Date of publication DEC-31, 2024, date of current version OCT-29, 2024
www.computingonline.net/ computing@computingonline.net

Print ISSN 1727-6209
Online ISSN 2312-5381
DOI 10.47839/ijc.23.4.3763

Research on the use of AI for selecting abstractions for natural language image generation tools

VOLODYMYR YAKYMIV[1], YOSYF PISKOZUB[1,2]
[1]Lviv Polytechnic National University, Institute of Applied Mathematics and Fundamental Sciences, Department of Applied Mathematics, Lviv, 79000 Ukraine
vvolodymyr.s.yakymiv@lpnu.ua, yosyf.z.piskozub@lpnu.ua
[2]Department of Applied Mathematics, Faculty of Computer Science and Telecommunications, Cracow University of Technology, Warszawska Str., 24, 31-155 Krakow, Poland (e-mail: yosyf.piskozub@pk.edu.pl)

Corresponding author: Volodymyr Yakymiv (e-mail: volodymyr.s.yakymiv@lpnu.ua).

ABSTRACT The article describes a method of image generation with Artificial Intelligence services using text abstraction retrieved using Artificial Intelligence services Dall-e, MidJourney and Stable Diffusion, that works with natural language. The implementation of the new approach gives a significant gain in image quality and consistency with analysed text. The methodology is based on using neural network API service instead of commonly used natural language algorithms to extract keywords or sentences. Proposed evaluation is applied to the generated images. An analysis of evaluation options is carried out depending on algorithm and Artificial Intelligence service, based on the tested book, length of result abstract and number of errors for each type. The evaluation results show that the new approach can provide better quality images that relate more with the text compared to natural language algorithms. For example, the average score of images generated by abstractions for GPT3 - 7.13 and GPT4 - 7.3, compared to natural language algorithms CO semantic - 5.43, TextRank - 4.98, TF-DF keywords - 4.74, WE spaCy - 3.04, WordNet - 4.34 for MidJourney generated images. Although results show most of the best results were generated for abstract with text length 20-40 words, meantime images generated for abstract with less or more words show much less consistency with text.

KEYWORDS artificial intelligence; computing; ai-generated images; text-to-image generation.

I. INTRODUCTION

ARTIFICIAL Intelligence (AI) tools are increasingly being introduced into everyday life and are being used in various fields. One such tool is a neural network model for generating images based on natural language. Such systems require input text that the model will convert into a graphic image. The input text can be any number of words that describe the desired result. Since such neural models are trained on image-text pairs, the selected input text is very important for understanding the task by the model at hand to reproduce the correct result [1]–[3].

Most of the input queries of such systems are a combination of words that are parameters, parts of objects that the user wants to see in the image [1]–[4]. But what if the task is to try to generate an image based on a whole text, for example, a page from a fiction book [5]? In this case, it would be more appropriate to use the main keywords or a summary of the page to identify the main objects or concepts that need to be depicted in the illustration.

This article will analyze the selection of keywords or abstraction from the text of a novel book to generate illustrations using AI tools. The study includes comparing a sample of the main key parameters from the text and the short abstract of the same text using prepared queries to AI services and comparing the generated illustrations based on the data obtained. Based on the results, it will be possible to conclude whether the description of the text of the book is better suited for AI image generation services, keywords or a summary of the text.

II. RELATED WORK

A. TEXT-TO-IMAGE GENERATION

The first publications on the topic of image generation by parameters began to appear in 2015-2016 [3, 6], although a concept for generating images by keyword phrase was presented in 2007 [7]. This paper describes a system for generating images using natural language text from children's books and scientific articles. The main difference between the approach presented in this article and the model described in [7] is that it uses a natural language model to select AI keywords instead of the standard keyword selection algorithm for text summarization [8, 9].

At the same time, in early studies, often the generated result was illustration with primitive objects of poor quality and often blurred [3, 6]. AI models available today can already produce photo-realistic quality images that are often difficult to distinguish from the real thing [10, 11].

B. PROMPT ANALYZING OF TEXT-TO-IMAGE GENERATION SYSTEMS

Many studies have been conducted on the analysis of input values for an AI model and output images, namely, which input text or object description can provide an image more relevant to the input query and of better quality [2, 12]–[16]. There have also been studies that aimed to investigate the reverse process: extracting parameters or object descriptions from AI-generated images to investigate the relationship of input parameters to the output to analyze which keywords or object descriptions can guarantee a better result [4, 17]. Meanwhile, this article is focused on analyzing the possibilities of automatic generation of an input query using an AI system.

C. KEYWORD EXTRACTION FROM NATURAL LANGUAGE TEXT

Research on keyword extraction from text analyses which algorithm should be used to get the best result. For example, in his article, the scholar Xiangdong You [18] uses the TextRank algorithm to extract key sentences and annotate them. In another article [19], the authors use the approach of removing individual stop words that are regularly matched and filtered by length, as well as the method of matching words through semantic relations to evaluate words. The results of this study showed that the use of the two methods simultaneously provides better results than the use of only one of the methods. Authors Enes Altuncu, Jason R.C. Nurse, Yang Xu, Jie Guo, and Shujun Li in their study [20] use an approach based on an improved level of semantic awareness with support for PoS tags and the use of entities from Wikipedia, which showed an increase in performance compared to other tested approaches.

A study on keyword extraction from multilingual texts [21] is worth mentioning, as it showed that the accuracy (matching with already selected keywords for texts) of the document frequency-inversion frequency (TF-IDF) algorithm is 80%, the graph-based algorithm is 60.65%, and the improved proposed algorithm is 91.3%. In the same article, the LTFIDF_POS algorithm based on a sliding window for the task of keyword extraction of short unlabeled news texts is proposed, which has shown its effectiveness because it fully takes into account unknown words and information about the distribution of words in the text.

III. METHODOLOGY

In this section, we will analyze the performance of AI systems for generating images from text using keywords or short descriptions from pages of classic literature obtained with the help of some general algorithms and the proposed methodology of using AI, and compare the results. For this purpose, a list of 20 books was prepared by different authors:

- The Count of Monte Cristo by Alexandre Dumas and Auguste Maquet;
- The Sign of the Four by Arthur Conan Doyle;
- Dracula by Bram Stoker;
- A Christmas Carol in Prose; Being a Ghost Story of Christmas by Charles Dickens;
- The Life and Adventures of Robinson Crusoe by Daniel Defoe;
- The Great Gatsby by F. Scott Fitzgerald;
- Metamorphosis by Franz Kafka;
- The Time Machine by H. G. Wells;
- At the mountains of madness by H. P. Lovecraft;
- Moby Dick; Or, The Whale by Herman Melville;
- The Call of the Wild by Jack London;
- Grimms' Fairy Tales by Jacob Grimm and Wilhelm Grimm;
- The Last of the Mohicans by James Fenimore Cooper;
- Pride and Prejudice by Jane Austen;
- Gulliver's Travels by Jonathan Swift;
- A Journey to the Centre of the Earth by Jules Verne;
- The Wonderful Wizard of Oz by L. Frank Baum;
- Alice's Adventures in Wonderland by Lewis Carroll;
- The Picture of Dorian Gray by Oscar Wilde;
- Treasure Island by Robert Louis Stevenson.

5 random pages were chosen by a general random algorithm from each book, so we received 100 test pages in total.

A. KEYWORDS AND TEXT EXTRACTION ALGORITHMS

The following methods were chosen for comparison:

- TF-IDF is a statistical indicator used to assess the importance of words in the context of a document that is part of a document collection or paragraph;
- WordEmbedding (WE) is a language modeling algorithm in natural language processing (NLP), in which words or phrases from a dictionary are converted into vectors of real numbers;
- TextRank is a graph-based text ranking model suitable for finding keywords in text [23];
- WordNet is an algorithm based on the semantic similarity of two words based on the WordNet database;

- Co-occurrence semantics is an algorithm with a random frequency of the ordered appearance of two related terms in the selected text;
- using the API of the AI system with a specific prepared query to describe the plot of the submitted text for the gpt-3.5-turbo model;
- using the AI system's API with a specific prepared request to describe the plot of the submitted text for the gpt-4-1106-preview model.

