The potential of Recall and Precision as interface design parameters for information retrieval systems situated in everyday environments

Ayman Moghnieh
Universitat Pompeu Fabra
C/Tanger 122-140, E-08018
Barcelona, Spain
ayman.moghnie@upf.edu

Josep Blat
Universitat Pompeu Fabra
C/Tanger 122-140, E-08018
Barcelona, Spain
josep.blat@upf.edu

ABSTRACT
In this paper, we investigate ways for a tighter integration of IR and HCI in new urban contexts, as HCI expands its reach outside the workplace towards environments where efficiency and performance no longer constitute the backbone of interaction requirements. In particular, we propose to use Recall and Precision as design parameters to describe the information settings and performance of situated interfaces acting as retrieval systems in these environments. To explore this notion, we follow an inductive design research process by which different prototypes are designed, developed, and evaluated. Our experience shows that Recall and Precision, as design parameters, help to reflect the information requirements onto the interface design, and contribute to adapting IR to the contemporary challenges it faces, although more work is needed to consolidate its role vis-à-vis the growing ubiquity of computer technologies.

Categories and Subject Descriptors
H.5.2 User Interfaces.

General Terms
Design, Experimentation, Human Factors, Theory.

Keywords
Information Retrieval, Human-Information Interaction, Situated Interfaces, Interface and Interaction Design

1. INTRODUCTION
As computer technologies become more ubiquitous and versatile, and get further integrated in human environments, several genres of situated information interfaces (e.g. interactive peripheral displays, ambient displays, and interactive surfaces) are starting to assume a mediating role between people and digital information spaces in different environments. From an HCI perspective, these situated interfaces, primarily found in public and semi-public environments such as malls, public transportation, building entrances, and public squares, represent new border zones that maintain connectivity and mutual presence between the real and the digital worlds, and actively sustain flows of useful or relevant information towards nearby people who in-turn search, discover, and interact with the displayed information.

The human interaction with information via situated interfaces creates new challenges for conventional information retrieval (IR) systems: first, the relationship between people and digital information spaces becomes more explicit and the technology that supports it more ubiquitous. Second, the human interaction with information spaces adopts a more direct approach supported by the coming of age of new interaction paradigms (e.g. touch, gesture, speech) that emulate the manipulation of objects. Third, the information space hosted by a situated interface tends to be specialized in subjects and themes befitting the environment where the interface is situated, and the goals and interests of the people present in it. Fourth, the interaction properties may vary considerably in terms of interaction duration and the amount of user attention delegated to the situated interface [1].

These challenges, among others [2], justify the search for a tighter coupling of interface and interaction design, and IR systems, by which IR as a supporting technology for interacting with information contributes to making the interface design more transparent and the human-information interaction more fluid and direct. Therefore, we reason that the performance of situated interfaces as IR systems ought to be attuned according to the nature of each specific interaction scenario, given that a maximization of IR performance, may not be adequate for answering the interaction design requirements in all kinds of user experiences with situated interfaces [5, 10]. Consequently, IR performance tilts towards becoming a design issue that determines some of the characteristics of situated interfaces that mediate this interaction.

Currently, two metrics (Recall and Precision) are used to assess the performance of IR systems in response to user queries [3]. Recall is the fraction of retrieved information elements from the entire existing set of elements that are relevant to the user query in the information space. Precision is the fraction of retrieved elements found relevant with respect to the user query, over the entire set of retrieved elements. However, the query as a middleman between humans and information spaces goes against the transparent design of situated interfaces that support a direct interaction with information spaces. In addition, the information spaces hosted by situated interfaces are usually predetermined or pre-queried in accordance with the specific interests of potential

users and the characteristics or nature of the environments where the hosting interfaces are situated. Instead of querying, the explicit momentarily needs of users are answered by direct interaction with the visualized information. This superlatively converts the relevance of the displayed information to the user interests from a performance factor to a design issue.

Therefore, we argue that the definition of Recall and Precision can be loosened or reinterpreted to respectively describe the quantity of retrieved information elements and their visual diversity as displayed on the interface, since relevance is no longer a performance factor from an HCI stance. These two metrics can consequently act as parameters that bind the design and performance of situated interfaces as retrieval systems to the informational expectations of users, by controlling the amount and diversity of visualized information in order to maximize the transparency of their designs to support a direct human-information interaction.

