Narrative and Text Visualization: A Technique to Enhance Teaching Learning Process in Higher Education

S. Anupama Kumar, M. N. Vijayalakshmi, Shaila H. Koppad and Andhe Dharani

Abstract Data visualization is the state-of-the-art technology that enables people from various domains to exhibit their work in a professional way. Visualization helps in understanding things in the past, present and predict the near future. The various tools and techniques of data visualization can be applied to different domains including health care, sales and marketing, forecasting, research, education, etc. This chapter intends the application of two different visualization techniques: narrative and text in the domain of education. Narrative and text visualization are generally considered as two different elements wherein this chapter intends to integrate both the techniques together and make it as a single component to help the education community. Narrative visualization is implemented using a storyboard that converts the text into a narrative format to help the student to understand the course in a simple way. Dashboards are also a narrative visualization technique which helps the student to have a glance of the course material. The content of the text is processed, frequent words are generated and visualized in the form of word cloud and other visualization methods to enable the student to understand and remember the course in a better way as a part of text analytics and visualization. Various techniques and tools including Canva storyboards, R, Tabulae and Power BI are employed in this work to implement narrative and text visualization to enable the tutor to teach better and make the teaching learning process easier.

Keywords Storyboard · Text extraction · Frequent words · Analysis · Word cloud · Stacked bar

1 Introduction

Data visualization is a state-of-the-art technology which enhances people to understand data by providing visual representation. Visualization is the graphical presentation of information, with the goal of providing the viewer with a qualitative understanding of the information contents [1]. It gives the data a completely new

S. Anupama Kumar (✉) · M. N. Vijayalakshmi · S. H. Koppad · A. Dharani
Department of MCA, R V College of Engineering, Bengaluru, India

S. M. Anouncia et al. (eds.), *Data Visualization*,
https://doi.org/10.1007/978-981-15-2282-6_1

meaning and reveals the hidden trends and information present in the data. The various domains ranging from education to research, advertising and marketing, all business set-ups, factories, banking sector and health care make use of data visualization techniques to bring out the new knowledge from it.

Narrative and text visualization are two important visualization techniques which can be applied over various domains to understand data better. Data comes in different formats like text, images and numerals. Narrative visualization techniques convert the data into a storyboard and help us to understand the data better.

Narrative visualization can be implemented by building a storyboard or a dashboard which explains the entire scenario in a better way to the user. Stories are a part of human communication structures, and storytelling enables to pass information and wisdom to the human community in a lively manner. Stories help us make sense of our past and reason about the future. Johnson [2] and Maclntyre [3] argue that story narrative also goes beyond communication; it is also a process of extracting meaning from events that is central to human experience and conduct.

Text visualization techniques are very important in the digital world since the documents are digitized in the current era. Therefore, it becomes essential to adopt a technique which can read the electronic document and help people to easily understand the content of the document without much effort. The visualization technique helps to summarize a document through two main aspects: (1) contents such as words and figures and (2) features such as average sentence length and number of verbs. The contents can further be visualized in the form of word blocks or cloud for explicitly.

The advanced visualization techniques can further create mind maps, visualization trees, blocks and other forms to make the user understand the theme easily. Further topic-based analysis can also be made using specific algorithms, and they can be understood easily.

This chapter intends to integrate narrative and text visualization techniques in the education domain. Narrative visualization will help the tutor to use visualization as a teaching aid to the student and help the student to understand the concepts in a better way. Text visualization will help to create a gist of the course material and help the student to use it as a material and learn the course in an enhanced way.

The following section gives an insight into the coupling of narrative and text visualization techniques.

2 Model for Integrating Narrative and Text Visualization

Narrative visualizations are data visualizations with embedded “stories” presenting particular perspectives using various embedding mechanisms [4]. Kosara and Mackinlay [5] and Lee et al. [6] exhibits storytelling in visualization as a primary way of communication apart from exploratory analysis.