1) General keywords extraction algorithms

The text of books is not divided into pages but is a continuous text. The developed script analyses the number of paragraphs and selects one arbitrary, and generates a page based on the average number of letters per page - the page that will be used for further analysis (Fig. 1). The average number of letters per page is assumed to be 3000 letters [22].

The next step is to use algorithms to select keywords. The algorithms used are based on the POS tagging algorithm [19].

In these algorithms, except for TF-IDF for selecting key sentences, the same initial steps are used to select candidate keywords (Fig. 2). In the TF-IDF algorithm for keyword selection, there is no step to filter out unique keywords. This algorithm is based on the frequency of word usage in the text, so filtering unique words is completely incompatible with this algorithm.

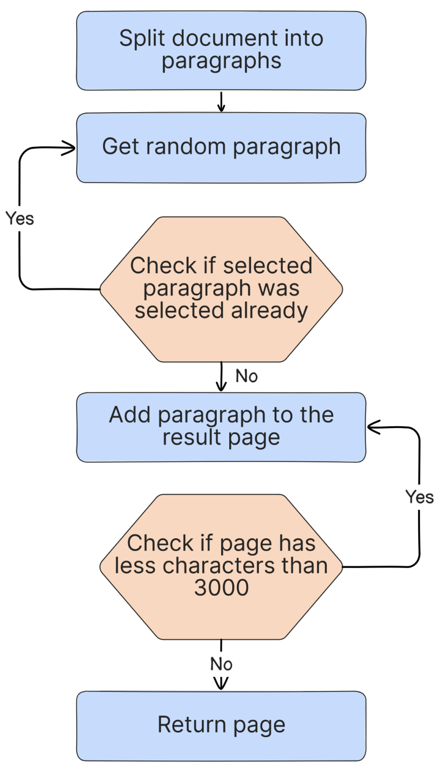

Figure 1. Algorithm to select random page from book

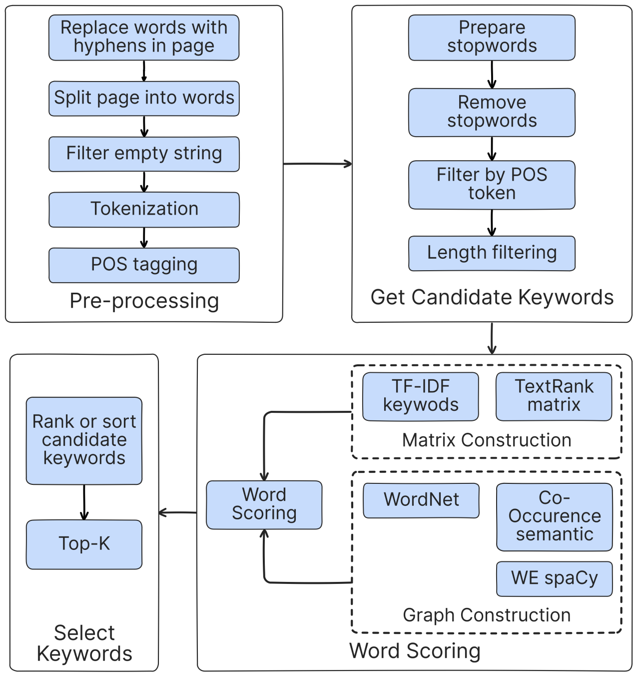

Figure 2. Diagram of TF-IDF, TextRank, WordNet, WE spaCy and Co-Occurence Semantic algorithms for keyword extraction

All algorithms for obtaining keywords start with splitting the text into separate words, removing empty values from the resulting set of words, tokenization, and POS tagging, which defines the part of speech of the word.

The next part of the algorithm is to obtain the candidate words. In this case, stop words from the prepared NLTK database are removed from the set of words obtained from the previous part. Stop words should not be removed before POS tagging, as they provide important information about the sentence structure and, as a result, the part of speech of the word in the sentence. The algorithm then filters the words by POS tags, filters out unique words (for all algorithms except TF-IDF), and filters out short words. The next steps for evaluating and selecting keywords from the resulting matrix or graph of word vectors differ in each algorithm.

2) CO-Occurrence keyword extraction algorithm

For this study, we used the approach proposed in [19], in which the algorithm is based on the joint appearance of words and semantic relations between them.

At the beginning of the algorithm, a graph of the occurrence of words G in the text is built:

$$G = (V, E) \tag{1}$$

where V - a set of words in a text, and, E - a set of dots that indicate connections between word occurrences.

The next step is to determine the semantic compatibility between words. It is calculated using a word embedding model, such as cosine similarity between word vectors:

$$sim(w_i, w_j) = cosine\_similarity(v_i, v_j) \tag{2}$$

where v_i, v_j - vectors of words w_i, w_j respectively.

Next, a score is determined for each word based on the proportions of its common usage in the G graphs:

$$score(w_i) = \sum_{w_j \in N(w_i)} w_i \tag{3}$$

where $N(w_i)$ - is the set of neighbors of w_i in the graph G.

The next step is to combine the scores of the two graphs: the graph of the joint occurrence of words and the graph of the importance of words in the document; which can be expressed as:

$$score_{com} = \lambda \cdot score_{co} + (1 - \lambda) \cdot score_{sem} \tag{4}$$

where λ - is a number in the range between 0 and 1. For this study, the value of 0.5 was chosen based on the study [19].

The last step is to select the keywords with the highest weight:

$$keywords = TopK(score_{com}) \tag{5}$$

where $TopK()$ - is a function that selects the words with the highest score.

3) WordNet keyword extraction algorithm

One of the models for calculating semantic compatibility graph points that will be used for this study is the model based on the WordNet database, namely, using the shortest path between words in WordNet to determine the distance between words in the graph.

First, we define and calculate the semantic compatibility matrix between words:

$$sim\,[i, j] = path\_similary(w_i, w_j) \tag{6}$$

where $path\_similarity()$ - is a function from the NLTK library [24] that determines the semantic distance between words.

The result is the following matrix:

$$SM = \begin{bmatrix} sim(w_1, w_1) & \cdots & sim(w_1, w_n) \\ \vdots & \ddots & \vdots \\ sim(w_n, w_1) & \cdots & sim(w_n, w_n) \end{bmatrix} \tag{7}$$

Next, we create a graph where the nodes represent potential keywords. We add edges between the nodes based on the semantic similarity threshold using equation (1).

The next step is to calculate the score of each node using the PageRank algorithm [25], which will result in a dictionary where each keyword will have its own score:

$$word\_score(w_i) = PageRank(w_i) \tag{8}$$

Next, we calculate the score of each potential keyword based on the graph structure and node scores. For each keyword w_i we calculate the sum of the scores of the word's neighboring nodes using the rejection factor:

$$score(w_i) = (1 - d) + d \cdot \sum_{j \in N(w_i)} word\_score(w_j) \tag{9}$$

where d - is the rejection factor.

4) WeSpacy keyword extraction algorithm

Another model for calculating the points of the semantic compatibility graph is the WeSpacy database model, which uses the cosine distance between the word vectors of the SciPy library [26].

The difference from the algorithm with the WordNet model is that instead of equation (6), the following formula is used to calculate the semantic distance:

$$sim(w, v) = \frac{w \cdot v}{\|w\| \cdot \|v\|} \tag{10}$$

where w and v - are the vectors of the words w and v respectively.

5) TextRank keyword extraction algorithm

Another algorithm used in the study is an algorithm based on the TextRank model [23].

For this algorithm, it is also necessary to build a graph of the occurrence of words G in the text based on expression (1). The model then performs the following calculations to obtain estimates of the graph nodes:

$$R(V_i) = (1 - d) + d \cdot \sum_{j : V_j \to V_i} \frac{w_{ji}}{\sum_{k : V_j} w_{jk}} R(V_j) \tag{11}$$

where w_{ji} - is the evaluation of the edge from node V_j to the current node V_i.

