than 90% users were happy with the UI. About 16% i.e. 2 out of 16 users faced difficulty in navigating to the output and these improvements are to be incorporated into the system. Another point highlighted after testing is that the classification models are susceptible to a high false negative rate. These classification models will be re-trained based on the feedback given by users. In the second phase, final testing of the pilot implementation [9] will be completed before deploying it into production.

5 Results and Conclusion

On testing with validation data from the dataset, the image classification obtained an accuracy of 75.8%, which is slightly lesser than the training accuracy of 87%. The text classification neural network achieved an accuracy of 95% on the training data and 92% on the test data. Results from usability testing also indicate that the image classification engine is biased towards hateful content. The reason for differences in the image

and text classification can be attributed to the datasets that both classifiers were trained on. We are currently working on developing a more generalized and robust classification system. The testing results also show that users prefer the design of such apps to be concise and simple with minimum amount of navigation to be needed. Color-coding of results with percentage of hate speech has also been an effective means of displaying the report as tested by the users. The input sections for the app can be completely isolated from each other and their respective reports can be displayed on a separate page specifically for results. This will reduce the number of clicks and make the user experience better in the latter versions of this application. The latter versions can also feature a more detailed classification report with in-depth analysis of the given input.

Acknowledgement

We thank all the users / participants of usability testing of Prototype Design of a Multi-Modal AI Based Web Application for Hateful Content Detection, in Pune, India, for their cooperation and support.

References

Kovács, G., Alonso, P. and Saini, R.: "Challenges of Hate Speech Detection in Social Media." SN COMPUT. SCI. 2, 95 (2021)

Percentage of teenagers who have encountered hate speech on social media platforms (2018):https://www.statista.com/statistics/945392/teenagers-who-encounter-hate-speech-online-social-media-usa/

Social Media and Teens, American Academy of Child and Adolescent Psychiatry:https://www.aacap.org/AACAP/Families_and_Youth/Facts_for_Families/FFF-Guide/Social-Media-and-Teens-100.aspx

Joni Salmine, Maximilian Hopf, Shammur A. Chowdhury,Soon-gyo Jung,Hind Almerekhi, Bernard J. Jansen : "Developing an Online Hate Classifier for Multiple Social Media Platforms". Published in Human Centric computing & information Sciences. **10,** Article1 (2020):

Barendt, E.: "What Is the Harm of Hate Speech?" Published in Ethical Theory and Moral Practices **22,** 539–553 (2019)

Pereira-Kohatsu JC, Quijano-Sánchez L, Liberatore F, Camacho-Collados M. "Detecting and Monitoring Hate Speech in Twitter." Published in Sensors 2019; 19(21):4654 (2019)

Alexander Brown: "What is so Special about Online (as compared to Offline) Hate Speech?" Ethnicities, Volume:18, Issue:13 (2017)

Kiranyaz S. , Avci O. , Abdeljaber O., Ince T. , Gabbouj M. , Inman D.J. : "1D Convolutional Neural Networks and Applications – A Survey" Published in Mechanical systems and Signal Processing (2021)

Hertzum, M., Bansler, J.P., Havn, E., Simonsen, J.: "Pilot implementation: Learning from field tests in IS development. Communications of the Association for Information Systems" 30(1), 313-328 (2012)

Bhutkar, G., Mehetre, A., Sagale, U., Kendre, A., Bhole, D., Jathar, N.: "Heuristic

Evaluation of an Eye-Free Android Application." 6th International Conference on M4D Mobile Communication Technology for Development, 206-215 (2018).

Driver-Vehicle Interaction Analysis: Pilot Implementation for Driver Behaviour Classification Using Smartphone Sensor Data

Pawan Wawage[1], Dr. Yogesh Deshpande[2]

Vishwakarma University, Pune, India
pawan.wawage-026@vupune.ac.in
yogesh.deshpande@vupune.ac.in

Abstract
Driving is viewed as one of the most complex task, as the driver is involved in several other tasks apart from only driving. The primary task which a driver should be performing is to safely drive a vehicle and focus on driving only. But, there are several secondary tasks which he/she has to perform at the same time. For example, handling the steering wheel and operating on the controls mounted on dashboard and steering wheel, operating on the pedals (brake, accelerator, and clutch) and at the same time shifting the gears as required, and so on. Researcher's and Scientists found it difficult to model the naturalistic driving behaviour. In this paper we discuss the need of driver behaviour analysis and the method using smart phone sensor data for better visualisation of driving behaviour and estimating driver behaviour patterns.

Keywords: Driving behaviour, Fuel Efficiency, Control Area Network (CAN), On-Board Diagnostics (OBD), ADAS, SVM, and MLA.

1 Introduction

Modeling driver behaviour has been the focus of many research studies, either for analyzing driving pattern, or awareness for safe driving. The goal of these studies is to identify the effect of driving pattern and other affecting driving behaviour through their computational model. Considering driving is a complex task characterized by a wide variety of inner and outer variables, and it has been proved that the majority of road accidents are caused by human faults such as traffic rule violations, distraction, inattention, drowsiness, fatigue, etc. Thanks to the progress in data analysis domain over the years, it is now possible to explore this area. The development of these approaches improved the quality of driving pattern and driver behaviour analysis and opened the door for new fields of applications [5].

However, we observed that, there is no standard model proposed in the literature and found lack of a unified framework for analyzing driver behaviour. We attempted to capture a set of quantitative and qualitative factors that are essential in the analysis of driver behaviour. These factors are either driver related, driving events executed, or

traffic conditions. These factors are the result of the study and literature survey that brings together the purpose of the study using different models, derives the data used, and different factors taken into account relative to driver behaviour.

In this paper, we are going to present a literature review, need for driver behaviour analysis, methods used for analysis, and the vital inner and outer factors affecting the driver behaviour. The remaining of this paper is organized as the following; we firstly discuss the research method we followed, then we present the driver behaviour classification based on smart phone sensor recorded parameters. Later, we introduce the experimental setup and the results of our study.

2 Literature Survey

Driving Behaviour while executing driving events gives the analysis of how he/she is driving including aggressively driving by over-speeding, sudden braking, sudden lane changing and sharp turning. Monitoring driving is important and necessary as it increases driver performance, efficiency and safety with less fuel consumption and servicing costs. It results in safe driving and lesser road accidents.

According to the survey report by Ministry of Petrol and Natural Gas about 70% of diesel and 99.6% of petrol are consumed in the transport sector alone. Out of the total diesel sale, the highest consumption is by cars, utility vehicles and three wheelers [13].

According to the data available on Open Government Data Platform India or data.gov.in, there has been a fearful increase in accidental deaths due to cars on Indian roads since 2016. In [7] the analysis shows that during the last ten years, deaths and injuries on roadways due to driver inattention or as a result of unintended maneuvers have increased by 5%.

In [9] how driver behaviour analysis is leading to the increase in the demand for smart infrastructure solutions like Connected Vehicle Technology (CVT), Vehicle to Infrastructure communication (V2I) and enabling the applications with live traffic data, and forward collision warning is discussed. The Connected Vehicle Technology will make aware the drivers with traffic conditions on the go resulting in safe driving.

The survey results in the need to address the aggressive driving behaviour and road rage incidents like speeding in heavy traffic, lane cutting, sudden braking of the drivers during short-term and long-term driving [1]. The efforts to find the risk posed by aggressive driving by considering the behavioral and emotional factors of a driver will result in decreasing road accidents and frequency of insurance claims. The outcome of this research would help the motor insurance companies to assess the driving

risk more accurately and to propose a solution to calculate the personalized premium based on the driving behaviour with most importance towards prevention of risk.

Analysis shows that the distribution of road accidental deaths and injuries in India varies according to age, gender, vehicle type and time of the day. Figure 1 presents the distribution of causes of road accidents in India. It can be observed that the drivers fault is considered to be the most important factor responsible for road accidents. Drivers fault resulting in highest of others with 79% of total accidents. Within the category of drivers fault, accidents caused due to over speeding, aggressive driving, sharp turning accounted for a high share [8].

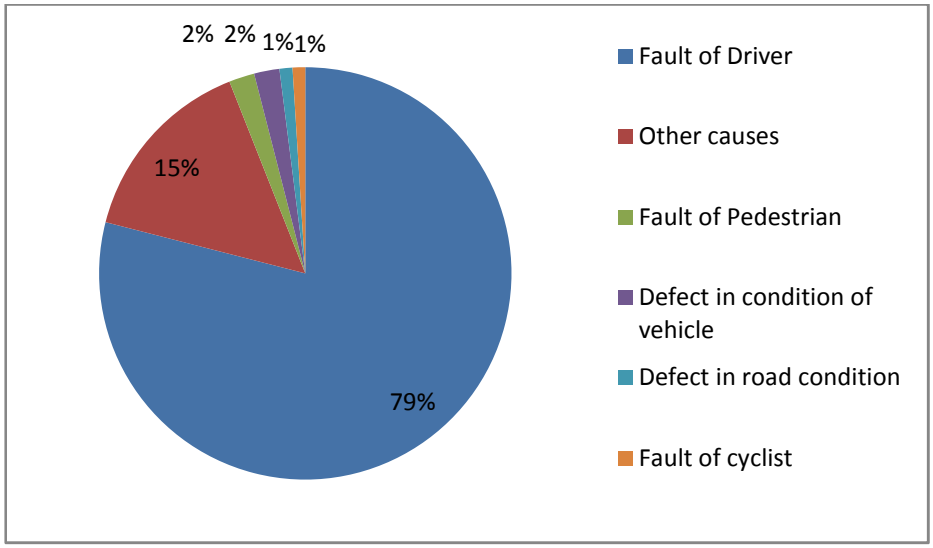


Fig.1 Causes of road accidents

Table 1 gives a summary of research done on driver behaviour classification, the purpose of the research, the data used for the analysis, and the algorithm used for analysis.

Purpose	**Data**	**Algorithm**
Driver Behaviour Recognition	Smartphone sensors	AI and ML
Intelligent Driving Style Analysis Systems	Accelerometer, Gyroscope, Magnetometer, and GPS sensor data	Classification algorithm
Driver Behaviour profiling	Telematics boxes	Fuzzy Logic
Modeling Driver distraction	Monitoring cameras	Hidden Markov Model
Recommendation systems	OBD-II parameters,	Support Vector Machines

to prevent accidents	Weather data, Traffic data	
Saving fuel consumption	OBD-II parameters	Dynamic Time Wrapping
Usage-Based Insurance	OBD-II parameters, smart phone sensor data.	Artificial Neural Networks

Table-1: Classification of research done on diver behaviour

3 Methodology

Many literatures have provided the study on various methods on driver behaviour analysis and the focus was on driver-oriented applications, with its three main sub-applications: Accident prevention, Driving pattern assessment, and Driver behaviour prediction [6]. The methods are classified according to the objective which might be one of the above sub- applications, and their input factors taken into consideration for the analysis phase. They can be either quantitative or qualitative factors. In our study we have considered both quantitative (driving data, sensor parameters, etc.) and qualitative (road condition, weather condition, etc.) factors such as smart phone sensor data, and drivers inner and outer factors.
First step in the driver behaviour analysis is to collect the driving data while driver is executing the driving events. But, before we collect the data we have to decide on the model driver (persona) who's going to execute these events in a naturalistic environment. So, we first create the driver persona to understand the potential driver representing the model car drivers.

3.1 Creating driver Persona

Personas are imaginary, comprehensive representations of your model car drivers. They help you understand your potential drivers and make it easier for you to modify content to the specific needs, behaviours, and concerns of different groups. Personas can be created through research, surveys, and interviews of the target audience to gather driver's information. Rather than using imaginary variables like gender, age, driving experience, and educational background for creating persona to analyze drivers behaviour, we looked at how people actually drove, including factors like phone usage, honking, infotainment system use, and more [17].
The participants were asked to answer a few questions with the help of a driver survey form. Participants were requested to fill up the details like-

- Personal background (Name, Age, Gender, Qualification),
- What type of Driver he/she is? (Driving lessons, Experience, Skills, Daily distance, hrs spent, Average speed).

- What are the Factors affecting Driving Behaviour? (Time of driving event, Use of AC, Traffic, Road, Honking, Lane Cutting, Environmental Conditions, in-car stereo system, distraction)
- Is he/she is aware of any Driving Assistance available in-car (Awareness, Usefulness, Navigation system used)?
- Apart from all the features (safety, infotainment, ADAS, etc.) available in the vehicle, what else is required/suggested for safe driving.

The responses received helped us to select one-person representing the model car driver. Short description of the persona is defined in the following section.