Text visualization can be considered as an information visualization technique of visualizing raw text data or the application of text mining algorithms to visualize textual data [7]. Text visualization can be implemented to visualize the document similarity, revealing the content of the text and depicting the content for easy understanding and usage.

This chapter intends to build a model to couple both narrative and text visualization to bring out various forms of text in a better way and help the tutors to enhance their teaching techniques.

The authors in [8] have coupled the text and narrative visualization manually to interpret the textual data. The authors in [9, 10] analysed text to narrative visualizations using annotations and other elements. The authors in [11] coupled the techniques to implement a technique that can use text documents as well as web links and depict the links in a better way. Different case studies were investigated and proved to ensure the coupling of text and narrative visualization.

In this chapter, the course material (e-book) of a particular course in a higher education system is taken as a data source. Figure 1 shows the coupling of the techniques.

The techniques are implemented as follows:

(i) The course material is available in pdf formats with six chapters. All the chapters are split into separate text documents so that it becomes easy to analyse and visualized in a sophisticated way. The e-book is converted into text files using functions for further process.

(ii) The second step is carried over in two methods.

   a. Storyboards are created using Canva studio. The textual materials in the chapter are converted into visual formats to understand the content in a better manner. Different types of visualizations are used to understand the data in a better way.

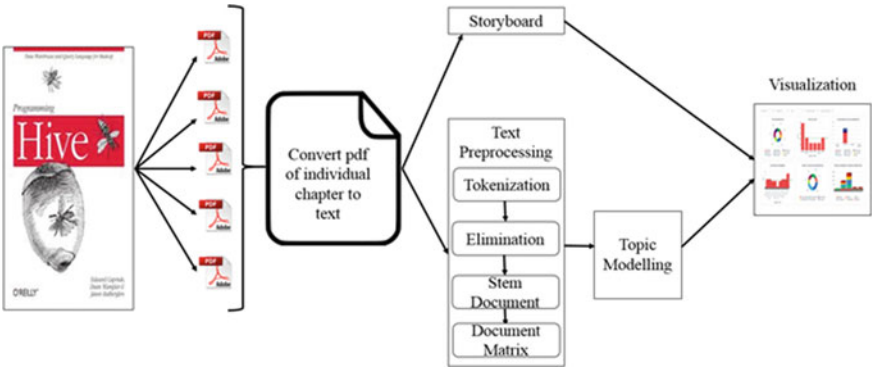

Fig. 1 Coupling of narrative and text visualization

b. The text file is preprocessed to extract relevant information. Then, topic modelling algorithms are implemented to generate frequent topics from each chapter. The above functions are carried out using R.

(iii) The final step is to apply visualization of the applied textual data using appropriate techniques. Narrative visualization is depicted using hierarchical charts, Bubbled charts and dashboards. In case of text visualization, the frequent topics are visualized using different charts/graph using Power BI which is one of the visualization tools.

The following section explains the implementation of narrative and text visualization over the higher educational data.

3 Application of Narrative and Text Visualization

Visualization is one of the most powerful tools that can be implemented over a variety of domains. Education is one of the important domains where visualization helps the tutor to teach better and the learner to understand better. The techniques can be used to enhance the teaching learning process in a better way. The authors in [12] have implemented visualization techniques to teach mathematics for engineering graduates. They have used 3D visualization techniques using JavaScript and WebGL to make the course understandable to learners. The application of visualization in understanding the geo spatial data with multiple dimensions has been well explained in [13]. Joris Klerkx, Katrien Verbert and Erik Duval have discussed the different models, methods and techniques of visualization that can improve and enhance their teaching [14]. They have also discussed various types of learning models and methods over which the visualization techniques can be applied. These methods have given insight to apply narrative and text visualization techniques over the course material.

3.1 Implementation Narrative Visualization

Narrative visualization is one of the effective techniques that can be used for textual representation. But there are few challenges and issues that are to be handled during the process of working. The authors in [15, 16] have listed out various challenges that can arise while implementing narrative visualization over textual data as stories written by authors. The following challenges were met while building a model for narrative visualization

(i) The foremost challenge is inferring the environment in which the events take place. The tutor should be comfortable with the course so that tutor can build the course material in a very sophisticated way.