6) TF-IDF text keyword extraction algorithm

Also, for comparison, we chose the TF-IDF algorithm, which is used to assess the importance of words in the text. The word importance score is proportional to the number of occurrences of the word in the text and inversely proportional to the frequency of the word in other parts of the text:

$$tf(w, d) = \frac{freq(w, d)}{\sum_{\acute{w} \in d} freq(\acute{w}, d)} \tag{12}$$

where $freq(w, d)$ - is frequency of occurrence of the word w in the text d, and $\sum_{\acute{w} \in d} freq(\acute{w}, d)$ - total number of words $\acute{w}$ in the document d.

$$idf(w, D) = log(\frac{N}{|\{d \in D : w \in d\}|}) \tag{13}$$

where N - is the total number of sections in the document $N = |D|$, $|\{d \in D : w \in d\}|$ - number of text sections in which the word appears w.

And the final word importance score is calculated:

$$tfidtf(w, d, D) = tf(w, d) \cdot idf(w, D) \tag{14}$$

7) TF-IDF text summary algorithm

Since the proposed method using AI services returns the result in the form of sentences rather than a set of words, the study also uses an algorithm based on TF-IDF, the result of which is the selected sentences for summarizing the text (Fig. 3).

In terms of calculating the importance of words, this algorithm uses the same equations (12), (13) and (14) of the TF-IDF score, only in this case the calculations are performed for sentences rather than for individual words.

8) AI text summary extraction

For the analysis, we used the API of the OpenAI AI system with a comparison of the gpt-3.5-turbo and gpt-4-1106-preview models, since at the time of writing and conducting the study, this is the only system with an open API to work with AI models. The AI system models were selected among the available and most productive and updated OpenAI models. The API accepts commands to be executed on the proposed text. For the study, after using and testing a list of various commands, the following command was selected "*You will be provided with a block of text, and your task is to return a short one sentence of what is happening in provided text*", because the result met the expected requirements, as the result was a brief generalized description of the events in the text.

B. IMAGE GENERATION

The next stage of the study is to generate images from the received list of keywords or short descriptions. At this stage, two AI systems that have an open API are used and compared:

– Dall-e;
– Stable Diffusion;
– MidJourney.

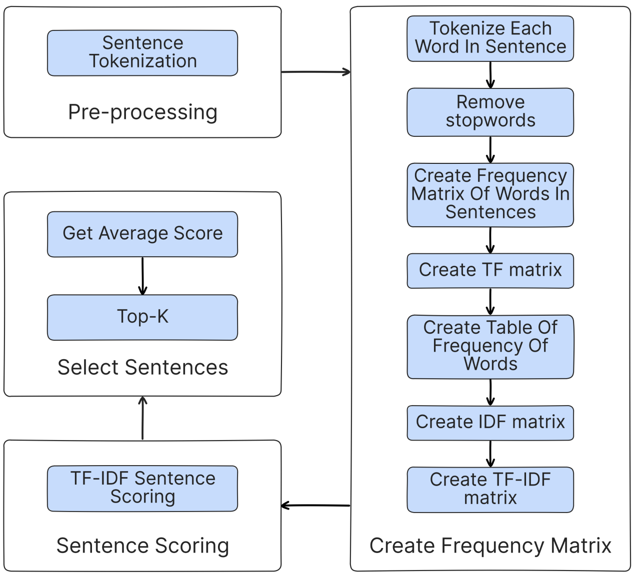

Figure 3. Diagram of TF-IDF algorithm to extract summarization

The result of using the API is a URL link to the image, which is stored along with the keywords and the text to which they refer.

C. IMAGE AND KEYWORD SCORING

The following scale was proposed to evaluate the received keywords or text (Table.1):

Table 1. Scale for assessing the relevance of keywords or descriptions to the text

Text mark	Number mark	Description
Not Relevant	1-2	The provided keywords or summary have minimal to no connection with the content of the text. There's little or no overlap in terms of theme or information.
Slightly Relevant	3-4	There are some minor connections between the keywords or summary and the text, but the correspondence is weak. The keywords may touch on peripheral aspects.
Somewhat Relevant	5-6	There is a moderate level of relevance. The keywords or summary capture some aspects of the text, but there are notable gaps or differences.
Moderately Relevant	7-8	The keywords or summary align well with the text, capturing the main ideas and themes. However, there may be some nuances or details that are not perfectly reflected.
Highly Relevant	9	The keywords or summary closely match the content of the text. They effectively encapsulate the main points and themes, with only minor variations.
Perfectly Relevant	10	The keywords or summary perfectly describe the text. Every important detail, theme, and nuance is accurately reflected.

The same scale was used to assess the relevance of images to text.

IV. RESULTS AND DISCUSSIONS

The key feature of this study, as well as the evaluation and analysis of the obtained results of the algorithms for keyword selection, is that this study analyses not just how well the obtained keywords correspond to the analyzed text, but how well they are suitable for further use as a query for image generation.

The evaluation was carried out only by the author of the study, so a certain subjective discrepancy in the obtained estimates should be taken into account.

A. KEYWORDS AND SUMMARY TEXT SCORE RESULTS ANALYSIS

From the data obtained, for each individual book, we calculated the average value for each algorithm separately

Table 2. Average score of algorithms and AI keywords and summary results

	CO Semantic	GPT3 Summary	GPT3 Keywords	GPT4 Summary	GPT4 Keywords	TextRank	TF-IDF Keywords	TF-IDF Summary	WE spaCy	WordNet
A. Dumas and A. Maquet - The Count of Monte Cristo	4.6	9.6	5.4	9.4	5.8	5.8	5.6	5.4	3.8	4
A. C. Doyle - The Sign of the Four	5.6	9.2	6	8.6	4	3	4.6	6.2	3	4.2
B. Stoker - Dracula	5.8	9.4	5.6	10	5.2	5.8	5.4	5.4	3	3.8
C. Dickens - A Christmas Carol in Prose	6.4	8.8	7.2	10	6.2	5.6	6.2	6.2	3.6	6.2
D. Defoe - The Life and Adventures of Robinson Crusoe	6.2	9	7	9.2	6.6	4.4	4	4	3.2	4.6
F. S. Fitzgerald - The Great Gatsby	6	8.4	6.4	9.2	6.8	4.8	4.4	4.4	4.2	3
F. Kafka - Metamorphosis	6.2	9.2	6	9.6	5.6	6	7	6.2	3	4.6
H. G. Wells - The Time Machine	5.4	8.2	6	8.8	5.2	5.8	5.8	4.8	3.6	4.4
H. P. Lovecraft - At the mountains of madness	6.4	9	6.6	9	6.8	5.8	5	5.4	4	4.6
H. Melville - Moby Dick; Or, The Whale	6.6	9.2	8	10	8	6.4	4.4	5.2	2.4	5.6
J. London - The Call of the Wild	7	9.4	8	9.8	8	6.8	4.6	6.2	3	5.8
J. Grimm and W. Grimm - Grimms' Fairy Tales	6.8	9.2	8.2	9.2	8.2	6.8	5.6	5.6	2.6	6.4
J. F. Cooper - The Last of the Mohicans	6	8.8	6.6	9.4	7.8	6.2	4.8	5	3.8	6
J. Austen - Pride and Prejudice	7.4	9	7.2	9.8	7.2	6.6	6	6	4.8	4.8
J. Swift - Gullivers Travels	6.8	8.6	6.8	8.8	6.4	6	6	5.8	4	6.6
J. Verne - A Journey to the Centre of the Earth	7.2	9.4	7.6	9.2	7.4	7	5	6.6	5.6	6
L. F. Baum - The Wonderful Wizard of Oz	8	9.4	7.8	10	7.8	6.2	7.2	5.8	6	5.6
L. Carroll - Alice's Adventures in Wonderland	7.6	9.4	7.6	9.6	7.8	6.4	5.4	4.8	6	6.6
O. Wilde - The Picture of Dorian Gray	6.2	9.2	6.2	9.8	6.6	6.2	6.4	6	5	5.4
R. L. Stevenson - Treasure Island	7.2	8.2	7.6	9.2	7.6	6.2	5.8	5.6	4.8	5.8
AVERAGE	**6.47**	**9.03**	**6.89**	**9.43**	**6.75**	**5.89**	**5.46**	**5.53**	**3.97**	**5.2**

for the results of evaluating the obtained keywords or text descriptions, as well as for the images generated by the AI. The average values of keyword or text description evaluation for the tested books are presented in Table. 2.

