Directions and Speeds of Mouse Movements on a Website and Reading Patterns

Ilan Kirsh
kirsh@mta.ac.il
The Academic College of Tel Aviv-Yaffo
Tel-Aviv, Israel

ABSTRACT

Mouse activity is known as an important indicator of user attention and interest on a web page. Many modern commercial web analytics services record and report mouse activity of users on websites. The position of the mouse cursor on the screen is the main source of information, as studies show a correlation between the cursor position during mouse activity and the user's eye gaze. This study focuses on mouse movement directions and speeds, and what they indicate, rather than on the mouse cursor position. Statistical analysis of mouse movements on a technical-educational website, which was selected for this study, sheds light on several interesting patterns. For example, most mouse movements in the examined usage data are either approximately horizontal or approximately vertical, horizontal mouse movements are more frequent than vertical mouse movements, and horizontal movements to the left and to the right are not equivalent in terms of moving time and speed. As this study shows, these statistical findings are related to the reading patterns and behaviors of web users. Associating mouse movements with text reading may potentially highlight content that most users tend to skip, and therefore, might not interest the website audience, and content that many readers read more than once or slowly, meaning it is possibly unclear. This could be useful in locating issues in textual content, in websites in general, and especially in online learning and educational technology applications.

CCS CONCEPTS

• **Information systems → Web mining**; **Traffic analysis**; *Browsers*;
• **Human-centered computing → Pointing devices**.

KEYWORDS

Web Usage Mining, Web Analytics, Web Pages, Websites, Mouse Movement, Human-Computer Interaction, Text Reading Patterns, Reading Behaviors, Online Learning, Education Technology

ACM Reference Format:
Ilan Kirsh. 2022. Directions and Speeds of Mouse Movements on a Website and Reading Patterns. In *Proceedings of the 10th International Conference on Web Intelligence, Mining and Semantics (WIMS'20). ACM, Biarritz, France.* https://doi.org/10.1145/3405962.3405982

1 INTRODUCTION

This study presents an analysis of mouse movements of users on a selected website. The analysis has two phases. In the first phase, statistical methods are used to explore directions and speeds of mouse movements on the website. In the second phase, examples of mouse movement activity of website visitors are examined.

The contributions of this paper include introducing a new approach for statistical analysis of mouse movements on a website, revealing interesting patterns of mouse movements, and explaining the statistical results (using examples of mouse movement activity of real users), as mainly associated with online reading behaviors.

This paper is organized as follows. Section 2 reviews related work. Section 3 introduces the website that was selected for this study and the data that were collected and used in this research. The statistical analysis phase is described in section 4, and examples of mouse movement activity of real users, which explain the statistical findings, are presented in section 5. Section 6 summarizes the results, discusses possible uses in web mining applications, and suggests possible further work.

2 RELATED WORK

Eye and gaze tracking technology can provide accurate information regarding which web page areas capture the user's attention, at any point in time [5]. This ability has been demonstrated in various applications, including in evaluating user attention to ads [15], identifying issues in a user interface [8], assessing web user interest on mobile devices [16], verifying that users pay attention when they authorize permissions [9], detecting user stress [23], and evaluating user enjoyment from online videos [18].

However, eye-tracking has its limitations. Although technically eye-tracking can be integrated into websites [5], this is usually impractical, because it requires user collaboration and may raise privacy concerns, as it makes use of cameras. In addition, accurate results require special equipment on the client-side. Therefore, many studies have examined analyzing user actions that can be tracked in modern browsers using JavaScript, such as page scrolling and mouse activity, as alternative implicit indicators of user attention [4, 12, 24].

Previous work has shown a correlation between eye gaze and mouse cursor positions on a screen [7], and the correlation is higher during mouse activity, i.e. when a user clicks or moves the mouse [3, 20]. Accordingly, mouse cursor position information was found to be useful in providing implicit relevance feedback on web search [6, 7, 19, 20], and in evaluating user attention in various other websites and applications, including in online surveys [2], web marketing [22], e-commerce [21], and task execution [17].

The cumulative user attention of all the visitors in different areas of a web page can be visualized by heatmaps [13, 14, 25] as follows. The web page is colored with several background colors. Hot background colors (e.g. red) are used in areas of a high frequency of mouse activity, and cold background colors (e.g. blue) are used in areas of a low frequency of mouse activity. Different frequency levels are represented by different shades of hot and cold colors. Heatmaps can help website maintainers by revealing the distribution of user attention to areas of the web pages, so that the structure and the content of web pages can be optimized and improved accordingly.

Significant resources are required to track, record, and store mouse movement data. Therefore, mouse movement tracking is currently not offered by free services, such as Google Analytics. However, many other commercial web analytics services track mouse clicks and movements [10].

As discussed above, mouse activity is considered an important source of information on user attention. Previous work, including in research and in commercial web analytics, has focused on the mouse cursor position as an implicit indicator of which page areas capture user attention. This study takes a different perspective and focuses on the directions and speeds of mouse movements, and what they indicate.

3 DATA IN THIS STUDY

This study is based on web usage data from the ObjectDB website (https://www.objectdb.com). 137 web pages with technical-educational information on ObjectDB and JPA (Java Persistence API) were used in this research. All these pages have a similar structure. Figure 1 shows the top of one of these pages.

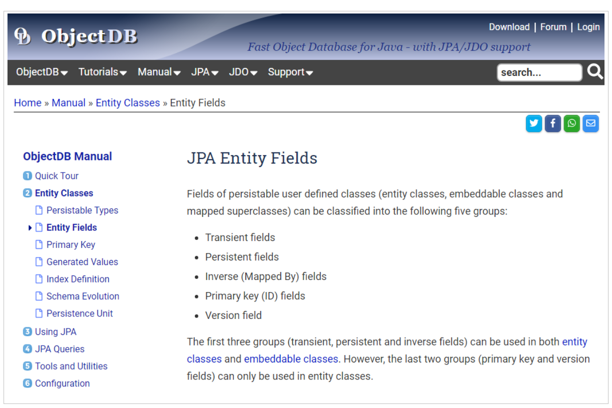

Figure 1: The ObjectDB Website

During a period of several months, ending in March 2020, mouse activity data of visitors to the website were recorded and stored anonymized in a database, adhering to industry standards of data anonymization and user privacy preservation. The implementation included client-side JavaScript code that was added to the website pages. When a page from the website was loaded into the user's browser, the JavaScript code captured the 'onmousemove' events. In order to save resources, the event rate was limited to one event per one-tenth of a second (or 10 events per second), as this was sufficient for the purposes of this research. The JavaScript code found for every cursor position the region on the page (e.g. menu, content, etc.) containing that position. This must be done on the client-side as page presentation is client dependent. All the collected data, including mouse event times, cursor positions, and their corresponding regions on the page, were reported back to the server and stored anonymized in a dedicated database. More information about the data collection implementation used, including privacy and data protection considerations, can be found in [11].

