2022 IEEE/WIC/ACM International Joint Conference on Web Intelligence and Intelligent Agent Technology (WI-IAT) | 978-1-6654-9402-1/22/$31.00 ©2022 IEEE | DOI: 10.1109/WI-IAT55865.2022.00045

What Really Matters in a Table? Insights from a User Study

Marco Cremaschi
Jessica Amianto Barbato
University of Milano - Bicocca
marco.cremaschi@unimib.it
jessica.amiantobarbato@unimib.it

Anisa Rula
University of Brescia
anisa.rula@unibs.it

Matteo Palmonari
University of Milano - Bicocca
matteo.palmonari@unimib.it

Rossana Actis-Grosso
University of Milano - Bicocca
rossana.actis@unimib.it

***Abstract*—Better understanding human visual attention during reading can provide valuable insights for developing user-centred computations models. A considerable amount of data, presented in a tabular form, is used in daily activities and is available on the Web nowadays. Several approaches have proposed an automated table summarisation method to improve the users' experience and give them succinct summaries of tables. However, there has been little attention to considering user behaviour in the design of automated table summarisation. In this paper, we present the findings of an empirical study, where we investigate, with the help of standard User Experience tools (eye-tracking technology and surveys), how users approach the reading of a table. We focus on evaluating how the domain knowledge and interest of the users influence their comprehension, eventually identifying four possible user-profiles and their different information needs. In order to show the impact of our findings on the selection of the information to keep in summary, we present and release a tool that, in addition to supporting the development of similar experiments, allows checking the information presented in summary in the form of Resource Description Framework (RDF) triples, by exploiting the semantic annotation of the table.**

***Index Terms*—Data Summarisation, Lexicalisation, Semantic Table Interpretation, Human Interpretation**

I. Introduction

Tables are everywhere and play a crucial role in creating, organising, and sharing information on the Web of Data. Tables contain high-value data, but they can be challenging to understand for humans and machines due to several factors, such as the ambiguity of the values contained therein and the lack of contextual information (e.g., metadata). Semantic Table Interpretation (STI) approaches provide explicit semantic annotations (e.g., identifying and annotating entities in cells, their types and the connections between entities), thus capturing knowledge from tables [4]. The annotations resulting from STI systems [9], [8] return a summary as a set of Resource Description Framework (RDF) triples which are difficult to be understood by people who are unfamiliar with the underlying technology. On the contrary, reading text is a much more accessible activity for humans. In the context of the Semantic Web, Natural Language Generation (NLG) is concerned with the implementation of textual interfaces that would make the information stored in the knowledge bases' triples more accessible to the potential users. Further development of NLG systems could be beneficial in a great range of application domains such as dialogue interfaces [14].

Different techniques have been proposed to compute table summaries that can be informative for the users [10], [18], [6], [11], [14]. To support table summarisation (e.g., as in [11]), a few of these works use triples and a background knowledge base to improve the process [14]. Further, these latter approaches may exploit the information hidden in the proposed patterns of the triples, e.g., *<Actors,spouseOf,Lawyer>*. All the above approaches lead to summaries which are often *not fit for purpose* [15]. Therefore, little attention has been dedicated so far to *users* and *their behaviour* when approaching some tables. To obtain table summaries that are understandable and suitable for them, a *user-centred* approach must be considered, which should include an empirical analysis of the characteristics of the users and their needs. When a *user* reads and analyses a table, cognitive processes are activated, which make users attribute a meaning (or semantics) to the table and retain some of its information. A better understanding of these cognitive processes may provide valuable insights into designing user-centred computational models that fully respond to the users' information needs and provide more effective data representations.

This paper presents an empirical user-centred study about how users behave when reading and analysing information contained in tables, especially those describing entities and their relationships. To collect insights on the users' cognitive processes and behaviours during table reading, we adopt user experience research methodologies such as eye-tracking technology and surveys. Based on the hypothesis that the approach of a user who reads and summarises the information in a table may be influenced by the degree of familiarity with and interest in its domain, we explicitly analyse visual attention patterns by controlling these variables. Our findings suggest that four main categories of users can be identified depending on their background knowledge and interests. Users in different categories show different visual attention patterns and answer differently to task-related questions. We extend an STI approach by including the findings of the cognition study in terms of rules which identify the most relevant elements of the tables for each profile, thus generating a profile-based summarisation. The triples are then verbalised

978-1-6654-9402-1/22/$31.00 ©2022 IEEE
DOI 10.1109/WI-IAT55865.2022.00045

through existing approaches [2]. Finally, the study has been conducted using a dedicated tool (integrated with the user experience techniques adopted in the study) which is released as open-source software.

