PJIM PARSONS JOURNAL FOR INFORMATION MAPPING

THE PARSONS INSTITUTE FOR INFORMATION MAPPING
68 Fifth Avenue New York, NY 10011
212 229 6825 piim.newschool.edu

Detecting End-User's Visual Model to Build a Visualization Tool Based on Online Reviews

ELIZABETH S. CARVALHO
MARCIRIO S. CHAVES

KEYWORDS *Data visualization, information visualization, visualization evaluation, online hotel reviews*

ABSTRACT To achieve a good level in information delivery it is important to know the end-user's visual model. On the other hand, it is also relevant to identify and characterize data accordingly to map it properly in visual terms. Although several guidelines are used in the Information Visualization field, there is still a lack of a solid and complete set able to give full support in the whole process. In this context, it is necessary to pre-evaluate, as much as possible, all the assumptions that are considered for its design and development. We present an exploratory study (n = 123) to detect the graphical preferences of travellers using accommodation portals of Web 2.0 (e.g., TripAdvisor.com). We took into account some of the most relevant ground rules applied in the field to map data visually and to design end-user interaction. Moreover, the evaluation process was completely oriented to data visualization. This paper introduces the main findings identified in terms of visual attributes and their combination. Although these findings were centered on the preferences of online hotel reviewers, we might extrapolate it to other domains. The goal of this work is twofold: To have a better knowledge of the online hotel reviewers' visual and cognitive model and to have a deeper and wider understanding of the model itself, detached from its specific domain. Finally, we also present a brief description of the first version of our prototype that was developed based on the findings of this work.

INTRODUCTION

How should we map and code data visually is always a hard question to answer. The visual mapping of data may be aesthetically beautiful but ineffective. It can also follow all the main guidelines and become aesthetically repulsive. To pursue a good balance between both aesthetics and utility is one of the golden rules, although it is not that simple. Information visualization techniques offer a more varied palette of characteristics than the typical interface. Characteristics such as color, brightness, contrast, surface area, line weight, and visual pattern, among others, can be used to represent differences between pieces of data.

Data visualization applied to online reviews has gained more attention in recent years (Gamon et al[1]; Draper and Riesenfeld[2]; Wu et al[3]). Currently, people are increasingly using websites to find out information about potentially anything. Almost all respondents of a survey performed by Gretzel et al[4] use the Internet as an information source for planning pleasure trips and most (82.5 percent), use it every time they plan a pleasure trip.

Most hotel portal websites display reviews generated by the reviewers, but this data is also generally poorly treated, analyzed, and offered with low level of interaction or user's specific needs before being made available in the website. Soukup and Davidson[5] describe several examples of how the use of visual data mining can help to expand business and improve customer satisfaction.

Although the use of advanced graphic techniques in online reviews is not recent[1], preliminary advanced graphical representation of online hotel reviews was first proposed by Carvalho and Chaves[6]. We present in this paper the development of this earlier study[6]. It provides contributions to the following topics:

- To understand the work, analysis, and information processing practices associated with the consultation of hotel booking information
- To elicit formal requirements for design
- To know what target users think of the different visualization techniques and what features are seen as useful
- To present the first version of a prototype developed taking into account the results from previous topics

The answers of the survey are further analyzed to identify and correlate possible patterns in the profiles and visual and interaction attributes of users. The relevant goal to be achieved is the identification of the visual model of the potential end-user that should be used to implement the first version of our prototype.

DATA VISUALIZATION IN ONLINE REVIEWS

Wu et al[3] designed and deployed OpinionSeer to address

the need to communicate opinion-mining results effectively and to facilitate the analytical reasoning process. The system was built on a new visualization-centric opinion mining technique that considers uncertainty for faithfully modeling and analyzing customer opinions. A new visual representation was developed to convey customer opinions by augmenting well-established scatterplots and radial visualization. Draper and Riesenfeld[2] presented a simple interactive visualization that allows users to construct queries on large tabular datasets from surveys and opinion polls and to view the results in real time. The results of two separate user studies suggested that their interface lowers the learning curve for naive users, while still providing enough analytical power to discover interesting correlations in the data.

Gamon et al[7] developed a system named Pulse, which contains a user interface based on the Tree Map technique to represent features from sentences in online reviews. The Pulse prototype system mines topics and sentiment orientation jointly from free text customer feedback.

Gamon et al[1] tracked thousands of blogs and the news articles that they cite, filing news articles that have highly overlapping content under the same icon (for inspection purposes). They then tagged each article with the number of blogs citing it, the political orientation of those blogs, and the level of emotional charge expressed in the blog postings that linked to the news article. They summarized and presented the results to the user via a novel form of visualization. The BLEWS system is a first foray into using a combination of text processing and link analysis to provide contextual information for news articles.

VISUALIZATION DESIGN & EVALUATION METHODOLOGY

Lam et al[8] proposed seven scenarios for evaluating visualization that were derived from an extensive literature review of over 800 publications. In the scope of our work, we considered that two of these scenarios should be used to guide the process of visualization design and evaluation: Evaluating Environments and Work Practices (EWP) and Evaluating User Experience (EU).

EWP helps to elicit the formal requirements for design. The goal of information visualization evaluations in this category is to seek an understanding of the work, analyses, and information processing practices of a given group of people with or without software in use. The outputs of studies in this category are often design implications based on a more holistic understanding of current workflows and work practices, the conditions of the working environment itself and of the potential of current tools in use.

In terms of design, visualization design requires taxonomy of techniques to guide automated algorithms, programmers, or users. It is an area still under debate and evolution. Most visualization taxonomies are based on the type of data involved[9]. Tweedie[10] describes several forms of data that may be represented visually: data values (described above), data structure, (e.g., file hierarchies or rectilinear vs. curvilinear grids) and metadata.