In total, 570,135 views of these 137 web pages have been tracked. Page views with no mouse movements at all (e.g. mobile users' page views and page views with no activity at all) were excluded, reducing the dataset to 509,800 page views. Table 1 provides more information on this dataset.

Table 1: General Details on the Web Usage Dataset

Web pages	137
Page views	509,800
Unique visitors (estimated)	207,195
Sampled mouse moves (max 10 per sec.)	47,390,544
Sampled mouse moves per page view (average)	92.96
Visibility time per page view (average)	518.3 sec.
Mouse movement time per page view (average)	9.3 sec.
% of the visibility time with mouse movements	1.8%

Some of the numbers in Table 1 are estimates. The number of unique visitors is based on browser fingerprint hash and counts a user that uses multiple computers or browsers (or even changed some browser settings) more than once. The main reason for the low ratio of page views per unique visitor is that this website is often used by occasional visitors, who arrive from a search engine after searching for particular information about JPA and then view only one page. Visibility time is based on the total time that a page was open in an active browser tab. There is no way to know how much of this time the user actually looked at the page (as opposed to looking at another screen, another window, etc.), so this is an upper bound.

Focusing on one specific website may have the typical advantages of a case study research, but in this case, it is also a constraint. There is a big difference between the main types of web mining. In web content and structure mining, the available data are virtually unlimited, as we can access public web content and web structures externally. In web usage mining, on the other hand, almost no data are available publicly. Obtaining data even from a single website could be a challenging barrier (due to policies of companies and organizations, user privacy issues, etc.). Regarding the size of the dataset in this study, it is considerably larger than in most previous studies on mouse activity. This is essential for a successful statistical analysis.

Although it is reasonable to expect that the results of this research are not unique for the selected website and can be extrapolated to other websites, due to the study focusing only on one website, further work on other websites is needed in order to establish and generalize these results.

4 STATISTICAL ANALYSIS

This section presents statistics on the mouse movements in this case study and highlights observations about mouse movement directions and mouse movement speeds. The results of the statistical analysis show general patterns of mouse movement behavior. The full picture becomes clear in section 5, which examines these statistical findings using examples of mouse activity of real users.

4.1 Directions of Mouse Moves

Directions of mouse movements are represented in this paper as angles measured in degrees in the polar coordinate system. Figure 2 shows the distribution of all the mouse move events in the case study dataset (47,390,544 in total, as shown in Table 1), in the following four directions: Right (0°±45°), Up (90°±45°), Left (180°±45°), and Down (270°±45°). Polar pie charts are very convenient for this purpose, as all four sectors have equal angles (90° as shown in Figure 2), so directions are preserved. Throughout this paper, right is represented by green, left by red, and up and down by yellow.

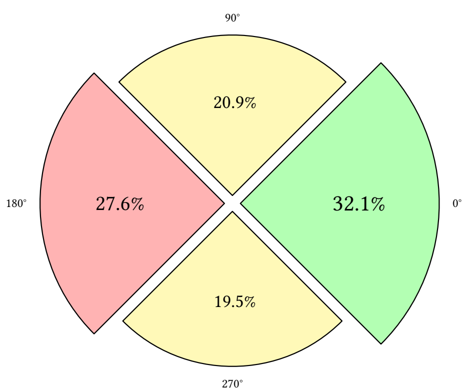

Figure 2: The Distribution of Mouse Moves in 4 Directions

Figure 2 provides the first indication that mouse moves in some directions are more common than mouse moves in other directions. We can see that horizontal moves are more frequent than vertical moves, and that moves to the right are more frequent than moves to the left. Note that due to the large amount of data, all the differences that are highlighted in this paper have a very high statistical significance. Even the modest difference between UP and DOWN (20.9% against 19.5%), which seems relatively small, has a very high statistical significance (p-value < 0.0001). Therefore, the interesting question is not about statistical significance, but which observed differences indicate something meaningful.

More detailed statistics are presented in Figure 3, where the data are divided into 12 narrower sectors of 30°, representing 12 directions: 0°±15°, 30°±15°, ..., 330°±15°.

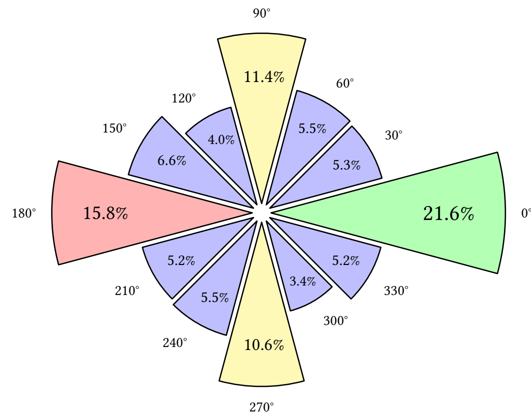

Figure 3: The Distribution of Mouse Moves in 12 Directions

Figure 3 exposes additional interesting details. The first notable observation is that most mouse moves are either relatively horizontal or relatively vertical (i.e. approximately parallel to the X and Y axes). The four sectors in green, red, and yellow constitute 33.3% of the circle sectors but represent 59.4% of the mouse moves. When users move the cursor from point A (current position) to point B (a target, e.g. a link), it is reasonable to expect them to prefer the shortest path, rather than move like a rook on a chessboard, only horizontally and vertically. Therefore, these differences may indicate that there are some common user activities or behaviors that involve moving the mouse in horizontal and vertical directions. In section 5, we look at relevant examples from the real usage data that can explain these behaviors.

Figure 3 also emphasizes the differences that we already saw in Figure 2, i.e. horizontal moves are more frequent than vertical moves, and moves to the right are more frequent than moves to the left. These more detailed statistics demonstrate considerably wider gaps.

A natural question arises as to the significant difference between moves to the right and moves to the left. If statistical data show that for a given period of time an elevator moved in total 21.6km up and 15.8km down, we are more likely to be skeptical about the data, rather than rush to look for that elevator in the sky. So is it possible to have this asymmetry between horizontal moves to the right and to the left? The answer is simple. The mouse events represent time rather than distance, or more precisely, each mouse move represents one-tenth of a second (due to the sampling rate used in the data collecting process of this case study). Therefore, Figures 2 and 3 expose differences in mouse movement times rather than differences in mouse movement distances. An interesting consequence of this discussion is that now we should expect to see that mouse movements to the left are faster than mouse movements to the right, in a way that equates the total traveling distances in these directions. This is examined in subsection 4.3.

4.2 From Moves to Movements

According to the terminology of this paper, mouse moves are individual mouse move events, sampled on the client-side using

JavaScript, with a frequency of up to 10 events per second. A mouse movement is a sequence of consecutive mouse moves.

