1. Submit to JESTEC (5 Nov 2020)
2. First Revision (1 Des 2020)
3. Submit Revised Paper (8 Des 2020)
4. Accepted Submission (15 Jan 2021)
5. Paper Publish (3 Feb 2021)

1. Submit to JESTEC (5 Nov 2020)

Dian Dharmayanti

<dian.dharmayanti@email.unikom.ac.id>

[JESTEC] Manuscript ID: JESTEC_16_1_2021_D Dharmayanti–Submission Received

1 message

Editorial Office <jjestec@gmail.com>   5 November 2020 at 10:45
To: Dian Dharmayanti <dian.dharmayanti@email.unikom.ac.id>

Dear Author,

Thank you very much for uploading the following manuscript to the submission system. One of our editors will be in touch with you soon.

Journal Name	Journal of Engineering, Science and Technology
Manuscript ID	JESTEC_16_1_2021_D Dharmayanti
Authors	Dian Dharmayanti, Adam Mukharil Bachtiar, Firman Nizammudin Fakhrul
Title	Data Visualization for Content Marketing Domain in Social Media

If you have any questions, please do not hesitate to contact editorial office.

Kind regards,

Editorial Office
Journal of Engineering, Science and Technology
E-Mail: jjestec@gmail.com

*** This is an automatically generated email ***

DATA VISUALIZATION FOR CONTENT MARKETING DOMAIN
IN SOCIAL MEDIA

Dian Dharmayanti* , Adam Mukharil Bachtiar , Firman Nizammudin Fakhrul

Informatics Engineering, Faculty of Engineering and Computer Science, Universitas Komputer Indonesia, Jalan Dipatiukur Nomor 112-116, Bandung, Jawa Barat, Indonesia
Corresponding Author dian.dharmayanti@email.unikom.ac.id

Email : *adam@email.unikom.ac.id

Abstract. This research will focus on producing proper data visualization in the content marketing domain in social media. The research method consists of nine steps and using techniques such as empathy map, statistics, data mining, data visualization technique, acceptance testing, and usability testing to generate proper data visualization. A prototype will be developed as the proof of concept of this research. The usability testing result of the prototype says that the effectiveness level is 100% and the efficiency level is 100%. By using the proper data visualization, it can increase the efficiency and effectiveness level of the content marketing research on social media for marketers.
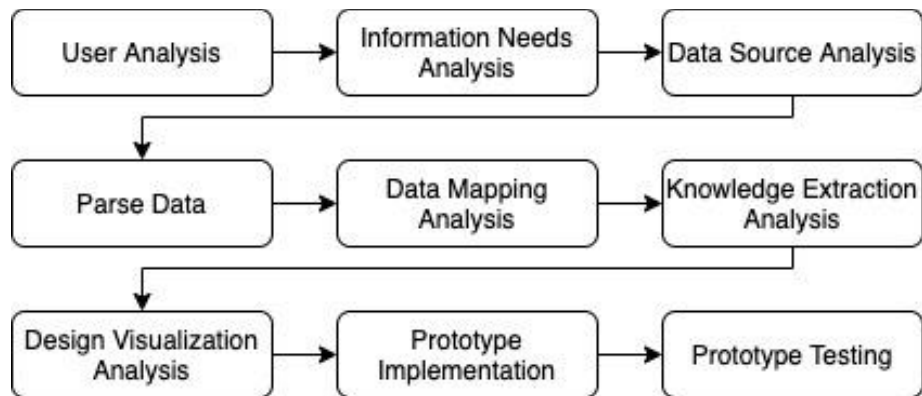
1. Introduction

Content marketing is now widely used by marketers to attract customers. Content marketing is a marketing process to create and distribute content that aims to give appeal to the intended target audience [1]. The crucial thing in content marketing is the content marketing strategy. But in fact, 62% of marketers do not have a solid strategy [2]. Several well- known brands use a data-driven approach to create a solid strategy [3]. One source of data that can be used to create a content marketing strategy is social media. Social media produce a very large number of data [4]. The huge data can be a tremendous opportunity for marketers to improve marketing content strategies, such as understanding the target audience and content discovery [5].

Based on this phenomenon, an interview was conducted with one of the marketers. The problem that arises is that marketers need a lot of time to read information from social media data needed for content marketing, such as target audience segmentation. The main factor of the difficulty that is due to the social media used by marketers provides an inappropriate form of data visualization so the data must be analyzed again to get the desired information. For example, audience research, marketers enter a keyword and analyze the characteristics of the target audience relevant to that keyword. In addition, errors in the process of obtaining such information cannot be predicted so that it can have an impact on misinformation.

An effective way to present detailed data into information that is easily accepted is by means of abstraction into visual information [6]. As explained earlier, marketers need a significant amount of time to read information from social media data relating to content marketing strategies due to inappropriate forms of data visualization or visual information. Therefore, further analysis is needed to find out what kind of data visualization is right in the content marketing domain on social media. The knowledge generated is used to create the right data visualization in the content marketing domain on social media for marketers.

2. Method

The research methodology was inspired by the book "Visualizing Data" which consists of 7 steps that acquire, parse, filter, mine, represent, refine, and interact [7]. The methodology adapted to the needs of this research. The research methodology can be seen in Figure 1.

Figure 1. Research Methodology

3. Results and Discussion
All stages in the research methodology are being explained in more detail in this section. The explanation consists of input, process, and output for each stage.
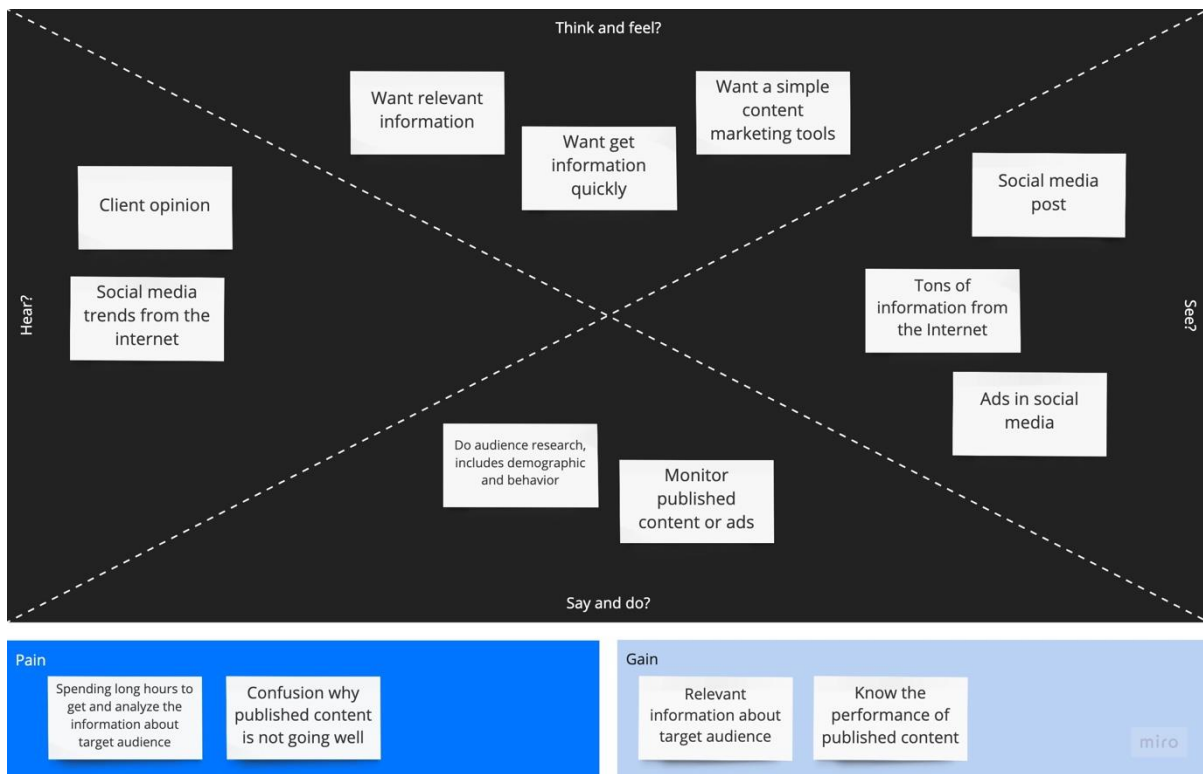
3.1. User Analysis
A good data visualization is well supported by identifying the user who will read the visualization itself [8]. To identify the user, the technique used is the interview. In this research, two marketers who are an actor in the field of content marketing are being interviewed. The interview is divided into three sections, those are opener question section, daily task section, and use of content marketing tools section. The example question of each section can be seen in Table 1.

Table 1. Interview Section

Section	Sample Question
Opener question	What is your role in the content marketing team?
Daily task question	What activities do you do every day associated with content marketing?
Use of content marketing tools section	What difficulty do you face when using content marketing tools?

The result of the interview is qualitative data. In this research, the empathy map is used as a tool to illustrate the interview result. The empathy map can be seen in

Figure **2**. The result shows that the user has two pain points, spending long hours to get and analyze the information about the target audience and confusion why the published content is not going well. Then, this empathy map is used as an archetype user to analysis, design, and testing the data visualization in this research.