The remainder of the paper is organised as follows: In Section II we compare and contrast our work with other contributions in the literature. Section III presents the design and methodology of the study. In Section IV we describe the experiments to investigate how people analyse a table. Section V describes the tool built to carry out the experiments and create summaries. In Section VI we discuss the main findings from the study. Finally, in Section VII, we outline future directions of research.

II. Background and related work

In the following we divide the Section into two areas: i) background on table understanding through cognitive tasks, and ii) literature on table summarisation.

Background on cognitive tasks. While elaborating the tables involves significant cognitive load, studies show that it is fairly easy to determine their domain, thanks to the homogeneous characterisation of the values within columns. Three cognitive processes are involved in the users' interaction with the table: (i) *reading and comprehension*, (ii) *searching for relevant information*, and (iii) *interpreting and verifying the information obtained* against the user's informative needs. In order to convert these processes into machine-executable ones, eye-tracking techniques and questionnaires, which are methodologies typical of the research domains of User Experience and Applied Linguistics, are leveraged. In the reading process, eye movement is not continuous and harmonious; the eyes make short and rapid movements, called *saccades*, during which no information is retained. They determine, for example, the reading direction (from left-to-right and from top-to-bottom or vice-versa) in linear texts, such as summaries or prosaic texts. Information processing occurs exclusively during *fixations* [16], when the eye observes a portion of text in between saccades over a time frame ranging from 220 to 280 milliseconds. Although evidence shows that the dynamic positioning of visual attention is influenced by the difficulty of lexical processing (e.g., the predictability of words given context), it is not entirely clear how saccadic programming and linguistic processing interact with each other. By combining information about eye movements, acquired by eye-tracking, and information about cognitive processing of the data being read, obtained via task-related questions in questionnaire form, a comprehensive pattern of how the user interacts with the table is derived.

Natural language generation and table summarisation. Approaches to NLG from structured data in current literature do not consider the users' needs and behaviour. This is particularly true in the case of table summarisation [1], [12], for which it is necessary to identify which data to include in the summary and how to generate it according to them. Structure-based methodologies commonly leverage automatically generated attribute-value taxonomies [6], single predefined schemas [7] or patterns [3]. Content-based techniques make use of value lattices [18] or table rows [13] but only partially account for the information conveyed by the table. Hybrid approaches generate a summary from a flattened table of key-value pairs (including serialisations of knowledge graph triples), representing as much information about the table's structure as the columns and their respective rows [11]. Finally, we mention two related but quite different problems in the semantic web community (with the main difference being in the input structure), namely, entity summarisation and verbalisation of RDF data. Our study contributes to shedding light on the content selection problem in table summarisation, not by introducing new computational methods but rather by analysing the behaviour that the user enacts when exploring the table. The only previous study we are aware of [21] that adopts a similar methodology focuses on statistical data (we consider tables with several numerical data but describing entity-related information), uses a previous generation technology, is carried out with a small group of users (10) and is aimed at investigating different phenomena (impact of format and numeracy). To the best of our knowledge, however, a similar study of user behaviour has not been considered in the related but different fields of entity summarisation and RDF verbalisation.

III. Study design and methodology

The study comprises three main experiments: in the first one, we investigate the user's visual attention and what is his/her initial understanding; in the second one, we investigate how users understand the tables concerning their background knowledge and interest; and in the third experiment, we investigate how users perceive the generated summary generated through our STI and verbalisation approach that consider the most relevant elements of the table according to the different users' profile.

Part 1. First experiment. The first experiment is an exploratory analysis which aims, through eye-tracking, to understand the user's cognitive processes through real cases and why specific mechanisms are chosen. This analysis allows us to study the user behaviour both during the reading and searching of data, and the execution of an interpretation process, such as creating a natural language summary of the information in the table. The first experiment is divided into two phases:

Part 1.1. Eye movement tracking uses an eye-tracking software and device. This phase of the experiment is divided into three steps: 1) *Calibration*: for a correct identification of the direction of the gaze, the tracking algorithm must learn the position in which the head is located (consequently, the eyes). This algorithm can associate an eye position with a point inside the screen. 2) *Table Viewing*: after the calibration process, a table that should be analysed and understood by the user is displayed. 3) *Heatmap*: at the end of the experiment, a map with *colour spots* is displayed, which uses a graded scale ranging from green to red. These spots visually show the number and length of fixations and the areas within the page that have been observed. The more the spots tend to red colour, the more the fixations have been prolonged on that area. The analysis of the colour spots suggests different ways to approach the tables.