(ii) Grammatical disfluencies, such as real-word spelling errors, missing punctuation, improper or omitted prepositions, incorrect verb tense and missing auxiliary verbs may cause the syntactic and semantic parsers.
(iii) The narrative text has a very complex theoretical structure which may be complicated for the learner to understand.
(iv) The sequence in which the narration (events) should happen should be depicted properly; otherwise, it may cause disproportionate storyboards to the learner.

Figure 2 shows the raw text material which is taken from the HIVE book of a higher education course.

The objective of this work is to convert this into a meaningful visualization format which will enable the learner to understand the concept easier. The process

CHAPTER 3

Data Types and File Formats

Hive supports many of the *primitive* data types you find in relational databases, as well as three *collection* data types that are rarely found in relational databases, for reasons we'll discuss shortly.

A related concern is how these types are represented in text files, as well as alternatives to text storage that address various performance and other concerns. A unique feature of Hive, compared to most databases, is that it provides great flexibility in how data is encoded in files. Most databases take total control of the data, both how it is persisted to disk and its life cycle. By letting you control all these aspects, Hive makes it easier to manage and process data with a variety of tools.

Primitive Data Types

Hive supports several sizes of integer and floating-point types, a Boolean type, and character strings of arbitrary length. Hive v0.8.0 added types for timestamps and binary fields.

Table 3-1 lists the *primitive* types supported by Hive.

Table 3-1. Primitive data types

Type	Size	Literal syntax examples
TINYINT	1 byte signed integer.	20
SMALLINT	2 byte signed integer.	20
INT	4 byte signed integer.	20
BIGINT	8 byte signed integer.	20
BOOLEAN	Boolean true or false.	TRUE
FLOAT	Single precision floating point.	3.14159
DOUBLE	Double precision floating point.	3.14159

Fig. 2 Raw text material

of creating narrative visualization for course material involves a lot of challenges. The creator of the visualization should be aware of the content of the book so that the visualization can be created accurately. The tutor should consider the following points while creating a storyboard:

(i) The content which the tutor wants to convey should be converted into accurate visualization format which otherwise will convey wrong meaning.
(ii) The important points of the content should be highlighted using appropriate colour, size, boldness and connecting elements like arrows and shaded trails should be properly maintained which otherwise will create trouble to the user.
(iii) The images should be aligned from the largest image to the smallest image with the appropriate caption given to them. The images should also have proper size and should be positioned in appropriate place so that content can be conveyed in proper manner.
(iv) The progress bar and other elements should be properly placed so that it enables the user to understand the importance of the content.
(v) The syntactic and sematic regulations of the content should be properly maintained.
(vi) The space dimensions in the storyboard should be maintained properly to enable the student to visualize the content.

Figure 3 depicts the hierarchical structure created for the given text.

Canva storyboards can be created using various templates provided in the software. The system is user friendly and allows the user to implement the narrative analytics in different ways. The challenge in using these softwares lies about the knowledge of the user who creates the storyboards. The user should also be aware of the scenes that should be created in a sequence. Figure 4 depicts another way of visualizing the same chapter to enable the learner to understand the concept.

The chart depicted in Figs. 3 and 4 will be integrated in the dashboard along with the text visualization to make the learner understand the concept in a better way. The following section explains the implementation of the text processing and topic modelling.