Figure 2. Empathy Map

3.2. Information Needs Analysis

Information needs analysis stage conducted to determine what information is needed by marketers in the domain of content marketing in social media. Empathy map from the previous stage is used to determine the information needs. Furthermore, the pain and gain outlined the information needs. The sample of information needs that are formed can be seen in Table 2.

Table 2. Information Needs

Pain	Gain	Information Needs
Spending long hours to get and analyze the information of the target audience	**Get relevant information about the target audience quickly**	How to know the interest of a segment of the target audience?
Confusion why the published content is not going well	**Know the performance of published content**	How to know the sentiment score from the post?

The result of this stage is nine information needs based on breakdown pain and gain of the empathy maps. These information needs are used as a reference for visualizing the data on the domain of content marketing in social media.

3.3. Data Source Analysis

Data source analysis conducted to analyze the characteristic of the data source. The data source used is data from Twitter. The data were obtained through a RESTful API provided by Twitter. Twitter data can be accessed through a specific endpoint. The complete list of Twitter API endpoint can be

accessed at https://developer.twitter.com/en/docs/api-reference-index. Example endpoint of Twitter API can be seen in Table 3.

Table 3. Example of Twitter API

Method	Endpoint Path	Output	URL Documentation
GET	users/show	User Object	https://developer.twitter.com/en/docs/accounts-and-users/follow-search-get-users/api-reference/get-users-show

Twitter API can be accessed through endpoints based on its features. The data obtained shaped as JSON (JavaScript Object Notation). The data contains Twitter data objects, where Twitter data objects are divided into several types, namely tweet objects, user objects, entities object, and geo objects.
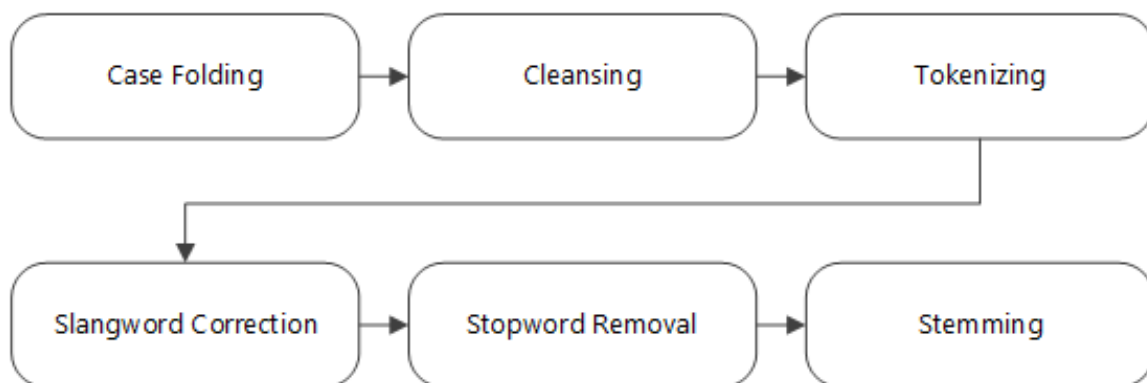
3.4. Parse Data

Parse data stage conducted to make the data easier to be processed. In this stage, text preprocessing and format conversion of data into a document-based database are done.

3.4.1. Text Preprocessing

Tweets data objects containing text that is not structured so it has to be preprocessed. The result of preprocessing will make the data mining stage easier [9]. Steps being taken can be seen in

Figure 3.

Figure 3. Text Preprocessing Steps

The sample result for a tweet that being preprocessed can be seen in Table 4.

Table 4. Sample Result of Text Preprocessing

Input	Output
Pagi yg menyenangkan. Cari bubur mantap utk sarapan. https://t.co	[pagi, senang, cari, bubur, mantap, sarapan]

3.4.2. Document and Data Structure Formation

Document-based database offers good performance and scalability [10]. To make it easier to process the data at the next stages, four types of documents formed. Documents that formed are user

documents, followers document, and mention document. Endpoint used and output objects of the document formed can be seen in Table 5.

Table 5. Document Formation

Method	Endpoint	Output

User Document	followers/list
	users/show
		{
		  "_id":String,
		  "user_profile":{
		    …
		    },
		  "user_tweets":[
		    ...
		    ]
		}

Followers Document	followers/list
	users/show
	statuses/user_timeline
		{
		  "_id":String,
		  "following":{
		    …
		    },
		  "follower_profile":{
		    ...
		    },
		  "follower_tweets":[
		    ...
		    ],
		}

Mention Document	users/show
	statuses/mention_timeline
		{
		  "_id":String,
		  "user_profile":{
		    …
		    },
		  "user_mentions":[
		    ...
		    ]
		}

This stage produces three documents. The documents are user documents, followers document, and mentions document. The documents will be used as a data source for the visualization.

3.5. Data Mapping Analysis
Data mapping stage conducted to choose and map data based on the information needs. Mapping data to the sample information needs can be seen in Table 6.

Table 6. Mapping Data to Information Needs

Gain	Information Needs	Document Used
Get relevant information on the target audience quickly	How to know the interest of a segment of the target audience?	User document, Followers document
Know the performance of	How to know the sentiment	User document,

published content	score from the post?	Mention document

3.6. Knowledge Extraction Analysis

The data that have been obtained and mapped are then performed statistical calculations or data mining methods to extract useful information. Data mining technique is used because it can be used to analyze the data and summarize it into useful information [11]. This stage will also determine whether the information needs of an exploratory analysis or explanatory analysis to understand the context of the needs of the information so that it can be concluded whether the information needs requires knowledge extraction method or not. The mapping of the knowledge extraction method is depended on the characteristic of the data too [8]. The mapping of the knowledge extraction method to sample information needs can be seen in Table 7.

Table 7. Mapping Knowledge Extraction Method

Gain	Information Needs	Exploration Type	Method
Get relevant information on the target audience quickly	How to know the interest of a segment of the target audience?	Exploratory	Classification, Statistics (distribution frequency and mode)
Know the performance of published content	How to know the sentiment score from the post?	Exploratory	Classification, Statistics (distribution frequency and mode)

3.7. Design Visualization Analysis

Design visualization analysis includes four steps that were inspired in the "storytelling with data" book. The steps are understanding the context, choosing the appropriate form of visualization, eliminating clutter, focusing attention, and acceptance testing [12]. In this stage, the information needs a sample that is used is "How to know the interest of a segment of the target audience?".

3.7.1. Understanding Context

Understanding the context of information needs is done by describing the context of who needs the information, what information will be delivered, and how the delivery of information. The example description of the context can be seen in Table 8.

Table 8. Understanding Context

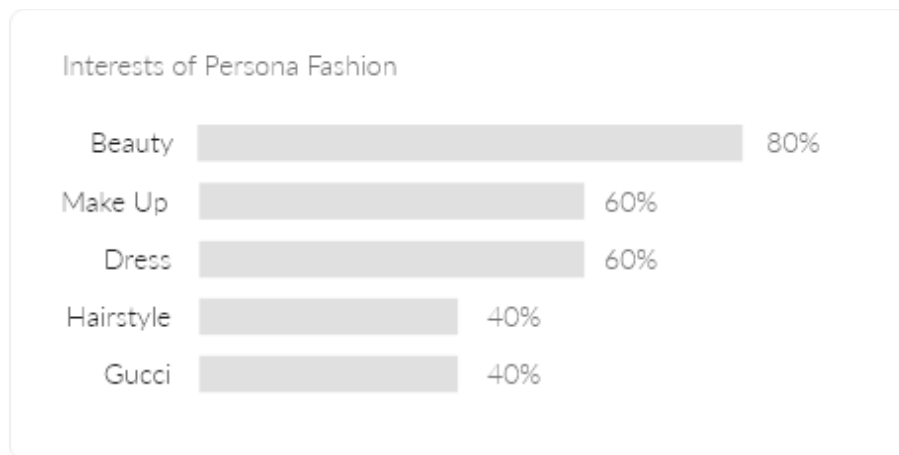| Who	What	How
A content marketer	Knowing the interest of a segment of the target audience	Displays a comparison of the list of interests of a target audience segment

3.7.2. Choosing Appropriate Type of Visualization

Choosing an appropriate type of visualization is done by analyzing how in understanding the context stage. The example of choosing the appropriate type of visualization can be seen in Table **9**.

Table 9. Choosing Appropriate Type of Visualization

How	Visualization Type	Why
Displays a comparison of the list of interests of a target audience	Horizontal bar chart	Compare several categories

segment

3.7.3. Eliminating Clutter

Data visualization is in accordance with principles that will not contain clutter. The sketch of the visualization should be made and can be seen in Figure 4. The sketch is analysed and should contain no clutter.

Figure 4. Sketch of Visualization

The example of clutter analysis can be seen in **Table 10**.