"Information visualization" can be similarly organized by data type[11]. Common categories are multi-dimensional databases (often containing more than three dimensions), text, graphs, and trees[12]. In addition to data type, some taxonomies[13] have organized visualization systems by display style (e.g., table or information landscape) or include generic tasks performed by users of the system (e.g., gaining an overview, drilling down on details or filtering).

Gretzel et al[4] made a survey administered during a four week period between January 5 and January 31, 2007. The respondents were from a TripAdvisor.com traveller panel. The questions were related to general travel experience/planning style, perceptions, use and impact of online travel reviews, motivations and barriers to posting online reviews, and it included demographic information.

Based on the conclusions of their study, it became clear that measurements such as: the experience of the reviewer, the similarities of the activities engaged in during a trip, and the purpose or the clarity of the comment made, need to be taken into account highly when considering the opinion of the reviewer. The age, gender, marital status, and the date the review was posted are dimensions that should be made available, although they do not have the same importance.

RESEARCH METHODOLOGY

This study has its background in user-generated data specifically related to the hotel sector. Online reviews collected from guests of small and medium hotels were used and cross-referenced to find patterns that could give a better insight to and support for effective decision-making. In order to classify and analyze the data, a methodology based on concepts of a domain-specific ontology—Hontology, which is described in detail in Chaves et al[14]—was applied.

A 25-item questionnaire was developed and pre-tested on a sample of 29 people to ensure its clarity and feasibility. It was divided into four main sections. Each was focused on an aspect considered important to identify the end-user visual model: visual attributes, interaction

attributes, dimensions and data measurements, and the end-user profile.

We plotted the data using standard graphics as well as mapping with more advanced visual techniques. These techniques were chosen taking into account the existing research work for this domain in the field. Moreover, we tried to propose something visually unusual and new to the end-users. The data was clustered and existing relationships identified. As a result, (see FIGURE 1) a rectangular tree-map (A), a radial graph (B), and a force-directed graph (C) mapping the relationship between the traveller profile, the ontology, and the polarity strength (e.g., very positive and very negative) were generated. A total of 24 concepts of ontology were available to be clustered and visualized. To guarantee the readability of the snapshots, we filtered data to produce the graphics, so that the number of items to be visualized was reduced. Thanks to that, the images of the graphics in FIGURE 1 present different number of concepts of an ontology in each case, although the graphic was produced considering all the 24 concepts.

We also wanted to evaluate the acceptance of the end-users of unusual visual representations, and included a stacked bar (D) and a pie (E) graphics. Although these visual representations are more limited in the number of dimensions they are able to plot simultaneously, we took this into account and mapped the same number of variables. In both cases, we adopted a red-blue color scale, using the red to indicate the negativity of the attribute while the blue its positivity. After a brief graphic orientation, respondents faced five graphics to evaluate, as depicted in FIGURE 1.

The evaluation of these graphs required the subjects to choose and, thus, to indicate which visual variables were more relevant to them. Each of these graphs offers a slightly different cognitive perception to the end-user. The questions aimed at them tried to take advantage of this and identify the most suitable visual attributes.

Although the number of questions in the interaction attributes section was less than the visual one, we tried to concentrate our questions in order to find out if data exploration tasks based on direct interaction or changing visual cues would be an asset to the end-user. The same shortcut was used in terms of identifying the most desirable dimensions and measurements of the data to be visually represented and the end-user profile. Finally, participants received an e-mail containing a link to the questionnaire web site. The survey was administered during a six-week period between April 2 and May 15, 2012.

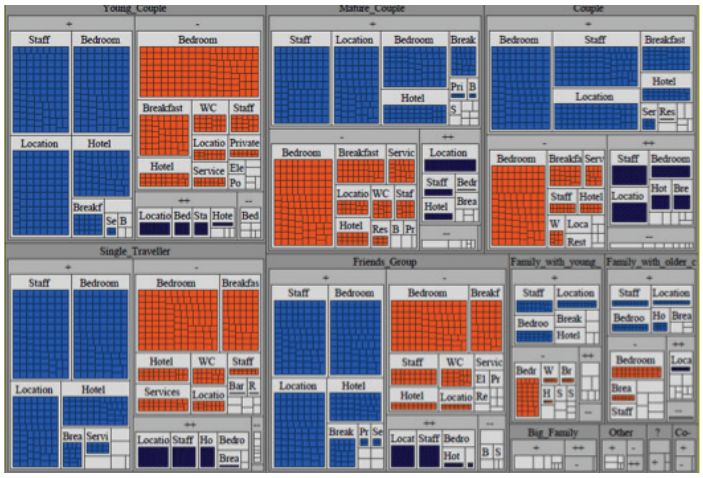
DATA ANALYSIS

The Web-based survey effort led to 123 usable responses, which represented a 21 percent response rate. Mean age was 38 (SD = 9 years) with the largest portion (47.8 percent) falling between the ages of 31 to 40. Most of the respondents (84 percent) had at least a college degree and 67 percent were men. 83 percent of the respondents usually read the reviews written in booking websites, while 77 percent of the respondents considered that reading the reviews was decisive or crucial for booking an accommodation. Regarding nationality, 57 percent of the respondents lived most of their lives in Portugal and 38.5 percent in Brazil. For a complete profile of the sample characteristics, see the appendix

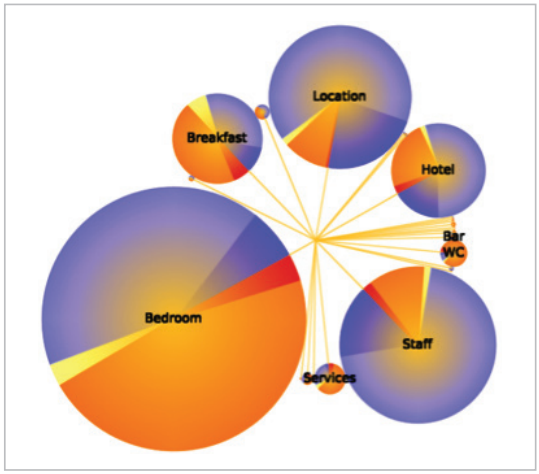
VISUAL ATTRIBUTES

Nine questions focused on identifying what visual attributes should be present in the redesigned version of the accommodation sector's data visualization. The goals and the frequency of the most popular graphic for questions 1 to 8 are summarized in TABLE 1. The G1 up to G5 graphs correspond to snapshots (A) up to (E) in FIGURE 1.

