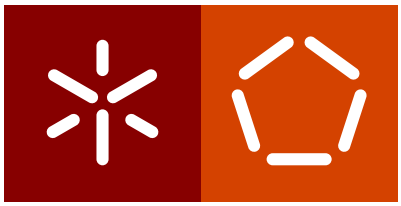
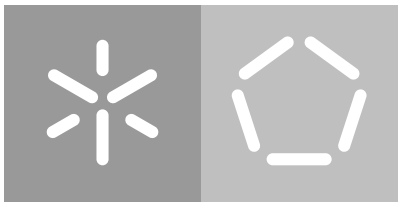


Universidade do Minho
Escola de Engenharia
Departamento de Informática

Rui Pedro Barbosa Rodrigues

NetLangEd, an editor to support Comment Analysis

March 2022



Universidade do Minho
Escola de Engenharia
Departamento de Informática

Rui Pedro Barbosa Rodrigues

NetLangEd, an editor to support Comment Analysis

Master dissertation
Integrated Master's in Informatics Engineering

Dissertation supervised by
Orientador: Pedro Rangel Henriques
Supervisor: Cristiana Araújo

March 2022

STATEMENT OF INTEGRITY

I hereby declare having conducted this academic work with integrity. I confirm that I have not used plagiarism or any form of undue use of information or falsification of results along the process leading to its elaboration.

I further declare that I have fully acknowledged the Code of Ethical Conduct of the University of Minho.

Rui Pedro Barbosa Rodrigues

_____

ACKNOWLEDGMENTS

I want to thank all the teachers in my academic career for having provided me with new knowledge and skills, which will be preponderant in my professional life.

To my supervisor Pedro Rangel Henriques and my supervisor Cristiana Araújo for having accepted me as a oriented and for their dedicated support and guidance throughout the development of this thesis.

All the volunteers who participated in the experiment for taking the time to test the editor and provide their feedback.

To all my friends for the help given in the difficult times I had and for motivating me to always be better.

All my family, especially my parents and brother, for their constant understanding and affection and for the opportunity they gave me to increase my knowledge. I am forever grateful to you!

ABSTRACT

This document formally reports a M.Sc. Thesis project needed to obtain the Master's degree in Informatics Engineering, focusing on the scientific areas of Digital Humanities, Social Networks and Inappropriate Social Discourse. The Master's work here presented was accomplished at Universidade do Minho in Braga.

The main objective of the referred Master's project was the development of an online editor that allows researchers to add their reflections and ideas to short sentences (usually called 'comments') that belong to a social dialogue triggered by a 'post' on a social network or a 'news' on social media. Those comments, to be analyzed by linguists or social science experts, are provided online and are extracted from the corpus created under the international project – NetLang. NetLangEd, the editor developed and here reported, is mainly a tool to allow the analysts to create their own notes to be associated in the right place of each comment while reading it. Basically, NetLangEd allows to highlight a multi-word term contained in the comment, using a color chosen by the user, and associate to that term an 'annotation'. An annotation is composed of two parts, a tag (also created and picked up at the user choice) and a text explaining the user idea. To make this 'annotation' process truly dynamic, NetLangEd provides, through a simple and user-friendly set of menus, three basic operations for adding, editing and removing annotations. Additionally, the editor also provides easy to use mechanisms to manage the tags so far created, as well as to view and locate the annotations.

This Master's dissertation also describes how NetLangEd was tested for usefulness and usability. For that purpose, an experiment was designed and conducted with end-users. The results will be presented and discussed.

Keywords: Text Editor, Online Annotation tool, Markup System, Inappropriate Social Discourse

RESUMO

Este documento reporta formalmente um projeto de mestrado necessário para obter o grau de Mestre em Engenharia Informática, focando-se nas áreas científicas das Humanidades Digitais, Redes Sociais e Discurso Social Impróprio. O trabalho de mestrado aqui apresentado foi realizado na Universidade do Minho em Braga.

O objetivo principal do referido projeto de mestrado era o desenvolvimento de um editor online que permitisse aos investigadores adicionar as suas reflexões e ideias a frases curtas (normalmente denominadas por 'comentários') que pertencem a um diálogo social desencadeado por uma 'publicação' numa rede social ou uma 'notícia' nas redes sociais. Estes comentários, a serem analisados por linguistas ou especialistas em ciências sociais, são disponibilizados online e extraídos do corpus criado no âmbito do projeto internacional – NetLang. NetLangEd, o editor desenvolvido e aqui relatado, é principalmente uma ferramenta para permitir que os analistas criem as suas próprias notas para serem associadas no lugar certo de cada comentário durante a sua leitura. Basicamente, o NetLangEd permite destacar um termo com várias palavras contido no comentário, usando uma cor escolhida pelo usuário, e associar a esse termo uma 'anotação'. Uma anotação é composta por duas partes, uma tag (também criada e escolhida de acordo com a preferência do utilizador) e um texto explicando a ideia do utilizador. Para tornar este processo de 'anotação' verdadeiramente dinâmico, o NetLangEd fornece, por meio de um conjunto de menus simples e de fácil utilização, três operações básicas para adicionar, editar e remover anotações. Adicionalmente, o editor também oferece mecanismos fáceis de usar para gerir as tags criadas até o momento, bem como visualizar e localizar as anotações.

Esta dissertação de mestrado também descreve como o NetLangEd foi testado quanto à sua utilidade e usabilidade. Para esse propósito, um experimento foi desenhado e conduzido com utilizadores finais. Os resultados serão apresentados e discutidos.

Palavras-Chave: Editor de texto, Ferramenta de anotação online, Sistema de marcação, Discurso social impróprio

CONTENTS

LIST OF FIGURES

LIST OF TABLES

1

INTRODUCTION

This document reports a Master's Thesis project in Informatics Engineering.

This first chapter introduces the project, along with the motivations, objetives, research hypothesis, methodology and document structure.

1.1 MOTIVATION

This Master Thesis fits the international project NetLang – The Language of Cyberbullying: Forms and Mechanisms of Online Prejudice and Discrimination in Annotated Comparable Corpora of Portuguese and English (PTDC/LLT-LIN/29304/2017).

In the last decades, we have witnessed an exponential growth of the Internet, more specifically in the way we communicate with each other. That brought many great things to our society from breaking distances between people to giving voice to those who didn't have one, allowing them to report the many injustices happening in the modern days.

However it also empowered anti-social behaviors like online harassment, cyberbullying and hate speech. This type of communication is usually hostile and malicious, expressing discrimination, intimidation and disapproval towards certain characteristics like sex, race, religion, ethnicity, colour, national origin, disability, or sexual orientation of a person or a group of people. The objective of this kind of speech is to injure, harass and degrade the targeted person or group in order to dehumanize them (Cohen-Almagor, 2014; Erjavec, 2012; Harvey, 2019; Neshkovska and Trajkova, 2017; Siegel, 2020; Ullmann and Tomalin, 2019).

In order to solve this problem, many big companies made available several options to address this type of speech like flagging, reporting, counter-speaking or simply censor certain words that are commonly used in this type of speech. However these strategies do not always work, generating many cases of outrage like the case of 2013 where several pages were found in Facebook with hateful content towards women like *Violently raping your friend just for laughs* and *Kicking your girlfriend in the fanny because she won't make you a sandwich*. In a matter of weeks a petition was created that aggregated 200,000 signatures and many important companies either removed their ads or threatened removing them from Facebook (Nobata et al., 2016).

In that context, the platform *NetLang* is being developed at Universidade do Minho to support researchers in the human and social sciences or psychologists collecting and making available for exploration a corpus of posts and comments that express this kind of hateful language. The posts/messages are gathered from different sources, such as Facebook, YouTube, or Web journal comment boards in various formats such as HTML, JSON, CSV, etc (Rangel Henriques et al., 2019). After the scrapping phase the documents collected are converted into a specially designed JSON format so that they be subsequently processed by some source-independent analysis tools. One of the objectives is to classify or categorize the comments according to sociolinguistic variables such as age, gender, ethnicity, nationality, social class, etc. This will pave the way for studies in language variation. Structural taggers can then be also used, providing information about the morphosyntactic organization of the texts, which may carry ideological intent.

To improve the usability of the platform it was decided to develop a tool, like an editor, that would allow the researchers to write their thoughts on specific parts of the comments creating annotations and saving them so they can continue the analysis later. After looking for tools that enable the addition of notes on a given document, several applications were found that offer this type of functionality like *Word*[1], *Adobe Reader*[2], *doccano*[3], *Hypothes.is*[4] and *Genius Web Annotator*[5], differentiating among them in the way the note is created, the method used to save/display them, the text formatting options available, etc.

After defining the context and motivation for this project, the objectives and research hypothesis are described below.

1.2 OBJECTIVES

This Master Thesis objectives are the following:

- A comprehensive study of the problem at hand by exploring related works and building a state of art about the topic. This provides the knowledge to identify the main problems and work on solutions;
- Develop an editor that allows its users to write notes related to the text of the comment lines that are being analyzed while also providing functionalities that allow their management, filter and location discovery;
- Elaborate a questionnaire to get feedback regarding the developed editor followed by an analysis and a discussion of the results obtained from them.

1 Available at https://www.groovypost.com/howto/annotate-in-word/, accessed in December 2020
2 Available at https://helpx.adobe.com/acrobat/using/commenting-pdfs.html, accessed in December 2020
3 Available at https://doccano.github.io/doccano/, accessed in December 2020
4 Available at https://web.hypothes.is/help/annotation-basics/, accessed in December 2020
5 Available at https://genius.com/web-annotator, accessed in December 2020

As final result, it is expected to have an editor that has a simple and easy to use interface. The editor should offer to the end user the possibility to keep in the same platform the corpus under study and the notes created along the study, ensuring better organization and accessibility.

1.3 Research Hypothesis

With this Master's work it is intended to prove that a tool that allows the annotation of texts will help analysts to achieve better results, this being achieved through better accessibility and organization of their comments.

1.4 Research Approach

To accomplish this Master Thesis, an iterative methodology based on literature revision, solution proposal, implementation and testing are followed.

To carry out this methodological approach, the work was structured and organized in the following steps:

- Bibliographic study to deeply understand the state of the art in the areas of Online Socially Unacceptable Discourse (SUD) and editors that have annotation components. The references Cohen-Almagor (2014); Erjavec (2012); Harvey (2019); Neshkovska and Trajkova (2017); Nobata et al. (2016); Siegel (2020); Ullmann and Tomalin (2019); OVSIANNIKOV et al. (1999); Kawase et al. (2009); Jindia and Chawla (2013); Fariza et al. (2013); Glover et al. (2007) were used as a starting point for this study;
- Requirement elicitation and definition of the features that the editor should have;
- Development of the editor and integrate it with the existing *NetLang* platform;
- Develop a questionnaire in order to obtain feedback regarding the editor;
- Analysis and discussion of the results obtained.

1.5 Document Structure

The structure of this dissertation is composed of six chapters and two appendixes with complementary materials.

Chapter 2 will cover topics such as the history of annotations, how people annotate on paper and on the web, annotation functionality and types, advantages and disadvantages

of digital annotations and an analysis of the existing solutions. To conclude this chapter, a summary of the subjects covered in it will be presented.

In Chapter 3 the system requirements will be listed and the system architecture will also be discussed and sketched using block diagrams.

Chapter 4 describes how the editor was developed, showing some screenshots of the final result.

Through Chapter 5 it will be possible to understand the process that was designed to test the developed editor.

Finally, in Chapter 6 an analysis of the results obtained will be made and conclusions will also be drawn regarding the work done. A list of suggestions to improve the editor, as future work, is also included.

2

STATE OF ART

In this chapter, the state of the art regarding annotation is portrayed. Its initial part is used to describe the evolution of the annotations over the years. Then, studies conducted by several researchers are analyzed in order to understand the behavior of readers when annotating on paper and online. Based on these studies and other articles, the purposes and types of annotations are described in general followed by the listing of the advantages and disadvantages of digital annotation compared to physical annotation. After that, several tools that allow the annotation of digital texts are listed for which their functionalities are described followed by a conclusion regarding their compatibility with the objectives that are intended to be achieved in this Master's work. To conclude this chapter, a summary is presented containing the most important lessons learned along the literature review. Also a comparison table is shown to sum up the most relevant features offered by the tools studied.

2.1 HISTORY OF ANNOTATIONS

The usage of text annotations became a prominent activity around 1000 AD in Talmudic commentaries and Arabic rhetorics treaties[1].

It was then used in the Medieval ages by scribes who took advantage of the interlinear spaces and margins of the manuscripts to discuss, critique and learn from annotations created by previous readers who also read the same manuscript. There were also situations where at the time the manuscripts were being copied, their annotations were included in the copy (Wolfe and Neuwirth, 2001).

However, the emergence of the printing press has made this use of annotations obsolete due to having facilitated the circulation of information and the ability to purchase individual copies of text (Wolfe and Neuwirth, 2001).

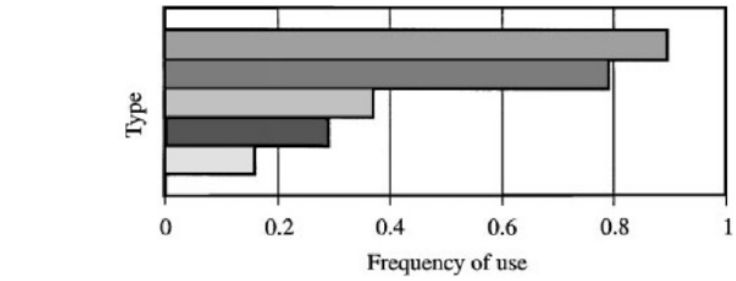
Nowadays annotation is an activity that is mostly done in private corresponding to the reader's interaction with the text being read. Computer-based technologies also provide many solutions for both individual and shared annotations, allowing to apply this method to online and offline digital documents (Wolfe and Neuwirth, 2001).

1 Available at https://en.wikipedia.org/wiki/Text_annotation, accessed in December 2020

2.2 How people annotate on paper

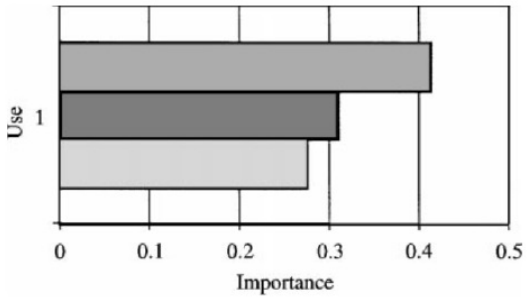
In this section, the behavior of readers when annotating on paper is studied. For this purpose, studies carried out by several articles are used in which conclusions were drawn on this topic.

The first study analyzed was developed by OVSIANNIKOV et al. (1999) in which the respondents were asked several questions, one of which was related to how they annotate the paper and another one was related to how the annotations were used by the readers. Through the responses obtained, the graphics displayed in Figures 1 and 2 were created.



Annotation types. Mark up; Write on margins; Write at the top; Write separately; Write between lines.

Figure 1: Type of annotations (OVSIANNIKOV et al., 1999).



Annotation applications. Remember; Think; Clarify.

Figure 2: Use of annotations (OVSIANNIKOV et al., 1999).

As can be seen in Figure 1 the highlight (Mark up) is the type of annotation that is predominantly applied and in terms of use the most popular was for remembering (Remember) as can be seen in the Figure 2. It is possible to state that there is a correlation between the purpose of remembering and the type of annotation that involves highlighting, which makes sense since highlights are usually used to mark the main ideas of the text so that they can be found more easily in the future in case the reader has forgotten something. In addition to this, the color of the highlights also proved to be very useful as it facilitates detection and

may contain additional meanings. Other identified uses that had an importance close to the purpose of remembering were to think and to clarify. The purpose of thinking is related to the fact that readers have to think when they make observations, questions, comments regarding the notes made by others, etc. Finally, the use of clarification refers to situations where the reader rewrites a sentence that he/she has read using his/her own words, normally not containing additional information, in order to facilitate its understanding in a future reading.

The article (Blustein et al., 2011) also analyzed documents that were annotated by participants in their study through which it was possible to observe other types of annotations as shown in Figure 3.

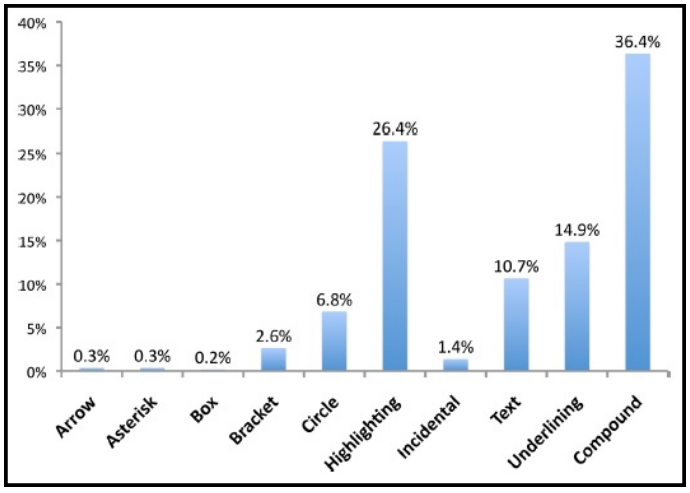


Figure 3: Distribution of types of annotation (Blustein et al., 2011).

As can be seen in Figure 3, the second most used type of annotation was *Highlighting* and the most used was *Compound*, which is a type of annotation that only this study takes into account, consisting of a combination of the other types of annotations so far identified. In this way, the study also sought to explore the distribution of the components of the *Compound* annotation type, resulting in the graphic visible in Figure 4.

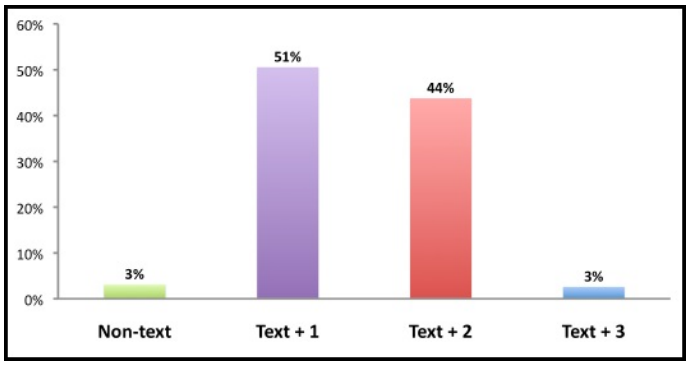


Figure 4: Distribution of the components of the *Compound* annotation type (Blustein et al., 2011).

As can be seen in Figure 4, the most frequent behavior when annotating with a combination of various types of annotation is to use text plus one type of annotation and text plus two types of annotation said from the least to the most frequent case.

This is an interesting observation because it shows that most of the time, when readers use textual annotations, they tend to use non-textual annotations to support them. Although not mentioned in the article, this case may be related to the need of linking the textual annotation to the part of the text to which it refers through one or more non-textual annotations.

Finally, the article (Kawase et al., 2009) also sought to explore the reader's behavior when annotating on paper, where in this case, the participants' research papers and articles were used as study material. After analyzing them, the table displayed in Figure 5 was created.

Annotation types		
Highlighting/Mark sections headings	153	8.6%
Highlighting/Mark text	1297	73%
Problem solving	2	0.1%
General notes (Notes in the margins)	326	18.3%

Figure 5: Average of each type of annotation per page (Kawase et al., 2009).

As can be see in Figure 5, most of the annotations used consist of highlighting parts of the text (Highlighting/Mark text), where the reason for this attitude could be related to the fact that the readers do not want to be constantly changing tasks preferring to focus on reading the text and simply mark the parts they consider to be important, since it is a quick method of execution compared to the others that require writing information. In addition to the annotations made in the document itself, pots-its attached to the document were also used and in other cases the annotations were written on a separate sheet or even in another copy of the respective document. Regarding the highlights made, the use of different colors was also observed in order to represent different levels of importance. As for the ways of signaling important parts of the text, it was noted that several participants created their own mechanisms, each one having its own meaning. A final observation is related to

the situations where the annotations were shared, where in this case the annotations were made in a more careful way so that the other reader could easily interpret them unlike the case of private notes, which often do not contain complete sentences or are made up of keywords. Another analysis that was carried out sought to understand whether the purpose of reading had any impact on the way readers choose to annotate texts. In this case, the same participants mentioned before were asked to say the purpose of the reading for each of their research papers and articles. After collecting all the answers, it was possible to draw the conclusions shown in Figure 6.

	Writing	Learning	Review	Other
Articles	31	23	9	3
Articles annotated	28	16	7	3
Annotations/Page	2.36	4.7	1.11	6.3
Annotation types				
Highlighting/Mark sections headings	10.5%	7.5%	9.4%	4.8%
Highlighting/Mark text	66.0%	82.9%	40.6%	72.2%
Problem solving	0.1%	-	0.9%	-
General notes (Notes in the margins)	23.3%	9.6%	49.1%	23.0%

Figure 6: Results by reading goal (Kawase et al., 2009).

Through the table in Figure 6 it is possible to identify the following three objectives for which a brief explanation of their purpose is given followed by an analysis of the behavior of the reader in each one of them.

- **Reading for learning:** This objective is based on the act of reading a text to learn something so that it can be applied in another situation or simply to gain more knowledge. After analyzing the data, it was possible to verify that a large number of annotations per page were related to the use of highlights over parts of the text (Highlighting/Mark text), concluding that their functionality in this situation was to support the memorization process.

- **Reviewing:** This case consists of reading a text with the sole purpose of providing feedback to its author. Based on the data obtained it is possible to observe a number of notes (General notes) slightly higher than highlights over parts of the text (Highlighting/Mark text), demonstrating that in this case the marks were not enough to express the opinion of the readers, having to resort to comments/responses to express themselves. Since in this case the objective is to carry out a revision of a text, it is very likely that many of these written comments will later be included in the review. In addition, the decrease in the use of highlight shows that in this case readers are not concerned with marking parts of the text to draw attention in future readings.

- **Reading for writing:** This case, which was not discussed in the article, consists of extracting information and ideas from texts to support later writing. Despite this, through the graphics provided, it is possible to verify that there is a greater use of highlights over parts of the text (Highlighting/Mark text) than written notes (General notes), demonstrating that in this case users also choose to mark important parts of the text to later be easier to locate and base themselves on them to write. However, it is possible to verify the existence of a significant number of written notes (General notes) being linked to the fact that in this situation it is common for the reader to write his interpretation of what he is reading to facilitate the later process.

2.3 How people annotate on web

In this section, it is explored how readers annotate in the web context. To do so, the experience described in (Kawase et al., 2009) is analyzed first in which the behavior of the participants when using the web annotation tool *SpreadCrumbs*(Figure 7) is studied. This tool allows its users to add post-its on the web pages and shared them with other users.

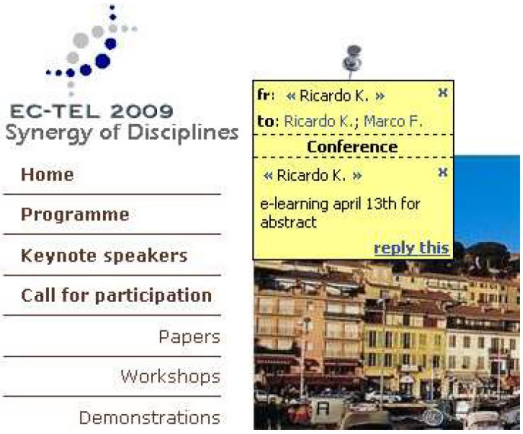


Figure 7: Web annotation tool *SpreadCrumbs* (Kawase et al., 2009).

The first phase of the experience consists of a set of questions to which participants would have to find the respective answers online. Upon finding the answers, the participants would have to annotate, with the *SpreadCrumbs* tool, on the respective source the answer that was found in it. After the experiment was completed, it was possible to observe that the participants were careful to placed the annotations near the place where the answers are present. Another observation is related to the fact that the majority of the participants wrote keywords of the questions in their annotations instead of the answers, having as a possible explanation the fact that these had the purpose of facilitating the rediscovery of the annotations made since the participants would have to re-find them in the second phase of

the experience. This way of writing annotations was also verified in an analysis carried out on the annotations made by the users that were not involved in the experiment, through which it was verified that the private annotations are usually shorter and more ambiguous than those that are shared, which are more developed and explicit as can be seen in Figure 8.

Personal	Shared
"Conference Deadline: October 29"	*"All artists are from Sweden, I think, and do Jazz music (quite soft) but nice..."*
"Flat 64m 2 rooms windthorststr. 8"	*"Let me know if there's anything else to be done."*
"TO DO!"	

Figure 8: Examples of private and shared annotations made online (Kawase et al., 2009).

Based on this evidence, it is possible to conclude that readers are not very concerned with the quality of the content of private annotations, but rather with what they represent for themselves unlike those that are shared since in this case it is evident the concern to be more explicit so that other readers will have no trouble understanding their content.

Another article that analyzes the behavior of readers when annotating on the web is Fariza et al. (2013). In this case, the participants used an annotation tool that was integrated in an online reading system. In this way, participants were assigned several texts that should be analyzed individually and together in some of them, as can be seen in the Figure 9.

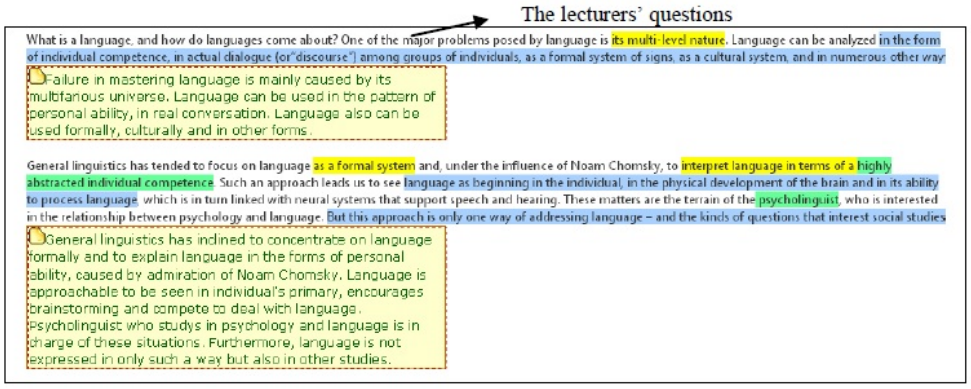


Figure 9: Example of the annotation tool usage (Fariza et al., 2013).

From the results of the experiment, it can be concluded that the three most popular uses of the annotations were to highlight parts of the text that were not understood, remember certain points through highlights and highlight important parts of the text, said from the most used to the least. This conclusion shows once again the importance of highlighting in the annotation process. The other three identified uses were for writing comments directly in the text concerning the reader's interpretation of what he/she read or a summary of it and sharing their understanding with other colleagues in order to build knowledge together.

These two having the same popularity value, followed by the use related to the suggestion of relevant websites for the text, this being the least popular. The purposes of the annotations made by the participants demonstrates not only that they were making them for their own use and for sharing, but also that they were involved in active reading. A final observation is related to the fact that the use of different colors in the highlight, which in this case are limited and already have a pre-assigned label, helped the participants to better structure their notes.

Finally, a last study reported in (Lu and Deng, 2013) also seeks to understand how readers annotate on the web. Another objective of this study is to compare the annotations made by a high-performance class (HPC) and an ordinary-performance class (OPC), however this will not be commented since it does not matter for the case of this thesis. In this way, the results achieved by the study will be interpreted in a general way.

A fact that makes this study more interesting is the use of the online annotation tool *Diigo* since it is one of the tools that will be analyzed in Section 2.7.

The experience consisted of annotating: articles chosen by the students themselves where they could annotate freely; and also articles selected by the teachers in which the students have to annotate according to some stipulated rules. After that, students shared their work in the *Diigo* group space where they could comment and review their colleagues' work. As an example of the work done by the students, Figure 10 is shown illustrating some of the annotations made.

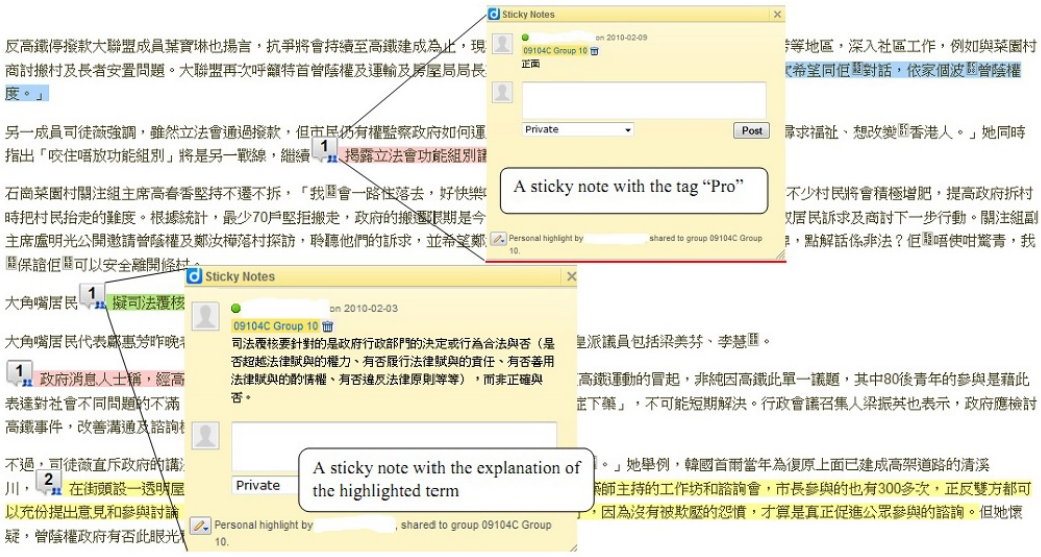


Figure 10: Example of some annotations made by one of the students (Lu and Deng, 2013).

Through the data collected from the experience, it was possible to identify various purposes for the sticky notes used by the students when annotating the articles. These had

the function of providing definitions, exposing the reader's position on issues, providing explanations and exposing the reader's opinions and conclusions.

Regarding the comments exchanged by students, these could be responses to highlights and sticky-notes made by other students or comments regarding the way of taking notes from other students. As for its content it could be an opinion, additional information, agreement or questions.

After analyzing the results, it was possible to observe several facts. One of them is related to the predominant use of highlights, confirming once again this trend that was also verified in the other studies already mentioned. Another observation was that the students annotated less in the articles that were chosen by them compared to those that were selected by the teachers. As for the sticky-notes made by the students, the two most frequent purposes were to provide definitions and additional information followed by providing explanations. Finally, regarding the interactions between students, the vast majority of them were responses to the highlights and sticky-notes made by other students which consisted mostly of opinions related to the issue or other comments.

2.4 Functionality of annotations

There are several benefits that are obtained through private and public annotations. According to the studies described in Sections 2.2 and 2.3 and the articles by Jindia and Chawla (2013), Gao (2013), Marshall (1997), O'Hara and Sellen (1997), O'Hara et al. (1998) and Agosti and Ferro (2003) it is possible to identify the following ones:

- Facilitate the current or future reading process with the clarification and interpretation of the text read;
- Facilitate the re-use of the key contents of the text for future activities;
- Understand the insights of another reader;
- Provide feedback to the text writer or other readers;
- Help with the memorization and recall process;
- Draw attention to certain parts of the text that are considered to be important for future reference or reading;
- Correct a specific part of the text;

2.5 Types of annotations

Taking into account the conclusions presented in Sections 2.2 and 2.3 and the articles by Jindia and Chawla (2013), Gao (2013), Marshall (1997), O'Hara et al. (1998) and Chen et al. (2019) it is possible to verify that annotations can take different forms, where some of them assume a textual representation and others consist of graphic effects. That being said, the types of annotations that have been identified are described below.

- **Mark:** The method of marking an important word/phrase through visual effects. These can be highlights, underlines, strikeouts, figures, etc.
- **Paraphrase:** It consists of reproducing in a simpler and more accessible way the central ideas of the original text, without changing its meaning.
- **Comment:** This type is based on the formulation of comments to specific parts of the text that are directed to the writer or self-directed. These may be of agreement/disagreement, questions, responses, connection to ideas from other texts, personal experience, adding explanation, etc.

Note that both the *Paraphrase* and *Comment* type can be combined with the *Mark* type in order to be able to contextualize them in the text.

2.6 Advantages and disadvantages of digital annotations

In this section the advantages and disadvantages of using digital annotations in comparison to paper-based annotations are discussed. This topic is of uttermost importance because one of the main reasons that led to the development of this work was the fact that the users of the platform *NetLang* were using paper to write their comments/analysis of the text being read. In addition to the existing knowledge regarding the properties that technologies offer, the discussions sustained in the previous sections, and the studies by Glover et al. (2007), Marshall (1997), O'Hara and Sellen (1997) and Schilit et al. (1999) it is possible to identify the pros and cons listed below.

Advantages

The advantages that have been identified are as follows:

- Better organization due to not having the notes spread over several sheets;
- Adding annotations wont damage the original text;
- Easier to change the content of annotations;

- Allows the removal of annotations without leaving marks in the document;
- It allows notes to be written with more extensive content without having to worry about the space they will occupy;
- Easier to locate annotations on the document through mechanisms;

Disadvantages
The disadvantages identified were the following ones:

- It is not possible to directly manipulate the document requiring to follow a specific process that is more complex than simply using the pencil to write and draw on the paper;
- The forms that digital annotations can take are limited to those offered by the tool used;
- Requires the user to learn how to use the tool;

2.7 EXISTING SOLUTIONS

In this section, some solutions that allow text annotation are discussed. This will consist of a brief description of the application followed by an analysis of the components it offers. It is important to note that, with the exception of the *Word* and *Adobe Reader* tools, all the other ones are completely analyzed through the available documentation. In the *Adobe Reader* case the documentation is only used for a small part of its functionalities.

2.7.1 *doccano*

doccano is an open source tool offers several annotation methods for *text classification*, *sequence labeling* and *sequence to sequence* that allow its users to label data in order to create datasets[2]. Bearing in mind the objectives of this thesis, only the first two will be discussed below.

The *text classification* functionality is used to classify texts with one or more labels from a set of these as shown in Figure 11.

2 Available at https://doccano.github.io/doccano/, accessed in January 2021

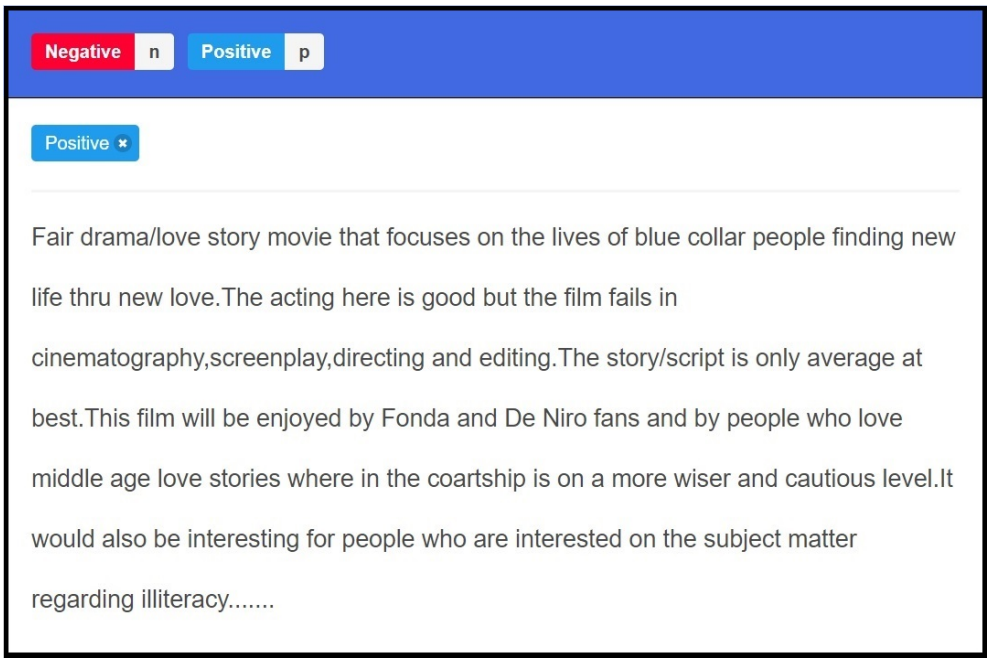


Figure 11: Example of the *text classification* functionality[3].

As can be seen from the Figure 11, the text can be classified by the *Positive* and *Negative* tags which in this case serve to reflect the user's feeling towards it. In the case of this functionality, the addition of tags is done by pressing on any of the tags present in the blue section or by clicking on the keyboard key corresponding to the letter that is in front of the tag name also present in the blue area, being then included in the area below the blue area which will allow the removal of these just by pressing the "x" in front of the tag name.

Regarding *sequence labeling*, its functionality is to allow the annotation of parts of the text with labels from a set of these as shown by the Figure 12.

3 Available at http://doccano.herokuapp.com/demo/text-classification/, accessed in January 2021

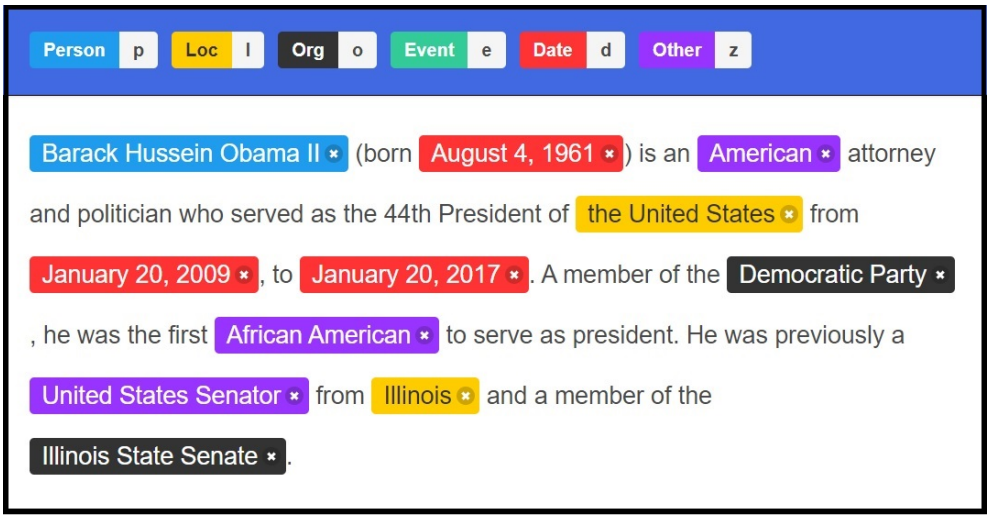


Figure 12: Example of the *sequence labeling* functionality[4].