Table 10. Clutter Analysis

Principle	Explanation	Checklist
Alignment	Alignment used is left-aligned because reading information will be easier	✔
White space	White space is only used as a margin between components so they don't overlap	✔
Closure	There is no unnecessary border on data visualization	✔
Continuity	Remove the y-axis border because the bar always starts from the same point so there is no interference between the label and the bar	✔

3.7.4. Focusing Attention

Focusing attention made to focus the user on the information that is important to the user. The example of focusing attention analysis can be seen in Table 11.
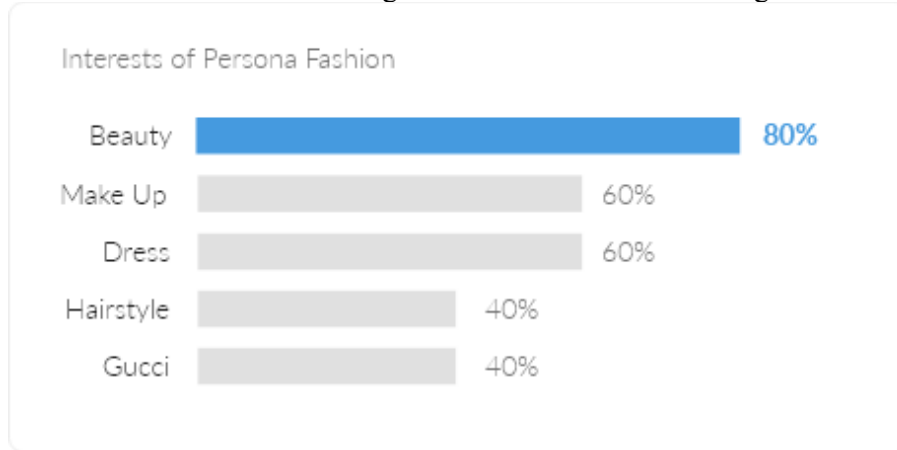
Table 11. Focusing Attention

How	Focused Information	Preattentive Attributes
Displays a comparison of the list of interests of a	Most dominant interest	Color

target audience segment

The result of visualization after focusing the attention can be seen in Figure 5.

Figure 5. Visualization that Has Been Focused

3.7.5. Acceptance Testing
Testing is done by acceptance testing. The result of acceptance testing can be seen in Table 12.

Table 12. Acceptance Testing

Actual Information	Obtained Information	Comment	Result
Comparison of the list of interests of a target audience segment	Comparison of the list of interests of a target audience segment	No comment	Accepted

A good data visualization can give the user the right information [13]. So, the acceptance criteria of this testing are the information obtained is correct and there is no additional comment. The conclusion from acceptance testing is that the form of visualization is appropriate.

3.8. Prototype Implementation
Data visualization is made in prototype form. Sequence information on the prototype is based on the results of card sorting and grouping by pain and gain of the empathy maps. The prototype can be seen in Figure 6 and Figure **7**.
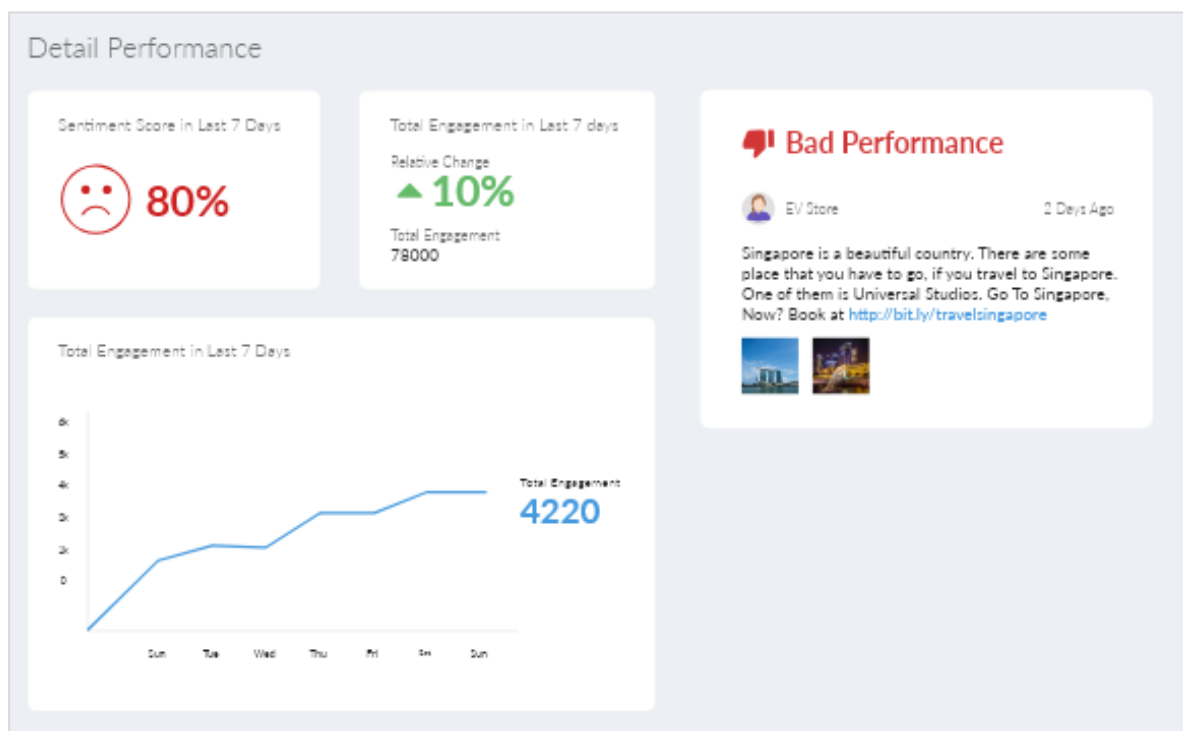
Figure 6. Prototype of First Group

Figure 7. Prototype of Second Group

3.9. Prototype Testing

Tests on a prototype using usability testing techniques. Scenario testing is divided based on the gain in the empathy map in the user analysis stage. There are 2 scenarios given to the user. The effectiveness level is 100% and the efficiency level is 100%.

4. Conclusion

Based on all processes in this research, the conclusion is that the data visualization in this research can increase the efficiency and effectiveness level of the content marketing research on social media for marketers. For the next research, so much data in social media that can be more explored and the data source can be anything besides Twitter.

References

[1] Chan A and Astari D, 2017 The Analysis of Content Marketing in Online Fashion Shops in Indonesia *Rev. Integr. Bus. Econ. Res.* **6**(2) p. 225–233.

[2] Beets, L. M., & Handley, A. 2018. B2C Content marketing-2018 Benchmarks, Budgets, and Trends—North America.

[3] Karczewski T, 2014, 6 Steps to a Data-Driven Content Marketing Strategy. [Online]. Available: http://contentmarketinginstitute.com/2014/06/data-driven-content-marketing-strategy/. [Accessed: 03-Mar-2018].

[4] Kharb, L. 2017. Exploration of Social Networks with Visualization Tools *Am. J. Eng. Res.* **6**(3), pp. 90–93.

[5] Litsa T, 2017, 5 Tips to Create a Data-Driven Content Marketing Strategy, *Search Engine Watch*. [Online]. Available: https://searchenginewatch.com/2017/09/19/5-tips-to-create-a-data-driven-content-marketing-strategy/. [Accessed: 03-Mar-2018].

[6] Syaripul, N. A., & Bachtiar, A. M. 2016. Visualisasi Data Interaktif Data Terbuka Pemerintah Provinsi DKI Jakarta: Topik Ekonomi dan Keuangan Daerah. *Jurnal Sistem Informasi*, **12**(2), pp. 82-89.

[7] Fry, B. 2008. Visualizing Data. 1005 Gravenstein Highway North, Sebastopol, CA 95472.

[8] Bachtiar A. M., & Dharmayanti D. 2018. Data Visualization at Trading Sector Using Government's Open Data *Int. J. Pure Appl. Math.* **118**, 20 p. 523–530.

[9] Rivki, M., & Bachtiar, A. M. 2017. Implementasi Algoritma K-Nearest Neighbor dalam Pengklasifikasian Follower Twitter yang Menggunakan Bahasa Indonesia. *Jurnal Sistem Informasi*, **13**(1), pp. 31-37.

[10] Nayak, A., Poriya, A., & Poojary, D. 2013. Type of NOSQL databases and its comparison with relational databases. *International Journal of Applied Information Systems*, **5**(4), pp. 16-19.

[11] Pippal, S., Batra, L., Krishna, A., Gupta, H., & Arora, K. 2014. Data mining in social networking sites: A social media mining approach to generate effective business strategies. *International Journal of Innovations & Advancement in Computer Science*, **3**(2), pp. 22-27.

[12] Knafli C N, 2015 *Storytelling with Data* **2** Wiley.

[13] Yuk, M., & Diamond, S. 2014. *Data visualization for dummies*. John Wiley & Sons..

2. First Revision (1 Des 2020)

Dian Dharmayanti

<dian.dharmayanti@email.unikom.ac.id>

[JESTEC] Manuscript ID: JESTEC_16_1_2021_D Dharmayanti – Minor Revision
1 message

Editorial Office <jjestec@gmail.com>   1 Desember 2020 at 08:10
To: Dian Dharmayanti <dian.dharmayanti@email.unikom.ac.id>

Dear Author,

Thank you for submitting the following manuscript to Journal of Engineering, Science and Technology:

Journal Name	Journal of Engineering, Science and Technology
Manuscript ID	JESTEC_16_1_2021_D Dharmayanti
Authors	Dian Dharmayanti, Adam Mukharil Bachtiar, Firman Nizammudin Fakhrul
Title	Data Visualization for Content Marketing Domain in Social Media

It has been reviewed by experts in the field and we request that you make major revisions before it is processed further. Please find the comments from reviewer in the attached file.