Working with movements, which contain more information than single moves, could be beneficial. Grouping mouse moves into meaningful mouse movements is not a trivial task, but for the purpose of this work the simplified method used was found to be effective. Dividing the polar coordinate system into 12 sectors, representing 12 directions (as in Figure 3), a movement in this work is defined as a sequence of moves that share the same direction (one of the 12 directions), with no time gaps longer than 5 seconds between every two moves (the 5 seconds threshold is sufficient to allow small pauses without breaking a movement, but it was chosen arbitrarily and other values could be used instead).

The clear advantage of this approach of grouping moves into movements is its simplicity, as the mouse moves can be easily grouped into movements, and each movement has a distinct direction (one of 12 in this configuration), which is the direction of all the moves that it consists of. This simplicity has its price, as complex movements are split into smaller movements by direction, though for the purpose of this work this grouping method is sufficient.

Every move is part of exactly a single movement, but since the motivation is to progress from examining individual moves to examining larger movements, small movements are excluded from the following data analysis. Table 2 compares several alternative thresholds for minimum movement size, based on the minimum number of moves, and by extension, the minimum moving time (not the elapsed time).

Table 2: Move and Movement Counting

	Moves	Movements		
Moves	1	≥ 3	≥ 5	≥ 10
Moving Time	100ms	≥ 300ms	≥ 500ms	≥ 1000ms
Covered Moves	47,390,544	11,402,534	6,605,021	2,622,996
Units	47,390,544	2,253,789	815,910	171,943
On Content	43,327,556	1,998,396	745,400	165,312
On Left Menu	1,234,069	53,206	15,917	1,390
On Top Menu	299,526	17,025	5,968	859
Elsewhere	2,529,393	185,162	48,625	4,382

Even with a very low threshold of three moves (i.e. filtering out movements made of one or two moves), we retain 2,253,789 movements, which in total consist of only 11,402,534 moves out of the total of 47,390,544 available moves, i.e. more than 75.9% of the moves are excluded. This highlights the limitation of this approach for grouping moves in strict directions. However, as the dataset used is large, aggregating moves into movements helps in focusing on meaningful movements, and provides a sample that is sufficiently large for further statistical analysis.

The bottom four rows in Table 2 show the distribution of mouse moves and movements on the main areas of the web pages. Clearly, most of the mouse movement activity on this website is in the content area, which contains mainly text, and up to 64 pixels in the left and right margins (depending on the resolution). The 'elsewhere' row represents the page header, the breadcrumbs, the search box, the left sidebar, the footer, and several other page elements.

Figures 4 and 5 show the distribution of mouse movements by direction, for two movement populations (the rightmost two columns in Table 2). Revisiting Figure 3, we can see a trend. The transition from moves to movements strengthens the main observations. Movements parallel to the axes are more frequent, horizontal movements are more frequent than vertical movements, and movements to the right are more frequent than movements to the left. The differences are much more extreme when considering longer movements (in Figure 5).

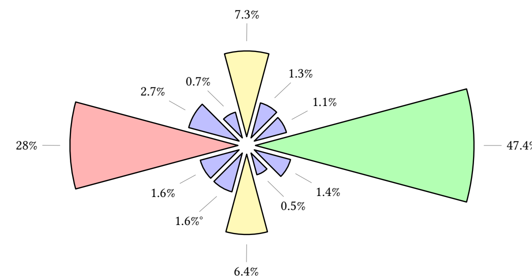

Figure 4: Mouse Movements ≥ ½ sec. by Direction

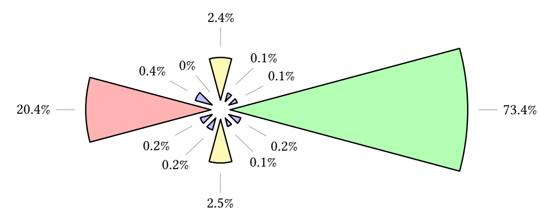

Figure 5: Mouse Movements ≥ 1 sec. by Direction

Table 3 summarizes all of the statistical data presented in subsections 4.1 and 4.2, and adds the absolute numbers of moves and movements in each direction and in total.

Table 3: Mouse Movements by Direction

Direction	Moves		Movements ≥ ½s		Movements ≥ 1s	
± 15°	Count	Share	Count	Share	Count	Share
0°	10,227,909	21.6%	387,149	47.4%	126,271	73.4%
30°	2,500,397	5.3%	9,308	1.1%	144	0.1%
60°	2,598,764	5.5%	10,206	1.3%	174	0.1%
90°	5,394,427	11.4%	59,186	7.3%	4,135	2.4%
120°	1,887,872	4.0%	5,856	0.7%	63	0.0%
150°	3,115,501	6.6%	21,783	2.7%	617	0.4%
180°	7,470,862	15.8%	228,564	28.0%	35,140	20.4%
210°	2,483,094	5.2%	12,683	1.6%	289	0.2%
240°	2,595,021	5.5%	13,431	1.6%	404	0.2%
270°	5,028,383	10.6%	52,175	6.4%	4,322	2.5%
300°	1,612,489	3.4%	4,466	0.5%	89	0.1%
330°	2,475,825	5.2%	11,103	1.4%	295	0.2%
Total	47,390,544	100%	815,910	100%	171,943	100%

4.3 Mouse Movement Speeds

Figures 6 and 7 show average speeds (in pixels per second) of movements in different directions. As expected, movements to the right are considerably slower than movements to the left. More complete data regarding speeds of moves and movements are included in Table 4.

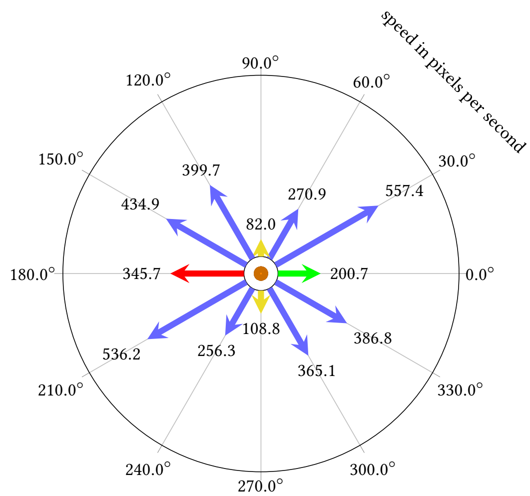

Figure 6: Average Speed of Movements ≥ ½ sec. by Direction

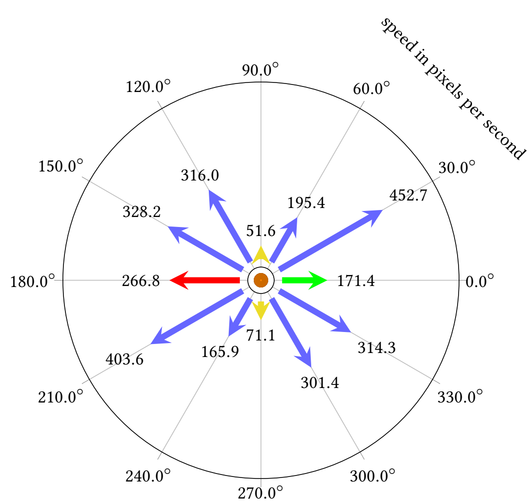

Figure 7: Average Speed of Movements ≥ 1 sec. by Direction

Note that the movement speed in this context is calculated as the 'air distance' from the beginning position of the first move event to the end position of the last move event, divided by the time elapsed from the first event to the last event. It is an average speed, and so, during that elapsed time there may be intervals in which the mouse is static.