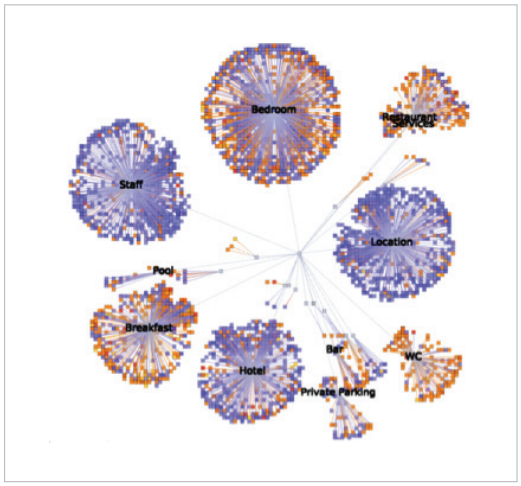
QUESTION NUMBER	QUESTION GOAL "IDENTIFY WHICH GRAPHIC...":	NUMBER & FREQUENCY OF THE MOST VOTED GRAPHIC
Q1	Is easier to compare values	G4 (44%)
Q2	Is easier to locate information without losing the general context	G2 (40%)
Q3	Is easier to obtain a summary overview of the information	G4 (33%)
Q4	Has color as the most relevant visual attribute to map information	G2 (29%)
Q5	Has shape as the most relevant visual attribute to map information	G2 (36%)
Q6	Has position (spatial location) as the most relevant visual attribute to map information	G2 (42%)
Q7	Is easier to quantify the number of comments about a certain aspect	G4 (41%)
Q8	Offers the better insight from the comments about the different aspects of the accommodation	G4 (34%)



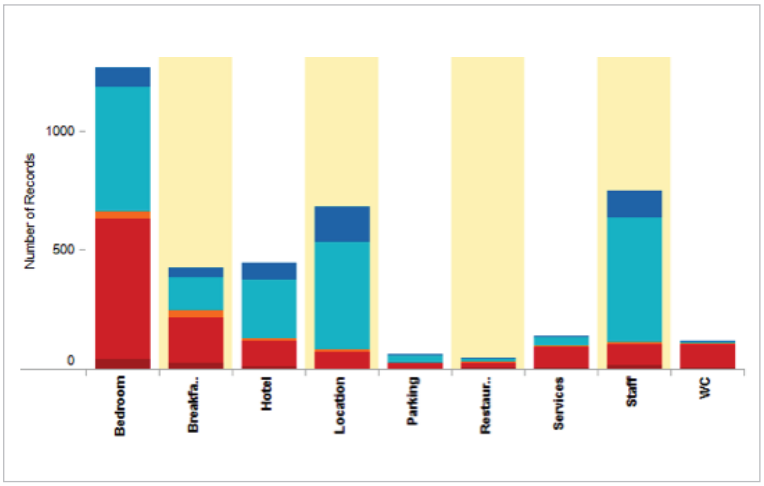
(A)



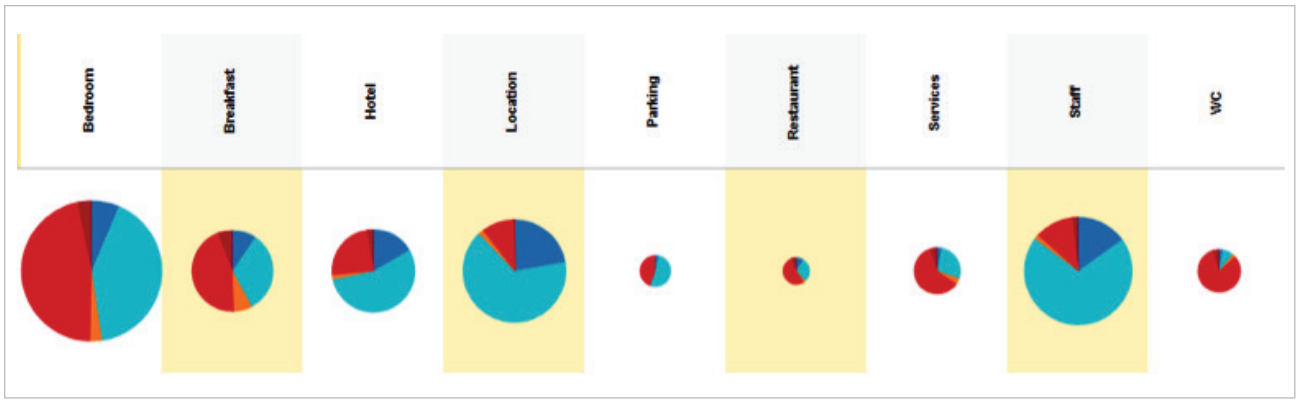
(B)



(C)



(D)



(E)

FIGURE 1: *The five graphics used in the questionnaire – (A) tree-map, (B) radial graph, (C) force-directed, (D) stacked bar and (E) pie.*

The most popular graphics were two and four. Graphic two was a radial graph, which is specially used when the hierarchy or existing dependency in data needs to be emphasized. In terms of relevant perceptual codification, this graph explores mostly position, size, proximity, connectivity, shape (circular), and color as visual attributes for mapping data.