Please revise the manuscript according to the reviewers' comments and upload the revised file within 10 days. Use the version of your manuscript found at the above link for your revisions, as the editorial office may have made formatting changes to your original submission.

If the reviewers have suggested that your manuscript should undergo extensive English editing, please address this during revision. We suggest that you have your manuscript checked by a native English-speaking colleague or use a professional English editing service.

Do not hesitate to contact us if you have any questions regarding the revision of your manuscript or if you need more time. We look forward to hearing from you soon.

Kind regards,

Editorial Office
Journal of Engineering, Science and Technology
E-Mail: jjestec@gmail.com

DATA VISUALIZATION FOR CONTENT MARKETING DOMAIN
IN SOCIAL MEDIA

Dian Dharmayanti* , Adam Mukharil Bachtiar , Firman Nizammudin Fakhrul

Informatics Engineering, Faculty of Engineering and Computer Science, Universitas Komputer Indonesia, Jalan Dipatiukur Nomor 112-116, Bandung, Jawa Barat, Indonesia
Corresponding Author dian.dharmayanti@email.unikom.ac.id

Email : *adam@email.unikom.ac.id

Abstract. This research will focus on producing proper data visualization in the content marketing domain in social media. The research method consists of nine steps and using techniques such as empathy map, statistics, data mining, data visualization technique, acceptance testing, and usability testing to generate proper data visualization. A prototype will be developed as the proof of concept of this research. The usability testing result of the prototype says that the effectiveness level is 100% and the efficiency level is 100%. By using the proper data visualization, it can increase the efficiency and effectiveness level of the content marketing research on social media for marketers.

1. Introduction

Content marketing is now widely used by marketers to attract customers. Content marketing is a marketing process to create and distribute content that aims to give appeal to the intended target audience [1]. The crucial thing in content marketing is the content marketing strategy. But in fact, 62% of marketers do not have a solid strategy [2]. Several well-known brands use a data-driven approach to create a solid strategy [3]. One source of data that can be used to create a content marketing strategy is social media. Social media produce a very large number of data [4]. The huge data can be a tremendous opportunity for marketers to improve marketing content strategies, such as understanding the target audience and content discovery [5].

Based on this phenomenon, an interview was conducted with one of the marketers. The problem that arises is that marketers need a lot of time to read information from social media data needed for content marketing, such as target audience segmentation. The main factor of the difficulty that is due to the social media used by marketers provides an inappropriate form of data visualization so the data must be analyzed again to get the desired information. For example, audience research, marketers enter a keyword and analyze the characteristics of the target audience relevant to that keyword. In addition, errors in the process of obtaining such information cannot be predicted so that it can have an impact on misinformation.

An effective way to present detailed data into information that is easily accepted is by means of abstraction into visual information [6]. As explained earlier, marketers need a significant amount of time to read information from social media data relating to content marketing strategies due to inappropriate forms of data visualization or visual information. Therefore, further analysis is needed to find out what kind of data visualization is right in the content marketing domain on social media. The knowledge generated is used to create the right data visualization in the content marketing domain on social media for marketers.

2. Method

The research methodology was inspired by the book "Visualizing Data" which consists of 7 steps that acquire, parse, filter, mine, represent, refine, and interact [7]. The methodology adapted to the needs of this research. The research methodology can be seen in Figure 1.
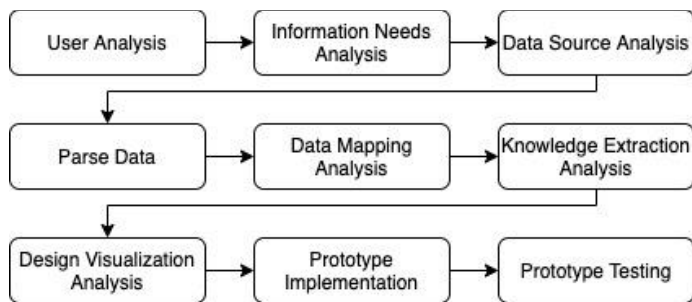
Figure 1. Research Methodology

3. Results and Discussion

All stages in the research methodology are being explained in more detail in this section. The explanation consists of input, process, and output for each stage.

3.1. User Analysis

A good data visualization is well supported by identifying the user who will read the visualization itself [8]. To identify the user, the technique used is the interview. In this research, two marketers who are an actor in the field of content marketing are being interviewed. The interview is divided into three sections, those are opener question section, daily task section, and use of content marketing tools section. The example question of each section can be seen in Table 1.

Table 1. Interview Section

Section	Sample Question
Opener question	What is your role in the content marketing team?
Daily task question	What activities do you do every day associated with content marketing?
Use of content marketing tools section	What difficulty do you face when using content marketing tools?

Commented [WU4]: Adjust all Tables to JESTEC format

The result of the interview is qualitative data. In this research, the empathy map is used as a tool to illustrate the interview result. The empathy map can be seen in Figure 2. The result shows that the user has two pain points, spending long hours to get and analyze the information about the target audience and confusion why the published content is not going well. Then, this empathy map is used as an archetype user to analysis, design, and testing the data visualization in this research.

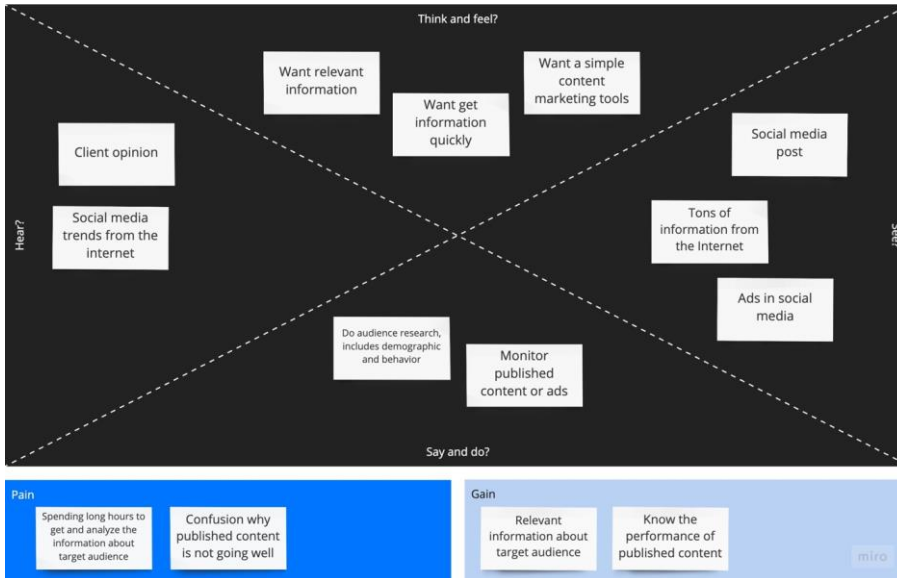Commented [WU5]: Make it to one pharagraph

Figure 2. Empathy Map

3.2. Information Needs Analysis

Information needs analysis stage conducted to determine what information is needed by marketers in the domain of content marketing in social media. Empathy map from the previous stage is used to determine the information needs. Furthermore, the pain and gain outlined the information needs. The sample of information needs that are formed can be seen in Table 2.

Table 2. Information Needs

Pain	Gain	Information Needs
Spending long hours to get and analyze the information of the target audience	**Get relevant information about the target audience quickly**	How to know the interest of a segment of the target audience?
Confusion why the published content is not going well	**Know the performance of published content**	How to know the sentiment score from the post?

The result of this stage is nine information needs based on breakdown pain and gain of the empathy maps. These information needs are used as a reference for visualizing the data on the domain of content marketing in social media.

3.3. Data Source Analysis

Data source analysis conducted to analyze the characteristic of the data source. The data source used is data from Twitter. The data were obtained through a RESTful API provided by Twitter. Twitter data can be accessed through a specific endpoint. The complete list of Twitter API endpoint can be

accessed at https://developer.twitter.com/en/docs/api-reference-index. Example endpoint of Twitter API can be seen in Table 3.

Table 3. Example of Twitter API

Method	Endpoint Path	Output	URL Documentation
GET	users/show	User Object	https://developer.twitter.com/en/docs/accounts-and-users/follow-search-get-users/api-reference/get-users-show

Twitter API can be accessed through endpoints based on its features. The data obtained shaped as JSON (JavaScript Object Notation). The data contains Twitter data objects, where Twitter data objects are divided into several types, namely tweet objects, user objects, entities object, and geo objects.
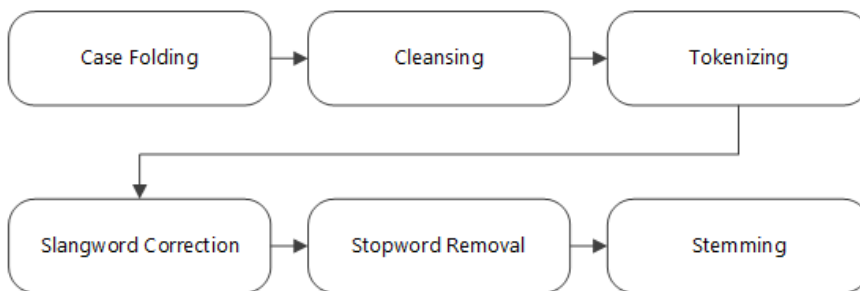
3.4. Parse Data
Parse data stage conducted to make the data easier to be processed. In this stage, text preprocessing and format conversion of data into a document-based database are done.

