

# WHO ARE THE USERS OF A VIDEO SEARCH SYSTEM? CLASSIFYING A HETEROGENEOUS GROUP WITH A PROFILE MATRIX

Max Kemman MSc, Martijn Kleppe MA, Prof. Dr. Henri Beunders

Erasmus University Rotterdam

## ABSTRACT

Formulating requirements for a video search system can be a challenging task when *everyone* is a possible user. This paper explores the possibilities of classifying users by creating a Profile Matrix, placing users on two axes: experience and goal-directedness. This enables us to describe the characteristics of the subgroups and investigate differences between the different groups.

We created Profile Matrices by classifying 850 respondents of a survey regarding a requirements study for a video search system. We conclude that the Profile Matrix indeed enables us to classify subgroups of users and describe their characteristics. The current research is limited to descriptions of subgroups and analysis of differences between these subgroups. In the future, we want to research what these differences mean with regard to the users' performance and acceptance of a video search system and explore the use of a profile matrix for other types of search systems.

## 1. INTRODUCTION<sup>123</sup>

In the last few decades, system design has shifted from system-focused to user-centred. Whereas previously, research focused on the information source and how it would or could be used, this focus has shifted to the users and their needs and behaviours [1]. Although this shift is sometimes received with scepticism [2], investigating users and their performance or acceptance of product design [3], as well as their capabilities and weaknesses [4], helps developers to create systems that are intuitive and helpful to the users.

However, behaviour of individuals is difficult to assess and generalize. A difficulty in requirements elicitation can arise when *everyone* is a potential user, making it difficult to choose users to investigate their behaviour. To address this difficulty, we try to distinguish several subgroups of users which we can then investigate in order to get a picture as complete as possible of all the users.

Groups can be distinguished in several ways, [1] describes two general approaches; by role (e.g. people with

the same occupation or people who are all patients) or by demographic characteristic (e.g. age, gender or socioeconomic status).

When we specify user groups, e.g. by choosing several (sub)groups based on age, we want to describe characteristics of these subgroups and investigate whether the chosen subgroups show differences in their needs and behaviours. To research this, we developed the *Profile Matrix*. The idea for such a Profile Matrix is inspired by the EuropeanaConnect project<sup>4</sup>, in which personas were characterized along two axes; a "search literacy" axis and a "navigational/explorative behaviour" axis [5]. We adjusted these axes to fit the goal of our project in which we develop a video search engine. The "search literacy" axis became a similar "experience" axis and the "navigational/explorative behaviour" axis became the "goal-directedness" axis.

By analysing the position of the specified user groups within this Profile Matrix, we can describe their characteristics and investigate their differences. With these group classifications, usability researchers can make better choices collecting respondents for requirements elicitation and system developers can get a better idea of the heterogeneous group of potential users, giving them a better idea of who might benefit from what functionality.

The research question of this paper is the following:

*"Can we classify (sub)groups of users in a heterogeneous group of people and describe their characteristics using the Profile Matrix?"*

We discuss the Profile Matrix within the context of the AXES project<sup>5</sup>, in which an online video search system is to be developed which should enable a heterogeneous group of people to find all sorts of videos.

In this paper we describe the method used to classify users in our research, we present findings from our survey among 850 respondents, discuss the research question and future directions for the Profile Matrix.

## 2. PROFILE MATRIX

### 2.1. Relevance of the axes

The vertical axis *experience* refers to skills or knowledge that users have, sometimes referred to as "user expertise"

<sup>1</sup> Corresponding author: kemman@eshcc.eur.nl

<sup>2</sup> We thank Allerd Peeters MA and Anna van der Zalm MSc for their insightful comments, and our colleagues from the AXES consortium for their help setting up the survey.

<sup>3</sup> We are grateful for the financial support from the EU FP7 project AXES – Access To Audiovisual Archives ICT-269980.

<sup>4</sup> <http://www.europeanaconnect.eu>

<sup>5</sup> <http://www.axes-project.eu>

[3], related to the task of searching and consuming online videos.

Previous research has shown that the experience or expertise of participants influences the results of usability studies. In [6] a distinction is made between “novices” and “experts”, based on how often a similar product was used before. They found that after an experiment, experts reported more usability problems than novices, although the problems reported by novices were rated more severe; while usability issues reported by experts were related to efficiency and functionality, issues reported by novices could prevent task completion.

According to [3], another difference between experts and novices is found in the way they form mental models of a system. Experts may have a mental model based on previous experience with similar systems, which can have either a positive or negative influence; such old mental models can help the expert to understand the system more easily, but can also lead to expectations that conflict with the current interface. In contrast, novice users learn a new interface without an existing old mental model.