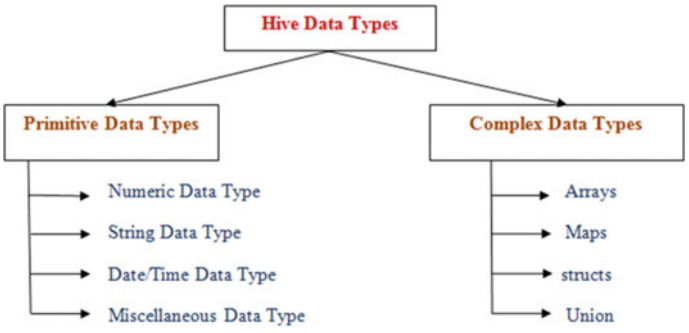

Fig. 3 Hierarchical structure of the raw text

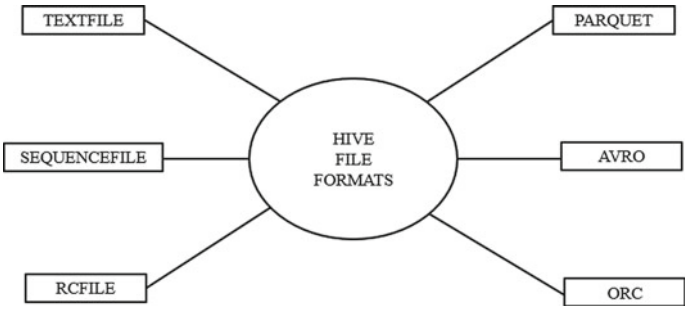

Fig. 4 Narrative text visualization of the chapter

3.2 Implementation of Text Visualization

Teaching learning process can be made effective if the learner is able to understand the concept by reading or visualizing the content of the course material in a simple way. Therefore, it becomes the duty of the tutor to make the material interactive and easy for the learner. If the course material is complicated to read and understand, the objective of the course will become diluted. Text processing and topic modelling play a major role in simplifying the course material and provide a gist of the contents to the learner in a better way. In [17], the author has discussed the application of NLP with scientific computer programs to enhance the process of education. Litman [18] has expressed various research areas in education where natural language processing can play a major role and bring in changes in teaching and learning the courses. Natural language processing involves processing the text data (structured or unstructured) and applies text analysis algorithms to get accurate output. The implementation of the text processing and topic modelling is carried using R. The output of topic modelling is depicted in the form of word cloud and other visual aids using R. Figure 5 is the overview of the text visualization process carried over using text processing, topic modelling and visualization.

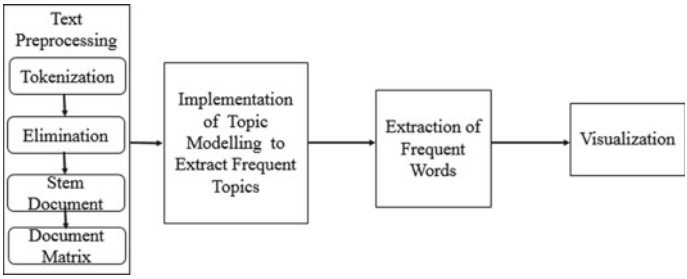

Fig. 5 Overview of text visualization

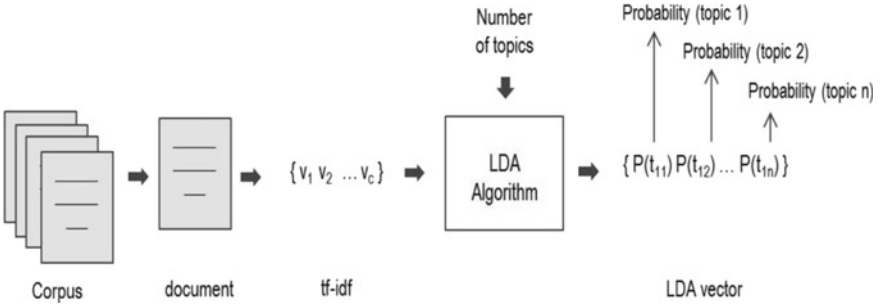

Fig. 6 Implementation of LDA algorithm

Text processing is carried over the course material to extract the raw text from the e-book. This is called as tokenization. Once the words are extracted, it is necessary to remove the stop words, stem the document and filter the document for effective extraction. The numerals and punctuations available in the text are also removed so as to get the plain text. The images and emoticons are not considered in the text processing. A term document matrix is created for each chapter, and the results are recorded.