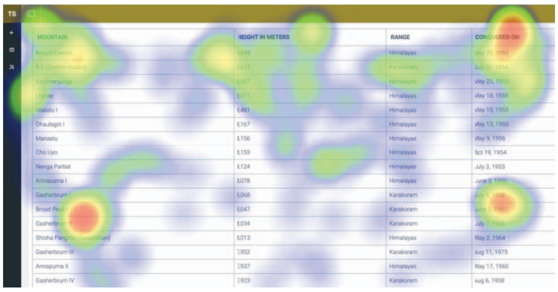

Fig. 1. Example of heatmap on the table containing mountains.

Part 1.2. Questionnaire is used to collect personal data and information on how to analyse and understand the table. In particular, the first experiment's questionnaire aims to learn how the user interacts with a table, which is his/her understanding, and which pieces of information are first remembered. The analysis of behaviours is intended to extrapolate the methodologies the user implements to define a mental model of the data and the psychological processes activated to learn more. The questionnaire, therefore, aims to: (i) verify the degree of interest of the participant w.r.t the domain of the table; (ii) gauge prior knowledge concerning the domain of the table; (iii) assess the understanding of the language by non-native English speakers; (iv) obtain from the participants a summary containing information about the structure and the content of each table presented in the experiment.

Hence, with the first questionnaire, we can obtain mainly two types of information: (i) how the user first approaches the table, asking, for example, which feature and/or information they most easily recall (e.g., *What is the first thing that you remember about the table?*); (ii) if the user is familiar with the domain of the table, asking, for example, to provide a rating to each column according to the relevance perceived by their understanding of the table content in order to justify his/her choices (e.g., *Why did you attribute these ratings to each column?*). Open-ended and evaluation questions using seven-point Likert scales are presented to the user.

Part 2. Second experiment. During the second experiment, the understanding of the data present in the table is analysed in detail and whether a given information is represented or described in the table. In particular, we investigate users' ability to understand whether or not a given sentence describes the data in a table. Initially, participants are presented with the table and given ample time to understand the structure and the data content. This experiment is carried out through a questionnaire divided into three parts.

Part 2.1. collection of the participant's data. Information on study degree, gender, age, and level of English proficiency is gathered to assess our sample's characteristics.

Part 2.2. the start and end time of the experiment. The start and end times of the experiment are collected for statistical purposes. To avoid discrepancies between native and non-native English speakers who may spend more time understanding the questions, both the response time to the questions and the exploration time are recorded.

Part 2.3. evaluation of the understanding of the data of different tables. The questionnaire used in the second experiment comprises three types of statements for which the users are asked to express agreement or disagreement: **q1** - statements related to information found directly within the table (e.g., *Videogames are the main topic of Table 1*); **q2** - statements about information that can be inferred from the contents of the table (e.g., *There have been collaborations between different developers*); **q3** - statements about information that can be obtained by performing operations with the data presented in the table (e.g., *The highest number of platforms on which any game has been distributed is 6*).

Response options vary from "strongly agree" to "strongly disagree" on a Likert scale. It should be noted that, at this stage, participants will be asked to remember the content of the tables they will be reading during the experiment. We prepared the statements in such a way as to be able to observe the users' operations performed in the table. Studies [17] show that in-depth cognitive processing of data can assist the user in memorising the data. In this work, it is observed that the user gives a correct answer to questions that require an extensive elaboration of the data (i) if he/she possesses the information in his/her cultural background, and therefore his/her fixation pattern does not show attention on the data in question or (ii) if he/she does not possess the information but, because he/she is interested, he/she performs an elaboration while exploring the table. In this second case, it is inferred that the interest prompted a cognitive elaboration of the content and that this allowed the user to remember not so much the individual data but the consequent inference that has been made. It is thus assumed that interest and background knowledge affect the memorisation of the table content with different levels of depth. Eye-tracking techniques, in fact, can only highlight a pattern of fixations from which it is not possible to infer, without further investigation, which cognitive processes are activated during the reading phase. The analytical depth of the facts represented in the table required from the experimental subject grows from group q1 to group q3. Similarly, for the subjects to show a high degree of agreement with the sentences of groups q2 and q3, they must remember more information in the table. We assume, then, that interest and domain knowledge also influence the extent and type of information remembered [19].