Graphic four was a regular stacked bar graph, which is normally used when we need to emphasize quantitative differences between data. It cannot map as much information as a radial graph, although it offers a simpler visualization of data. Moreover, it has been around for a long time, which makes it readable for almost everyone. Its significant perceptual codifications are position, size, color, and shape (rectangular).

According to TABLE 1, 50 percent of the questions pointed out graph two as being the best to locate information without losing the general context or as having used the shape visual attribute better, for instance. Meanwhile, 50 percent of the questions pointed out graph four as being more suitable in other aspects. Another result was that women mostly voted for graph four, while men mostly voted for graph two.

The other question that aimed at identifying visual attributes was specially centered at the color issue. Considering that color is one of the most important visual cues, we thought that allowing the subject to identify which color palette should be used to map data from online reviews from the accommodation sector should be included in the survey. In this question, nine colors were made available (including the primary ones) to choose and order, as well as the option of adding a comment suggesting other ones. The answers to these questions were analyzed according to two perspectives: what is the preferred order of colors and what combination is most suitable.

FIGURE 2 illustrates the results achieved. The blue and green were the most popular for the first position (64 percent and 40 percent, respectively), while yellow was the most preferred for the second and third positions (37 percent and 29 percent, respectively). In terms of color combination, we did not achieve a clear result because most of the subjects chose all of the colors (15 percent), except the grey one. Following this result, the most significant result (9.2 percent) indicated that a combination between green, red, yellow, blue, and orange should be used to map the information conveniently.

FIGURE 3 shows another perspective achieved with the color issue. The green and red were the most popular among subjects aged between 26 and 45. The blue and yellow appear slightly more significant than green only for the group aged between 31 and 35. Subjects with ages between 51 and 63 tend to choose yellow or blue as preferred colors.

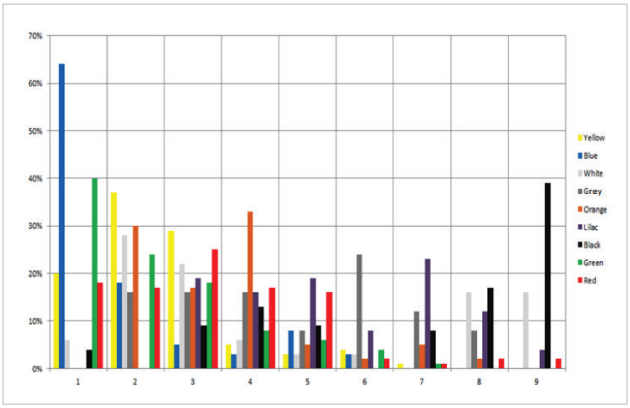


FIGURE 2: *Color palette results sorted by position (1–9) in which they appear in user preference.*

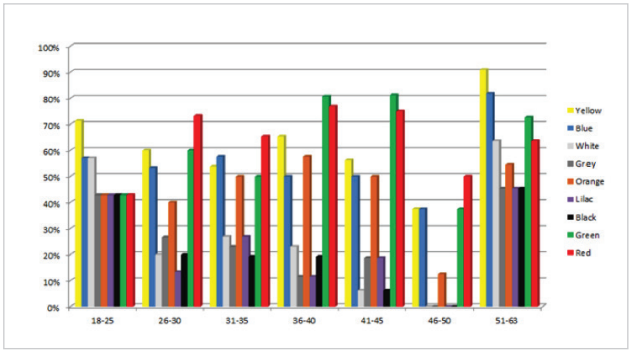


FIGURE 3: *Color palette results according to age clusters.*

INTERACTION ATTRIBUTES

These attributes are essential, because they give support to the data exploration. In our case, it is applicable because our data can achieve high density and thus become voluminous. The accommodation sector holds thousands of hotels, which have and generate a large quantity of possible qualifications. Because of the large amount of information to be visualized, suitable interactivity to selectively visualize and gradually expose data according to the user's particular interest at the time of use will help to reduce potential information overload on a static visualization scheme. Hence, the data visualization must be properly supported by different interaction techniques.

In our survey, we included a multiple-choice question with a list of eight interaction techniques and attributes. Some of the options were actually visual attributes being used to support interaction on data and, thus, worked

as interaction ones—transparency, color and change of brightness. The other options were straightforward interaction techniques—dragging, zooming, scale change, point and select, and animation. This question also included a comment field where the subjects would add other interaction attributes or leave a comment. Although most of the interaction techniques are largely used and available in most software, to avoid any misunderstanding of their meaning (based on the techniques names), we inserted a small movie illustrating their use on two of the graphs.

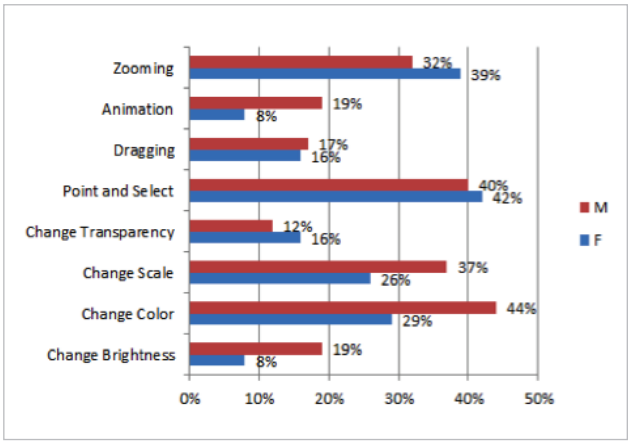


FIGURE 4: *Interaction attributes by gender.*

FIGURE 4 illustrates the results obtained individually for the male and female groups. Women and men indicated that the point and select technique must be included, while men considered the color change very useful to support data interaction. Clearly, women did not consider change of brightness and the use of animation helpful. On the other hand, men seem not to value transparency. FIGURE 5 shows individual values obtained for each attribute.