3.4.1. Text Preprocessing
Tweets data objects containing text that is not structured so it has to be preprocessed. The result of preprocessing will make the data mining stage easier [9]. Steps being taken can be seen in

Figure 3.

Figure 3. Text Preprocessing Steps

The sample result for a tweet that being preprocessed can be seen in Table 4.

Table 4. Sample Result of Text Preprocessing

Input	Output
Pagi yg menyenangkan. Cari bubur mantap utk sarapan. https://t.co	[pagi, senang, cari, bubur, mantap, sarapan]

3.4.2. Document and Data Structure Formation
Document-based database offers good performance and scalability [10]. To make it easier to process the data at the next stages, four types of documents formed. Documents that formed are user

---

Commented [WU7]: Check your grammar and mistype

documents, followers document, and mention document. Endpoint used and output objects of the document formed can be seen in Table 5.

Table 5. Document Formation

Method	Endpoint	Output

User Document	followers/list
	users/show
		{
		  "_id":String,
		  "user_profile":{
		    …
		    },
		    "user_tweets":[
		     ...
		    ]
		}

Followers Document	followers/list
	users/show
	statuses/user_timeline
		{
		  "_id":String,
		  "following":{
		    …
		    },
		    "follower_profile":{
		     ...
		    },
		    "follower_tweets":[
		     ...
		    ],
		}

Mention Document	users/show
	statuses/mention_timeline
		{
		  "_id":String,
		  "user_profile":{
		    …
		    },
		    "user_mentions":[
		     ...
		    ]
		}

This stage produces three documents. The documents are user documents, followers document, and mentions document. The documents will be used as a data source for the visualization.

3.5. Data Mapping Analysis

Data mapping stage conducted to choose and map data based on the information needs. Mapping data to the sample information needs can be seen in Table 6.

Table 6. Mapping Data to Information Needs

Gain	Information Needs	Document Used
Get relevant information on the target audience quickly	How to know the interest of a segment of the target audience?	User document, Followers document
Know the performance of	How to know the sentiment	User document,

published content	score from the post?	Mention document

3.6. Knowledge Extraction Analysis

The data that have been obtained and mapped are then performed statistical calculations or data mining methods to extract useful information. Data mining technique is used because it can be used to analyze the data and summarize it into useful information [11]. This stage will also determine whether the information needs of an exploratory analysis or explanatory analysis to understand the context of the needs of the information so that it can be concluded whether the information needs requires knowledge extraction method or not. The mapping of the knowledge extraction method is depended on the characteristic of the data too [8]. The mapping of the knowledge extraction method to sample information needs can be seen in Table 7.

Table 7. Mapping Knowledge Extraction Method

Gain	Information Needs	Exploration Type	Method
Get relevant information on the target audience quickly	How to know the interest of a segment of the target audience?	Exploratory	Classification, Statistics (distribution frequency and mode)
Know the performance of published content	How to know the sentiment score from the post?	Exploratory	Classification, Statistics (distribution frequency and mode)

3.7. Design Visualization Analysis

Design visualization analysis includes four steps that were inspired in the "storytelling with data" book. The steps are understanding the context, choosing the appropriate form of visualization, eliminating clutter, focusing attention, and acceptance testing [12]. In this stage, the information needs a sample that is used is "How to know the interest of a segment of the target audience?".

3.7.1. Understanding Context

Understanding the context of information needs is done by describing the context of who needs the information, what information will be delivered, and how the delivery of information. The example description of the context can be seen in Table 8.

Table 8. Understanding Context

Who	What	How
A content marketer	Knowing the interest of a segment of the target audience	Displays a comparison of the list of interests of a target audience segment

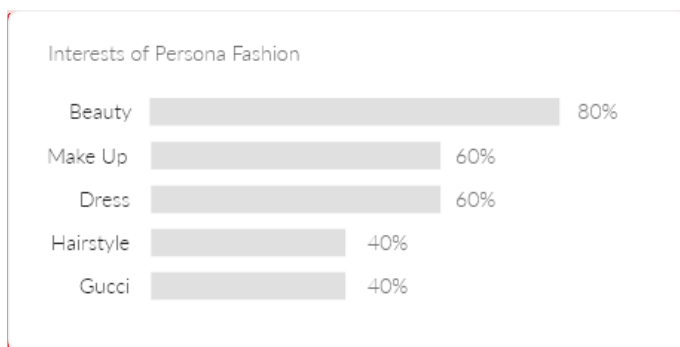#### 3.7.2. Choosing Appropriate Type of Visualization

Choosing an appropriate type of visualization is done by analyzing how in understanding the context stage. The example of choosing the appropriate type of visualization can be seen in Table **9**.

Table 9. Choosing Appropriate Type of Visualization

How	Visualization Type	Why
Displays a comparison of the list of interests of a target audience	Horizontal bar chart	Compare several categories

segment

3.7.3. Eliminating Clutter

Data visualization is in accordance with principles that will not contain clutter. The sketch of the visualization should be made and can be seen in Figure 4. The sketch is analysed and should contain no clutter.

Figure 4. Sketch of Visualization

The example of clutter analysis can be seen in **Table 10**.

Table 10. Clutter Analysis

Principle	Explanation	Checklist
Alignment	Alignment used is left-aligned because reading information will be easier	✔
White space	White space is only used as a margin between components so they don't overlap	✔
Closure	There is no unnecessary border on data visualization	✔
Continuity	Remove the y-axis border because the bar always starts from the same point so there is no interference between the label and the bar	✔

3.7.4. Focusing Attention

Focusing attention made to focus the user on the information that is important to the user. The example of focusing attention analysis can be seen in Table 11.

Table 11. Focusing Attention

How	Focused Information	Preattentive Attributes
Displays a comparison of the list of interests of a	Most dominant interest	Color

target audience segment	

The result of visualization after focusing the attention can be seen in Figure 5.

Figure 5. Visualization that Has Been Focused

3.7.5. Acceptance Testing
Testing is done by acceptance testing. The result of acceptance testing can be seen in Table 12.

Table 12. Acceptance Testing

Actual Information	Obtained Information	Comment	Result
Comparison of the list of interests of a target audience segment	Comparison of the list of interests of a target audience segment	No comment	Accepted

A good data visualization can give the user the right information [13]. So, the acceptance criteria of this testing are the information obtained is correct and there is no additional comment. The conclusion from acceptance testing is that the form of visualization is appropriate.

3.8. Prototype Implementation
Data visualization is made in prototype form. Sequence information on the prototype is based on the results of card sorting and grouping by pain and gain of the empathy maps. The prototype can be seen in Figure 6 and Figure **7**.
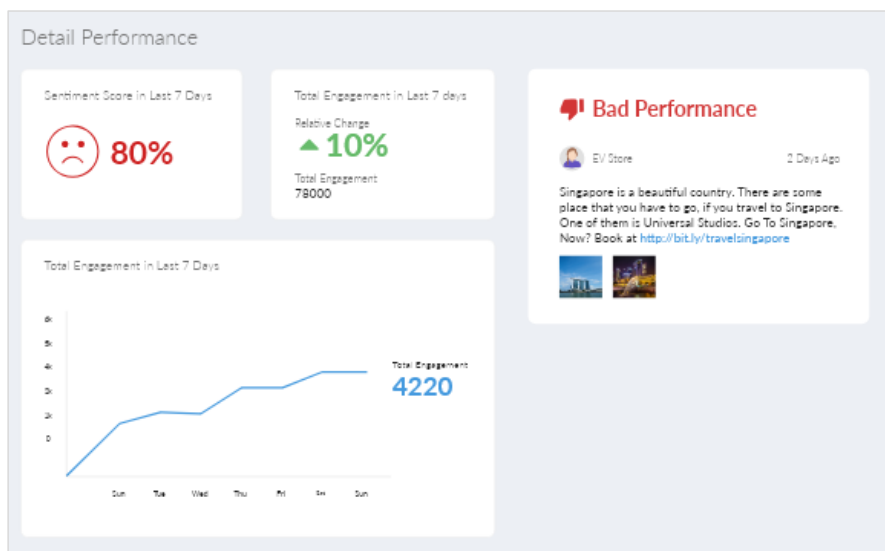
Figure 6. Prototype of First Group

Figure 7. Prototype of Second Group

3.9. Prototype Testing

Tests on a prototype using usability testing techniques. Scenario testing is divided based on the gain in the empathy map in the user analysis stage. There are 2 scenarios given to the user. The effectiveness level is 100% and the efficiency level is 100%.

4. Conclusion

Based on all processes in this research, the conclusion is that the data visualization in this research can increase the efficiency and effectiveness level of the content marketing research on social media for marketers. For the next research, so much data in social media that can be more explored and the data source can be anything besides Twitter.

References

[1] Chan A and Astari D, 2017 The Analysis of Content Marketing in Online Fashion Shops in Indonesia *Rev. Integr. Bus. Econ. Res.* **6**(2) p. 225–233.

[2] Beets, L. M., & Handley, A. 2018. B2C Content marketing-2018 Benchmarks, Budgets, and Trends—North America.

[3] Karczewski T, 2014, 6 Steps to a Data-Driven Content Marketing Strategy. [Online]. Available: http://contentmarketinginstitute.com/2014/06/data-driven-content-marketing-strategy/. [Accessed: 03-Mar-2018].