Through the Figure 12 it is possible to observe several parts of the text being classified by the labels present in the label set. In this functionality, the addition of tags requires the user to select the part of the text that he wants to annotate being then the same as the addition process in the case of the *text classification* and its removal is done by pressing the "x" that is in front of the annotated part of the text.

Note that in both cases it is possible to import and export the work done and also to create the tags which can be customized in different ways such as their background color, the font color, etc. Another functionality that this tool provides is collaborative work in which several users can annotate the same document together.

Taking into account all the functionalities mentioned above and the purpose of this thesis, the form that this tool allows for annotation would be quite restrictive since it is based mainly on labels, not allowing analysts to freely write their thoughts regarding specific parts of the text. However, being an effective and easy-to-use tool, it would be very useful for cases of automatic detection of hate speech through machine learning methods for which it is necessary to provide datasets that contain their content properly labeled. That being said, it is possible to conclude that this tool is not compatible with the objectives that are intended to be achieved in this thesis.

4 Available at http://doccano.herokuapp.com/demo/named-entity-recognition/, accessed in January 2021

2.7.2 *Word*

Word is a graphical word processor created by *Microsoft* whose purpose is to allow its users to type and save documents[5]. Although this tool offers several features, it will only be discussed the one that allows its users to annotate texts since this is the focus of this thesis.

The way that this tool handles the annotation of texts is very simple, starting by requiring the user to select the part of the text he wants to annotate. Then the user can create a comment in several ways, such as through the pop-up that appears after finishing the selection as shown in Figure 13, through the "Review" section that is visible in Figure 14, through the "Comments" button shown in Figure 15 and through the options that result from right-clicking on the selection.

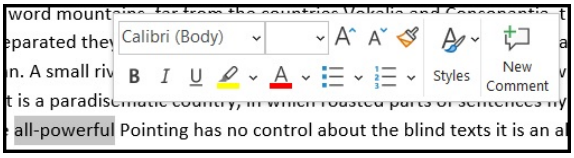


Figure 13: Pop-up that appears after the selection.

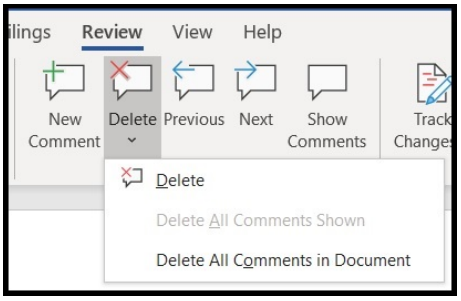


Figure 14: "Comments" subsection located in the "Review" section.

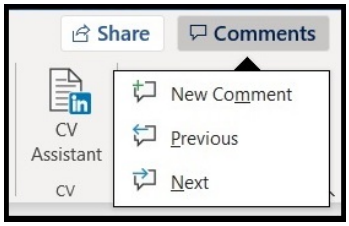
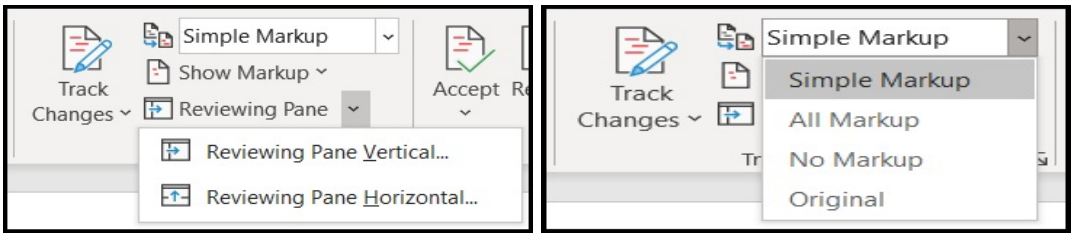


Figure 15: "Comments" button.

After the option "New Comment" has been picked in any of the options described above, a text area will be displayed to insert the comment which can be formatted in many ways such as choosing the font, choosing the font color, inserting lists, etc. These can then be viewed

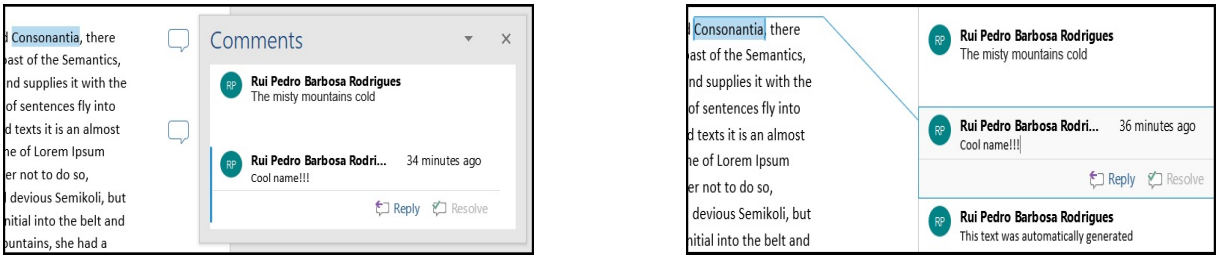
5 Available at https://simple.wikipedia.org/wiki/Microsoft_Word, accessed in January 2021

in different ways through the comment display options that are present in the subsection "Tracking" shown in Figure 16, which can be "Simple Markup" with and without the option "Show Comments" of Figure 14 activated, "All Markup" and the options "Reviewing Pane Vertical" and "Reviewing Pane Horizontal", all of which effects can be seen in Figure 17a, Figure 17b, Figure 18, Figure 19a and Figure 19b respectively.



(a) (b)

Figure 16: "Tracking" subsection located in the "Review" section.



(a) With the "Show Comments" option disabled.

(b) With the "Show Comments" option enabled.

Figure 17: "Simple Markup" option.

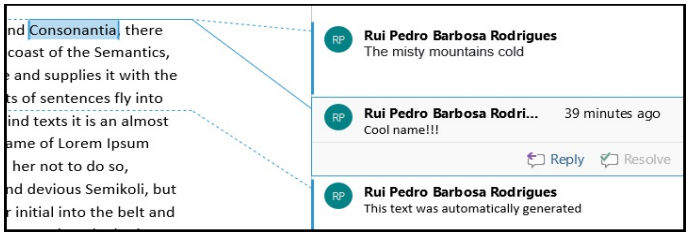
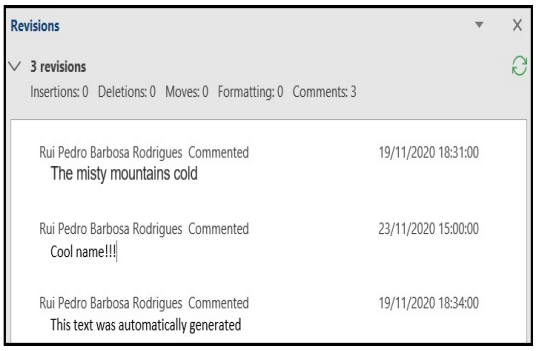


Figure 18: "All Markup" option.



(a) "Reviewing Pane Horizontal" option.

(b) "Reviewing Pane Vertical" option.

Figure 19: "Reviewing Pane" button options.

Another way to view the comments made on the parts of the text of the document is through the pop-up that can be seen in the Figure 20 which is displayed when the mouse cursor is placed on top of them for a few moments.

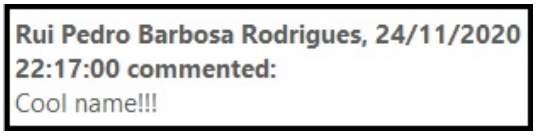


Figure 20: Pop-up of a highlight.

The *Word* tool also allows its users to search for annotations whose comments contain a certain word/phrase. For that, the user will have to use one of three possible ways, these being clicking on the magnifying glass, clicking on the "Find" sub-option or by pressing *Ctrl+F*, being possible to see the first two options in Figure 21.

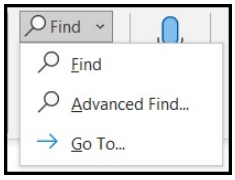


Figure 21: "Find" option and its sub-options.

After executing one of the three methods mentioned above, the navigation menu that is visible in Figure 22 will be displayed. In this case, after the user writes the content he wants to search, the option "Results" should be selected through which he will be able to find, at the end of the list, the comments that contain the searched content.

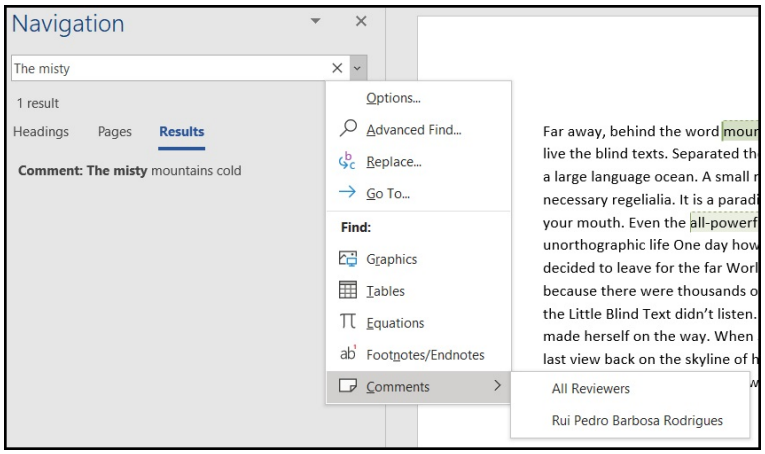


Figure 22: Normal search menu.

Through Figure 22, its also possible to see the option "Options..." which will display the window visible in Figure 23, through which the user can manage some aspects of the search.

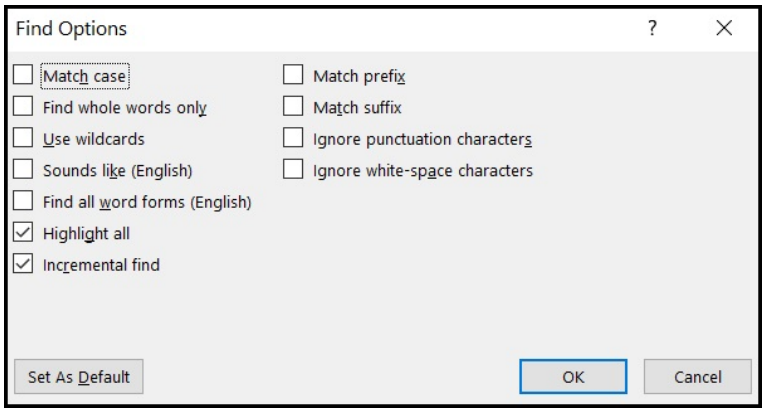


Figure 23: Search parameters.

Finally, the user can perform a more advanced search that can be done through the "Advanced Find..." option that is visible in Figures 21 and 22. The result of both actions will be the window shown in Figure 24 through which the user will have more options to improve his search, including some options that are present in the window visible in the Figure 23 that was already addressed. Note that the user will have to choose the option "Comments" in the field "Find In", that is visible in Figure 24, in order to search only in the comments of the annotations.

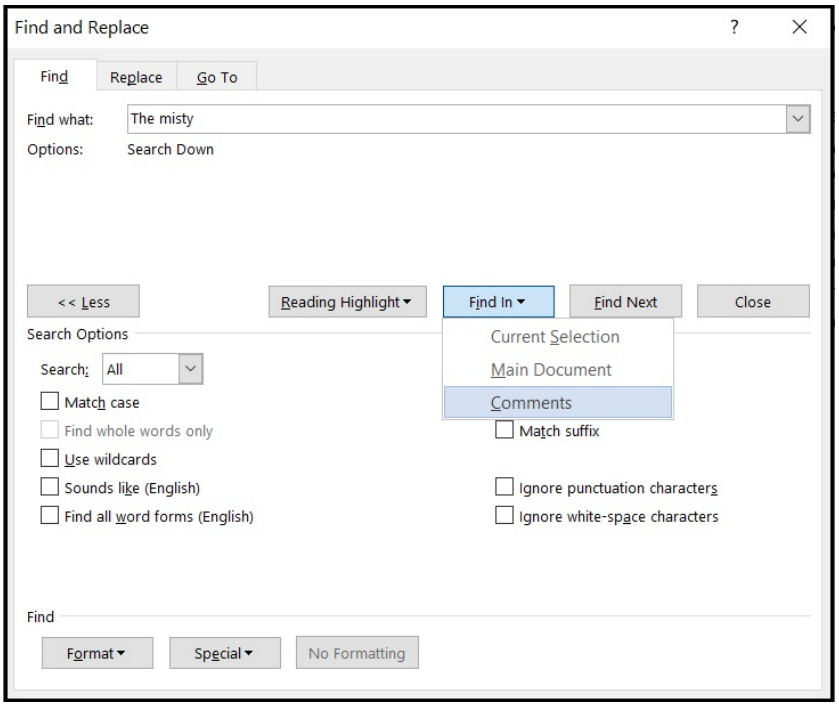


Figure 24: "Advanced Find..." option result.

Regarding the comment displays, these can also be very useful to help users locate their annotations more easily. In the case of the Figure 17a, the parts of the text that are annotated are marked through a "talking box" that has two functions, one of which highlights all the annotations that it covers in case the user hovers the mouse over it and the other one exposes all the comments of the annotations that it covers in case it is clicked, where if any of these are selected it will highlight the part of the text that is commenting on. Through the case of Figure 17b, the user will be able to see on the side all the comments made in the document, where clicking on one of them will link it through a line to the part of the text that it refers to which will also be highlighted. In the case of the Figure 18 the user will be able to see on the side all comments made in the document in which case all of them will be linked through a line to the part of the text to which they refer and if one of then is selected it will only be differentiated from the rest through small visual differences. Finally, in the cases of the Figure 19 the user will also be able to see all the comments made in the document but when these are clicked it will only highlight the part of the text that they refer to.

Word also provides some ways to make it easier to navigate through the comments present in the documents. One of them is made in the comment display options "Reviewing Pane Horizontal" and "Reviewing Pane Vertical" shown in the Figure 19a and Figure 19b respectfully, in which clicking on a comment causes the document to move to the position where the part of the text to which it refers is located. Another one is through the "Previous" and "Next" buttons shown in Figure 14 and Figure 15, having the same effect as the method

mentioned before. The last one is through the option "Go To..." that is visible in Figures 21 and 22, which were discussed previously, or by pressing *Ctrl+G*, resulting in the window present in Figure 25. In this window the user must select the option "Comment" to be able to circulate through all annotations or just through the annotations of a specific author, with the same effect that was explained in the first method that was mentioned before, being all this visible in the Figure 25.

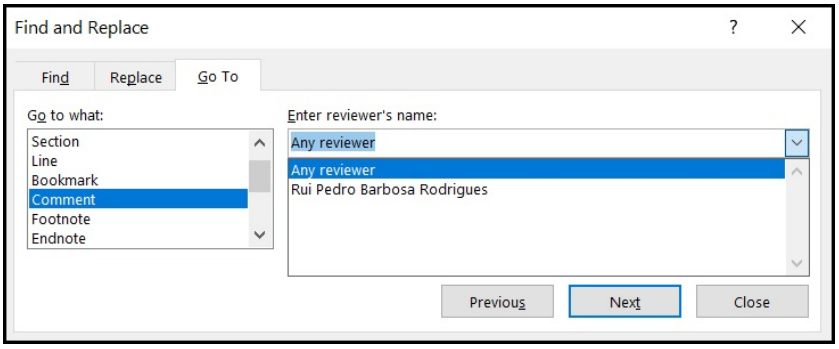


Figure 25: Result of the "Go To..." option.

Regarding editing comments, users have two ways to do this. The first is done through the comment box that is visible in Figure 17a, Figure 17b, Figure 18, Figure 19a and Figure 19b. The other one is through the options that arise from right-clicking on the part of the text to which the comment to edit refers.

Another important function is to eliminate comments, which can be done in several ways that aim at the same objective. One of these requires the user to select the part of the text to which the comment to be removed refers to or the comment box, like the ones visible in Figures 17a, 17b, 18, 19a and 19b, that contains the comment to be removed, after which can be deleted through the delete options present in the "Review" section that is visible in Figure 14. The other way is through the options that arise from right-clicking on the part of the text to which the comment to be removed refers to or on the comment box that contains the comment to be removed. There is also the option to remove all comments present in a document at once through one of the removal options present in the "Review" section, as can be seen in Figure 14.

Although it is not an objective of the work to be developed in this thesis, it is important to mention that this tool also provides mechanisms that allow collaborative work in which it is possible for multiple users to comment on the same document and respond to the comments of others, having two ways to perform this last case. One of them is through the options that arise from right-clicking on the part of the text to which the comment to be replied refers to or on the comment box, like the ones visible in Figures 17a, 17b, 18, 19a and 19b, that contains the comment to be replied. The other one is through the "Reply" button present in the comment boxes as can be seen through the Figures 17a, 17b and 18. Another feature

that *Word* provides and that is useful in collaborative work situations is the possibility of identifying the author of the highlight in the text as shown in Figure 26, through which it is also possible to see the option that activates this functionality. Note that this option only works if the comment display option "All markup" selected, being restricted to the comment displays "Reviewing Vertical Pane" and "Reviewing Horizontal Pane", that are shown in Figures 19a and 19b respectfully, to write the comments.

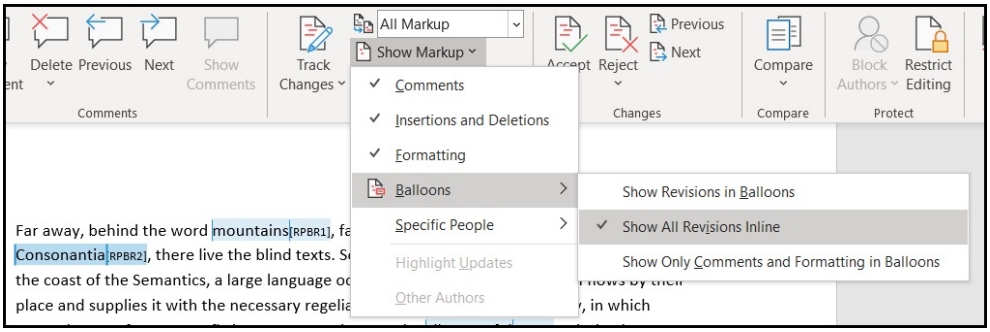


Figure 26: Functionality to identify the authors of the highlights.

Word also allows the customization of the highlights and comments. To do this, the user must follow the process that can be seen in Figure 27, which consists of clicking on the icon found in the lower right corner of the "Tracking" subsection of the "Review" section followed by selecting the advanced options which will then open the customization window that can be seen in Figure 27, through which it is also possible to see the various customizations that can be made. However, it is important to note that all highlights will have the same color, which goes against the approach of most annotation tools since they allow the user to chose the color of each highlight.

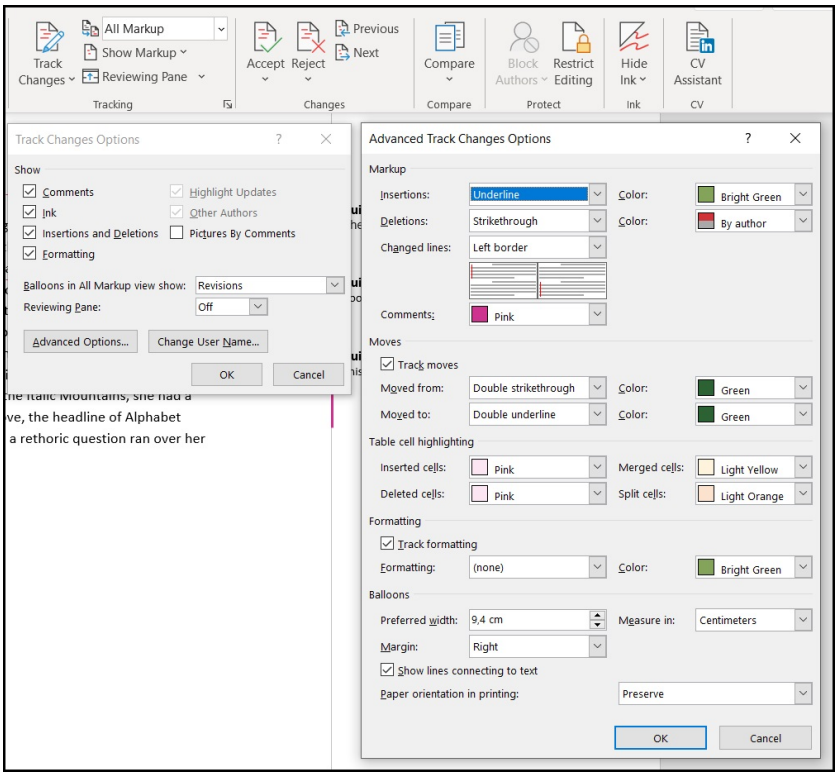


Figure 27: Options for customizing comments.

To conclude the analysis of the *Word* tool, it is important to mention that although there are several methods that allow the exportation of the work done, none of these allows it to be imported later in order to be continued.

After having analyzed the various features that *Word* offers, it is possible to state that it is a very good tool for annotating texts and managing annotations in a simple and effective way with the additional advantage of being a tool that is already widely used. Having this said, it is possible to conclude that *Word* is not compatible with what is intended for the work to be developed in this thesis since it is not a web based tool.

2.7.3 *Adobe Reader*

Adobe Reader is a software that allows its user to view, create, manipulate, manage and print files in the PDF format[6]. Of all the features it offers, only those related to the annotation process will be discussed since this is the objective to be achieved in this thesis.

When opening the tool, there are several ways to start the annotation process right away. One of them requires the user to right-click on the document and activate the "Select Tool" option. After that, the user can select the part of the text that he wants to annotate. After

6 Available at https://en.wikipedia.org/wiki/Adobe_Acrobat, accessed in January 2021

completing the selection the pop-up that is present in Figure 28 will be displayed, through which the user can perform various actions such as "Highlight Text", "Underline Text", "Strikethrough Text", "Copy Text" and "Edit Text & Images"[7].

Figure 28: Actions after text selection[7].

Another type of selection allowed is the selection of images where after its completion the pop-up shown in Figure 29 will be displayed, providing several actions such as "Add Sticky Note", "Highlight Text", "Edit Text & Images" and "Copy Image"[7].

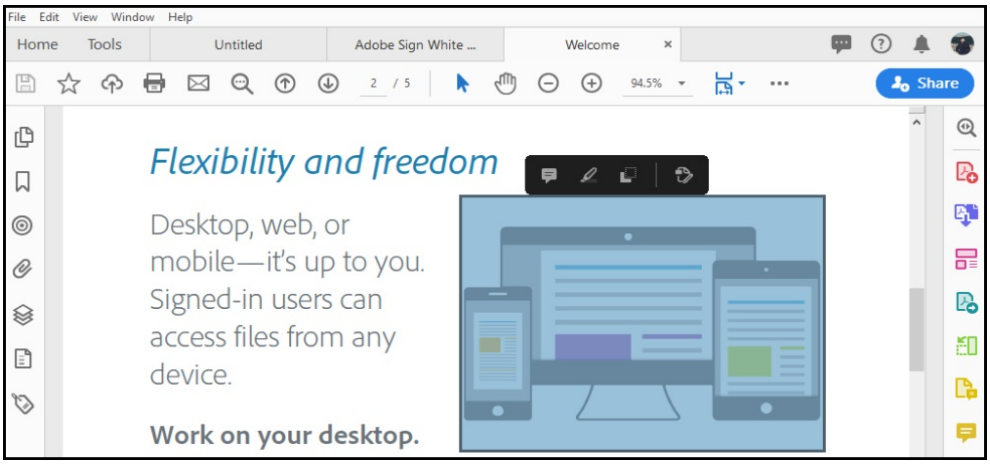


Figure 29: Actions after image selection[7].

In case options "Highlight Text", "Underline Text" and "Strikethrough Text" are used the user can also click the annotation to display the pop-up visible in Figure 30, through which it will be possible to perform the actions "Add Note" and "Delete"[7].

7 Available at https://helpx.adobe.com/sea/acrobat/using/commenting-pdfs.html, accessed in January 2021
7 Available at https://helpx.adobe.com/sea/acrobat/using/commenting-pdfs.html, accessed in January 2021

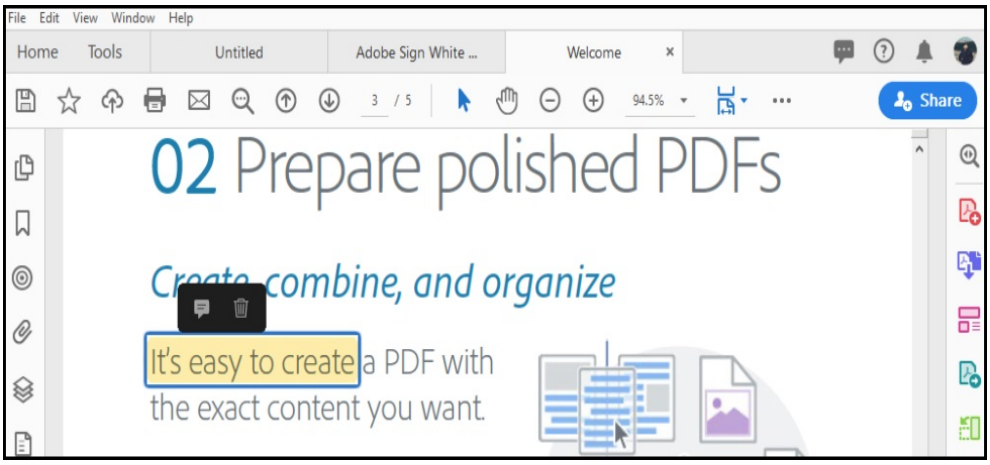


Figure 30: Actions after selecting certain types of annotations[7].

Another quick solution would be through the "Add Sticky Note" option present in the options window obtained from performing a right-click on the document.

A last quick way would be through the two options present in the initial toolbar that can be seen in the Figure 31, one of which is the "Add sticky note" and the other one is the "Highlight text".

Figure 31: Quick annotation methods present in the initial toolbar.

It should be noted that all the options mentioned above that affect the text of the document will be explored in more detail below, some of which with different names.

That being said, in order to use all the annotation features provided by *Adobe Reader*, the user will have to select the "Comment" tab found in the sidebar on the right side as shown in Figure 32.

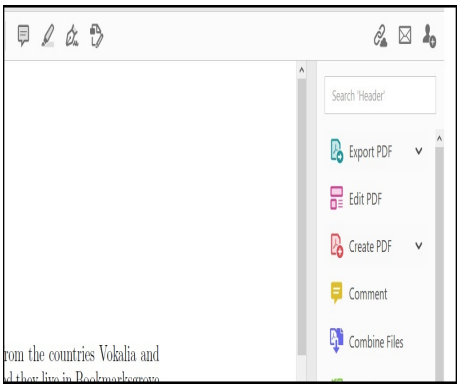


Figure 32: Sidebar with the "Comment" tab.

After that, a toolbar will appear with all the annotation options and auxiliary options, as can be seen in the Figure 33.

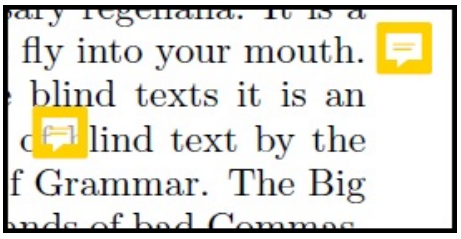
Figure 33: Toolbar with the annotation options and auxiliary options.

Below, each of the options that are present in the Figure 33 will be explored.

- **Add sticky note**

  This option is represented in the toolbar by the icon in the Figure 34a and allows the user to place sticky notes anywhere in the document as shown in the Figure 34b.

(a) In the Figure 33.



(b) In the document.

Figure 34: Graphical representation of the "Add sticky note" option.

The appearance properties of the annotations shown in the Figure 34b can be edited through the options visible in the Figure 35 that arise from right clicking on them or on the respective comments that are present in the comments list that is visible in Figure 74.

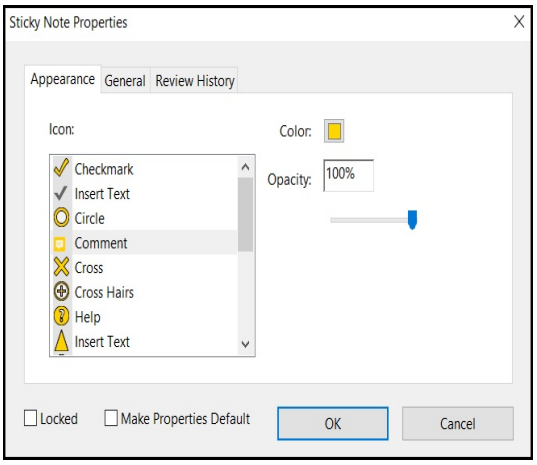
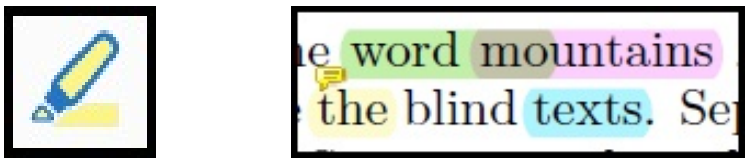


Figure 35: Properties of the "Add sticky note" option in the document.

- **Highlight text**

  This type of annotation is present in the toolbar with the visual shown in the Figure 36a and allows the user to highlight parts of the text as shown in the Figure 36b.



(a) In the Figure 33. (b) In the document.

Figure 36: Graphical representation of the "Highlight text" option.

The appearance properties of the annotations shown in the Figure 36b can be edited through the options visible in the Figure 37 that arise from right clicking on them or on the respective comments that are present in the comments list that is visible in Figure 74.

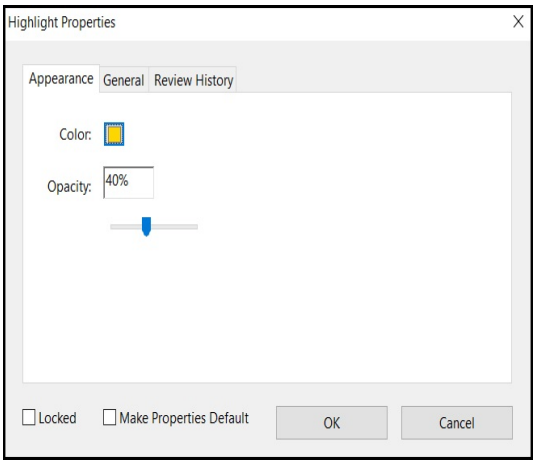
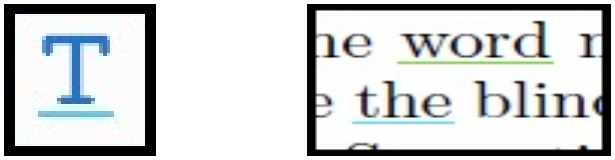


Figure 37: Properties of the "Highlight text" option in the document.

Another interesting feature to mention is the effect that occurs when these highlight annotations overlap. In this case, the highlight color of the part of the text were the overlap occurred will be different from the colors of the annotations that were made over it, as can be seen in the Figure 36b. This feature will help the user to detect more easily the situations where the same part of the text is being annotated by more than one annotation of the highlight type.

- **Underline text**

  This option is accessible on the toolbar through the icon shown in Figure 38a and allows the user to underline parts of the text as shown in the Figure 38b.



(a) In the Figure 33. (b) In the document.

Figure 38: Graphical representation of the "Underline text" option.

The appearance properties of the annotations shown in the Figure 38b can be edited through the options visible in the Figure 39 that arise from right clicking on them or on the respective comments that are present in the comments list that is visible in Figure 74.

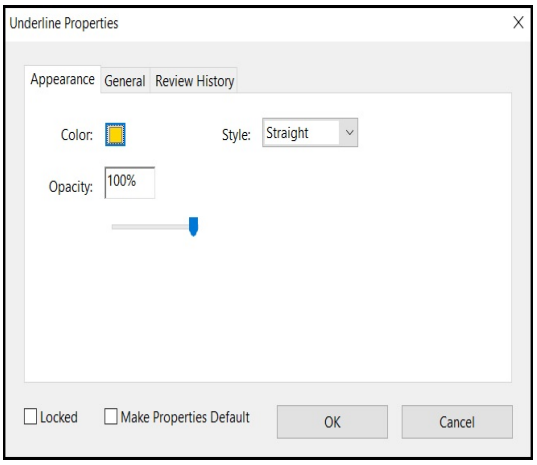
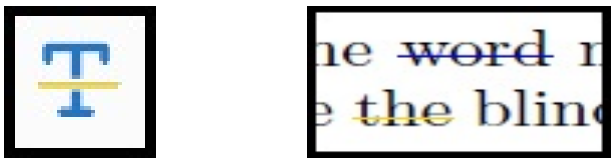


Figure 39: Properties of the "Underline text" option in the document.

- **Strikethrough text**

  This type of annotation is represented in the toolbar by the icon in the Figure 40a and allows the user to strikethrough parts of the text as shown in the Figure 40b.



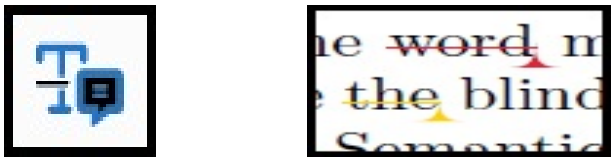
(a) In the Figure 33. (b) In the document.

Figure 40: Graphical representation of the "Strikethrough text" option.

  The appearance properties of the annotations shown in the Figure 40b can be edited through the same window presented in the Figure 37 that arise from right clicking on them or on the respective comments that are present in the comments list that is visible in Figure 74.

- **Add note to replace text**

  This type of annotation is represented in the toolbar and in the document by the icon shown in the Figure 41a and in the way shown in the Figure 41b respectively, with the main objective of allowing the user to write notes to replace the parts of the text to which they are referring.



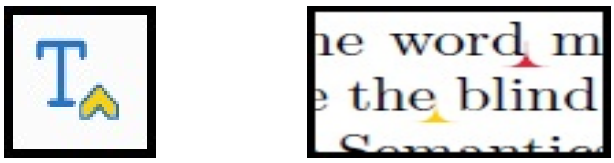
(a) In the Figure 33. (b) In the document.

Figure 41: Graphical representation of the "Add note to replace text" option.

The appearance properties of the annotations shown in the Figure 41b can be edited through the same window presented in the Figure 37 that arise from right clicking on them or on the respective comments that are present in the comments list that is visible in Figure 74.

- **Insert text at cursor**

  This type of annotation is represented in the toolbar and in the document by the icon shown in the Figure 42a and in the way shown in the Figure 42b respectively, with the main objective of allowing the user to indicate the location in the text to which the notes are referring.



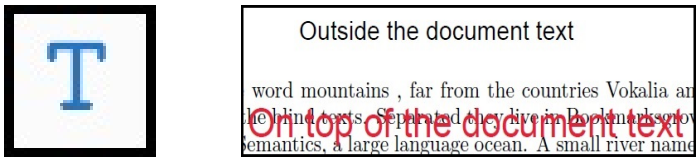
(a) In the Figure 33. (b) In the document.

Figure 42: Graphical representation of the "Insert text at cursor" option.

The appearance properties of the annotations shown in the Figure 42b can be edited through the same window presented in the Figure 37 that arise from right clicking on them or on the respective comments that are present in the comments list that is visible in Figure 74.

- **Add text comment**

  This option is represented in the toolbar by the icon in the Figure 43a and allows the user to add comments to any location in the document as shown in the Figure 43b.

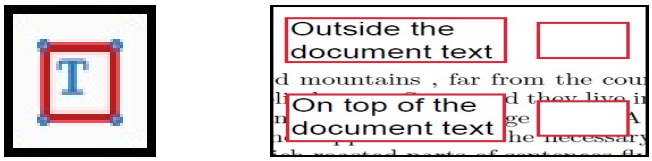


(a) In the Figure 33. (b) In the document.

Figure 43: Graphical representation of the "Add text comment" option.

- **Add text box**

  This option is represented in the toolbar and in the document by the icon shown in the Figure 44a and in the way shown in the Figure 44b respectively, with the purpose of allowing the user to add boxes in the document in which comments can be written.



(a) In the Figure 33. (b) In the document.

Figure 44: Graphical representation of the "Add text box" option.

The appearance properties of the annotations shown in the Figure 44b can be edited through the options visible in the Figure 45 that arise from right clicking on them or on the respective comments that are present in the comments list that is visible in Figure 74.

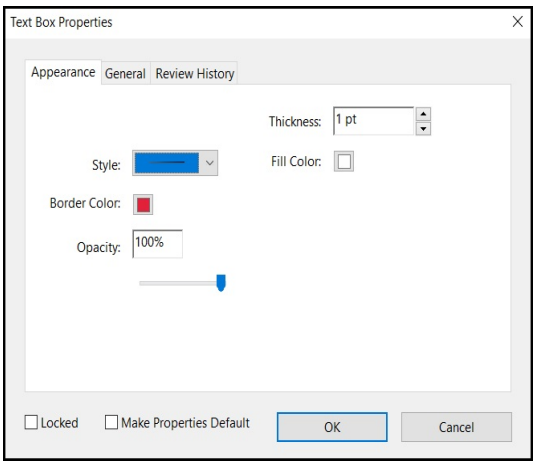
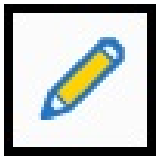


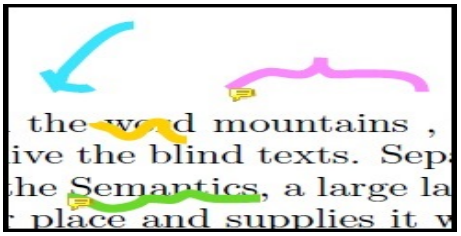
Figure 45: Properties of the "Add text box" option in the document.

- **Use drawing tool**

  This option is represented in the toolbar by the icon in the Figure 46a and allows the user to design the notes with the visual aspect that he wants as shown in the Figure 46b.



(a) In the Figure 33.



(b) In the document.

Figure 46: Graphical representation of the "Use drawing tool" option.

The appearance properties of the annotations shown in the Figure 46b can be edited through the options visible in the Figure 47 that arise from right clicking on them or on the respective comments that are present in the comments list that is visible in Figure 74.