Based on the results presented in Table. 2 and the diagram (Fig. 4), one can immediately notice the difference between the scores for text description generated by AI and other methods. It can also be seen that the keywords generated by AI have similar results as the keywords obtained using algorithms.

However, in a few individual cases, AI-generated words have slightly better results. The text description has both more words and more specific words, such as adjectives, which actually gives this approach better and more expected results. However, this study is aimed at analyzing which approach would be better as an input text for an AI image generation service. Therefore, these results are rather superficial, as there is a possibility that the AI service will be able to determine by itself whether the keywords belong to a certain novel, or a place or character from a certain book, and generate the corresponding image. While for a more accurate description, you can ignore part of the description and focus only on some details of the description.

If we consider only the results of the NLP algorithms, the CO Occurrence Semantic algorithm obtained the best results, TextRank slightly worse, while WE spaCy was the worst. It is also worth noting that the TF-IDF algorithm for keyword extraction and the TF-IDF algorithm for key sentence extraction showed quite similar results.

B. IMAGE GENERATION ERRORS ANALYSIS

At the outset of this analysis, it is worth noting that generating AI keywords or summaries for the description of the analyzed page contains problems that rarely occurred when using text analysis algorithms.

In short, AI services have a list of unacceptable words that are used to check the queries sent by users. Examples of such unacceptable words include racist, politically incorrect, or swear words. Most of these errors were received when using the DALL-e service (Fig. 5). DALL-e also refused to process requests that contained descriptions of bloody scenes. For example, most of these cases occurred when generating images for randomly selected pages from "In the Mountains of Madness" by H.P. Lovecraft.

The rest of the services did not indicate the cause of the errors. In particular, in some cases, Stable Diffusion resulted in a solid black image, and the API response did not contain details of the errors that occurred. At the same time, the API is configured to repeat the request in case of an error in order to minimise the risk of possible Internet outages, internal system errors, or false alarms (in the case of Dall-e's input word filters).

The largest total number of errors in image generation was obtained when using the Stable Diffusion service.

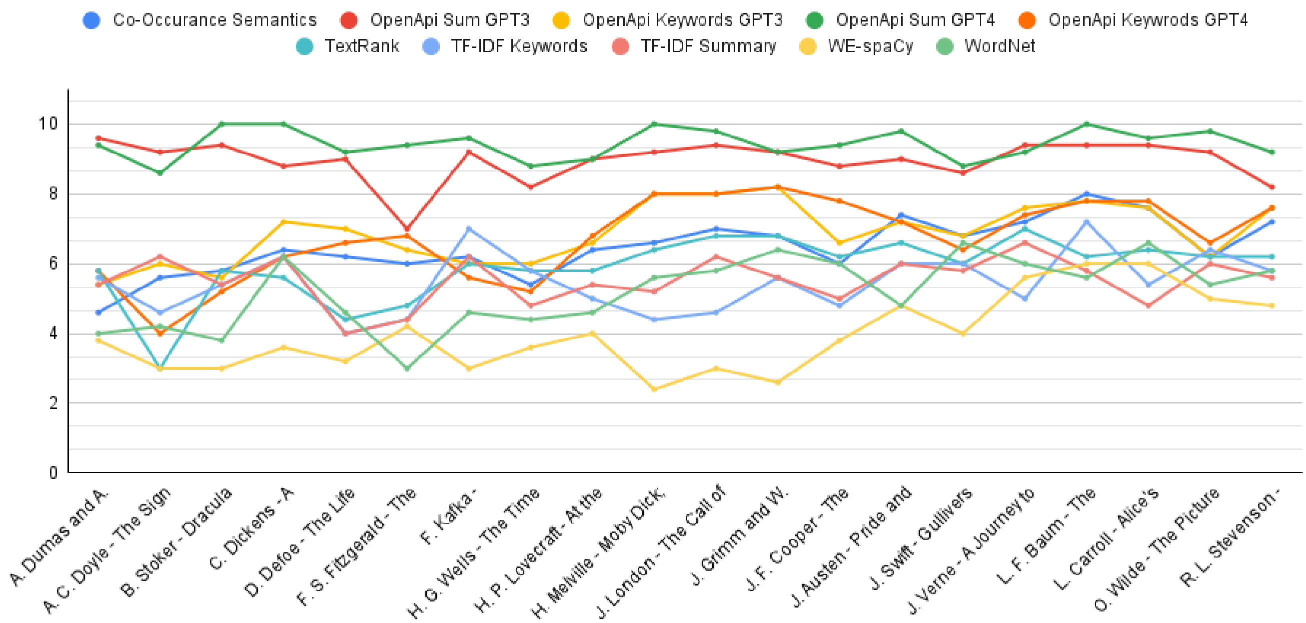

Figure 4. Diagram of the average score for the result keywords and text summary by algorithm for the tested books

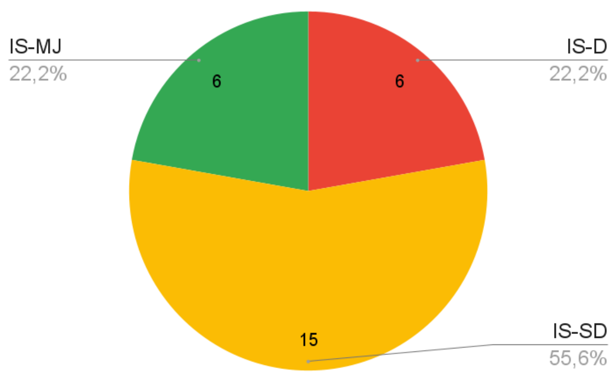

Figure 5. Diagram of the number of errors when using keywords or abstraction depending on the AI image generation service

The summary statistics of errors obtained during image generation by Dall-e, Stable Diffusion, and MidJourney services are presented in Table. 3.

C. IMAGE GENERATION SCORE RESULTS ANALYSIS

Figs. 6-15 show examples of generated images for each service used (Dall-e, MidJourney, Stable Diffusion), which, according to the evaluation results, best match the selected text. That's the randomly selected page from the book "Alice's Adventures in Wonderland", by L. Carrol. Tested page describes Alice having a conversation with a caterpillar, which is smoking hookah sitting on the mushroom.

It is important to note that the illustrations generated by MidJourney contain four images at once. This is a feature of the service's generation. The API allows you to select and save each image separately. However, since it is difficult to identify which image best matches the test page of text in automatic mode, the score for this illustration is the average score for all four images, as well as the saved image result having all generated 4 images.

As a matter of fact, the list of keywords retrieved from using CO-Occurrence Semantic algorithm contain some of the main keywords that can describe the analysed page, but also contain some others that have no specific relation to the text.

Table 3. Number of errors for tested algorithms

	IS-D	IS-SD	IS-MJ
CO Semantic	0	2	0
OpenApi Summary GPT3	1	1	0
OpenApi Summary GPT4	2	0	0
OpenApi Keywords GPT3	0	2	1
OpenApi Keywords GPT4	1	0	1
TextRank	0	4	0
TF-IDF Keywods	0	1	2
TF-IDF Keywods	1	2	0
WE spaCy	1	1	2
WordNet	0	2	0

For instance, the image generated by Dall-e (Fig. 6) contains all described characters mentioned in the text, but overall, poorly reflects the tested page. As to, the image generated by MidJourney, doesn't contain Alice at all.

Figure 6. Examples of generated images by keywords of the Co-Occurrence semantic algorithm ("alice", "caterpillar", "bit", "side", "last", "mushroom", "mouth", "time", "hookah", "minute")

Figure 7. Examples of generated images by GPT3 extracted keywords ("Alice", "Caterpillar", "Size", "Mushroom", "Height", "Hookah", "Change", "Shoulders", "Neck", "Green")

Finally, the StableDiffusion image doesn't contain Alice and overall has very moderate relevance.

Results of using OpenAI GPT API provides a list of keywords very similar to the list of keywords retrieved from using CO-Occurrence Semantic algorithm results. In this case, generated images (Fig. 7 and Fig. 8) also have the same composition and details, except Dall-e images have much higher relevance, but still not fully accurate.

Furthermore, the abstraction generated by OpenAI GPT API provides a general description of the analysed text, clarifying the main events, but seems missing some of the important details of the current scene (i.e. caterpillar doesn't smoke hookah). Therefore, images generated by these abstractions (Fig. 9 and Fig. 10) have similar high relevance for the text, but are still missing some of the details (even though they fully represent abstraction text). In addition, Stable Diffusion images have now higher relevance compared to previous images.