3.2 Driver Persona

Persona Name: Mr. Jivan Thakare
Demographics: Male, 35 years
Background:
He's a Doctor by profession who's passionate about his profession and takes care of everyone in his family. He is experienced both from professional life and as a car driver. He’s a self-confident driver with 5+ years of driving experience. Other people see him as a good human being, practical and self-confident man with authority. He’s basically from Amravati district in Maharashtra, India. He completed his Bachelor's Degree of Ayurvedic Medicine and Surgery (BAMS) study and started his practice since 2012. He’s a father of 2 lovely kids Aarav and Arnav. His wife is also a Doctor by profession. There are 8 members in his family including him. His father is a retired Bank Manager and mother is a housewife. His elder brother is working in a Finance Company as a Manager at Nagpur, Maharashtra, India.
As a part of his daily routine, Mr. Jivan used to travel approximately 20 km to reach his clinic and back to home. He spends his entire day at the clinic and starts his return journey between 4pm to 5pm using the same National Highway.

4 Pilot Implementation

4.1 Objective

To record the smart phone sensor data while the driver executed particular driving events in a naturalistic traffic.

4.2 Experimental setup

Once the model driver (persona) is finalized, the next task is to design the experiment and the experimental conditions. We performed the experiment for 7 days, where the

driver performs 2 trips per day. Each one-way trip was of approximately 17.2 km of 25 minutes on average. In this experiment, we used Android application for recording smart phone sensor (accelerometer, gyroscope) data. The smart phone was horizontally fixed on the car's utility box while driver was executing the driving events. Table 2 gives the details of experimental setup.

Setup	**Description**
Android application used for recording Smartphone sensor data	Sensor Record (ver. 2.3.0)
Smartphone sensor	Accelerometer, Gyroscope
Vehicle type	LMV
Vehicle Model	New Maruti Suzuki Swift VXI
Smartphone used for recording sensor data	Redmi 4 with Android version 7.1.2
Smartphone location and position	Horizontally fixed in the car's utility box
Sampling Rate	50 Hz default
Driving experience	5+ years
Weather condition	Sunny

Table 2. Experiment Setup

4.3 Inner Variables

To understand the driver behaviour we have to study the other factors affecting his/her driving performance. Inner variables or in-car environment is one such factor that needs to be analyzed for driver behaviour classification. In our case, inner variables such as driving context, secondary tasks performed, phone usage, emotional status, etc. are taken into consideration while executing the driving events as the factor affecting the driver behaviour.

1. Driver context (passengers, infotainment system, in-car environment etc.)
   a. Driver Alone, No co-passenger.
   b. Radio ON.
2. Driver secondary tasks (phone conversation, chat with co- passenger, eating/drinking, etc.)
   a. Rare phone conversation (mostly patient issues).
3. Driver objective/s (safety, on time, comfort, enjoyment etc.)
   a. Reaching Clinic safely with comfortable drive.
4. Driver status (stress, fatigue, drowsiness, experience, self-confident etc.)
   a. At-ease
   b. A self-confident regular drive.
5. Driver primary goal (momentary)
   a. Keep consistent speed, maintain safe distance to vehicles in front and stay inside lane.

4.4 Outer Variables

Outer variables such as traffic conditions, road situations, environmental conditions, etc. also need to consider while analyzing driver behaviour. In our experiment we have taken care of these variables while recording the sensor data.

1. Traffic condition (nearby vehicles, rush hour, traffic- jam, etc.)
   a. Moderate traffic.
2. Road situation (highway, rural, urban)
   a. National Highway.\\
3. Road environment (wet surface, daylight, sun, darkness, fog etc.)
   a. Good weather conditions,
   b. Summer, mostly sunny,
   c. Dry road surface,
   d. Early morning start and afternoon return journey.\\
4. Main task (overall)
   a. Highway driving.

4.5 Results

The experiment was conducted for 7 days. Driver was executing the driving events every day for 2 trips. Smartphone sensors like Accelerometer and Gyroscope data were recorded with time stamps in a CSV file and used for further processing of driver profiling.

Driver behaviour is the responsible factor in traffic safety, fuel/energy consumption and environmental pollution. Driver behaviour profiling tries to understand how driver performs behind the wheels and how driving behaviour (normal, rash, etc.) impact other factors. Usually driver behaviour profiling tasks involve automated collection of driving data, calibration, and application of computational models to generate a classification that characterizes the driver driving profile. Driver behaviour profiling is the process of:

- Automated collection of the driving data (e.g., speed, acceleration, breaking, steering, and location) and applying a Computational Model to them.
- Input sensor (Accelerometer, Gyroscope) data (CSV file).
- Data collecting phase, driving a car, and gathering data from several different sensors (accelerometer, gyroscope, magnetometer, GPS, etc.).
- Pre-processing.
- Events mainly detected by calculating mean of successive accelerometer readings on certain axes and comparing them to empirical fixed/dynamic thresholds.
- Calculating Mean of certain axes.
- Applying MLA (SVM and Logistic Regression) for classification of driver behaviour (Normal/Aggressive).
- Comparing values with threshold and applying SVM algorithm.

- Comparing values with threshold and applying Logistic Regression algorithm.
- Driver Behaviour Classification (Normal/Aggressive)

5 Conclusion and Future Scope

In this work we presented our method for driver behaviour classification and evaluated the performance of two MLAs (Support Vector Machine and Logistic Regression) using data collected from Android smart phone sensors (accelerometer and gyroscope). We collected samples of these event types in a real-world for 7 days with a driver executing the driving events.
Our experiment shows that:
(i) Driver behaviour classification is possible using smart phone sensors; (ii) the accelerometer and the gyroscope are the best sensors to detect driving events; (iii) MLA such as SVM and Logistic Regression can be used for driver behaviour classification; and (v) the behaviour classification accuracy with available dataset for SVM is 99.98% and for Logistic Regression is 99.81% respectively.

As a future work, we intend to collect a more number of driving events samples with different vehicles, on various road conditions, in different weather, and temperature. We also expect to perform the experiment with more number of drivers. Finally, we intend to develop an Android smart phone application which can detect driving events in real-time and calculate the driver behaviour score.

References

1. Hamid Reza Eftekhari & Mehdi Ghatee, “A similarity-based neuro-fuzzy modeling for driving behaviour recognition applying fusion of smartphone sensors”, Journal of Intelligent Transportation Systems, 2019.
2. Bhakti Baheti, Suhas Gajre, Sanjay Talbar, “Detection of Distracted Driver using Convolutional Neural Network”, IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops, 2018.
3. Hamid Reza Eftekhari, Mehdi Ghatee, “Hybrid of discrete wavelet transform and adaptive neuro fuzzy inference system for overall driving behaviour recognition”, Transportation Research Part F 58 (2018).
4. Dang-Nhac Lu, Duc-Nhan Nguyen, Thi-Hau Nguyen, ID and Ha-Nam Nguyen, “Vehicle Mode and Driving Activity Detection Based on Analyzing Sensor Data of Smartphones” Article, Sensors 2018.
5. Kawtar Zinebi, Nissrine Souissi, Kawtar Tikito, "Driver Behaviour Analysis Methods: Applications oriented study", The 3rd International Conference on Big Data, Cloud and Applications - BDCA'18, Morocco, 2018.

6. Jair Ferreira Junior, Eduardo Carvalho, Bruno V. Ferreira, Cleidson de Souza, Yoshihiko Suhara, Alex Pentland, Gustavo Pessin, “Driver behaviour profiling: An investigation with different smartphone sensors and machine learning”, PLOS ONE | https://doi.org/10.1371/journal.pone.0174959 April 10, 2017.
7. Sanjay Kumar Singh, “Road Traffic Accidents in India: Issues and Challenges”, Transportation Research Procedia 25C (2017).
8. Anjali, Dr. V. K. Ahuja, “Traffic and Road Safety Management in India”, International Journal of All Research Education and Scientific Methods (IJARESM) ISSN: 2455-6211, Volume 4, Issue 3, March- 2016.
9. Caleb Ronald Munigety, Tom V. Mathew, “Towards Behavioural Modeling of Drivers in Mixed Traffic Conditions”, Springer International Publishing Switzerland 2016.
10. Minglin Wu, Sheng Zhang, and Yuhan Dong, “A Novel Model-Based Driving Behaviour Recognition System Using Motion Sensors”, Article, Sensors 2016.
11. Zhenyu Liu, MengfeiWu, Konglin Zhu, and Lin Zhang, “SenSafe: A Smartphone-Based Traffic Safety Framework by Sensing Vehicle and Pedestrian Behaviours”, Hindawi Publishing Corporation Mobile Information Systems Volume 2016.
12. Gys Albertus Marthinus Meiring and Hermanus Carel Myburgh, “A Review of Intelligent Driving Style Analysis Systems and Related Artificial Intelligence Algorithms”, Article, Sensors 2015.
13. Press Information Bureau, Government of India, Ministry of Petroleum & Natural Gas, https://pib.gov.in/newsite/printrelease.aspx?relid=102799.
14. Johannes Paefgen, Flavius Kehr, Yudan Zhai, Florian Michahelles, “Driving Behaviour Analysis with Smartphones: Insights from a Controlled Field Study”, ACM, 2012.
15. H. Eren, S. Makinist, E. Akin, and A. Yilmaz, “Estimating Driving Behaviour by a Smartphone”, Intelligent Vehicles Symposium, 2012.
16. Per Amdahl & Per Chaikiat, “Personas as drivers: An alternative approach for creating scenarios for ADAS evaluation”, Master Thesis in Cognitive Science, Linköping University, Sweden, 2007.

University Digital Engagement of Students

Maria C. Pereira[1][0000-0002-9792-3682] and João C. Ferreira[1,2][0000-0002-6662-0806] and Sérgio Moro[1][0000-0002-4861-6686] and Frederica Gonçalves[3][0000-0002-8857-6244]

[1] Instituto Universitário de Lisboa (ISCTE-IUL), ISTAR, Lisboa, Portugal
[2] Inov Inesc Inovação—Instituto de Novas Tecnologias, 1000-029 Lisbon, Portugal
[3] ITI/LARSyS, Universidade da Madeira - Escola Superior de Tecnologias e Gestão
maria_c_pereira@iscte-iul.pt

Abstract. Most of the prospective university students, especially from abroad, search in the online social networks what other students, present and former, are saying regarding the universities and the programs. Students express their sentiments regarding universities, programs, and courses they attended, or they are attending, in informal conversations in online social networks. We aim to understand how universities' and programs' enrolment numbers, relate to students' sentiments expressed in online social networks, and how this electronic Word of Mouth relate to the dissemination of the universities and the programs in online social networks. The research goal of this study is to create a prototype of a decision support system, based on information extracted from online social networks and data collected in the universities' informatic systems, to assist universities' managers in the decision-making process regarding internationalization.

Keywords: Universities, Social Media, Sentiment Analysis.

1 Introduction

Nowadays, universities try to attract students and are fighting for high rank positions. In the databases we queried, Scopus and Web of Science (WoS), we found 104 articles between 2008 and 2020, highlighting the importance of the use of online social networks to internationalize universities [1,2]. One of the most recent studies we found analyzes the use of Chinese social networks Weibo and WeChat, by UK universities to advertise and to interact with Chinese students in Mandarin [3]. In that study, it was possible to detect a positive association between the UK universities' reputation showed in Chinese social networks and the number of Chinese students enrolling in those universities.

In our research work, we extract students' satisfaction from social networks and promote a digital process to capture future students. We are investigating universities' reputation, performing sentiment analysis on data collected on online social networks, to understand if electronic Word of Mouth (eWoM) is an important source of candidates' recruitment. We want also to investigate the impact of universities' dissemination in social networks into the enrollment numbers. We are conducting a pilot test,

disseminating a program from ISCTE, Mestrado em Sistemas Integrados de Apoio à Decisão (MSIAD), since July 2020. In the years before 2020, MSIAD received a very low number of new students what eases the study of the effect of the dissemination.

Objectives

Objectives for the pilot test: 1st improve the number of new students enrolling in MSIAD; 2nd understand which platforms and which type of contents achieve more interactions.

Objective of the research: propose a prototype of a decision support system for higher education internationalization based on organization and social media data to improve students' satisfaction and universities attractiveness, for home and foreign students.

2 Proposal Concept – Action Planning & Action Tacking

We are following an action research methodology (ARM) [4], depicted in Fig. 1.

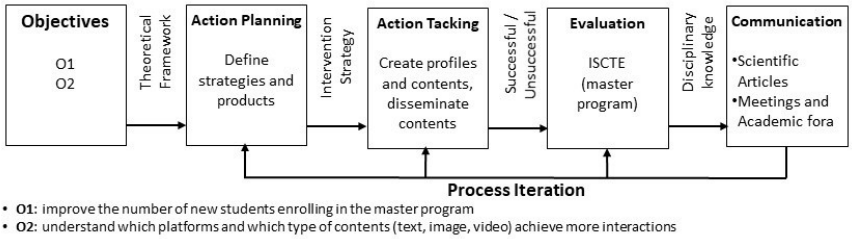


Fig. 1. Action Research Methodology

Planning and design

In the first phase, dissemination was made by two persons from July to August 2020. Plan: 1st create two blogs (PT & EN); 2nd publish in each blog in alternate days; 3rd share the posts in social networks[1]; 4th posts with text and one image; 5th collect statistics and the number of new students enrolled in MSIAD.