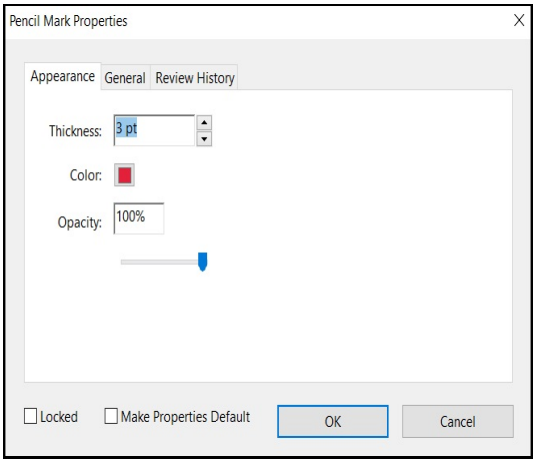


Figure 47: Properties of the "Use drawing tool" option in the document.

- **Erase drawing**

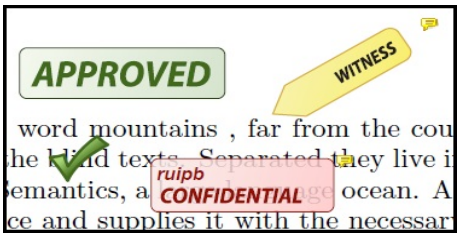
  This type of annotation is represented in the toolbar by the icon in the Figure 48 and allows the user to erase the drawings he made.

Figure 48: Icon of the "Erase drawing" option in Figure 33.

- **Add stamp**

  This option is represented in the toolbar by the icon in the Figure 49a and allows the user to place stamps with different purposes on the document as shown in the Figure 49b.

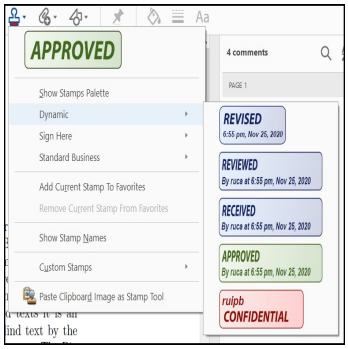
(a) In the Figure 33.



(b) In the document.

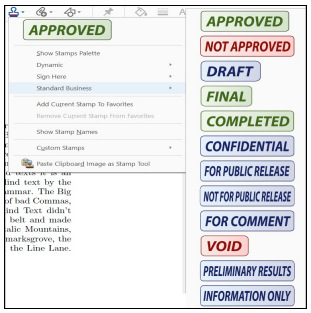
Figure 49: Graphical representation of the "Add stamp" option.

Since there are too many stamps to be shown here, it will only be possible to demonstrate them through Figure 50. However, the graphical representations that are visible in Figure 50 are the same when the stamps are placed on the document.



(a) In "Dynamic" sub-option.

(b) In "Sign Here" sub-option.



(c) In "Standard Business" sub-option.

Figure 50: All stamps available.

The appearance properties of the annotations shown in the Figures 49b and 50 can be edited through the same window presented in the Figure 37 that arise from right

clicking on them or on the respective comments that are present in the comments list that is visible in Figure 74.

- **Add a new attachment**

  This option is represented in the toolbar by the icon in the Figure 51.

Figure 51: Icon of the "Add a new attachment" option in Figure 33.

It is however divided into the sub-options "Attach File" and "Record Audio", both of which are visible in Figure 52.

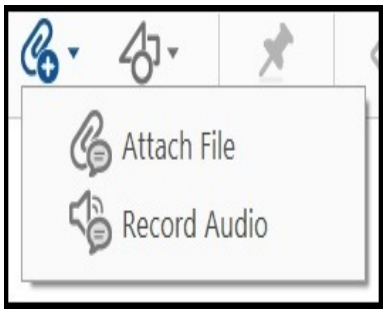


Figure 52: Sub-options of the "Add a new attachment" option.

That being said, each of them will be explored below individually.

Attach File

This option allows the user to insert files in any place of the document as shown by the Figure 53.

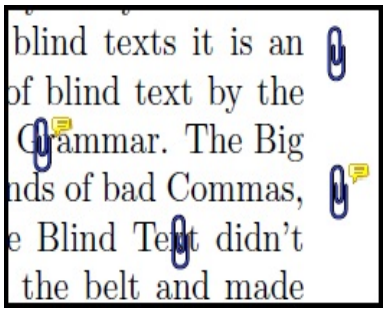


Figure 53: Application of sub-option "Attach File" in the document.

The appearance properties of the annotations shown in the Figure 53 can be edited through the options visible in the Figure 54 that arise from right clicking on them or on the respective comments that are present in the comments list that is visible in Figure 74.

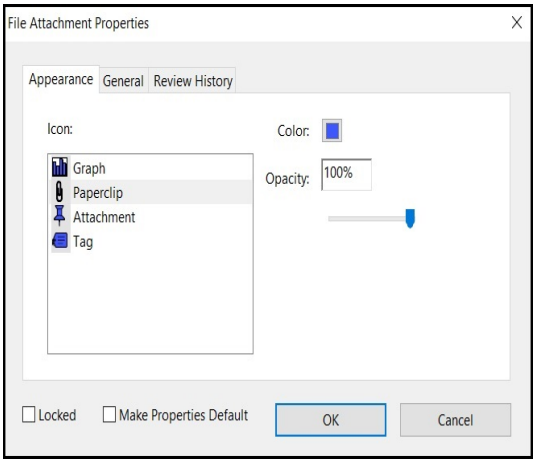


Figure 54: Properties of the "Attach File" sub-option.

Record Audio

This option allows the user to record audio and insert it in any place of the document as shown by the Figure 55.

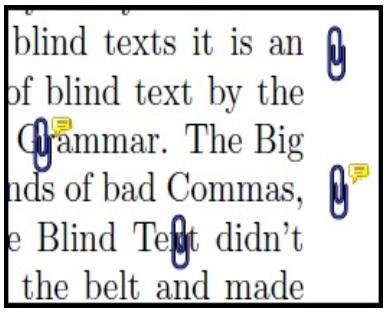


Figure 55: Application of sub-option "Record Audio" in the document.

The appearance properties of the annotations shown in the Figure 55 can be edited through the options visible in the Figure 56 that arise from right clicking on them or on the respective comments that are present in the comments list that is visible in Figure 74.

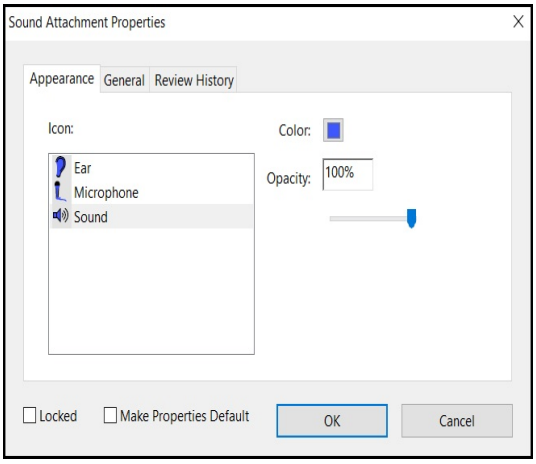


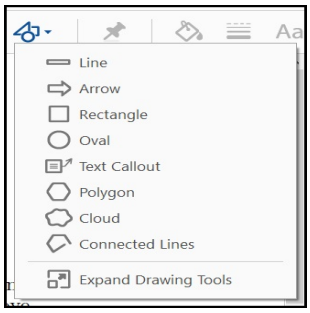
Figure 56: Properties of the "Record Audio" sub-option.

- **Drawing tools**

  This option is represented in the toolbar by the icon in the Figure 57a.

  When selecting this option, several sub-options will be presented to annotate the text as shown by the Figure 57b. In order to explain more clearly each one of them we will use the option "Expand Drawing Tools" which, as the name implies, will place each one of these options in the toolbar shown in Figure 33 as it is possible to verify by the Figure 58.

(a) In the Figure 33.



(b) Sub-options of the "Drawing tools" option.

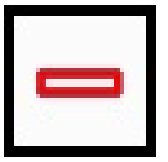
Figure 57: Graphical representation of the "Drawing tools" and its content.

Figure 58: Result of the "Expand Drawing Tools" option.

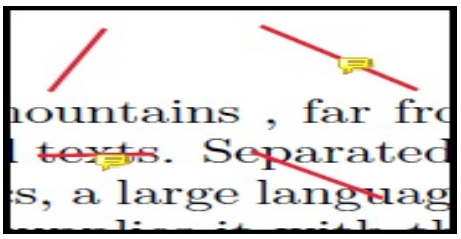
Based on Figure 58, the sub-options of the "Drawing tools" option will be described below.

Draw line

This option is represented in the toolbar by the icon in the Figure 59a and allows the user to draw lines in the document as shown in the Figure 59b.



(a) In the Figure 58.



(b) In the document.

Figure 59: Graphical representation of the "Draw line" sub-option.

The appearance properties of the annotations shown in the Figure 59b can be edited through the options visible in the Figure 60 that arise from right clicking on them or on the respective comments that are present in the comments list that is visible in Figure 74.

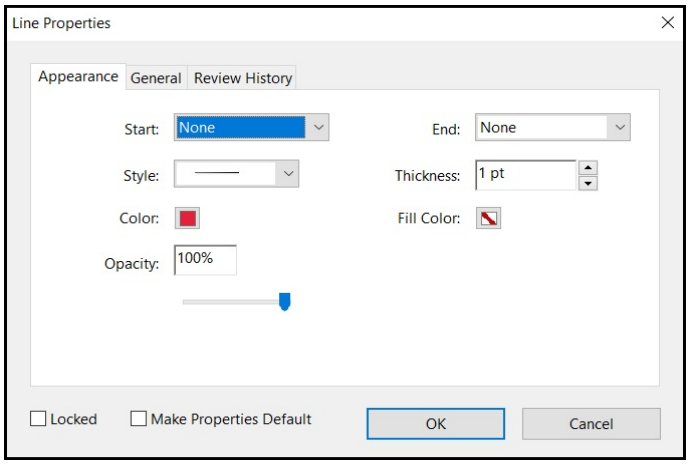
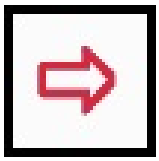


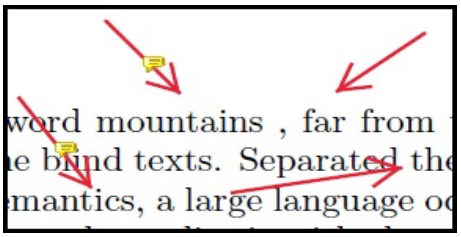
Figure 60: Properties of the "Draw line" sub-option.

Draw arrow

This type of annotation is represented in the toolbar by the icon in the Figure 61a and allows the user to draw arrows in the document as shown in Figure 61b.



(a) In the Figure 58.



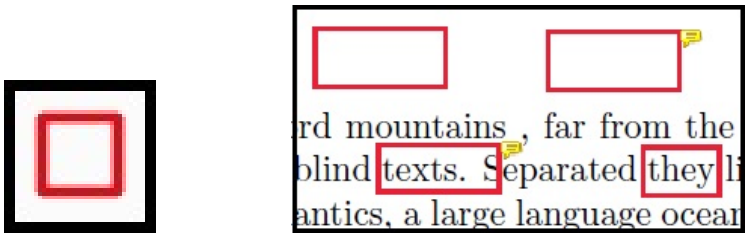
(b) In the document.

Figure 61: Graphical representation of the "Draw arrow" sub-option.

The appearance properties of the annotations shown in the Figure 61b can be edited through the same window presented in the Figure 60 that arise from right clicking on them or on the respective comments that are present in the comments list that is visible in Figure 74.

Draw rectangle

This option is accessed in the toolbar through the icon in the Figure 62a and allows the user to draw rectangles in the document as shown in Figure 62b.



(a) In the Figure 58. (b) In the document.

Figure 62: Graphical representation of the "Draw rectangle" sub-option.

The appearance properties of the annotations shown in the Figure 62b can be edited through the options visible in the Figure 63 that arise from right clicking on them or on the respective comments that are present in the comments list that is visible in Figure 74.

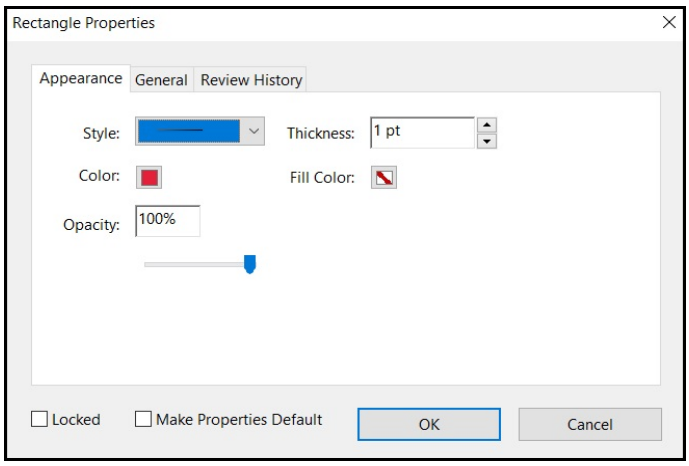
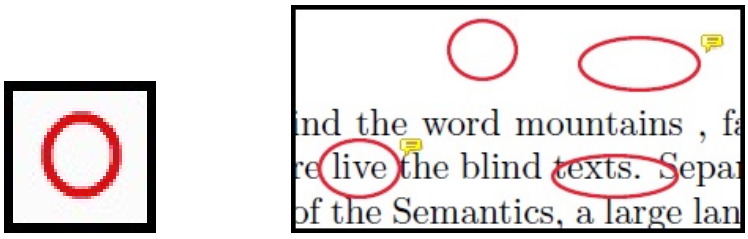


Figure 63: Properties of the "Draw rectangle" sub-option.

Draw oval

This type of annotation is represented in the toolbar through the icon in the Figure 64a and allows the user to draw oval figures in the document as shown in Figure 64b.



(a) In the Figure 58. (b) In the document.

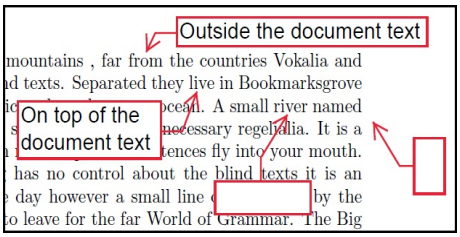
Figure 64: Graphical representation of the "Draw oval" sub-option.

The appearance properties of the annotations shown in the Figure 64b can be edited through the same window presented in the Figure 63 that arise from right clicking on them or on the respective comments that are present in the comments list that is visible in Figure 74.

Draw text callout

This option is represented in the toolbar by the icon in the Figure 65a and allows the user to draw text boxes, each of which is connected to an arrow as shown in Figure 65b. This type of annotation has the main purpose of indicating with the arrow the part of the document to which the content inside the text box refers.

(a) In the Figure 58.



(b) In the document.

Figure 65: Graphical representation of the "Draw text callout" sub-option.

The appearance properties of the annotations shown in the Figure 65b can be edited through the options visible in the Figure 66 that arise from right clicking on them or on the respective comments that are present in the comments list that is visible in Figure 74.

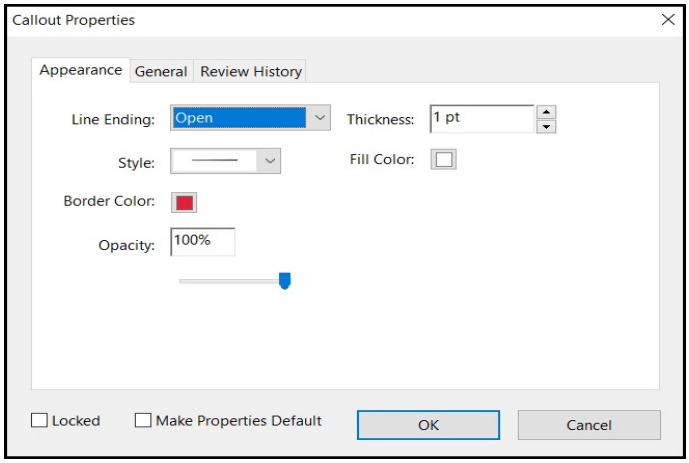
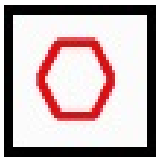


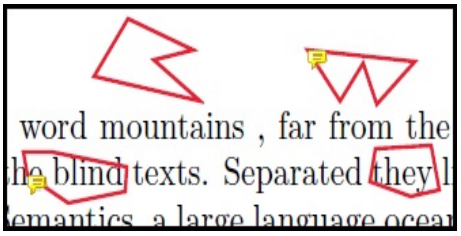
Figure 66: Properties of the "Draw text callout" sub-option.

Draw polygon

This option is represented in the toolbar by the icon in the Figure 67a and allows the user to draw polygons in the document as shown in Figure 67b.



(a) In the Figure 58.



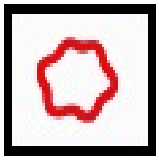
(b) In the document.

Figure 67: Graphical representation of the "Draw polygon" sub-option.

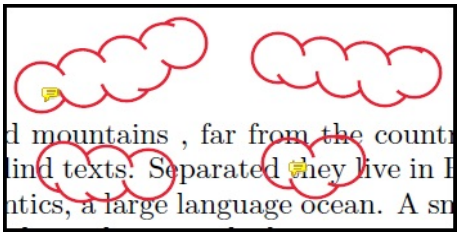
The appearance properties of the annotations shown in the Figure 67b can be edited through the same window presented in the Figure 63 that arise from right clicking on them or on the respective comments that are present in the comments list that is visible in Figure 74.

Draw cloud

This type of annotation is represented in the toolbar by the icon shown in the Figure 68a and allows the user to draw clouds in the document as shown in Figure 68b.



(a) In the Figure 58.



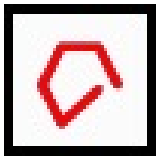
(b) In the document.

Figure 68: Graphical representation of the "Draw cloud" sub-option.

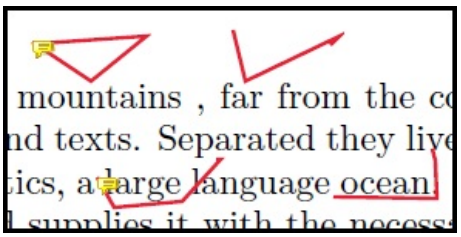
The appearance properties of the annotations shown in the Figure 68b can be edited through the same window presented in the Figure 63 that arise from right clicking on them or on the respective comments that are present in the comments list that is visible in Figure 74.

Draw connected lines

This type of annotation is accessed in the toolbar through the icon shown in the Figure 69a and allows the user to draw connected lines in the document as shown in Figure 69b.



(a) In the Figure 58.



(b) In the document.

Figure 69: Graphical representation of the "Draw connected lines" sub-option.

The appearance properties of the annotations shown in the Figure 69b can be edited through the same window presented in the Figure 60 that arise from right clicking on them or on the respective comments that are present in the comments list that is visible in Figure 74.

- **Keep tool selected**

  This option is represented in the toolbar by the icon shown in the Figure 70 and when its activated it will keep the type of annotation that is being used selected so that the user can use it several times without having to be always selecting it whenever he uses it in the document.

Figure 70: Icon of the "Keep tool selected" option in Figure 33.

- **Change color**

  This option is represented in the toolbar by the icon shown in the Figure 71 and allows the user to change the color of some types of annotations.

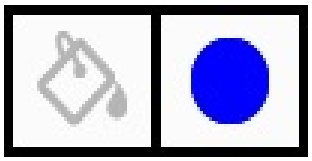


Figure 71: Icon of the "Change color" option in Figure 33.

- **Change line thickness**

  This option is represented in the toolbar by the icon present in the Figure 72 and allows the user to change the line thickness of some types of annotations.

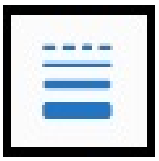


Figure 72: Icon of the "Change line thickness" option in Figure 33.

- **Text properties**

  This option is represented in the toolbar by the icon shown in the Figure 73 and allows the user to change text formatting. The formatting options depend on the type of annotation and only allow basic formatting such as changing the font color, font, italic, underline, etc.

Figure 73: Icon of the "Text properties" option in Figure 33.

After any type of annotation has been inserted in the document, a box will appear in the comments list to write its content as shown in Figure 74.

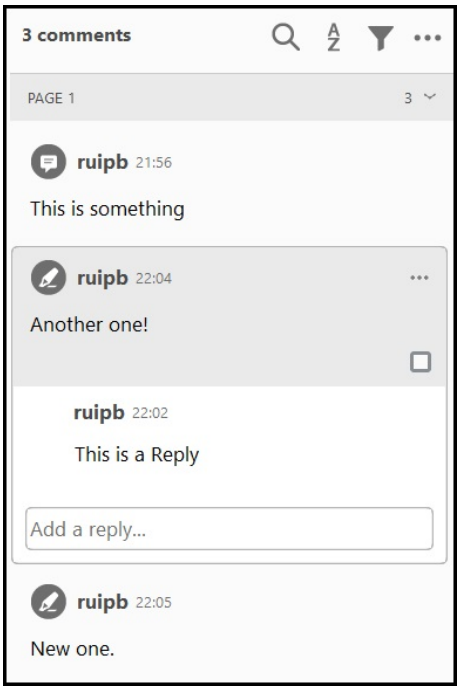


Figure 74: Comments list.

The only ways that *Adobe Reader* makes available to view the content of the annotations made is through the comments list visible in the Figure 74, through the option "Open Pop-Up Note" obtained from right-clicking on them which results in Figure 75 or through the pop-up, visible in Figure 76, which appears after hovering the mouse cursor over them.

Figure 75: "Open Pop-Up Note" option.

Figure 76: Pop-up of an annotation.

Another feature that only works on the comments list visible in Figure 74 is the ability to automatically locate in the text the annotations that are present in it after being clicked.

Adobe Reader also provides methods to search for a word/phrase in the content of the annotations comments. Note that this method is different from the "Search comments..." option, visible in Figure 79 and which will be discussed later, as it does not filter comments. That said, to use this feature the user will have three options, these being through the use of the "Find" sub-option of the "Edit" option, clicking on the magnifying glass or by pressing *Ctrl+F*, where the first two options can be seen in Figure 28. After the user executes one these methods, the small window that is shown in Figure 77 will be displayed, after which it is necessary to select the sub-option "Include Comments" in order to search in the comments of the annotations.

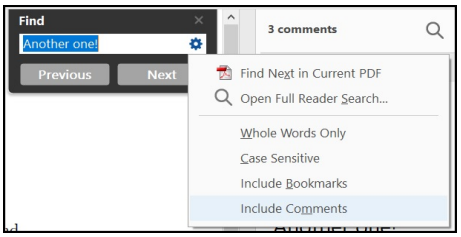


Figure 77: Normal search window.

It is important to note that the user can perform a more advanced search through three methods, these being through the "Advanced Search" sub-option present in the "Edit" option, through the "Open Full Reader Search..." option or by pressing *Shift+Ctrl+F*, where the first two options can be seen in Figures 28 and 77 respectfully. The result of these options will be the window present in Figure 78, which will present more options that the user can use to improve his search.

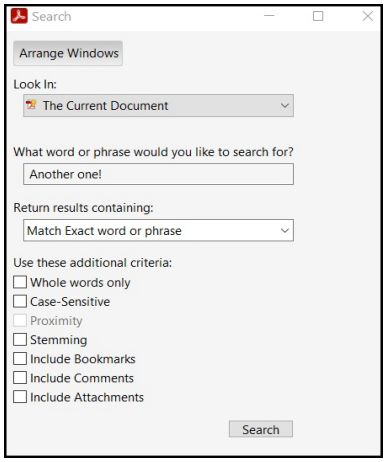


Figure 78: "Advanced Search" option window.

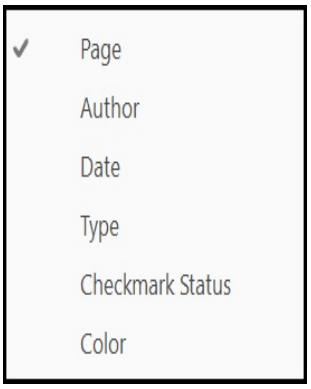
In order to help the user to search content in its annotations and organize them in a more desirable way, *Adobe Reader* offers some options that make it possible to achieve these objectives. This options can be seen in the top right corner of Figure 74 and will be analyzed below.

The first option to be described is the "Search comments..." which is represented in the comments list by the icon visible in the Figure 79. This option allows the user to search for the occurrence of a word/phrase in the content of the comments present in the comments list shown in the Figure 74.

Figure 79: Icon of the "Search comments..." option in Figure 74.

As for the "Sort Comments" option represented in the comments list by the icon shown in Figure 80a is used to organize the annotations according to one of the several criteria visible in Figure 80b.

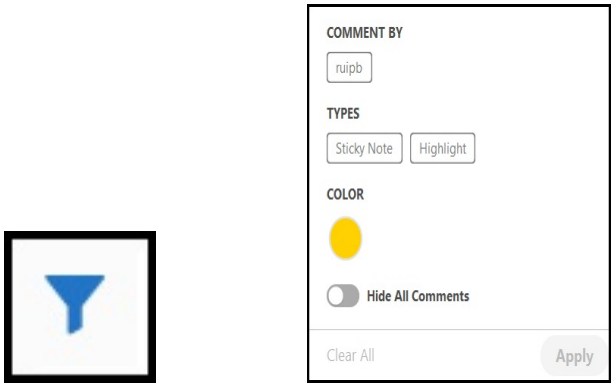
(a) Icon in the Figure 74.



(b) Organization criteria.

Figure 80: "Sort Comments" option.

Regarding the option "Filter comments" represented in the comments list by the icon of the Figure 81a has the function of allowing the filtering of the comments through the norms visible in the Figure 81b.



(a) Icon in the Figure 74. (b) Filtration standards.

Figure 81: "Sort Comments" option.

Finally, the option "Options", which is represented in the comments list by the icon of the Figure 82a allows to export some annotations, export all annotations and import annotations as can be seen in the Figure 82b. Note that the way of exporting the work done allows it to be imported afterwards so that it can be continued.

(a) Icon in the Figure 74. (b) Sub-options of the "Options" option.

Figure 82: "Options" option.

To edit the content of the annotations the user can do it in several ways, one of which involves double-clicking on the respective annotation in the document or in the respective comment in the comments list visible in Figure 74. The other is done through the options that arise from right-clicking on the respective comment present in the comments list visible in Figure 74. The last method is done through the result of the option "Open Pop-Up Note", shown in the Figure 75, obtained from right-clicking on an annotation located on the document.

The user can also remove annotations, which like editing can be done in several ways. One of them is done through the options that arise from right-clicking on the respective annotation in the document or on the respective comment present in the comments list

visible in Figure 74. The last method is achieved with one of the many options that arise from right-clicking on the window shown in Figure 75 that results from choosing the option "Open Pop-Up Note" that is accessed by right-clicking on top of the respective annotation in the document.

Regarding commenting, *Adobe Reader* provides some customization options. For that, the user must select the sub-option "Preferences..." of the option "Edit", visible in Figure 28, or by simply pressing *Ctrl+K*. The result of both actions will be the window visible in Figure 83 in which it will be necessary to select the "Commenting" option, through which the user can customize various aspects of the pop-ups and comments.

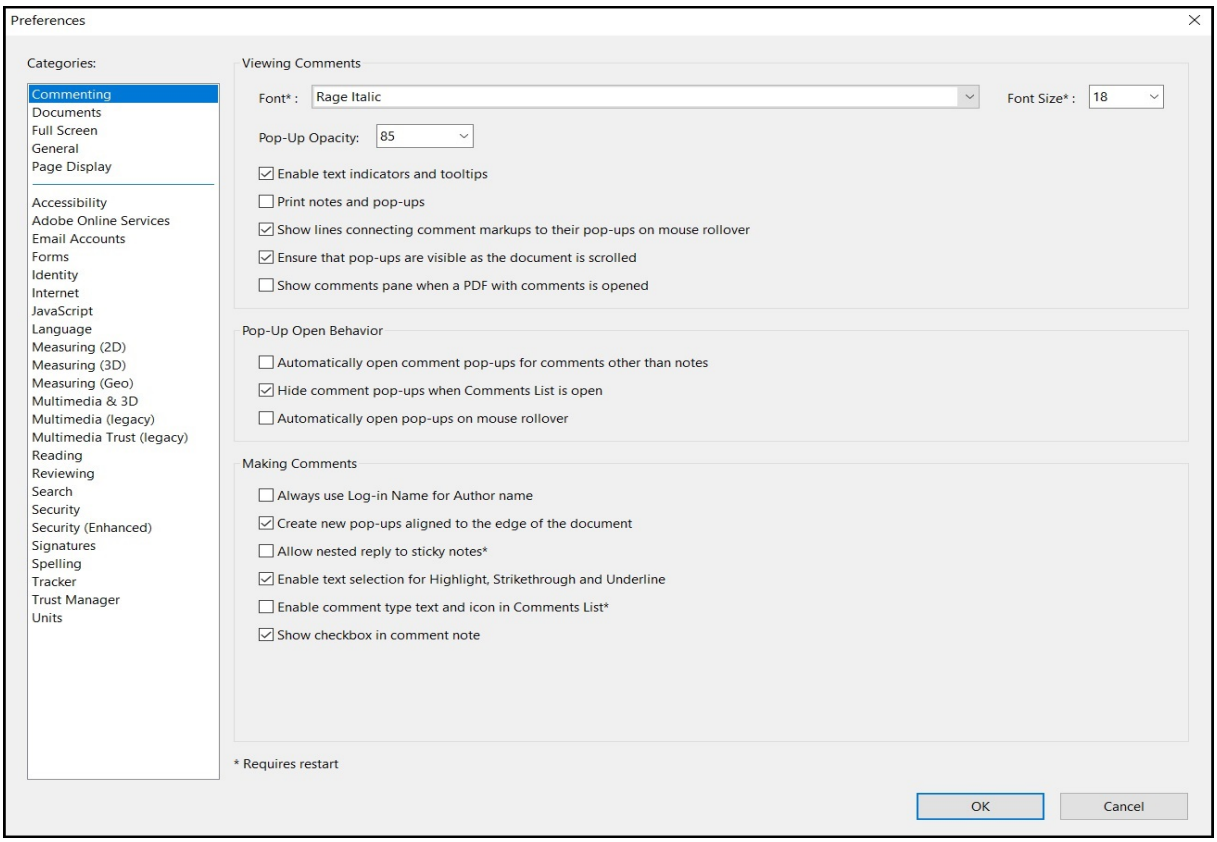


Figure 83: Commenting preferences.

Adobe Reader also supports collaborative work. In addition to the classic annotation sharing method of exchanging the entire document, this tool offers two more elegant ways to carry out this process. One of them is through the options that allow importing and exporting annotations, all of which are present in the option in Figure 82b. The other method is through the option "Send for Comments" present in the sidebar, as shown in Figure 84, through which it is possible to send a link to the recipients so that they can see and comment on the document online.

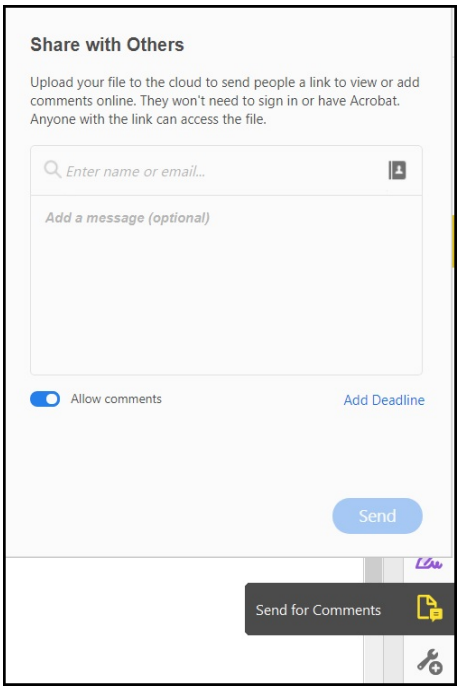


Figure 84: "Send for Comments" option.

In addition to being able to annotate the documents together in these two methods mentioned above, the *Adobe Reader* also allows its users to respond to comments made by other users, which can be done in several ways. One of them is done in the comments list through the text area that has the text "Add a reply ..." as a placeholder as it is visible in the Figure 74. The other method is through the "Reply" button present in the window shown in Figure 75 that results from choosing the option "Open Pop-Up Note" that is accessed by right-clicking on top of an annotation present in the document. The final method is done through the options that arise from right-clicking on an annotation that is present in the document.

Taking into account all the features discussed above, it is possible to conclude that *Adobe Reader* is an excellent tool for annotation purposes as it is a tool that, although it appears to be complex, is actually simple and easy to understand. In addition to providing various forms of annotation, this being its strongest point, it also offers several ways to manage them. However, this tool does not fit into what it is intended to achieve with this thesis since it is not a web based tool.

2.7.4 *Web annotation tools*

Below is a list containing the names of some web annotators that were found after conducting a specific search.

- *Weava*[8]
- *LINER*[9]
- *Sciwheel*[10]
- *Beanote*[11]
- *Note Board*[12]
- *Note Anywhere*[13]
- *Additor*[14]
- *Markit*[15]
- *JustClip*[16]
- *Highlights*[17]
- *NotesAlong*[18]
- *Pundit Annotator*[19]
- *Simple Webpage Note*[20]
- *Diigo*[21]
- *RevNote*[22]

8 Available at https://www.weavatools.com/, accessed in January 2021
9 Available at https://getliner.com/, accessed in January 2021
10 Available at https://chrome.google.com/webstore/detail/sciwheel-browser-extensio/ljfhgpiambpnabgpnaihcebebmoijfci, accessed in January 2021
11 Available at https://chrome.google.com/webstore/detail/beanote-note-taking-on-we/nikccehomlnjkmgmhnieecolhgdafajb, accessed in January 2021
12 Available at https://chrome.google.com/webstore/detail/note-board-sticky-notes-a/goficmpcgcnombioohjcgdhbaloknabb, accessed in January 2021
13 Available at https://chrome.google.com/webstore/detail/note-anywhere/bohahkiiknkelflnjjlipnaeapefmjbh, accessed in January 2021
14 Available at https://additor.io/product, accessed in January 2021
15 Available at https://chrome.google.com/webstore/detail/markit-highlighter-and-no/oilpcbohncpdjdadofhbldfmojneciop, accessed in January 2021
16 Available at https://justclip.co/, accessed in January 2021
17 Available at https://www.learningpaths.io/, accessed in January 2021
18 Available at https://notesalong.com/, accessed in January 2021
19 Available at https://thepund.it/annotator-web-annotation/, accessed in January 2021
20 Available at https://chrome.google.com/webstore/detail/simple-webpage-note/akfajpjakckpkbhjmollinmpojhndafp, accessed in January 2021
21 Available at https://www.diigo.com/, accessed in January 2021
22 Available at https://www.revnote.io/, accessed in January 2021

- *Sticky Note*[23]
- *Web Page Sticky Notes*[24]
- *Sticky Note*[25]
- *Anywhere stickers*[26]
- *Self-stick notes*[27]
- *WAFFLEPEN Highlighter*[28]
- *ClipTo*[29]
- *Hypothes.is*[30]
- *Genius Web Annotator*[31]
- *WorldBrain's Memex*[32]
- *Scrible*[33]
- *Remarq*[34]

Of these tools listed, only the features of *Weava*, *LINER*, *Diigo* and *Hypothes.is* are analyzed below because they are the most popular and because they fit the objective of this thesis.

23 Available at https://chrome.google.com/webstore/detail/sticky-note/khmmaepcamdkhdikdaphneenognobngd, accessed in January 2021
24 Available at https://chrome.google.com/webstore/detail/web-page-sticky-notes/alpjieidnmmkljnceakgpeajlngabnee, accessed in January 2021
25 Available at https://chrome.google.com/webstore/detail/sticky-note/cmabpemmbmhlgmnlemmkelphankfclcf, accessed in January 2021
26 Available at https://chrome.google.com/webstore/detail/anywhere-stickers-simple/joiomjhjkacipamidllnbicjcdmoheha, accessed in January 2021
27 Available at https://chrome.google.com/webstore/detail/self-stick-notes/hcdihnnnfbdeinkamoggggghnmcfaebca?hl=pt-BR, accessed in January 2021
28 Available at https://chrome.google.com/webstore/detail/wafflepen-highlighter/dboaglbnegpafcgeabpfblnmgelejmlc, accessed in January 2021
29 Available at https://chrome.google.com/webstore/detail/clipto-notes-highlights-a/ngcfglcfnghkmiihacggclkdcijilhnh, accessed in January 2021
30 Available at https://web.hypothes.is/#features, accessed in January 2021
31 Available at https://genius.com/web-annotator, accessed in January 2021
32 Available at https://getmemex.com/, accessed in January 2021
33 Available at https://www.scrible.com/, accessed in January 2021
34 Available at https://remarqable.com/web/index.html, accessed in January 2021

Weava

According to the ***Weava*** website,

> Weava is a Chrome extension that allows you to highlight PDFs and websites with any colour you'd like. Weava is a workspace for your research and studies, offering tools that help you highlight, annotate, organize and collaborate on your research[35].

The main functionality of this tool is the ability to make highlights that are made through the selection of the text to be annotated, completing this process through the options that arise from performing a right-click on the selection or by choosing a color for the highlight through the pop-up that appears after completing the selection as shown in Figure 85. There is also the possibility to remove this highlights as shown in the Figure 86.

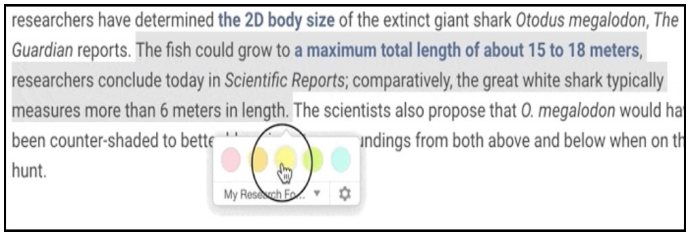


Figure 85: Possible highlight colors[36].

Figure 86: Removing a highlight[36].

The user also has the possibility to annotate the highlights, this being done through the box shown in Figure 87 that appears after clicking on the highlighted text. After that, the respective highlight will have a gray underline, as shown in Figure 87, in order to make it easier to see that it has already been annotated. Note that although these annotations have no length limit, it is not possible to use complex formatting in them.

35 Available at https://www.weavatools.com/the-weava-manual/, accessed in January 2021
36 Available at https://www.weavatools.com/highlight-with-weava/, accessed in January 2021

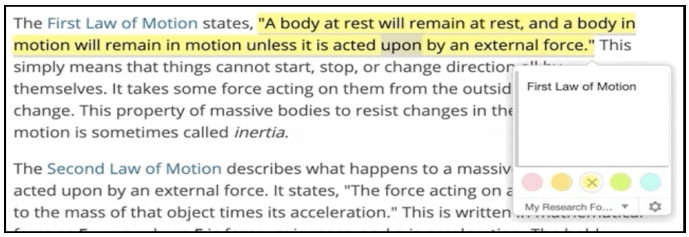


Figure 87: Annotation of a highlight[37].