Figures 6 and 7 and Table 4 show also that vertical movements are much slower than movements in other directions, and that movements in the directions of 30° and 210° are faster than movements in other directions.

Table 4: Average Mouse Speed (pixels per sec.) by Direction

Direction	Moves		Movements ≥ ½s		Movements ≥ 1s	
± 15°	Count	Speed	Count	Speed	Count	Speed
0°	10,227,909	449.1	387,149	200.7	126,271	171.4
30°	2,500,397	754.8	9,308	557.4	144	452.7
60°	2,598,764	531.4	10,206	270.9	174	195.4
90°	5,394,427	455.2	59,186	82.0	4,135	51.6
120°	1,887,872	579.8	5,856	399.7	63	316.0
150°	3,115,501	586.8	21,783	434.9	617	328.2
180°	7,470,862	587.4	228,564	345.7	35,140	266.8
210°	2,483,094	736.3	12,683	536.2	289	403.6
240°	2,595,021	497.3	13,431	256.3	404	165.9
270°	5,028,383	451.6	52,175	108.8	4,322	71.1
300°	1,612,489	565.2	4,466	365.1	89	301.4
330°	2,475,825	574.7	11,103	386.8	295	314.3
Total	47,390,544		815,910		171,943	

We already know that in this dataset horizontal movements are much more common than movements in other directions. It is interesting to compare the proportions of movements to the right and to the left out of all the movements, across different movement speeds, as shown in Figures 8 and 9.

On a side note, comparing the right side (speeds of 1,000 pixels per second and above) of these two figures demonstrates the cost of using a higher threshold of 1 second when grouping moves into movements. The curves in Figure 9 are less smooth, probably due to the relatively small amount of data available at these speeds, as Table 6 in appendix A shows.

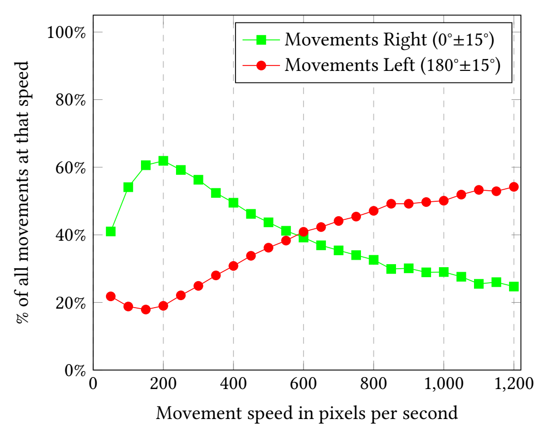

Figure 8: Proportion of Movements ≥ ½ sec. by Speed

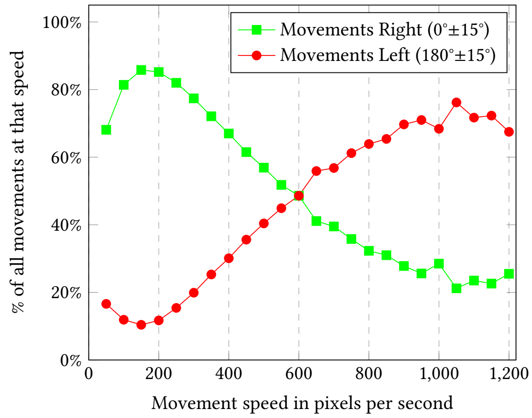

Figure 9: Proportion of Movements ≥ 1 sec. by Speed

Movements to the right peak at around 150-200 pixels per second and are much less dominant at higher speeds. At speeds higher than 600 pixels per second, movements to the left are more common than movements to the right.

5 EXAMPLES OF MOUSE ACTIVITY

This section uses mouse movements of real users from the dataset of this study to demonstrate user behavior that may explain the main statistical observations made in section 4. There is no attempt to cover all possible behaviors.

Mouse cursor positions are displayed on the web pages as orange circles. The lines connecting adjacent circles represent mouse move events (usually during one-tenth of a second, except after mouse movement stops).

5.1 Horizontal Movements and Reading

The main statistical findings in section 4 regarding horizontal mouse movements are as follows:

- Horizontal movements are more frequent than movements in other directions, indicating that they may be related to some sort of user activity or behavior.
- Movements to the right are more frequent in general than movements to the left (in terms of the number of sampled mouse move events and the total movement time, but not necessarily in terms of total distance).
- On average, movements to the left are faster and movements to the right are slower.
- At lower speeds movements to the right are more frequent than movements to the left, with a peak at around 150-200 pixels per second. At higher speeds (above 600 pixels per second), movements to the left are more frequent than movements to the right (see Figures 8 and 9).

Figure 10 shows intensive horizontal mouse movements in a page view of a real user, which closely align with the text. As discussed in section 2, previous work has shown a correlation between eye gaze and mouse cursor positions on a screen [7], and the correlation is higher during mouse activity, i.e. when a user clicks or moves the mouse [3, 20]. Therefore, the mouse activity in Figure 10 probably indicates that the mouse cursor was moved along the text while the user was reading it.

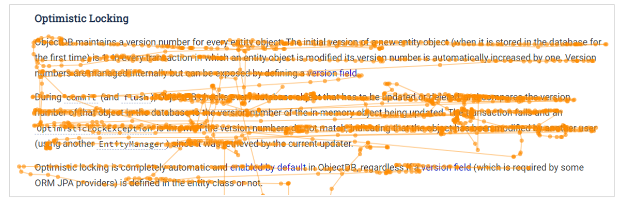

Figure 10: Nearly Horizontal Movements

Figures 11 and 12 show a separation of the movements in Figure 10 to movements to the right and to the left (respectively). Clearly, there are more move events to the right than to the left. This is logical as movements to the right (the reading direction in this website's language, English) are expected to be slower, at a reading rate. Movements to the left (to the beginning of the next text lines) are faster, and therefore, contain less sampled move events.