Technical configuration

The layouts for the blogs were selected to cause visual impact. Dissemination on social networks was done by two persons.

Organizational adaptation

In the second phase, dissemination is being made by one person from January to August 2021. New plan: 1st create two new blogs (PT & EN), with new layouts; 2nd new posts on blogs are published occasionally; 3rd other contents[2], are published on other platforms[3].

---

1 Facebook, LinkedIn, Twitter, Instagram, Pinterest, Tumbir

2 https://drive.google.com/drive/folders/1ntADiiO-BcCvmiJW4eVVVh0_S1JPymsV?usp=sharing

3 https://drive.google.com/file/d/1c82YdLGAO7_UAGaKZf5mm3UYY1h7-TT2/view?usp=sharing

Use
The PT blog[4] is primarily used to inform about MSIAD. The EN blog[5] is used to disseminate scientific articles authored by students of MSIAD.
Learning
First lesson: impacting images get more public interactions.
Second lesson: blogs are important by the extension and diversity of contents that they permit, but social networks have far more interactions.

3 Evaluation - Results from the application to a PT master program

First phase - July and August 2020
New 10 students enrolled during the dissemination. A total of 27 students enrolled in 2020/2021 versus a total of 18 students enrolled in 2019/2020. Increase of 50% from 2019/2020 to 2020/2021. We observed that Facebook was the primary source for blogs visitors. The peak of visitors happened when we published about the opening of the academic year 2020/2021, under the COVID-19 pandemic. Fig. 2., shows only a few statistics of the blogs compiled from Google Analytics.

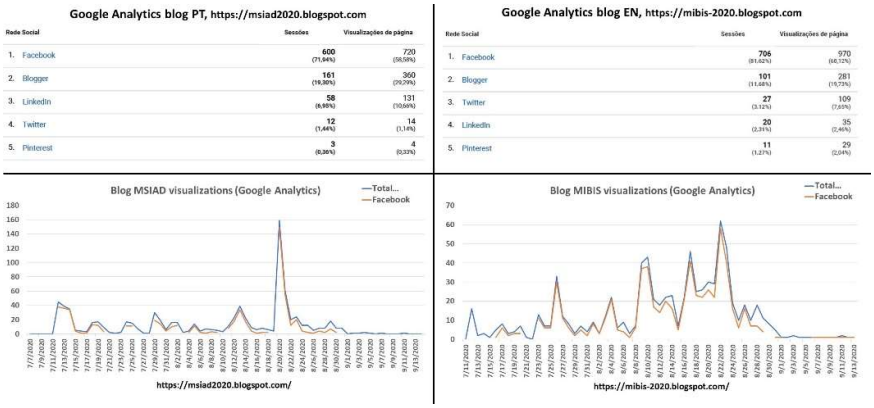


Fig. 2. Statistics of the blogs compiled from Google Analytics, regarding July and August 2020.

Second phase – February to April 2021
A total of 23 students enrolled to 2021/2022 versus a total of 16 students enrolled to 2020/2021, in the same period. Increase of 44% in this phase. Disseminating on a broader range of platforms allowed us to get a better idea of the preferred contents. Statistics shown in Fig. 3., are only about this phase. We see that MSIAD has the

4 https://msiad-iscte.blogspot.com
5 https://mibis-iscte.blogspot.com

largest percentage of conversions although has modest numbers in the other parameters, in the statistics compiled by ISCTE, from Google Analytics.

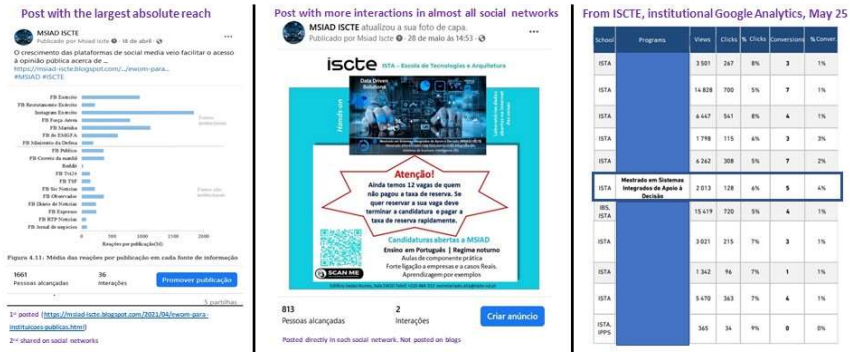


Fig. 3. Posts with more interactions and statistics from ISCTE.

4 Conclusions

We still do not have enough data to make a solid statement but, considering that no other factors considerably changed in MSIAD, we conclude that our dissemination positively contributed to the increase of students enrolling in MSIAD and could justify the apparent discrepancy in the percentage of conversions on Fig. 3. These observations agree with the studies pointing to social networks as good sources of students' recruitment. This ongoing study shows that the dissemination of programs in social networks deserves more attention from the universities' managers. We are collecting more data and implementing other types of analysis, e.g. sentiment analysis, to understand the universities' reputation in eWoM. Social networks are indicators of university and master program reputations.

Acknowledgement: This work is partially funded by national funds through FCT - Fundação para a Ciência e Tecnologia, I.P., under the project FCT UIDB/04466/2020.

References

1. Choudaha, R.: Social media in international student recruitment. Association of International Education Administrators (AIEA) Issue Brief no. March, 1–12 (2013).
2. Asderaki, F., Maragos, D.: The internationalization of higher education: the added value of the european portals and social media pages for the national and the institutional internationalization strategies. International Conference on Information Communication Technologies in Education, 498–510 (2012).
3. Zhu, Y.: Social media engagement and Chinese international student recruitment: understanding how UK HEIs use Weibo and WeChat. Journal of Marketing for Higher Education 29(2), 173–190 (2019). doi: 10.1080/08841241.2019.1633003
4. Baskerville, R. L.: Investigating Information Systems with Action Research. Communications of the Association for Information Systems 2(19), (1999). doi: 10.17705/1cais.00219

Pilot Implementation in Virtual Reality: Pitfalls, Tensions and Organizational Issues

Cristiano França[1,2], Michael Sousa[1,2,] Frederica Gonçalves,[2,3] [0000-0002-8857-6244] Élvio Rúbio Gouveia [2,4] [0000-0003-0927-692X] Bruna Gouveia [2,5][0000-0001-7706-190X] and Pedro Campos[1,2] [0000-0001-7706-5038]

[1] Exact Sciences and Engineering Faculty, University of Madeira
[2] Interactive Technologies Institute/LARSyS,
[3] Escola Superior de Tecnologias e Gestão, University of Madeira
[4] Department of Physical Education and Sport, University of Madeira
[5] Saint Joseph of Cluny Higher School of Nursing and Regional Directorate of Health, Secretary of Health of the Autonomous Region of Madeira
pedro.campos.pt@gmail.com

Abstract. Pilot implementations are, essentially, field tests of properly engineered yet unfinished systems. Contrasting with usability evaluations, pilot implementations are typically conducted in the field, and involve users performing their real work using the pilot system. There are many open research questions regarding pilot implementations. In this position paper, we briefly reflect on work-in-progress of a virtual reality installation for physical rehabilitation of professional football athletes. Concretely, we debate on the pitfalls, organizational issues, stakeholders' tensions and overall open questions about this particular pilot implementation. Results can be used to drive further work, as well as to foster discussion about pilot implementation challenges and research agenda.

Keywords: Pilot Implementations, Virtual Reality, Human Work Interaction Design, Field Tests, Rehabilitation.

1. Pilot Implementation in Virtual Reality

1.1. A VR system for Rehabilitation of Professional Football Athletes

Throughout their career, professional football athletes suffer many types of physical injuries, and are naturally affected by them. Whenever that happens, a rapid, swift recovery is paramount to maintaining their career level strong. The longer the recovery, the more difficult it is for the athlete to overcome it, and motivation can quickly become a serious issue, even when physical recovery is attained.

Typically, the rehabilitation process of an athlete is coordinated by a trained physiotherapist as well as other medical staff, depending on the football club's dimension. However, this process is always tedious and frustrating to the athlete. Exercises are relatively simple to perform, yet they can become extremely tedious.

On the other hand, Virtual Reality (VR) and its derived technologies have been increasingly popular in a variety of rehabilitation domains, as well as entertainment. Challenging the player or promoting exploration makes the plan of rehabilitation more exciting and more enjoyable than the traditional rehabilitation plan [3].

The rationale of using VR in soccer players is also based on its exceptional features. The possibility of creating different scenarios with an infinite number of repetitions where it is possible to manipulate the visual environment, for example, brightness of objects, their location, different perspective, temporal and spatial distortions of the movement trajectory, and feedback, may influence the performance in a way that is hard to achieve in the real world [3]. Besides that, VR can invoke embodied simulations of the human in the world, used to represent and predict actions, concepts, and emotions [4].

In this pilot implementation, we created, designed and rolled-out five different rehabilitation exercises built in a virtual environment in order to understand if the stimulation of the virtual environment promoted better psychological and social influence on the athletes. These exercises were performed in two different scenarios to give some freedom to the athlete performing the rehabilitation exercise.

Fig. 1. A user testing out the implementation of a series of physical rehabilitation exercises

Figure 1 provides an idea of the environment.

The athletes learn to execute the exercises when accompanied with someone from the medical team that guides them to the steps they must execute. This member of the

medical team helps and teaches the athlete to perform the exercises in the most appropriate way, so that the exercise helps in the rehabilitation, instead of having the opposite effect. Nevertheless, athletes do not always have a close contact, control or guidance with the medical staff when they are realizing the rehabilitation exercises from their plan. This makes it impossible to check if they are doing the exercise with the best technique and get instant feedback about that [3]. Consequently, this support and supervision that is usually realized by the medical staff can be done using Virtual Reality since it is possible to give instant feedback to the athlete from the immersive world where they are, regarding the correct technique needed for each exercise, and without the need for an extra person supervising the athlete.

1.2. Implementation Issues

The development of this system was made using the Unity3D game engine in conjunction with a VR headset and body tracking off-the-shelf equipment (VIVE) for the sake of bringing more immersive and realistic experiences to the rehabilitation process. We used VIVE Trackers, which give us the possibility of creating full-body experiences. To create this full body avatar, we track the athlete's skeleton using three different trackers that work synchronously with the headset and two controllers, creating the six points necessary for the effective, smooth tracking of the athlete's body. The three trackers are placed in each foot and in the athlete's waist, the headset in his head and the controllers in their hands. We also used VR inverse kinematics to help make the tracking more accurate, smooth and rapid. When injured athletes try these scenarios with the full body tracking, it is possible to have a more immersive experience since the athlete sees his body inside the scenario performing exactly what he is doing in real life and with very low latency [2].

The development of these exercises for the rehabilitation of injured athletes was implemented using state machines. Each exercise has its own state machine where we have states for each temporal part of the exercise and states for when all the repetitions are made. There is also a game-over state, triggered when the athlete fails to complete a given exercise. Using state machines gives us more control in each exercise, since we know in which state the user is when performing the exercise. It also provides the possibility of developing additional feedback to deliver a more immersive and responsive experience.

2. Pitfalls, Tensions and Organizational Issues

Pilot implementations are, essentially, "field tests of properly engineered yet unfinished systems" [1]. Contrasting with usability evaluations, pilot implementations are typically conducted in the field, and involve users performing their real work using the pilot system. There are many open research questions regarding pilot implementations [1].

It is important to reflect on this work-in-progress of a virtual reality installation for physical rehabilitation of professional football athletes. Concretely, we debate on the

pitfalls, organizational issues, stakeholders ’tensions and overall open questions about this particular pilot implementation.

Profiling our project with Hertzum’s five activities of pilot implementations [1], we detail the following:

Planning and design. We employed a multidisciplinary approach to this pilot’s planning and design activity. The main tension was to accurately depict the actual VR system behavior in a visual way, to achieve a common understanding by all stakeholders, namely the medical staff, athletes, experts in rehabilitation and sport sciences, as well as the engineers.

Technical configuration. The configuration of such systems always faces the extra challenge of calibration due to different users’s body sizes, arms’ length, etc. In a pilot implementation of a “sole”-digital system this does not occur. The parts of the VR system necessary for the pilot implementation had to be properly configured to fit the pilot site, interfaces to the users ’other systems were developed and tested. Adapting a novel space to such a highly-technological equipment also generates tension, but it helped some users to avoid withholding their opinion about the system or even try to derail pilot implementations by means of counter-implementation strategies.

