

# Text Analysis to Identify and Examine the Themes for Research in Selected Paulo Coelho's Novels

**\*Mrs. Vinodini S.**

Research Scholar – Faculty of Science & Humanities (English),  
Sathyabama Institute of Science and Technology,  
Chennai – 600119. Mail ID – vinodini.adi@gmail.com  
[ORCID: 0000-0002-2605-4811]. (**\*Corresponding Author**)

**Dr. Avinash M.**

Assistant Professor – Department of English,  
Dharmamurthi Rao Bahadur Calavala Cunnan Chetty Hindu College,  
Chennai – 600072. Mail ID – dravinash1981@gmail.com

## Abstract

This study presents a comprehensive analysis of the themes of travel, journey, spiritual journey, and physical journey to achieve one's destiny in selected novels of Paulo Coelho. By utilizing text analysis, this study reveals that these themes are consistently present throughout Coelho's work and are explored in a variety of ways. The quantitative analysis of raw and relative frequencies of words related to these themes offers insight into their significance in the novels, while the collocates provide contextual information and correlations suggest the relationships between them. The study emphasizes the importance of using text analysis as a guide for further research and analysis, allowing researchers to delve deeper into the nuances of these themes and their representations in Coelho's work. Researchers can use this analysis to investigate the use of travel as a metaphor for personal growth and self-discovery, or to explore the role of spiritual journeys in achieving one's destiny. Additionally, the analysis can be used to compare and contrast the representations of these themes across different novels or to examine how they interact with other themes and motifs in Coelho's works. Overall, this text analysis of Paulo Coelho's novels provides a rich source of data for researchers of literature who are interested in studying the themes of travel, journey, spiritual journey, and physical journey to achieve one's destiny. It provides a starting point for more detailed analysis and offers a framework for exploring the complex and multifaceted ways in which these themes are represented in Coelho's work.

**Keywords:** Text analysis, Travel, Journey, Physical, Spiritual, Voyant Corpus, Voyant Terms, Voyant Trends

## 1. Introduction

Text analysis [1] is crucial in researching literary texts in English literature as it helps researchers to gain a deeper understanding of their meaning, structure, and significance. It enables them to examine the literary techniques and devices used in a text, identify themes and motifs, and compare and contrast different literary works. Computational text analysis provides several advantages over manual analysis as it enables researchers to analyze large volumes of texts quickly and efficiently, offers more objective results, and allows for more precise and accurate analysis.

Paulo Coelho's novels [18] have gained worldwide popularity due to their universal themes of self-discovery, spirituality, and personal growth. Researching his works can provide a deeper understanding of his creative process, literary techniques, and cultural and historical contexts. Text analysis is essential in this research as Coelho's novels are complex, multilayered, and contain multiple themes, symbols, and motifs. It enables researchers to examine the literary techniques [19] used by Coelho, explore the cultural and historical contexts in which his works were written, and gain insights into his worldview and the influences that shaped his writing.

Manual text analysis can be challenging when researching Paulo Coelho's novels as his works are often multilayered and complex, making it difficult to interpret through manual analysis. Additionally, manual analysis can be time-consuming, and the interpretation can be subjective and prone to personal biases. Computational text analysis can solve these problems by enabling researchers to process a large volume of texts quickly and efficiently, identifying patterns and themes in the text, and offering objective and unbiased analysis.

Text analysis is a crucial tool for researching literary texts in English literature, and computational text analysis offers several advantages over manual analysis. Researching Paulo Coelho's novels using text analysis is essential as it enables researchers to gain a deeper understanding of his works' complex themes and literary techniques. Computational text analysis can provide a more systematic, comprehensive, and objective approach to analyzing Coelho's novels, complementing and enhancing traditional manual analysis methods.

## 2. Selected Novels of Paulo Coelho

The below is the six novels of Paulo Coelho selected for this research work. The overall summary of each novel is explained in below sub sections.

### 3.1 Aleph

Paulo Coelho's novel Aleph [3] is a story of self-discovery and spiritual awakening. The protagonist embarks on a physical journey across Asia to find a lost friend and heal past wounds. During his journey, he reflects on his past and confronts his own ego to find inner peace and enlightenment. The novel explores the themes of spirituality, destiny, and self-discovery through the protagonist's journey.

### 3.2 Brida

Brida [4] is a novel about a young Irish girl who is on a spiritual journey to discover her true self. She is introduced to the world of magic and witchcraft by her mentor and lover, and her journey leads her to explore her spirituality and innermost desires. Brida's journey is a combination of physical and spiritual elements, with her travels taking her to different parts of the world. The novel explores

the themes of spirituality, love, and self-discovery through Brida's journey.

### 3.3 By the River Piedra I Sat Down and Wept

This novel [5] follows the journey of a woman who sets out on a physical and spiritual journey to find her lost love. Her journey takes her to the Pyrenees, where she meets a spiritual teacher who helps her to reconnect with her faith and spirituality. The novel explores the themes of love, spirituality, and personal growth through the protagonist's journey.

### 3.4 The Alchemist

This novel [6] follows the journey of a shepherd boy who embarks on a physical and spiritual journey to fulfill his destiny. His travels take him across the desert, where he meets various people who teach him valuable lessons about life, love, and the pursuit of one's dreams. The novel explores the themes of destiny, spirituality, and personal growth through the protagonist's journey.

### 3.5 The Fifth Mountain

This novel [7] is a fictionalized account of the prophet Elijah's journey. Elijah is on a physical and spiritual journey to find his destiny, and his travels take him through the mountains of Israel. Along the way, he confronts various challenges and learns valuable lessons about faith, spirituality, and the human condition.

### 3.6 The Pilgrimage

This novel [8] follows the journey of a man who embarks on a physical and spiritual pilgrimage across Spain to find spiritual enlightenment. His journey takes him through various towns and cities, where he meets people who help him to overcome his fears and confront his inner demons. The novel explores the themes of spirituality, self-discovery, and personal growth through the protagonist's journey.

All six novels explore the theme of journeys [2] in various forms. The characters in these novels embark on physical, spiritual, or both journeys to reach their destinies. The novels cover travel to different countries and the protagonists' internal journeys. In Aleph, the protagonist embarks on a journey across Siberia and Russia to find enlightenment and himself. Brida follows the story of a young Irish girl's journey to find her soulmate and her path in life through the teachings of a wise woman. In By the River Piedra I Sat Down

and Wept, the protagonist goes on a spiritual journey to find her lost love and to discover her true self. The Alchemist [11] follows a shepherd boy's journey from Spain to Egypt in search of his treasure and his personal legend. The Fifth Mountain depicts the story of a prophet's physical and spiritual journey in ancient Israel, and The Pilgrimage covers the journey of a modern-day pilgrim on the Camino de Santiago.