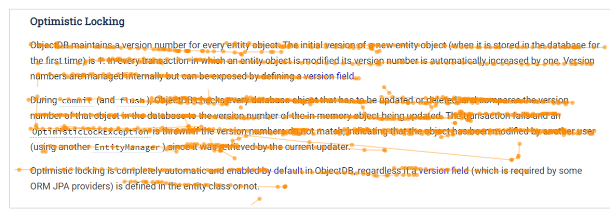

Figure 11: Movements Right (From Figure 10)

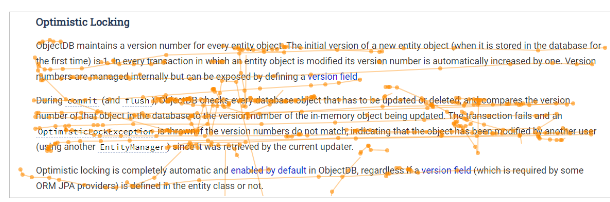

Figure 12: Movements Left (From Figure 10)

Figure 13 shows horizontal movements of another user. In this example, the cursor is moved more freely but still seems to accompany the reading pattern of the user.

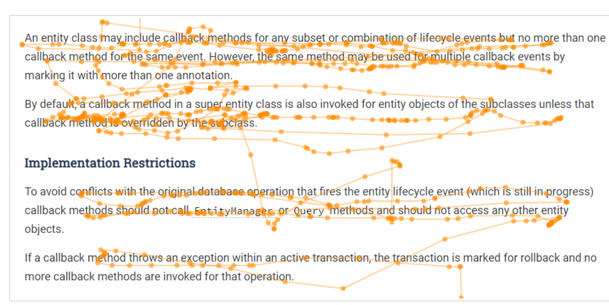

Figure 13: Loose Horizontal Movements

The peak of movements to the right in Figures 8 and 9 at around 150-200 pixels per second, can be explained in the context of the average reading speed. According to a rough estimate (shown in appendix B), the average word width on this website is about 50 pixels. A reading rate of 150-200 pixels per second is equivalent to 3-4 words per second, or 180-240 words per minute (wpm). A recent review that examined 190 different studies (17,887 participants) evaluated that most adults read non-fiction English at a rate of 175–300 wpm, with an average of 238 wpm [1]. This may explain why movements to the right, if used mainly for reading, are relatively more frequent at these speeds. Note that reading speeds on this website may be slightly slower than the average reading speed due to various reasons, including the complexity of the technical material, and the audience, which consists mainly of non-native English speakers. Movements in the direction of 210° may be faster as this direction is related to moving from the end of one text line to the beginning of the next line (i.e. not restricted by reading speed). Movements in the opposite direction (30°) are also relatively fast, which requires further investigation.

5.2 Vertical Movements and Reading

Vertical movements, according to the statistical findings in section 4, are less frequent than horizontal movements, more frequent than movements in other directions, and slower than horizontal movements. Figures 14, 15, 16, 17, and 18 demonstrate vertical movements of several users. A reasonable explanation of this activity is that users mark the currently read line of text with the cursor. Marking text lines using vertical movements seems like a less demanding version of marking words with horizontal movements. It gives the user a weaker indication of the reading position, but less effort is needed in moving the mouse cursor along the text lines to keep it approximately synchronized with the exact reading position. The fact that this activity requires fewer and slower movements may possibly support the statistical findings that vertical movements are less frequent than horizontal movements (although further investigation is required).

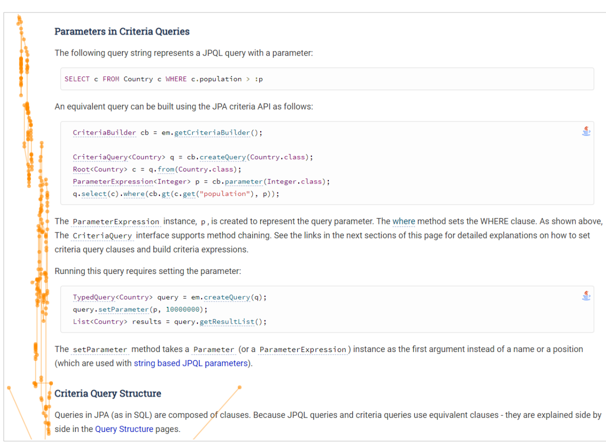

Figure 14: Vertical Movements on the Left Margin

The repeating movements up and down in Figure 14 are confusing. We expect users to read mainly in one direction, from top to bottom, occasionally backtracking upwards. Down movements are obviously needed to synchronize the cursor position with the reading position, which progresses down the page. A possible explanation of the prevalence of up movements is related to page scrolling. When a page is scrolled *n* lines down, the cursor position in the window is unchanged, but the cursor position relative to the document is moved down *n* text lines that have not been read yet. Therefore, moving the cursor back to the reading position is needed.

Figure 15 demonstrates much less movement activity, showing that it is possible to avoid the routine adjustments of the cursor upwards after scrolling, by scrolling down a line at a time (using the mouse wheel or the down key on the keyboard). This way, the cursor can mark the current reading line without having to move the mouse frequently up and down. We can still see some mouse movements in Figure 15, possibly in regions where the user spent more time (the table of contents and the two code fragments).

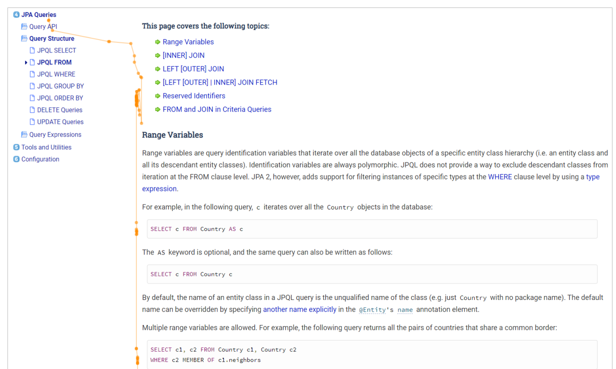

Figure 15: Scrolling with Vertical Movements

Moving the mouse cursor to mark the reading position is more common on both margins (left or right), but some users move the mouse vertically on the text, as shown in Figure 16, possibly to mark the reading position.

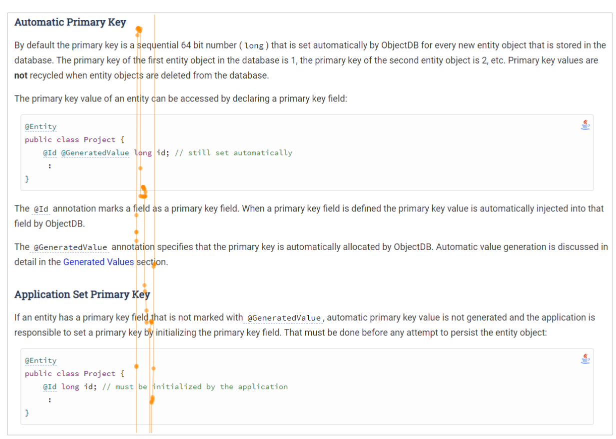

Figure 16: Vertical Movements on the Content

Figures 17 and 18 show vertical movements on the right margin.