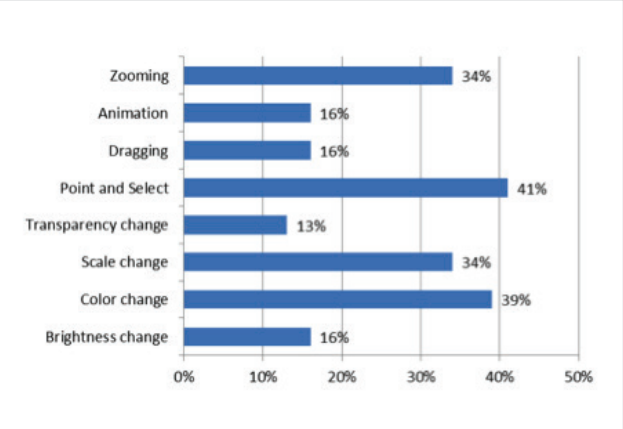


FIGURE 5: *Totals of interaction attributes.*

DIMENSIONS AND MEASURES

The data considered are composed and defined by several variables. Some of them are, for example, the name of the hotel, its number of stars and rating, the facilities that it offers, the concept of ontology and its strength, or the traveller travelling condition (if alone or in a group, with or without family, etc.) and his/her origin. These variables define potential dimensions to be used to create the visual representations of the data. Most of these dimensions are likely to be measured somehow or even to be classified as measures themselves. For instance, the average of the rating or the maximum and minimum strengths achieved by a concept of ontology would be interesting to make available to the end-user in the visual representations as measurements. The number of stars can be considered both as a dimension (ordinal) and a measurement itself.

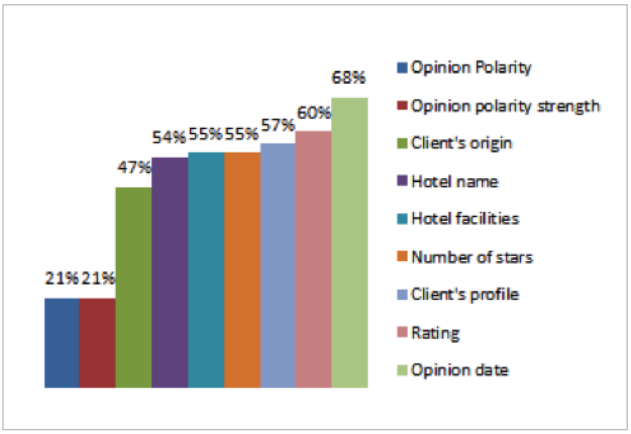


FIGURE 6: *Distribution of frequencies by preferred graphic dimensions.*

As mentioned in the Introduction, one of our concerns in the survey was to find out as much as possible, which dimensions and measures should be included in the visual mapping of the data. To evaluate this, we included a multiple choice question, where the user had to choose which dimensions would be offered to him/her and also considered mostly relevant for supporting his/her decision. As before, a field to receive the comments and suggestions of the subjects about other dimensions to be considered was also included. FIGURES 6 and 7 illustrate some of the results achieved.

FIGURE 6 indicates that the opinion date and rating were the most important dimensions to be included in the visualization. In addition, the profile of the traveller, the number of stars, and the hotel facilities were also considered relevant. FIGURE 7 shows how these dimensions were chosen according the age range of the subjects.