### **3. Text analysis on selected Novels**

Text analysis allows researchers to identify patterns and trends within the selected novels that may not be immediately apparent through traditional close reading methods, and to quantify the frequency of certain words or phrases. Six significant novels are to be explored through text analysis, each centering on characters who embark on physical and/or spiritual journeys in order to find their destinies. The novels showcase the importance of self-discovery and finding one's purpose in life, and text analysis can help reveal the language, imagery, and larger themes present in each novel. Text analysis is an effective and valuable method for exploring the complex themes present in these novels related to travel, journey, spiritual journey, and physical journey to reach destiny in further sections.

In this research, the Voyant tool [9] was used to analyze six novels by Paulo Coelho, focusing on identifying the most frequent words and themes across the texts. The tool provided insights into the main themes and concepts explored in the novels and allowed for efficient analysis of the large corpus of text data. The Voyant tool is a powerful web-based text analysis tool that enables users to analyze large volumes of text data quickly and easily. It provides a range of analytical tools and generates easy-to-read visualizations, making it an essential tool for researchers and professionals in various fields. It allows users to upload large volumes of text data, making it possible to analyze multiple texts simultaneously. This saves time and increases efficiency, especially when compared to manual text analysis techniques. The tool provides users with a range of analytical tools, including word frequency, keyword extraction, and collocation analysis, which enables users to extract meaningful insights from the texts. These tools help users to identify patterns, themes, and significant

information within the texts, which would otherwise be difficult to identify through manual analysis. This tool generates easy-to-read visualizations, which allow users to quickly understand and interpret the analyzed data. These visualizations include word clouds, trends graphs, and concordance charts, which provide users with a quick overview of the most frequent terms and their distribution across the texts. The Voyant tool [10] is easy to use, and users do not require advanced technical skills or programming knowledge to use it. This makes it accessible to a wide range of users, including students, researchers, and professionals. Voyant tool is an efficient text analysis tool that provides users with numerous benefits, including the ability to analyze large volumes of text data, a range of analytical tools, easy-to-read visualizations, and ease of use. These advantages make it an indispensable tool for text analysis in various fields, including literature, social sciences, and business.

### **4. Computational text analysis on selected Novels using Voyant tool**

#### **5.1 Research framework using Voyant tool**

The table 1 outlines the research framework used for analyzing the six novels by Paulo Coelho using the Voyant tool [20]. The six selected novels were downloaded in PDF format and ensured their authenticity by checking for any unwanted watermarks in first step. The next step is to upload the PDFs into the Voyant tool to create a corpus, which is a collection of texts that will be analyzed together. Once the corpus is created, the text analysis process can begin using the various analytical tools available in the Voyant tool. In the third step, a range of analytical tools are applied to the text data, including word frequency analysis, keyword extraction, and collocation analysis. In the fourth step, the Voyant tool generates visualizations of the analyzed data, such as word clouds, trends graphs, and concordance charts. In the fifth step, the analyzed data is interpreted to identify the most frequent words and themes in the novels. In the sixth step, the results of the analysis are evaluated against the research question to determine the effectiveness of the Voyant tool in answering the research question. Finally, in the seventh step, the research is concluded by summarizing the main findings and providing

recommendations for future research using the Voyant tool. This framework provides a structured approach for analyzing the text data and enables

the efficient and effective use of the Voyant tool for text analysis.

**Table 1. Research Framework for doing text analysis on selected Paulo Coelho’s novels**

S.No	Step	Explanation
1	Data Preparation	Collect the text data from the novel and clean it for consistency.
2	Data Upload	Upload the text data into the Voyant tool in a compatible format.
3	Text Analysis	Apply word frequency analysis, keyword extraction, and collocation analysis tools to analyze the text data.
4	Data Visualization	Generate word clouds, trends graphs, and concordance charts to visualize the analyzed data.
5	Interpretation	Interpret the analyzed data to identify the most frequent words and themes in the novel.
6	Evaluation	Evaluate the results of the analysis against the research question to determine the effectiveness of the Voyant tool in answering the research question.
7	Discussion	Conclude the research by summarizing the main findings and providing recommendations for future research using the Voyant tool.

## 5.2

### Text analysis parameters for this research work

The selected parameters [12] in Table 2 are crucial for analyzing text as they provide valuable insights into the characteristics of the text, such as its length, complexity, readability, richness, and diversity of vocabulary. Word Count provides an idea of the scope of the text and how much time and effort it would take to read and analyze it. Vocabulary Density helps to understand the uniqueness and diversity of the vocabulary used in the text, which is important for identifying the unique features of the text and the author's style. Average Words Per Sentence provides insights into the complexity of the text and its readability, which is important for understanding the target audience and the purpose of the text. Readability Index measures how easy or difficult it is to read and

comprehend the text, which is crucial for evaluating the effectiveness of the text and its suitability for the intended audience. Distinctive Words are important for identifying the unique features of the text and the author's style, which can help to distinguish the text from other texts in the corpus. Raw Frequency and Relative Frequency provide insights into the prominence and importance of different words in the text, which can help to identify the most common themes and topics in the text. Collocates are important for identifying the common patterns of word usage in the text, which can help to understand the context and meaning of individual words. Correlations help to identify the relationships between different words in the text, which can help to understand the underlying themes and ideas in the text.

**Table 2. Text analysis parameters considered in this research work on selected Paulo Coelho’s novels**

S.No	Parameter	Explanation
1	Word Count	The total number of words present in the text. It gives an idea about the length of the text.
2	Vocabulary Density	The ratio of the unique words to the total number of words in the text. It helps to understand the richness and diversity of the vocabulary used in the text.
3	Average Words per Sentence	The average number of words per sentence in the text. It helps to gauge the complexity of the text and its readability.
4	Readability Index	A measure of how easy or difficult it is to read and comprehend the text. Voyant tool uses different algorithms to calculate the readability index, such as the Flesch-Kincaid Grade Level and the Gunning-Fog Index.