Another interesting feature is the ability to clip images in which it is only necessary to drag the desired image to a specific area as shown in Figure 88, making it easier to access them later.

Figure 88: Clip images functionality[36].

In addition to allowing the annotation of online websites and online PDFs, *Weava* also makes it possible to annotate local PDFs files on the user's computer. This can be done by opening the PDF document with *Google Chrome* or by uploading the PDF file from the option present in the *Weava* sidebar or in the *Weava* dashboard which will both be discussed below.

Regarding the viewing and management of the annotations, *Weava* offers two ways to do this, one of them through the *Weava* sidebar, this being an extension in the *Google Chrome*, and the other through the *Weava* dashboard, this being the web app. To begin we will analyze the features of the *Weava* sidebar since all of these are also present in the *Weava* dashboard and note that all images shown hereinafter were taken from the *Weava* dashboard perspective.

In both cases, the user can copy and delete highlights as well as add, copy, edit and delete notes as shown in Figure 89.

36 Available at https://www.weavatools.com/highlight-with-weava/, accessed in January 2021
37 Available at https://www.weavatools.com/annotate-with-weava/, accessed in January 2021

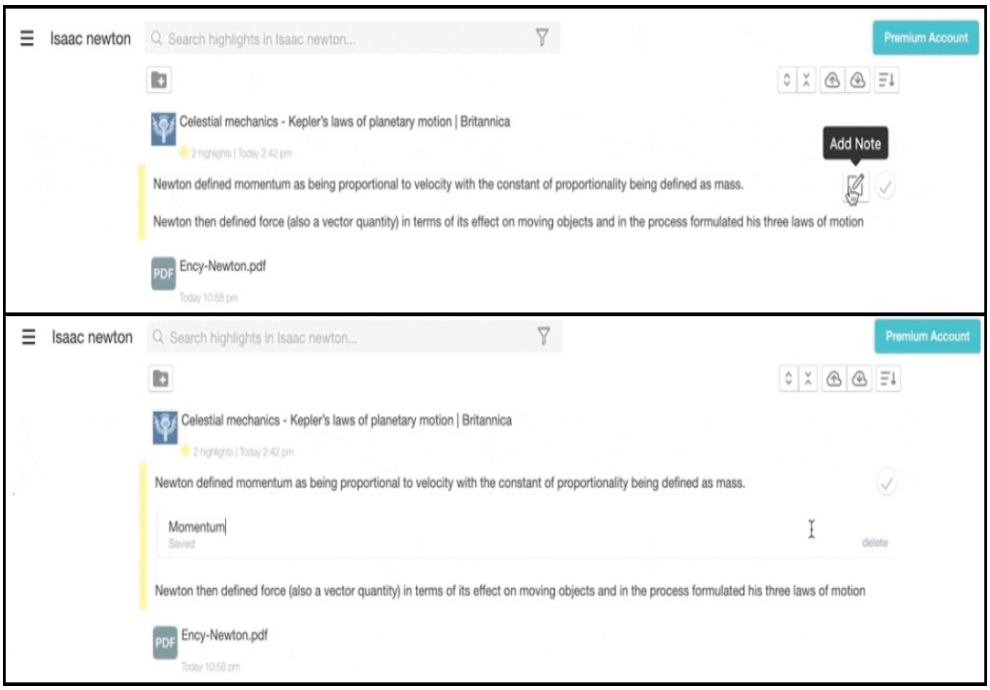


Figure 89: Highlights and annotations management[38].

It is also possible to search for highlights through a search bar, show/hide highlights and filter/sort highlights according to the options shown in Figure 90.

Figure 90: Filter/sort options[38].

In order to better organize the annotations, *Weava* allows the user to add, edit and delete folders and sub-folders. Having these created, the user can move highlights between them and add, edit and delete color labels so that they have some meaning within the folder or sub-folder where they are inserted as shown in Figure 91. The user can also drag folders into other folders as shown in Figure 92.

38 Available at https://www.weavatools.com/organize-with-weava/, accessed in January 2021

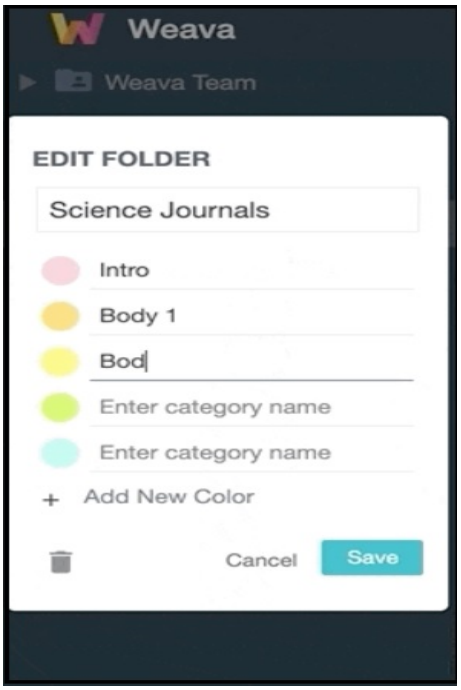


Figure 91: Color management of labels within folders/sub-folders[38].

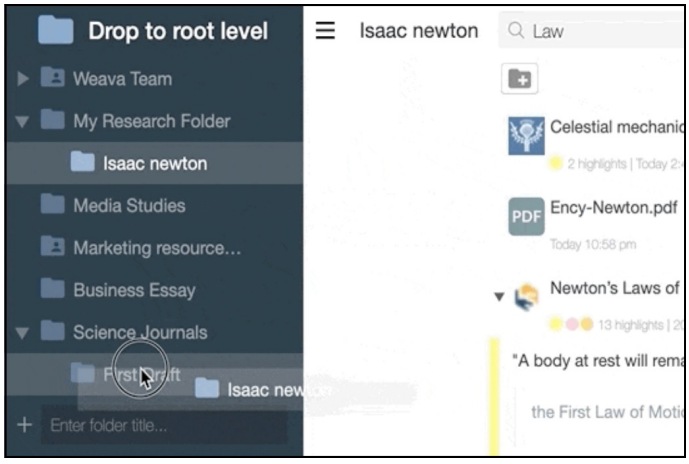


Figure 92: Drag folders functionality[38].

Finally, let's look at some features exclusive to the *Weava* dashboard. These allow the user to create and copy citations of the sources used, create *Weava* documents that will be saved in their folders and sub-folders, move several highlights at once from one folder to another, update URLs of websites/PDFs that have changed file paths and apply the filter/sorting options shown in Figure 90 on the highlights that were made on various websites and PDFs.

38 Available at https://www.weavatools.com/organize-with-weava/, accessed in January 2021

The exportation of highlights/annotations is also possible, however this method does not allow them to be imported later.

Another ability of the *Weava* dashboard is to allow the user to continue to highlight a website/PDF that was previously highlighted. To do so, the user just has to select an highlight present on the dashboard that has been made on the website/PDF that he wants to continue to annotate, opening it in an iFrame on the dashboard as shown in the Figure 93.

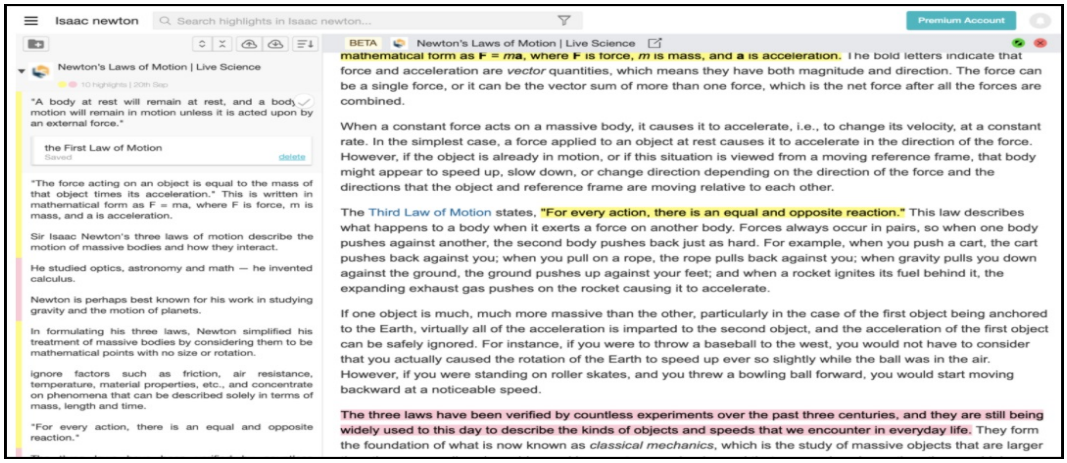


Figure 93: Editing on iFrame[38].

Weava also allows collaborative work by sharing the folders and sub-folders, which in this case will have a "person" icon to distinguish them from those that are not being shared, thus allowing multiple users to annotate the same websites and PDFs. In addition to being able to check with whom the folder is being shared, it is also possible to remove users from the share. Unfortunately, only the owner of the shared folder can remove any of its sub-folders. Note that accept folder invitations are only accessible through the *Weava* dashboard, since only this one receives notifications.

In conclusion, *Weava* is a very useful web-based annotation tool that allows its users to annotate online websites and online PDFs in a simple and intuitive way. Although it only provides two ways to view and manage annotations, it compensates by offering several options that allow users to manage annotations in a variety of ways. Taking all this into account it is possible to affirm that this tool fits well with the objectives of this thesis, despite the fact that some of the features and feature options that it offers are only accessible through a premium account that costs money thus limiting the use of the application for free users.

38 Available at https://www.weavatools.com/organize-with-weava/, accessed in January 2021

LINER

LINER is another web-based annotator that allows users to annotate websites and PDFs. For this, the user will have to select the text he wants to annotate followed by clicking on the icon that appears after the selection is finished in order to complete the highlight as shown in Figure 94.

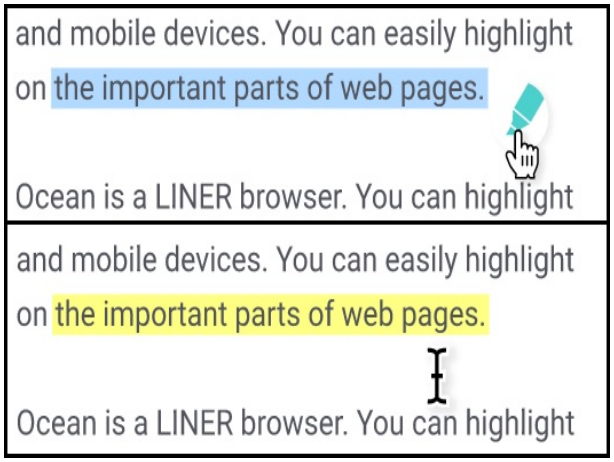


Figure 94: Highlighting process[39].

After that, the user can change the highlight color or annotate the highlight simply by clicking on the respective highlight as shown in Figure 95.

Figure 95: Options after performing a highlight[39].

When the user chooses the annotation option, a text box will appear as shown in Figure 96.

39 Available at https://getliner.com/, accessed in January 2021

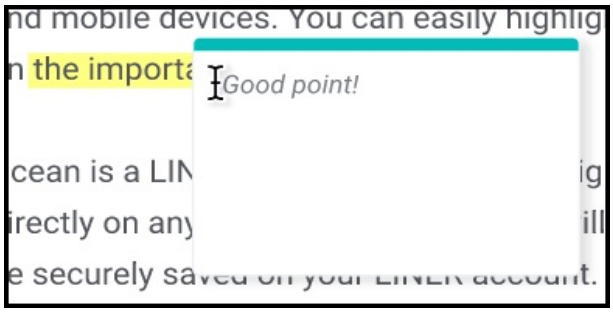


Figure 96: Annotation option text box[39].

Another feature that *LINER* makes available after it is added to the browser is the ability to recommend pages that appear in user search results that have been highlighted by other users.

In order to allow a better management of the elaborated work, *LINER* allows its users to add tags to the highlighted pages in order to find them more easily in the future.

This management can only be done on the *LINER* web app, this being the only place where the user can view all his annotations as can be seen in the Figure 97. Here the users can filter by tags, by the highlight color, by page type and filter annotations whose title, URL, highlighted text, description and/or written comment contain a certain word/phrase, add tags to the page, open the pages to continue making highlights on them and export annotations through a method that does not allow them to be imported later. As the user makes annotations, *LINER* will also recommend other articles that may be of his interest based on the annotations he made.

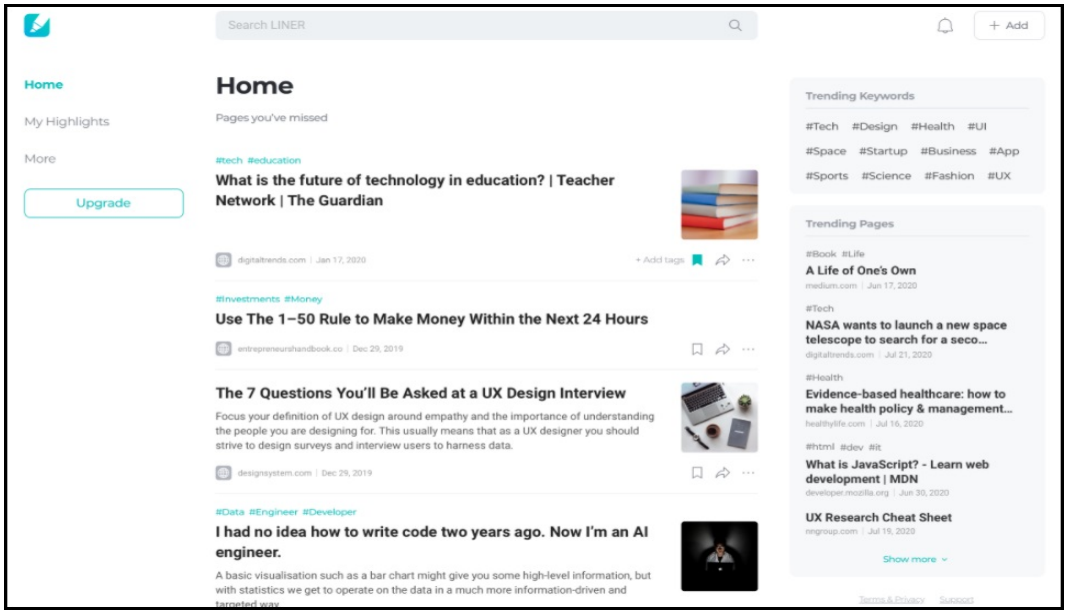


Figure 97: *LINER* web app[39].

Like the tools already analyzed, *LINER* also allows collaborative work. To do this, the user can share a page with all the highlights and comments that were made in it with another person by simply sending a link.

After analyzing all its features, it is possible to conclude that *LINER* is a very simple yet versatile web-based annotation tool that fits with the objectives that this thesis is intended to achieve. However, this tool also distinguishes its users through free and premium accounts offering some more features and feature options to the latter type, thus limiting the use of the tool in the case of free users.

Diigo

Diigo is an all-in-one web-based tool that allows its users to highlight and annotate, archive, take screenshots and bookmarking web pages.

Through the bookmarking functionality the users will be able to save web pages with the possibility of using the several options shown in Figure 98 that will allow them to find the web pages that were bookmarked more easily in the future. Of all the options shown in Figure 98, it is worth highlighting the option to associate tags to a web page and adding the web page to a list, since both will assist in managing the work done.

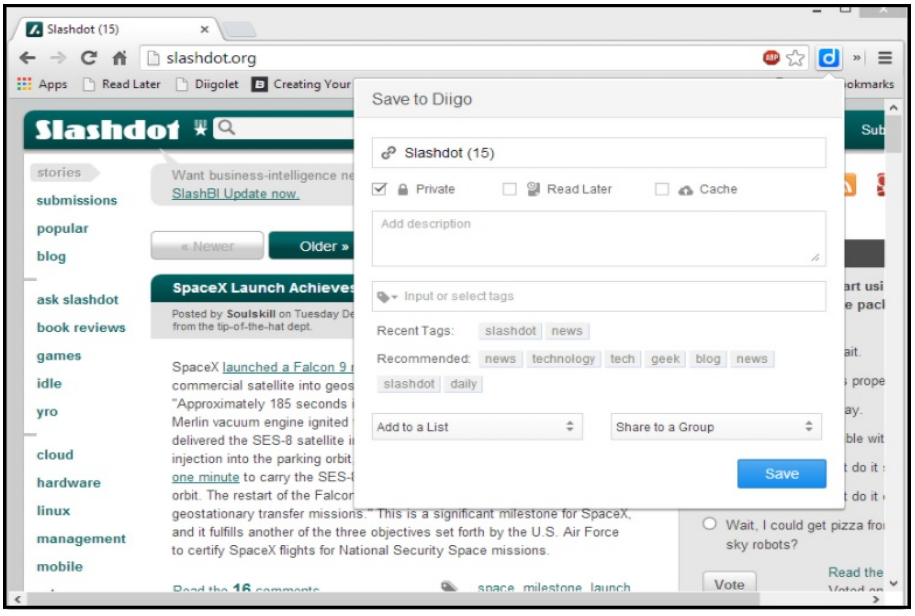


Figure 98: Bookmarking functionality options[40].

39 Available at https://getliner.com/, accessed in January 2021
40 Available at http://help.diigo.com/how-to-use-chrome-extension/bookmarking-chrome, accessed in January 2021

In addition, the user can make highlights on the web pages. This can be done through the pop-up that appears after selecting the desired text part or through the options that are accessible from the extension's icon in the browser, both cases shown in Figure 99. Note that the user has some options for managing highlights such as removing them and changing their color as shown in Figure 100.

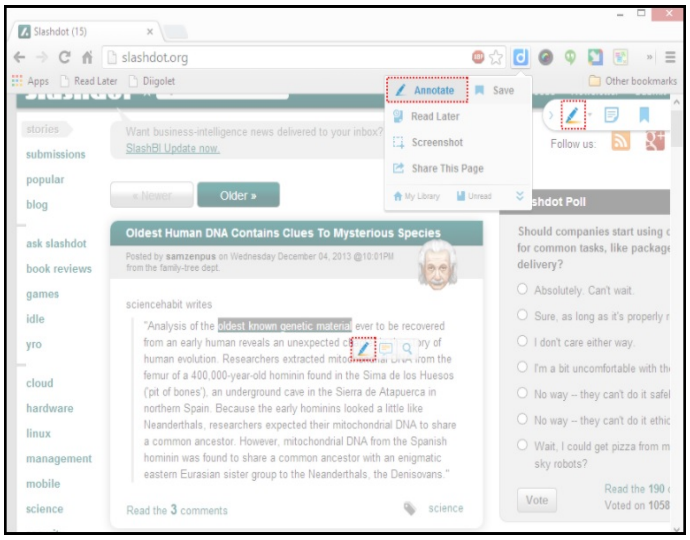


Figure 99: Ways to perform Highlights[41].

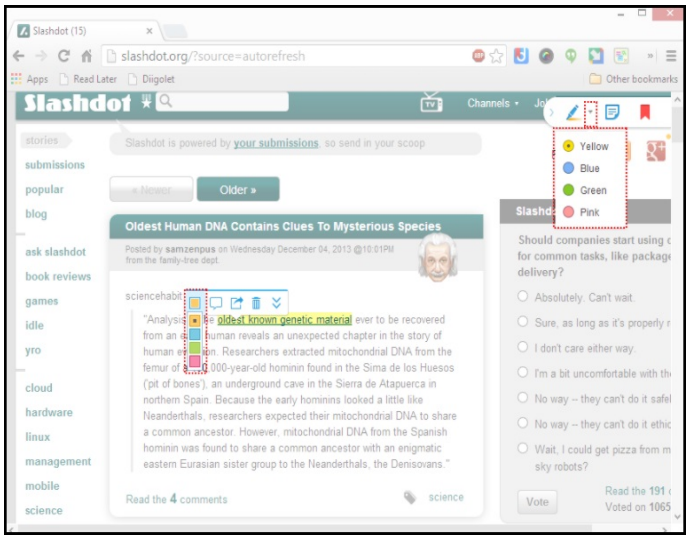


Figure 100: Highlights management options[41].

Another option that the user has when using the *Diigo* tool is the ability to annotate the highlights through sticky notes. This can be done by clicking on the highlight to be

41 Available at http://help.diigo.com/how-to-use-chrome-extension/highlighting-chrome, accessed in January 2021

annotated or through the options that are accessible from the extension's icon in the browser as shown in Figure 101. Note that these sticky notes have two types, one that is attached to a highlight and the other that can be placed anywhere on the web page, where in both cases it is possible to remove them and change their color.

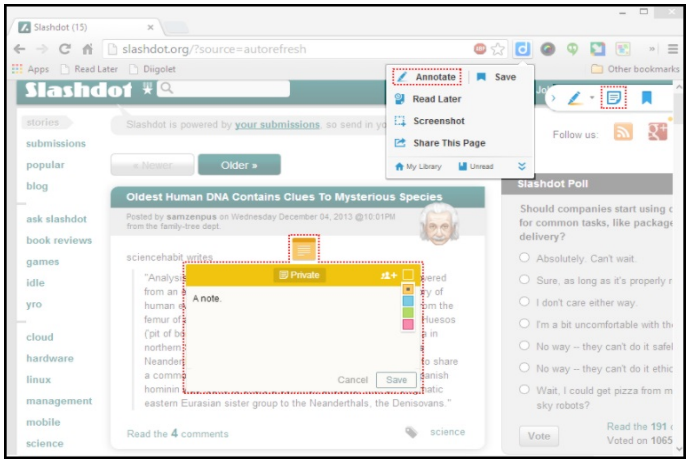


Figure 101: Options for creating a sticky note[42].

In addition to the traditional features already mentioned, it is important to mention the ability to save images. This can be done through the options that arise from performing a right-click on the desired image as shown in Figure 102 or through the options that are accessible from the extension's icon through which it will also be possible to mark-up the image as shown in Figure 103.

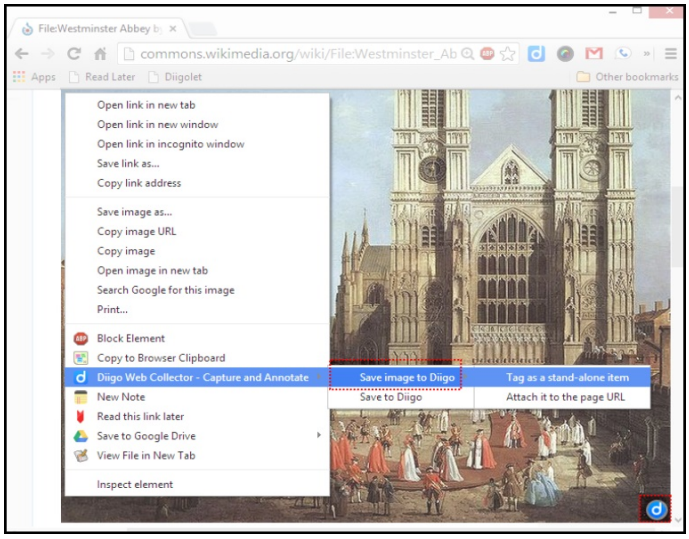


Figure 102: Options for creating a sticky note[43].

42 Available at http://help.diigo.com/how-to-use-chrome-extension/sticky-notes-chrome, accessed in January 2021

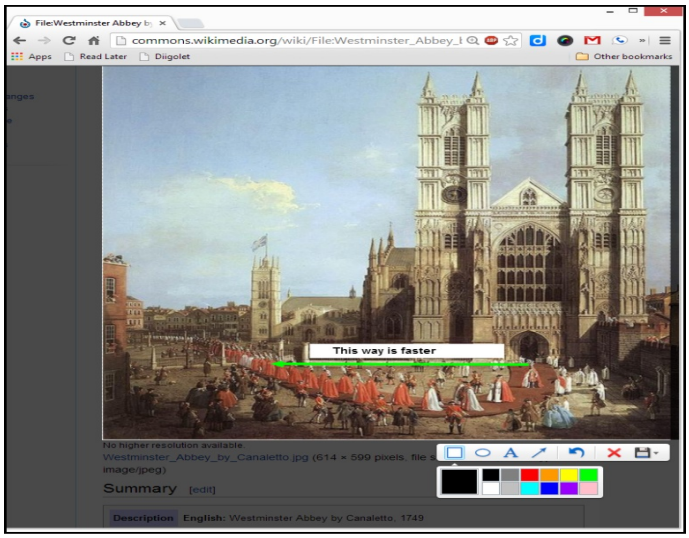


Figure 103: Options for creating a sticky note[43].

In order to manage and view the work performed the user will have at his disposal the *Diigo* web app shown in Figure 104.

43 Available at http://help.diigo.com/how-to-use-chrome-extension/collect-and-capture-images-chrome, accessed in January 2021

43 Available at http://help.diigo.com/how-to-use-chrome-extension/collect-and-capture-images-chrome, accessed in January 2021

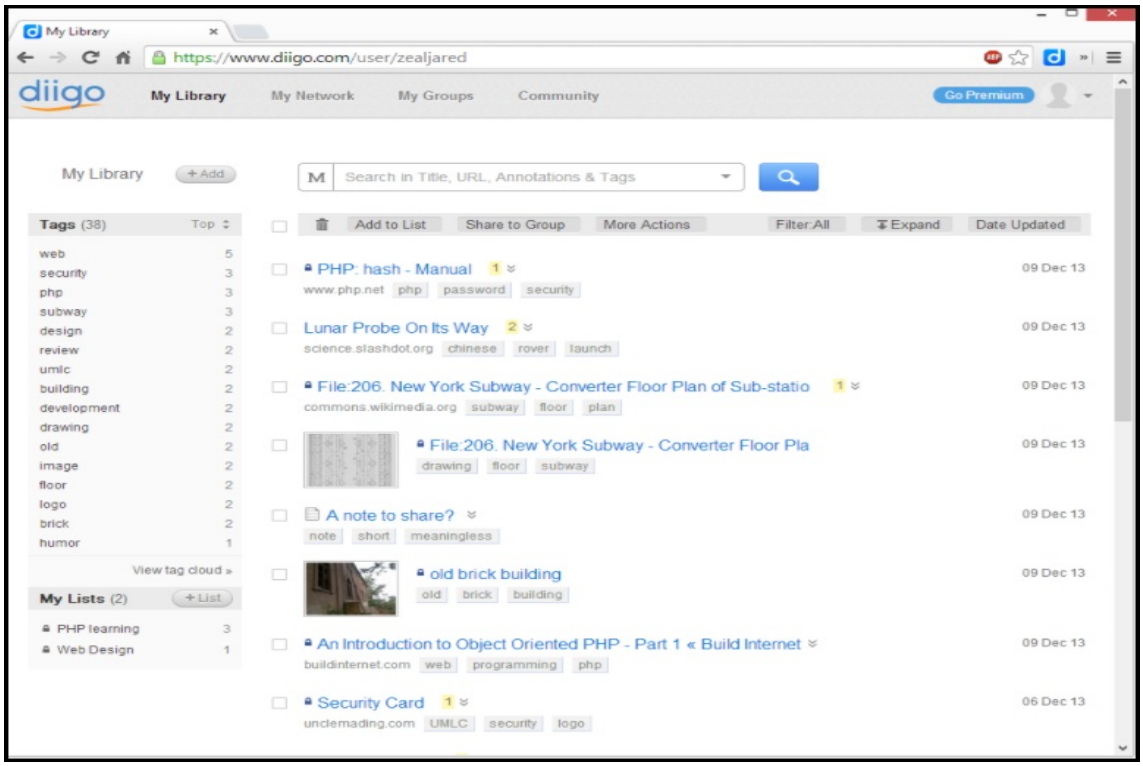


Figure 104: *Diigo* web app[44].

Through this platform, the user will be able to manage his work in several ways. One of them is by searching the pages saved through three methods, that is, through a word or phrase that is contained in the title of the web pages, in the URLs, in the tags or in the annotations, through the tags or looking for a word/phrase that is in the pages web. The user will also be able to sort the web pages he has saved and filter them through tags or their type. As already mentioned, it is also possible to create lists in which the saved web pages can be inserted in order to better organize the elaborated work. Through this platform it will also be possible to delete the web pages that have been saved, add more tags to them, export the work done or even continue it just by clicking on the link to be redirected to the original web page.

Diigo also supports collaborative work, this being achieved by marking the saved web page as public in the options shown in the Figure 98 or by creating/joining groups, shown in Figure 105, in which saved web pages can be shared with other users of the *Diigo* tool. In these groups the users are able to interact with each other and even highlight and annotate web pages together. Note that it is possible to manage everything that concerns the administrative part of these groups.

44 Available at `http://help.diigo.com/my-library-tutorial/general-tutorial`, accessed in January 2021

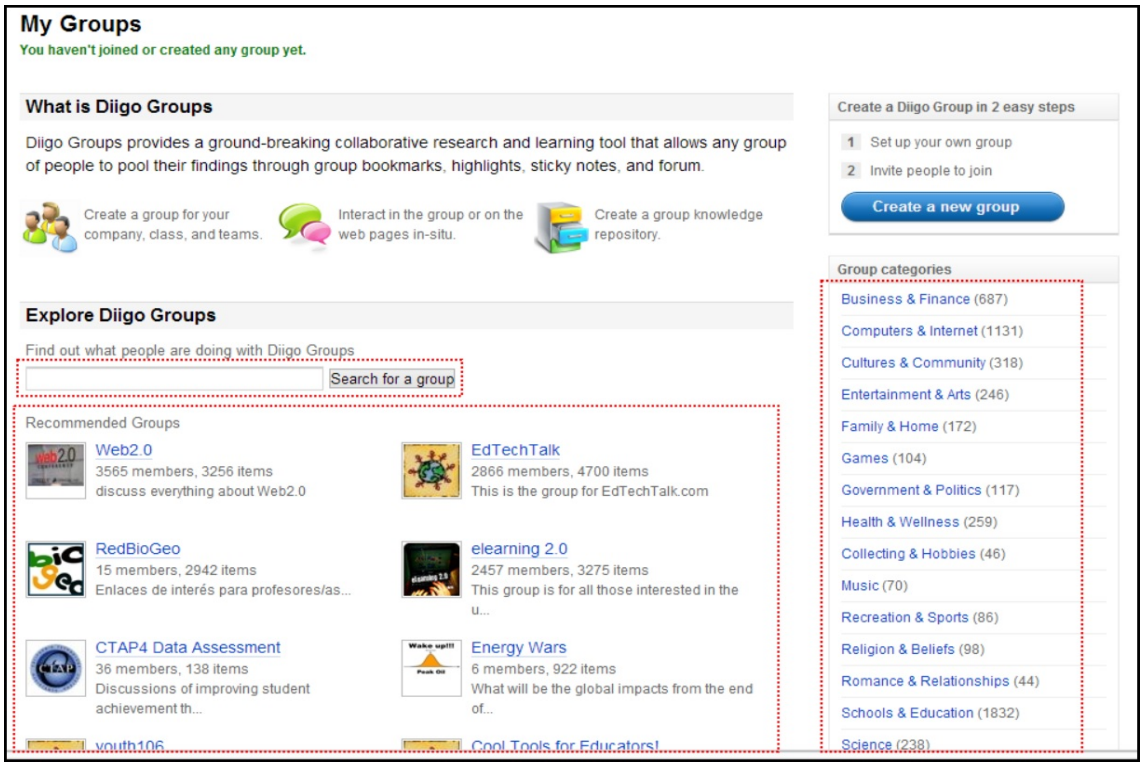


Figure 105: *Diigo* tool groups[45].

Taking into account all the features discussed above, it can be said that *Diigo* is an excellent web-based solution that allows you to make highlights and annotate web pages in a simple but effective way offering several ways to manage them. From this, it is possible to conclude that this tool fits with the objectives of the thesis, although it also blocks certain features and feature options from free users limiting them to some extent the use of the tool.

Hypothes.is

Finally, we have the ***Hypothes.is*** tool which, like other web annotators, also allows users to annotate web pages. To do this, the *Hypothes.is* tool uses the traditional highlighting and annotation methods that all other tools use being this through a pop-up, which is shown in Figure 106, that appears after the selection of the text to be annotated. In case the selection made contains a part of the text that has already been annotated, the number of annotations already existing in it will appear in the pop-up as shown in Figure 106 which can all be viewed by clicking on the button that contains the number.

45 Available at `http://help.diigo.com/my-groups-turorial/general-tutorial`, accessed in January 2021

Figure 106: Pop-up with highlight and annotation options[46].

In case the user wants to annotate, the text box in Figure 107 will be displayed in the comment can be written with the option of using formatting options more complex than plain text, being even possible to add video and images. The user can also associate tags to the annotations in order to organize them better and to be more easily found in the future.

Figure 107: Annotation text box[46].

In addition to being able to annotate parts of the text, the user can also make comments regarding the web page in general.

Regarding the visualization and management of the work performed, the user will have two options. One of them is through the extension menu of *Hypothes.is* present in the browser, visible in Figure 108, through which is only possible to visualize all the work done on the currently open web page. The other way will be through the *Hypothes.is* dashboard, in which it will be possible to view all the work done as shown in Figure 109.

46 Available at `https://web.hypothes.is/quick-start-guide-for-students/`, accessed in January 2021

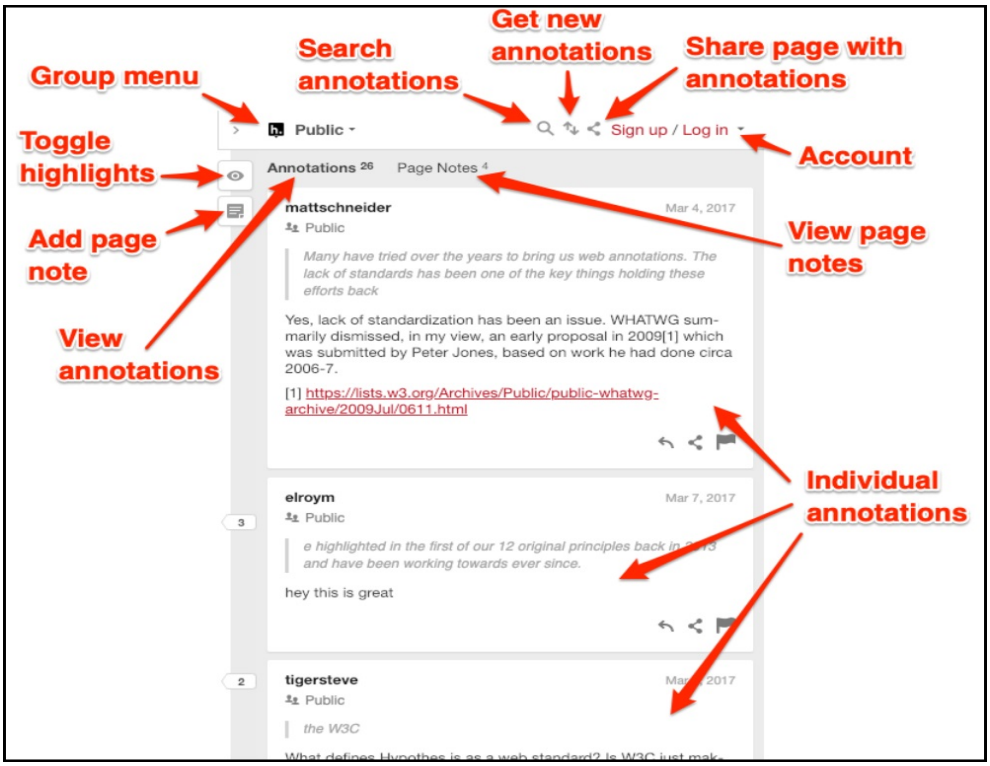


Figure 108: *Hypothes.is* extension menu[47].

47 Available at `https://web.hypothes.is/help/annotation-basics/`, accessed in January 2021

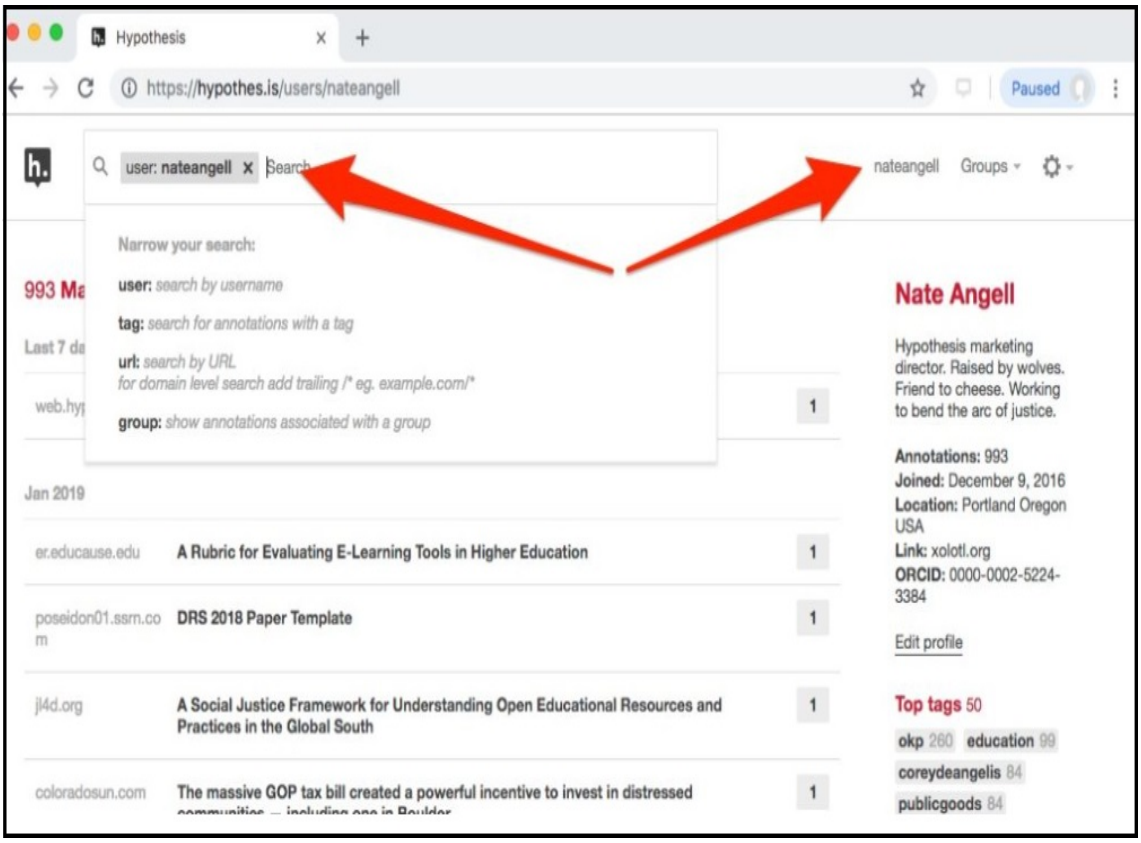


Figure 109: *Hypothes.is* dashboard[47].