Part 3. Third experiment. The third experiment aims to test the hypothesis that the user will perceive the summary created as more adequate for the group of users who share his/her same level of reported interest and background knowledge of the table domain. The questionnaire tests whether, and to what extent, the summaries produced based on user profiles obtained from previous experiments are perceived as preferable by the user over those created for different user groups. Four tables, also common to the first and second experiments, are then presented, each with four alternative summaries, among which the user is asked to choose.

IV. Study implementation and results

In this Section, we analyse the results on how users with different background examine a table.

A. Configuration of the Experiments

Participants. The experimental subjects were recruited among people with normal or corrected-to-normal vision (i.e., with the support of glasses or lenses), non-native English speakers who use tables during their working or daily life and adopt a left-to-right (LTR) reading direction. For this reason, we cannot exclude entirely that the tables had been seen before, but topics had been chosen with a low probability of being explored with everyday web use. Two partially overlapping groups of 28 different participants (age range 20-40 y.o.) volunteered for the first two experiments, 24 for the third. They were heterogeneous by gender (19 males and 9 females for the first experiment, 12 females and 16 males for the second one, of which 16 males and 9 females overlapped between the first two experiments; 13 females and 11 males, none overlapping with the previous experiments) and by educational qualifications (from diploma to master's degree and doctorate). The only prerequisites were (i) the use of tables in their working or daily life and (ii) a basic understanding of the English language. The number of participants has been chosen based on previous literature on eye-tracking studies in reading [21], [19].

Datasets. The tables[1] used are extracted from the web and from the T2Dv2 Gold Standard dataset[2]. Tables were chosen to present data of the same type (e.g., date ranges, years, names) at different positions throughout the experiment and thus curb any possible bias due to participants' reading direction.

Set Up of the Experiments. To perform the experiments using eye-tracking, it was decided to use a software. The software chosen, GazeRecorder[3] monitors eye movements in real-time, using only a PC and the integrated webcam. GazeRecorder is one of the best software available online; in fact, an accuracy and comparison test showed that the software measurements are higher than the desired level of accuracy indicated by the literature, significantly when the head can move freely.

Some visualisation and usability changes have been implemented at GazeRecorder. The calibration step has been changed to make it more understandable and usable for the participant. Although the experiments were performed remotely[4], the setting of the experiments was ensured to be the same, requiring the users to use a desktop because the instruments used are not fully responsive. The calibration process aims to adapt the measurements in relation to the characteristics of the machine used. The factors that may have affected the legibility of the tables, such as brightness and contrast of the screen fall into the group of settings defined by the user to improve the usual exploration of their content, thus having a minimal impact on the performance. The map provides a visual representation of the user's eye movements. It is displayed only at the end of the visualisation process not to distract the user while analysing the table. The eye movements of each user are saved on a database so that different analyses can be carried out.

B. Analysis of the Results.

From the results of the experiments described in the previous Section, we can identify four approaches to reading and understanding a table that are closely related to two user traits: (i) background knowledge; (ii) interest in the table's domain.

From a qualitative inspection of the heat maps, we derive different patterns of ocular behaviour. By combining those patterns with questionnaires' answers related to prior knowledge and interest, we observe that: (i) those who do not possess prior knowledge of a given domain pay more attention to the first columns on the left, where the subject is usually placed; (ii) those who have previous knowledge adopt a different behaviour according to the number of columns: if the table has less than six columns, the user tends to observe all of them. On the contrary, if the number of columns is greater or equal to six, the user focuses only on the subject column and then moves on to the columns on the right of the table. Regarding participants' interest in the content of each table, those who are not interested in the table's domain mainly observe the header and, at most, the first eight rows; and those who are interested in the table's domain analyse all or most of its content. These observations lead to the identification of *four user profiles*: **Profile 1:** those who have a domain knowledge (background: 1) and have interest in the domain (interest: 1): when reading the table, the user adopts a different behaviour according to the number of columns while still observing the content in its entirety, driven by an interest in the topic; **Profile 2:** those who have a domain knowledge (background: 1) but have no interest in the domain (interest: 0): the disinterested user who has prior knowledge, observes what is necessary to contextualise the content of the table, hence its header and first few rows. The resulting analysis is superficial and will be supported by inferences from the domain knowledge; **Profile 3:** those who do not have a domain knowledge (background: 0) but have an interest (interest: 1): the user who does not know the domain of the table observes with more attention the columns on the left, among which is the one related to the subject, focusing however on the single data expressed in each row. It is a necessary behaviour to create the basic knowledge for the domain of reference and to learn as much as possible about it; **Profile 4:** those who do not have domain knowledge (background: 0) and have no interest (interest: 0): the absence of knowledge and interest determine a superficial behaviour, with a focus of attention only on the header and the first few