The list of keywords retrieved from the TextRank algorithm looks a bit similar to previous lists of keywords with moderate relevance to the text, but is missing a bit more important detail. As a result, generated images (Fig. 11) have similarity with images generated by GPT abstraction with difference in a few missing details.

Moreover, the next list of keywords obtained from the TF-IDF algorithm is now missing “Alice“ as a keyword and as a main character of the text. Accordingly, images (Fig. 12) have even smaller relevance, with losing main perception of the text in the StableDiffusion image.

TF-IDF algorithm provides a sentence with main focus on Alice but missing even more important details of the text. Despite the “mushroom” keyword not mentioned in the result sentence, MidJourney still added it to the image (Fig. 13).

WeSpacy list of keywords has the lowest relevance score among all others. It has some minor connection to the text, but overall lost the essence of the text, as well as image (Fig. 14) slightly touches some of the aspects of the text.

Figure 8. Examples of generated images by GPT4 extracted keywords ("Alice", "Caterpillar", "Size", "Mushroom", "Growing", "Shrinking", "Temper", "Hookah", "Contradiction", "Transformation")

Figure 9. Examples of generated images by GPT3 generated abstraction (“Alice is conversing with a Caterpillar about her desire to change her size and then later experiences a sudden change in her body.”)

Figure 10. Examples of generated images by GPT4 generated abstraction (“Alice is having a confusing and frustrating conversation with a Caterpillar about her size and is trying to use pieces of a mushroom to alter her height, which leads to her experiencing a sudden and alarming transformation.”)

Figure 11. Examples of generated images by TextRank algorithm keywords ("alice", "hand", "little", "mushroom", "side", "moment", "inches", "time", "hookah", "mouth")

Figure 12. Examples of generated images by TF-IDF algorithm keywords ("caterpillar", "inches", "hookah", "minute", "mushroom", "moment", "shoulders")

Figure 13. Examples of generated images by TF-IDF algorithm sentences ("This time Alice waited patiently until it chose to speak again. And where have my shoulders got to?")

Figure 14. Examples of generated images by WE-spaCy algorithm keywords ("life", "temper", "sir", "mind", "high", "hookah", "minute", "aloud", "violent blow", "room")

Figure 15. Examples of generated images by WordNet algorithm keywords ("side", "sides", "caterpillar", "inches", "time", "hookah", "mouth", "minute", "mushroom", "moment")

Finally, keywords obtained from the WordNet algorithm are very similar to the TF-IDF list of keywords. Therefore, images (Fig. 15) also look the same, with less than moderate relevance to the text.

Table. 4 shows the average results for all tested pages of evaluating the images generated by the Dall-e service for each approach and book. The diagram (Fig. 16) represents this data shows that it is difficult to single out the approach that provided the best scores of the most appropriate image.

Only a few of the results have a score value greater than 8. Although the chart shows that the best scores are for images generated using data from ChatGPT versions 3 and 4. However, even among them, for some books the score falls below the average and below most of the other approaches used.

Table. 5 contains the average evaluation results for images generated by the Stable Diffusion service. In turn, the following diagram (Fig. 17) shows a similar trend as in the results of evaluating Dall-e images, namely, the vast majority of the best results were obtained using data from the ChatGPT service. The diagram also demonstrates that the overall score of StableDiffusion images is slightly lower compared to Dall-e images.

The results of MidJourney image evaluation are presented in Table. 6. The diagram (Fig. 18) shows generally better results compared to the Stable Diffusion data, but at the same level as the Dall-e data. The image evaluation values obtained from the ChatGPT short description are higher than those obtained from the keywords and are also more uniform. This means that this approach demonstrates more stable results, regardless of the work.

Analysing the data on the dependence of the average image evaluation of different services on the applied algorithm shown in the diagram (Fig. 19), we can see that for all the services used, the sequence of algorithmic dependence is almost the same. Dall-e and MidJoruney have almost identical results, differing by 0.2-1 points: MidJourney has better results when using OpenAPI GPT3 and GPT4 for

Table 4. Average score of images generated by Dall-e per algorithm

	CO Semantic	GPT3 Summary	GPT3 Keywords	GPT4 Summary	GPT4 Keywords	TextRank	TF-IDF Keywords	TF-IDF Summary	WE spaCy	WordNet
A. Dumas and A. Maquet - The Count of Monte Cristo	5.4	7.6	5	7.8	5.4	5.2	6.2	4.6	2.2	3.8
A. C. Doyle - The Sign of the Four	5.2	6	6.2	4.8	3.6	3.4	3.8	5.4	1.8	2.8
B. Stoker - Dracula	5.6	6.2	5	6.6	5	5.4	4.8	3.2	2.6	3.4
C. Dickens - A Christmas Carol in Prose	6.8	8	7.4	7	5.8	5.4	5.2	6	3	6
D. Defoe - The Life and Adventures of Robinson Crusoe	5.6	6.4	5.8	6.2	6	4.8	3.8	3.4	2.6	4
F. S. Fitzgerald - The Great Gatsby	5.6	5.6	6	5.2	6.6	4.4	4.6	2.2	2.6	1.6
F. Kafka - Metamorphosis	4.6	5.8	3.8	5.8	3.4	5.4	4	4	1.6	3.8
H. G. Wells - The Time Machine	5.2	6.2	5.2	5.8	4.4	4.6	4.8	3.2	3.2	3.6
H. P. Lovecraft - At the mountains of madness	6	6.75	6.6	7	6.6	5.8	4.4	4.2	3.4	4
H. Melville - Moby Dick; Or, The Whale	6.4	7.6	7.6	7.8	7.2	6.6	4.4	5.2	1.75	5
J. London - The Call of the Wild	5	7.4	7.2	7.6	7.6	3.8	3	5.8	1.4	4.8
J. Grimm and W. Grimm - Grimms' Fairy Tales	6.6	6.6	7	6.2	7.6	6.6	4.6	4.6	1.2	5.6
J. F. Cooper - The Last of the Mohicans	4.8	7	6.2	7.4	7	5.2	4.4	5	2.6	4.2
J. Austen - Pride and Prejudice	7.4	7.6	6	8.4	6	6.4	5.2	6.4	2.2	2.6
J. Swift - Gullivers Travels	6.2	6.6	5.6	7.24	5.8	4.6	5.2	4.6	2.6	5
J. Verne - A Journey to the Centre of the Earth	6.6	7	6.8	7	7.4	5.8	5.4	5.4	4.6	5.6
L. F. Baum - The Wonderful Wizard of Oz	7.8	7.6	6.8	7.6	7	6.2	7.2	5.2	5.6	5.6
L. Carroll - Alice's Adventures in Wonderland	6.4	8.2	7.2	7.2	7.4	6.2	5.2	4.6	5.8	6.2
O. Wilde - The Picture of Dorian Gray	5	7.6	5.6	7.4	6.2	5.8	6.6	6	4.6	3.8
R. L. Stevenson - Treasure Island	6.8	7.6	7.2	7.2	7	6.2	5.6	5.4	4	5.2
AVERAGE	**5.95**	**6.97**	**6.21**	**6.86**	**6.15**	**5.39**	**4.92**	**4.72**	**2.97**	**4.33**