5	Distinctive Words	The words that occur more frequently in a particular text compared to other texts in the corpus. It helps to identify the unique features of a text.
6	Raw Frequency	The number of times a particular word occurs in the text. It gives an idea about the prominence of a word in the text.
7	Relative Frequency	The frequency of a particular word normalized by the total number of words in the text. It helps to compare the importance of different words in the text.
8	Collocates	The words that are commonly used together with a particular word in the text. It helps to identify the common patterns of word usage in the text.
9	Correlations	Identifies the relationship between different words in the text. It helps to understand the context in which a particular word is used and its association with other words.

The selected parameters are particularly relevant for analyzing Paulo Coelho's selected novels, as his writing style is characterized by a unique vocabulary and distinct themes. Understanding the richness and diversity of the vocabulary he uses is crucial for identifying the unique features of his texts. Moreover, his writing often includes complex sentences that can affect the readability of his texts. Therefore, Average Words per Sentence and Readability Index are important for understanding his target audience and purpose of the text. Distinctive Words are important for identifying the unique features of Coelho's texts, which often include spiritual and philosophical themes. Raw Frequency and Relative Frequency can help to identify the most common themes and topics in his novels. Collocates and

Correlations can provide insights into the patterns of word usage and relationships between different words in his texts, helping to understand the underlying themes and ideas in his writing.

### 5. Text analysis results of selected Novels using Voyant tool

After opening the Voyant's website (<https://voyant-tools.org/>). Click on "Upload" and select the PDF files for the six novels: "Aleph", "Brida", "By the River Piedra I Sat Down and Wept", "The Alchemist", "The Fifth Mountain", and "The Pilgrimage". Then, Wait for Voyant to process the texts. Once the texts are processed, the exploration starts on different features using Voyant's various tools. Here are some of the results which were observed during the analysis of the novels in Voyant tool.

### 6.1 Corpus Summary

The screenshot shows the 'Summary' page of the Voyant Tools interface. It provides a detailed overview of a corpus containing 6 documents with 329,030 total words and 19,134 unique word forms. Key statistics include document length, vocabulary density, average words per sentence, and readability index, all broken down by individual novel. A list of the most frequent words and a list of distinctive words (compared to the rest of the corpus) are also displayed.

**Document Length:**

- Longest: *The Pilgrimage* (68868); *Aleph* (65536); *Brida* (58063); *The Fifth Mountain* (53146); *By the River Piedra I...* (41749)
- Shortest: *The Alchemist* (41668); *By the River Piedra I...* (41749); *The Fifth Mountain* (53146); *Brida* (58063); *Aleph* (65536)

**Vocabulary Density:**

- Highest: *Aleph* (0.153); *By the River Piedra I...* (0.141); *The Pilgrimage* (0.091); *The Alchemist* (0.091); *The Fifth Mountain* (0.087)
- Lowest: *Brida* (0.087); *The Fifth Mountain* (0.087); *The Alchemist* (0.091); *The Pilgrimage* (0.091); *By the River Piedra I...* (0.141)

**Average Words Per Sentence:**

- Highest: *The Pilgrimage* (17.8); *Brida* (16.4); *Aleph* (14.7); *The Alchemist* (14.4); *The Fifth Mountain* (14.4)
- Lowest: *By the River Piedra I...* (13.9); *The Fifth Mountain* (14.4); *The Alchemist* (14.4); *Aleph* (14.7); *Brida* (16.4)

**Readability Index:**

- Highest: *Aleph* (9.505); *By the River Piedra I...* (8.285); *The Fifth Mountain* (7.935); *The Alchemist* (7.262); *Brida* (7.257)
- Lowest: *The Pilgrimage* (7.118); *Brida* (7.257); *The Alchemist* (7.262); *The Fifth Mountain* (7.935); *By the River Piedra I...* (8.285)

Most frequent words in the corpus: **said** (1511); **time** (863); **know** (773); **boy** (677); **people** (676)

**Distinctive words** (compared to the rest of the corpus):

- Aleph*: **hila** (187), **yao** (121), **rst** (78), **aleph** (75), **nd** (73).
- Brida*: **wicca** (237), **brida** (367), **mate** (76), **lorens** (62), **magus** (160).
- By the River Piedra I...*: **padre** (52), **zaragoza** (21), **savin** (21), **bilbao** (17), **bernadette** (17).
- The Alchemist*: **englishman** (66), **camel** (36), **oasis** (53), **fatima** (29), **omens** (45).
- The Fifth Mountain*: **elijah** (420), **governor** (142), **akbar** (174), **assyrians** (59), **jezebel** (48).
- The Pilgrimage*: **petrus** (390), **agape** (55), **dog** (97), **ram** (50), **exercise** (118).

Figure 1. Corpus Summary of Six selected Novels using Voyant tool

One of Voyant's most potent features is the Corpus Summary tool [14], which is depicted in Figure 1. It gives a comprehensive overview of the six novels analyzed as a single corpus. The tool provides essential details such as the total number of words utilized in the six texts, the unique words featured in them, and the richness of the vocabulary in each of the six novels. This information is beneficial in comparing and contrasting different texts or identifying unique linguistic and thematic features unique to a particular author or genre. By utilizing the Corpus Summary tool, users can gain a better understanding of the linguistic diversity and complexity of the analyzed texts. The tool can also help in determining the author's writing style and the thematic content of the texts. The Corpus Summary tool is a valuable tool that provides users with an insight into the overall structure and content of the analyzed texts, enabling them to extract meaningful information and derive useful conclusions.

### 6.2 Voyant Terms

The Voyant Terms tool [15], as shown in Figure 2, is highly effective in identifying the key themes, motifs, and topics present in the six analyzed texts. By analyzing the most frequently occurring words, users can gain a comprehensive understanding of the language utilized by Paulo Coelho and the crucial ideas and concepts conveyed through the texts. The tool can also identify any patterns or trends in the language, such as repeated words or phrases across the six novels. Moreover, the Terms tool can filter words based on their part of speech, enabling users to focus on specific types of words and obtain a more targeted understanding of the language employed in the texts. The Terms tool's power lies in its ability to provide insights into the linguistic and thematic features of the six novels. By analyzing the most frequently occurring words and filtering by part of speech, users can gain a deeper comprehension of the language used in the texts and the significant themes and ideas that Coelho presents.