In order to explore this idea further, we followed a line of inductive design research by conceptualizing, designing, and evaluating experimental prototypes. We first introduce two sets of prototypes devised to understand how users perceive the quantity and visible diversity of information objects. We then define parameterization scales for Recall and Precision based on these experiments. In order to develop a thoughtful understanding of how Recall and Precision, which we will consecutively refer to as R and P, can act as design parameters for situated interfaces, we use them in the analysis, design, and evaluation of five different situated interfaces. Next, we investigate how these two parameters can be dynamically controlled by users through the design of two interactive interfaces for searching and browsing news articles. We conclude by assessing our experience and discuss the viability and implications of our approach.

2. RECALL AND PRECISION FROM A PERCEPTUAL STANCE

Figure 1. An instance of the InformationCasserole prototypes

InformationCasserole is a series of video prototypes (figure 1) designed to study the effect that the number of visualized elements (R) has on the way humans perceive the information revealed on the interface. They show classified ads from magazines and newspaper floating on different levels in a glass container filled with slowly moving water. Therefore, their settings emulate a transparent interface design and foster a direct relationship between the human and digital information spaces.

Miller's Law argues that the total number of different objects that humans can simultaneously hold in their working memory is approximately seven [4]. This affects the manner by which information is perceived when the cardinality of the visualized set of objects increases. In particular, there is a natural observable tendency to perceptually cluster or group these objects recursively whenever the perceivable number exceed Miller's threshold. To observe this phenomenon, eight 10 minutes long think-aloud sessions were organized with eight different university students that watched InformationCasserole showing magazine ads progressively being added to the water container, and commented on how the number of ads shown in the casserole affects the way they perceive the set of visualized ads.

We observed that when one object is shown, it tends to engage the subjects in a prolonged and detailed examination. This changes when two to seven objects are displayed since subjects become more interested in identifying relations among the objects and comparing them. The interest in object relations abates with a higher object number, and instead the relations among clusters or collections of objects start to proportionally grab attention. When the number of visualized objects crosses a certain threshold, which we estimate at Miller's number squared, the casserole becomes perceptually saturated and the subjects begin to treat the set of ads as a space, reasoning about different regions in it. In conclusion, we find that the quantity of visualized objects (R) is perceived in four different density thresholds, and to each we accord a parameter value: R=0 for visualizing no or a single object; R=1 for a single collection of seven or less objects; R=2 for seven or less collections; and R=3 for single information space or more than seven squared objects. This is reflected in figure 2.

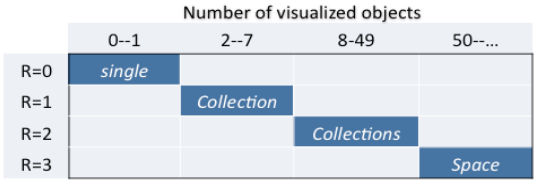

Figure 2. R as a design parameter

In order to study the effects that the visible diversity of information objects (P) has on the manner by which people perceive information, eight paper-based prototypes similar to the InformationCasserole were conceived. Each prototype shows a combination of twelve to fifteen information objects from different genres (e.g. classified ads, news headlines, blog posts, news pictures, movie posters, youtube videos, secondhand goods, and city events). The object genre was emphasized and differentiated by aesthetic design. The visible object diversity encourages people to search for relations among visualized objects [6]. Therefore, the combinations, ranging from one to eight genres, were designed to encourage subjects to search for patterns and relations among the objects. Six twenty minutes think-aloud sessions were organized with subjects whom were asked to search for and identify different genres of objects in each of the eight combinations presented in random order.

As expected, the subjects perceptually clustered the objects primarily in accordance to their genre. However, they sometimes tended to search for inner-divisions in objects of the same genre (e.g. clustering movies according to their cinematic kind or news articles in familiar news categories), or to merge related genres as a single genre (e.g. news articles and blog posts, or movie posters and news pictures). In total, the subjects perceived the diversity of

objects (P) in four different levels, and to each level we accord a corresponding parameter value inversely proportional to the number of visible object genres: the first level is a single-genre diversity (P=3); the second level is a diversity of two to three genres (P=2); the third level refers to diversity of three to four genres (P=1); the fourth level describes a diversity of five to seven genres of objects (P=0). Figure 3 shows the number of visible genres of objects in each of the eight combinations as seen by the subjects, and the P value of each of the four identified diversity levels.