[1] https://bit.ly/supplementary-material

[2] http://webdatacommons.org/webtables/goldstandardV2.html

[3] https://gazerecorder.com/

[4] Because of the pandemic the experiment was performed remotely and the software-assisted calibration process was necessary to make researcher intervention unnecessary.

rows; the only column that is properly taken into consideration is the one related to the subject.

To support this qualitative impression we run a statistical analysis (described in Section IV-C) on both saccades amplitude and frequency ($p<0.001$).

The answers to the first questionnaire presented in Figure 2, regarding which columns in the table are most taken into account and considered important for understanding, show a different focus of attention depending on whether or not the user has interest and background knowledge, confirming the results of the eye-tracking study. Those who have background knowledge are prone to consider and observe more columns than those who are not familiar with the domain; on the contrary, those who are interested in the domain tend to scan the table vertically, reading the data in each cell and focusing, in particular, on the column of the subject.

Reflecting what emerged with the eye-tracking tests and reported above, the four approaches to reading, analysing, and understanding tables result in a different ability to correctly recall the information they contain. In particular, we highlight that:

- statements in group q1 obtain the highest number of correct responses across all user groups, while those in group q3, which require active tabular data processing, lead to worse performance overall. In cases where subjects are then asked to calculate a physical or a time distance or duration, they do not process the data with adequate depth;
- individuals who are driven by an interest in the table domain, regardless of their background knowledge, make the most errors when statements involve the last rows and the rightmost columns. Conceivably, in the absence of domain knowledge, the individual focuses on the subject's column (the second from the left in *Videogames*), and pays less attention to the last columns, thus making mistakes on the statements related to the ordering (i.e., the first column) and the last rows. On the contrary, when a previous knowledge of the domain is present, like the subject whose eye-tracking pattern is presented. There is no difficulty in agreeing with the correct statements on sorting, leading to the lowest number of errors among the groups (even if present on groups q2 and q3);
- individuals who are not motivated by interest are those who, of all people, pay the least attention to the information presented in the table. Especially in the case where a domain knowledge related to the domain is present, the individual will leave out the data in favour of what they already know about the topic; this leads to the highest number of incorrect responses in the entire experiment, especially on responses that cannot be inferred from general prior knowledge (e.g., the largest number of platforms on which a game in the table was distributed). More general information, such as sorting and table subject, is considered in the absence of background, while more detailed information is entirely left out.

C. Statistical Analysis

In order to verify if there is a significant difference between the four user categories described above, we run a chi-squared test χ^2 on the amplitude and frequency of saccades for each category of users. To perform the analysis on amplitude, we identify the values of the abscissa and ordinate of each saccade for each participant. In contrast, for the frequency, the number of frames corresponding to each saccade is calculated for each participant. Those values are then averaged for the four different categories of participants. The average number of frames is divided by 100 to get the number of frames per tenth of a second. A 4X3 contingency table is thus created for each table, where the four user categories are present in the columns, while the measures under evaluation are present in the three rows (Table I).

TABLE I
EXAMPLE OF THE STATISTICAL ANALYSIS FOR VIDEOGAMES DATA.

	background 1 interest 1	**background 1 interest 0**	**background 0 interest 1**	**background 0 interest 0**
x	27.5	28.9	21.7	22.0
y	23.3	29.4	19.8	27.5
#frame	29.2	27.7	20.0	23.1

The value of χ^2 is:

$$\chi^2 = 27.26, df = 6, p - value = 0.0001294 \qquad (1)$$

This means we can reject the null hypothesis with a very high probability (the p-value, which is < 0.001, denotes the probability that the null hypothesis is true). Therefore, the four categories identified by both prior knowledge and interest result in four different patterns of ocular behaviours, denoting different attentional and cognitive strategies when reading a table. Subsequently, the t-test for paired samples is conducted because the measurements are carried out on the same subjects. Low p-values are mainly obtained, so there is no reason to claim that the difference between the two users is insignificant. The data and R scripts used in these tests are publicly available[5].