As such, it is important to understand how experienced the potential users are and where they acquired previous experience; this allows us to better understand their comments during usability studies, as well as their expectations of the system under development.

The horizontal axis *goal-directedness* refers to the search behaviour of users. In [7], a distinction is made between “navigators” and “explorers” based on their behaviour during search tasks. Characteristics of navigators are that they: 1) exhibit few deviations from the search path, 2) appear to tackle problems sequentially and 3) are more likely to revisit domains. Overall, it appears that navigators follow a more or less direct path from problem specification to the problem resolution.

In contrast, characteristics of explorers are that they: 1) tend to branch the search path frequently, 2) submit many queries during a search session and 3) visit many new domains. Explorers do not follow a direct path from problem specification to the problem resolution. As these users employ different search strategies, they need different support from the search systems; while navigators might benefit the most from support to follow the path as quickly as possible and revisit previous websites easily, explorers might benefit the most by showing recommendations and allowing the user to follow several search paths simultaneously [7].

The relation between these two axes is shown in [8], where it was found that search experience can influence search behaviour, as more experienced users plan more and are as such more goal-directed.

## 2.2. Creation of the axes

The creation of the axes of the Profile Matrix consisted of two steps. First, we constructed questions for a survey based

on literature as described below. We then evaluated these questions based on the results of the survey described in section 3. The result of this evaluation is the Profile Matrix as described in this paper and in [9].

The experience axis focuses on the experience with searching online video as well as with consumption of online video. In order to provide a measurement of experience, we ask for the use of several features of the most popular video website YouTube<sup>6</sup>, where functionality overlapped with the functionality of the British iPlayer<sup>7</sup> and the Dutch equivalent Uitzending Gemist<sup>8</sup>. The scale of the vertical axis is computed using the following items, all of which are rated on a Likert scale of 0-4:

- Features
  1. *Share*: Sharing with others (e.g. via Twitter, Facebook)
  2. *Subtitles*: Use of subtitles
  3. *Reuse*: Reusing (parts of) video for own purposes
  4. *Full-screen*: Use of full-screen video
  5. *Quality*: Switching video quality
- Search
  6. *Success*: Frequency of successful finding
  7. *Ease*: Ease of finding

This results in a total figure on a scale of 0-28. We tested the reliability of the scale using Cronbach's  $\alpha$ , with which we found that  $\alpha=0.711$ <sup>9</sup>, which is considered an acceptable value [10]. From this we conclude that these items provide a statistically reliable scale.

To present the result graphically, we transform the figure of experience linearly to a scale of -2 (inexperienced/novice) to 2 (experienced/expert).

The goal-directedness axis focuses on the search behaviour with the extremes navigational versus explorative behaviour, to which we refer with goal-directed search versus non-goal-directed search. In order to analyse this axis, we used the '*dimensional typology of browsing*' as a basis [11], in which the behaviour of browsing is characterized according to six dimensions: purpose, goal, content knowledge, structure knowledge, location knowledge and resource focus.

From these six dimensions, we created four items for our scale; the horizontal axis is computed with the following items, for which all items were rated with a semantic differential scale of 0-6:

---

<sup>6</sup> <http://www.youtube.com>  
[http://www.youtube.com/t/about\\_essentials](http://www.youtube.com/t/about_essentials)

<sup>7</sup> <http://www.bbc.co.uk/iplayer>

<sup>8</sup> <http://www.uitzendinggemist.nl>

<sup>9</sup> To address concerns of endogamy, we randomly separated the dataset in S1 (n=200) and S2 (n=650) to analyse independent values for Cronbach's  $\alpha$ . We found  $\alpha=0.670$  for S1 and  $\alpha=0.720$  for S2; which we both consider acceptable.

1. *Content*: Knowing what to search
2. *Location*: Knowing where to search
3. *Structure*: Knowing how to search
4. *Resource*: Adjusting query or following links

This resulted in a total figure on a scale of 0-24. We tested the reliability of the scale using Cronbach's  $\alpha$ , with which we found that  $\alpha=0.753$ <sup>10</sup>, which is considered an acceptable value. From this we conclude that these items provide a statistically reliable scale.

The 'purpose' dimension did not provide a meaningful measure as we found in our survey that it did not correlate with the other dimensions: those with clear information needs can both navigate or browse; the same goes for those with an entertainment need.

We do not use the 'goal' dimension as a measurement, but as the scale itself on which we position users. The definition of browsing is very broad and a matter of discussion [1], [11], while our main interest is how goal-directed people search. Therefore we chose to use the 'goal' dimension as the scale on which we position people, avoiding the semantic discussion of 'browsing'. We assume that people who know what, where and how to search and who adjust queries to search further are goal-directed, whereas people who do not know what, where and how to search and who follow links to search further are not goal-directed.