Organizational adaptation. Introducing a VR cave-like, immersive environment in such a large and diverse organization such as a Premier league football club is far from trivial in terms of organizational adaptation. Many different people, with many different backgrounds intervene in the pilot implementation in one way or another. Additionally there is the continuous, almost realtime pressure of results. We consider that avoiding unrealistic expectations was the main pitfall in this activity.

Use. The actual usage of the pilot implementation was not yet tried by the actual athletes, but current results are showing that users often consider that an inspirational scenario in a different environment rather than a football stadium can lead them to a better mood to perform some of the rehabilitation exercises.

Learning. This last activity is still ongoing. According to the physiotherapist, the athlete in a long and exhaustive rehabilitation process may want to choose a football stadium scenario, so they don’t feel like they are out of work for a long time, unlike an athlete in a short-term rehabilitation who may want to choose a different scenario in order to perform the exercise in a completely new environment. But the actual preferences need to be measured in-loco and should be triangulated with physiological metrics of heart rate variability, galvanic skin response or even EEG real time data. Qualitative information plays a useful role as well and we plan on collecting a significant amount until the end of the pilot.

In conclusion, we hope this position paper can be used to drive further work, as well as to foster discussion about pilot implementation challenges and research agenda. We are particularly interested in contributing to an immediate research agenda revolving around the pitfalls and challenges that come up with pilot implementations when a multidisciplinary project is at stake, but the results can probably scale to any other type of VR project for physical rehabilitation.

Acknowledgments

MTL - Marítimo Training LAB, ProCiência 14-20, Instituto Desenvolvimento Empresarial da Região Autónoma da Madeira / Marítimo da Madeira - Futebol SAD, (12/2019-11/2021) 1.116.027,14€.

References

1. Hertzum, M., Bansler, J.P., Havn, E., Simonsen, J.: Pilot implementation: Learning from field tests in IS development. *Communications of the Association for Information Systems* 30(1), 313-328 (2012).
2. Akbaş, Anna & Marszałek, Wojciech & Kamieniarz, Anna & Polechoński, Jacek & Kajetan, Słomka & Juras, Grzegorz. (2019). Application of Virtual Reality in Competitive Athletes - A Review. *Journal of Human Kinetics*. 69. 5-16. 10.2478/hukin-2019-0023.
3. Mestre, Daniel & Ewald, Marine & Maiano, Christophe. (2011). Virtual reality and exercise: Behavioral and psychological effects of visual feedback. Studies in health technology and informatics. 167. 122-7.
4. Bossard, Cyril & Kermarrec, Gilles & Bénard, Romain & De Loor, Pierre & Tisseau, Jacques. (2009). Virtual Reality for Research and Training in Sport: an Illustration with CoPeFoot.

Socio-technical and Collaborative Design of Pilot Implementations of Digital Twins

Barbara Rita Barricelli[1][0000−0001−9575−5542]

Department of Information Engineering
Università degli Studi di Brescia, Brescia, Italy
barbara.barricelli@unibs.it
https://barbara-barricelli.unibs.it/

Abstract. The purpose of this short article is to stimulate a conversation around the importance of domain expert involvement in the design of digital twin pilot implementations. A socio-technical approach that sees the combined use of Human Work Interaction Design and End-User Development techniques is briefly illustrated.

Keywords: Socio-technical design · Collaborative Design · End-User Development · Human Work Interaction Design

1 Digital Twins Pilot Implementations

In the past twenty years, the advent of the Internet of Things (IoT) [3, 10] is changing the way data are exchanged among different sources, especially in the Industry field. Indeed, the diffusion of technologies such as sensors and actuators connected through the Internet, allows a continuous exchange of data. The integration of Artificial Intelligence models and Big Data Analytics for processing IoT data [4, 16] motivates one of the latest, and probably one of the most important advancement in the field of technology, that is, the Digital Twin.

Digital Twins can be defined as physical and/or virtual machines that are simulating, emulating, mirroring, or twinning the life of a physical entity [5]. A Digital Twin is more than a simple model or simulation [14, 17, 9, 19]: it is a living, intelligent, and evolving model, being the virtual counterpart of a physical entity or process.

Digital Twins are gaining more and more interest for their potentials and strong impact in application fields such as manufacturing [21, 18], aerospace [12, 22], and healthcare and medicine [20, 11, 6].

One of the strategic roles of Digital Twins is allowing the domain experts to monitor, control, and optimize processes. They are widely used to predict future statuses (e.g., defects, damages, failures), allowing to put in place predictive maintenance interventions. Indeed, Digital Twins require a professional human intervention, particularly in scenarios where they are used to test new features

and modifications of physical assets, or when they are exploited to provide answers such as diagnosis and treatments.

Given the variety, complexity, and increasing scale of Digital Twins design projects, it is not easy nor recommended to shift directly from prototypes to full-developed systems: pilot implementations can be seen as opportunities for engaging the domain experts in the design and testing. Pilot implementations are defined in [15] as "a field test of a properly engineered, yet unfinished system, in its intended environment, using real data, and aiming, through real-use experience, to explore the value of the system, improve or assess its design, and reduce implementation risk".

The involvement of the domain experts, acting in concert with multidisciplinary teams (e.g. developers, interaction designers, database administrators), would allow to apply to the design phase multi-faceted knowledge: this collaborative and socio-technical approach can be seen as promising but of course will add complexity to the process.

2 A Socio-Technical and Collaborative Design Approach

The goal of applying a socio-technical and collaborative approach to the design of Digital Twins pilot implementation, is to satisfy needs and requirements that cannot be fully anticipated at prototyping time. However, The need of finding new strategies to support the collaboration in multidisciplinary teams can be seen as an open issue: to bridge the communication gaps among stakeholders with diverse cultural and professional backgrounds requires the development of open-ended software environments that can be evolved and tailored to tackle the co-evolution of users and systems [23, 2].

So the answer to this need of bridging the communication gaps, can be a socio-technical design approach [24]. Such approach would support the designers to collaborate with the domain experts for exploiting Implicit Information and Tacit Knowledge of domain experts .

2.1 Human Work Interaction Design and End-User Development

This approach can be framed into Human Work Interaction Design (HWID) [1], a lightweight version of Cognitive Work Analysis, addressing the concept of Work in Human-Computer Interaction.

To enable end users, even if domain experts, in modifying and extending features of Digital Twin systems, End-User Development (EUD) methods and tools can be implemented. As defined in [7], EUD is "the set of methods, techniques, tools, and socio-technical environments that allow end users to act as professionals in those ICT-related domains in which they are not professionals, by creating,

modifying, extending and testing digital artifacts without requiring knowledge in traditional software engineering techniques". EUD is well established both in literature and in practice when dealing with collaborative systems for domain experts, and is currently mostly used in the IoT domain [8, 13]. Unfortunately, disciplines like Human Work Interaction Design and End-User Development are often neglected by developers and engineers, who prefer to concentrate on the development and usage of the digital twins, sometimes even forcing the end users to deal with an unusable and error-prone systems.

In the specific case of collaborative design of digital twin pilot implementation, the domain experts can be involved at different stages: requirements analysis, prototyping, evaluation of the prototypes and, later, the adjustment of the digital twin pilot implementation through EUD tools that the developers provide them.

Unlike the final digital twin, its pilot implementation will not be linked to the physical counterpart. This choice is related to the need of preserving a functioning system from being harmed by trial and error procedures. However, along the design and development of the pilot implementation, tests on the actual physical system have to be performed, to assess its effectiveness and efficacy, and EUD techniques could be adopted for this as well.

An Example of Digital Twin Pilot Implementation: Manufacturing
In the manufacturing field, digital twins are used to give a virtual but entirely mirrored representation of a production plant, in its entirety or limited to a specific section. The introduction of this technology usually takes place in the innovation of previously existing plants of which all workers and stakeholders, know their functioning, purpose, strengths, and weaknesses.

This industrial re-conversion goes hand in hand with the introduction of IoT devices that can sometimes require major changes in production processes. This aspect has a strong impact on workers and involving them in the design of the new plant version can serve to minimize the problems of its future use. The workers need therefore to be involved in both the first requirements collection phase and the prototyping activities. At the same time the workers should be involved as experts of the earlier version of the manufacturing plant, but also need to be formed to be able to fully understand the technological changes that are going to be applied to the plant and to the working environment and processes as a whole.

The role of EUD in this case, is in the creation of a digital twin pilot implementation that offers the workers to change the way information are viewed, to add new features (or ask for them) and to discuss among them about the features that are already in place. Obviously, with such a socio-technical approach, the workers will be also involved in proper evaluations of the implementation, aimed at collecting all those requests and feedbacks that were not gathered through EUD.

3 Conclusion

Like any system designed to support human work, Digital Twins need to be well designed to support a natural, effective, and efficient interaction experience. The introduction of a socio-technical and collaborative approach to Digital Twins Pilot implementations design is focused on enabling all stakeholders to actively contribute in the fundamental decisions that have to be taken since the early-stage of prototyping.

References

1. Abdelnour-Nocera, J., Barricelli, B.R., Lopes, A., Campos, P., Clemmensen, T.: Preface. In: Human work interaction design: Analysis and interaction design methods for pervasive and smart workplaces. Springer, New York, NY, USA (2015)
2. Ardito, C., Barricelli, B.R., Buono, P., Costabile, M.F., Piccinno, A., Valtolina, S., Zhu, L.: Visual mediation mechanisms for collaborative design and development. In: Stephanidis, C. (ed.) Universal Access in Human-Computer Interaction. Design for All and eInclusion. pp. 3–11. Springer Berlin Heidelberg, Berlin, Heidelberg (2011)
3. Ashton, K.: That 'internet of things' thing. RFiD Journal **22**, 97–114 (01 2009)
4. Banerjee, A., Dalal, R., Mittal, S., Joshi, K.P.: Generating Digital Twin models using Knowledge Graphs for Industrial Production Lines. In: Workshop on Industrial Knowledge Graphs, co-located with the 9th International ACM Web Science Conference 2017. Troy, NY, USA (June 2017)
5. Barricelli, B.R., Casiraghi, E., Fogli, D.: A survey on digital twin: Definitions, characteristics, applications, and design implications. IEEE Access **7**, 167653–167671 (2019). https://doi.org/10.1109/ACCESS.2019.2953499
6. Barricelli, B.R., Casiraghi, E., Gliozzo, J., Petrini, A., Valtolina, S.: Human digital twin for fitness management. IEEE Access **8**, 26637–26664 (2020). https://doi.org/10.1109/ACCESS.2020.2971576
7. Barricelli, B.R., Cassano, F., Fogli, D., Piccinno, A.: End-user development, end-user programming and end-user software engineering: A systematic mapping study. Journal of Systems and Software **149**, 101–137 (2019). https://doi.org/https://doi.org/10.1016/j.jss.2018.11.041, https://www.sciencedirect.com/science/article/pii/S0164121218302577
8. Barricelli, B.R., Valtolina, S.: A visual language and interactive system for end-user development of internet of things ecosystems. Journal of Visual Languages and Computing **40**, 1–19 (2017). https://doi.org/https://doi.org/10.1016/j.jvlc.2017.01.004, https://www.sciencedirect.com/science/article/pii/S1045926X16300295, semiotics, Human-Computer Interaction and End-User Development
9. Boschert, S., Rosen, R.: Digital twin – the simulation aspect. In: In P. Hehenberger, D.B. (ed.) Mechatronic Futures Challenges and Solutions for Mechatronic Systems and their Designers, pp. 59–74. Springer-Verlag (2016)
10. Canedo, A.: Industrial iot lifecycle via digital twins. In: Proceedings of the Eleventh IEEE/ACM/IFIP International Conference on Hardware/Software Codesign and System Synthesis. p. 29. ACM (2016)