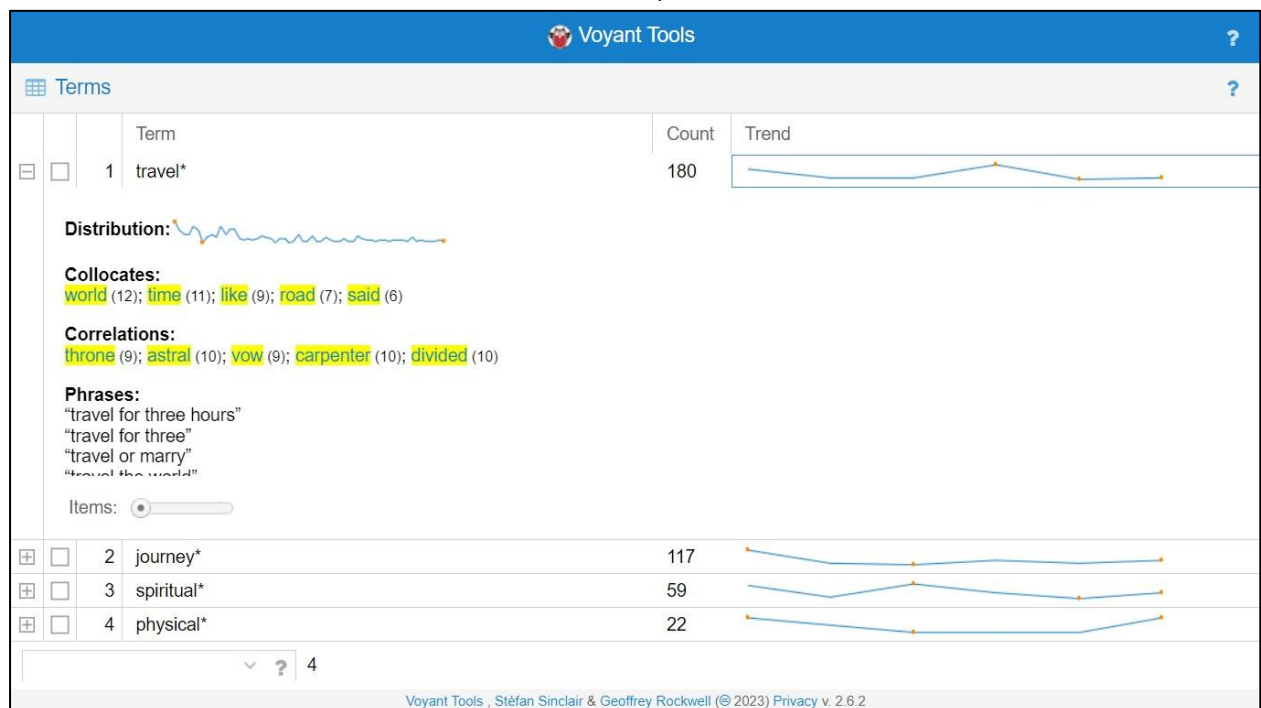


Figure 2. Terms of Six Selected Novels using Voyant tool

### 6.3

#### Voyant Trends

The Voyant Tool's Trends [13] feature as in Figure 3, is an essential resource for analyzing the frequency and context of specific words or phrases within the six novels. With this tool, users can

identify patterns or trends in the use of words or phrases throughout the texts, highlighting recurring themes, ideas, and motifs present within them. The Trends feature represents this data visually, presenting a graphical overview of word or phrase

frequency throughout the six novels. This visualization simplifies the interpretation of text

trends, enabling users to draw meaningful insights from the data.



Figure 3. Trends of Six Selected Novels using Voyant tool

By analyzing the Trends feature, users can also identify significant shifts or changes in the usage of words or phrases across different sections of the six novels. This feature's ability to analyze specific words or phrases within the texts provides valuable insights into their linguistic and thematic features. By recognizing the frequency and context of words or phrases, users can gain a more comprehensive understanding of the author's writing style, the themes conveyed in the texts, and the significance of particular words or phrases within the overall narrative. Overall, the Voyant Tool's Trends feature is an indispensable tool for identifying and interpreting patterns or trends in word or phrase usage across the six novels. Its visual representation of the data provides users with an accessible means of analyzing trends in the texts, making it easier to understand and interpret their linguistic and thematic features.

6. Writing style of Coelho on selected Novels

The table 3 provides key parameters of Paulo Coelho's selected novels that are relevant to literature research. The word count ranges from

41,668 to 68,868, indicating that the selected novels are of moderate to long length. The vocabulary density ranges from 9% to 15%, which is relatively low, suggesting that the author uses a limited range of vocabulary. However, this could also be due to the fact that the novels are translated into English from the author's native language. The average words per sentence range from 13.9 to 17.8, indicating that the texts are moderately complex and may require a certain level of reading comprehension. The readability index ranges from 7.118 to 9.505, indicating that the texts are moderately difficult to read and may require a certain level of reading comprehension. The table provides valuable insights into the key parameters of Paulo Coelho's selected novels that are relevant to literature research, including word count, vocabulary density, average words per sentence, and readability index. These parameters can be used to evaluate the complexity, richness, and readability of the texts, as well as to identify the unique features of the author's writing style.

Table 3. Writing style of Paulo Coelho on selected novels

S.No	Novel	Word Count	Vocabulary Density	Average Words Per Sentence	Readability Index
1	Aleph	65,536	15%	14.7	9.505

2	Brida	58,063	9%	16.4	7.257
3	By the River Piedra I Sat Down and Wept	41,749	14%	13.9	8.285
4	The Alchemist	41,668	9%	14.4	7.262
5	The Fifth Mountain	53,146	9%	14.4	7.935
6	The Pilgrimage	68,868	9%	17.8	7.118

### 7. Distinctive words on selected Novels

Table 4 gives a detailed analysis on the top five distinctive words from each of the six novels with its relative themes.

In the novel "Aleph," the word "hilal" is an Arabic word meaning "crescent moon" and is often associated with spiritual growth and transformation. In the context of the novel, it may symbolize the protagonist's personal journey towards self-discovery and spirituality. The word "yao" is a Chinese word that means "friendship" or "friend." This word is likely related to the theme of personal growth, as the protagonist's journey involves meeting new people and forming meaningful relationships. The word "aleph" is a Hebrew word that means "the first letter of the

alphabet" and is often used symbolically to represent the beginning of a journey or the search for knowledge. This word is likely related to the themes of self-discovery, destiny, and travel in the novel. The word "rst" could refer to the acronym for "rapid serial visual presentation," which is a method of speed-reading. This word could be related to the theme of personal growth, as the protagonist in the novel is on a journey to learn more about himself and the world around him. The word "nd" could refer to the verb "find," which is often associated with the themes of self-discovery and personal growth. It may symbolize the protagonist's search for meaning and purpose in his life.