In both scenarios the user will be able to remove and edit all of his highlights, annotations and comments related to the web page and search annotations through the search bar using the criteria visible in Figure 109.

A feature present in the *Hypothes.is* extension is the ability to automatically locate in the text the annotation that present in the list show in the Figure 108 when the user clicks on it.

On the other hand, a feature that only the dashboard in Figure 109 provides is the possibility to open the original web page thus allowing the user to view the notes in their original context and even continue to annotate it.

Like all the tools that have already been analyzed, *Hypothes.is* also allows collaborative work through which several people can interact with each other and annotate web pages together, being this done in several ways. One of them is by marking the annotation as public at the time of its creation, which can then be seen by anyone who visits the web page where it was inserted while also using *Hypothes.is*. Another way is through the sharing options through which it is possible to share a link that allows another person to have access to the annotated web page or annotation. A final method would involve creating or joining a group where multiple users can share their work with each other.

47 Available at `https://web.hypothes.is/help/annotation-basics/`, accessed in January 2021

Once again, taking into account all the mentioned features, it is possible to state that *Hypothes.is* is a simple and easy to use web-based annotation tool. Although it does not have a direct way to export the work done and no way to change the highlight color, it is the only web-based annotation tool from those that have been analyzed in this section that does not restrict its use depending on the type of user since there are no premium accounts. With all this in mind, we can conclude that the *Hypothes.is* tool fits the objectives of the thesis.

2.8 SUMMARY

Since Chapter 2 was quite long, the main topics covered in it will be summarized in this section.

The first theme that was explored was the history of annotations in Section 2.1, through which it was possible to observe that they evolved from a collective use that had the main objective of sharing and generating more knowledge to a mostly individual use that reflects the experience of the reader during his reading. However, collaborative use has not completely disappeared since it is predominant in digital contexts through the use of web annotation systems.

After that, several experiences carried out by different studies were analyzed in order to verify how readers annotate on paper and on the web, this being done in Sections 2.2 and 2.3.

In the case of paper annotation, it was verified that highlights were the most used type of annotation and the most common purpose was to remember, thus being able to relate these two since highlights are normally used to help in the memorization process and to make it easier to find the important parts of the text in a later reading. Another reason for this predominant use of highlights is because they allow the reader to stay focused on the task of reading since it is a method of quick execution compared to the other ones. Still regarding the highlights, another important characteristic is their colors, which may have additional meanings and can facilitate detection. In one of the reported experiments, it was also found that the purpose of reading greatly influences the way in which readers annotate their documents, namely in the types of annotations they use. A final observation is that the annotations with text that are shared are written in a more explicit way so that other readers understand them more easily.

Regarding annotation on the web, it is difficult to determine which types of annotations are most used since it depends on the features that the systems provide. That being said, in all cases, readers were careful to place the notes as close as possible to the respective parts of the document to which they referred. Another interesting observation is that, in cases where it was possible to highlight, this tended to be the most common choice, confirming the popularity of this type of annotation. In this context, it was discovered that they had

the purpose of signaling parts of the text that were not understood, that they wanted to remember or because they were important. As in the paper case, here the colors of the highlights are also of great importance, allowing readers to better structure their annotations. Regarding the notes that were made with the purpose of being shared, it was also possible to verify the same situation that was described in the paper case. These are more developed and explicit than the private notes, which are shorter and ambiguous. Through this, it is possible to conclude that the readers when writing private notes are only concerned with their significance to themselves while in the case of shared notes they are written so that the other readers have no problems in understanding them.

Through the observations made in Sections 2.2 and 2.3 and through other articles it was possible to identify the various purposes and types of annotations in a general way and the advantages and disadvantages of digital annotation compared to physical annotation. However, since the respective outcomes were summarized in Sections 2.4, 2.5 and 2.6, it does not make sense to mention them again.

To conclude this chapter, in Section 2.7, several tools that have annotation mechanisms were analyzed. For its choice, three criteria were used, such as its compatibility with what is intended to be achieved with this thesis, its popularity and whether it has interesting characteristics. In order to sum up the lessons learned and to make easy to compare those tools concerning the most relevant feature from the perspective of this Master´s project, Table 1 was built.

Table 1: Tool comparison.

Tool	Annotation of web pages	Free	Multiple forms of annotation	Text formatting	Search facilities	Export and import
doccano	No	Yes	Limited	No	No	Yes
Word	No	No	No	Yes	Limited	No
Adobe Reader	No	Yes	Yes	Limited	Yes	Yes
Weava	Yes	Limited	No	No	Yes	No
LINER	Yes	Limited	No	No	Yes	No
Diigo	Yes	Limited	Limited	No	Yes	No
Hypothes.is	Yes	Yes	No	Yes	Yes	No

Observing Table 1 it is possible to conclude that, although there is no perfect tool, the most complete is the *Adobe Reader* since it is the one that satisfies most of the requirements in a satisfactory way. However, it does not fulfill the most important requirement identified in the context of *NetLang* community to annotate web pages.

Thus, if we look at the tools that fulfill this requirement, the best one would be *Hypothes.is* due to fulfilling the greatest number of the remaining requirements. Another factor that should be mentioned is that this tool is the only one of the four web based tools that have been analyzed here that does not restrict some of its functions through free and paid accounts. The only negative aspects of this tool is that it only allows the use of highlights, to which it is possible to associate comments, as a form of annotation and does not have any mechanism that allow to export the work done so that it can be imported later to continue it.

To conclude this summary, it is important to mention that the characteristics of the tools that were analyzed served as inspiration for the features of the developed editor. These include the *Adobe Reader* interface, import/export mechanism, filtering options and method of handling the overlap of annotations and the *Hypothes.is* formatting options and location discovery mechanism.

3

NETLANGED, THE REQUIREMENTS AND ITS ARCHITECTURE

In this chapter, the Master's project is described in detail, starting with a list that contains the features that the editor to be developed must have, and followed by the diagrams of the system architecture.

3.1 REQUIREMENTS

Based on the information collected during the literature review carried out to prepare the *State of the Art* presented in Chapter 2, the features that the editor should provide are defined below.

Functional properties

Below are listed the functional properties that NetLanEd must offer.

- **Marks on the text:** The part of the text that the user notes should be marked in some way.
- **Highlight color options:** The user must be able to customize the color of the highlight.
- **Tooltips in annotated text:** When hovering the cursor over the annotated text, a tooltip containing at least the respective comment must be presented.
- **Text formatting:** The user must be able to format the comment text in different ways, such as changing the font size, changing the font family, creating lists, etc.
- **Annotated comments list:** The user must be able to see all comments on the annotations made in the document he is analyzing.
- **Removal:** There should be options that allow the removal of annotations and their comments.

- **Editing:** As with removals, there should also be options that allow editing annotation comments.
- **Location discovery:** There must be a mechanism that allows a quick location of the annotation of the respective comment in the document.
- **Search:** The user should be able to search a word/phrase in the comments of the annotations.
- **Filter:** The user must be able to filter comments according to various criteria.
- **Import the work done:** The user must be able to import from his computer the status, that were previously saved, of the document that is currently being analyzed.
- **Export the work done:** The user must be able to export to his computer the current status of the document.
- **Export the notes made:** The user must be able to export the comments of the annotations that he made in the document to his computer.
- **Clear the document:** The user must be able to remove all annotations in the document at once.

Non-functional properties

Here, the non-functional properties that *NetLangEd* must provide are listed below.

- **Simple and clear interface:** The interface should not take up too much space so as not to distract the user and its content must be simple and explicit so that the user does not feel confused when using it.
- **Simple and clear functionalities:** The functionalities must be easy to understand and to execute so that the user does not have difficulties in using the tool.
- **Quick add functionality:** This functionality should be quick to perform, being done in the smallest number of steps possible.
- **Quick edit functionality:** Editing annotation comments should be possible both in the annotation comments list and in the document, thus allowing the user to remove them in any situation. This functionality should also be quick to perform, being done in the smallest number of steps possible.
- **Quick removal functionality:** As in the case of editing, the option to remove comments from annotations should be possible to execute both in the comment list of annotations

and in the document, thus allowing the user to remove them in any situation. This functionality should also be quick to perform, being done in the smallest number of steps possible.

- **Annotation overlay:** The parts of the text where annotations overlap should be properly treated so that the annotations involved can be easily distinguished.
- **Annotation comment representation:** An annotation comment must be presented in the same way both in the annotation comment list and in the document's tooltips.

3.2 System Architecture

In this section, several diagrams are presented to better understand the architecture of the system. Thus, through Figure 110 is possible to have a better idea of how the web tool developed in this Master's project, *NetLangEd*, is integrated into the *NetLang* Corpus Search Engine.

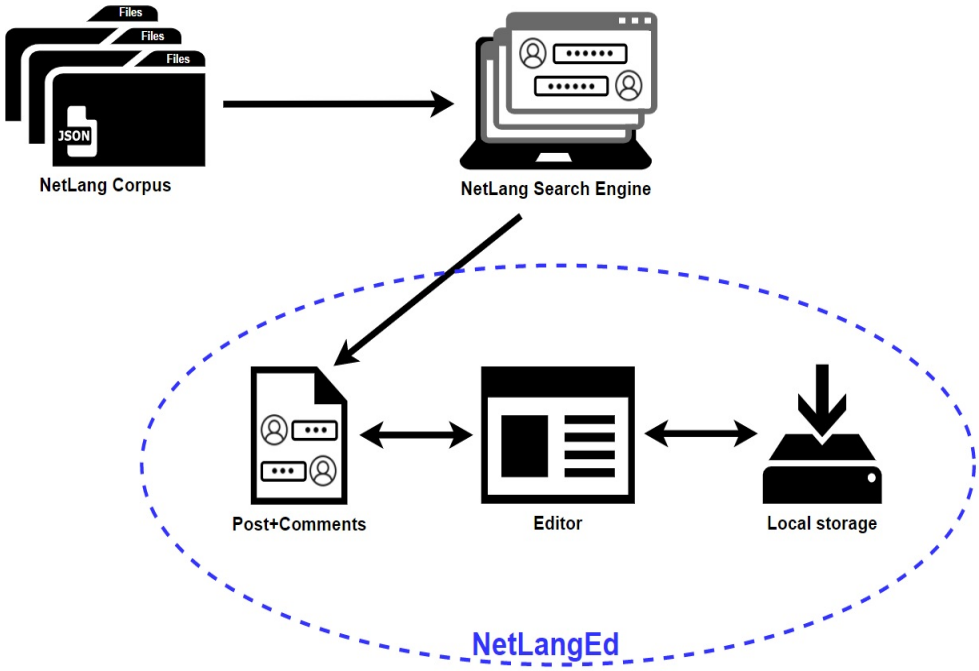


Figure 110: *NetLangEd* Architecture and Integration with *NetLang* platform.

Looking at Figure 110 it is possible to see that the editor will be accessed from the pages that contain the posts and comments that are stored in the *NetLang* repository. These pages contain a button, in the upper right corner, called "Annotate" that opens the NetLangEd editor that receives the HTML of these pages as input. Another aspect that can be seen in Figure 110 is the possibility of exporting and importing the work done, these being done

to and from the user's computer respectively. It is important to mention that the work produced by each user is completely personal and cannot be seen by other users of the platform.

Finally, Figure 111 presents how the functionalities are organized in the system and how they can be accessed.

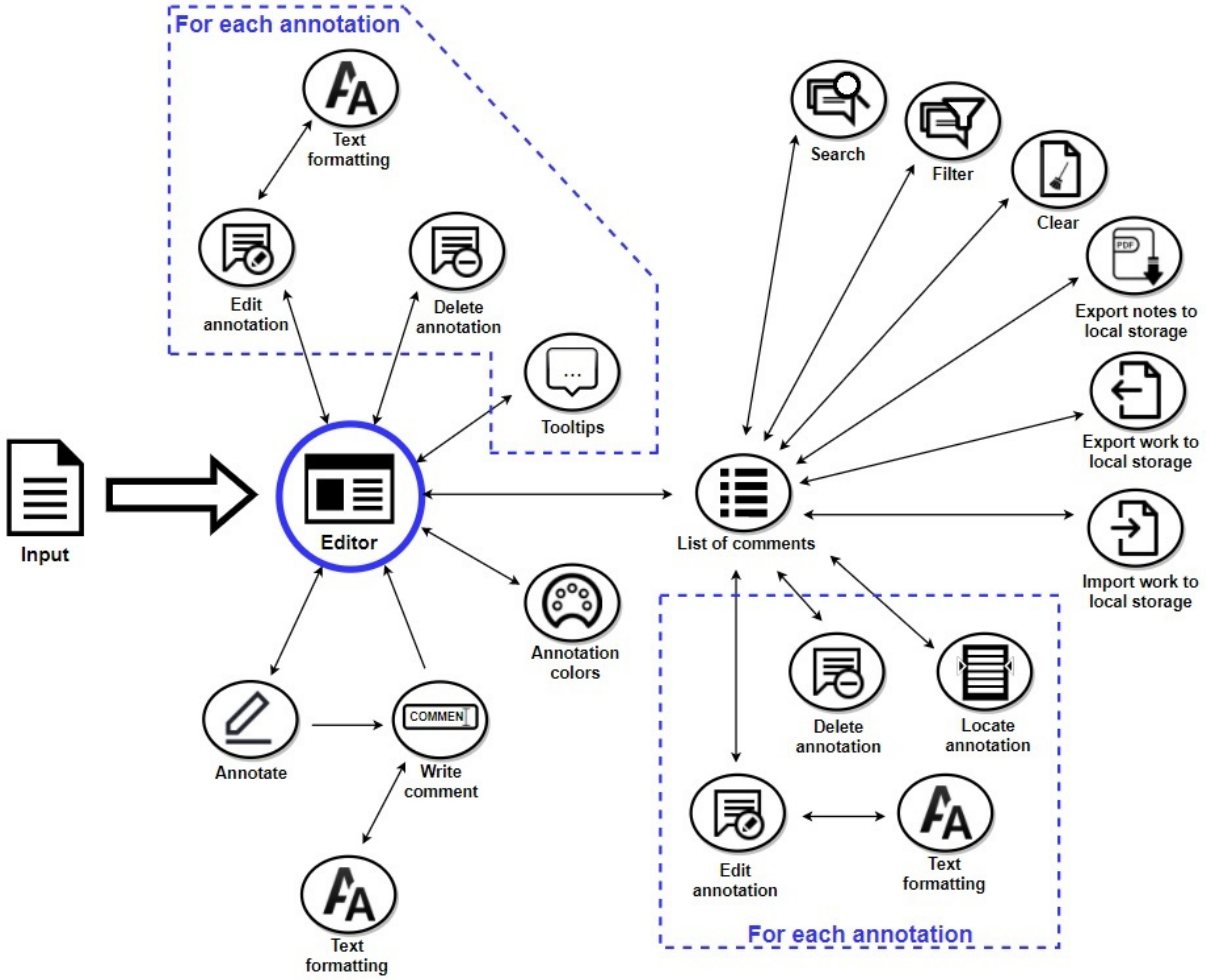


Figure 111: *NetLangEd* usage diagram.

Figure 111 shows the functionalities that can be performed in the overall document and those that can be performed in each of the annotations present in it. This cases are represented in the "Editor" side of the diagram.

In Figure 111 it is also possible to see the menu, that is represented by the "List of comments", which has the list of all comments made in the document. In addition off having several functionalities that allow their management and filtration, also has some that can be performed in each of them.

Although the editor will be integrated into the *NetLang* platform, it is not exclusive to it, as it can easily be used in different scenarios.

3.3 SUMMARY

As a recap, this chapter listed the requirements for the editor, more specifically the functional and non-functional ones. After the enumeration of the desired features for *NetLanEd*, the system architecture is presented depicting how the editor is integrated into the *NetLang* platform. Also a diagram is shown to explain how the described features are organized in the system.

4

NETLANGED, THE EDITOR AND ITS DEVELOPMENT

This chapter is used to describe the implementation of the work done in this Master's project. To this end, it is described how each of the requirements that were listed in Section 3.1 were achieved and how the integration process with the *NetLang* project carried out.

It is important to mention that the *JavaScript* language was used to develop the editor along with the *jQuery* and *Rangy* libraries and the *xcolor* plugin.

4.1 FUNCTIONAL PROPERTIES

In this section, it is explained how the functional requirements were implemented in the final work.

4.1.1 *Marks on the text*

This requirement, that consists of allowing users to annotate HTML text without affecting its original formatting, was the most important objective to be achieved in this project.

The initial approach was to surround the selection made by the user with span tags, but the implementation of this approach failed in some cases. The first one occurred when removing an annotation that was overlapped by another one. In this case, removing one of the annotations would also removed the part of the other annotation that overlapped it. The second occurs when removing a simple annotation that contains content that does not cover the closing tags. Here, removing the annotation would cause that part of the HTML to be broken.

Taking these problems into account, the only solution that avoided most of them was to surround each word, with the exception of HTML tags, with a *span* tag, after which it was decided to surround each character in order to be able to annotate any part of the words instead of being forced to annotate it entirely. These *span* tags have a unique identifier in order to be able to change their properties more easily, and for that it was necessary to

develop a function that returned the identifiers of the *spans* that are present in the selection made by the user.

It should be noted that this approach is not trouble-free, being dependent on the text to be annotated to remain unchanged so that the identifiers remain constant. Other problems arose during the integration phase with the *NetLang* platform as will be discussed in Section 4.3.

Having said that, the adding process begins with the user selecting the part of the text that he wants to annotate followed by clicking on the pop-up that appears after completing the selection, in order to confirm the intention to annotate the selected part of the text. After that, the addition box will appear in the sidebar in which the user can enter the comment, with the possibility of associating tags, completing the process using the button to save the comment. This whole process can be seen in Figure 112.

These notes are completely private, being impossible to be seen by other users of the platform when analyzing the same document. The decision to only allow private notes came from the fact that the users of the *NetLang* platform were not very keen on sharing their notes with other users.

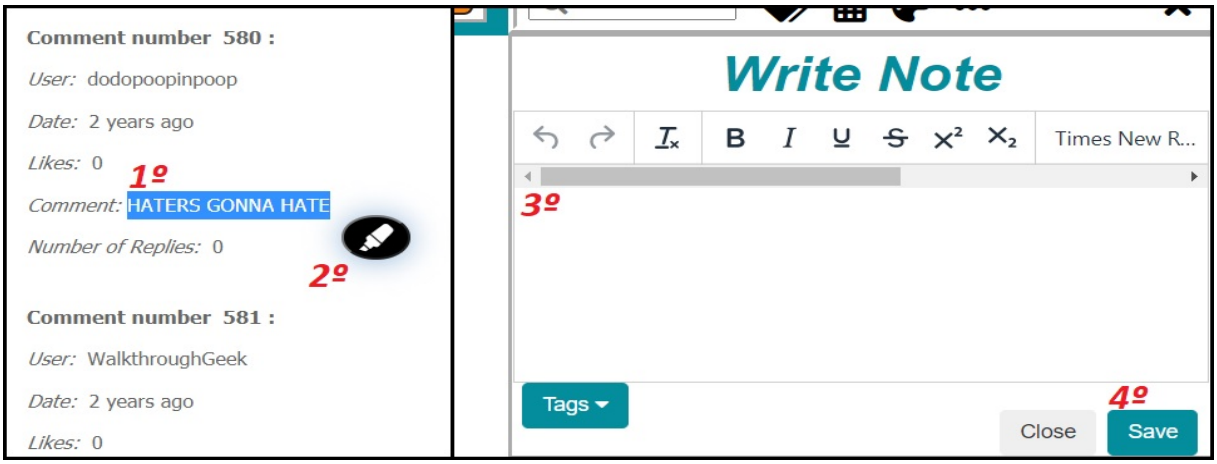


Figure 112: Text annotation functionality.

It is important to note that when writing comments, users will not be able to use the characters "<", ">" and "&" and will not be able to perform the paste operation in order to prevent them from writing malicious code and interfere with the proper functioning of the editor.

4.1.2 *Highlight color options*

This feature was implemented as a palette with the colors that are available, as can be seen in Figure 113.

Initially it was possible to pick any color, however this method would complicate not only the process of filtering the annotations by their color because it would lead to too many options but it would also complicate the process of picking a color since the user would have to either find the desired color or would have to save the color he picked for later use.

Figure 113: Color palette.

4.1.3 *Tooltips in annotated text*

This feature can be seen whenever the user places the mouse cursor on top of an annotation presenting the result visible in Figure 114.

In addition to containing the comment that is associated with the annotation, the number of the note is also shown to make it possible to identify it later easily.

However, there are some restrictions regarding the display of comments in the tooltips. In order to make sure that the tooltip is always visible, it can only be shown up to six lines of the comment. In cases where there is more content to show, an ellipsis is displayed.

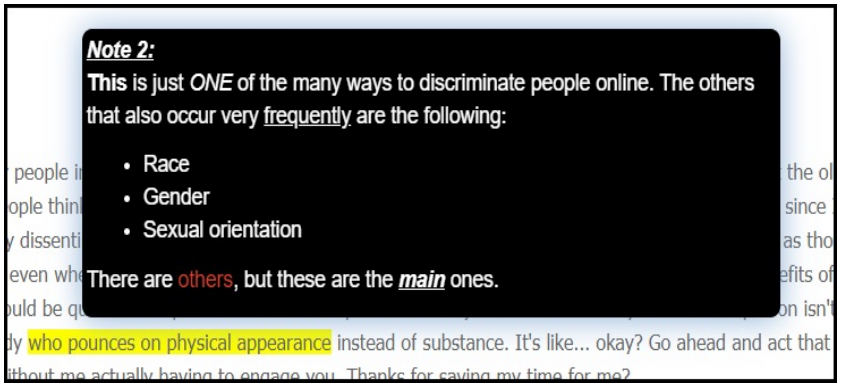


Figure 114: Example of a tooltip.

4.1.4 *Text formatting*

Concerning text formatting, there are many options that can be included, as *undo*, *redo*, *clear formatting*, *bold*, *italic*, *underline*, *strikethrough*, *superscript*, *subscript*, *fonts*, *font sizes*, *numbered list*, *bullet list*, *text color*, *background color*. From a technical point of view, there were others that could be included. However only fifteen (as can be seen in Figure 115) were implemented because the others were difficult to display in the tooltips.

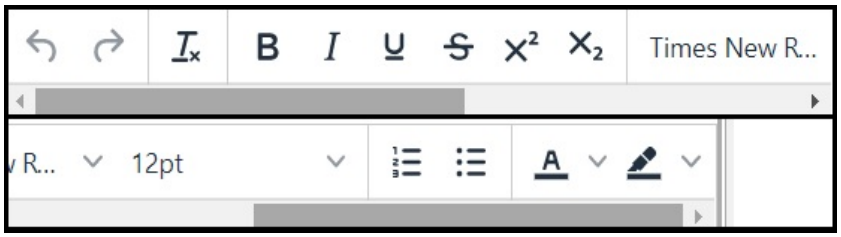


Figure 115: Formatting options.

4.1.5 *Annotated comments list*

As can be seen in Figure 116, all the annotations that were made on the document that is currently being analysed are listed on a sidebar.

Another important aspect of this sidebar was its simplicity, clarity and convenience. To this end, it was decided to keep at the top of it all the options that are used on a recurring basis like the filters, and group in one button all the options that would not be used as regularly, like exporting, importing, etc. Another decision made was to place all the options that work on comments (Remove, Edit, Locate) on top of each one of them. Finally, it was considered useful to include information in each note regarding its number, the date of its creation, the number of the page on which the annotation was made and the tags that were associated to it.

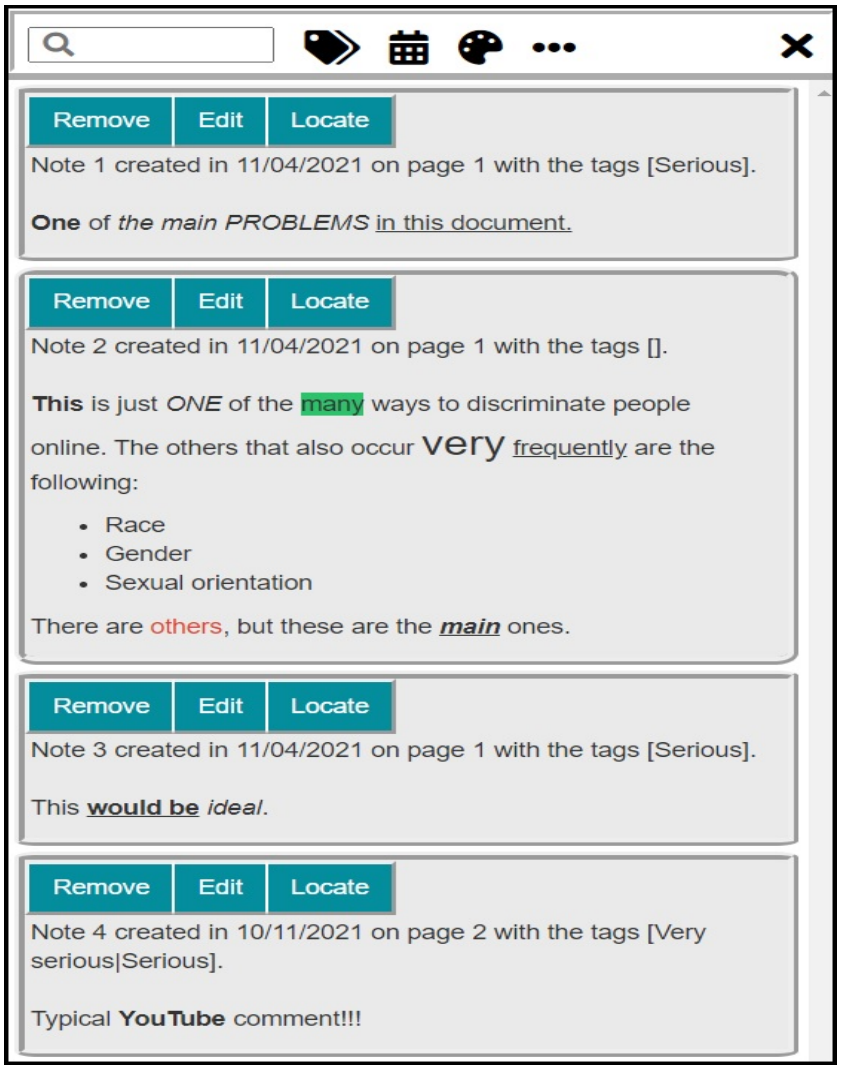


Figure 116: List of comments in the sidebar.

4.1.6 *Removal*

This requirement has been implemented and can be executed in two ways. One is performed by clicking on the desired annotation, that is present in the document, showing the edition box that has the button that allows its removal. The other way is through the removal button that is located over the annotations present in the sidebar. Both alternatives can be seen in Figures 117 and 118.

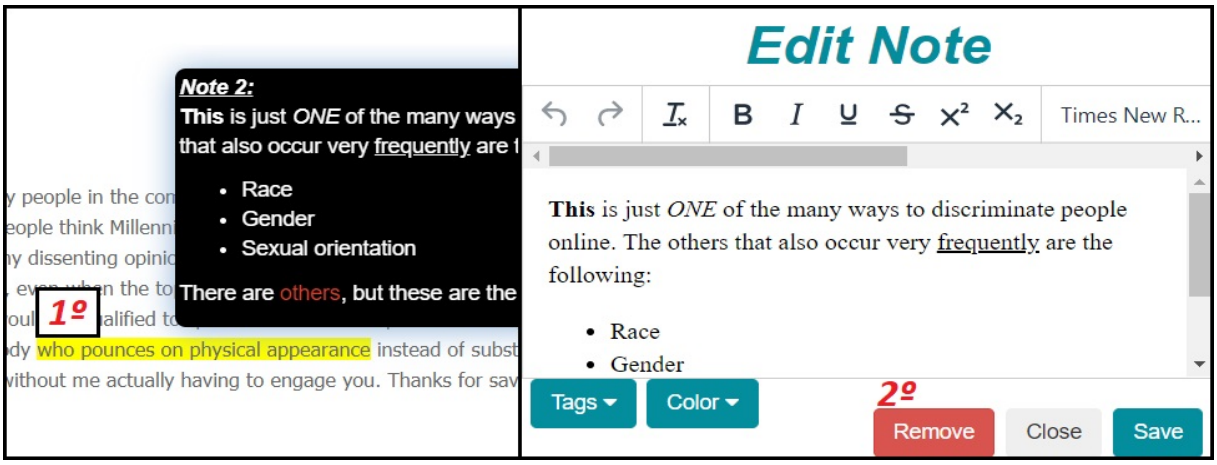


Figure 117: Annotation removal functionality from the document.

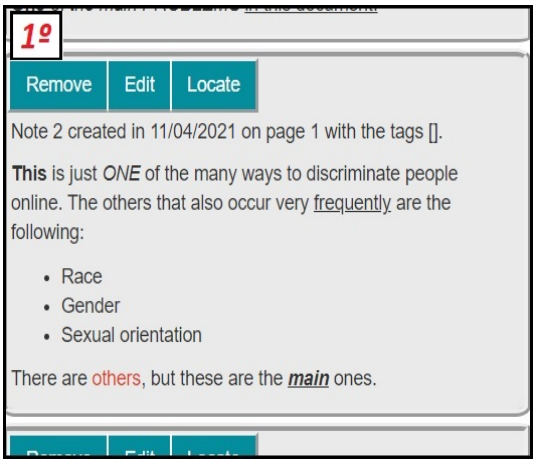


Figure 118: Annotation removal functionality from the sidebar.

Note that, in both ways, the text box will be precariously filled with the current comment of the respective annotation so that the user can make the decision to remove the annotation safely.

4.1.7 *Editing*

As in the previous *functional property*, the editing process can also be done in two ways. The first consists of clicking on the annotation, present in the document, whose comment is to be edited. The second way is through the button that allows editing that is located on top of each of the annotations present in the sidebar. In both cases the result of the clicks is the display of the edit box where the user can edit the comment, as well as the tags and color associated with it, concluding the process by clicking in the button to save the comment. Both ways can be seen in Figures 119 and 120.

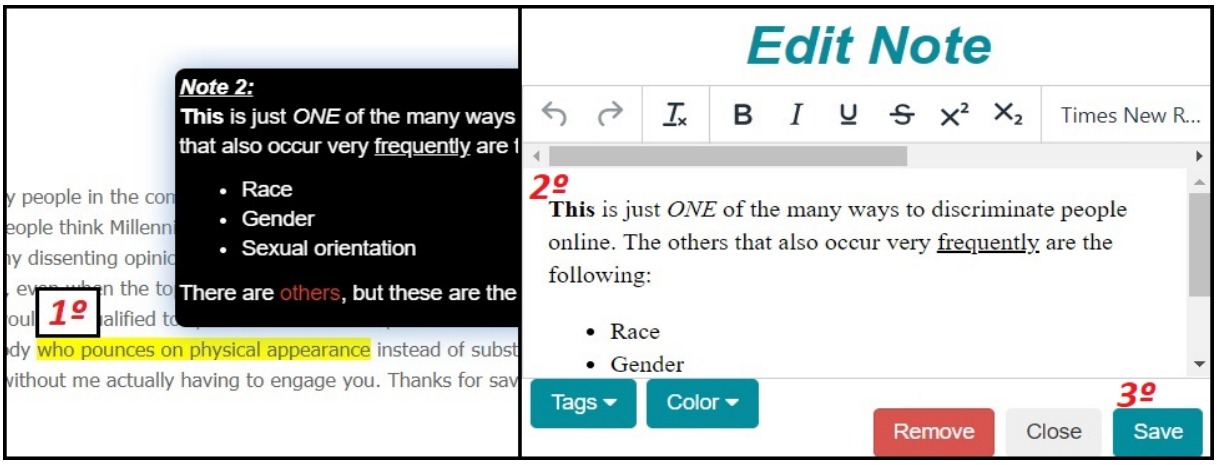


Figure 119: Annotation editing functionality from the document.

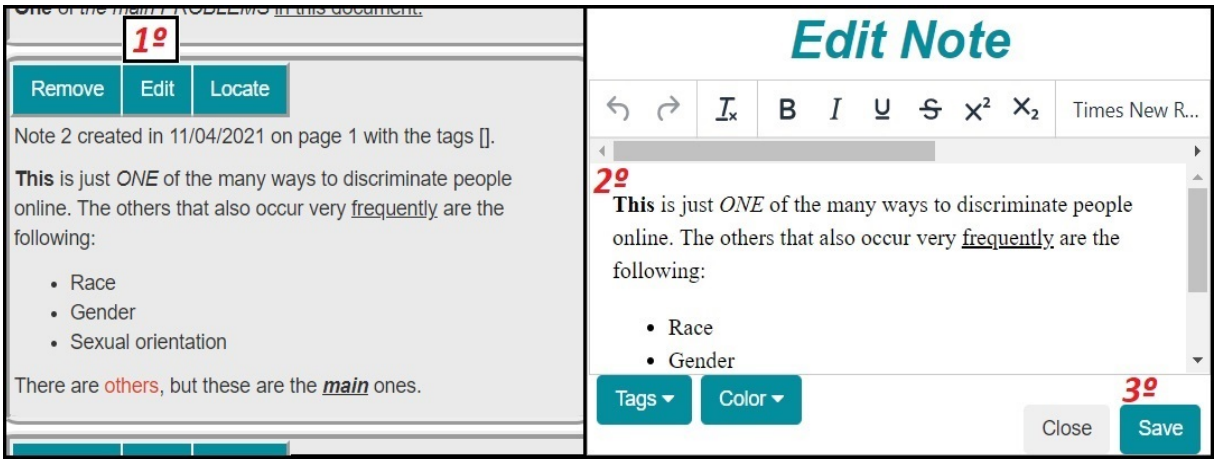


Figure 120: Annotation editing functionality from the sidebar.

It should be noted that, in this case, the text box will be precariously filled with the current comment of the respective annotation, so that the user does not have to rewrite the entire text when he only wants to make small changes. As in the case of adding annotations, in this one, users are also unable to write the characters "<", ">" and "&" and are unable to perform the paste operation in order to prevent them from writing malicious code and interfere with the proper functioning of the editor.

4.1.8 *Location discovery*

The purpose of this feature is to help the user to easily locate the comment in the document, being implemented through the button at the top of each of the comments present in the sidebar visible in Figure 116.

Initially, this feature only scrolled the page to the place where the annotation was located, making it the only visible annotation in order to be more easily detectable. However, after making the decision to display the documents in pages, due to the problems that were detected in the process of integrating the editor with the *NetLang* platform and which will be discussed later, this functionality now has the ability to switch to the page where the annotation is located.

4.1.9 *Search*

This feature is present at the top of the sidebar that can be seen in Figure 121. It is important to mention that only the annotations whose comments contain exactly what is entered by the user or a substring of it will be filtered, that is, the search takes into account upper and lower case letters, blank spaces, etc. It is also important to note that the result of applying the filter is visible both in the document and in the sidebar.

Figure 121: Search bar.

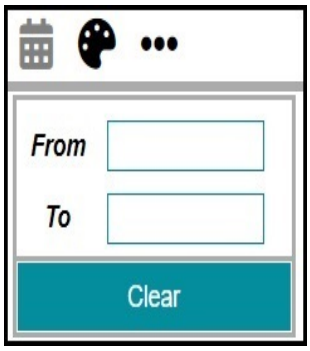
4.1.10 *Filter*

This requirement was implemented through the three buttons that are to the right of the search bar present at the top of the sidebar visible in Figure 122, in order to allow filtering the annotations by tags, date and color. In addition to these three types of filters, others were discussed, however these seemed the most useful.

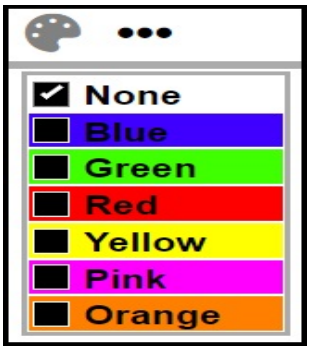
Note that these three filters can be used together, including with the search functionality. That said, it was necessary to make another decision on how the filters will act together, that is, if it is enough that the annotations comply with one of the filters or if they have to comply with all of them. After some reflection and taking into account that the purpose of filtering is to specify as much as possible a characteristic of something, it was decided that the annotations have to comply with all the filters to be shown (the logical operator 'and' is used to connect the parts).

Another characteristic that is worth mentioning is that, as in the previous requirement, the result of applying the filters is visible both in the document as in the sidebar.

(a) Filter by tags.



(b) Filter by date.



(c) Filter by color.

Figure 122: Filter options.

Regarding the management of the tags, this will be done through the button visible in Figure 122a, through which it will be possible to open the pop-up menu shown in Figure 123 where the user will be able to create, edit and remove tags.

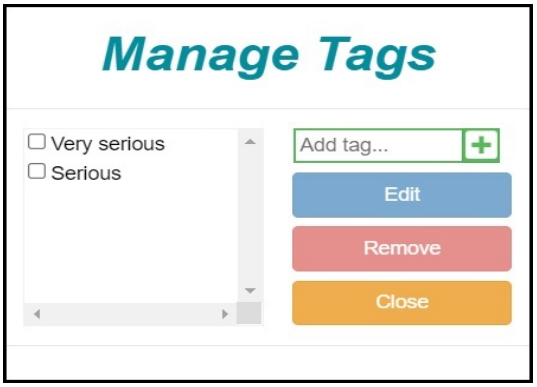


Figure 123: Tags manager.

4.1.11 *Import the work done*

This requirement was implemented as a feature that is present in the dropdown that appears clicking on the button represented by the three dots visible in Figure 124.

Implementing this functionality a security mechanism was applied to prevent the user from being able to write malicious HTML code in the part of the comments that would later be run when loading the file. Although it works and prevents the import of a file that has been changed, the user can access the "keyword" and replicate the hash process in order to overcome this barrier. However, it can be said that at least this process cannot be carried out in such an easy way.

Figure 124: More options window.

4.1.12 *Export the work done*

This functionality also appears in the dropdown that results from clicking on the button that is represented by the three dots visible in Figure 124.