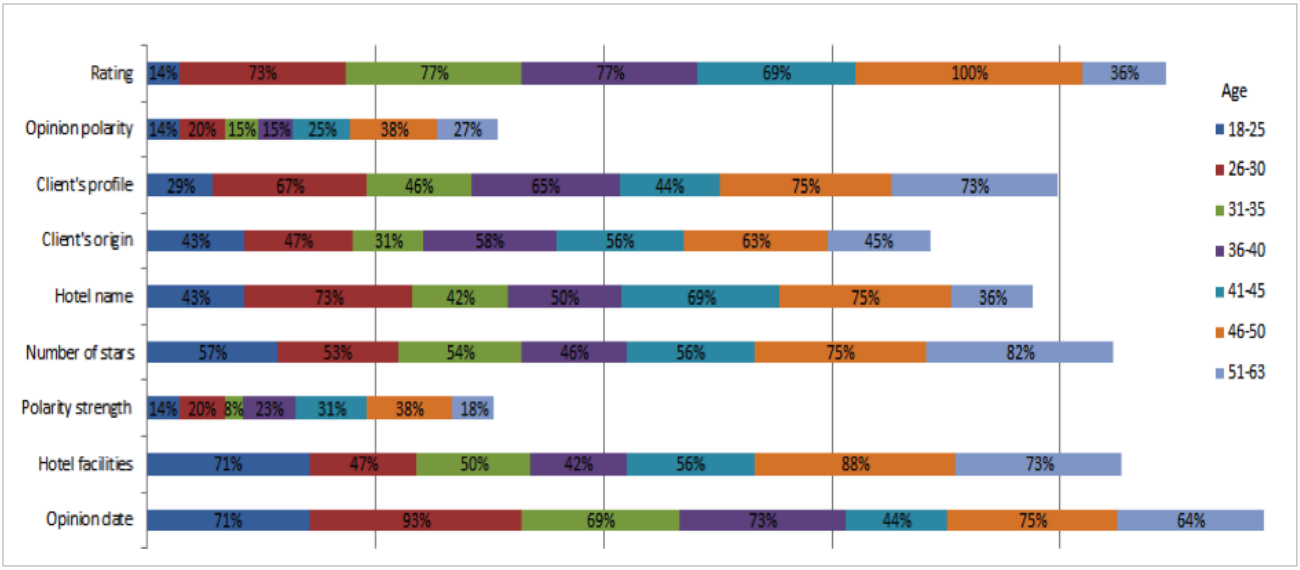


FIGURE 7: *Frequency of choice of dimensions versus age range of subjects.*

Those between the ages of 46 and 50 preferred the rating dimension (or measure) (100 percent), followed by those with ages between 31 and 35 and 36 and 40 (each with 77 percent). The senior subjects with ages between 51 and 63 mostly cited the number of stars, while the subjects with ages between 36 and 40 gave lesser relevance to this dimension.

To evaluate the measures to be potentially offered, we included another multiple-choice question relating all the dimensions that were listed before. Because we intended to make the kinds of measures clearly understood by every subject independently of his literacy, we included only the more standard ones, such as the average, the minimum and maximum values, and the standard deviation. As usual, we also allowed the subjects to leave comments and suggestions about the measures to be included in the visualizations.

The average (AVG) was the most highly scored measure for all the dimensions followed by the total (Count), maximum (Max), and minimum values (Min). For instance, for the rating dimension, the average followed by the maximum value and the total (the counting) were the most chosen (with 33, 16 and 18 votes, respectively), as independent measures to be made available, although some of the subjects (6) indicated that both the standard deviation (SD) and the average should be made available in the visualization of this dimension. A smaller number of subjects opted for the possibility of the data being presented with more than one measure. For almost all of the dimensions just one subject tended to indicate a more complex range of measures.

ANALYSIS OF THE RESULTS

Almost two thirds of the respondents always or frequently use Web 2.0 portals for booking hotels, while 84 percent of the respondents usually read the reviews. Moreover, most respondents (77 percent) consider the content of the online reviews decisive, when booking a hotel. In the questions one up to eight, most women preferred Graphic 4, while men opted mostly for Graphic 2. These questions were directly related with the identification of relevant visual attributes and, thus, visual tendencies. Huang and Pashler[14] have presented a more recent model of low-level vision, while Huang et al[16] proposed that the visual system is capable of dividing a scene into exactly two parts: selected elements and excluded elements. This is the "Boolean map" that underlies their theory. Based on this, we may conclude that women and men possibly create slightly different visual Boolean maps of the data, thus giving that result.

Blue and green are the most dominant colors. Green is closely related to Portuguese culture, and blue is an omnipresent presence, because of the colors of the ocean. According to Wright et al[17], individuals develop a language of color, as they mature, based on common and cultural usage. Due to this fact, existing symbolism and cultural use of color should be considered, when designing an interface. This is especially meaningful when looking at colors according to gender and age range. Red and green are especially expressive ones and both colors are also Portuguese national colors. Unfortunately, little theoretical[18] or empirical work has been conducted to date on

the influence of color on psychological functioning and the work that has been done is driven mostly by practical concerns not scientific rigor.

The strength of the polarity and opinion polarity were the less popular dimensions. We considered a possible cause of their lack of popularity to be a fairly straightforward fact: their meaning was not fully clear and comprehended. According to Hirst[19], writer-based and reader-based views of textual meaning are reflected by the respective questions "What is the author trying to tell me?" and "What does this text mean to me personally?" Contemporary computational linguistics, however, generally takes neither view.

Another interesting finding was that all subjects aged between 51 and 63 reported the rating dimension. This dimension is an ordinal one. Hotel ratings are often used to classify hotels according to their quality. We think that this possibly happened, because people in this age range are much more experienced and have already concluded that the rating system is trustworthy.

In terms of measures applied on the dimensions, most of the subjects tended to choose the average as the main one. In this case, maybe this happens because it is a kind of measure that is openly adopted and known. It is learned (or how to be calculated) in the first years of school. Moreover, it gives people a rough clue about the general data tendency. In short, it is simple and widely used.

Although we did not evaluate all the same parameters that Gretzel et al[4] did, some of ours may be directly correlated to theirs such as the profile of the client and the opinion date. These dimensions were also significantly indicated as being desirable for visual mapping.

One of the most surprising findings was that users preferred traditional ways of visualizing online reviews of the accommodation sector, such as those offered by TripAdvisor.com or Booking.com (plain text, basic bar graphic, and extremely summarized information), instead of more advanced graphical proposals. According to Hullman et al[20], the sense of a user about what is expected from them in using a visualization, including whether to use analytical or intuitive reasoning and the extent to which they should "let the graph do the work", can be a powerful influence on how they structure their interaction. Introducing novel visualizations to an end-user can sometimes induce unproductive levels of frustration, thus negating the influencing premise and making people tend to refuse primarily more advanced data visualization. The fact that data is not being presented in a summarized way can explain this result. On the other hand, the novel visualization was probably not preferred because the unfamiliarity dictates some limitations to the end-user. After good training and familiarization with the real benefits of such visualizations, their usefulness will be felt and voiced by people.

We also analyzed all of the comments received and we were able to detect some patterns and interesting suggestions including:

- Use of another color combination with high contrast and avoiding any kind of gradation
- Use of color according to international standards
- Horizontal legends instead of vertical ones
- Inclusion of 3D
- Inclusion of more labels with text
- Use of guiding lines in graphs
- Making available the age of the person, who gave an opinion and all of the comments
- Use of glyphs to map data (such as Chernoff[21] faces, for instance)

We consider that this discussion is relevant to both scholars and practitioners, because the graphical representation of online reviews is crucial for helping the decision-making process.