11. Chen, X., Feng, Y., Zhong, W., Kleinstreuer, C.: Numerical investigation of the interaction, transport and deposition of multicomponent droplets in a simple mouth-throat model. Journal of Aerosol Science **105**, 108–127 (2016)
12. Force, U.A.: Global horizons final report: United states air force global science and technology vision. In: AF/ST TR 13-01. United States. Air Force (2013 - last accessed: April 7, 2019), https://www.hsdl.org/?view&did=741377
13. Ghiani, G., Manca, M., Paternò, F., Santoro, C.: Personalization of context-dependent applications through trigger-action rules. ACM Trans. Comput.-Hum. Interact. **24**(2) (Apr 2017). https://doi.org/10.1145/3057861, https://doi.org/10.1145/3057861
14. Grieves, M.: Digital twin: Manufacturing excellence through virtual factory replication. Digital Twin White Paper p. 7 (2015), https://research.fit.edu/media/site-specific/researchfitedu/camid/documents/1411.0_Digital_Twin_White_Paper_Dr_Grieves.pdf
15. Hertzum, M., Bansler, J.P., Havn, E.C., Simonsen, J.: Pilot implementation: Learning from field tests in is development. Communications of the Association for Information Systems **30** (2012)
16. Ikävalko, H., Turkama, P., Smedlund, A.: Value creation in the internet of things: Mapping business models and ecosystem roles. Technology Innovation Management Review **8**, 5–15 (03 2018). https://doi.org/10.22215/timreview/1142
17. Kritzinger, W., Karner, M., Traar, G., Henjes, J., Sihn, W.: Digital twin in manufacturing: A categorical literature review and classification. IFAC-PapersOnLine **51**(11), 1016–1022 (2018)
18. Lu, Y., Liu, C., Wang, K.I.K., Huang, H., Xu, X.: Digital twin-driven smart manufacturing: Connotation, reference model, applications and research issues. Robotics and Computer-Integrated Manufacturing **61**, 101837 (2020). https://doi.org/https://doi.org/10.1016/j.rcim.2019.101837, https://www.sciencedirect.com/science/article/pii/S0736584519302480
19. Negri, E., Fumagalli, L., Macchi, M.: A review of the roles of digital twin in cps-based production systems. Procedia Manufacturing **11**, 939 – 948 (2017). https://doi.org/https://doi.org/10.1016/j.promfg.2017.07.198, http://www.sciencedirect.com/science/article/pii/S2351978917304067, 27th International Conference on Flexible Automation and Intelligent Manufacturing, FAIM2017, 27-30 June 2017, Modena, Italy
20. Scharff, S.: From digital twin to improved patient experience (September 2010), https://www.siemens-healthineers.com/news/mso-digitaltwin-mater.html
21. Tao, F., Cheng, J., Qi, Q., Zhang, M., Zhang, H., Sui, F.: Digital twin-driven product design, manufacturing and service with big data. The International Journal of Advanced Manufacturing Technology **94**, 3563–3576 (2018)
22. Zakrajsek, A.J., Mall, S.: The development and use of a digital twin model for tire touchdown health monitoring. In: 58th AIAA/ASCE/AHS/ASC Structures, Structural Dynamics, and Materials Conference. p. 0863 (2017)
23. Zhu, L., Barricelli, B.R., Iacob, C.: A meta-design model for creative distributed collaborative design. Int. J. Distrib. Syst. Technol. **2**(4), 1–16 (2020). https://doi.org/10.4018/jdst.2011100101
24. Zhu, L., Mussio, P., Barricelli, B.R., Iacob, C.: A habitable space for supporting creative collaboration. In: 2010 International Symposium on Collaborative Technologies and Systems. pp. 617–625 (2010). https://doi.org/10.1109/CTS.2010.5478455

Co-designing Workshop as Pilot Implementation for Complex Workplaces.

Parisa Saadati* , José Abdelnour-Nocera*, and Torkil Clemmensen‡

University of West London, UK; † University of West London and ITI/Larsys Portugal,
‡ Copenhagen Business School, Denmark

parisa.saadati@uwl.ac.uk,
abdejos@uwl.ac.uk,tc.digi@cbs.dk

ABSTRACT. Here we present a case study to explore the implications of the co-design of future autonomous technologies for user experience (UX) and engagement. Given the high demand for automation in daily life and workplaces, there is a need to assess the value of co-design with the end-users to evaluate users' experiences and engagements and employee's social sustainability in multiple contexts such as work, health, entertainment, and learning. The term automation in this paper also covers some of what has been called AI or more sophisticated automation. This case is driven by a member of the innovation department of the airport and UX researchers. Our main objective was to employ participatory design and work domain analysis (WDA) as part of a framework to co-design and plan trials, i.e. pilots, for a future automated systems for smart work in airport terminal operations. Over two weeks in two workshops in a London-based airport, we run code-designed workshops to help the decision-makers understand workplace needs and employee welfare while selecting future automated systems.

Keywords: HCI, Participatory Design, Automation, Personas, Interaction Design, Abstraction Hierarchy, Trials, Pilot Implementation

1 Introduction

Automation and the introduction of Industry 4.0 interactive technologies in industrial work systems have brought new ambiguities in the challenges and burdens on interactive systems designers [13]. Socio-Technical System Design (STSD) has identified and addressed several problems in understanding and developing complex autonomous systems [1, 3, 8, 10]. Therefore, service design and co-design processes are utilized to facilitate collaboration in new ways [16]. Designing future technologies needs more cross-organizational collaborations to produce innovative and creative outcomes. The trial planning or pilot implementation process in designing such scenarios can help organizations understand less visible socio-technical factors. In addition, more studies on UX and innovation are needed in light of advances in Artificial Intelligence (AI) and the growing use of more sophisticated automation technologies [10].

Simultaneously, there is a need for cooperation and better communication in human teams and individuals and AI systems to achieve better UX goals for future automation

scenarios, namely AI/UX goals[1, 15]. Hence, developments towards future autonomous systems need to be carefully co-designed and tests in the field before or during the design. This should be properly engineered and bring some lessons learned to improve the system. In this paper, we report on a case study that employs participatory design and work domain analysis (WDA) [6] as a means for planning, co-designing, and choosing trials for future automated systems considering AI/UX goals in the context of airport terminal operations.

This paper describes the process of a team of researchers and operational decision-makers of an airport in co-designing two scenarios supporting smart work in the terminals. Workshop sessions resulted in two future work scenarios leading to automated systems' prototyping using a WDA tool known as Abstraction Hierarchies (AH) [5] for each scenario. The objective is to explore and illustrate how a London-based airport uses participatory design, interaction design, and WDA methods to decide future autonomous systems trials. The paper concludes with lessons we learned using such workshops as a pilot implementation phase before selecting the new technologies.

2 Participatory Design and Work Domain Analysis

Approaches that have emerged to promote user engagement in system development have been implemented in different industries, such as aviation [17, 19, 20]. In addition, the participatory design approach has gained interest in the development of commercial and business applications [9]. The participatory design process emphasizes mutual learning, as none of the participants, either ´designer' or 'user' knows everything. WDA is part of the broader Cognitive Work Analysis framework that supports and structures the analysis needed when designing a flexible and adaptive system [21]. Using WDA has two distinct advantages. First, WDA is a multi-dimensional analysis that incorporates the physical and the social environment to provide a detailed description. Secondly, WDA can be paired with interface design [11] to generate new information system designs.

We decided to use AH as a design tool, as it provides a discreet, discreet, and complete description of a work domain at different abstraction and concrete levels. Furthermore, AHs are suitable to use by stakeholders from different backgrounds with little or no experience of WDA [2, 5]. AH is a cognitive engineering approach to human-machine systems design supported by empirical studies of operators' fault-finding strategies. In the Ecological Interface Design framework, the AH is used to determine what kinds of information should be displayed on the system interface and how the information should be arranged [11]. Using WDA allows us to identify some of the socio-technical and contextual aspects early, which are generally not visible to the designers anticipating some of the benefits typically only found in pilots or trials.

3 Case Study questions

This study focused on understanding people, directly and indirectly, working towards selecting the new technologies or dealing with such systems' UX goals. We offer

insights based on our empirical studies that can be used to analyze and evaluate the lesson learned in co-designing for future scenarios. We co-created scenarios, personas, and AH as design methods to answer several questions: can co-designing workshops lead to optimal design/selection of autonomous future technologies?

3.1 Workshop's settings

Our observations on the trials and the shortfalls seen regarding the use, usability, and UX of these tested technologies helped us suggest the airport user-centered design tools and methods for introducing new automated systems. We suggested that implementing co-design workshops with their employees before selecting future technologies can help the organizations select an efficient automated system for emergent future in a complex organization. This system can have better user engagement and interaction than their trials. We noticed that the airport selects its future technologies mainly by running trials and observing how the trials proceed. However, customer and employee satisfaction and UX in the trials were not easily assessed and articulated. This suggestion then triggered two co-design workshops with 14 researchers, industry decision-makers, and employees from different departments.

Various topics and design techniques were introduced in the workshop 1 by the design researchers. Participants engaged in prototyping system concepts for autonomous systems needed in their departments. They considered the customer demands and value proposition for their systems. All participants had enough experience with autonomous technologies and leading Industry 4.0 implementation. Each group produced two future technologies concepts in the first round after refining the categories into the future's short and concrete experiences. Finally, we had a sharing and voting session to discuss and converge on the most appealing alternative future scenarios. From a list of suggested scenarios (automated helpers, autonomous tugs, context-aware guest invites system, baggage tracing, and smart asset management.), the participant selected two; Autonomous Tug and Pushback Taxi (ATPT) and Automated Asset Management and Maintenance (AAMM).

In workshop 2 we used customer job (part of value proposition canvas), persona's template, AI/UX goals [15] and AI guidelines [1], sketching, and AH [7] to prototype the short-listed future scenarios voted in the workshop 1. Researchers fulfilled a facilitator role supporting participants unfamiliar with the design tools and were active designers while engaging in creative tasks. In the last stage of workshop 2, each group, walkthrough their concept [23] using their low fidelity prototype and received feedback from other colleagues and decision-makers on how likely the idea is feasible. Finally, group members evaluated their design with UX goal templates [1, 15].

4 What we learned

This case study's main contribution is helping the airport co-design automation scenarios incorporate workers' UX and work. During the design process in the workshops, we used WDA to contextualised the systems and assessed employees' acceptance for

these systems. We believe this should be part of a pilot implementation phase for the complex organisations before introducing new systems. Our main learning insights are based on the interactions and outputs from the field observations, focus group, brainstorming session, co-design workshops transcripts, and researchers' notes.

We have learned that combining academics and employees can bring more efficiency to these workshops by providing recommendations viable on the industry that can academically back up. We also learned that employees are the best consultant for analysing the functions' allocation for an optimal distribution of both functions and tasks between a partly autonomous system and the user [4]. Also, we observed that in complex jobs with a high volume of data (e.g. airport), employees show more interest in providing support of human workers by robots or machines or any other AI systems as an essential aspect of the new technologies. Precisely, for tasks with a range of unpleasant, repetitive, too exhausting, or unsafe nature [14, 22]. However, our field observation showed that low-skilled jobs have a different perspective and fear losing their jobs when it comes to AI systems. We also learned that trust and safety considerations are more important in aviation than other sectors as aviation insurance policies associated AI technologies with risks and uncertainty as it takes control from humans.

Towards the end of workshop 2, most participants understood how a future automation prototype could fit better in their work environment through increased awareness of the work domain. Furthermore, they have agreed that there is a relation between modes of discovery, design improvements, interaction, and socio-spatial aspects. At the end of the session, one senior manager praised co-designing and WDA methods to select and design new trials of future automation use cases and prototypes for the airport.

4.1 Co-designing workshop settings

We noticed that using such workshops for better employees' engagement in the emergent future system design or even selecting the trials is beneficial for complex organizations. For a better result in the ideation process, providing a context, scenario, and better facilitation is required. Previous studies cover how the ideation process can be more comfortable and quicker for the participants [12, 18]. We observed that selecting the right design tool can provide a better reflection of the participants' knowledge. For example, participants found AH complicated to use without the design researchers' help. Adding more tangible elements [18], such as pre-structured cards and easy-to-use collaborative tools, can always be beneficial.

Moreover, there is a need to use data-driven design tools for the co-design workshops for future scenarios. Preparing the environment to ideate is another important factor; participants should understand the scenario, products, and the future system's domain and environment. The whole process provided the participants with an exemplary scenario to generate and manage their ideas. Otherwise, we must provide sufficient (internal/external) data for the context and persona creation. In line with the environment preparation and facilitator's instructions, it is essential to plan for a scenario in advance for every co-design workshop for future systems. This can help the participant access context and a road map to look forward and share more relevant information.

5 Conclusion

The main aim of the presented case study is to explore and illustrate how a London-based airport uses participatory design, interaction design, and WDA methods to make decisions on trials for future autonomous systems that incorporate workers' experiences. The paper presented that these co-design workshops and WDA design tools can yield early contextual socio-technical insights typically only found during pilot implementations. We empirically assess our hypothesis on the ATPT and AAMM scenarios in an airport domain based on customer profiles, personas, scenarios, AH, and a rough prototype to achieve this goal. We observed that this method can help designers or decision-makers foreseen the factors that may not be visible in the design. Producing the five layers for AH more efficiently with the help of Academics and providing more tangible examples was one of the main motivations in producing the final designs. After the workshop, it was highlighted in the senior participants' responses that using this method helps them better understand workers and the aviation sector's needs and perspectives for selecting the new autonomous systems, which was not part of their practices.