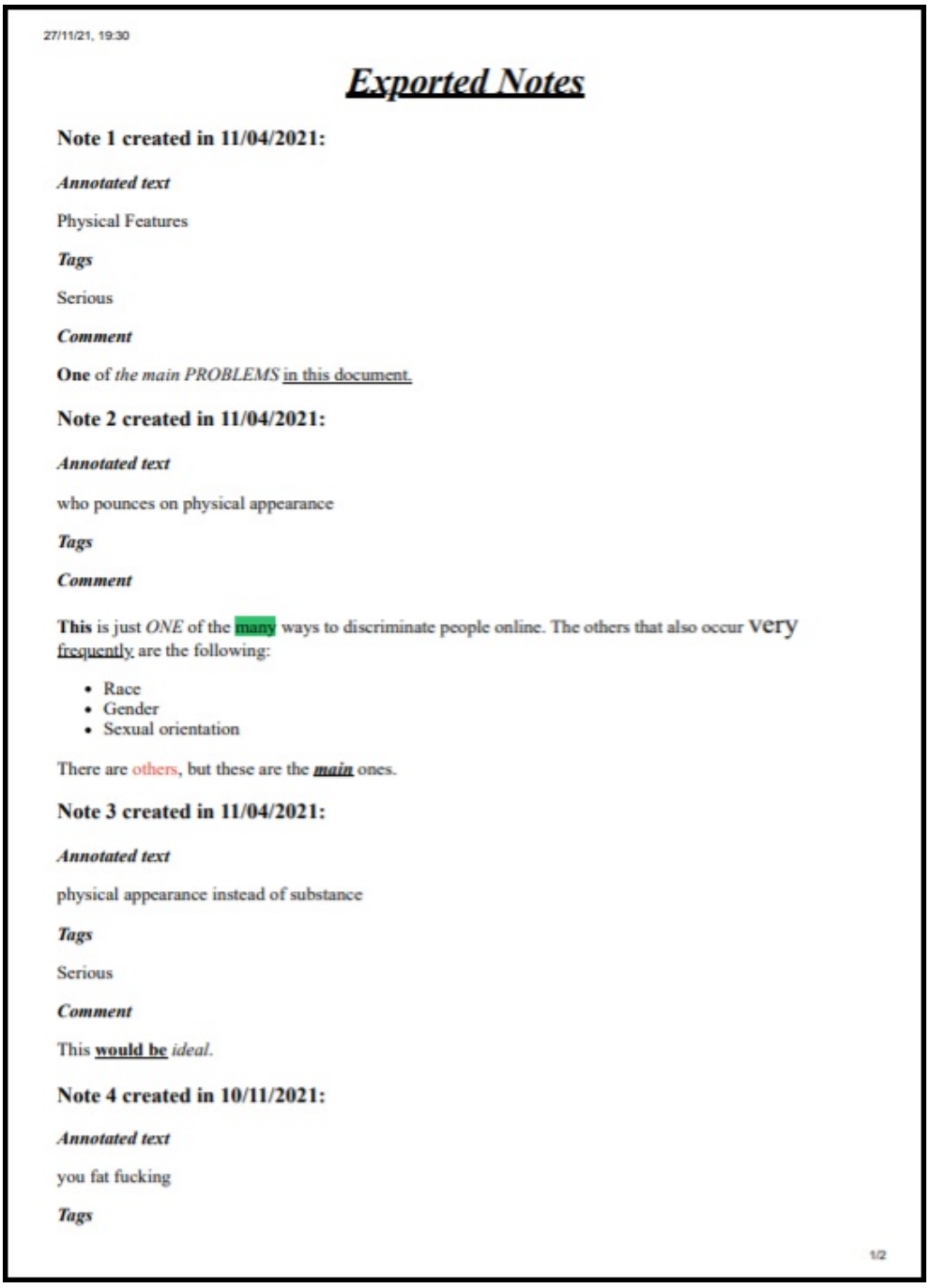
Export option writes in a text file the necessary information so that the user can restart the work, in the next session, in the point where it was left. It is important to mention that the user is in charge of identifying the file so that afterwards he can find out which document it belongs to, since loading the wrong file will lead to the editor's malfunction. Note that the user must click on the *save button* to save the current status of the work done, before exporting it. Otherwise, the work will be lost.

4.1.13 *Export the notes made*

This requirement was implemented as one of the options that is present in the window that results from clicking on the button that is represented by the three dots visible in Figure 124.

In order to implement this functionality, some libraries were initially used to generate *PDFs*, however, all of them presented problems related to situations where some formatting was ignored or, in cases where the *PDF* was composed of images of the notes, these were not only slightly blurred but also did not allow text selection. In view of this scenario, the solution involved generating a new web page containing only the user's notes followed by the use of the print functionality that browsers have to save it in the PDF format.

The generated *PDF* will contain all the annotations that were made in the document, where for each one of them, the text that was selected from the document, the tags that were selected, the written comment, the number of the annotation and the date of its creation will be shown. All this can be seen in Figure 125.



27/11/21, 19:30

Exported Notes

Note 1 created in 11/04/2021:

Annotated text

Physical Features

Tags

Serious

Comment

One of *the main PROBLEMS* in this document.

Note 2 created in 11/04/2021:

Annotated text

who pounces on physical appearance

Tags

Comment

This is just *ONE* of the many ways to discriminate people online. The others that also occur very frequently are the following:

- Race
- Gender
- Sexual orientation

There are others, but these are the ***main*** ones.

Note 3 created in 11/04/2021:

Annotated text

physical appearance instead of substance

Tags

Serious

Comment

This **would be** *ideal*.

Note 4 created in 10/11/2021:

Annotated text

you fat fucking

Tags

1/2

Figure 125: PDF with exported notes.

Note that, also this second export functionality takes into account the last saved status of the document. Therefore, the user must click on the *save button* to save the current status of the work done before exporting it to a PDF file.

4.1.14 *Clear the document*

This feature is also present in the drop-down that results from clicking on the button that is represented by the three dots visible in Figure 124.

This functionality simply removes all annotations made in the document as well as in the sidebar.

4.2 NON-FUNCTIONAL PROPERTIES

In this section, it is explained how the non-functional requirements were implemented in *NetLangEd* tool.

4.2.1 *Simple and clear interface*

This requirement was achieved through several variables that were taken into account in the development of the interface.

The first one is related to the way in which the buttons were placed at the top of the sidebar. Those that will probably be used more frequently were displayed explicitly. Those that are used less frequently were grouped in a dropdown menu accessible trough the three dots special button, as can be seen in Figure 116.

Another important aspect is related to the comments of the annotations present in the sidebar. These were embedded inside compartments to better distinguish where a comment starts and ends. Another important decision was to place the operations that can be executed on a comment directly over it in order to be clear which comment is applying the operations on. All these decisions are shown in Figure 116.

The placement of the color palette was also subject of reflection. It was initially planed to be placed in the sidebar, however since it is a tool that the user may want to use at any time, it was decided to keep it fixed next to the sidebar as shown in Figure 126.

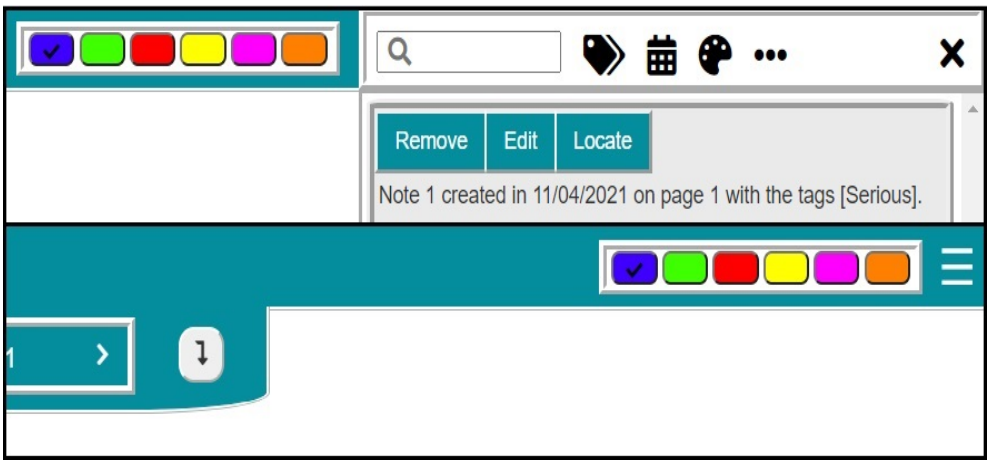


Figure 126: Color palette with and without sidebar.

Regarding the boxes that appear to edit or add a comment, it initially appeared in the center of the screen with the capability to be moved to any other part of the screen. However, since this task of always moving the box is boring and as it is preferable to always be able to see the part of the text that is being commented for contextualization, it was decided that in both cases these boxes would be displayed in the sidebar, as can be seen in Figure 127.

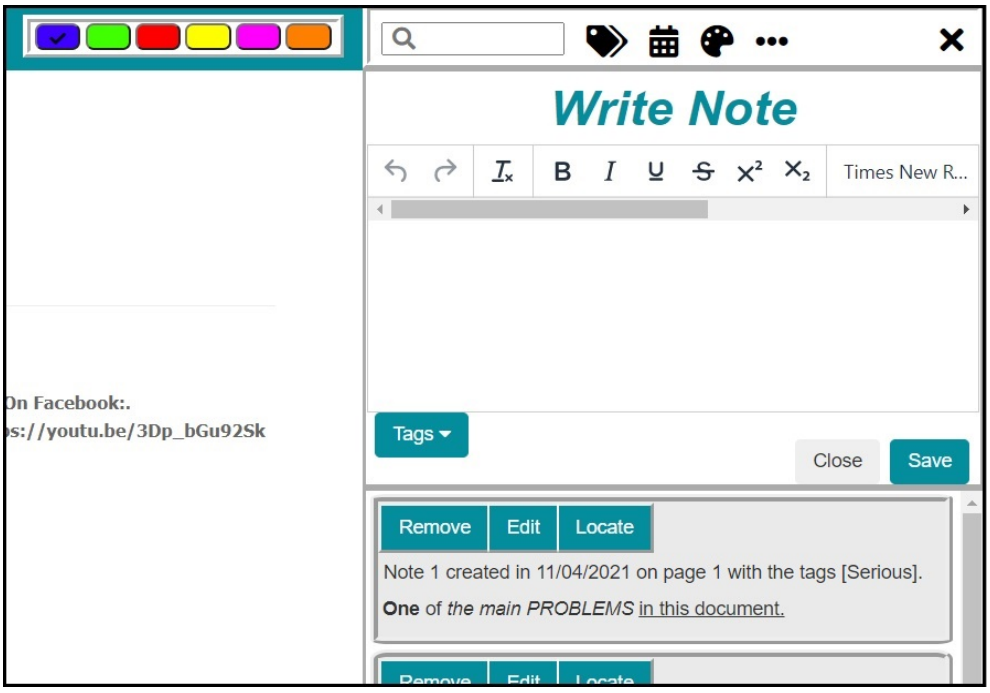


Figure 127: Add/edit box in the sidebar.

Finally, the last issue considered was the sidebar itself. It was decided to allow it to be hidden or expanded since when it is open it could distract the user. Another detail, inspired

by the article Wolfe (2008), is related to the fact that the sidebar pushes the text that is being annotated to the side instead of overlapping it, thus not covering the text, allowing the user to continue to annotate even with it expanded. All this decisions can be seen in Figures 128 and 129.

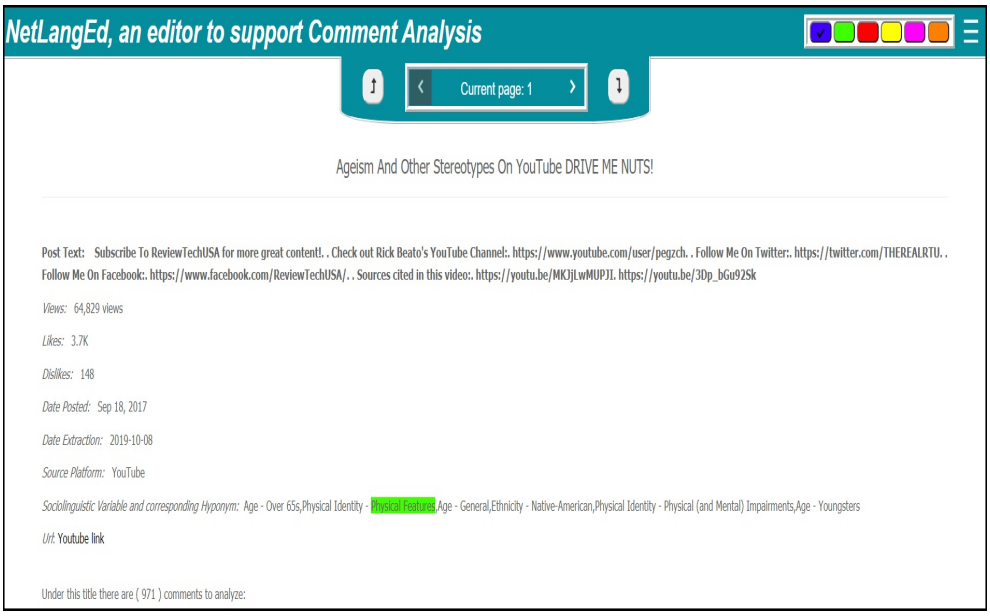


Figure 128: Sidebar hidden.

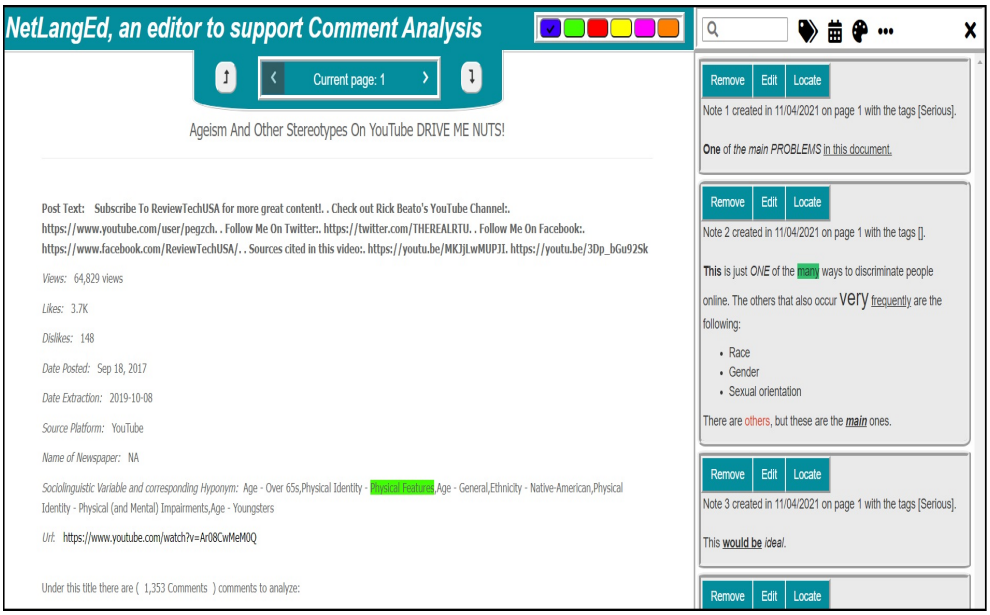


Figure 129: Sidebar expanded.

4.2.2 *Simple and clear functionalities*

Most features have a word on the respective button that clearly describes their purpose, as for example, in the case of Figure 116.

There are some exceptions, such as the functionalities at the top of the sidebar that have icons as shown in Figure 129. However, care was taken to use icons that represent the respective functionalities. Thus, only the operations that are executed from the document, these being the tooltips, removal and editing, are not so obvious to execute. However, in this case it was decided to sacrifice this aspect a little in order to benefit other requirements that will still be discussed here.

4.2.3 *Quick add functionality*

This functionality was implemented being only necessary to perform four actions (select text, click on pop-up, write comment and save comment) as can be seen in Figure 112.

There is a way to reduce the number of necessary steps to three, which consisted in removing the pop-up part and display the addition box after completing the selection, however this method could create complications in cases where the user accidentally selects something that he didn't want to annotate. Note that this process may require a greater number of steps if the user wants to associate tags to the annotation, however it is not a mandatory step in the process.

4.2.4 *Quick edit functionality*

To fulfill this requirement, it was important to make sure that the editing functionality could be performed both in the document and in the sidebar. In this case, both situations only require three actions (in the document it is necessary to click on the annotation, edit and save changes and in the case of the menu it will be necessary to click on the edit, edit and save changes button) to complete the process, both of which are visible in Figures 119 and 120, without having found any way to reduce this number. Note that, as in the case of addition, this process may require a greater number of steps if the user wants to edit the tags or the color that are associated with the annotation, however these are not mandatory steps in the process.

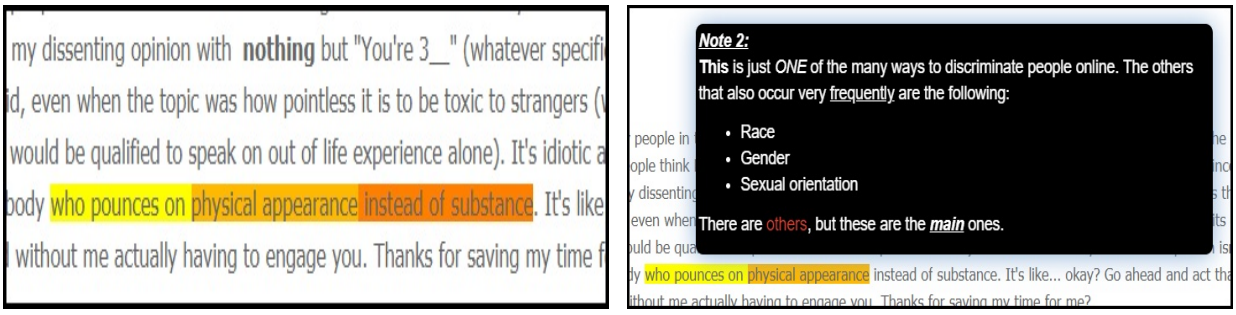
4.2.5 *Quick removal functionality*

This requirement, just like in the previous one, was achieved by not only allowing the user to use the removal functionality on the document and on the sidebar but also execute it with the shortest number of clicks.

Regarding this last aspect, in the document it is only required two steps (click on the annotation and click on the remove button) and on the list of comments its required only one (click the remove button), both of this cases being visible in Figures 117 and 118. In the case of executing this functionality in the document, it is possible to be done in just one click by right-clicking on the annotation to be removed, however this method could lead to situations where the user could unintentionally remove an annotation.

4.2.6 *Annotation overlay*

This feature was implemented using a color effect, as can be seen in Figure 130a. Observe that the color effect is present in the part of the text in which the overlap has occurred, which will have a color that will be the result of combining the color of the last annotation made on that part of the text with the color of the new annotation also made on that part of the text. However this effect does not work when there are several overlaps. In order to resolve these situations, placing the cursor over an annotation will only show that specific annotation, as can be see in Figure 130b, returning to normal as soon as the mouse is moved out of it.



(a) Without effect.

(b) With effect.

Figure 130: Example of an annotation overlay.

4.2.7 *Annotation comment representation*

This feature can be observed in Figure 131, where it is possible to see that the annotation content and its formatting are displayed in the same way in both cases.

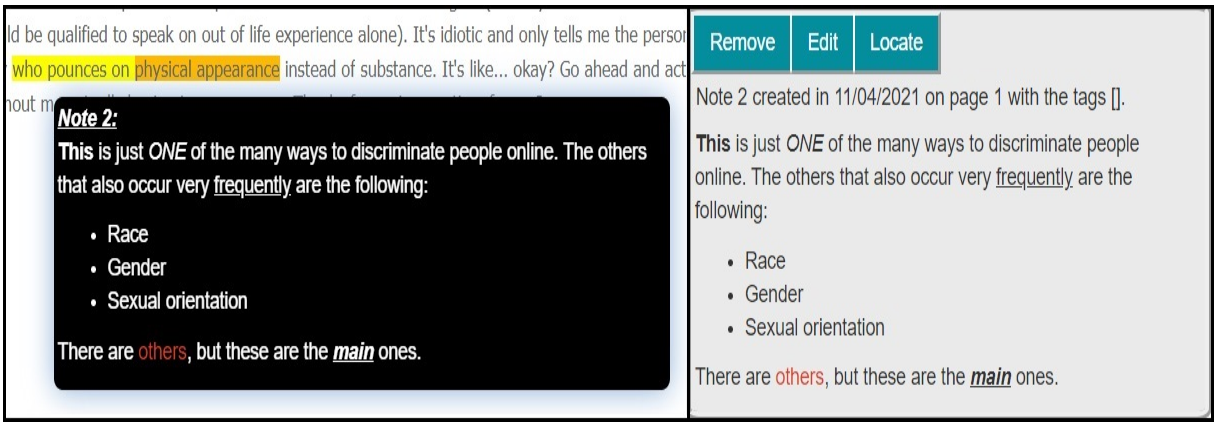


Figure 131: Representation of comments and their formatting.

4.3 Integration with the NetLang project

To wrap up this chapter, where the tool implementation is discussed, the problems that were encountered during the integration process with the existing *NetLang* platform are identified and the solutions found are also presented.

The first problem was related to the fact that there are very long texts on which it would be applied the process of placing *span* tags in all of its characters, causing the browser to jam and not load the page. The solution to this problem was to load part of the document instead of loading it all at once and to load more whenever the user pressed a button. Although the previous solution eliminated the problem of not being able to load the page, it did not avoid the second problem that occurred when the page size started to become very large, causing the editor to function slowly. In view of this, the final solution was to load the document as pages that would not exceed a maximum limit, where the user could navigate backwards or forwards one page at a time or even directly open a specific page, this being visible in the Figure 132.

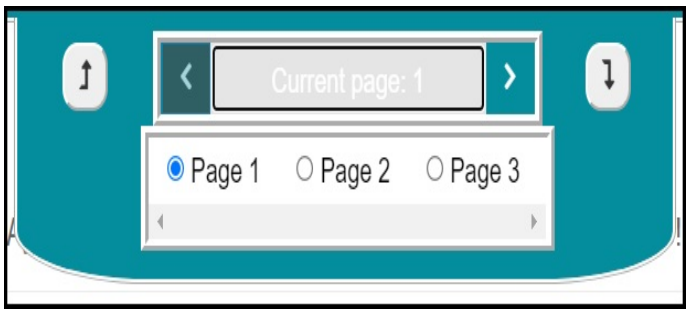


Figure 132: Page browser.

After overcoming all this problems, the planned tool is operational and can be accessed at https://bit.ly/AccessToNetLangEd.

It is important to mention that the editor can easily be applied in other scenarios that are not related to the *NetLang* platform, since the editor *NetLangEd* only requires a text in the HTML format as an input to work properly.

4.4 SUMMARY

As a recap, in this chapter it was reported how the functional and non-functional requirements were implemented, as well as the various challenges faced during the integration of the editor in the *NetLang* platform and how these were overcome.

5

NETLANGED, THE EXPERIMENTS

After developing the editor, an experiment was developed in order to collect feedback from its future users and people with computer skills regarding its usability.

Thus, this chapter is divided into two sections. The first one is used to describe the experiment, how it was set up. In the second one the results obtained are analyzed and discussed.

5.1 THE SETUP

This section describes the participants in the experiment and details its design and conduction. The script used to guide participants testing the editor is presented in Appendix A. The Questionnaire prepared and data collected with it are included in Appendix B.

The experiment was carried out by students, professors and researchers from the University of Minho in the areas of Linguistics, Informatics, Humanities, Social Sciences and Engineering. Of these, 62.5% are male and 37.5% are female, aged between 22 and 53, where 50% have a Master's degree, 25% have a degree and 25% have a doctorate.

To help testers, a script was developed. The script is divided in two parts. In the first section a *NetLangEd* overview is provided and some aspects of the editor are clarified; the second section describes the various tasks that need to be performed to complete the experiment.

In the first phase of the experiment, it is intended that testers use the browser and mark the input text according to pre-defined rules, that is, spelling errors with the orange color, lack of verbal agreement with the purple color and use of slang with the red color. To complete this phase, the tester must save the work done and export it as well as the notes.

The second phase consists of importing a pre-prepared work and filtering it by one of the tags that are present in it and by the date "16/09/2021". After the filters are applied, it is intended that the tester analyze the tag/tags, comment and color of the notes present in the menu and use the function that locates the annotated text in the document. After that, if the tester does not agree with some aspect of the annotation, he is expected to edit it as he thinks it is more correct.

The third and last phase of the experiment consists of answering a questionnaire where several aspects regarding the usability of the editor will be assessed.

5.2 THE RESULTS AND DISCUSSION

In this section, the results of the questionnaire answered by the testers, at the end of the experiment, are analyzed. Questions and answers of the questionnaire used are shown in Appendix B.

The first question intends to know if testers are used with corpus analysis, resulting in 50% of them having done so, being possible to conclude that only half of the participants will give answers from the point of view of the editor's target user.

The next question allowed us to verify that 75% of the participants agreed that the interface is simple and clear. The reasons that led the others to disagree were also analyzed. Furthermore, the editor web page was considered by all to be easy to understand and use.

Regarding the colors, 75% of the participants agreed that the colors available are enough. Those that disagreed prefer the ability to create their own colors. The reason for not including this feature was to simplify the usage of the editor by avoiding the management of another aspect in addition to the tags management.

Another aspect of the editor that was questioned was the formatting options, where 87.5% agreed that the existing options were sufficient. As for the others who did not agree simply stated that the options to copy and paste should be present. However the reason for this limitation was to avoid security breaches.

In relation to the filtering options, 87.5% of the participants considered them to be sufficient. That being said, the only participant who disagree did not fully understand the issue since its justification did not relate to the problem in question.

As for the information present in the tooltips that appear when placing the mouse cursor over an annotation, 87.5% of the participants agreed that it is enough. The others who did not agree stated that the tags that were associated with the annotation should also be present. However it was decided not to put them as it would further limit the content shown in the tooltips since its only possible to display six lines of the comment plus the line that displays the number of the note.

In regards to whether the information displayed in the notes present in the menu is sufficient, as can be seen in Figure 133, 50% of the testers concluded that the information was not enough, stating that it should also have the tags that were associated with the comment.

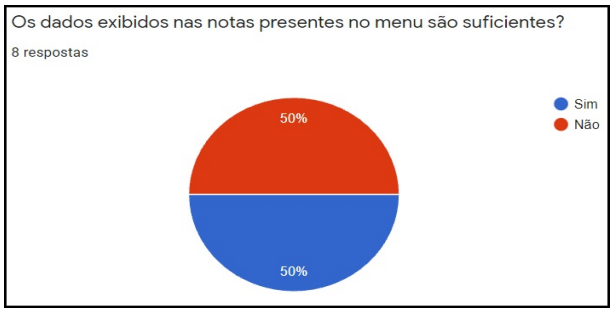


Figure 133: Question regarding the information present in the menu notes.

The reason for not including them initially was to avoid overload the notes with too much information. However, after this feedback, this information is now present in the notes that are visible in the menu.

Another question, which is similar to the above, is whether the information displayed in the exported PDF is sufficient. As in the previous question, 50% of the testers concluded that the information presented is not enough, as can be seen in Figure 134, stating that it should also contain the tags that were associated with the comment and the text that was annotated.

Figure 134: Question regarding the information present in the exported PDF.

Technological issues justify the information missing in the exported PDF. However, after this feedback and after exploring solutions to this issue, the information is also now present in the exported PDF.

Regarding the number of steps that are required to perform the actions of adding, editing and removing annotations, from the document and from the menu, most participants agreed that none of them require too many steps. For the processes of adding annotations, edit from document, edit from menu, remove from document and remove from menu resulted in the values of 87.5%, 75%, 75%, 100% and 87.5% respectively. Unfortunately, the number of steps that these processes require were already reduced as much as possible, not having discovered another way to reduce them even more.

One of the most important questions was whether the editor's performance was pleasing. 75% of the participants said that they are happy with the performance; however, 25% of

them claim that the editor takes a long time to change pages. The next question intends to understand whether the structure of the document in pages prejudice the process of reading. Only 37.5% of the participants said that the presentation of the comments split into different pages is not comfortable, as shown in Figure 135.

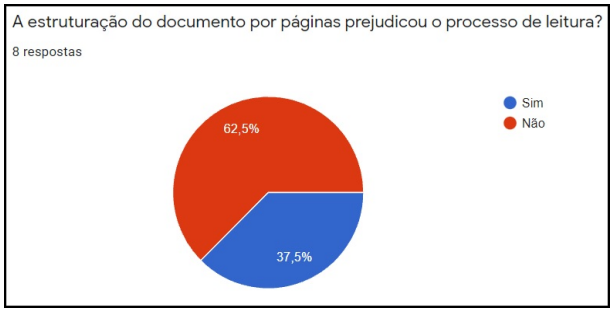


Figure 135: Question concerning the structuring of the document in pages.

Both the problems mentioned in the previous two questions are related to the way in which the *span* tags are placed in the document; unfortunately an alternative way was not found. This situation is one of several aspects that are part of the future work that will be mentioned later.

In regards to the possibility to use the editor without training, 37.5% of testers concluded that it cannot, as can be seen in Figure 136.

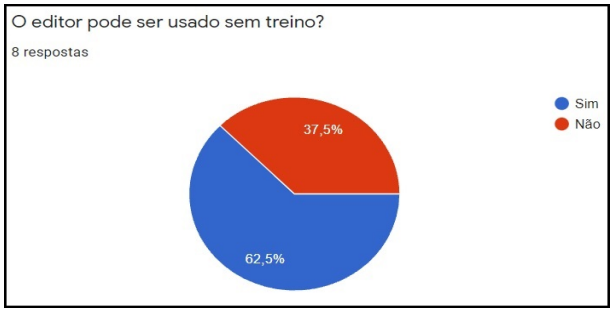


Figure 136: Question regarding if the editor can be used without training.

The reasons given by the testers were that the names of some buttons were not intuitive and the icons used at the top of the sidebar are not clear. Both situations have already been fixed by giving clearer names to the buttons and by adding tooltips to the icons that appear after placing the cursor over them for one second.

Finally, all participants agreed on the questions of whether the editor facilitated the ability to analyze the text and whether it is useful in the process of analyzing the corpus, being possible to conclude that the participants found value in using the editor.

The analysis carried out on the results of the questionnaire, allowed to affirm that most of the defects pointed out by the participants are cases of personal preference and not editor faults. Thus it can be concluded that the objectives that were set for the editor were achieved.

5.3 SUMMARY

As a recap, this chapter described how the experiment to test the editor was conceived. First, participants were characterized. Then the process carried out was explained. At last, an analysis and discussion of the results was presented.

6

CONCLUSION

Along the development of the Master's work here reported it was possible to have a better notion about the evolution that occurred in the use of annotations, and how did they evolved from a strict use for sharing knowledge to a mostly personal use. Studying carefully the work carried out by several authors it was possible to observe several details of the readers' behavior when annotating on paper and online, being possible to identify the types of annotations commonly used and also their purposes.

It was also possible to enumerate some advantages and disadvantages that online annotations have compared to paper annotations, allowing to conclude that there are two main factors that lead readers to not prefer the use of online annotations. The first is related to the original format of the document to be annotated. In case of being paper it will be very unlikely that it will be digitized to be digitally annotated. The second factor is related to the functionalities that the annotation system provides; when the tool does not cover properly the readers' main needs of annotation will lead the users to opt for printing the document and annotate it on paper.

In order to understand the characteristics that a web annotator should have, several existing solutions have been analyzed, being possible to observe some characteristics that can be considered essential and others that are either not very useful or simply do not fit the objectives to be achieved. When comparing these solutions it was possible to conclude that, although none was perfect, the best would be *Hypothes.is* since it allows for the annotation of web pages, as well as it fulfills the greatest number of the identified requirements.

This analysis was then used as a basis to create a list of functional and non-functional requirements that the editor to develop must fulfill. Taking into consideration the full list of requirements elicited, the diagrams of the system architecture were developed. In this project phase, it was planned how the editor would be integrated into the *NetLang* platform and how the required features would be organized in it.

The development of the editor's functional and non-functional requirements was successful, with some sacrifices being made for the sake of safety and aesthetics. The complicated part of the development of the editor occurred at the time of integrating the editor in the *NetLang* platform causing the editor to work very slowly or making the browser to crash.

The solution found was implemented but it did not overcome all the performance problems, however a pleasant usability level was achieved.

After finishing the development of the editor, an experiment was designed to test the editor. In order to obtain a more realistic feedback from it, potential users and people with computer skills were chosen as participants. Using a questionnaire after the experiences with the editor, it was possible to obtain information about the quality of the editor, potential features to be added in the future, and about errors detected. It is worth highlight the result of two questions that were essential to validate the editor. One question asked whether the editor facilitated the ability to analyze the text; the second one asked whether the editor is useful in the process of analyzing the corpus documents. All the participants agreed, this is, answered that the editor was helpful on both cases.

During this testing phase, it was also possible to write the "*NetLangEd, The Web Editor to Support Online Comment Annotation*" paper based on the work developed, which was published at the conference 10th Symposium on Languages, Applications and Technologies (SLATE 2021) (Rodrigues et al., 2021).

Finally, after completing all the steps mentioned above, it is possible to affirm that all the requirements that were stipulated for the editor were fulfilled. As a final result, an editor was obtained that has a combination of all the positive aspects that were identified in the other editors that were explored here, thus dispensing their use by the users of the *NetLang* platform. However, there are some aspects that can be improved or even added in the future, such as the following ones:

- Improve the visual aspect of the editor, such as the way the editor is structured, the color of the editor, etc.
- Allow filters to act as a conjunction.
- Provide more variants of annotation.
- Allow the customization of the highlight colors.
- Develop the ability to move the highlight on the text to another part of the text.
- Develop a more efficient way to make annotations so that it is possible to load the document in its entirety instead of loading it in pages.

A

EXPERIMENT SCRIPT

Below, the script that was developed to test the *NetLangEd* editor will be presented, having been written in the Portuguese language.

EDITOR ONLINE NETLANGED
GUIÃO DO TESTE DE USABILIDADE

Este guião visa orientar o utilizador do editor online ***NetLangEd*** em 2 experimentos pré-concebidos para que no fim seja possível recolher através de um questionário, o feedback de cada utilizador que tenha aceite participar neste experimento, que se realiza todo online e de forma assíncrona. É importante referir que o editor foi desenvolvido e testado com o browser ***Google Chrome***, não podendo garantir o seu correto funcionamento com outros browsers.

A.1 ESCLARECIMENTOS SOBRE O EDITOR

Antes de iniciar o experimento, é importante referir algumas decisões que foram tomadas durante o desenvolvimento do editor NetLangEd, o qual foi concebido para ser usado por qualquer Analista não-informático permitindo-lhe escrever as suas observações e ideias livremente, e essencialmente com carater pessoal, à medida que lê os documentos do corpus NetLang em português e inglês.

A primeira observação que é importante referir tem a ver com o conceito de *tag* (o que é, para se usa) e, mais concretamente, como é feita a gestão das tags.

As *tags* são simplesmente nomes de *etiquetas* que poderão ser atribuídos às anotações (notas escritas pelo utilziador) para depois facilitar o processo de procura e organização destas mesmas. Na verdade a tag funciona como uma etiqueta que identifica um tipo de problema digno de nota.

Para aceder à janela que permite gerir as tags basta abrir o menu, clicar no icon das tags (a primeira figura a contar da esquerda) e depois clicar no botão `'Tags'`. Nesta janela poderão ser efetuadas três operações.

Criar tags, basta escrever o nome da tag onde diz `'Add tag...'` e clicar depois no botão `'+'` que se encontra à frente.

Editar tags, será necessário selecionar uma das tags que se encontra na lista de tags criadas, o que irá levar a que o seu nome apareça na caixa que foi referida na operação acima, podendo depois alterar o nome e concluir o processo clicando no botão `'Edit'`.

Remover tags, basta selecionar uma ou mais tags e depois clicar no botão `'Remove'`.

Outra observação importante está relacionada com o facto de *não ser possível associar as cores às tags criadas*. São dois marcadores visuais distintos e complementares que o ajudam a enriquecer o apontamento das suas reflexões. Esta decisão foi tomada para não limitar a utilização do editor, permitindo assim que os utilizadores anotem os documentos da forma que preferirem, isto é, permitindo associar uma tag a múltiplas cores ou associar uma tag a uma só cor ou também associar a mesma cor a diferentes tags.

Um último aspecto que importa referir é a existência de tooltips (notas curtas) com informação de ajuda relativa aos icons presentes na barra de ferramentas. Para as visualizar será necessário manter o cursor do rato sobre um dos icons durante pelo menos um segundo.

A.2 EXPERIMENTO

Nesta secção serão descritas as três fases do experimento, duas propondo tarefas de utilização do editor para que se possa aperceber das funcionalidades oferecidas e de como o mesmo funciona e uma terceira que se destina a recolher a sua opinião através de um inquérito.

A.2.1 *Primeira Fase*

Nesta fase pede-se que leia comentários de um documento do corpus do NetLang e tome as suas próprias notas para criar o seu ficheiro de anotações da análise que fez.

1. Abra o documento (um texto base (post/article) seguido de uma longa lista de comentários) que se encontra disponível no link `https://bit.ly/Documentoaanotar` e clique em "Annotate".

2. Explore o documento através do navegador de páginas de modo a ir vendo os vários comentários.

3. Leia o texto base e comece a ler alguns dos comentários.

4. Por cada comentário lido, cada vez que detete um termo (1 ou mais palavras) problemático (erro ortográfico, ausência de concordância verbal, utilização de calão)...
   a) Se a tag desse tipo de problema ainda não existe, crie a tag.
   b) Caso a situação problemática encontrada seja:
      i. Um erro ortográfico, anote o termo com a cor laranja; associe-lhe a respetiva tag do problema; e escreva a justificação.
      ii. Ausência de concordância verbal, anote o termo com a cor roxo; associe-lhe a respetiva tag do problema; e escreva a justificação.
      iii. A utilização de calão, anote o termo com a cor vermelho; associe-lhe a respetiva tag do problema; e escreva a justificação.
   c) Se no mesmo comentário encontrar outros termos problemáticos que pretenda igualmente anotar, retorne ao ponto ***4*** da ***Primeira Fase*** e repita o processo; Quando terminar esse comentário retorne ao ponto **2** da ***Primeira Fase*** e repita o processo.
5. Guarde o trabalho elaborado através da funcionalidade "Save".
6. Exporte o trabalho elaborado, para que este possa ser continuado numa outra altura.
7. Exporte as notas que elaborou nas etápas anteriores.
8. Avance para a ***Segunda Fase***.