In terms of visual attributes and techniques, the users preferred stacked bars and radial graphs to review the data. Due to that, the most significant visual attributes that were identified were the position, the size, the shape, the color, the proximity, and the animation. The treemap technique and the containment visual attribute are not recommended. The pie chart was also refused possibly because it was difficult to compare the size of the slices and their angles. In terms of colors, green, red, yellow, blue, and orange were the most preferred, blue and green being the most dominant. Because red and green are recurrent predominant colors they should be used in the visual interface.

In the case of the interaction attributes, the users indicated that the possibility of pointing, selecting, and zooming the visualization in addition to the control of the color scale, would be very useful. Women did not find the use of the animation effect on the data or the control over the color brightness helpful, while the men did not value the control of the visual transparency (change the alpha channel).

Finally, we found that the date of the opinion is an essential dimension and its temporal visualization would be very helpful (thus, temporal visualization of data must be attended). The average, counting, maximum, and minimum values of the data are valuable and important measures to the end-users.

VISUALIZATION TOOL BASED ON ONLINE REVIEWS

ViTOR is a prototype primarily developed to allow testing of the visualization premises assessed by the survey described in this paper. In addition, it also aims to test potential new proposals of visual representation of its input data and evaluate it against online reviews from other domains (e.g., restaurants and books). The sources are generally provided from Web 2.0 websites (e.g., online portals, social networks and micro blogging). Although the main goal at this stage is to identify and validate the end-user data visualization, this prototype is actually the starting point to the implementation of a complete application. An important goal of the architecture of the application is to be fully abstract to receive multi-domain and multilingual UCG (User-Generated Content).

FIGURE 8 depicts the architecture of the application in its higher abstraction. It receives data from multi-domain and multilingual information sources that are used as input to the processing and analysis module. This module classifies the data according to ontology or by applying a group of predefined features, identifies clusters, relations and patterns, and visually delivers its results using a set of information visualization techniques. In order to guarantee an easy sharing and reuse of the knowledge achieved, the processed data (classified and treated) is stored at a knowledge base. The end-user interfaces are tailored to attend two different potential end-users: customers and managers.

We started the implementation of the ViTOR application by its second module, as it was fundamental in testing

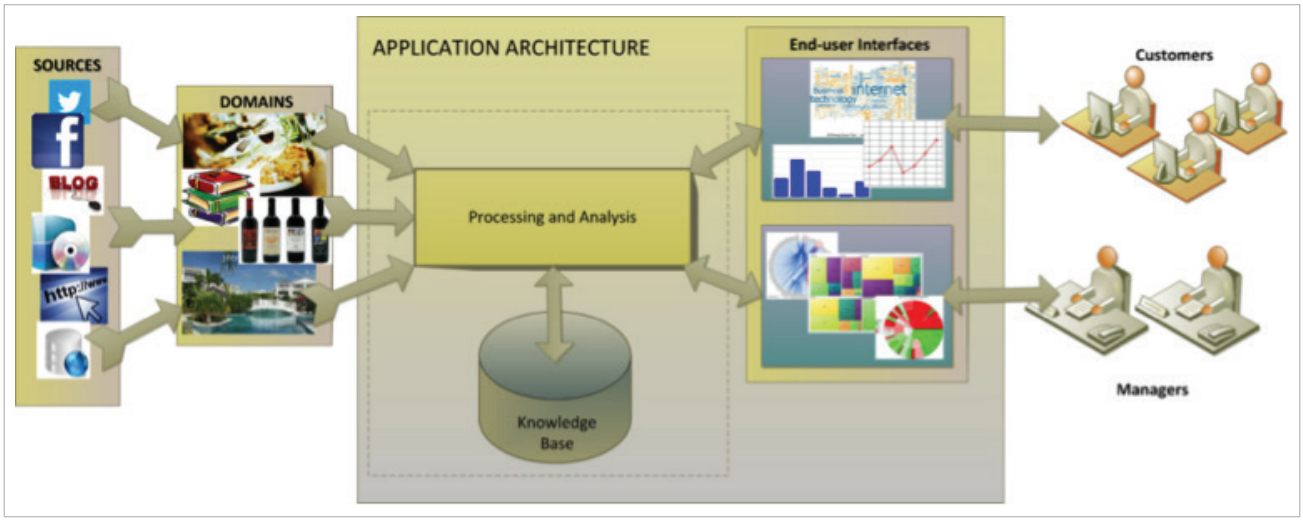


FIGURE 8: *ViTOR general architecture.*

FIGURE 9: *Classification of the concepts of ontology with color scale and lexical cloud (color scale indicates good or bad, while text size, the counting of votes)*

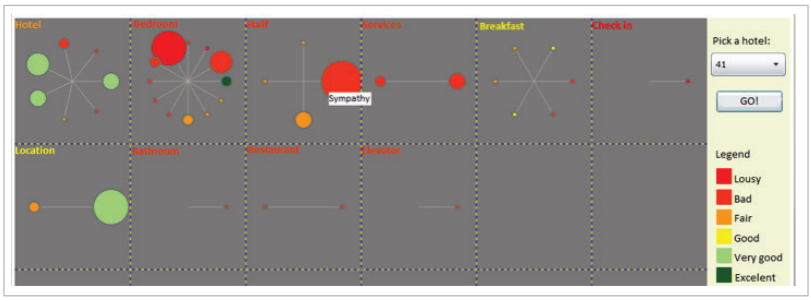


FIGURE 10: *Classification of the concepts of ontology with a radial graph*

our visualization and interaction premises. Our prototype implements both the Data Analyzer and Visualizer sub-modules. To implement it, we developed a fully web-based application, using the Microsoft Silverlight 5. Silverlight is a free plug-in, powered by the .NET framework and compatible with multiple browsers, devices and operating systems, bringing a good level of interactivity wherever the Web works. FIGURE 9 and 10 illustrate some of the visualizations that are offered.