**Table 4. Distinctive Words on selected Paulo Coelho's novels**

S.No	Novel	Distinctive Words (Top 5)	Themes
1	Aleph	hilal (187), yao (121), rst (78), aleph (75), nd (73)	Personal growth, spirituality, self-discovery, destiny, travel
2	Brida	wicca (237), brida (367), mate (76), lorens (62), magus (160)	Witchcraft, spirituality, love, destiny, self-discovery
3	By the River Piedra I Sat Down and Wept	padre (52), zaragoza (21), savin (21), bilbao (17), bernadette (17)	Love, faith, religion, spirituality, pilgrimage
4	The Alchemist	englishman (66), oasis (53), omens (45), camel (36), fatima (29)	Personal legend, destiny, self-discovery, spirituality, travel
5	The Fifth Mountain	elijah (420), akbar (174), governor (142), assyrians (59), jezebel (48)	Faith, perseverance, religion, spirituality, leadership
6	The Pilgrimage	petrus (390), exercise (118), dog (97), agape (55), ram (50)	Spiritual awakening, personal transformation, self-discovery, faith, pilgrimage

In "Brida," the relationship between the distinctive words and themes is a key aspect of the story. The term "wicca" is closely tied to the theme of witchcraft, which is a significant part of the novel's plot. As the story unfolds, Brida, the main character, embarks on a journey of self-discovery and spiritual growth, making the theme of self-discovery highly correlated with the novel. The

traditional South American drink "mate" is also mentioned in the novel and is somewhat correlated with the theme of spirituality, as it is often used in spiritual and ceremonial contexts. The name "lorens" is another distinctive word in the novel and is somewhat correlated with the theme of love, as it is the name of one of the main character's love interests. The word "magus" is highly correlated



with the theme of spirituality, as it is a term used to describe a magician or sorcerer, often associated with spiritual or mystical practices.

In "By the River Piedra I Sat Down and Wept," the word "padre" is highly correlated with the theme of religion and faith, as it is the Spanish word for father and often used to refer to Catholic priests. This theme is further reinforced throughout the novel as the main character undertakes a pilgrimage to reconnect with her faith and spirituality. The word "zaragoza" is somewhat correlated with the theme of pilgrimage, as it is a city in Spain that is home to several important religious sites, including the Basilica of Our Lady of the Pillar, which is a popular destination for pilgrims. The novel's setting in Spain and its focus on religious pilgrimage also reinforces this theme. The words "savin," "bilbao," and "bernadette" are less strongly correlated with the novel's themes. "Savin" may be somewhat associated with the theme of spirituality, as it is a type of evergreen tree that has been used in traditional medicine and spiritual practices. "Bilbao" is a city in Spain and does not have an obvious connection to the novel's themes. "Bernadette" may be somewhat associated with the theme of religion, as it is the name of Saint Bernadette Soubirous, a French Catholic saint who is associated with visions of the Virgin Mary.

In "The Alchemist," the word "englishman" is correlated with the theme of travel, as he is a character who is traveling through the desert in search of knowledge. This theme is reinforced throughout the novel as the main character travels from Spain to Egypt in pursuit of his personal legend. The word "oasis" is highly correlated with the theme of destiny and self-discovery, as it is a place of rest and rejuvenation that often symbolizes a turning point in the main character's journey towards fulfilling his personal legend. The main character's encounter with an oasis also leads him to meet Fatima, a character who becomes integral to his self-discovery and spiritual growth. The word "omens" is highly correlated with the theme of spirituality, as it represents the signs and symbols that the main character uses to guide him on his journey. Omens serve as a reminder of the interconnectedness of all things and the power of the universe to guide individuals towards their

destiny. The word "camel" is somewhat correlated with the theme of travel, as it is the primary mode of transportation used by the characters in the novel to traverse the desert. The word "Fatima" is highly correlated with the themes of personal legend and love, as she is the woman that the main character falls in love with and who helps him to understand the importance of pursuing one's dreams. Her presence in the novel serves to remind the main character that love is a powerful force that can guide individuals towards their true purpose in life.

In "The Fifth Mountain," the word "Elijah" is highly correlated with the theme of faith and spirituality, as it is the name of the novel's protagonist, who is a prophet in the Hebrew Bible. This theme is further reinforced throughout the novel as Elijah struggles with his faith and attempts to find his place in the world. The word "Akbar" is somewhat correlated with the theme of leadership, as it is a title meaning "great" that was used by several rulers in India and Iran. In the novel, Akbar is the name of a wise and influential character who helps guide Elijah on his journey. The word "governor" is highly correlated with the theme of leadership, as it refers to a person who is in charge of a territory or region. In the novel, the governor is a powerful figure who challenges Elijah's faith and presents obstacles for him to overcome. The words "Assyrians" and "Jezebel" are less strongly correlated with the novel's themes. "Assyrians" may be somewhat associated with the theme of perseverance, as they were a powerful civilization that endured for centuries. "Jezebel" is a biblical name that may be somewhat associated with the theme of religion, as Jezebel was a queen who worshiped foreign gods and was considered a symbol of evil in the Hebrew Bible.

In "The Pilgrimage," the word "Petrus" is highly correlated with the theme of spiritual awakening and personal transformation, as he is the mentor and guide for the main character on his pilgrimage journey. The theme of self-discovery is also prominent throughout the novel, as the main character embarks on a journey of self-discovery and transformation. The word "exercise" is somewhat correlated with the theme of self-discovery, as physical exercise is often used as a way to improve mental and emotional well-being

and aid in self-discovery. The word "dog" is less strongly correlated with the novel's themes, but it may be associated with the theme of companionship and loyalty, as the main character has a dog who accompanies him on his journey. The word "agape" is highly correlated with the theme of faith and spirituality, as it is a Greek term used in Christianity to describe a selfless and unconditional love. Finally, the word "ram" may be somewhat associated with the theme of pilgrimage, as the ram is a symbol of sacrifice and surrender, which are important elements of many religious pilgrimage traditions.

### 8. Raw and Relative frequency of themes in selected Novels

Raw and relative frequencies can help literary researchers in analyzing six novels by providing a

quantitative basis for identifying and comparing recurring themes across the novels. Raw frequency refers to the number of occurrences of a specific theme in each novel, while relative frequency refers to the proportion of the occurrences of that theme relative to the total number of themes in each novel. By calculating the raw and relative frequencies of themes such as travel, journey [17], spiritual journey [16], and physical journey, researchers can gain insight into the thematic preoccupations of the novels and compare the frequency of these themes across the different works. This allows for a more systematic and objective analysis of the texts and can help researchers identify patterns and trends in the portrayal of certain themes and motifs.