Topic modelling is one the interesting text mining technique which enables the user to extract topics from a huge corpus. Paul and Girju [19] applied LDA algorithm in interdisciplinary research areas and proved that it can be applied to extract topics from education field also. The working of the LDA algorithm is depicted in Fig. 6.

The course material is converted into the text document and preprocessed in the previous phase. Then, a term document matrix is generated. The term document matrix is fed as an input to the LDA algorithm, and a set of topics are generated depending on the user's demand. The most frequent words from each topic are then extracted from all the chapters. Figure 7 is the resultant sample of frequent words generated using LDA algorithm.

Figure 8 depicts the list of frequency words that are extracted using the algorithm in all the chapters and visualization of common words from the chapter in the form of a cluster. Bar chart is created using Power BI. The most frequent word in the course material is 'hive', the second highest is 'data', and the third is 'Hadoop' and it goes on. The top most words are depicted in green colour in the stacked graph.

The frequent words are imported to Power BI for further visualization. Figure 9 depicts the total number of frequent words in each chapter, and the Bowtie Chart helps to visualize the frequent terms present in the chapter.

The hierarchical structure in Fig. 10 is a sample which shows the visualization of the total frequent words present in the course material and how it is broken down under each chapter.

From Fig. 10, it depicts the occurrence of the words in each chapter. The bubble in the end of first structure states that the process is incomplete, and the words are still to be generated.

hive	69	1	Chapter 1 : Introduction
data	56	1	Chapter 1 : Introduction
hadoop	48	1	Chapter 1 : Introduction
use	41	1	Chapter 1 : Introduction
word	39	1	Chapter 1 : Introduction
mapreduc	33	1	Chapter 1 : Introduction
java	26	1	Chapter 1 : Introduction
count	26	1	Chapter 1 : Introduction
can	25	1	Chapter 1 : Introduction
queri	24	1	Chapter 1 : Introduction
provid	21	1	Chapter 1 : Introduction
languag	21	1	Chapter 1 : Introduction
sql	20	1	Chapter 1 : Introduction
job	18	1	Chapter 1 : Introduction
one	18	1	Chapter 1 : Introduction
key	18	1	Chapter 1 : Introduction