[4] Kharb, L. 2017. Exploration of Social Networks with Visualization Tools *Am. J. Eng. Res.* **6**(3), pp. 90–93.

[5] Litsa T, 2017, 5 Tips to Create a Data-Driven Content Marketing Strategy, *Search Engine Watch*. [Online]. Available: https://searchenginewatch.com/2017/09/19/5-tips-to-create-a-data-driven-content-marketing-strategy/. [Accessed: 03-Mar-2018].

[6] Syaripul, N. A., & Bachtiar, A. M. 2016. Visualisasi Data Interaktif Data Terbuka Pemerintah Provinsi DKI Jakarta: Topik Ekonomi dan Keuangan Daerah. *Jurnal Sistem Informasi*, **12**(2), pp. 82-89.

[7] Fry, B. 2008. Visualizing Data. 1005 Gravenstein Highway North, Sebastopol, CA 95472.

[8] Bachtiar A. M., & Dharmayanti D. 2018. Data Visualization at Trading Sector Using Government's Open Data *Int. J. Pure Appl. Math.* **118**, 20 p. 523–530.

[9] Rivki, M., & Bachtiar, A. M. 2017. Implementasi Algoritma K-Nearest Neighbor dalam Pengklasifikasian Follower Twitter yang Menggunakan Bahasa Indonesia. *Jurnal Sistem Informasi*, **13**(1), pp. 31-37.

[10] Nayak, A., Poriya, A., & Poojary, D. 2013. Type of NOSQL databases and its comparison with relational databases. *International Journal of Applied Information Systems*, **5**(4), pp. 16-19.

[11] Pippal, S., Batra, L., Krishna, A., Gupta, H., & Arora, K. 2014. Data mining in social networking sites: A social media mining approach to generate effective business strategies. *International Journal of Innovations & Advancement in Computer Science*, **3**(2), pp. 22-27.

[12] Knafli C N, 2015 *Storytelling with Data* **2** Wiley.

[13] Yuk, M., & Diamond, S. 2014. *Data visualization for dummies*. John Wiley & Sons..

Commented [WU9]: Adjust the reference to the JESTEC format

3. Submit Revised Paper (8 Des 2020)

Dian Dharmayanti

<dian.dharmayanti@email.unikom.ac.id**>**

[JESTEC] Manuscript ID: JESTEC_16_1_2021_D Dharmayanti –Revised VersionReceived

1 message

Editorial Office <jjestec@gmail.com>   8 Desember 2020 at 13:17
To: Dian Dharmayanti <dian.dharmayanti@email.unikom.ac.id>

Dear Author,

Thank you very much for resubmitting the modified version of the following manuscript:

Journal Name	Journal of Engineering, Science and Technology
Manuscript ID	JESTEC_16_1_2021_D Dharmayanti
Authors	Dian Dharmayanti, Adam Mukharil Bachtiar, Firman Nizammudin Fakhrul
Title	Data Visualization for Content Marketing Domain in Social Media

A member of the editorial office will be in touch with you soon regarding progress of the manuscript.

Kind regards,

Editorial Office
Journal of Engineering, Science and Technology
E-Mail: jjestec@gmail.com

*** This is an automatically generated email ***

DATA VISUALIZATION FOR CONTENT MARKETING DOMAIN IN SOCIAL MEDIA

DIAN DHARMAYANTI[1]*, ADAM MUKHARIL BACHTIAR[2], FIRMAN NIZAMMUDIN FAKHRUL[3]

[1,2,3]Informatics Engineering, Faculty of Engineering and Computer Science,
Universitas Komputer Indonesia,
Jalan Dipatiukur Nomor 112-116, Bandung, Jawa Barat, Indonesia
*Corresponding Author : dian.dharmayanti@email.unikom.ac.id

Abstract

The purpose of this research is to determine the proper data visualization on the domain of content marketing in social media. Based on interviews with marketers, they need a lot of time to read information from social media data to create content marketing strategy. It occurs due to inappropriate forms of data visualization, so that the data must be analyzed again to get the desired information. To overcome this problem, further analysis is needed to know what kind of data visualization is suitable in the domain of content marketing on social media. The research method consisted of nine steps and used techniques such as empathy map, statistics, data mining, data visualization technique, acceptance testing, and usability testing to generate proper data visualization. To know if the data visualization is correct, acceptance testing was carried out on data visualization and usability testing on the prototype. Acceptance testing results indicated that 100% of the designed visualization forms are accepted. In addition, the results of usability testing on the prototype showed that effectiveness and efficiency reaches 100%. From these results, it was concluded that visualization of data in the domain of content marketing on social media is made correctly.

Keywords: Data Visualization, Content marketing, Social media, Statistics, Data mining.

1. Introduction

Content marketing is now widely used by marketers to attract customers. It is a marketing process to create and distribute content that aims to give appeal to the intended target audience [1]. The crucial thing in content marketing is the content marketing strategy. However, 62% of marketers do not have a solid strategy [2]. Several well-known brands use a data-driven approach to create a solid strategy. One source of data that can be used to create a content marketing strategy is social media. Social media produce a very large number of data. The huge data can be a tremendous opportunity for marketers to improve marketing content strategies, such as understanding the target audience and content discovery [3].

Based on this phenomenon, an interview was conducted with one of the marketers. The problem that arises is that marketers need a lot of time to read information from the required social media data for content marketing, such as target audience segmentation. The main difficulty factor is because social media used by marketers provides an inappropriate form of data visualization. Therefore, the data must be reanalyzed to get the desired information. The use of several techniques such as personas, statistics, data mining, and data visualization techniques were used to produce accurate data visualizations. To find out that the data visualization is correct, acceptance testing is carried out on the data visualization and usability testing on the prototype.

An effective way to present detailed data into information that is easily accepted is through the process of abstraction into visual information [4]. As explained earlier, marketers need a significant amount of time to read information from social media data relating to content marketing strategies. Therefore, further analysis is needed to determine what kind of data visualization is suitable on the domain of content marketing in social media. The resulting knowledge is used to create data visualizations right on the domain of content marketing in social media for marketers. The purpose of this research is to produce appropriate data visualization in the content marketing domain on social media for marketers.

2. Method

The research methodology was inspired by the book "Visualizing Data" which consists of seven steps namely acquire, parse, filter, mine, represent, refine, and interact [4]. The methodology adapted to the needs of this research is shown in Fig. 1.
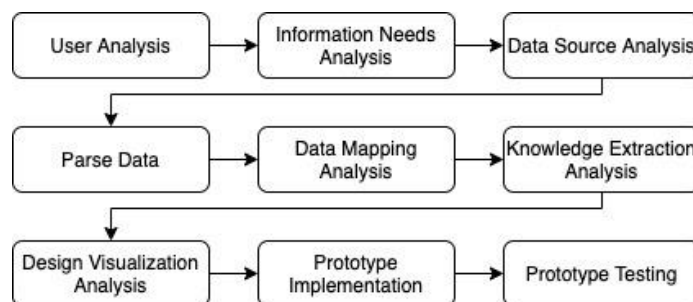

Fig. 1. Research Methodology.

3. Results and Discussion

All stages in the research methodology are explained in more detail in this section. The explanation consists of input, process, and output for each stage.

3.1. User Analysis

A good data visualization is well supported by identifying the users who read the visualization itself [4]. To identify the user, the technique used was interview. In this research, two marketers were interviewed. The interview is divided into three sections namely opener question, daily task, and use of content marketing tools section. The example question of each section is shown in Table 1.

Table 1. Interview Section

Section	Sample Question
Starter question	What is your role in the content marketing team?
Daily task question	What activities do you do every day associated with content marketing?
Use of content marketing tools question	What difficulty do you face when using content marketing tools?

The result of the interview is qualitative data. In this research, the empathy map is used as a tool to illustrate the interview result which is shown in Fig. 2. The result shows that the user has two pain points namely spending too much time to get and analyze the information about the target audience as well as finding a reason why the published content is not going well. Then, this empathy map is used as an archetype user to analyze, design, and test the data visualization in this research.
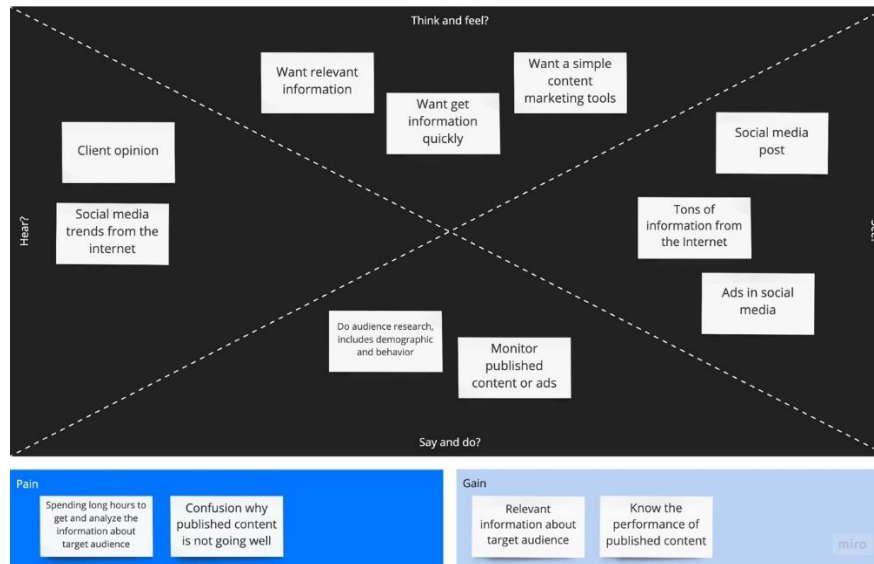
Fig. 2. Emphaty Map.