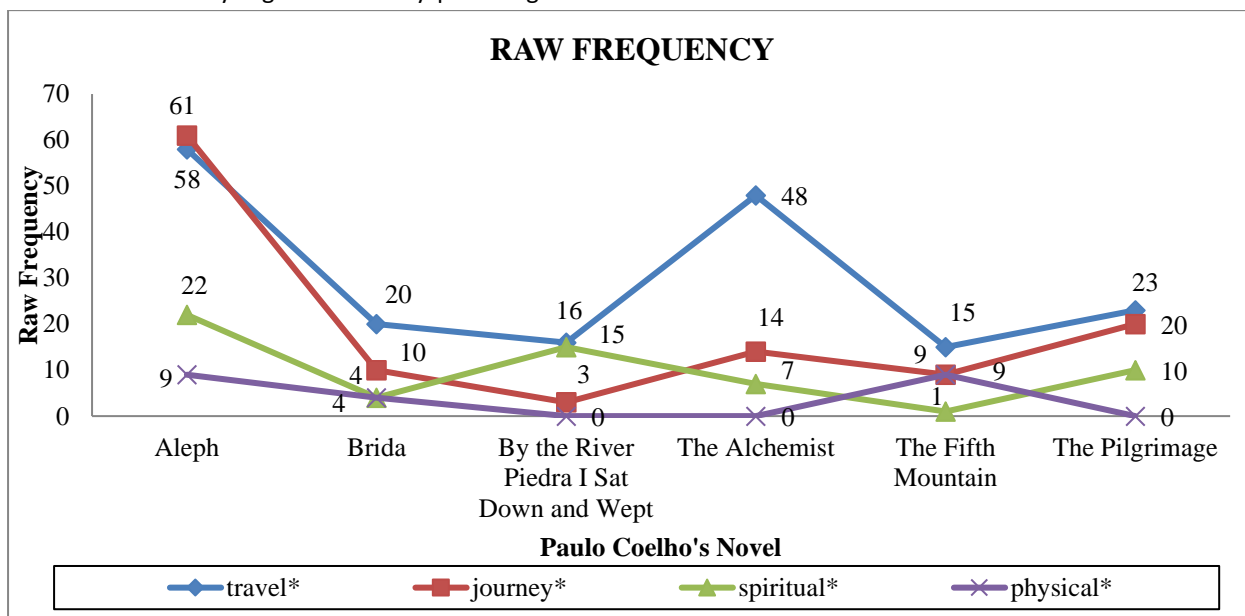


Figure 4. Raw frequency of themes in Selected Novels

The raw frequency of the themes of travel, journey, spiritual journey, and physical journey in the six selected novels is illustrated in Figure 4. The data demonstrates that the most prevalent theme across all six novels is travel, with a total raw frequency of 180. Journey is the second most common theme, with a raw frequency of 117, followed by spiritual journey, which has a total raw frequency of 59, and physical journey, which has a total raw frequency of 22. Aleph has the highest frequency for the theme of travel, with a raw frequency of 58. The novel centers around the author's journey through various countries as he seeks spiritual enlightenment, so this result is unsurprising. The Alchemist also has a relatively

high frequency of travel, which is not unexpected given that the protagonist travels from Andalusia to the Egyptian desert in search of treasure. The theme of journey is most commonly found in Aleph and The Pilgrimage, with a raw frequency of 61 and 20, respectively. Both novels portray the protagonist's journey towards self-discovery and spiritual growth. Brida and The Alchemist also have a significant frequency of the journey theme, which is to be expected, as the protagonists in these novels embark on personal journeys towards self-discovery and fulfillment. The theme of spiritual journey is most commonly found in By the River Piedra I Sat Down and Wept, with a raw frequency of 15. The novel describes the protagonist's

spiritual journey towards love and faith. Aleph and The Pilgrimage also have a relatively high frequency of spiritual journey, which aligns with the novels' themes of self-discovery and spiritual growth. The theme of physical journey is most common in The Fifth Mountain, with a raw frequency of 9. The novel depicts the physical journey of the prophet Elijah as he flees persecution and seeks refuge. The

results demonstrate that travel and journey are the most common themes in the six novels, followed by spiritual journey and physical journey. These themes reflect the characters' quests for personal growth, self-discovery, and fulfillment, and the physical and spiritual obstacles they encounter on their journeys.