As with the vertical axis, we transform the figure of goal-directedness linearly to a scale of -2 (goal-directed) to 2 (non-goal-directed) to present the result graphically.

### 3. PROFILE MATRIX RESULTS<sup>11</sup>

In our research of users of online video systems, we chose groups based on demographic characteristics: gender, age, education and nationality, and on frequency of online video consumption. For this research, 318 respondents from the UK, 311 from the Netherlands and 341 from Germany completed our online survey, summing up a total N of 970<sup>12</sup>, representative in distribution of age, gender and education [9]. For this paper, we present the results for the groups based on frequency of online video consumption and age. The other results can be found in [9]. Of the 970 respondents, 850 watched video online. The overall distribution of these 850 respondents in the Profile Matrix is shown in figure 1, from which we can see that overall, users

appear to be best represented by the goal-directed inexperienced quadrant of the matrix.

When reporting the Profile Matrix for (sub)groups, we use the averages of the groups specified.

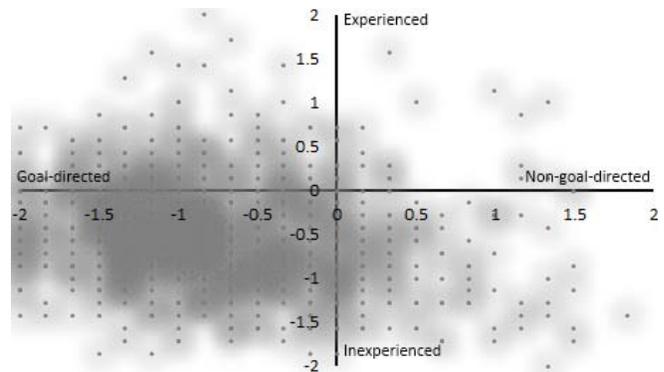


Figure 1 – Overall distribution of respondents on Profile Matrix

#### 3.1. Frequency of online video consumption differences

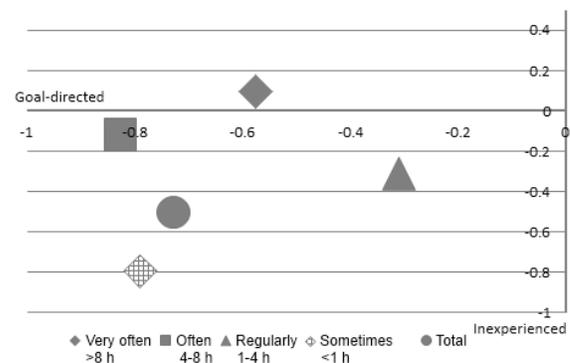


Figure 2 – Zoomed-in Profile Matrix for frequency of online video consumption, in hours per week

The results of the Profile Matrix for the frequency of online video consumption can be seen in figure 2. Here we can see that with increased frequency, respondents are more experienced, with the highest frequency of consumption on the experienced-side of the matrix. We tested this with an ANOVA, with which we found a significant difference between the frequency of use on the experience axis with  $F(3,846)=113.459$ ,  $p<0.05$ , but no significant difference on the goal-directedness axis with  $F(3,846)=2.220$ ,  $p>0.05$ . From the post-hoc Bonferroni-test, we found that all subgroups were significantly different with each other with  $p<0.05$ . Every group with higher frequency of consumption is thus significantly more experienced, which shows that the experience axis gives a meaningful indication of experience.

#### 3.2. Age differences

The results of the Profile Matrix for the chosen age-groups can be seen in figure 3. Here we see that all age-groups are

<sup>10</sup> To address concerns of endogamy, we analysed independent samples of Cronbach's  $\alpha$  for S1 and S2. We found  $\alpha=0.786$  for S1, and  $\alpha=0.744$  for S2, which we both consider acceptable.

<sup>11</sup> The survey and respondent data relevant to this paper are available open access through <http://persistent-identifier.nl/?identifier=urn:nbn:nl:ui:13-jo0i-q4>

<sup>12</sup> We had 1265 dropouts and 31 invalid responses. The high number of dropouts is due to our efforts to get a sample representative of the general population.

more or less inexperienced and goal-directed, as the averages fall in that quadrant of the matrix. As all groups fall within the scale of -1 to 0 on both the experience and goal-directedness axes, we only provide a zoomed-in version of the Profile Matrix to better discern the differences.