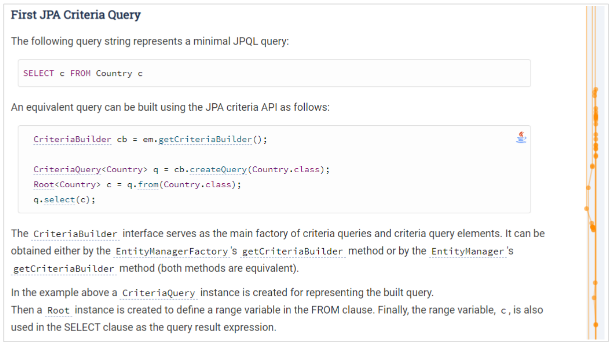

Figure 17: Vertical Movements on the Right Margin

Figure 18 shows vertical movements on the right margin, probably to mark the reading position, with occasional horizontal movements to mark words.

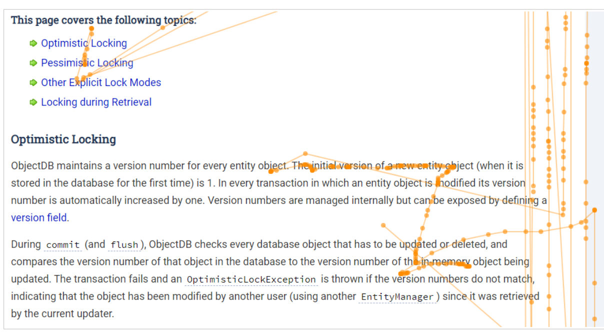

Figure 18: Vertical and Horizontal Movements Combined

Figure 19 shows a different pattern of vertical movements. The movements up and down are too extreme for cursor adjustments. It is possible that the user fidgeted with the mouse while reading.

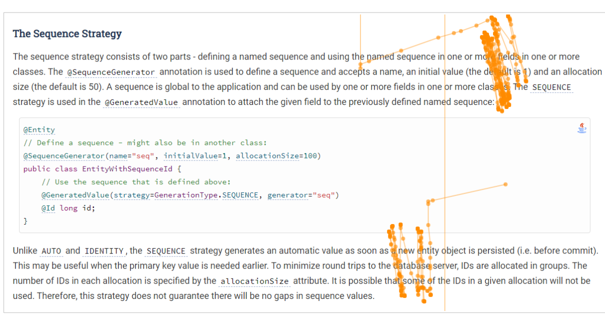

Figure 19: Dense Vertical Movements on Specific Areas

5.3 Menu Mouse Movement

Mouse movements are also used for navigation. The website in this case study has a nested menu bar at the top and an additional menu on the left. Figure 20 shows navigation in the main menu.

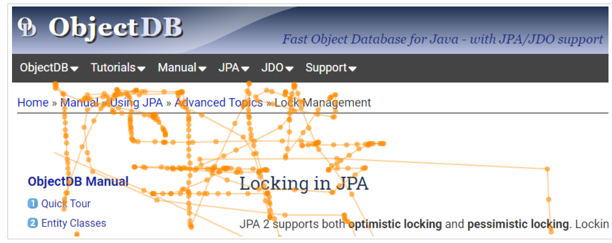

Figure 20: Using the Top Menu

The visualization in Figure 20 is confusing because all of the popup menus on the menu bar are shown closed. The mouse cursor moved on the menu when it was open. Using this menu also requires horizontal and vertical movements, but because mouse movements outside the content area of the page are relatively uncommon (as shown in Table 2), this probably had a minor effect on the statistical results in section 4. Figure 21 shows mouse movements on the left side menu.

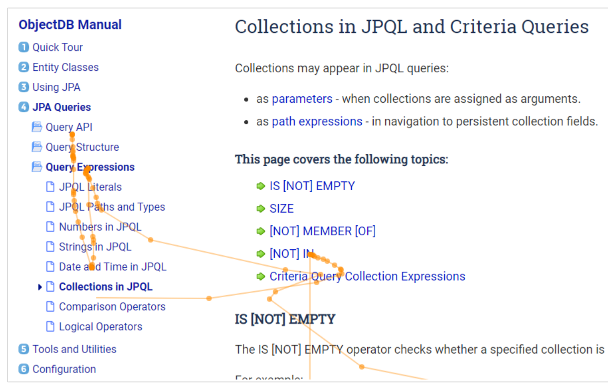

Figure 21: Using the Left Side Menu

The user, whose mouse movements are shown in Figure 21, arrived at the page shown using the left side menu (the "Collections in JPQL link" on another page). After reading the page content (which is out of the frame of this screenshot), the user scrolled up to the top of the page and used the left side menu again to move to another page. Figure 21 demonstrates mouse movements in different directions: horizontal, vertical, and others.

6 CONCLUSIONS AND FURTHER WORK

This study presents a two-phase analysis of mouse movements on a selected website. The first phase reveals several statistical phenomena in the examined data, most notably that most of the mouse movements are either approximately horizontal or approximately vertical, horizontal mouse movements are more frequent than vertical mouse movements, and movements to the right are slower than movements to the left. The second phase shows examples of mouse movements of real users, which indicate that these statistical phenomena may be related to the way that some readers use the mouse cursor as a reading assistant tool, marking the reading position while reading. Using the cursor of a pointing device (e.g.

a mouse or a touchpad) as an assistant tool while reading, can be referred to as **Pointer Assisted Reading (PAR)**.

Besides the scientific interest, these findings may contribute to various applications of web usage mining, web analytics, and online reading research. For example, by tracking the reading position it may be possible to estimate user reading speeds and assess user reading behaviors. This could replace eye-tracking in applications where it is not viable, including on most websites, as discussed in section 2. In addition, by tracking the reading position, we may be able to learn about the content of a website by exploring the reading patterns of users. For example, text that most users tend to skip might not interest the website audience. Similarly, sentences that many readers read more than once or more slowly, may be insufficiently clear. This could be useful in locating issues in textual content, in websites in general, and in particular in specific applications such as online learning.

It is important to note that this reading technique (moving the mouse while reading) is not practiced by everyone all the time. There are many page views in this case study dataset with no sign of this mouse movement activity at all. As a rough indication, the mouse was moved (for any purpose) less than 2% of the time that the page was visible in the browser (on average, according to Table 1). Therefore, the extent of this reading behavior may be significant to form these statistical results but not sufficient to follow every individual user. However, the applications that are mentioned above do not require data from every user. Data from sample users that use PAR could be sufficient in web usage mining, as reviews of a product on a shopping website or comments on an article on a news website, even from some users, could be sufficient in relevant applications of web content mining.

There are many open questions that require further work. First of all, how many visitors use the mouse cursor as a reading assistant tool, and how are they divided between the horizontal movers who mark words and the vertical movers who mark lines of text? In addition, before adopting these users' data as a sample that represents the whole website audience, we need to investigate possible biases. Therefore, we have to study which people are more likely to use PAR and which are less likely.