References

1. Amershi, S. et al.: Guidelines for Human-AI Interaction. In: Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems. pp. 1–13 Association for Computing Machinery, New York, NY, USA (2019). https://doi.org/10.1145/3290605.3300233.
2. Barcellini, F. et al.: Designers' and users' roles in participatory design: What is actually co-designed by participants? Applied Ergonomics. 50, 31–40 (2015). https://doi.org/10.1016/j.apergo.2015.02.005.
3. Barricelli, B.R. et al. eds: Human Work Interaction Design. Designing Engaging Automation: 5th IFIP WG 13.6 Working Conference, HWID 2018, Espoo, Finland, August 20 - 21, 2018, Revised Selected Papers. Springer International Publishing (2019). https://doi.org/10.1007/978-3-030-05297-3.
4. Barricelli, B.R. et al.: Human Work Interaction Design. Designing Engaging Automation: 5th IFIP WG 13.6 Working Conference, HWID 2018, Espoo, Finland, August 20-21, 2018, Revised Selected Papers. Springer (2019).
5. Bodin, I. et al.: Work domain analysis of an intensive care unit: An Abstraction Hierarchy based on a bed-side approach. In: Proc. Human Factors Ergonom. Soc. Eur. Annu. Conf. pp. 109–118 (2016).
6. Burns, C.: Cognitive Work Analysis: New Dimensions. In: Campos, P. et al. (eds.) Human Work Interaction Design. Work Analysis and HCI. pp. 1–11 Springer Berlin Heidelberg, Berlin, Heidelberg (2013). https://doi.org/10.1007/978-3-642-41145-8_1.
7. Burns, C.M. et al.: Evaluation of Ecological Interface Design for Nuclear Process Control: Situation Awareness Effects. Hum Factors. 50, 4, 663–679 (2008). https://doi.org/10.1518/001872008X312305.

8. Cabrero, D.G. et al.: A hermeneutic inquiry into user-created personas in different Namibian locales. In: Proceedings of the 14th Participatory Design Conference: Full papers-Volume 1. pp. 101–110 ACM (2016).
9. Chin, G., Rosson, M.: A case study in the participatory design of a collaborative science-based learning environment. Presented at the (2004).
10. Dikmen, M., Burns, C.: Trust in autonomous vehicles: The case of Tesla Autopilot and Summon. In: 2017 IEEE International Conference on Systems, Man, and Cybernetics (SMC). pp. 1093–1098 (2017). https://doi.org/10.1109/SMC.2017.8122757.
11. Hajdukiewicz, J., Burns, C.: Strategies for Bridging the Gap between Analysis and Design for Ecological Interface Design. Proceedings of the Human Factors and Ergonomics Society Annual Meeting. 48, 3, 479–483 (2004). https://doi.org/10.1177/154193120404800344.
12. Inie, N., Dalsgaard, P.: How Interaction Designers Use Tools to Manage Ideas. ACM Trans. Comput.-Hum. Interact. 27, 2, 7:1-7:26 (2020). https://doi.org/10.1145/3365104.
13. Kadir, B.A., Broberg, O.: Human-centered design of work systems in the transition to industry 4.0. Applied Ergonomics. 92, 103334 (2021). https://doi.org/10.1016/j.apergo.2020.103334.
14. Kirk, A.K., Brown, D.F.: Employee assistance programs: a review of the management of stress and wellbeing through workplace counselling and consulting. Australian Psychologist. 38, 2, 138–143 (2003). https://doi.org/10.1080/00050060310001707137.
15. Kymalainen, T. et al.: Evaluating Future Automation Work in Process Plants with an Experience-Driven Science Fiction Prototype. In: 2016 12th International Conference on Intelligent Environments (IE). pp. 54–61 IEEE, London, United Kingdom (2016). https://doi.org/10.1109/IE.2016.17.
16. Mugglestone, M. et al.: Accelerating the improvement process. Clinical Governance: An Intl J. 13, 1, 19–25 (2008). https://doi.org/10.1108/14777270810850599.
17. Oostveen, A.-M., Lehtonen, P.: The requirement of accessibility: European automated border control systems for persons with disabilities. Technology in Society. 52, 60–69 (2018). https://doi.org/10.1016/j.techsoc.2017.07.009.
18. Rygh, K., Clatworthy, S.: The Use of Tangible Tools as a Means to Support Co-design During Service Design Innovation Projects in Healthcare. In: Pfannstiel, M.A. and Rasche, C. (eds.) Service Design and Service Thinking in Healthcare and Hospital Management. pp. 93–115 Springer International Publishing, Cham (2019). https://doi.org/10.1007/978-3-030-00749-2_7.
19. Simonsen, Jesper and Robertson, Toni: Routledge international handbook of participatory design. Routledge (2012).
20. Tan, W., Boy, G.A.: Tablet-Based Information System for Commercial Aircraft: Onboard Context-Sensitive Information System (OCSIS). In: Harris, D. (ed.) Engineering Psychology and Cognitive Ergonomics. pp. 701–712 Springer International Publishing, Cham (2018). https://doi.org/10.1007/978-3-319-91122-9_55.
21. Vicente, K.J.: Cognitive Work Analysis: Toward Safe, Productive, and Healthy Computer-Based Work. CRC Press (1999).

22. Yerkes, R.M., Dodson, J.D.: The relation of strength of stimulus to rapidity of habit-formation. Punishment: Issues and experiments. 27--41 (1908).
23. Tools | Service Design Tools, https://servicedesigntools.org/tools.html, last accessed 2021/01/27.

Mental Happiness – A Pilot Implementation for Elderly People

Arminda Guerra Lopes[1][0000-1111-2222-3333] Inês Cerejo[2][1111-2222-3333-4444]

[1] LARSyS/ITI and Polytechnic Institute of Castelo Branco, Portugal
[2] Polytechnic Institute of Castelo Branco, Portugal
aguerralopes@gmail.com

Abstract. Aging is a worldwide problem in rapid acceleration. As far as people get older their cognitive and functional capacities progressively decrease. To reinforce a positive behavior for those that are alone or isolated Information and Communication Technologies (ICTs) play an important role.
As a contribution to extend elderly quality of life we propose a solution to promote an active aging by stimulating and reinforcing the cognitive and functional capacities. The results of this pilot prototype show that elderly people benefited to improve their skills and desires of socializing, moving, training the brain, using memory resources to gain and maintain confidence. Conversely, the interviewed psychologists considered that the online application contributed to decrease and delay situations of depression.

Keywords: Elderly, Cognitive Stimulation, Caregivers, Pilot Prototype, Application Design.

1 Introduction

According to the Pordata [1], the number of elderly in Portugal will double in the next ten years. The aging rate, which in 2013 was 133 elder people to 100 young people, is estimated to be 225 elder in 2033 and in 2060 will be 309 elder people. This means that in the next decade, the world will have more seniors than children. This problem is especially important in Western societies where the increasing average of life expectancy and the low birth rates have led to more and more elderly people being isolated, alone, and without family support.

Although new technologies allow and facilitate communication and the fight against isolation, in practice, these tools are only available to those who still maintain a high standard of intellectual abilities and learned in a timely manner to deal with such tools, creating a kind of cycle. Those who would most need access to cognitive stimulation are precisely the least able to access the media that provides it.

The access to Internet and other social platforms works well when cognitive skills and communication skills are still high. After this phase, a true recessive spiral of loss of faculties, loss of communication capacity, loss of references, loss of motivation,

isolation, mental inactivity, depression, which in turn leads to life disinterest, among other situations. Rapidly, these people's quality of life will decrease.

For some elderly people to stay at home rather than going to nursing homes is a way to increase the time spent in their spaces, their things and their memories. However, what we see today is that some of them no longer have their peers, the children and grandchildren nearby, and in some cases, the towns or villages where they live are depopulated and far from the large housing centers.

The proposed application intended to contribute to: strength human abilities, to collaborate and share information with others, to exercise their skills, to find an occupational therapy and to remember lost memories of the past. Conversely, it provides psychologist support, about mental and physical activity of their patients.

This paper is organized as follows: first we introduce the theme. Then, we outline, according to literature, the importance of cognitive stimulation. Next, the study is introduced focusing on the users. The methodological approach is presented and afterwards the application design is summarized. Finally, the main results are presented as well as conclusions.

2 Cognitive Stimulation

People who are affected by dementia often face cognitive, behavioral and psychosocial difficulties, including impairment and degeneration of memory and of perceptions of identity. As a result, many have reduced physical activities or social engagement, or are unable to work. Along the years, scientists have identified ways to minimize age-related changes and improve everyday memory function: socialize, get moving, train the brain, use memory aids to gain and maintain confidence [2, 3].

Some studies, for example [4, 5] show that brain aging, and loss of cognitive and memory functions can be partially counteracted by exercise and by the practice of intellectual activity. The capacity for communication, for reasoning and memory is kept longer when stimulated and exercised. There are other studies that indicate the benefits of interventions with the goal of training communication skills especially in terms of reinforcing positive behavior and satisfying interactions [6, 7]. Some examples of communication strategies are giving an idea, asking closed questions, giving direct instructions, paraphrasing repetitions, using simple phrases, using communication based on personalities, that is, considering the life history, values and personal preferences of individuals with dementia [8, 9].

3 The Context for the Pilot Prototype

For the development of this application, it was important to analyze the elderly's routines and to monitor them at home. They had very low level of personal satisfaction, with monotonous routines and activities. After this analysis, it was decided to produce a website capable of bringing them some emotional and physical wellbeing conditions through new challenges. The purpose of this application will be the entertainment of the elderly population, providing them with moments of fun and wellbeing. This

application will be also implemented in nursing homes. It will contribute to allow elderly people to have access to an account, with the available data, after starting a session, to enjoy the entire website content.

3.1 Personas

This application is for elderly people since they are in a less privileged situation in society. It will provide the elderly people with the possibility of improving their mental conditions, ending loneliness, as well as improve their health problems, including depression, or anxieties. Obviously, each type of person has different needs, tastes, and preferences, depending on the profile of each user. The selection of the multimedia content will be made depending on various data, such as age, gender, professional training, level of education social class, purchasing power, location, consumption habits, among others. So, considering these variabilities, it was decided to design the type of users, creating personas.

The role of the persona is, in this case, to take the application into a real context, adapting it to the various situations of each person, and the various preferences, in order to optimize the application design and to assign a profile to the developed application (Figure 1).

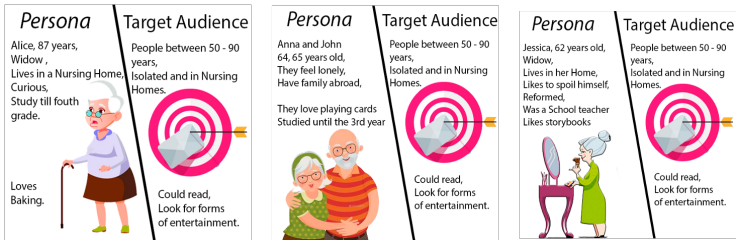


Figure 1 - Personas

4 Methodological Approach

This section describes the methods and participants involved in a preliminary study. The study served as a guide to support the proposed interactive application. We followed a user centered design approach which focused on the user necessities to avoid future errors and to design positive user's experiences.

Interviews were a research method of data collection and analysis. Interviews were conducted with caregivers (12 participants) and with elderly people (22 participants - 8 male and 14 female participants). The questions were mainly open questions and the issues discussed with caregivers were about the number of elderly they had in charge to follow; the difficulties and challenges they had with them; types of tools and equipment they had available to use; personal suggestions to include in an interactive application to help the collaboration and communication processes among the intervenient. The content of the end-users' questions was about: if they lived alone or with family; ways of spending time; visiting people per day; access to technological tools – laptop

and mobile phone; internet access. Interviews were face-to-face and with interviewee authorization. They were video recorded. Subsequently, a transcript of the interview was made. After, memo-writing was elaborated to highlight the main interviewee' s concerns. The data analysis results were considered on the application development since it was our goal to approach a user centered design strategy. User centered design is a design approach where the user has a close involvement in order to meet users' expectations and requirements. The design process sketches the phases of the design and development and it explicitly present users, tasks and environments. The data collected in the study's first stage, from questionnaires and interviews were analyzed and the results were considered on the application design.

5 Application Design

This section outlines the design and architecture, of the proposed application. There are four main activities which have included different thematic components. The user will choose whatever they want. However, caregivers may suggest those that are more appropriated for each user. We got information about the favorite themes that users suggested and wanted to share with others. The obtained data was considered on the application design. The user has access to different activities. For cognitive, physical, entertainment and social interaction development. The components functionalities permit to accomplish different tutorials about dance, gymnastic and handwork. For cognitive stimulation they can read stories, play with puzzles, practice language skills (grammar, orthography and words correction) and math exercises. The physical activities are accompanied through videos. To improve their thinking, users can answer questions about music, mathematics, Portuguese language, history of Portugal and geography. The application authorizes also to listen to different radio channels and see each family photo gallery. The pilot implementation has been in test at a nursery home.