3.2. Information Needs Analysis

Stages in analyzing information need to be implemented to determine what information is needed by marketers in the content marketing domain in social media [5]. Empathy map from the previous stage is used to determine the information needs. Furthermore, the pain and gain points outlined the information needs. The sample of information needs is shown in Table 2.

Table 2. Information Needs

Pain	Gain	Information Needs
Spending too much time to get and analyze the information of the target audience	**Getting relevant information about the target audience quickly**	How to know a segment interest of target audiences?
Finding a reason why the published content is not going well	**Knowing the performance of published content**	How to know the sentiment score from the post?

Based on the results of pain and gain analysis sections from the empathy map, nine information needs are generated. These information needs are used as a reference for visualizing the data on the domain of content marketing in social media.

3.3. Data Source Analysis

Data source analysis was conducted to analyze the characteristics of the data used [6]. The data source used is from Twitter. The data were obtained through a RESTful API provided by Twitter. Twitter data can be accessed through a specific endpoint [7]. The complete list of Twitter API endpoint can be accessed at https://developer.twitter.com/en/docs/api-reference-index. Example endpoint of Twitter API is shown in Table 3.

Table 3. Example of Twitter API

Method	Endpoint Path	Output	URL Documentation
GET	users/show	User Object	https://developer.twitter.com/en/docs/accounts-and-users/follow-search-get-users/api-reference/get-users-show

Twitter API can be accessed through endpoints based on its features. The data obtained shaped as JSON (JavaScript Object Notation). It contains Twitter data objects, where the data objects are divided into several types, namely tweet objects, user objects, entities object, and geo objects.

3.4. Parse Data

Parse data stage is carried out to make the data easier to be processed [7]. In this stage, text preprocessing and format conversion of data into a document-based database are done.

3.4.1. Text Preprocessing

Text preprocessing stage is carried out because tweets data objects cointain text is not structured. Data from a more structured preprocessing facilitates data mining stage [8]. The steps is shown in Fig. 3.
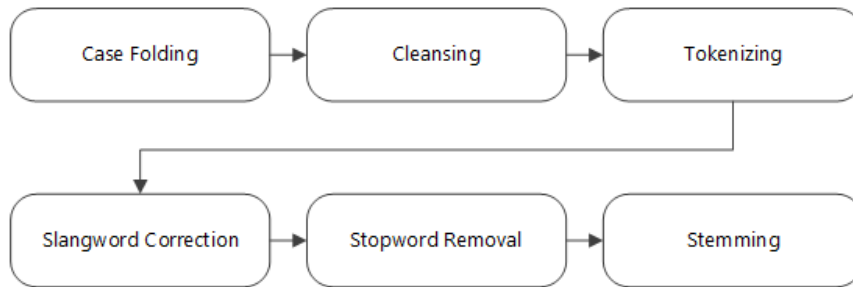

Fig. 3. Text Preprocessing Steps

The sample result for a tweet that being preprocessed is shown in Table 4.

Table 4. Sample Result of Text Preprocessing

Input	Output
Pagi yg menyenangkan. Cari bubur mantap utk sarapan. https://t.co	[pagi, senang, cari, bubur, mantap, sarapan]

3.4.2. Document and Data Structure Formation

Document-based database offers a good performance and scalability [9]. To make it easier to process data in the next stage, four types of documents were formed. Documents that formed are user documents, followers document, and mention document. Endpoint used and output objects of the document formed is shown in Table 5.

Table 5. Document Formation

Method	Endpoint	Output

User Document	followers/list
	users/show
		{
		  "**_id**":String,
		  "**user_profile**":{
		    …
		    },
		  "**user_tweets**":[
		    ...
		    ]
		}

Followers Document	followers/list
	users/show
	statuses/user_timeline
		{
		  "**_id**":String,
		  "**following**":{
		    …

Method	Endpoint	Output

		},
		 "follower_profile":{
		 ...
		 },
		 "follower_tweets":[
		 ...
		],
		 }

Mention Document	users/show
	 statuses/mention_timeline
		{
		 "_id": String,
		 "user_profile":{
		 …
		 },
		 "user_mentions":[
		 ...
		]
		 }

This stage produces three documents. The documents are user documents, followers document, and mentions document. The documents was used as a data source for the visualization.

3.5. Data Mapping Analysis

Data mapping stage was conducted to choose and map data based on the information needs [9]. Mapping data to the sample information is shown in Table 6.

Table 6. Mapping Data to Information Needs

Gain	Information Needs	Document Used
Getting relevant information on the target audience quickly	How to know the interest of a segment of the target audience?	User document, Followers document
Knowing the performance of published content	How to know the sentiment score from the post?	User document, Mention document

3.6. Knowledge Extraction Analysis

The data that have been obtained and mapped are then performed statistical calculations or data mining methods to extract useful information. Data mining technique is used because it can be used to analyze the data and summarize it into a useful information [9]. This stage also determine whether the information needs of an exploratory analysis to understand the context of the information needs. Therefore, it can be concluded if the information needs requires knowledge extraction method. The knowledge extraction method mapping is also depended on the characteristic of the data [10]. The knowledge extraction method mapping to sample information needs is shown in Table 7.

Table 7. Mapping Knowledge Extraction Method

Gain	Information Needs	Exploration Type	Method
Getting relevant information on the target audience quickly	Information to determine the interests of the target audience segment	Exploratory	Classification, Statistics (distribution frequency and mode)
Knowing the performance of published content	Information to find out the sentiment score of the post	Exploratory	Classification, Statistics (distribution frequency and mode)

3.7. Design Visualization Analysis

Design visualization analysis includes four steps that were inspired in the "storytelling with data" book. The steps are understanding the context, choosing the appropriate form of visualization, eliminating clutter, focusing attention, and testing the acceptance [11]. In this stage, the information needs used a sample, which is "How to know the interest of a segment of the target audience?".

3.7.1. Understanding Context

Understanding the context of information needs is done by describing the context of who needs the information, what information will be delivered, and how the information is delivered. The example of context description is shown in Table 8.

Table 8. Understanding Context

Who	What	How
A content marketer	Knowing the interest of a segment of the target audience	Displays a comparison of the interests list of a target audience segment

### 3.7.2. Choosing Appropriate Type of Visualization

Choosing an appropriate type of visualization is done by analyzing how to understand the context stage. The example of choosing the appropriate type of visualization is shown in Table 9.

Table 9. Choosing Appropriate Type of Visualization

How	Visualization Type	Why
Displays a comparison of the interests list of a target audience segment	Horizontal bar chart	Compare several categories

3.7.3. Eliminating Clutter

Data visualization is in accordance with principles that are not contain clutter. The visualization sketch should be made (See Fig. 4). The sketch is analyzed and should not contain clutter.
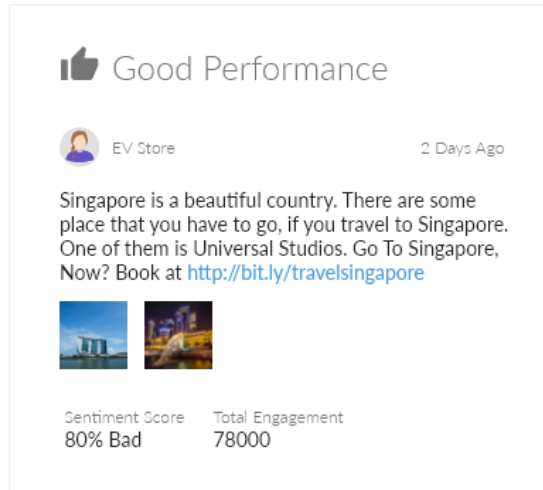
Fig. 4. Sketch of Visualization.

The example of clutter analysis is shown in Table 10.

Table 10. Clutter Analysis

| Principle	Explanation	Checklist
Alignment	Alignment used is left-aligned because reading information can be easier	✓
White space	White space is only used as a margin between components so they do not overlap	✓
Closure	There is no unnecessary border on data visualization	✓
Continuity	Remove the y-axis border because the bar always starts from the same point so there is no interference between the label and the bar	✓

3.7.4. Fosucing Attention

Focusing attention was made to focus the user on the information. It is important to the user. The example of focusing attention analysis is shown in Table 11.

Table 11. Focusing Attention

How	Focused Information	Preattentive Attributes
Displays a comparison of the interests list of a target audience segment	Most dominant interest	Color

The result of visualization after focusing the attention is shown in Fig. 5.
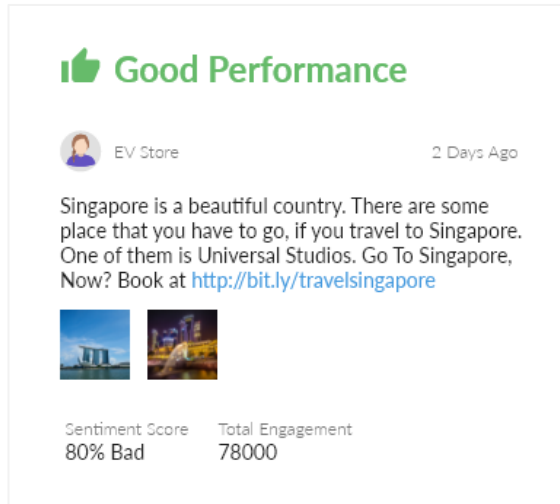
Fig. 5. Visualization that Has Been Focused.