As discussed, obtaining data for web usage mining research is usually much more challenging than for web content mining research, and therefore, most of such studies focus on a single website. This study is no different, and further work is required on web usage data from other websites. For example, it would be interesting to see if statistical results regarding the right and the left directions are reversed on websites that use RTL languages (such as Hebrew and Arabic).

Further work may include developing methods for locating and recognizing PAR activity. A PAR recognizer may help in solving the questions above, as well as in exploring the discussed usage opportunities.

REFERENCES

[1] Marc Brysbaert. 2019. How many words do we read per minute? A review and meta-analysis of reading rate. *Journal of Memory and Language* 109 (2019), 84. https://doi.org/10.1016/j.jml.2019.104047

[2] Catia Cepeda, Joao Rodrigues, Maria Camila Dias, Diogo Oliveira, Dina Rindlisbacher, Marcus Cheetham, and Hugo Gamboa. 2018. Mouse Tracking Measures and Movement Patterns with Application for Online Surveys. In *Machine Learning and Knowledge Extraction*, Andreas Holzinger, Peter Kieseberg, A Min Tjoa, and Edgar Weippl (Eds.). Springer International Publishing, Cham, 28–42.

[3] Mon Chu Chen, John R. Anderson, and Myeong Ho Sohn. 2001. What Can a Mouse Cursor Tell Us More? Correlation of Eye/Mouse Movements on Web Browsing. In *CHI '01 Extended Abstracts on Human Factors in Computing Systems* (Seattle, Washington) *(CHI EA '01)*. Association for Computing Machinery, New York, NY, USA, 281–282. https://doi.org/10.1145/634067.634234

[4] Mark Claypool, Phong Le, Makoto Wased, and David Brown. 2001. Implicit Interest Indicators. In *Proceedings of the 6th International Conference on Intelligent User Interfaces* (Santa Fe, New Mexico, USA) *(IUI '01)*. Association for Computing Machinery, New York, NY, USA, 33–40. https://doi.org/10.1145/359784.359836

[5] Sukru Eraslan, Yeliz Yesilada, and Simon Harper. 2020. "The Best of Both Worlds!": Integration of Web Page and Eye Tracking Data Driven Approaches for Automatic AOI Detection. *ACM Trans. Web* 14, 1, Article 1 (Jan. 2020), 31 pages. https://doi.org/10.1145/3372497

[6] Qi Guo and Eugene Agichtein. 2008. Exploring Mouse Movements for Inferring Query Intent. In *Proceedings of the 31st Annual International ACM SIGIR Conference on Research and Development in Information Retrieval* (Singapore, Singapore) *(SIGIR '08)*. Association for Computing Machinery, New York, NY, USA, 707–708. https://doi.org/10.1145/1390334.1390462

[7] Jeff Huang, Ryen White, and Georg Buscher. 2012. User See, User Point: Gaze and Cursor Alignment in Web Search. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (Austin, Texas, USA) *(CHI '12)*. Association for Computing Machinery, New York, NY, USA, 1341–1350. https://doi.org/10.1145/2207676.2208591

[8] Shamsi Iqbal and Brian Bailey. 2004. Using Eye Gaze Patterns to Identify User Tasks. *The Grace Hopper Celebration of Women in Computing* 04 (01 2004), 6 pages.

[9] Yousra Javed and Mohamed Shehab. 2017. Look before you Authorize: Using Eye-Tracking to Enforce User Attention towards Application Permissions. *Proceedings on Privacy Enhancing Technologies* 2017, 2 (2017), 23–37. https://content.sciendo.com/view/journals/popets/2017/2/article-p23.xml

[10] Ilan Kirsh and Mike Joy. 2020. A Different Web Analytics Perspective Through Copy to Clipboard Heatmaps. In *Proceedings of the 20th International Conference on Web Engineering (ICWE 2020), Lecture Notes in Computer Science, vol 12128*. Springer International Publishing, Cham, 543–546. https://doi.org/10.1007/978-3-030-50578-3_41

[11] Ilan Kirsh and Mike Joy. 2020. Exploring Pointer Assisted Reading (PAR): Using Mouse Movements to Analyze Web Users' Reading Behaviors and Patterns. In *Proceedings of the 22nd HCI International Conference (HCII 2020), Lecture Notes in Computer Science*. Springer International Publishing, Cham, 18 pages.

[12] Ilan Kirsh and Mike Joy. 2020. Splitting the Web Analytics Atom: From Page Metrics and KPIs to Sub-Page Metrics and KPIs. In *Proceedings of the 10th International Conference on Web Intelligence, Mining and Semantics (WIMS 2020), Biarritz, France*. Association for Computing Machinery, New York, NY, USA, 33–43. https://doi.org/10.1145/3405962.3405984

[13] Fabrizio Lamberti and Gianluca Paravati. 2015. VDHM: Viewport-DOM Based Heat Maps as a Tool for Visually Aggregating Web Users' Interaction Data from Mobile and Heterogeneous Devices. In *Proceedings of the 2015 IEEE International Conference on Mobile Services (MS '15)*. IEEE Computer Society, USA, 33–40. https://doi.org/10.1109/MobServ.2015.15

[14] F. Lamberti, Gianluca Paravati, Valentina Gatteschi, and Alberto Cannavò. 2017. Supporting Web Analytics by Aggregating User Interaction Data From Heterogeneous Devices Using Viewport-DOM-Based Heat Maps. *IEEE Transactions on Industrial Informatics* 13 (08 2017), 1989 – 1999. Issue 4. https://doi.org/10.1109/TII.2017.2658663

[15] JooWon Lee and Jae-Hyeon Ahn. 2012. Attention to Banner Ads and Their Effectiveness: An Eye-Tracking Approach. *International Journal of Electronic Commerce* 17, 1 (2012), 119–137. https://doi.org/10.2753/JEC1086-4415170105 arXiv:https://doi.org/10.2753/JEC1086-4415170105

[16] Yixuan Li, Pingmei Xu, Dmitry Lagun, and Vidhya Navalpakkam. 2017. Towards Measuring and Inferring User Interest from Gaze. In *Proceedings of the 26th International Conference on World Wide Web Companion* (Perth, Australia) *(WWW '17 Companion)*. International World Wide Web Conferences Steering Committee, Republic and Canton of Geneva, CHE, 525–533. https://doi.org/10.1145/3041021.3054182

[17] Alexandre Milisavljevic, Kevin Hamard, Coralie Petermann, Bernard Gosselin, Karine Doré-Mazars, and Matei Mancas. 2018. Eye and Mouse Coordination During Task: From Behaviour to Prediction. In *International Conference on Human Computer Interaction Theory and Applications* (Funchal, Madeira, Portugal). SciTePress, Setúbal, Portugal, 86–93. https://doi.org/10.5220/0006618800860093