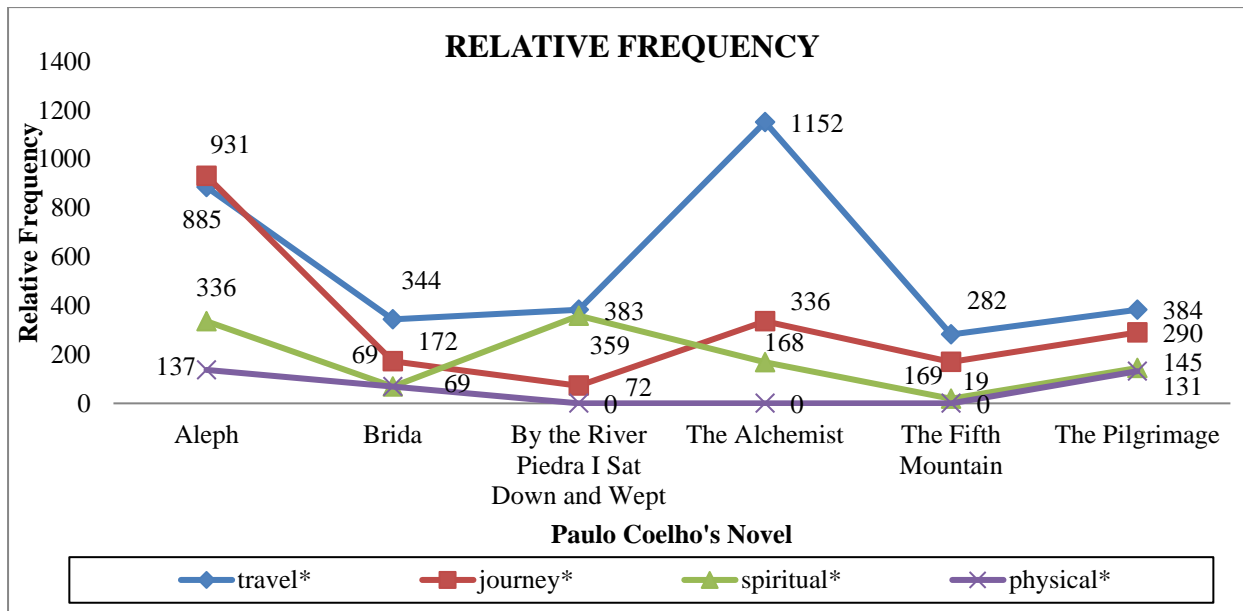


Figure 5. Relative frequency of themes in Selected Novels

The figure 5 shows the relative frequency of the themes of travel, journey, spiritual journey, and physical journey in six selected novels. For instance, in Aleph, the theme of travel has a relative frequency of 88.5%, which means that travel is the most dominant theme in the novel, occurring in almost 9 out of 10 themes. Similarly, in Brida, the theme of journey has a relative frequency of 34.4%, which indicates that journey is the most common theme in the novel, but it is not as dominant as travel in Aleph. By comparing the relative frequencies across the six novels, we can observe that the dominant themes in each novel differ. For example, while travel is the most common theme in Aleph, The Alchemist, and The Pilgrimage, it is journey that dominates in Brida and The Fifth Mountain. In By the River Piedra I Sat Down and Wept, spiritual journey is the most dominant theme, with a relative frequency of 35.9%, which is considerably higher than the other novels. The relative frequency table provides a more nuanced perspective on the themes in the novels, showing not only the most common themes but also the

proportion of these themes to the total number of themes in each novel.

### 9. Collocates & Correlations of the themes in the selected Novels

The Table 5 shows the analysis of the raw frequency, relative frequency, collocates, and correlations of the theme words "travel", "journey", "spiritual" and "physical" or its variations in Paulo Coelho's novels. This provides insight into the importance of travel, journey, spiritual journey and physical journey as a theme in his writing and the various ways in which it is explored throughout his works.

The theme of travel is present in all selected Paulo Coelho's novels, with the word "travel" or its variations appearing multiple times in each of his novels. The raw frequency of the word "travel" or its variations ranges from 15 to 58, and the relative frequency ranges from 282 to 1152. This suggests that travel is an important theme in Coelho's writing. In terms of collocates, we can see that the word "travel" frequently appears alongside words like "world," "time," "road," "throne," "astral," "vow," "carpenter," and "divided." These collocates

provide insight into the context in which the word "travel" is used and suggest that the theme of travel is often intertwined with themes of personal growth, destiny, and spirituality in Coelho's writing. In terms of correlations, it is seen that the entire corpus have strong correlations between the word "travel" and other words. The word "travel" is strongly correlated with the words "throne," "astral," "vow," "carpenter," and "divided." This suggests that the theme of travel in this book is particularly focused on spiritual growth and exploration.

The theme of journey appears to be present in all of his selected novels, with the word "journey" or its variations appearing multiple times in each novel. The raw frequency of the word "journey" or its variations ranges from 3 to 61, while the relative frequency ranges from 72 to 931, indicating that

journey is an important theme in Coelho's writing. The collocates listed in the table are the words that frequently appear alongside the word "journey" or its variations. The word "journey" is frequently accompanied by words like "make," "time," "end," "like," and "continue." These collocates provide insights into the context in which the word "journey" is used and suggest that the theme of journey is often intertwined with themes of personal growth, exploration, and destiny in Coelho's writing. In terms of correlations, it is seen that there are certain strong correlations between the word "journey" and other words in specific books. For instance, in "Aleph," the word "journey" is strongly correlated with the words "said," "time," "know," "people," and "god." This suggests that the theme of journey in these novels is particularly focused on self-discovery and spiritual growth.

**Table 5. Collocates and Correlations for the themes on selected Paulo Coelho's novels**

Theme – Travel				
Paulo Coelho's Novel	travel* (Raw Frequency)	travel* (Relative Frequency)	Collocates	Correlations
Aleph	58	885	world(12) time(11) like(9) road(7) said(6)	throne(9) astral(10) vow(9) carpenter(10) divided(10)
Brida	20	344		
By the River Piedra I Sat Down and Wept	16	383		
The Alchemist	48	1152		
The Fifth Mountain	15	282		
The Pilgrimage	23	384		
Theme – Journey				
Paulo Coelho's Novel	journey* (Raw Frequency)	journey* (Relative Frequency)	Collocates	Correlations
Aleph	61	931	make(7) time(6) end(5) like(5) continue(4)	said(1511) time(863) know(773) people(676) god(477)
Brida	10	172		
By the River Piedra I Sat Down and Wept	3	72		
The Alchemist	14	336		
The Fifth Mountain	9	169		
The Pilgrimage	20	290		
Theme – Spiritual				
Paulo Coelho's Novel	spiritual* (Raw Frequency)	spiritual* (Relative Frequency)	Collocates	Correlations
Aleph	22	336	life(7) world(6) search(5) love(3)	said(1511) time(863) boy(675) people(676)
Brida	4	69		
By the River Piedra I Sat Down and Wept	15	359		

The Alchemist	7	168	physical(3)	elijah(420)
The Fifth Mountain	1	19		
The Pilgrimage	10	145		
Theme – Physical				
Paulo Coelho's Novel	physical* (Raw Frequency)	physical* (Relative Frequency)	Collocates	Correlations
Aleph	22	336	pain(4) spiritual(3) contact(2) just(2) world(2)	said(1511) time(863) boy(675) know(773) man(607)
Brida	4	69		
By the River Piedra I Sat Down and Wept	15	359		
The Alchemist	7	168		
The Fifth Mountain	1	19		
The Pilgrimage	10	145		

The theme of spirituality appears to be present in all of his selected novels, with the word "spiritual" or its variations appearing multiple times in each novel. The raw frequency of the word "spiritual" or its variations ranges from 1 to 22, while the relative frequency ranges from 19 to 336, indicating that spirituality is an important theme in Coelho's writing. The word "spiritual" is frequently accompanied by words like "life," "world," "search," "love," and "physical." These collocates provide insights into the context in which the word "spiritual" is used and suggest that the theme of spirituality is often intertwined with themes of personal growth, exploration, and self-discovery in Coelho's writing. The word "spiritual" is strongly correlated with the words "said," "time," "boy," "people," and "Elijah." This suggests that the theme of spirituality in this novels is particularly focused on the search for meaning and spiritual awakening.