A.2.2 *Segunda Fase*

Nesta fase pede-se que leia as notas criadas por um outro utilizador que analisou previamente os comentários de um documento do corpus do NetLang e exportou as suas anotações.

1. Importe o documento anotado que se encontra disponível no link `https://bit.ly/Notasaimportar`.
2. Filtre por uma das tags que já foram criadas e foram usadas pelo autor das anotações que está a ler.
3. Filtre os comentários de modo a mostrar apenas os que foram elaborados na data ***"16/09/2021"***.
4. Leia um dos comentários filtrados.
5. Use a função que localiza o comentário no documento de modo a poder ler com atenção o comentário e compreender o seu contexto.

6. Caso...
    a) Não concorde com a tag atribuída a um determinado termo, por achar que se trata de um problema diferente:
        i. Se a tag do problema pela qual pretende alterar não existe, crie a tag (Menu - Icon das tags - Tags).
        ii. Altere a cor de acordo com o que foi estipulado no ponto ***4-b*** da ***Primeira Fase***; associe a tag do problema que corresponde à cor; e altere a justificação.
    b) Não concorde com a justificação, apesar de achar que a tag está bem atribuída, altere apenas essa justificação.
7. Se não analisou pelo menos 6 comentários, volte ao ponto ***4*** da ***Segunda Fase***.
8. Avance para a ***Terceira Fase***.

A.2.3 *Terceira Fase*

Terminadas as duas tarefas—anotação de um documento desde o inicio e a análise de um documento previamente anotado—por favor,

1. responda ao questionário que se encontra disponível em `https://bit.ly/NovoquestionariodeusabilidadedoeditorNetLangEd`.

Muito obrigado pelo seu esforço e tempo dispendido!

B

QUESTIONNAIRE AND ANSWERS

B.1 QUESTIONNAIRE

In this section the questions of the questionnaire will be presented.

B.1.1 *Section 2*

This subsection presents the questions of the questionnaire that aim to collect information related to the participant.

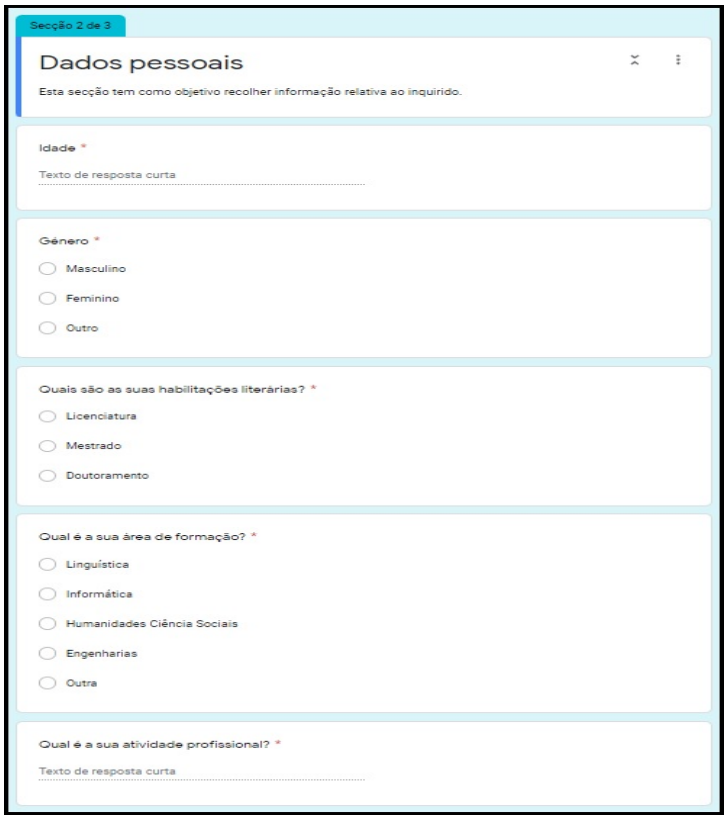


Figure 137: Section number 2 of the questionnaire.

B.1.2 *Section 3*

This subsection presents the quiz questions that are intended to gather information about the quality of various aspects of using the *NetLangEd* editor.

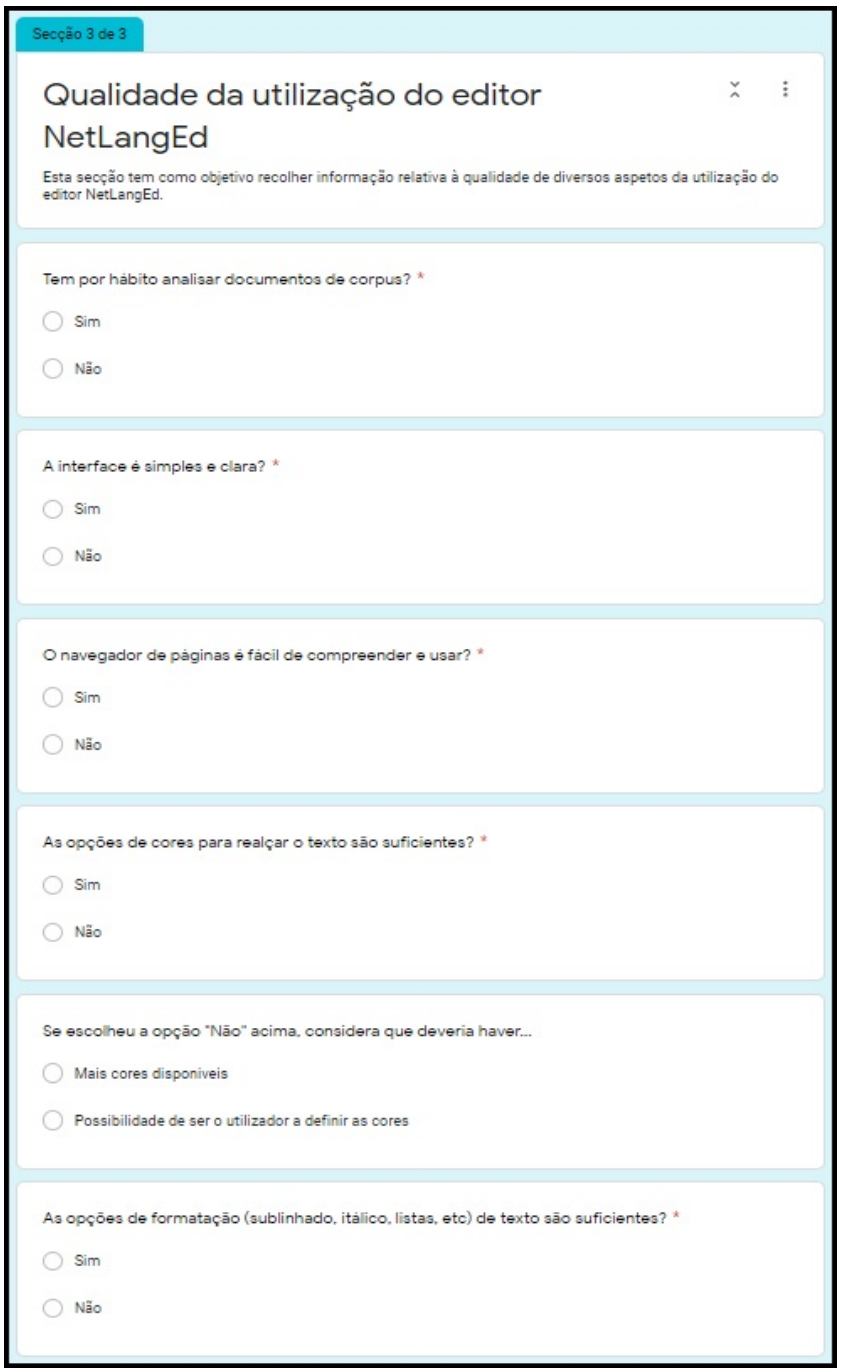


Figure 138: First part of the section 3.

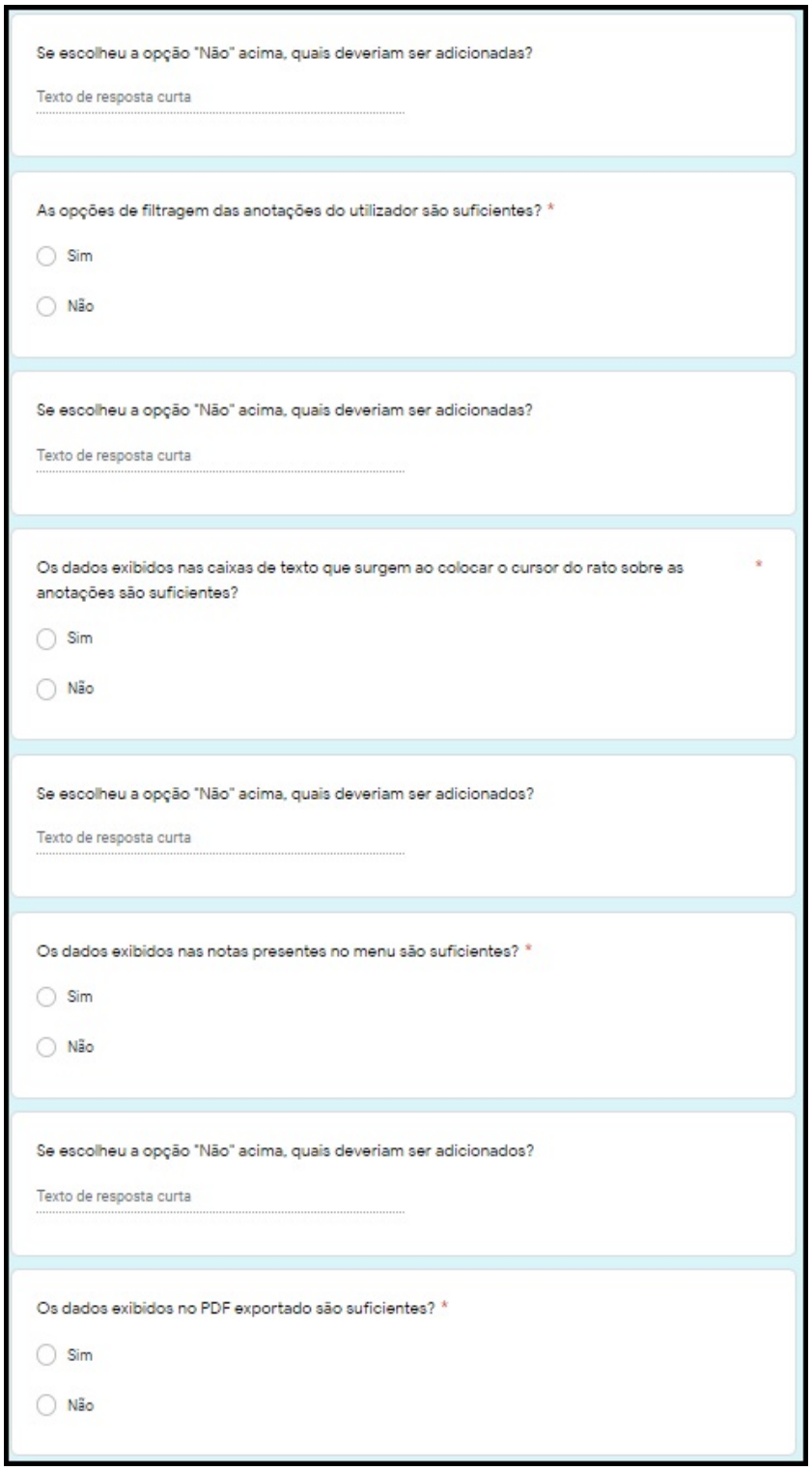


Figure 139: Second part of the section 3.

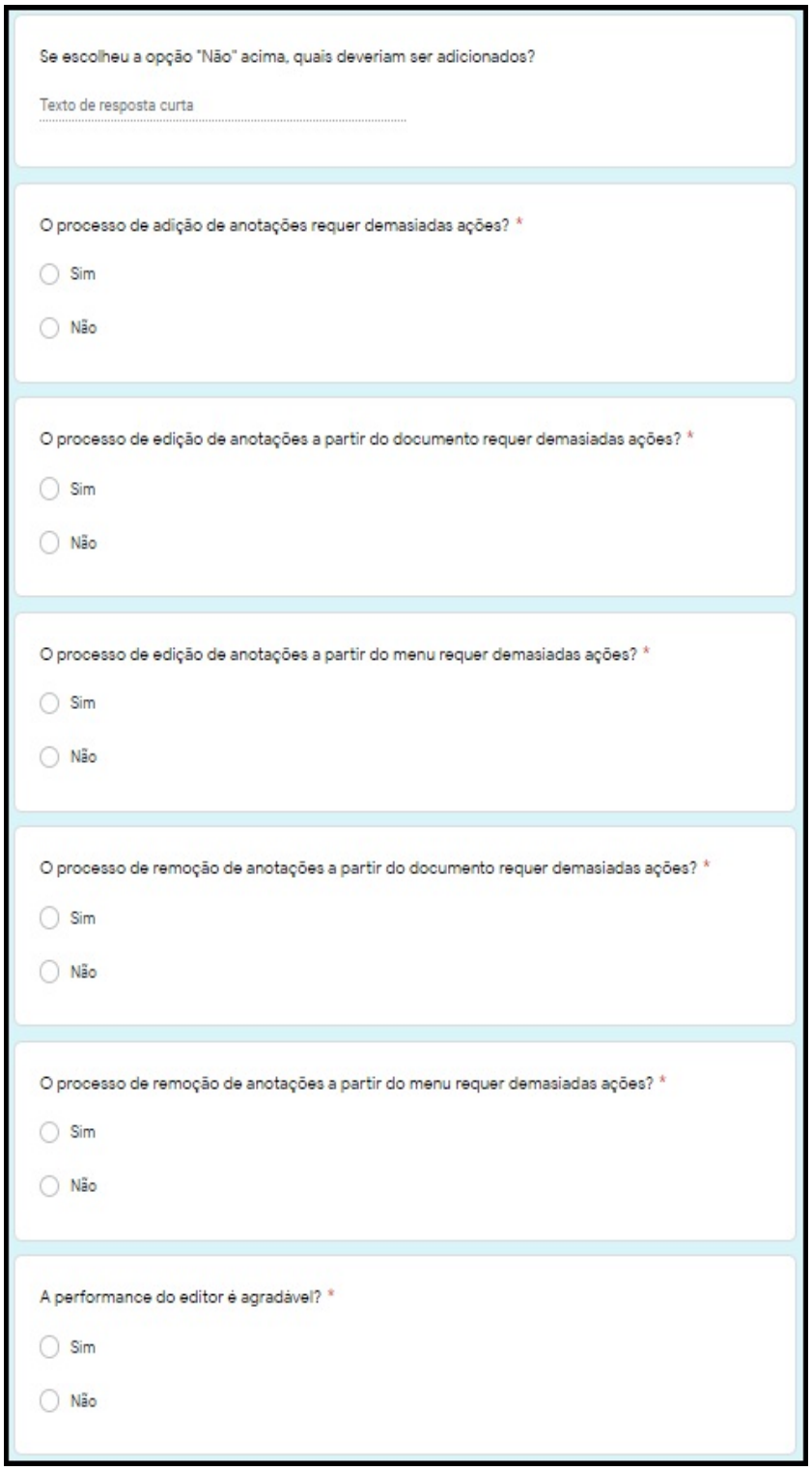


Figure 140: Third part of the section 3.

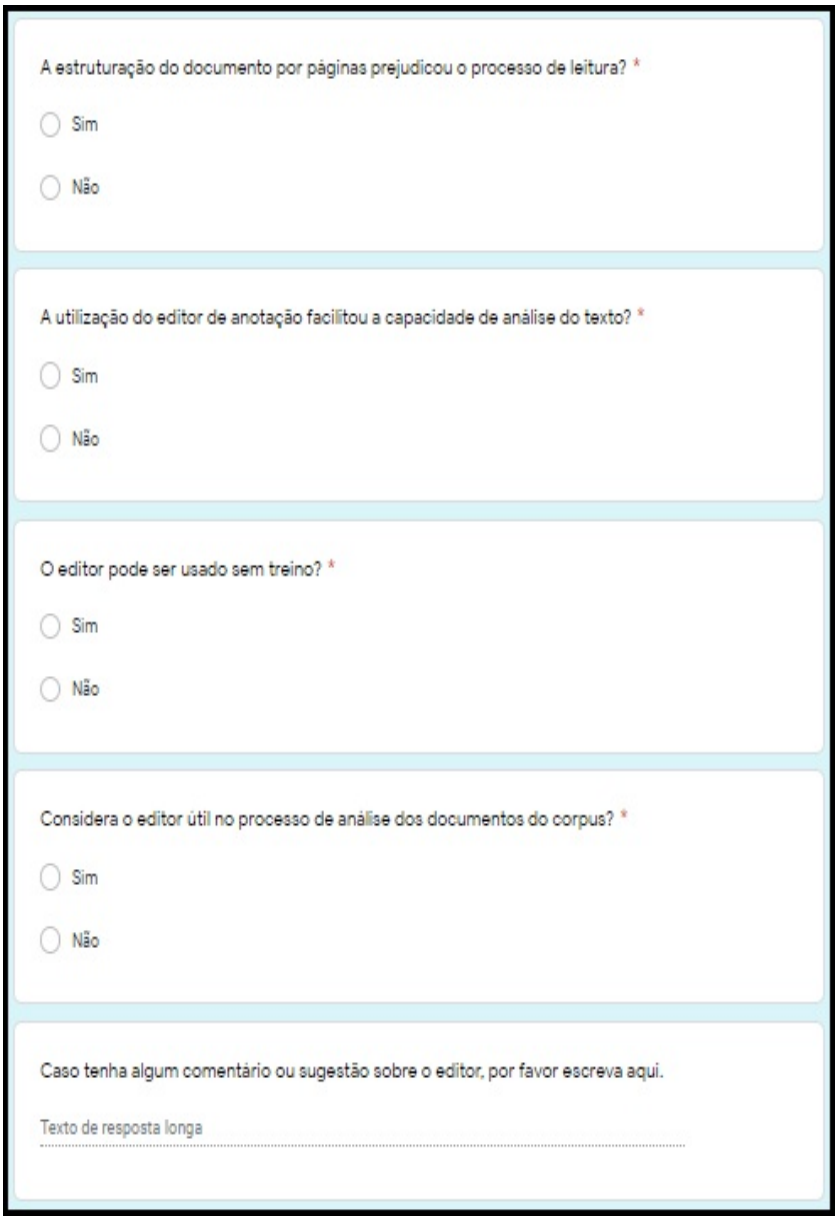


Figure 141: Fourth and final part of the section 3.

B.2 ANSWERS

In this section the answers to the questions of the questionnaire will be presented.

B.2.1 *Age*

The ages of the participants.

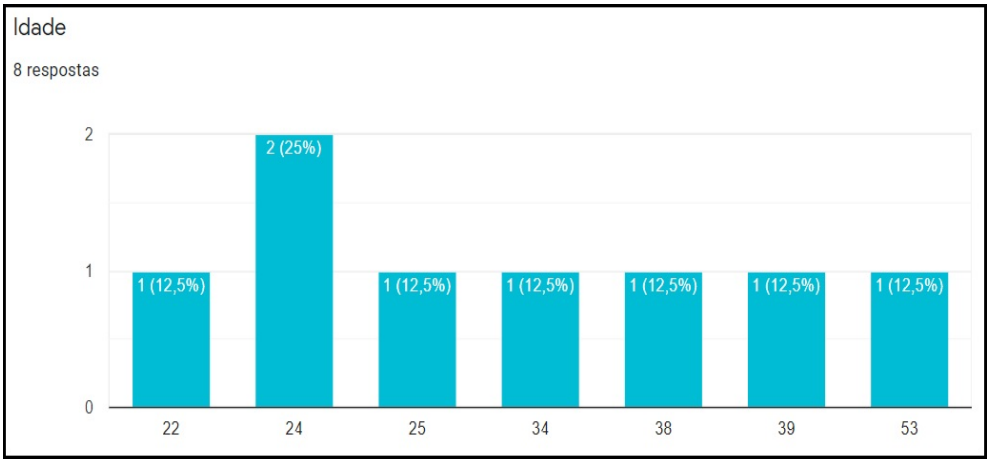


Figure 142: Question 1 of the questionnaire.

B.2.2 *Gender*

The genders of the participants.

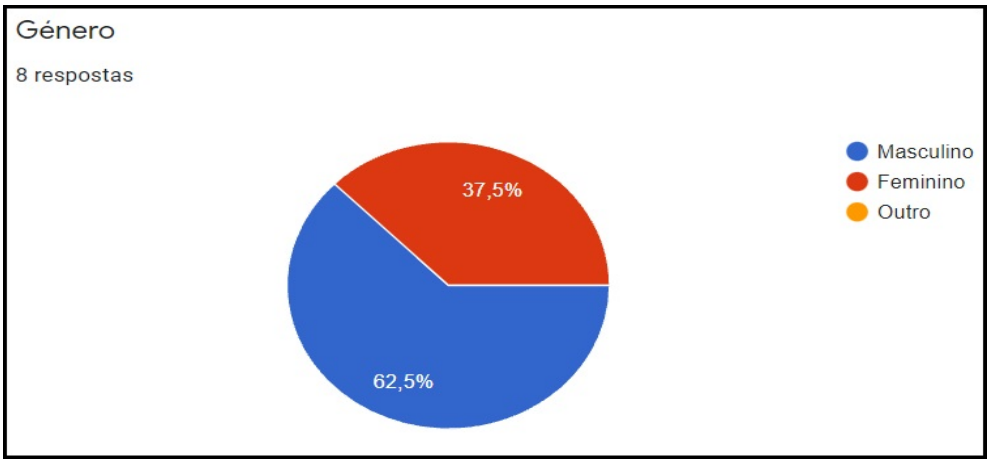


Figure 143: Question 2 of the questionnaire.

B.2.3 *What are your educational qualifications?*

The educational qualifications of the participants.

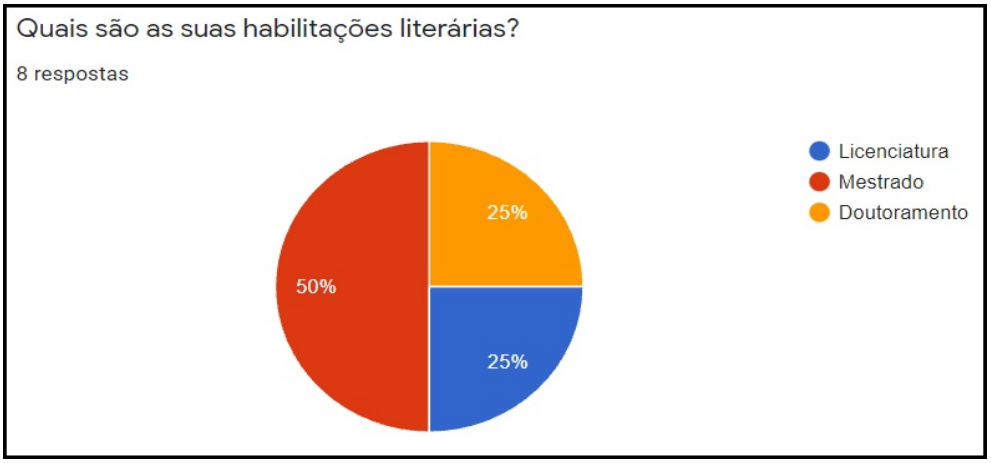


Figure 144: Question 3 of the questionnaire.

B.2.4 *What is your training area?*

The areas of training of the participants.

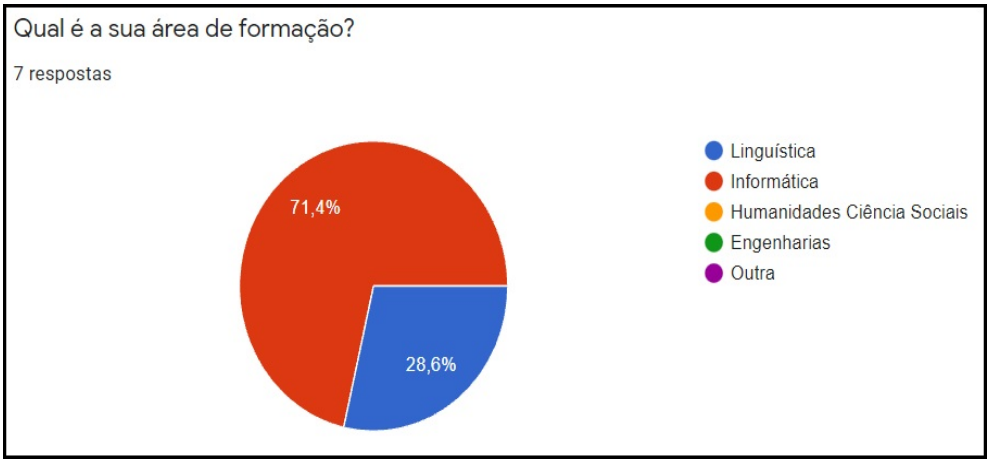


Figure 145: Question 4 of the questionnaire.

B.2.5 *What is your professional activity?*

The professional activities of the participants.

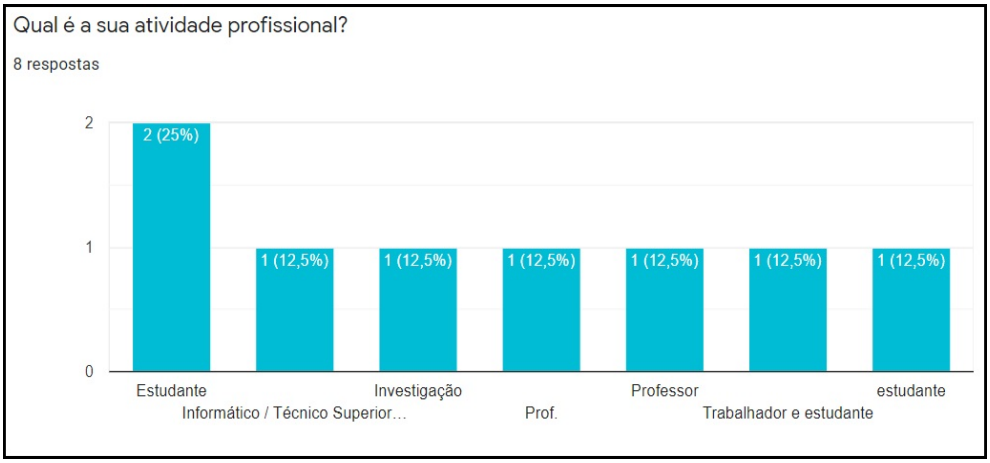


Figure 146: Question 5 of the questionnaire.

B.2.6 *Do you usually analyze corpus documents?*

The percentage of participants who usually analyze corpus documents.

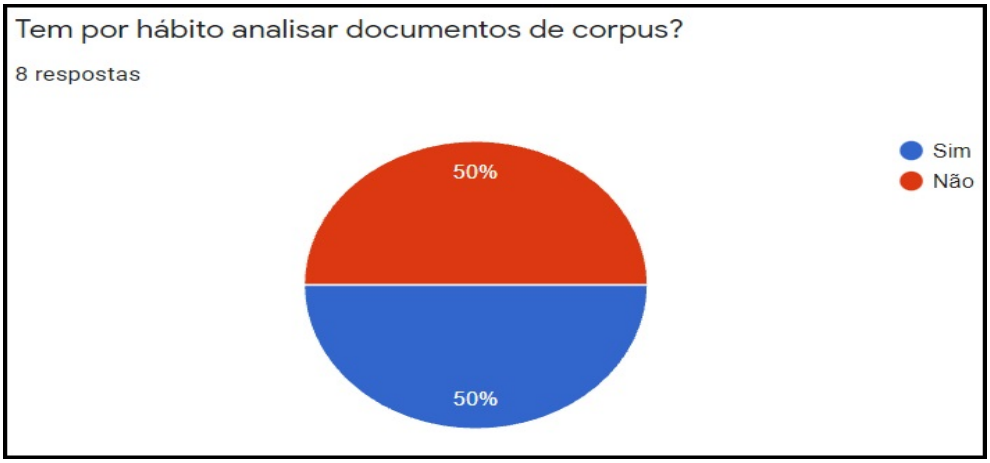


Figure 147: Question 6 of the questionnaire.

B.2.7 *Is the interface simple and clear?*

The percentage of participants who found the interface simple and clear.

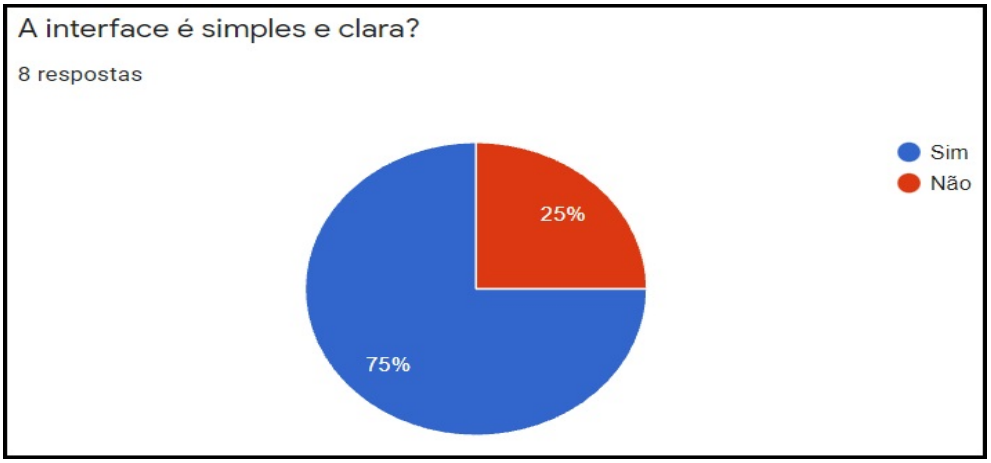


Figure 148: Question 7 of the questionnaire.

B.2.8 *Is the page browser easy to understand and use?*

The percentage of participants who find the page browser easy to understand and use.

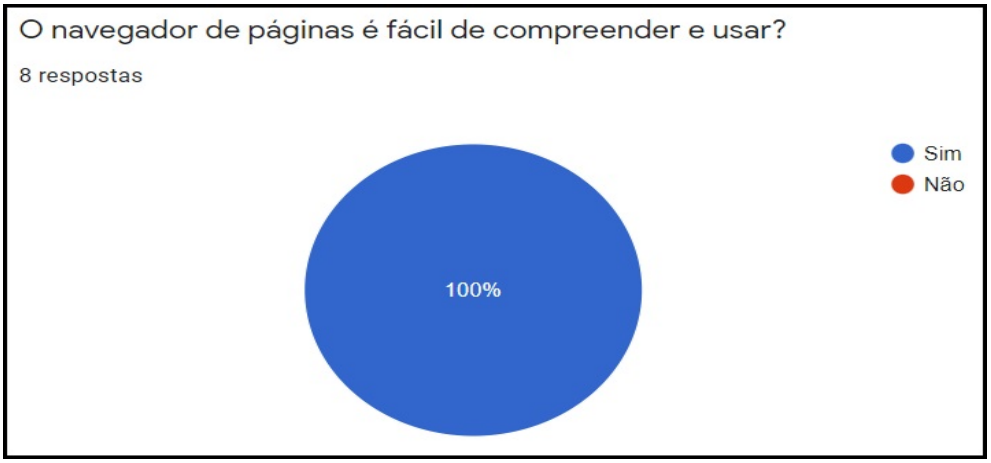


Figure 149: Question 8 of the questionnaire.

B.2.9 *Are the color options for highlighting text sufficient?*

The percentage of participants who agree that the color choices for highlighting the text are sufficient.

Figure 150: Question 9 of the questionnaire.

B.2.10 *If you chose the "No" option above, do you consider that there should be...*

The reasons why the participants do not agree with the above question.

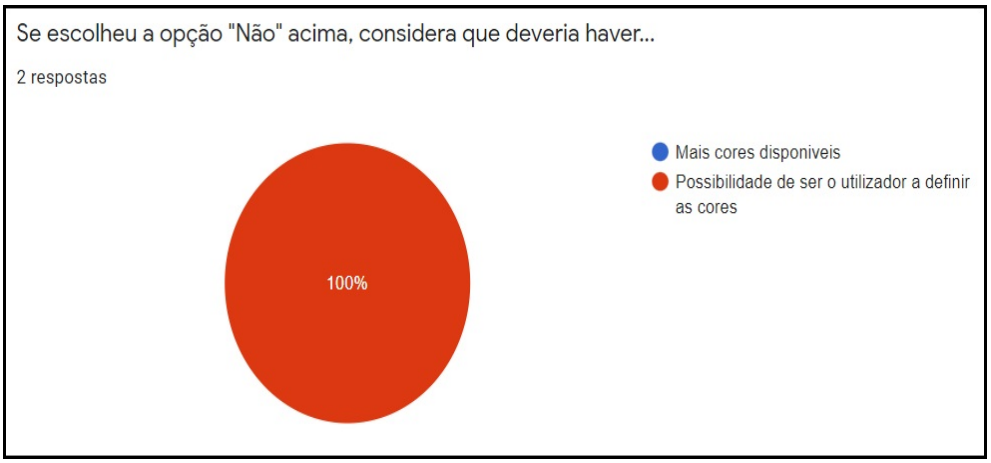


Figure 151: Question 10 of the questionnaire.

B.2.11 *Are the text formatting options (underline, italics, lists, etc) sufficient?*

The percentage of participants who agree that text formatting options are sufficient.

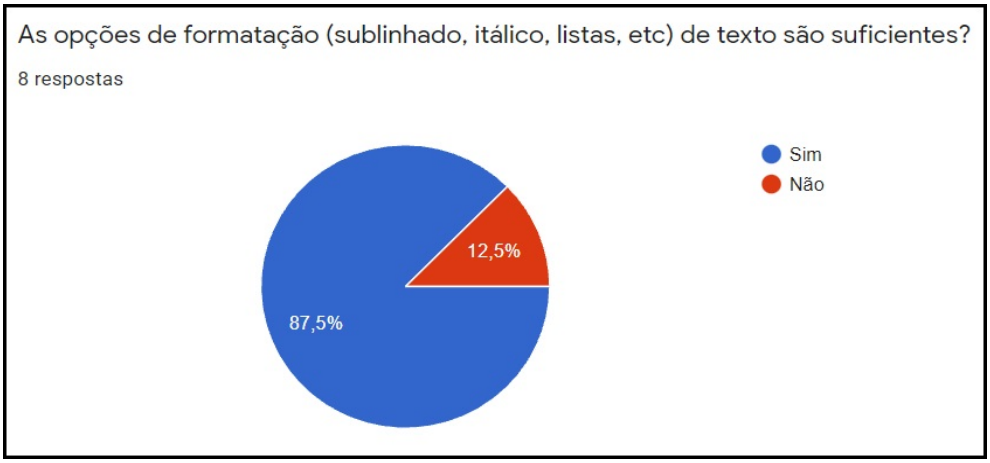


Figure 152: Question 11 of the questionnaire.

B.2.12 *If you chose the "No" option above, which ones should be added?*

The reasons why the participants do not agree with the above question.

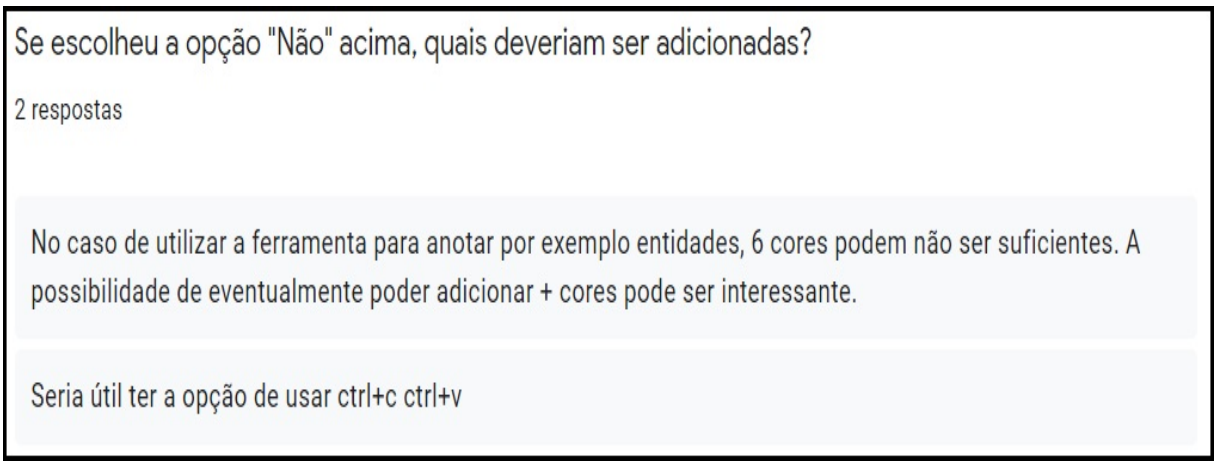


Figure 153: Question 12 of the questionnaire.

B.2.13 *Are the filtering options for user annotations sufficient?*

The percentage of participants who agree that the user annotations filtering options are sufficient.

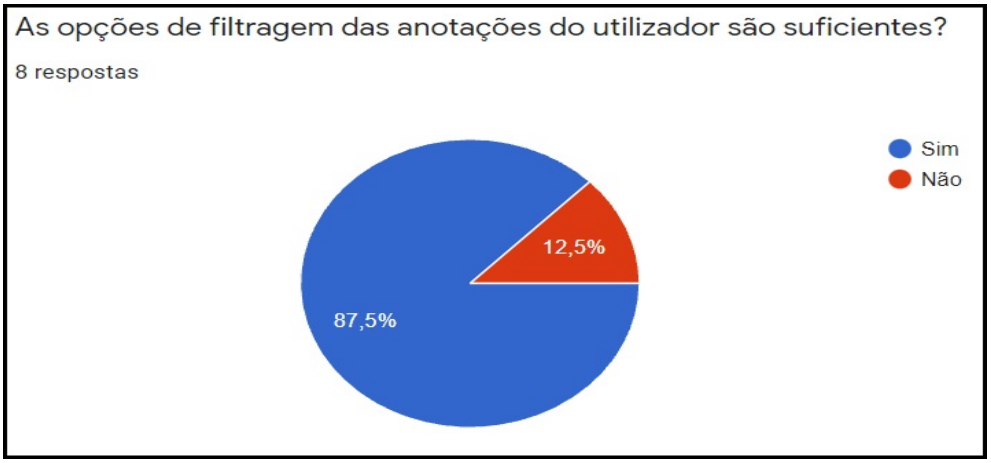


Figure 154: Question 13 of the questionnaire.