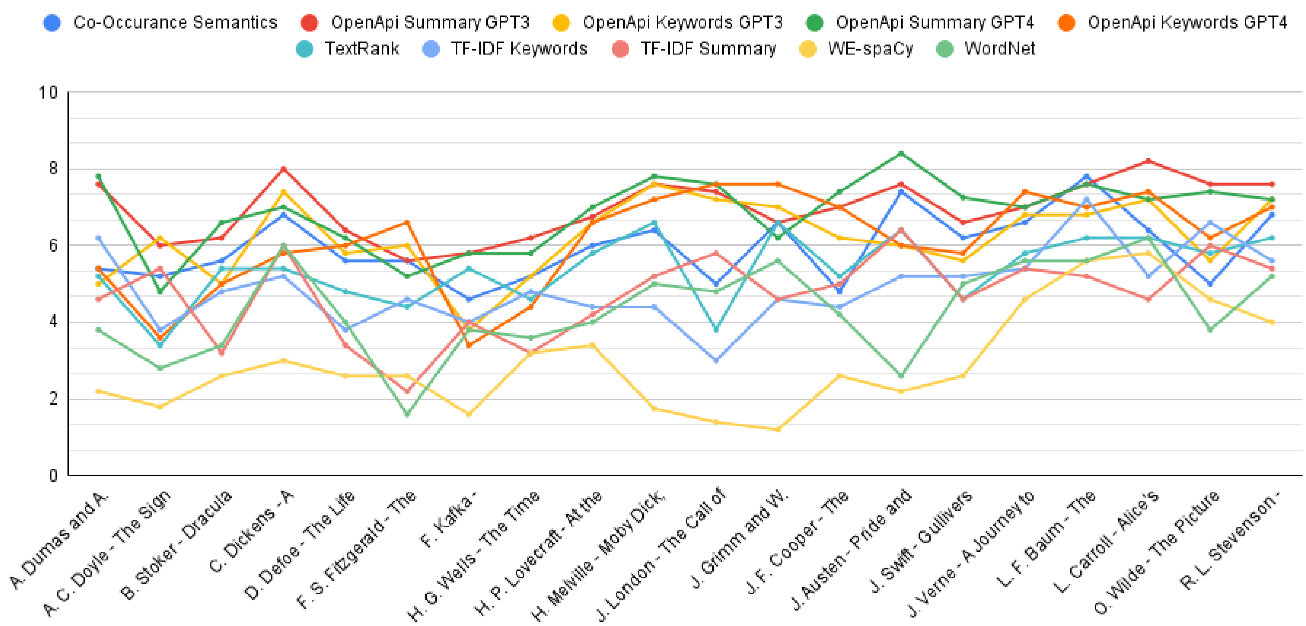

Figure 16. Diagram of the average score of the Dall-e service generated images for the tested books

Table 5. Average score of images generated by Stable Diffusion per algorithm

	CO Semantic	GPT3 Summary	GPT3 Keywords	GPT4 Summary	GPT4 Keywords	TextRank	TF-IDF Keywords	TF-IDF Summary	WE spaCy	WordNet
A. Dumas and A. Maquet - The Count of Monte Cristo	3.8	7	4	7.2	4	3.6	4.75	3.8	2.6	2.75
A. C. Doyle - The Sign of the Four	2.4	4.6	4.8	4.4	2.4	1.8	2.8	3	1.2	1.8
B. Stoker - Dracula	4.6	4.6	4.4	6.6	4.6	4.8	4	2.6	2	2.4
C. Dickens - A Christmas Carol in Prose	4.4	5.8	5.4	6.4	5	4.4	4.6	4.6	2.2	5.4
D. Defoe - The Life and Adventures of Robinson Crusoe	4.75	5	4.2	5	5.4	3.4	3.8	2.4	1.6	4
F. S. Fitzgerald - The Great Gatsby	5	6.2	5.8	5.6	5.4	3.6	4.2	1.8	2	1.25
F. Kafka - Metamorphosis	2.8	4.25	3.2	4.6	2.8	4.6	3.8	3	1.8	3.2
H. G. Wells - The Time Machine	5	6.4	4.8	6.2	4.8	4.6	4.8	2.5	2.6	3.4
H. P. Lovecraft - At the mountains of madness	4.2	6	5	5.4	5.4	4.2	2.75	2.4	2.8	3.6
H. Melville - Moby Dick; Or, The Whale	4.8	6.4	6.2	6.8	5.6	6.33	2	4	1.2	5.25
J. London - The Call of the Wild	6	7.6	7	7.8	6.6	4.4	3	3.6	1.25	4.2
J. Grimm and W. Grimm - Grimms' Fairy Tales	5.4	7.2	6.2	7	6	5.6	4.6	4.4	1.2	5.4
J. F. Cooper - The Last of the Mohicans	4.2	6.4	6.2	6	6.8	4.25	4	4.4	2.8	4.2
J. Austen - Pride and Prejudice	6.2	6.8	6.4	7.4	6.8	6	5	5	2.8	4.8
J. Swift - Gullivers Travels	5.2	5.4	6	5.8	5.2	4.8	5.4	4	2.4	4.8
J. Verne - A Journey to the Centre of the Earth	5.6	5.8	5.8	6.8	5.8	5.8	4.8	5.4	3.4	5
L. F. Baum - The Wonderful Wizard of Oz	5.8	6.2	6.2	6	6.4	5	5.6	4.4	4.6	4.8
L. Carroll - Alice's Adventures in Wonderland	5.6	7	5.5	6.8	5.8	5.6	4	4.6	4.4	5.2
O. Wilde - The Picture of Dorian Gray	5.4	6.8	5.4	6.2	5.8	5.2	5.6	5.6	3.6	4.8
R. L. Stevenson - Treasure Island	5.6	6.6	6.2	6.4	6	4.2	4.2	4.6	3.6	4
AVERAGE	**4.84**	**6.1**	**5.44**	**6.22**	**5.33**	**4.61**	**4.19**	**3.81**	**2.5**	**4.01**

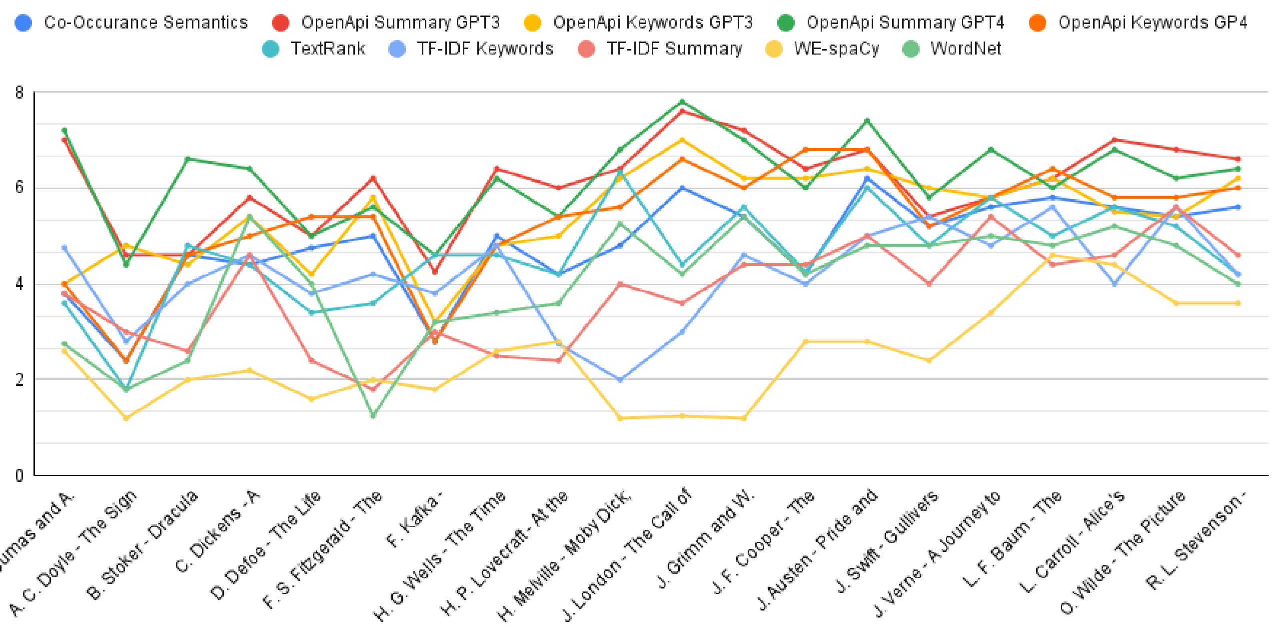

Figure 17. Diagram of the average score of the Stable Diffusion service generated images for the tested books