V. A TOOL TO SUPPORT PROFILE-BASED TABLE SUMMARIES

The experiments described in the previous sections were carried out through a web application[6]. In addition to the functionality for viewing the table, the system for tracking eye movements and questionnaires has been integrated into this application. The analysis of the experiments allowed the extraction of the requirements, thus leading to the implementation of the system for the summary generation. It consists of four main steps: (i) in the first phase, the system allows the display of a list of tables, and it is possible to add other tables through an upload function. (ii) In the second step, viewing a specific table and loading the semantic annotations is possible. (iii) In the third step, the user indicates the degree

[5]https://bit.ly/supplementary-material
[6]https://table-summarisation.ml/

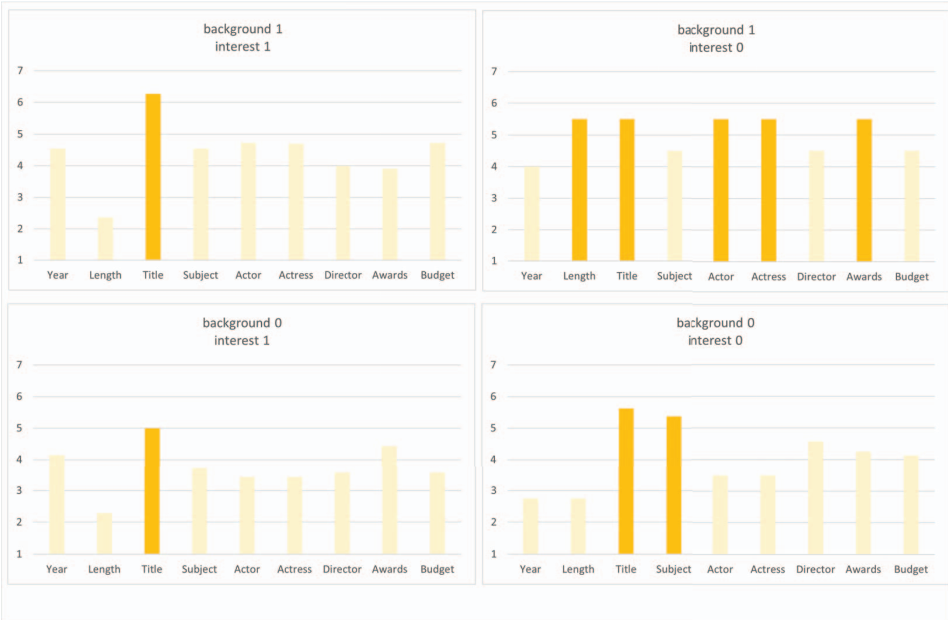

Fig. 2. Results from the *Film* table presented in the first experiment. The graph shows the mean of participant's results during the first experiment. Participants are distributed as follows: 11 users in the background 1/interest 1 (HH) group; 8 users in the background 0/interest 0 (LL) group; 7 users in the background 0/interest 1 group (LH); 2 users in the background 1/interest 0 (HL) group.

of familiarity with the knowledge and interest in the table's domain. According to the user's choice, the system highlights different rows (Table II) and columns based on the experiments described in the previous sections. (iv) The next step shows a set of triples generated concerning the selection.

TABLE II
HEURISTIC RULES FOR CONTENT SELECTION IN RELATION TO THE FOUR USER PROFILES.

background	interest	header	columns	rows
1	**1**	yes	col <6, all col >6, subject + last three	all
1	**0**	yes	col <6, all col >6, subject + last three	max 8
0	**1**	yes	subject + first two	all
0	**0**	yes	subject + first one	max 8

The tool is available through a Git repository[7]. It is developed with the Django framework. The tool has been encapsulated in a Docker container to facilitate the deployment and scalability by replication. Message management is performed using Task Queues (i.e., Celery Workers).