6 Results and Conclusion

This application was preceded by a preliminary study, conducted in three villages of Castelo Branco, a Portuguese region. Our source of information was an elderly association, which was the key motivational basis for the proposed work and two nursery homes. We used questionnaires and interviews research methods to understand the context for the application and the intervenient' s desires. In the second phase we designed the application. The prototype was tested with caregivers and elderly users.

Presently, we are collecting new data from users' observations and comments which will be included in the final application design. The feedback comments are very positive, and the users are extremely satisfied with this application.

Acknowledgments

This research was funded by LARSyS (Projeto - UIDB/50009/2020). The authors would like to thank the anonymous interviewees for their valuable comments and helpful suggestions.

References

1. Pordata, https://www.pordata.pt/Pesquisa/numero%20de%20idosos%20em%20portugal, last accessed 2021/04/02.
2. Kausler D. H. (1991). Experimental psychology, cognition, and human aging (2nd ed). New York, NY: Springer-Verlag.
3. Hulicka I. M., & Grossman J. L. (1967). Age-group comparisons for the use of mediators in paired-associate learning. Journal of Gerontology, 22, 46–51. [PubMed]
4. Mandolesi, L., Polverino, A., Montuori, S., Foti, F., Ferraioli, G., Sorrentino, P., & Sorrentino, G. (2018). Effects of Physical Exercise on Cognitive Functioning and Wellbeing: Biological and Psychological Benefits. Frontiers in psychology, 9, 509. doi:10.3389/fpsyg.2018.00509
5. Lautenschlager, Nicola. KayCox, Elizabeth, Cyarto. V. (2012). The influence of exercise on brain aging and dementia. In Biochimica et Biophysica Acta (BBA) - Molecular Basis of Disease. Edited by B.T. Gold, J.N. Keller Volume 1822, Issue 3, March 2012, Pages 474-481
6. Woolham JG, Steils N, Fisk, M, Porteus J, Forsyth K. (2018). The UTOPIA project. Using Telecare for Older People In Adult social care: The findings of a 2016-17 national survey of local authority telecare provision for older people in England. London: Social Care Workforce Research Unit, King's College London.
7. Eva Eggenberger, Katharina Heimerl and Michael I. Bennett Communication skills training in dementia care: a systematic review of effectiveness, training content, and didactic methods in different care settings. International Psychogeriatrics, Available on CJO doi:10.1017/S1041610212001664
8. Wilson R, Rochan E, Mihilidis A, et al. (2012). Examining success of communication strategies used by formal caregivers assisting individuals with Alzheimer's disease during an activity of daily living. Journal of Speech, Language, and Hearing Research. 55: 328-341.
9. Small JA, Gutman G, Makela S, et al. (2003). Effectiveness of communication strategies used by caregivers of persons with Alzheimer's disease during activities of daily living. Journal of Speech, Language, and Hearing Research. 46: 353-367. https://doi.org/10.104 4/1092- 4388(2003/028)

Usability Evaluation of Mobile App to Support Road Safety

Shrikant Salve[1] and Ganesh Bhutkar[2]

[1]MIT Academy of Engineering, Pune, India
shrikant.salve@gmail.com
[2]Vishwakarma Institute of Technology, Pune, India

Abstract. A road safety is one of the major concerns, which is leading cause of deaths and injuries in India. In the year of 2019, around 440k of road accidents happened, killing about 150k people and 450k injured. However, the report also discussed that, 4,775 accidents reported due to potholes on the roads. The potholes also cause a problem of traffic congestions especially in rainy season and the people take one to two hours extra to reach their homes from offices. There are few mobile apps available dealing with potholes and the most of them are developed by local governments. Majority of them are failed due to poor usability. To address these issues, we have developed a prototype of mobile app, named as *G-connect*. This research paper briefly discusses about *G-connect* and mainly focused on usability evaluation of it. Ten participants were involved in usability evaluation of this prototype using SUS scale. The mean SUS score of all participants is 74 out of 100, which is good. This suggests that *G-connect* is acceptable app, but having minor usability issues like some users unable to identify the use of 'support' button and some users have faced problem to pin exact location.

Keywords: Road Safety, Potholes, Mobile Application, Usability Evaluation, SUS, *G-connect*

1 Introduction

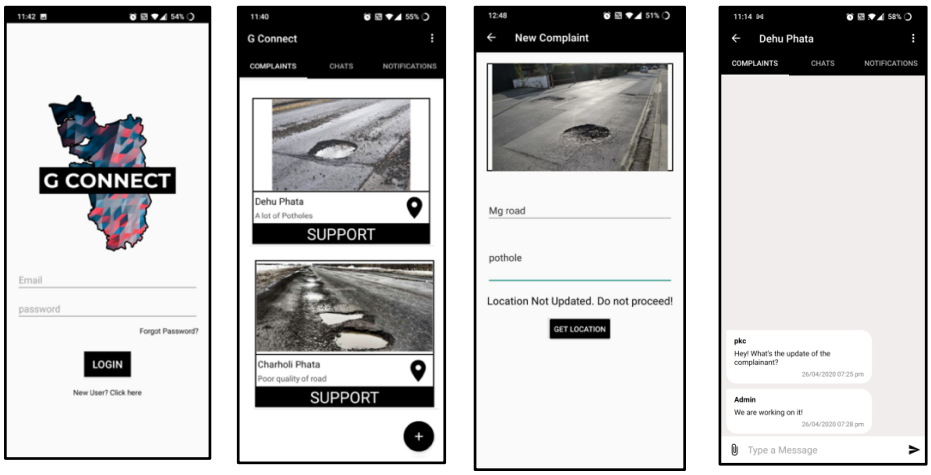
Road accidents and injuries due to potholes on the roads, are major safety issues for the developing country like India. According to the Government of India report in 2019, about 4,775 accidents were reported due to potholes [1]. Such accidents damage the government properties and vehicles, having thousands of casualties. The potholes also cause problem of traffic congestions especially in rainy season. People take one to two hours extra to reach their homes from offices.

There are several existing mobile applications like *Meri Sadak* [2], *RoadDetector* [3], *IntentsGo* [4], *MyBMC24x7*[5] for road safety related issues in India, including for potholes and related reporting. These apps are having many usability issues like inappropriate notifications [2], improper error handling [3], incompatibility with mobile devices [4] and inability of users to register [5]. Also, some functionalities like

'Map' do not work. The mobile app for automatic pothole identification using ultrasonic sensors developed by [6] can be life-saving tool. The machine learning and deep learning techniques like least squares support vector machine, artificial neural networks [7] and convolution neural networks [8, 9], have been used for potholes classification [10]. The automatic pavement distress recognition system [11] can perform real-time identification of potholes. The literature has suggested complex machine learning techniques to solve the potholes related problems. The existing mobile applications related with road safety have discussed about many usability issues. Therefore, we have proposed mobile app through which people can report their issues related to potholes directly to local government. The design and working of mobile app is explained in the next section. The later section discusses about the usability evaluation of the mobile app using SUS.

2 Proposed Prototype *G-connect* Mobile App

We have named the proposed app for the potholes as '*G-connect'* in which 'G' stands for the Government and 'connect' refers to getting connected to the Government. This prototype is developed by using software tool - 'Adobe XD'.



(a) (b) (c) (d)

Fig.1. Screen-shot of *G-connect* app (a) Login screen (b) Main screen displaying three functionalities (c) New complaint screen (d) Chat screen for getting status of complaint.

The fig.1 depicts the four screen-shots of prototype of *G-connect* mobile app. Once the user logged into the app [fig.1 (a)], he/ she can see three sections- complaints, chats and notifications as shown in fig.1 (b). In complaints section, the user can view the existing complaints raised by other people. The user can also support the complaints raised by others, by clicking on 'support' button related with the complaint. Also, the user can give new complaints by pressing '+' icon. After that, the user has to

take a photograph of the pothole and has to mention the physical location of to give a new complaint as depicted in fig.1 (c). In chat section shown fig.1 (d), the user can chat with administrator, who is a local government person or a member of special task force, regarding his/her complaint. The special task force has been prepared by local government to address pothole issues received through this app. Through the notification section, the administrator can notify to the user regarding acknowledgement or necessary action initiated.

This app is specifically developed for local city government. The user complaints are received by administrator or special task force. According to complaint location, the pothole is identified, repaired and the photo of repaired work is uploaded onto the app and the user or complainant is informed about complaint follow-up through the notification section. The next section discusses about the usability evaluation of the mobile app using SUS.

3 Usability Evaluation of *G-Connect*

We have used System Usability Scale (SUS) [12] for usability evaluation of *G-connect*. SUS is widely used standardized questionnaires for assessment of perceived usability. The user rate (or score) all ten questions, that range from 0 (poorest rating) to 4 (best rating). Based on user responses, SUS score is calculated. For odd-numbered questions, a value - 1 is subtracted from the intermediate score and for the even-numbered questions, the score is subtracted from 5. A standard score is calculated using the sum of the above adjusted scores, then multiplied by 2.5.

For the pilot testing, ten participants including six females and four males, from different professions and age groups have been selected. The six participants are undergraduate students, two are faculty members of engineering college, one is housewife and remaining one is retired teacher. The participants have used the prototype of *G-connect* app and filled the SUS questionnaire through 'Google Form'. Because of COVID'19 situation in India, we have used this online questionnaire to record participant responses for SUS score.

The SUS questionnaire is investigated for assessing usability of *G-connect*. The graph in fig.2 presents the SUS score of all ten participants. The mean of all SUS scores of ten participants is about 74/100. There are rules for interpretation of the SUS score [13]. If the SUS score is over 85, the system/product is highly usable, in the range from 70 to 85, it is characterized from good system. If the SUS score is in the range of 50 to 70, it shows that the system is acceptable, but it has some usability problems and needs improvement. Therefore, according to the criteria of Bangor et al. [14], the usability of the evaluated *G-connect* prototype is a good level, though it has some usability problems. Few participants have problem in identifying all facilities related to complaint section. In complaint section, user can give new complaint and also support other user's complaints by pressing 'support' button. But, some of the

users are unable to identify usage of ‘support’ button. Some users have faced problem to pin exact location of the pothole. The graph shown in fig.2 below depicts the SUS scores of all ten participants, which indicates that developed app is good usable app. The Y-axis represents the SUS score and the X-axis is used for participants.

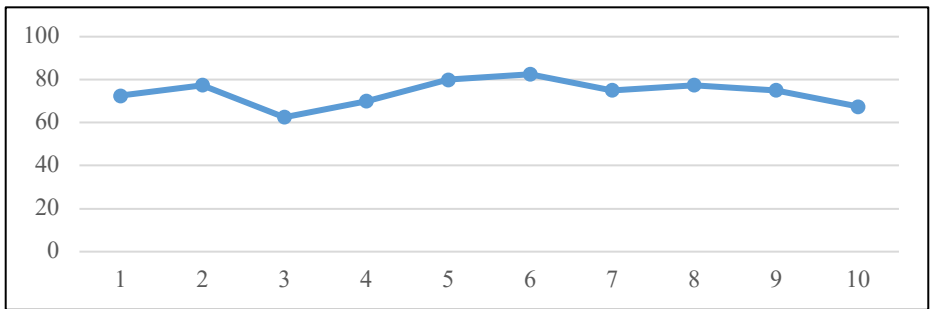


Fig.2. System Usability Score (SUS) Graph

Conclusion

G-connect is simple mobile app to deal with issue of potholes on the roads. The mean SUS score for this app is about 74 out of 100, which suggests that it is a good and usable system [14]. This app have few minor usability issues. Few participants have faced a problem to identify the use of ‘support’ button in complaints section and some users are not able to pin the exact location of the pothole.

In future, we will rectify the identified usability issues and build the final version of *G-connect*. Also, this is a pilot study conducted with just ten users, which can be extended to large sample size for effective usability evaluation.

Acknowledgement

We are thankful to Saumya Rashu, Vaishnavi Ghule, Prajyot Chinchane and Omkar Rakshe, the final year computer engineering students at MIT AOE Pune, India, who have helped in developing this mobile app. We also appreciate the users, who voluntarily participated in usability evaluation process.

References

1. Road Accidents in India 2019, https://morth.nic.in. Last accessed 25 May 2021.
2. *“Meri Sadak”*,v.6.0.6, (Pradhan Mantri Gram Sadak Yojana, 16 March 2021), Android Application, Government of India.
3. *“RoadDetector”*, v.3.0, (Us Project, 5 February 2020), Android Application, Us Project.
4. *“Intents Go”*, v.1.12, (Intents Mobi Private Limited, 21 May 2021), Android Application. Intents Mobi Private Limited.