[18] A. Marthe Möller, Susanne E. Baumgartner, Rinaldo Kühne, and Jochen Peter. 2019. The effects of social information on the enjoyment of online videos: An eye tracking study on the role of attention. *Media Psychology* 0, 0 (2019), 1–22. https://doi.org/10.1080/15213269.2019.1679647

[19] Vidhya Navalpakkam, LaDawn Jentzsch, Rory Sayres, Sujith Ravi, Amr Ahmed, and Alex Smola. 2013. Measurement and Modeling of Eye-Mouse Behavior in the Presence of Nonlinear Page Layouts. In *Proceedings of the 22nd International*

Conference on World Wide Web (Rio de Janeiro, Brazil) *(WWW '13)*. Association for Computing Machinery, New York, NY, USA, 953–964. https://doi.org/10.1145/2488388.2488471

[20] Kerry Rodden and Xin Fu. 2007. Exploring How Mouse Movements Relate to Eye Movements on Web Search Results Pages. In *Proceedings of ACM SIGIR 2007 Workshop on Web Information Seeking and Interaction*. Association for Computing Machinery, New York, NY, USA, 29–32. http://research.microsoft.com/~ryenw/proceedings/WISI2007.pdf

[21] Johannes Schneider, Markus Weinmann, Jan vom Brocke, and Christoph Schneider. 2017. Identifying Preferences Through Mouse Cursor Movements – Preliminary Evidence. In *Proceedings of the 25th European Conference on Information Systems (ECIS)*. Research-in-Progress Papers, Guimarães, Portugal, 2546–2556.

[22] Katerina Tzafilkou, Nicolaos Protogeros, and C.G. Yakinthos. 2014. Mouse Tracking for Web Marketing: Enhancing User Experience in Web Application Software by Measuring Self-Efficacy and Hesitation Levels. *International Journal on Strategic Innovative Marketing* 01 (2014). https://doi.org/10.15556/IJSIM.01.04.005

[23] Jun Wang, Eugene Yujun Fu, Grace Ngai, Hong Va Leong, and Michael Xuelin Huang. 2019. Detecting Stress from Mouse-Gaze Attraction. In *Proceedings of the 34th ACM/SIGAPP Symposium on Applied Computing* (Limassol, Cyprus) *(SAC '19)*. Association for Computing Machinery, New York, NY, USA, 692–700. https://doi.org/10.1145/3297280.3297345

[24] Saniya Zahoor, Mangesh Bedekar, and Pranali K. Kosamkar. 2014. User Implicit Interest Indicators Learned from the Browser on the Client Side. In *Proceedings of the 2014 International Conference on Information and Communication Technology for Competitive Strategies* (Udaipur, Rajasthan, India) *(ICTCS '14)*. Association for Computing Machinery, New York, NY, USA, Article 57, 4 pages. https://doi.org/10.1145/2677855.2677912

[25] Oleg Špakov and Darius Miniotas. 2007. Visualization of eye gaze data using heat maps. *Elektronika ir Elektrotechnika - Medicine Technology* 115 (01 2007).

A THE DISTRIBUTION BY SPEED & DIRECTION

Tables 5 and 6 contain the data that Figures 8 and 9 are based on (respectively), including absolute values.

Table 5: Proportion of Movements ≥ ½ sec. by Speed

Speed ±25	All	Right (0°±15°)		Left (180°±15°)	
(pixel/sec.)	Movements	Count	Share	Count	Share
50	90,750	37,163	41.0%	19,807	21.8%
100	89,308	48,325	54.1%	16,786	18.8%
150	86,116	52,199	60.6%	15,454	17.9%
200	74,115	45,856	61.9%	14,105	19.0%
250	61,249	36,229	59.2%	13,553	22.1%
300	49,064	27,644	56.3%	12,201	24.9%
350	40,128	21,020	52.4%	11,251	28.0%
400	33,621	16,647	49.5%	10,342	30.8%
450	28,376	13,113	46.2%	9,589	33.8%
500	24,214	10,591	43.7%	8,776	36.2%
550	20,582	8,474	41.2%	7,881	38.3%
600	18,220	7,136	39.2%	7,443	40.9%
650	15,877	5,853	36.9%	6,716	42.3%
700	13,630	4,819	35.4%	6,007	44.1%
750	12,160	4,129	34.0%	5,521	45.4%
800	10,686	3,480	32.6%	5,038	47.1%
850	9,206	2,750	29.9%	4,530	49.2%
900	8,172	2,461	30.1%	4,017	49.2%
950	7,092	2,048	28.9%	3,528	49.7%
1,000	6,345	1,840	29.0%	3,181	50.1%
1,050	5,618	1,548	27.6%	2,915	51.9%
1,100	5,065	1,292	25.5%	2,699	53.3%
1,150	4,488	1,168	26.0%	2,376	52.9%
1,200	4,081	1,008	24.7%	2,210	54.2%

Table 6: Proportion of Movements ≥ 1 sec. by Speed

Speed ±25	All	Right (0°±15°)		Left (180°±15°)	
(pixel/sec.)	Movements	Count	Share	Count	Share
50	16,066	10,940	68.1%	2,670	16.6%
100	24,700	20,097	81.4%	2,931	11.9%
150	28,144	24,136	85.8%	2,932	10.4%
200	24,270	20,669	85.2%	2,841	11.7%
250	18,314	15,014	82.0%	2,820	15.4%
300	13,007	10,072	77.4%	2,587	19.9%
350	9,240	6,665	72.1%	2,336	25.3%
400	7,022	4,702	67.0%	2,117	30.1%
450	5,325	3,275	61.5%	1,894	35.6%
500	4,163	2,368	56.9%	1,680	40.4%
550	3,117	1,616	51.8%	1,400	44.9%
600	2,527	1,227	48.6%	1,227	48.6%
650	1,861	765	41.1%	1,040	55.9%
700	1,533	606	39.5%	870	56.8%
750	1,176	421	35.8%	720	61.2%
800	993	321	32.3%	635	63.9%
850	768	238	31.0%	502	65.4%
900	683	190	27.8%	476	69.7%
950	472	121	25.6%	335	71.0%
1,000	424	121	28.5%	290	68.4%
1,050	307	65	21.2%	234	76.2%
1,100	230	54	23.5%	165	71.7%
1,150	195	44	22.6%	141	72.3%
1,200	157	40	25.5%	106	67.5%

B PIXELS PER WORD EVALUATION

The rough estimate of the average width in pixels, of words on the case study website, is based on:

- One of the most frequently viewed pages was examined (https://www.objectdb.com/java/jpa/entity/generated), on a 1920 pixel width screen, which is the most commonly used resolution on that website.
- The number of words in the full-text lines on that page (15 in total) have been counted: 22, 22, 22, 22, 24, 22, 25,24, 28, 23, 26, 27, 22, 24, 24, and the average is 23.8 words per line.
- Division of the width of a line in pixels, 1182, with the average number of words, 23.8, results in an estimate of 49.7 pixels per word, on average.