B.2.14 *If you chose the "No" option above, which ones should be added?*

The reasons why the participants do not agree with the above question.

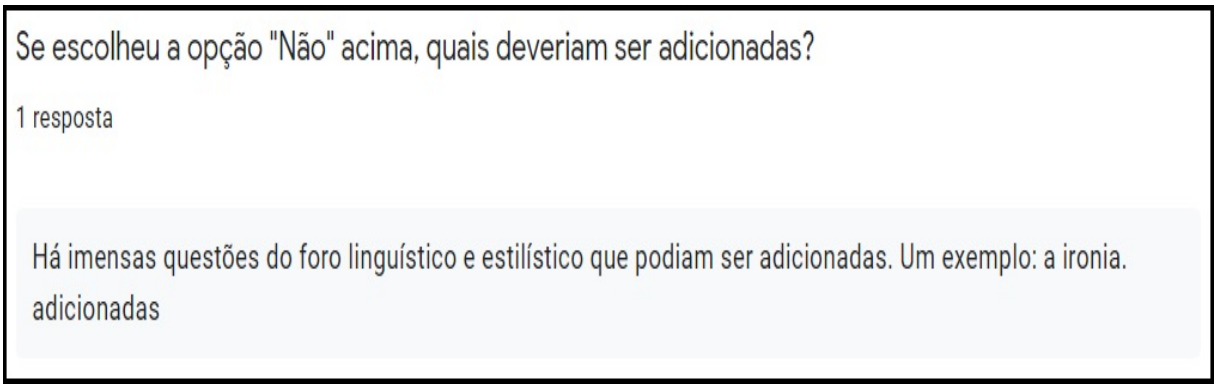


Figure 155: Question 14 of the questionnaire.

B.2.15 *Is the data displayed in the text boxes that appear when placing the mouse cursor over the annotations sufficient?*

The percentage of participants who agree that the data displayed in the text boxes that appear when placing the mouse cursor over the annotations is sufficient.

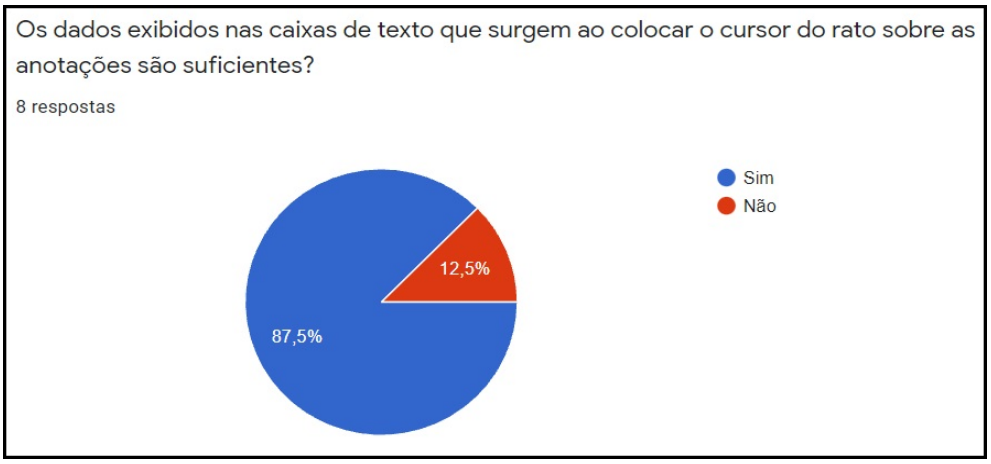


Figure 156: Question 15 of the questionnaire.

B.2.16 *If you chose the "No" option above, which ones should be added?*

The reasons why the participants do not agree with the above question.

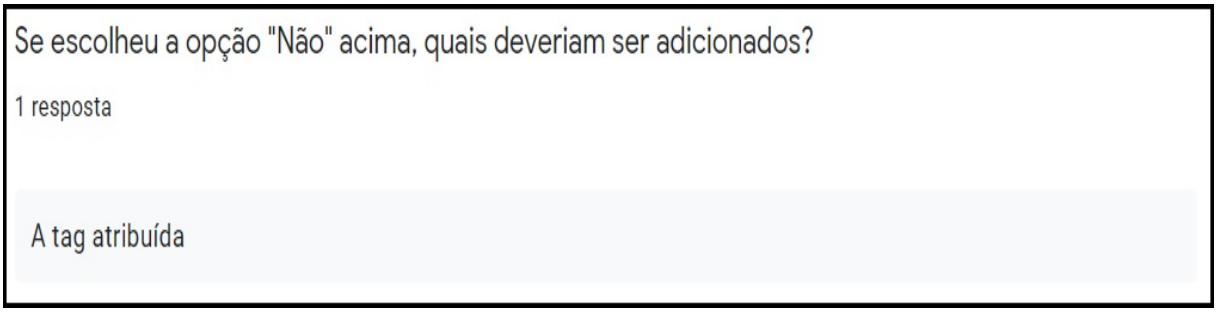


Figure 157: Question 16 of the questionnaire.

B.2.17 *Is the data displayed in the notes present in the menu sufficient?*

The percentage of participants who agree that the data displayed in the notes on the menu is sufficient.

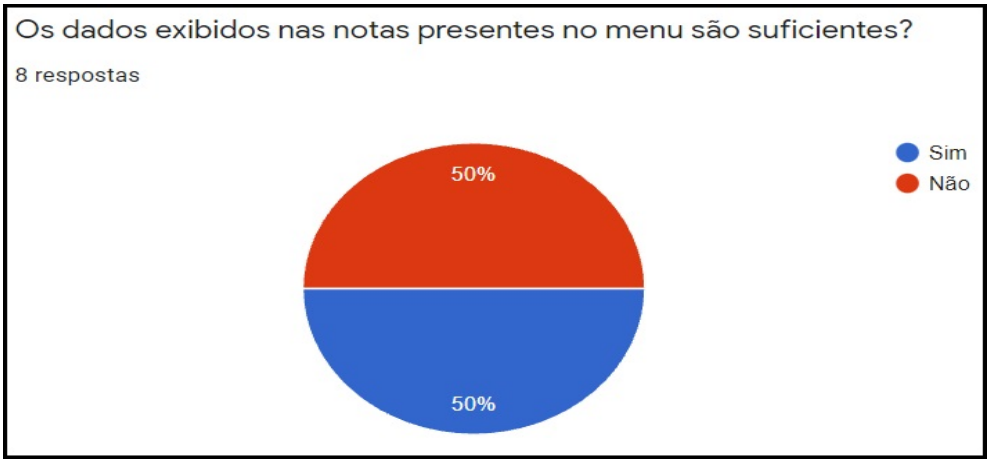


Figure 158: Question 17 of the questionnaire.

B.2.18 *If you chose the "No" option above, which ones should be added?*

The reasons why the participants do not agree with the above question.

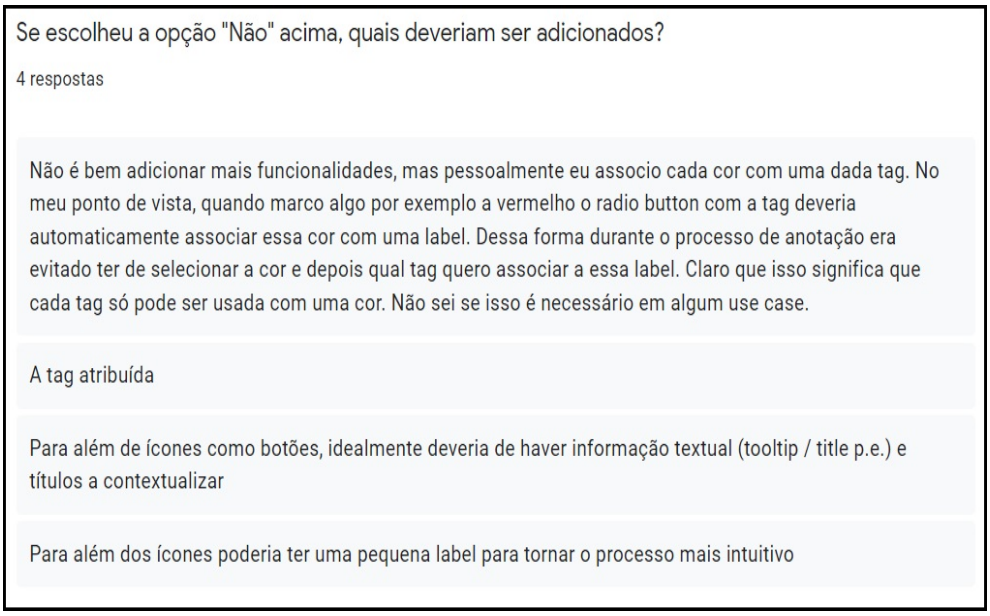


Figure 159: Question 18 of the questionnaire.

B.2.19 *Is the data displayed in the exported PDF sufficient?*

The percentage of respondents who agree that the data displayed in the exported PDF is sufficient.

Figure 160: Question 19 of the questionnaire.

B.2.20 *If you chose the "No" option above, which ones should be added?*

The reasons why the participants do not agree with the above question.

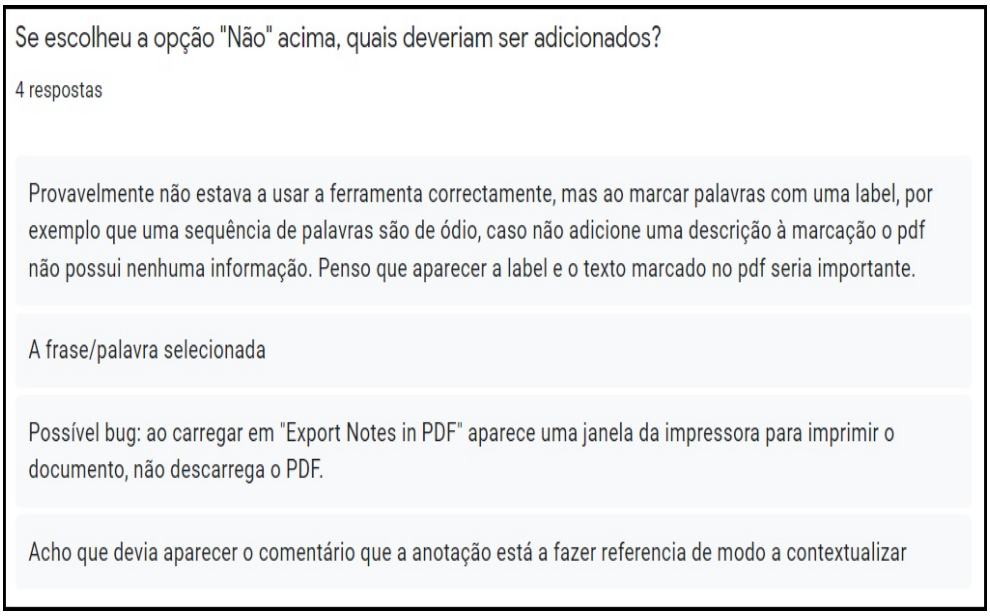


Figure 161: Question 20 of the questionnaire.

B.2.21 *Does the process of adding annotations require too many actions?*

The percentage of participants who do not agree that the process of adding notes requires too many actions.

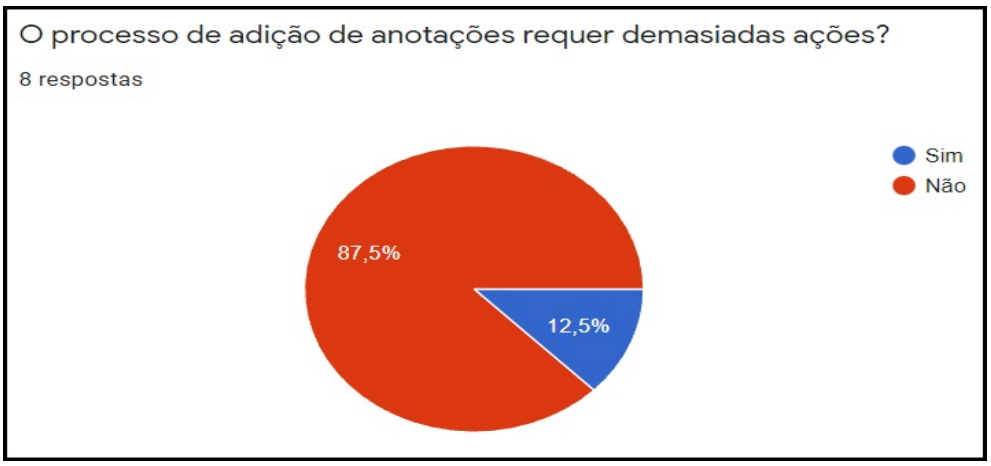


Figure 162: Question 21 of the questionnaire.

B.2.22 *Does the process of editing annotations from the document require too many actions?*

The percentage of participants who do not agree that the process of editing annotations from the document requires too many actions.

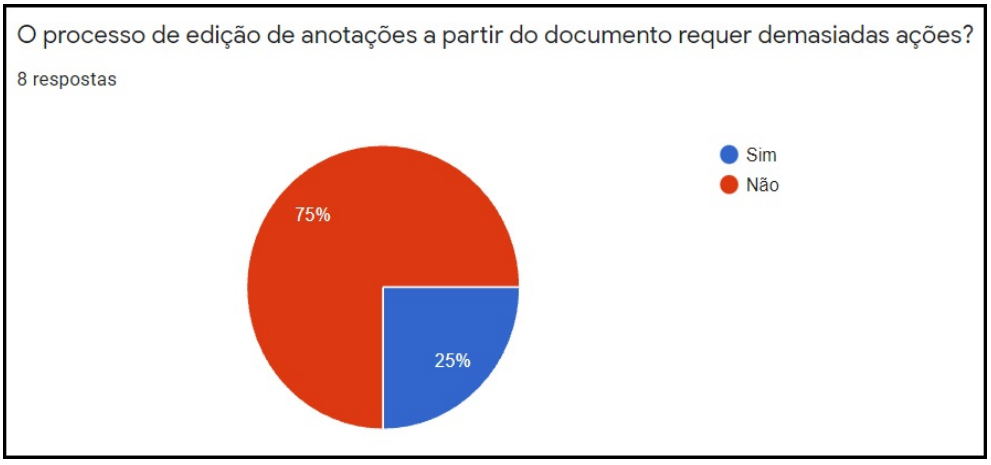


Figure 163: Question 22 of the questionnaire.

B.2.23 *Does the process of editing annotations from the menu require too many actions?*

The percentage of participants who do not agree that editing annotations from the menu requires too many actions.

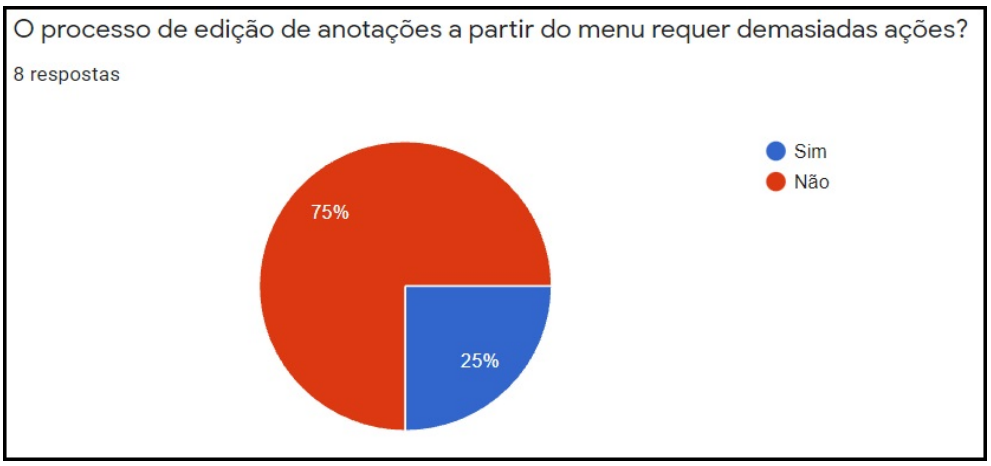


Figure 164: Question 23 of the questionnaire.

B.2.24 *Does the process of removing annotations from the document require too many actions?*

The percentage of participants who do not agree that the process of removing annotations from the document requires too many actions.

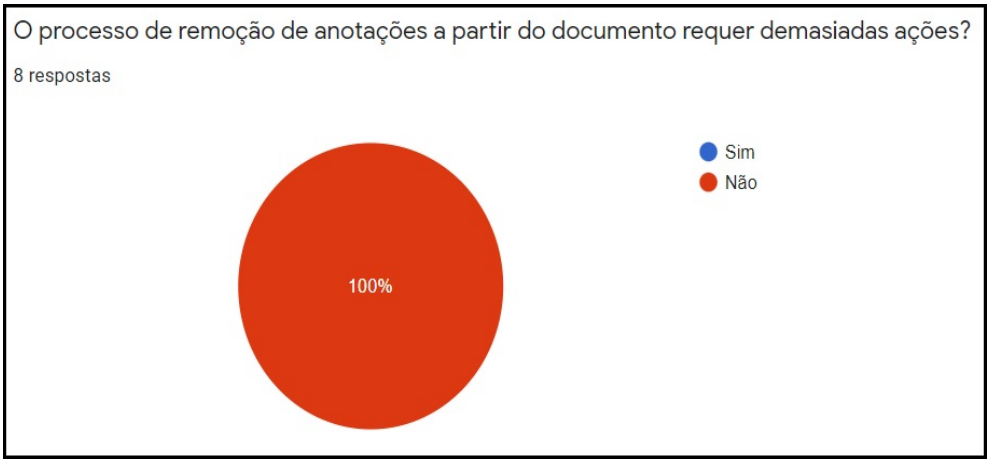


Figure 165: Question 24 of the questionnaire.

B.2.25 *Does the process of removing annotations from the menu require too many actions?*

The percentage of participants who do not agree that the process of removing annotations from the menu requires too many actions.

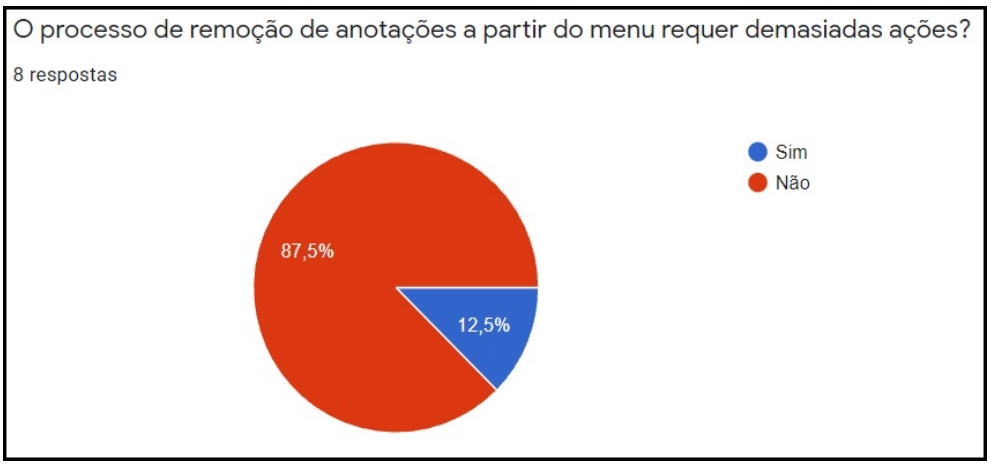


Figure 166: Question 25 of the questionnaire.

B.2.26 *Is the editor's performance pleasant?*

The percentage of participants who agree that the editor's performance is pleasant.

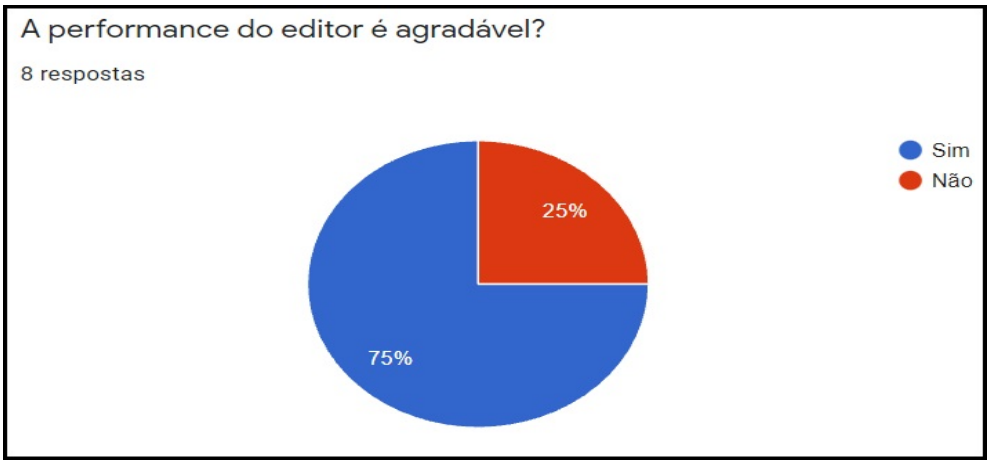


Figure 167: Question 26 of the questionnaire.

B.2.27 *Does the document's page structure affect the reading process?*

The percentage of participants who agree that structuring the document by pages hindered the reading process.

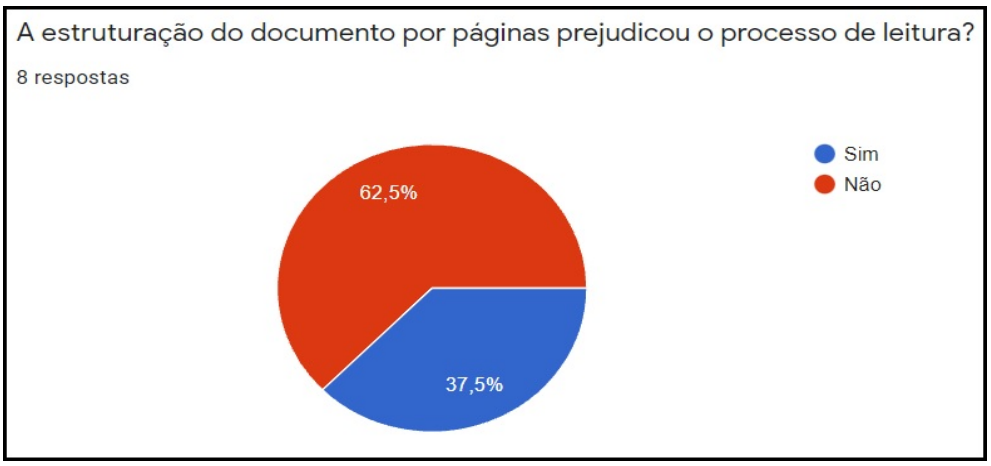


Figure 168: Question 27 of the questionnaire.

B.2.28 *Did the use of the annotation editor facilitate the ability to analyze the text?*

The percentage of participants who agree that using the annotation editor facilitated the ability to analyze the text.

Figure 169: Question 28 of the questionnaire.

B.2.29 *Can the editor be used without training?*

The percentage of participants who agree that the editor can be used without training.

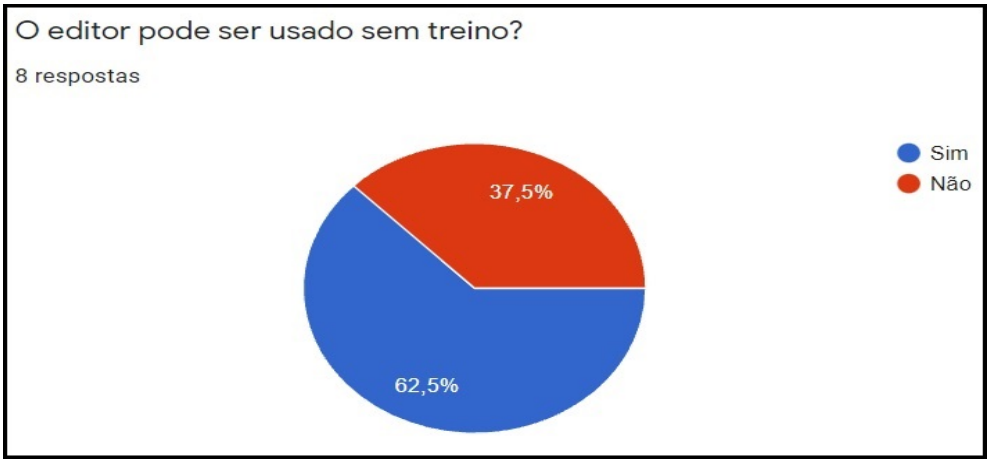


Figure 170: Question 29 of the questionnaire.

B.2.30 *Do you consider the editor useful in the process of analyzing the documents in the corpus?*

The percentage of participants who agree that the editor is helpful in the process of reviewing corpus documents.

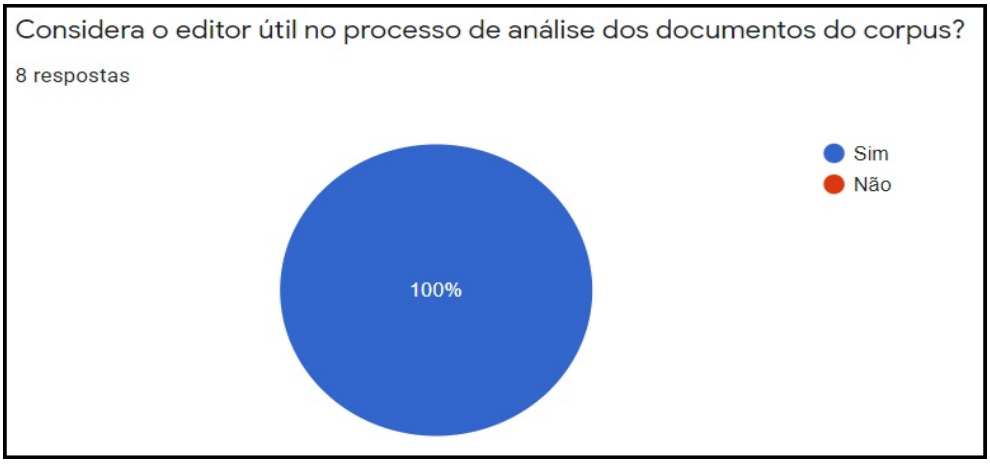


Figure 171: Question 30 of the questionnaire.

B.2.31 *If you have any comments or suggestions about the editor, please write here.*

Comments/suggestions given by the participants regarding the editor.

Caso tenha algum comentário ou sugestão sobre o editor, por favor escreva aqui.

4 respostas

[Tinha anotado observações num txt enquanto navegava. Colo aqui caso sejam úteis]

No geral funciona como esperado - parabéns pelo trabalho. Os pormenores/bugs a seguir são secundários, para contribuir rumo a aperfeiçoar não para criticar:

Geral:
- É mencionado que apenas foi testado no Chrome. Segui o guião usando o browser Firefox (+ recente): funcionou sem problemas (melhor que no Chrome).

- Botão "Download PDF":
a) Demora imenso a descarregar (23s no meu pc) e dá impressão de não funcionar (inicialmente pensei que não estava a funcionar pois carregava e não havia feedback, vim anotar isso aqui, quando voltei já tinha descarregado) - o idela seria alguma indicação/gif de loading e que demorará, para o utilizador perceber que está a processar.

b) Quando descarrega no browser Chrome, o ficheiro vem com o nome "_Youtube_extraction_portuguese_55.pdf_". O underscore final faz com que em máquinas windows não seja recohecido por default como PDF (para utilizadores menos experientes pode ser problemático).
Por oposição: quando descarrego no Firefox vem sem underscores ("Youtube_extraction_portuguese_55.pdf") e já é reconhecido automaticamente pelo OS do demónio (windows).

- No menu direito, na parte principal ("save", "import work" etc.) quando carrego em "Export Notes in PDF" abre-me janela de impressão e obriga a imprimir, não me dá um PDF.

Fase 1:
(Testado no Firefox e Chrome)
- A navegação entre as páginas (as setas no topo central "Current page") é um bocado lenta, há um lag considerável no carregamento (+/- 8s entre o clique na seta e o loading do novo grupo de comentários).
Nota: testei no Chrome e parece igual.

- No topo central permite saltar para o topo ou fim da página: para uma navegação interna com dados estáticos sem AJAX nem manipulação do DOM o feedback é um bocado lento (1.5s a saltar).
Nota: testei no Chrome e já funcionava rápido, o lag era só no Firefox.

Figure 172: First part of the question 31 of the questionnaire.

- O guião não é claro sobre como etiquetar uma palavra com uma tag: só por tentativa erro percebi que deveria seleccionar a palavra no texto com o rato e apareceria um botão em forma de pincél que permitia identificar tag. Inicialmente pensei que seria pelo painel direito (que se abre ao carregar nos 3 traços). Nota: esse pincel nem sempre aparece à primeira seleção, às vezes tenho de repetir a seleção.

- Encontrei uma palavra que era mais do que um tipo de tag: era ao mesmo tempo calão e ao mesmo tempo erro ortográfico ("bue" - é informal, mas ao mesmo tempo está mal escrito costuma ser "bué" com acento). Mas a interface parece-me só permitir identificar um tipo de cor para cada palavra. Coloquei vermelho mas também podia ter colocado laranja, mas defini internamente as duas tags. Não sei se devia haver orientanção no manual nesse tipo de casos (ou uma cor para conteúdos mistos).

- Foi necessário criar uma lista de tags antes de etiquetar - não era possível ter já algumas default mais comuns ou até fixas? Dei nomes à minha maneira ("erro_ortográfico", mas podia ser "erro" ou "erroOrto" p.e. etc.), não sei se nomes incongruentes dará algum problema mais tarde na identificação de dados na DB.

- Sempre que crio uma nova tag (quando ela não existe; no gestor de tags), não me é possível definir a cor por defeito associada. Isto resulta que para cada palavra tenho que indicar qual a tag e ao mesmo tempo mudar a cor no painel (às vezes esqueço-me e começo a anoto com a tag, só depois reparo na cor e tenho de apagar e recomeçar com a cor certa). Como já temos de criar as tags manualmente, nessa altura se desse para definir uma cor por defeito agilizaria na marcação.

- O sistema de cores não é claro: no topo existe já uma barra de cores; no entanto abrindo o menu direito de gestão das tags (que está colapsado por defeito) volta a aparecer outro botão para seleção de cores (o símbolo da paleta de cores) - ter a seleção de cores duas vezes repetidas (e não estão em sincronia, mudando num sítio não muda no outro) pode confundir.

- PS: sobre o ponto anterior: depois de explorar melhor descobri que a paleta era apenas um sistema de filtragem para exibir comentários já feitos só daquela cor. Por isso funciona como pretendido. Mas esta confusão inicial pode ser indicativo que outros utilizadores pensarão o mesmo - talvez colocando um título nesse menu lateral a indicar que é apenas um menu de navegação, não de marcação? Como não tem título não estava a perceber para que servia inicialmente.

- No seguimento do ponto anterior: em UI/UX usar apenas imagens não é muito intuitivo, p.e. os símbolos da paleta, ou do menu para criar novas tags, ou o central em forma de calendário, a funcionalidade não é imediatamente clara apenas pelo símbolo. O ideal seria p.e. uma tootip textual descritiva ao passar o rato por cima (p.e. usando o atributo "title" com estilização de tooltip).

- Sobre a barra de seleção de cores fixa, se p.e. tiver o vermelho ativo, seleccionar uma palavra e carregar no pincel para criar uma tag, mas só nessa altura reparar que afinal quero outra cor, mudando no painel superior não aplica na palavra - é necessário cancelar ("close"), mudar a cor no topo e recomeçar o processo da craição da tag. Idealmente deveria ser possível alterar a cor a meio do processo de marcação, não apenas antes de o iniciar.

- PS: sobre o ponto anterior, afinal permitia alterar a cor no botão ao lado das tags, eu é que tinha ficado demasiado focado na barra superior das cores. Mas deixo a indicação na mesma apenas para registar a primeira impressão da UI.

Figure 173: Second part of the question 31 of the questionnaire.

Fase 2:

- Os menus de filtragem funcionam bem, todos correram como esperado. Se fosse sugerir algo, seria apenas sobre o encerramento dos menus: depois de abrir um filtro, o menu só fecha se voltarmos a pressionar o botão do mesmo. Não encerra quando p.e. carregamos na parte vazia da página (que é o mais comum e a primeiro intuição que segui), só por último tentei carregar novamente no botão. No entanto funciona perfeitamente como está.

- Sobre localizar o comentário: seria mais intuitivo se este saltasse para o topo da página em vez do fundo. Mas funciona na mesma, não é relevante.

Sobre o guião:
- A primeira nota termina com "um tipo de problema digo de nota" -> penso que seria "digno de nota"

Demasiado tempo ao selecionar um botão. O tutorial solicita a cor roxo para ausência de concordância, mas essa cor não está listada. Botão edit de um texto já com tag, não funcionou.

> As opções de cabeçalho do menu lateral direito seriam mais claras com a adição de tooltips;
> Seria útil ter a possibilidade de ajustar a área realçada na anotação, após a sua criação (sem ter que eliminar a nota e criar uma nova com a área desejada).

Gostei do editor, mas considero-o muito pobre, pois só aponta três marcadores analíticos. O da falta de concordância verbal é capaz de ser o mais útil, para estudos de sintaxe. A questão dos erros talvez desse para um estudo sobre iliteracia. E o calão, claro, para uma abordagem sociolinguística. Mas há muitos mais aspetos que podem interessar este ou aquele. Portanto, talvez fizesse mais sentido deixar a ferramenta em aberto, permitindo ao analista adicionar não só a tag, mas a categoria da tag, consoante o que lhe interessasse analisar. É pena, também, que a ferramenta não tenha uma versão em inglês, uma vez que os linguistas da equipa são estrangeiros. Independentemente destes ligeiros óbices, obrigada aos alunos pelo esforço de criar esta interface! Isabel Ermida :)

Figure 174: Third and final part of the question 31 of the questionnaire.

BIBLIOGRAPHY

Maristella Agosti and Nicola Ferro. Annotations: Enriching a digital library. In *Research and Advanced Technology for Digital Libraries*, 2003.

James Blustein, David Rowe, and Ann-Barbara Graff. Making sense in the margins: A field study of annotation. In *Proceedings of the 15th International Conference on Theory and Practice of Digital Libraries: Research and Advanced Technology for Digital Libraries*, 2011.

Chih-Ming Chen, Ming-Chaun Li, and Tze-Chun Chen. A web-based collaborative reading annotation system with gamification mechanisms to improve reading performance. *Computers & Education*, 2019.

Raphael Cohen-Almagor. Countering hate on the internet. *Annual Review of Law and Ethics*, 2014.

Karmen Erjavec. "you don't understand, this is a new war!" analysis of hate speech in news web sites' comments. *Mass Communication and Society*, 2012.

Nor Fariza, Hazita Azman, and Afendi Hamat. Investigating students' use of online annotation tool in an online reading environment. *3L: Language, Linguistics, Literature*, 19:87–101, 2013.

Fei Gao. A case study of using a social annotation tool to support collaboratively learning. *The Internet and Higher Education*, 2013.

Ian Glover, Zhijie Xu, and Glenn Hardaker. Online annotation – research and practices. *Computers & Education*, 49(4):1308 – 1320, 2007.

David John Harvey. Dangerous speech - some legislative proposals. *SSRN Electronic Journal*, 2019.

Anjali Jindia and Sonal Chawla. Annotations in e-learning. *INTERNATIONAL JOURNAL OF COMPUTERS & TECHNOLOGY*, 12:3852–3859, 2013.

Ricardo Kawase, Eelco Herder, and Wolfgang Nejdl. A comparison of paper-based and online annotations in the workplace. In *Learning in the Synergy of Multiple Disciplines. EC-TEL 2009. Lecture Notes in Computer Science*, pages 240–253, 2009.

Jingyan Lu and Liping Deng. Examining students' use of online annotation tools in support of argumentative reading. *Australasian Journal of Educational Technology*, 2013.

Catherine C. Marshall. Annotation: From paper books to the digital library. In *Proceedings of the Second ACM International Conference on Digital Libraries*, 1997.

Silvana Neshkovska and Zorica Trajkova. The essentials of hate speech. *International Journal of Education TEACHER*, 2017.

Chikashi Nobata, Joel Tetreault, Achint Thomas, Yashar Mehdad, and Yi Chang. Abusive language detection in online user content. *Proceedings of the 25th International Conference on World Wide Web - WWW '16*, 2016.

Kenton O'Hara and Abigail Sellen. A comparison of reading paper and on-line documents. In *Proceedings of the ACM SIGCHI Conference on Human Factors in Computing Systems*, 1997.

Kenton O'Hara, Fiona Smith, William Newman, and Abigail Sellen. Student readers' use of library documents: Implications for library technologies. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, 1998.

ILIA A. OVSIANNIKOV, MICHAEL A. ARBIB, and THOMAS H. MCNEILL. Annotation technology. *International Journal of Human-Computer Studies*, 50(4):329 – 362, 1999.

Pedro Rangel Henriques, Cristiana Araújo, Isabel Ermida, and Idalete Dias. Scraping news sites and social networks for prejudice term analysis. In *Proceedings of the 16th International Conference on Applied Computing*, pages 179–189, 2019.

Rui Rodrigues, Cristiana Araújo, and Pedro Rangel Henriques. NetLangEd, A Web Editor to Support Online Comment Annotation. In Ricardo Queirós, Mário Pinto, Alberto Simões, Filipe Portela, and Maria João Pereira, editors, *10th Symposium on Languages, Applications and Technologies (SLATE 2021)*, volume 94 of *Open Access Series in Informatics (OASIcs)*, pages 15:1–15:16, Dagstuhl, Germany, 2021. Schloss Dagstuhl – Leibniz-Zentrum für Informatik. ISBN 978-3-95977-202-0. doi: 10.4230/OASIcs.SLATE.2021.15. URL https://drops.dagstuhl.de/opus/volltexte/2021/14432.

Bill Schilit, Gene Golovchinsky, and Morgan Price. Beyond paper: Supporting active reading with free form digital ink annotations. *Conference on Human Factors in Computing Systems - Proceedings*, 1999.

Alexandra A. Siegel. *Social Media and Democracy: The State of the Field, Prospects for Reform*, pages 56–88. Cambridge University Press, 2020.

Stefanie Ullmann and Marcus Tomalin. Quarantining online hate speech: technical and ethical perspectives. *Ethics and Information Technology*, 2019.

Joanna Wolfe. Annotations and the collaborative digital library: Effects of an aligned annotation interface on student argumentation and reading strategies. *I. J. Computer-Supported Collaborative Learning*, 2008.

Joanna Wolfe and CM Neuwirth. From the margins to the center - the future of annotation. *Journal of Business and Technical Communication*, 2001.