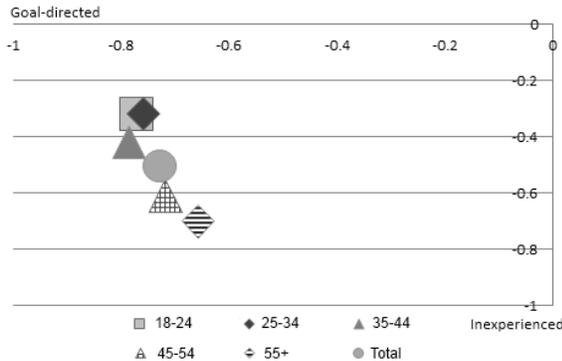


Figure 3 – Zoomed-in Profile Matrix based on age-groups

From this zoomed-in version, it appears that the older respondents are, the less experienced as well as less goal-directed they are.

To test whether these differences are significant, we performed an ANOVA, with which we found a significant difference between groups, but only for the experience axis with  $F(4,845)=16.115$ ,  $p<0.05$ , while for the goal-directedness axis we found  $F(4,845)=0.966$ ,  $p>0.05$ . From the post-hoc Bonferroni-test, differences were found with  $p<0.05$  between:

- 18-24 and 45-54
- 18-24 and 55+
- 25-34 and 45-54
- 25-34 and 55+
- 35-44 and 45-54
- 35-44 and 55+

From this we see that a significant difference seems to exist between respondents below 45 years and respondents of 45 years and older, where the latter is less experienced than the former.

#### 4. CONCLUSION & DISCUSSION

From our research, we conclude that we can answer the research question positively; with the Profile Matrix, we can classify (sub)groups of users and describe their characteristics.

The research discussed here is limited to descriptions of subgroups and analysis of differences between these subgroups. In the future, we want to research what these differences mean with regard to the users' performance and their acceptance of a video search system. Moreover, the current research is limited to classifications based on demographics. In the future, we will use the Profile Matrix to compare groups based on roles, namely broadcast

professionals, journalists, academic students, researchers & educators and home users [9], to further test and develop the Profile Matrix, as well as to investigate the characteristics of these user groups and the differences between these groups.

While our focus was on consumption of online audiovisual content, experience and goal-directedness are not only relevant for video search; we can imagine specifically adjusted Profile Matrices for other search systems, focusing on e.g. websites or photos.

Creating a Profile Matrix will help in choosing groups of participants during usability research, as well as help the developers to better understand the end-users of the system. It does not provide a guideline of how to develop a system, its value lies in providing a better understanding of the heterogeneous group of users, as we can classify subgroups within this large group of people. With these subgroups, a better understanding can be obtained of the different needs, capabilities and weaknesses of the users.

#### REFERENCES

- [1] D. O. Case, *Looking for Information - A Survey of Research on Information Seeking, Needs, and Behavior*, vol. 29, no. 3. San Diego Academic Press, 2002.
- [2] D. A. Norman, "Technology first, needs last: the research-product gulf," *interactions*, vol. 17, no. 2, pp. 38-42, 2010.
- [3] Y. Liu, A. Osvalder, and M. Karlsson, "Considering the importance of user profiles in interface design," no. May, 2010.
- [4] J. Preece, Y. Rogers, and H. Sharp, "Interaction Design: Beyond Human-Computer Interaction," *Design*, vol. 18, no. 1, pp. 68-68, 2002.
- [5] K. G. Rasmussen, R. Iversen, and G. Petersen, "M3.2.3 Personas Catalogue," *EuropeanaConnect*, 2010. <http://www.europeanaconnect.eu/>
- [6] J. Sauer, K. Seibel, and B. Rüttinger, "The influence of user expertise and prototype fidelity in usability tests.," *Applied ergonomics*, vol. 41, no. 1, pp. 130-40, Jan. 2010.
- [7] R. W. R. White and S. M. Drucker, "Investigating behavioral variability in web search," *Proceedings of the 16th international conference on World Wide Web - WWW '07*, pp. 21-30, 2007.
- [8] C. Hölscher and G. Strube, "Web search behavior of Internet experts and newbies," *Computer networks*, vol. 33, no. 1, pp. 337-346, 2000.
- [9] M. Kleppe, M. Kemman, and H. Beunders, "D1.2 User Requirements Report V1," AXES, 2011. <http://www.axes-project.eu/>
- [10] A. Field, *Discovering Statistics Using SPSS*, vol. 2. Sage Publications Ltd, 2009, p. 675.
- [11] S. J. Chang and R. E. Rice, "Browsing: A Multidimensional Framework.," *Annual review of information science and technology (ARIST)*, vol. 28, pp. 231-76, 1993.