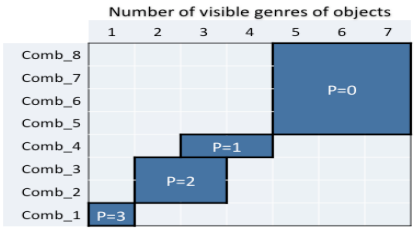

Figure 3. P as a design parameter

3. SITUATED INTERFACES AS IR SYSTEMS

In order to assess how R and P act as design parameters for the information settings of situated interfaces, the following five interfaces that act as retrieval systems in real-world environments were analyzed, and for each a corresponding design was developed and evaluated in settings that resemble or emulate its deployment environment.

The *Arts&Movies* is a situated interface intended for movie theatre lobbies to support the search and discovery of new interesting movies through an animated visualization that draws attention to relationships between movies and concepts. The *DigiJuke* is installed inside a bar to allow people to browse and select music songs on the touch-screen, and play their video clips accompanied by related images on the projection display. The *YouServe* prototype is collocated in a university library lobby to assist people in familiarizing themselves with the available library services, and finding a service relevant to specific needs. The *NewsWall* is a large display situated in the news production room of a broadcasting corporation. The prototype subtly visualizes the constantly evolving news information space on the web. The *MetroWindow* is designed for metro wagons and broadcasts summarized local news about cultural and civic events in the city of Barcelona.

In related works [7, 8] we have argued how R and P, as design parameters, can be quantified during requirement analysis and used alongside other aspects to conceptualize the design of information interfaces. For each situated interface, a couple of designers analyzed the characteristics of three entities being: the deployment environment, the humans present in it, and the adequate information space, which was defined based on an understanding of the needs and goals of the humans alongside the nature of the environment and the information and activity flows that it hosts. Based on this analysis, the designers qualified the values of R and P for each situated interface, and consequently described its information settings, being the quantity of information to visualize and its visible diversity. This qualification of R and P was defined in accordance with several non-disjoint or co-dependent situational aspects of human-information interaction such as:

- The amount of available user attention (e.g. MetroWindow disposes of little attention in contrast with DigiJuke).
- The duration of human interaction with information (e.g. NewsWall remains in contact for prolonged durations, while the interaction with YouServe is more momentarily).
- The convergence or divergence of the information seeking tasks (e.g. YouServe supports finding a specific library service, while Arts&Movies is designed to acquaint people with many movies).

Table 1. Values of R and P parameters for each interface

Situated interface	Recall	Precision
Arts&Movies	2	1
DigiJuke	3	3
YouServe	1	2
NewsWall	1	1
MetroWindow	0	3

The results of this R and P qualification are summarized in table 1. They show how R and P can characterize, from a perceptual stance, the role of a situated interface as an information retrieval engine, and parameterize the design of its information settings accordingly. For example, when the user objectives are to search for specific objects (e.g. *YouServe*), R is minimized, while P can be maximized when the search converges on specific genres (e.g. *MetroWindow*) or minimized when it diverges to cover many genres (e.g. *NewsWall*). A maximized R signals that the interaction tackles a large number of objects. In this case, when P is maximized (e.g. *DigiJuke*), it determines that this large number is a single collection of similar objects, or, when it is minimized (e.g. *Arts&Movies*), it signals that this large number of objects is a visually diversified information space.

The designers also developed the interfaces information architecture and aesthetic design, but these activities lies outside the scope of this paper. The final designs are shown in figure 4.

Figure 4. The situated interfaces final designs

4. USER CONTROL OVER R AND P

Based on the discerned ability of R and P to describe the information settings of situated interfaces and consequently their performance as information retrieval systems, we explored the possibility of allowing users to control them dynamically in classic search and retrieval scenarios. Therefore, we designed two experimental prototypes (figure 5) for querying a large information space of news articles, by which users can set and control the values of both R and P. The prototypes were evaluated to assess the feasibility of this approach and its utility.

The NewSearch prototype collocates two slide-bars adjacently to the query textbox for setting R and P explicitly, and returns an equivalent clustered visualization of news articles. Users control the number of clusters (discerned by color) by P and their average cardinality by R. The 3DQuery prototype uses a tag-map as a new concept for defining user queries, and shows a corresponding map of news articles. The tag-map is a rectangular box where users can place different tags of distinct sizes. The position of each tag determines that of the corresponding cluster of news articles, and the tag size the cluster cardinality.