5. *"MyBMC24x7*", v.3.10, (Publisher mobile.support@mcgm.gov.in, 10 March 2021), Android App, Bumbai Muncipal Corpporation, Mumbai.
6. Madli, R., Hebbar, S., Pattar, P., Golla, V.: Automatic Detection and Notification of Potholes and Humps on Roads to Aid Drivers. IEEE Sensors Journal, 15, 4313-4318 (2015).
7. Hoang, N-D.: An artificial intelligence method for asphalt pavement pothole detection using least squares support vector machine and neural network with steerable filter-based feature extraction. Advance Civil Engineering (2018).
8. Aparna, Bhatia, Y., Rai, R., Gupta, V., Aggarwal, N., Akula, A. et al.: Convolutional neural networks based potholes detection using thermal imaging. Journal of King Saud University- Computer and Information Sciences (2019).
9. Varona, B., Monteserin, A., Teyseyre, A.: A deep learning approach to automatic road surface monitoring and pothole detection. Pers Ubiquit Comput 24, 519–534 (2020).
10. Thomas, A., Harikrishnan, P.M., Nisha, J.S., Gopi, V.P., Palanisamy, P.: Pothole and Speed Bump Classification Using a Five-Layer Simple Convolutional Neural Network. In: Gunjan V.K., Zurada J.M. (eds) Proceedings of International Conference on Recent Trends in Machine Learning, IoT, Smart Cities and Applications. Advances in Intelligent Systems and Computing, vol 1245. Springer, Singapore (2021).
11. Tedeschi, A., Benedetto, F.: A real-time automatic pavement crack and pothole recognition system for mobile android-based devices. Advanced Engineering Informatics 32:11–25 (2017).
12. Brooke, J.: SUS: A "quick and dirty" usability scale. In Usability evaluation in industry. Taylor & Francis, 189-194 (1996).
13. Konstantina, O., Nikolaos T., Christos K.,: Perceived Usability Evaluation of Learning Management Systems: Empirical Evaluation of the System Usability Scale, International Review of Research in Open and Distributed Learning, 16(2), 227-246 (2015).
14. Bangor, A., Kortum, P., Miller, J.: An Empirical Evaluation of the System Usability Scale. International Journal of Human-Computer Interaction, 24(6), 574–594 (2008) doi:10.1080/10447310802205776.

Pilot implementation: Organizational alignment when implementing an IT-system

Linnea Krista Herbæk, Carl Emil Derby Hansen, and Torkil Clemmensen[0000-0002-0934-2336]

Copenhagen Business School, Howitzvej 60, DK-2000, Frederiksberg, Denmark
{Lihe18ad; Caha18aj}@student.cbs.dk, tc.digi@cbs.dk

Abstract. Pilot implementation can be seen as a socio-technical design approach. This paper presents a design case that focused on the optimal implementation and organizational change process when a new IT system is to be implemented in an organization. The case was the pilot implementation of new self-service-oriented IT system for both customers and employees in a Danish SME 'proptech' company. 'Proptech' denote companies dealing with property and technology. The underlying epistemic view of the design case was that of functional pragmatism, and the collection and interpreting of data was aimed at analyzing how the case company managed to implement a new IT system from technical, organizational, and economical perspectives. The analysis revealed that though the case company did not do what current theory in the area prescribes, they still thrived with their pilot implementation. Our position is thus that organizational alignment is an important aspect of pilot implementation, and that it can be analyzed by analyzing various management practices related to the pilot implementation.

Keywords: Pilot implementation, Organizational alignment, Organizational structure.

1 Introduction

Organizational alignment is the process of aligning the organizational structure, culture, and resources with the individuals in the organization. In connection with pilot implementation, it is the process of giving the team behind the pilot implementation the best environment and resources that compliments the development, feedback loop and implementation of the new IT system. This paper argues for the importance of organizational alignment for a pilot implementation to thrive when implementing an IT system. Current theory about pilot implementation presents it as temporary opportunity to let a part of the target users (employees) experiences how the new system will be like to work with [1–3], see also [4–7]. This paper adds to the current theory about pilot implementation by focusing on the organizational alignment that is necessary for pilot implementation.

This paper builds on prior research done by Herbæk & Hansen [8]. They did a case study in a small company that implemented a new IT system through pilot

implementation. The findings from the case study supports this paper's points, and the case study highlights strengths and limitations in the studied company's organizational structure and how it affected their implementation process. Herbæk & Hansen's [8] design case focused on the optimal implementation and change process when a new IT system is to be implemented in an organization. The case was the pilot implementation of new self-service-oriented IT system for both customers and employees in a Danish SME 'proptech' company. 'Proptech' denote companies dealing with property and technology. The underlying epistemic view of the design case was that of functional pragmatism, and the collection and interpreting of data was aimed at analyzing how the case company managed to implement a new IT system from technical, organizational, and economical perspectives. The analysis revealed that though the case company did not do what current theory in the area prescribes, they still succeeded with their pilot implementation. Below we unfold the findings and the argument from the case.

2 Objectives

There are three main objectives for this position paper:

- To emphasize the importance of ensuring alignment between the technical side and organizational side of a pilot implementation.
- To contribute to the design of the optimal pilot implementation framework/process/model.
- To contribute with the findings of a conducted case study on pilot implementation

3 Organizational alignment

3.1 Why is organizational alignment crucial for conduction pilot implementation?

Pilot implementation is defined by "a field test of a properly engineered, yet unfinished system in its intended environment, using real data, and aiming – through real-use experience – to explore the value of the system, improve or assess its design, and reduce implementation risk" [2]. With the introduction of an unfinished system comes certain obstacles which the management must focus on to make the development and implementation process thrive.

We use the notion of organizational alignment to describe interventions that ensure that the pilot implementation and the current and the long-term business goals of the organization are aligned. It is well established that one of the key factors for successful organizations is the close linkage of its IT strategy and business strategy [9], and this is also true for socio-technical design and business strategy. Different conceptualizations of organizational strategies exists, and we follow the idea of strategy as practice [10] where the study object is how management practices are used to put strategy into

practice. Thus, the practices of managers should be aligned with the practices of designers to achieve successful pilot implementation.

However, if the organization is not aligned with the pilot implementation it could lead to stagnation of the implementation process, and to inefficiency in the work, test and development flow [8]. Herbæk & Hansen analyzed organizational alignment initiatives based on the theory of the three legged stool [11] that focuses on decision-making rights, performance evaluation and reward systems, and supported by the theories of Kotter's 8 steps for organizational change [12, 13] and the principles for lean startups of companies [14, 15]. The three-legged stool stipulates that decision-making rights, performance evaluation and reward systems must be in balance and aligned to the company's current situation to create success in what is desired to be accomplished. A finding from analysis of these three areas was that the initiatives of the case company management affected the pilot implementation in both positive and negative ways, and that though the case company did not do what current theory in the area prescribes, they still succeeded with their pilot implementation [8].

3.2 The importance of data management in pilot implementation

In the case of the company studied by Hansen and Herbæk [8], the company desired to develop and implement a new IT system through pilot implementation. In theory, pilot implementation requires iterative testing and flowing communication between the developers and employees participating in the testing of the new unfinished system [2]. The developers require the decision rights to develop and test anything they find valuable for the new system, if it is based on the feedback generated by the. It is important to minimize the bureaucracy and allow developers to test as this generates feedback about the system in development which is crucial in pilot implementation. For the feedback loop to generate sufficient and valid feedback about the system in development, good communication tools and data management is required.

However, the design case revealed a significant amount of freedom in the communication and reporting of feedback between the developers and employees, but a lack of structured data management when multiple employees reported feedback about the system in development to the developers. The lack of structured data management led to confusion among the developers, as the feedback data got unmanageable as there was a lack of overview [8]. The case study finding is thus that pilot implementation requires a refined data management structure to get the most out of the feedback generated from the employees.

3.3 Alignment of organizational reward systems to design practices

The use of an unfinished system in real-use experience tends to be accompanied by lower efficiency in the workflow in which the system is used. This can lead to employees choosing to not utilize the new system in full, as it does not generate the same outcome as the already established system. This resistance needs to be addressed by the management. If the organizational performance evaluation is not aligned with the need

for testing the new system, developers will have a harder time getting the necessary feedback in which they use to further develop the new system.

In the case of the company studied by Hansen & Herbæk [8], the management lowered their performance evaluation goals as a means to empower the employees' engagement in testing the new system. The lowering of the organization's performance evaluation goals furthered the development of the new system through organizational alignment and ensured that the employees' individual needs and the organization's vision and strategy was aligned [8]. The engagement of the employees used for testing as well as the developers is crucial. The allocated decision rights and the freedom that follows these decision rights needs support to sustain the level of engagement. This can be done through financial incentives. In the case of the company studied by Hansen & Herbæk [8], the organization allowed for the full time workers to obtain warrants, which represents an equity in the firm, as part of their salary. The inclusion of warrants generates a sense of ownership with the employees and further boosts the engagement, as they will know that whatever obstacles they face now are worth it, as it benefits the organization as well as them as individuals. The organizational alignment of organizational reward systems to design practices therefore benefits the development, testing and implementation process of a pilot implementation [8].

4 Discussion and conclusion

This position paper asks what management can do to support a successful pilot implementation and answer the question by pointing to organizational alignment. The major insight from the design case was that though the case company management did not do exactly what current theory in the area prescribes, their initiatives were still important to the pilot implementation. Thus, the practices of managers were aligned with the practices of designers and users to achieve successful pilot implementation. This contrasts somewhat current theory that do not mention management except as project management [2] or as a stop/go decision maker for pilot implementation [1]. Our position is that organizational alignment is an important aspect of pilot implementation, and that it can be analyzed as Herbæk & Hansen [8] did it by analyzing various management practices related to the pilot implementation. Future research on pilot implementation may analyze management not as external stakeholders but as co-designers.

We provide two starting points for further research on organizational alignment in pilot implementation. First, the management of feedback data should be studied by looking at how the feedback data management systems needs to align with the feedback requested and tests that are being conducted. Furthermore, innovative accounting practices should be studied with an eye on measuring progress using test and quantitative data derived from the tests conducted to supplement the qualitative descriptions of the progress that the development team has done. Second, the design of reward structures and performance evaluations related to the individuals participating in the pilot implementations should be aligned to generate motivation to generate feedback, as the individuals participating in the are likely to not utilize the system in development as much as needed for pilot implementation if it affects their performance bonuses.

5 References

1. Mønsted, T., Hertzum, M., Søndergaard, J.: A Socio-Temporal Perspective on Pilot Implementation: Bootstrapping Preventive Care. Comput. Support. Coop. Work. 29, 419–449 (2020).
2. Hertzum, M., Bansler, J.P., Havn, E.C., Simonsen, J.: Pilot implementation: Learning from field tests in IS development. Commun. Assoc. Inf. Syst. 30, 20 (2012).
3. Hertzum, M.: Organizational Implementation: The Design in Use of Information Systems. Synth. Lect. Human-Centered Informatics. 14, i–109 (2021).
4. Bansler, J.P., Havn, E.: Pilot implementation of health information systems: Issues and challenges. Int. J. Med. Inform. 79, 637–648 (2010).
5. Hertzum, M., Simonsen, J.: Effects-driven IT development: an instrument for supporting sustained participatory design. In: Proceedings of the 11th Biennial Participatory Design Conference. pp. 61–70 (2010).
6. Hertzum, M., Manikas, M.I., á Torkilsheyggi, A.: Grappling with the future: The messiness of pilot implementation in information systems design. Health Informatics J. 25, 372–388 (2019).
7. Torkilsheyggi, A. á, Hertzum, M.: User participation in pilot implementation: Porters and nurses coordinating patient transports. In: Proceedings of the 26th Australian Computer-Human Interaction Conference on Designing Futures: the Future of Design. pp. 290–299 (2014).
8. Hansen, C.E.D., Herbæk, L.K.: Proper - A study of a change and implementation process (unpublished bachelor thesis), (2021).
9. Baets, W.: Aligning information systems with business strategy. J. Strateg. Inf. Syst. 1, 205–213 (1992).
10. Jarzabkowski, P.: Strategy as practice: recursiveness, adaptation, and practices-in-use. Organ. Stud. 25, 529–560 (2004).
11. Brickley, J., Smith, C., Zimmerman, J.: Managerial economics and organizational architecture. McGraw-Hill Education (2015).
12. Kotter, J.P.: Leading change. Harv. Bus. Rev. 2, 1–10 (1995).
13. Appelbaum, S.H., Habashy, S., Malo, J., Shafiq, H.: Back to the future: revisiting Kotter's 1996 change model. J. Manag. Dev. (2012). https://doi.org/10.1108/02621711211253231.
14. Ries, E.: The lean startup: How today's entrepreneurs use continuous innovation to create radically successful businesses. Currency (2011).
15. Eisenmann, T.R., Ries, E., Dillard, S.: Hypothesis-driven entrepreneurship: The lean startup. (2012).