Fig. 7 Sample processed output of LDA algorithm

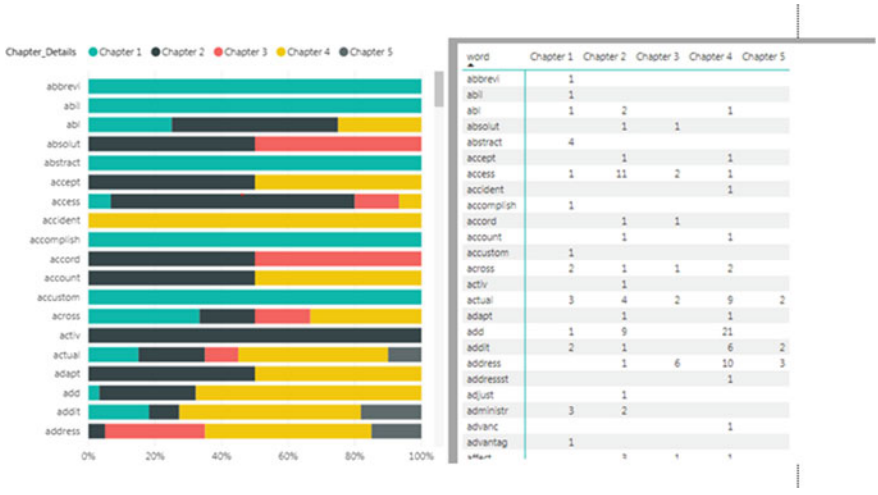

Fig. 8 Frequent words using LDA and its visualization

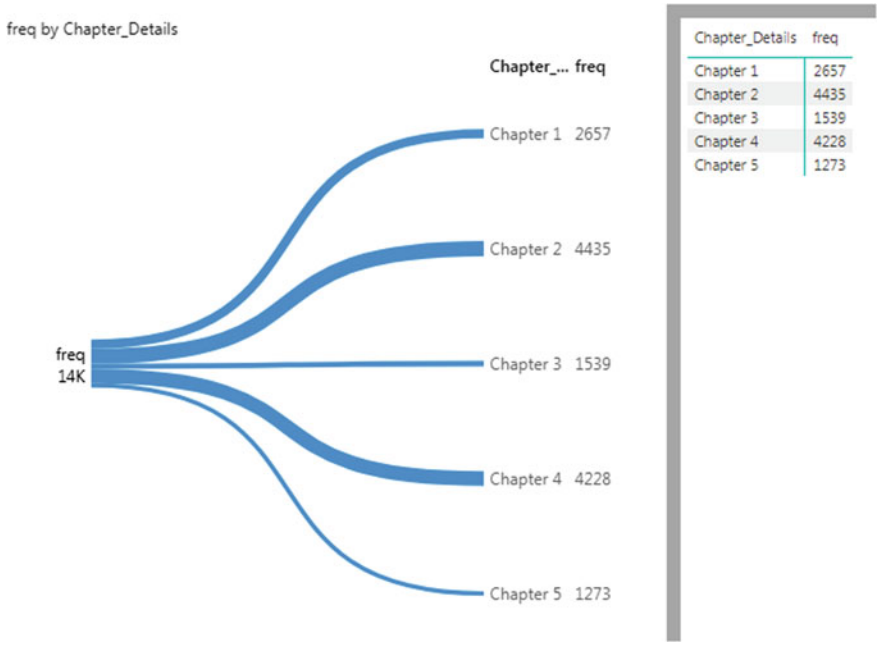

Fig. 9 Bowtie chart representing total frequent words in each chapter

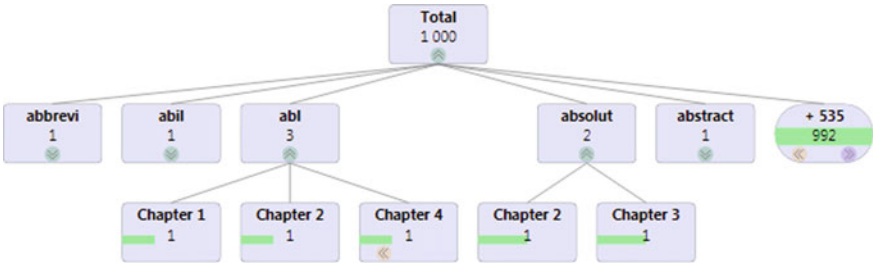

Fig. 10 Visualization of hierarchical structure of frequent words

Figure 11 is a pie chart that explains the total number of frequent words present in each chapter. The total frequent words present in chapter one is 869, chapter two 1003, chapter three 495, chapter four 881 and chapter five 402.

3.3 Integration of Narrative and Text Visualization

The final step is coupling all the visualizations together using a dashboard which will enable the learner to understand the concept of the course. The tutor should take