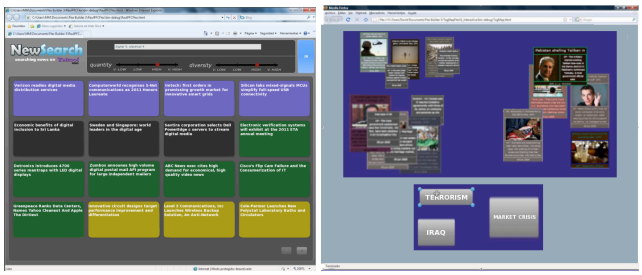

Figure 5. NewSearch (left) and 3DQuery (right) prototypes

Each prototype was evaluated by a different group of ten subjects in the lab. The subjects were asked to browse and read the collection of news articles for fifteen minutes, and then answer a set of open-ended questions concerning their utility and usability. The user evaluations of both prototypes showed that their learning curve is not negligible. Subjects were not naturally inclined to use the slide-bars of NewSearch to control the information settings. An explanation for this may well be that they are accustomed to a given query paradigm and the difficulty lies in making the paradigm change [9]. However, this issue requires further investigations. Subjects found it easy to use the tag-map paradigm in general, but it was deemed too complicated for simple queries and more useful for prolonged search and exploration since it allows users to dynamically adjust queries and therefore eliminates or reduces the need for re-querying.

The experience and knowledge gathered with the design and evaluation of these two prototypes would be used for developing future prototypes that intent to delegate more intuitively a dynamic control over the information settings of information retrieval interfaces to their users.

5. CONCLUSIONS

During the course of this paper we have explored ways to tightly integrate IR and HCI in a variety of human-information interaction scenarios where interfaces act as information retrieval systems. In particular, we studied how R and P as design parameters can describe the information settings of these interfaces. Both aspects were parameterized on a 0-3 scale on the basis of conducted experiments to analyze different possible information settings. Consequently, five situated interfaces were designed and analyzed to discern how R and P are qualified during requirement analysis, and how together they describe the information settings of situated interfaces, and therefore help reflect the interaction requirements onto the interface design.

Finally, we investigated the feasibility and utility of delegating control of R and P dynamically to users during classic search and retrieval scenarios, and concluded that while this approach is clearly advantageous for exploration tasks and tasks that require re-querying, a more profound study should be conducted for further analysis. Such endeavor will constitute the essence of our future work.

6. DISCUSSION

The approach that we presented in this paper demonstrates that a tighter integration of HCI and IR is possible, by exploring the potential of R and P as design parameters for the information settings of situated interfaces. The use of these two performance metrics as design parameters may be seen as controversial, however, it is justified given that efficiency and information relevance no longer constitute the backbone of user expectations in all cases of human-information interaction. Instead, new aspects of human-information interaction (e.g. emotional, cognitive, experiential, situational, and cultural) are affecting the manner by which we conceptualize information systems. Our approach does not comprehensively address all these aspects, and therefore can be complemented by introducing new parameters to reflect with a higher affinity the aspects of human-information interaction onto the system design.

7. ACKNOWLEDGEMENTS

The authors would like to thank Oriol Galimany and other members of the Interactive Technology Group at Universitat Pompeu Fabra for their support.

8. REFERENCES

[1] Vogel, D. and Balakrishnan, R. 2004. Interactive public ambient displays: transitioning from implicit to explicit, public to personal, interaction with multiple users. Proceedings of UIST '04, pp. 137- 146.

[2] NJ Belkin. Some (what) grand challenges for information retrieval. ACM SIGIR Forum, 2008

[3] R.A. Baeza-Yates and B. Ribeiro-Neto. 1999. *Modern Information Retrieval*. Addison-Wesley Longman Publishing Co., Inc., Boston, MA, USA.

[4] Miller G. The Magical Number Seven, Plus or Minus Two: Some Limits on Our Capacity for Processing Information. The Psychological Review, 1956.

[5] L. Hallnäs and J. Redström. 2001. Slow Technology, Designing for Reflection. *Personal Ubiquitous Comput.* 5, 3 (January 2001), 201-212.

[6] Koffa, K. (1935): *Principles of Gestalt Psychology*. London, Routledge & Kegan Paul Ltd.

[7] Moghnieh, A., & Blat, J. (2009). A basic framework for integrating social and collaborative applications into learning environments. Proceedings of m-ICTE'09 Vol. 2 (pp. 1057-1061), 2009.

[8] Moghnieh, A., Sayago, S., Arroyo, E., Sopi, G., and Blat, J. Parameterized User-Centered Design for Interacting with Multimedia Repositories. In Proc. MMEDIA '09, IEEE.

[9] B. Buxton. 2007. Sketching User Experiences: Getting the Design Right and the Right Design. Morgan Kaufmann Publishers Inc. CA, USA.

[10] S. Bødker. 2006. When second wave HCI meets third wave challenges. In Proceedings of NordiCHI '06.