The occurrences of the word "physical" or its variations are in selected Paulo Coelho's novels. The highest raw frequency being 22 in Aleph and the lowest being 1 in The Fifth Mountain is observed. The relative frequency of "physical" or its variations ranges from 19 to 336. The collocates of "physical" include words such as "pain," "spiritual," "contact," "just," and "world," providing insight into the context in which the word is used. The correlations of "physical" show a strong association with words such as "said," "time," "boy," "know," and "man," suggesting that physicality is often intertwined with themes of personal growth, self-discovery, and identity in Coelho's writing.

**10. Conclusion**

This research work gives a clear overview of the study's focus, which is to identify and analyze the research themes in selected novels of Paulo Coelho using text analysis. The study highlights the significance of identified themes in Coelho's work and how text analysis can be used to explore their representations in greater depth. The study also emphasizes the importance of using text analysis as a guide for further research and analysis, enabling researchers to delve deeper into the nuances of these themes and their representations in Coelho's work. The study also explores advantages of computational text analysis over manual analysis in researching literary texts, particularly for analyzing large volumes of texts quickly and efficiently and providing objective and unbiased analysis. It is especially relevant in the case of Coelho's novels, which are often multilayered and complex, making manual analysis challenging and subjective. Computational text analysis can provide a more systematic and comprehensive approach to analyzing Coelho's novels, enhancing traditional manual analysis methods. This study's findings can be useful to researchers of literature who are interested in studying the themes of travel, journey, spiritual journey, and physical journey to achieve one's destiny in Coelho's novels. The study provides a starting point for more detailed analysis and offers a framework for exploring the complex and multifaceted ways in which these themes are represented in Coelho's work. The study also highlights the importance of using text analysis as a guide for further research and analysis, enabling researchers to delve deeper into the nuances of

these themes and their representations in Coelho's work.

### References

- [1] Aneri Arya (2020): An Overview of Textual Analysis as a Research Method for Cultural Studies, *International Journal for Innovative Research in Multidisciplinary Field*, Volume - 6, Issue - 3, pp. 173 - 177.
- [2] Anmol D (2021): Journey Motif in Paulo Coelho's *The Alchemist: A Study of the Role and Relevance of Archetypal Figures*, *Contemporary Literary Review India*, Volume - 8, Issue - 2, pp. 1-9.
- [3] Coelho, Paulo (2012): *Aleph*, Translated by Margaret Jull Costa, Vintage International, 2012.
- [4] Coelho, Paulo (2014): *Brida*, Translated by Margaret Jull Costa, HarperCollins Publishers, 2014.
- [5] Coelho, Paulo (2014): *By the River Piedra I Sat Down and Wept*, Translated by Alan R. Clarke, HarperCollins Publishers, 2014.
- [6] Coelho, Paulo (2014): *The Alchemist*, Translated by Alan R. Clarke, HarperCollins Publishers, 2014.
- [7] Coelho, Paulo (2014): *The Fifth Mountain*, Translated by Clifford E. Landers, HarperCollins Publishers, 2014.
- [8] Coelho, Paulo (2014): *The Pilgrimage*, Translated by Alan R. Clarke, HarperCollins Publishers, 2014.
- [9] Ella Alhudithi (2021): Review of Voyant Tools: See through your text, *Language Learning & Technology*, Volume - 25, Issue - 3, pp. 43–50.
- [10] Hendrigan H (2019): Mixing Digital Humanities and Applied Science Librarianship: Using Voyant Tools to Reveal Word Patterns in Faculty Research, *Issues in Science and Technology Librarianship*, Volume - 91.
- [11] Sarfraz M (2022): Stylistic Analysis of Coelho's novel *The Alchemist*, *International Journal of Linguistics, Literature and Translation*, Volume - 5, Issue - 3, pp. 58–66.
- [12] Sinclair S, and Rockwell G (2015): *Voyant Tools Documentation*, University of Alberta, Canada.
- [13] Sinclair Stéfan and Geoffrey Rockwell (2023): Trends, *Voyant Tools, Web*, <[https://voyant-tools.org/?query=travel\\*&query=journey\\*&query=spiritual\\*&query=physical\\*&mode=&chartType=line&corpus=e09710074420ba36ad4374f9201d3343&view=Trends](https://voyant-tools.org/?query=travel*&query=journey*&query=spiritual*&query=physical*&mode=&chartType=line&corpus=e09710074420ba36ad4374f9201d3343&view=Trends)>.
- [14] Sinclair Stéfan and Geoffrey Rockwell (2023): Summary, *Voyant Tools, Web*, <<https://voyant-tools.org/?corpus=e09710074420ba36ad4374f9201d3343&view=Summary>>.
- [15] Sinclair Stéfan and Geoffrey Rockwell (2023): Terms, *Voyant Tools, Web*, <[https://voyant-tools.org/?query=physical\\*&query=spiritual\\*&query=journey\\*&query=travel\\*&corpus=e09710074420ba36ad4374f9201d3343&view=CorpusTerms](https://voyant-tools.org/?query=physical*&query=spiritual*&query=journey*&query=travel*&corpus=e09710074420ba36ad4374f9201d3343&view=CorpusTerms)>.
- [16] Sunil Kumar (2019): *Spiritual Journey in Paulo Coelho's Novel the Alchemist*, *Journal of Advances and Scholarly Researches in Allied Education*, Volume - 16, Issue - 2, pp. 224 - 229.
- [17] Syed Abthaheer K and Dr. M Arunachalam (2021): *Journey To The Quest: The Path Of Self Actualization In Paulo Coelho's "The Alchemist"*, *Journal of Language and Linguistic Studies*, Volume - 17, Issue - 3, pp. 2189-2194.
- [18] Ullah Z, Tahreem S, and Amjad S (2023): Stylometric Appraisal of Coelho's Selected Translated Novels: A Digital Humanities Study, *Pakistan Languages and Humanities Review*, Volume - 7, Issue - 1, pp. 117–132.
- [19] Vinodini S and Dr. Avinash M (2023): Exploring Military Strategies and Attitudes for Personal Growth and Self-Discovery in Paulo Coelho's Novels, *Res Militaris*, Volume - 13, Issue - 2, pp. 4912 - 4923.
- [20] Vinodini S and Dr. Avinash M (2023): Enhancing Literary Appreciation and Critical Thinking: Exploring the Potential of Digital Initiatives in Teaching Paulo Coelho's Novels, *European Chemical Bulletin*, E-ISSN. 2063-5346, Vol. 12, Special Issue. 4, pp. 12570 - 12585.
- [21] Welsh Megan E (2014): Review of Voyant Tools, *Collaborative Librarianship*, Volume - 6, Issue - 2, Article - 8.