FINAL REMARKS

The main contribution of this paper is to present the findings concerning the visual model and the cognitive response of users of hotel booking websites. These findings are useful for designing web interfaces according to the preferences of real users. The ViTOR prototype was used based on the findings delivered by the survey described here.

Even adopting current Information Visualization modeling methodologies and taking into account their best practices, our results must be taken only as guidelines, which do not completely guarantee success in future visual representations that are subsequently implemented.

Further research includes the evaluation of the prototype described in this paper. We also intend to use this prototype to run tests to reach the finest tuning between the end-user visual model and the data mapping and display offered.

BIOGRAPHY

Elizabeth Simão Carvalho is a researcher at CIAC—Research Center for Arts and Communication of the University of Algarve, Portugal and an Assistant Professor at University Aberta, Lisbon. Her research is concerned with the area of computer graphics with special focus on information and scientific visualization of data.

Marcírio Silveira Chaves is a researcher and professor at Universidade Nove de Julho – UNINOVE, Brazil. He works in the area of Project Management with a project titled Managing Information Technology 2.0. He has also published papers in the areas of Information Systems, Social and Semantic Web, Software Engineering, Natural Language Processing and Information Visualization.

NOTES

1 M. Gamon, S. Basu, D. Belenko, D Fisher, M. Hurst, and A. C König. "BLEWS: Using Blogs to Provide Context for News Articles". (Presented at AAAI Conference on Weblogs and Social Media 2008), 60–67.

2 G. Draper and R. Riesenfeld. "Who Votes for What? A Visual Query Language for Opinion Data". *IEEE Transactions on Visualization and Computer Graphics Vol 14*(6) (2008): 1197–1204.

3 Y. Wu, F. Wei, S. Liu, N. Au, W. Cui , H. Zhou, and H. Qu. "OpinionSeer: Interactive Visualization of Hotel Customer Feedback". *IEEE Transactions on Visualization and Computer Graphics Vol 16*(6) (2010): 1109-1118.

4 U. Gretzel, K. H. Yoo, and M. Purifoy. "Online Travel Review Study: Role & Impact of Online Travel Reviews". (Working paper, Laboratory for Intelligent Systems in Tourism Texas A&M University, 2007).

5 T. Soukup and I. Davidson. *Visual Data Mining: Techniques and Tools for Data Visualization and Mining.* (New York: John Wiley & Sons, 2002).

6 E. Carvalho and M. Chaves. "Exploring User-Generated Data Visualization in the Accommodation Sector". (Presented at the 16th International Conference Information Visualisation. Montpellier, France, 10-13 July, 2012), 198-203.

7 M.Gamon, A. Aue, S. Corston-Oliver, and E. Ringger. "Pulse: Mining Customer Opinions from Free Text". In: Proceedings of the International Symposium on Intelligent Data Analysis IDA, 2005, Nr. 3646, 121-132.

8 H. Lam, E. Bertini, P. Isenberg, C. Plaisant, and S. Carpendale. "Seven Guiding Scenarios for Information Visualization Evaluation". TR 2011-992-04, Department of Computer Science, University of Calgary, 2011.

9 C. Ware. *Information Visualization: Perception for Design* (Boston: Morgan Kaufmann 2012).

10 L. Tweedie. "Characterizing Interactive Externalizations", (Proc. Conf. Human Factors in Computing Systems, 1997), 375-382.

11 E.H Chi. "A Taxonomy of Visualization Techniques Using the Data State Reference Model". (Proc. IEEE Symposium on Information Visualization, 2000) 69-75.

12 R. Spencer. *Information Visualization: Design for Interaction* (Harlow, England: Prentice Hall. 2007).

13 S.K. Card, J.D. Mackinlay, and B. Shneiderman, *Readings in Information Visualization: Using Vision to Think.* (San Francisco: Morgan Kaufmann, 1999).

14 M. Chaves, L.A. Freitas, and R. Vieira "Hontology: a multilingual ontology for the accommodation sector in the tourism industry". (Proceedings of the 4th International Conference on Knowledge Engineering and Ontology Development, Barcelona, Spain, 4-7 October, 2012).

15 L. Huang and H. Pashler. "A Boolean Map Theory of Visual Attention". *Psychological Review* Vol 114(3) (2007), 599–63.

16 L. Huang, A. Triesman, and H. Pashler "Characterizing the limits of human visual awareness". *Science* 2008, 317, 823–825.

17 P. Wright, D. Mosser-Wooley, and B. Wooley. "Techniques & Tools for Using Color in Computer Interface Design". *Crossroads* -Vol 3(3) March 1997, 3-6.

18 A. J Elliot, A. C., Moller, R. Friedman, M.A. Maier, and J. Meinhardt. "Color and Psychological Functioning: The Effect of Red on Performance Attainment". *Journal of Experimental Psychology* 136, (2007), 154-168.

19 G. Hirst. "The Future of Text-Meaning in Computational Linguistics". Lecture Notes in *Computer Science*, Vol 5246 (2008): 3-11.

20 J. Hullman, E. Adar, and P. Shah "Benefiting Infovis with Visual Difficulties". (Proceedings of InfoVis'11, Providence, RI, October 23-28, 2011).

21 M.D. Lee, R.E. Reilly, M.E. Butavicius. "An empirical evaluation of Chernoff faces, star glyphs, and spatial visualizations for binary data". ACM International Conference Proceeding Series; Vol. 142, Proceedings of the Asia-Pacific symposium on Information Visualization - Volume 24, 2003.