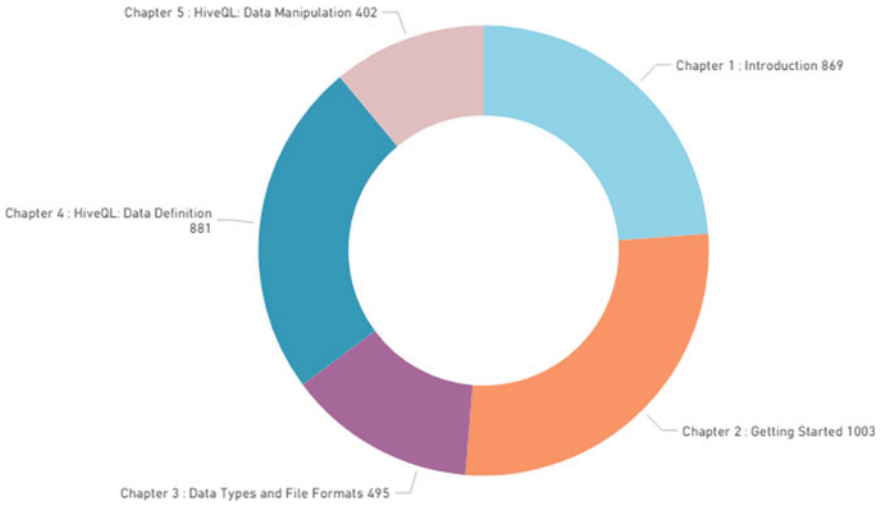

Fig. 11 Donut chart depicting the frequent words in each chapter

care that all the concepts available in the course material are provided to the student in a simple way. Figure 12 is the dashboard created through the storyboard of Canva. The tutor has taken care to explain the chapters by giving a title to the chapter which will enable the learner to have a glance on what is present in the chapter.

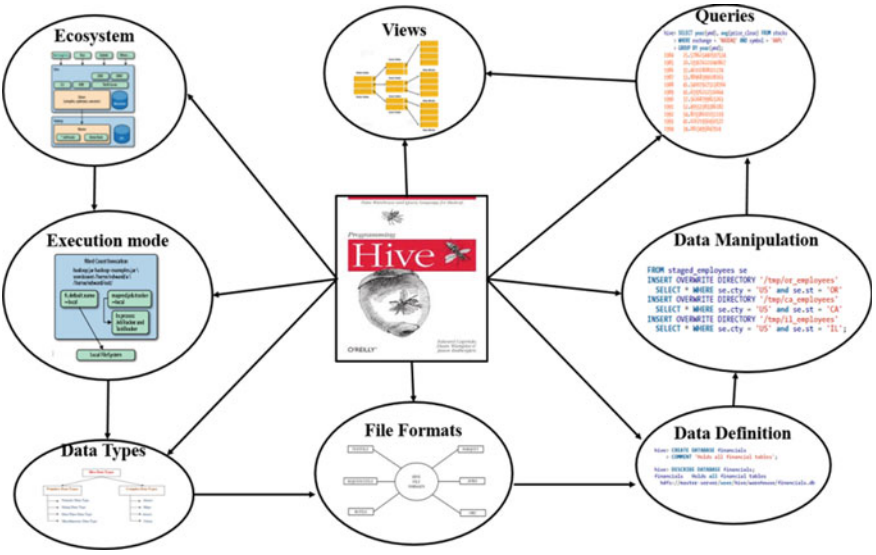

Fig. 12 Dashboard using Canva for the storyboard—course material

Fig. 13 Integrated dashboard of narrative and text visualization

Figure 13 depicts the integration of the text visualization along with the storyboard to enable the student to understand the key words present in each chapter so that student can understand the concepts available in each chapter. Therefore, it becomes the responsibility of the tutor to depict the appropriate keyword in the storyboard to explain the concept to the student. The word ecosystem gives an insight into the student that chapter one will be helping them to understand the Hive Ecosystem. Figure 13 depicts the integration of the dashboard integrating the narrative and text visualization of the course material.

In Fig. 13, the dashboard of narrative and text visualization has been coupled together to make the learner understand the concept easier. The wordcloud gives a list of frequent keywords present in each chapter. The other visualization charts have been already discussed earlier. The integrated visualization helps the tutor to give an overview of the course in a better way that enables the learner to understand the concepts better.

4 Conclusion

Data visualization is one of the imperative techniques in the current era which can be applied on various domains. This chapter has provided an insight into the application of data visualization using narrative and text in the field of education. Narrative visualization has been implemented to extract the text and depict in the form of charts and dashboard to enable the students to understand the content of the chapters in a sophisticated manner. The implementation of narrative visualization has been done using Canva storyboard and coupled with Power BI. Text visualization has been applied in two phases: processing the text to tokenize, removing the punctuation, stop words, numerals, etc. LDA algorithm was applied to generate topics based on each chapter, and frequent words were extracted from each chapter. Text processing and analysis are implemented using R tool. The visualization of the same is implemented using Power BI. Narrative and text visualization are coupled together using dashboards

which has given a clear visual module of the course material to the learner. The learner is able to understand the content and flow of the course, important key words in each chapter and understand the course in a better way. In future, visualization techniques can be further implemented to other domains of education also.