3.7.5. Acceptance Testing

Testing is done by acceptance testing. The result of acceptance testing is shown in Table 12.

Table 12. Acceptance Testing

Actual Information	Obtained Information	Comment	Result
Comparison of the interests list of a target audience segment	Comparison of the interests list of a target audience segment	No comment	Accepted

A good data visualization can give the user a right information [12,13]. Therefore, the acceptance criteria of this testing are the correct information obtained with no additional comment. The conclusion from acceptance testing is that the form of visualization is appropriate.

3.8. Prototype Implementation

Data visualization was made in prototype form. Sequence information on the prototype is based on the card sorting results. It was grouping by the empathy maps pain and gain [12]. The prototype is shown in Fig. 6 and Fig. 7.
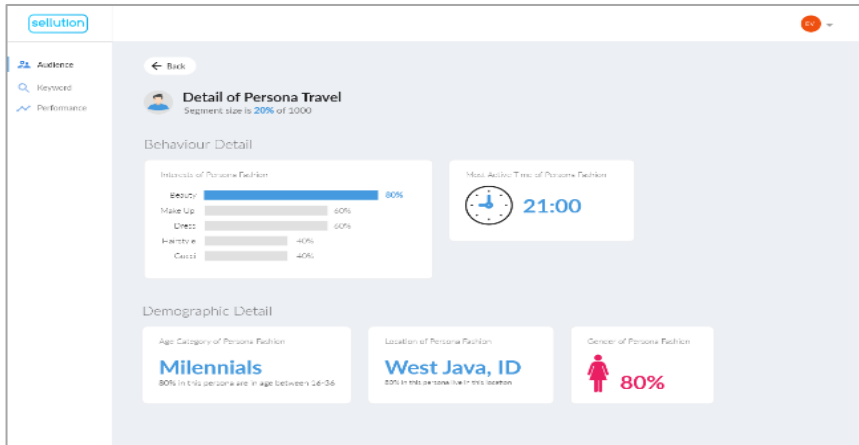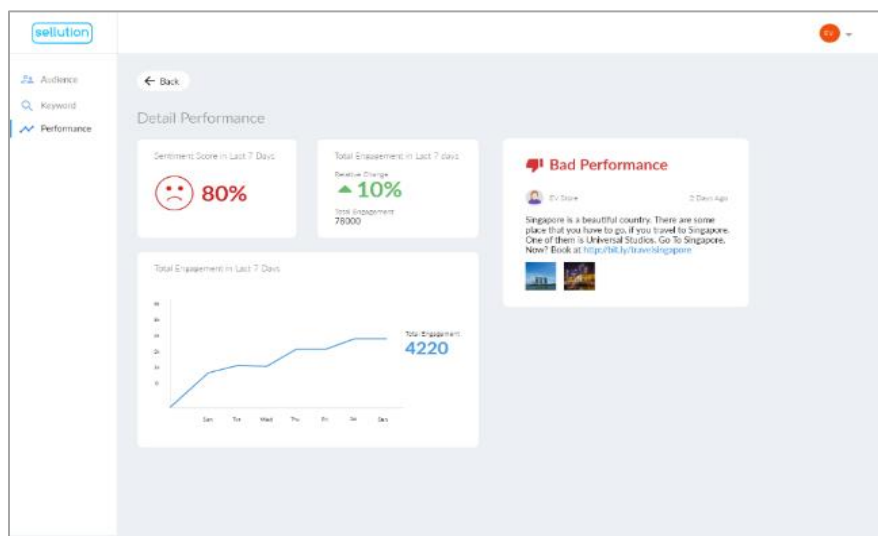
Fig. 6. Prototype of First Group.

Fig. 7. Prototype of Second Group.

3.9. Prototype Testing

Tests on a prototype was using usability testing techniques [13]. Scenario testing is divided based on the gain in the empathy map. There are two scenarios given to the user. The effectiveness and efficiency level are 100%, respectively.

4.   Conclusion

Based on this research, the data visualization can increase the efficiency and effectiveness level of the content marketing research on social media. For the next research, there are a lot of data in social media that can be explored and the data source are various.

References

1. Chan, A.; and Astari, D. (2017). The analysis of content marketing in online fashion shops in Indonesia. *Review of Integrative Business and Economics Research,* 6(2), 225.
2. Järvinen, J.; and Taiminen, H. (2016). Harnessing marketing automation for B2B content marketing. *Industrial Marketing Management*, 54, 164-175.
3. Lieb, R. (2017). *Content-the atomic particle of marketing: the definitive guide to content marketing strategy*. Kogan Page Publishers.
4. Baltes, L. P. (2015). Content marketing-the fundamental tool of digital marketing. *Bulletin of the Transilvania University of Brasov. Economic Sciences. Series V*, 8(2), 111-118.
5. Kharb, L. (2017). Exploration of Social Networks with Visualization Tools. *American Journal of Engineering Research (AJER)*, 6(3), 90-93.
6. Hansen, D.; Shneiderman, B.; and Smith, M. A. (2010). *Analyzing social media networks with NodeXL: Insights from a connected world*. Morgan Kaufmann.
7. Kandogan, E. (2012, October). Just-in-time annotation of clusters, outliers, and trends in point-based data visualizations. In *2012 IEEE Conference on Visual Analytics Science and Technology (VAST)*, pp. 73-82.
8. Pippal, S.; Batra, L.; Krishna, A.; Gupta, H.; and Arora, K. (2014). Data mining in social networking sites: A social media mining approach to generate effective business strategies. *International Journal of Innovations & Advancement in Computer Science*, 3(2), 22-27.
9. Nayak, A.; Poriya, A.; and Poojary, D. (2013). Type of NOSQL databases and its comparison with relational databases. *International Journal of Applied Information Systems*, 5(4), 16-19.
10. Bachtiar, A. M.; and Dharmayanti, D. Data visualization at trading sector using government's open data. 118, 523-530.
11. Jaeschke, G.; Leissler, M.; and Hemmje, M. (2005). Modeling interactive, 3-dimensional information visualizations supporting information seeking behaviors. In *Knowledge and information visualization*. pp. 119-135.
12. Diamond, M.; and Mattia, A. (2017). Data visualization: An exploratory study into the software tools used by businesses. *Journal of Instructional Pedagogies*, 18, 1-7.
13. Mifsud, J. (2015). Usability metrics–A guide to quantify the usability of any system. Usability Geek. [Online]. Available: https://usabilitygeek.com/usability-metrics-a-guide-to-quantify-system-usability/.

4. Accepted Submission (15 Jan 2021)

M Gmail

Dian Dharmayanti

<dian.dharmayanti@email.unikom.ac.id>

[JESTEC] Manuscript ID: JESTEC_16_1_2021_D Dharmayanti –Accepted

1 message

Editorial Office <jjestec@gmail.com>   15 January 2021 at 15:05
To: Dian Dharmayanti <dian.dharmayanti@email.unikom.ac.id>

Dear Author,

We are pleased to inform you that the following paper has been officially
accepted for publication:

Journal Name	Journal of Engineering, Science and Technology
Manuscript ID	JESTEC_16_1_2021_D Dharmayanti
Authors	Dian Dharmayanti, Adam Mukharil Bachtiar, Firman Nizammudin Fakhrul
Title	Data Visualization for Content Marketing Domain in Social Media

We will now make the final preparations for publication, then return the
manuscript to you for your approval.

If, however, extensive English edits are required to your manuscript, we will
need to return the paper requesting improvements throughout.

Kind regards,

Editorial Office
Journal of Engineering, Science and Technology
E-Mail: jjestec@gmail.com

5. Paper Publish (3 Feb 2021)

M Gmail

Dian Dharmayanti

<dian.dharmayanti@email.unikom.ac.id>

[JESTEC] Manuscript ID: JESTEC_16_1_2021_D Dharmayanti –Published Online

1 message

Editorial Office <jjestec@gmail.com>   3 February 2021 at 14:23
To: Dian Dharmayanti <dian.dharmayanti@email.unikom.ac.id>

Dear Author,

We are pleased to inform you that your article:

Journal Name	Journal of Engineering, Science and Technology
Manuscript ID	JESTEC_16_1_2021_D Dharmayanti
Authors	Dian Dharmayanti, Adam Mukharil Bachtiar, Firman Nizammudin Fakhrul
Title	Data Visualization for Content Marketing Domain in Social Media

is available online:

https://jestec.taylors.edu.my/Vol%2016%20issue%201%20February%202021/16_1_24.pdf

Please take a moment to check that everything is correct. You can reply to
this email if there is a problem. Note that at this stage we will not accept
further changes to the manuscript text.

Thank you for choosing our journal to publish your work, we look forward
to receiving further contributions from your research group in the future.

Kind regards,

Editorial Office
Journal of Engineering, Science and Technology
E-Mail: jjestec@gmail.com