Table 6. Average score of images generated by MidJourney per algorithm

	CO Semantic	GPT3 Summary	GPT3 Keywords	GPT4 Summary	GPT4 Keywords	TextRank	TF-IDF Keywords	TF-IDF Summary	WE spaCy	WordNet
A. Dumas and A. Maquet - The Count of Monte Cristo	4	9	4	8.4	4.6	4.2	5	5	2.2	3.6
A. C. Doyle - The Sign of the Four	4.2	6.6	5.6	5.6	3.6	2.4	3.2	5	2.2	3.2
B. Stoker - Dracula	5.6	6	4.6	7.2	4.8	5.4	5	3.6	2.6	3.4
C. Dickens - A Christmas Carol in Prose	5.8	7.6	6.6	8.6	5.4	4.8	6.4	6	2.8	4.8
D. Defoe - The Life and Adventures of Robinson Crusoe	5.2	6.8	6.6	6.6	6.4	3.6	4.2	3.6	2.2	3.6
F. S. Fitzgerald - The Great Gatsby	5.6	6.8	6	6.6	7	4.2	4.8	2.2	2.4	1.4
F. Kafka - Metamorphosis	3.2	5.6	3.4	6.4	3.2	4.8	4.4	5	1.8	4.2
H. G. Wells - The Time Machine	5	6.4	5.4	5.8	4.4	5	4.6	3.2	3.2	3.6
H. P. Lovecraft - At the mountains of madness	5.4	7.2	6.25	7.2	6.25	4.8	4	4.4	3.4	3.8
H. Melville - Moby Dick; Or, The Whale	5.6	7.4	6.6	7	6.8	6.4	2.8	6	2	5
J. London - The Call of the Wild	5	7.4	7.8	8.2	7.4	4.4	3.25	3.8	1.4	4.4
J. Grimm and W. Grimm - Grimms' Fairy Tales	5.8	7.6	6.8	7.4	7.6	6	5	4.6	1.2	5.6
J. F. Cooper - The Last of the Mohicans	4.6	7.4	6.6	7.6	7.2	4.6	4.8	5.4	3.25	5
J. Austen - Pride and Prejudice	6.2	7.6	6.8	8	6.8	6.4	5.4	6.2	3.6	3.2
J. Swift - Gullivers Travels	5.6	6	5.8	6.8	5.8	5.2	5.4	4.8	3	5.8
J. Verne - A Journey to the Centre of the Earth	6.6	8	6.8	7.4	6.8	6	5.2	6.6	4	5.4
L. F. Baum - The Wonderful Wizard of Oz	7	7	6.6	7.4	6.4	5	5.8	5.6	5.4	5.2
L. Carroll - Alice's Adventures in Wonderland	6.4	7.4	6.2	7.6	6.6	6	4.4	4.8	5.6	5.6
O. Wilde - The Picture of Dorian Gray	5.8	7.6	5.2	8.2	5.8	5	5.8	5.6	4.6	4.2
R. L. Stevenson - Treasure Island	6	7.2	7	8	6	5.6	5.4	4.4	4	5.8
AVERAGE	**5.43**	**7.13**	**6.03**	**7.3**	**5.94**	**4.98**	**4.74**	**4.74**	**3.04**	**4.34**

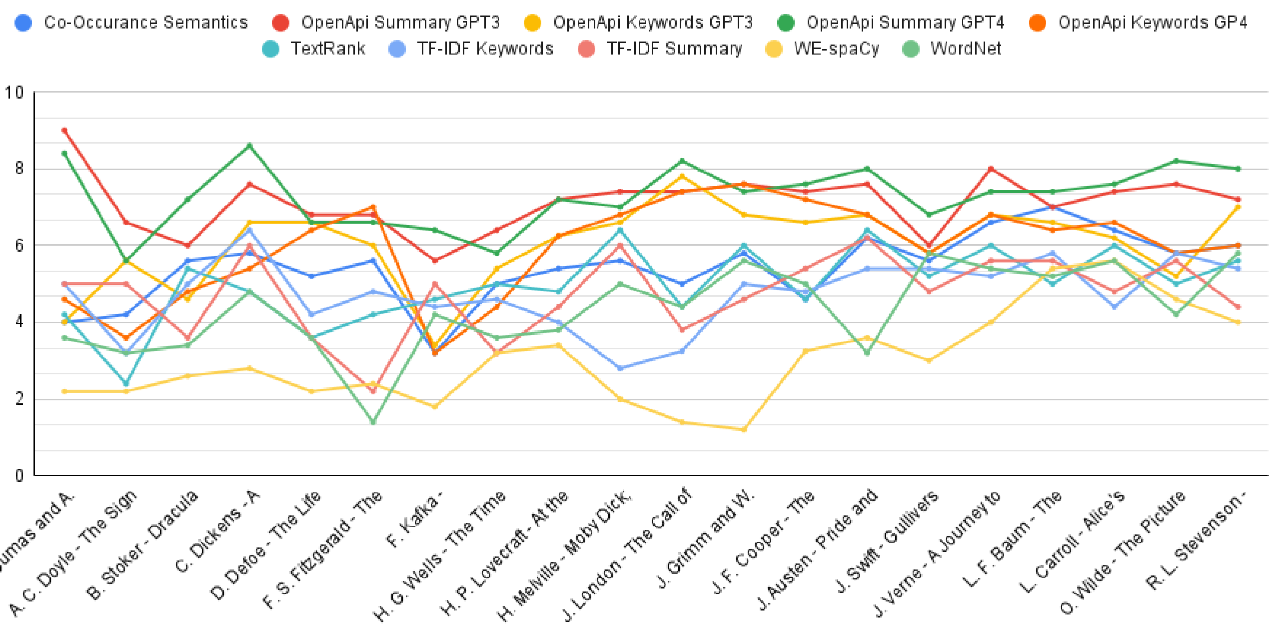

Figure 18. Diagram of the average score of the MidJourney service generated images for the tested books

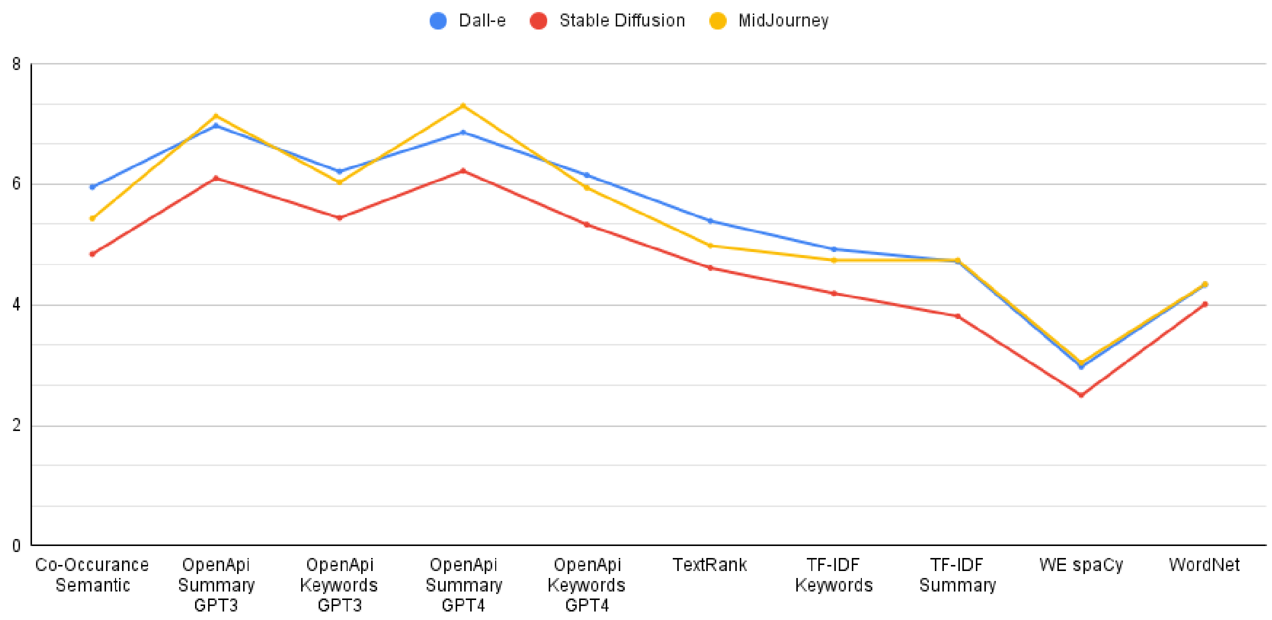

Figure 19. Diagram of average score between Dall-e, Stable Diffusion and MidJourney by algorithm