The results obtained with the third experiment, conducted by leveraging the summaries created by the tool, confirm the experimental hypothesis that users tend to find the summary created for a group of individuals with the same background knowledge and interest in the domain of interest more adequate. Table V shows the p-values resulting from applying Fisher's exact test for correlation evaluation. A correlation between summary choice and user category is detectable, although not always strong. From the significance values presented in the table, it is also possible to conclude that in the most polarised categories, with both parameters of knowledge and interest high or low (LL, HH), the user's choices do not strictly correspond to those hypothesized by the experimenters. This could be due to the propensity to provide biased self-assessments, a condition that, in the study of user behaviour, is difficult to avoid completely.

[7]https://bit.ly/mantistablex

VI. Discussion and Limitations

The eye-tracking data indicates how the saccades are programmed on specific elements of the table (i.e., the subject column and the header) considered informative by all four categories. Domain knowledge and low-interest users analyse these informative elements and a few elements of tables. Users with domain knowledge and high interest start from these elements and then analyse the whole table. Users with low domain knowledge and high interest analyse the entire table by making short saccades. Users without knowledge and interest analyse only informative elements.

In the current implementation, the tool allows the user to indicate their level of prior knowledge of the domain. However, we believe this method may lead to biased self-assessments. The user may be unable to assign themselves one or the other degree of knowledge, as the final summary is not produced as expected and/or needed. Based on the studies that we have conducted, we can therefore plan to implement an eye-tracking system within the tool itself. When utilised by the end-user, the tool can identify the pattern of observations and assign them to one of the four approaches identified in Section

TABLE III
THE RESULTING P-VALUES FOR EACH CATEGORY/PREDICTED ANSWER COUPLE ARE REPORTED

Mountains	LL	LH	HL	HH
O1	**0.047 (0.053)**	-	-	-
O2	-	**0.022 (0.022)**	-	-
O3	-	-	**0.004 (0.004)**	-
O4	-	-	-	**0.022 (0.022)**

Emperors	LL	LH	HL	HH
O1	**0.004 (0.007)**	-	-	-
O2	-	**0.001 (0.001)**	-	-
O3	-	-	**0.018 (0.018)**	-
O4	-	-	-	0.061 (0.061)

Organisations	LL	LH	HL	HH
O1	0.082 (0.082)	-	-	-
O2	-	**0.038 (0.038)**	-	-
O3	-	-	**0.007 (0.007)**	-
O4	-	-	-	**0.000 (0.000)**

Games	LL	LH	HL	HH
O1	**0.020 (0.020)**	-	-	-
O2	-	**0.002 (0.002)**	-	-
O3	-	-	**0.018 (0.018)**	-
O4	-	-	-	0.091 (0.091)

IV and consequently give them a personalised summary of the table to have the best user experience.

VII. Conclusions

We have presented a study that reveals how different users focus their attention on different aspects of a table they read based on their background and interest. In our study, eye-tracking has highlighted interesting visual attention patterns, which can inform guidelines to select the elements of a table that capture the user's interest and should be retained in a table summary. By complementing eye-tracking with questionnaires, we also found that the user profile, especially background knowledge and interests, influence the users' visual attention. The latter observation suggests that table summaries may be generated by considering these two variables. Finally, we presented an open-source tool to help understand profile-based table summaries more intuitively by using semantic technologies to express the table content. The data and tool used in our experiments are available for future research.

To the best of our knowledge, this is the first study that looks into the cognitive processes activated when approaching entity-related tables and focusing on their content using recent eye-tracking technology. In future work, we plan to move from proof-of-concept rules for profile-based summarisation, as used in our tool, to softer content selection mechanisms. Although we focused on the table summarisation task in this paper, we believe insights from visual attention patterns activated in reading a table can benefit other domains. For example, the advent of neural table interpretation models [5] paves the way to adapt cognitively-inspired saliency mechanisms applied to text comprehension and summarisation systems [19], [20] to systems devoted to the interpretation and summarisation of structured data, e.g., in the field of STI [9], [8], and entity summarisation.