References

1. https://web.cs.wpi.edu/~matt/courses/cs563/talks/datavis.html.
2. Johnson, M. (1993). *The moral imagination*. Chicago: University of Chicago Press.
3. Maclntyre, A. (1981). *After virtue*. Notre Dame, IN: University of Notre Dame Press.
4. Segel, Edward, & Heer, Jeffrey. (2010). Narrative visualization: Telling stories with data. *IEEE Transactions on Visualization and Computer Graphics, 16*(6), 1139–1148.
5. Kosara, R., & Mackinlay, J. (2013). Storytelling: The next step for visualization. *Computer, 46*(5), 44–50.
6. Lee, B., Riche, N. H., Isenberg, P., & Carpendale, S. (2015). More than telling a story: Transforming data into visually shared stories. *IEEE Computer Graphics and Applications, 35*(5), 84–90.
7. Kucher, K., & Kerren, A. (2015). Text visualization techniques: Taxonomy, visual survey, and community insights. In *2015 IEEE Pacific Visualization Symposium (PacificVis)*. ISSN: 978-1-4673-6879-7.
8. Kwon, B. C., Stoffel, F., Jäckle, D., Lee, B., & Keim, D. (2014). VisJockey: Enriching data stories through orchestrated interactive visualization. In *Poster Compendium of the Computation + Journalism Symposium* (Vol. 3).
9. Hullman, J., Diakopoulos, N., & Adar, E. (2013). Contextifier: Automatic generation of annotated stock visualizations. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (pp. 2707–2716). ACM.
10. Bryan, C., Ma, K.-L., & Woodring, J. (2017). Temporal summary images: An approach to narrative visualization via interactive annotation generation and placement. *IEEE Transactions on Visualization and Computer Graphics, 23*(1), 511–520.
11. Metoyer, R., Zhi, Q., Janczuk, B., & Scheirer, W. (2018). Coupling story to visualization: Using textual analysis as a bridge between data and interpretation. IUI 2018, March 7–11, 2018, Tokyo, Japan. Copyright © 2018 ACM. ISBN: 978-1-4503-4945-1/18/03.
12. Pócsová, J., Mojžišová, A., & Takáč, M. (2016). Application of the visualization techniques in engineering education. In *17th International Carpathian Control Conference (ICCC)*, IEEE, INSPEC Accession Number: 16107928.
13. Teaching with Visualizations, Created by Bob MacKay, Clark College, https://serc.carleton.edu/sp/library/visualizations.
14. Klerkx, J., Verbert, K., & Duval, E., Enhancing learning with visualization techniques. Katholieke Universiteit Leuven, Belgium. https://core.ac.uk/download/pdf/34578672.pdf.
15. Baikadi, A., Goth, J., Mitchell, C. M., Ha, E. Y., Mott, B. W., & Lester, J. C., Towards a Computational Model of Narrative Visualization, AAAI Technical Report WS-11-18.
16. Segel, E., & Heer, J., Narrative visualization: Telling stories with data.
17. Alhawiti, K. M. (2014). Natural language processing and its use in education. *International Journal of Advanced Computer Science and Applications, 5*(12).
18. Litman, D. (2016). Natural language processing for enhancing teaching and learning. In *Proceedings of the Thirtieth AAAI Conference on Artificial Intelligence (AAAI-16)*.
19. Paul, M., & Girju, R. (2009) Cross-cultural analysis of blogs and forumswith mixed-collection topic models. In *Proceedings of the 2009 Conference on Empirical Methods in Natural Language Processing* (pp. 1408–1417), Singapore, 6–7 August 2009.