References

[1] Bao, J., Tang, D., Duan, N., Yan, Z., Lv, Y., Zhou, M., Zhao, T.: Table-to-text: Describing table region with natural language. AAAI **32**(1) (2018)

[2] Castro Ferreira, T., Gardent, C., Ilinykh, N., van der Lee, C., Mille, S., Moussallem, D., Shimorina, A.: The 2020 bilingual, bi-directional WebNLG+ shared task: Overview and evaluation results (WebNLG+ 2020). In: Proceedings of the 3rd International Workshop on Natural Language Generation from the Semantic Web (WebNLG+). pp. 55–76. Association for Computational Linguistics, Dublin, Ireland (Virtual) (12 2020)

[3] Chen, J., Pan, J.Y., Faloutsos, C., Papadimitriou, S.: Tsum: Fast, principled table summarization. In: 7th Workshop on Data Mining for Online Advertising (2013)

[4] Cremaschi, M., Paoli, F.D., Rula, A., Spahiu, B.: A fully automated approach to a complete semantic table interpretation. FGCS (2020)

[5] Deng, X., Sun, H., Lees, A., Wu, Y., Yu, C.: Turl: table understanding through representation learning. VLDB (2020)

[6] Ienco, D., Pitarch, Y., Poncelet, P., Teisseire, M.: Knowledge-free table summarization. In: Data Warehousing and Knowledge Discovery. pp. 122–133. Springer (2013)

[7] Jain, P., Laha, A., Sankaranarayanan, K., Nema, P., Khapra, M.M., Shetty, S.: A mixed hierarchical attention based encoder-decoder approach for standard table summarization. In: Proceedings of the 2018 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 2 (Short Papers). pp. 622–627. Association for Computational Linguistics, New Orleans, Louisiana (Jun 2018)

[8] Jimenez-Ruiz, E., Hassanzadeh, O., Efthymiou, V., Chen, J., Srinivas, K., Cutrona, V.: Results of semtab 2020. CEUR Workshop Proceedings (2020)

[9] Jiménez-Ruiz, E., Hassanzadeh, O., Efthymiou, V., Chen, J., Srinivas, K.: Semtab 2019: Resources to benchmark tabular data to knowledge graph matching systems. In: The Semantic Web. pp. 514–530. Springer (2020)

[10] Kaffee, L., Vougiouklis, P., Simperl, E.: Using natural language generation to bootstrap missing wikipedia articles: A human-centric perspective. Semantic Web **13**(2), 163–194 (2022)

[11] Laha, A., Jain, P., Mishra, A., Sankaranarayanan, K.: Scalable micro-planned generation of discourse from structured data. Comput. Linguistics (2019)

[12] Lebret, R., Grangier, D., Auli, M.: Neural text generation from structured data with application to the biography domain. In: Proceedings of the 2016 Conference on Empirical Methods in Natural Language Processing. pp. 1203–1213. ACL (2016)

[13] Lo, M.L., Wu, K.L., Yu, P.S.: Tabsum: A flexible and dynamic table summarization approach. In: ICDCS. pp. 628–635 (2000)

[14] Mishra, A., Laha, A., Sankaranarayanan, K., Jain, P., Krishnan, S.: Storytelling from structured data and knowledge graphs : An NLG perspective. In: ACL Tutorial Abstracts. pp. 43–48. Association for Computational Linguistics (2019)

[15] Neumaier, S., Umbrich, J., Polleres, A.: Automated quality assessment of metadata across open data portals. J. Data and Information Quality **8**(1) (oct 2016)

[16] Rayner, K.: Eye movements in reading and information processing: 20 years of research. Psychological bulletin **124**(3), 372 (1998)

[17] Rouet, J.F., Vidal-Abarca, E., Erboul, A.B., Millogo, V.: Effects of information search tasks on the comprehension of instructional text. Discourse Processes **31**(2), 163–186 (2001)

[18] Selçuk Candan, K., Cao, H., Qi, Y., Sapino, M.L.: Alphasum: Size-constrained table summarization using value lattices. In: EDBT (2009)

[19] Sood, E., Tannert, S., Frassinelli, D., Bulling, A., Vu, N.T.: Interpreting attention models with human visual attention in machine reading comprehension. In: CoNLL. pp. 12–25 (2020)

[20] Sood, E., Tannert, S., Mueller, P., Bulling, A.: Improving natural language processing tasks with human gaze-guided neural attention. NeurIPS **33**, 6327–6341 (2020)

[21] Xu, A.: Eye tracking study on identifying and analyzing user behavior-eye movements, eye fixation duration and patterns-when processing numeric table data in paper or pdf format. Tech. rep., The University of North Carolina at Chapel Hill (2000)