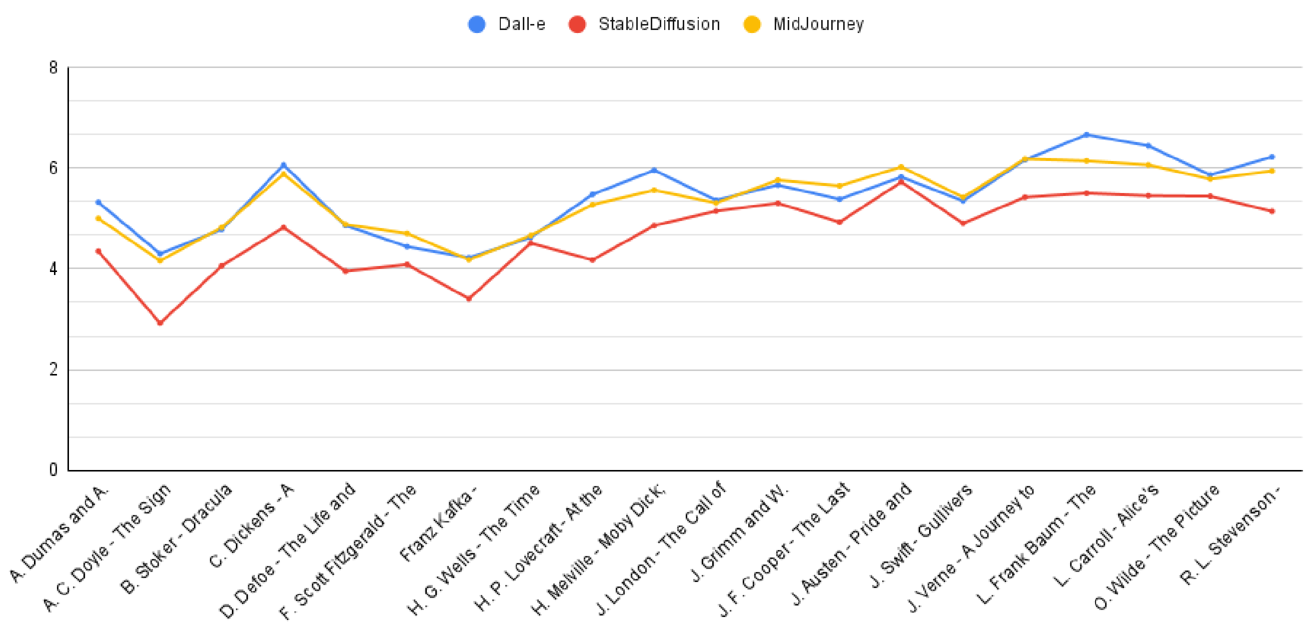

Figure 20. Diagram of average score between Dall-e, Stable Diffusion and MidJourney by book

extracting abstractions, while Dall-e has better results when using OpenAPI GPT3 and GPT4 for extracting keywords, as well as for Co-Occurrence Semantic, TextRank, and TF-IDF algorithms for extracting keywords. The results for the TF-IDF generalization, WE spaCy and WordNet algorithms are the same between Dall-e and MidJourney. At the same time, StableDiffusion image evaluation results are one unit worse than the median of Dall-e and MidJourney.

A similar sequence can be seen in another diagram (Fig. 20), which shows the dependence of the AI service evaluation results on the book. In this case, Dall-e still has superior results compared to MidJourney and StableDiffusion. MidJourney showed better results only for the books The Great Gatsby, Grimms' Fairy Tales, The Last of the Mohicans, and Pride and Prejudice. Compared to the other books, the test pages from these books contained more people in the scenes and more general descriptions characterized by distinctive clothing styles and general appearance typical of the respective years and locations. The graph sequence for StableDiffusion results remains comparable to that of Dall-e and MidJourney as in previous diagram.

At the beginning of the data analysis, it was noticeable that the generated image does not always correspond to the quality and relevance of the description, which better conveys the essence of the analyzed page. In this case Dall-e still has more higher results compared to MidJourney and Stable Diffusion.

On the contrary, the shorter the text, the better the image, and the less descriptive phrases the text contains and the more keywords it contains, the better and more relevant the image. Also, the images obtained did not always contain all the objects that were specified in the list of keywords or description.

The generated images mostly contained only a few of the specified objects, while the rest were ignored by the AI service.

From this we can conclude that at this stage of development of image generation systems, services have limitations in terms of full compliance of the generated images with the provided input query.

The following diagrams (Figs. 21-26) show the dependence of the number of words in the text descriptions generated by OpenAI GPT3 and GPT4 on the obtained image matching scores, as well as the number of times a certain number of words is obtained in the text. Accordingly, the more abstractions with the same length were retrieved, the larger the circle on the diagram will be.

The indicator of the number of used word lengths is necessary for a qualitative analysis of the dependency under study, since several estimates obtained for the best or worst results may be more likely to be a coincidence than a pattern. The purpose of this analysis was to find out whether there is a strong correlation between the number of words in a text and the quality and relative relevance of images to the text. It can be seen that most of the obtained descriptions contain an average of 20-40 words. In particular, the largest number of the obtained results have a score from 6 to 10. According to the data obtained, it is not possible to clearly define the dependence of the number of words on the quality of the obtained images for the Dall-e and Stable Diffusion services. But for the MidJourney results, the dependence is more noticeable. In particular, when using GPT3, the number of words in the range of 20-40, and for GPT4 the range of 25-45 words, provides greater relevance of images to the text.

It can be also noticed the difference between algorithms for analyzing pages that contain only dialogues between characters. In this case, conventional algorithms for selecting keywords do not cope well with their task, because they are able to isolate only certain words that do not have key semantic meaning.

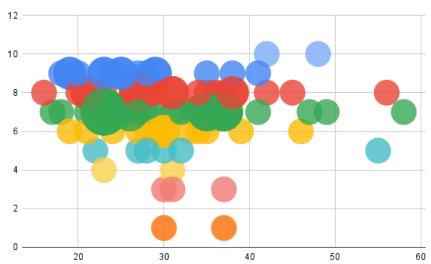

Figure 21. Diagram of GPT3 summary text length for per image score ratio for Dall-e images

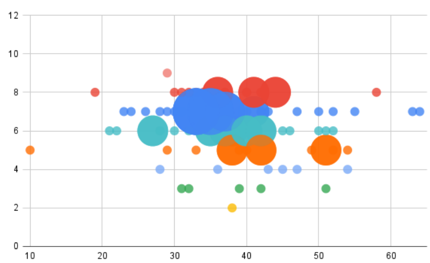

Figure 23. Diagram of GPT4 summary text length for per image score ratio for Dall-e images

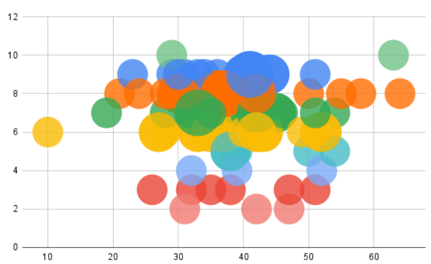

Figure 22. Diagram of GPT3 summary text length for per image score ratio for Stable Diffusion images

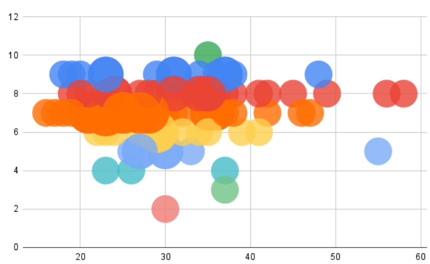

Figure 24. Diagram of GPT4 summary text length for per image score ratio for Stable Diffusion images

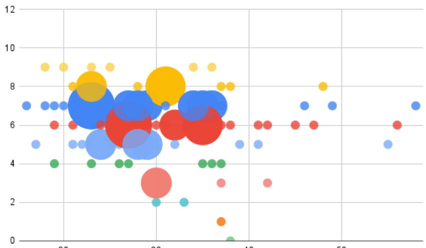

Figure 25. Diagram of GPT3 summary text length for per image score ratio for MidJourney images

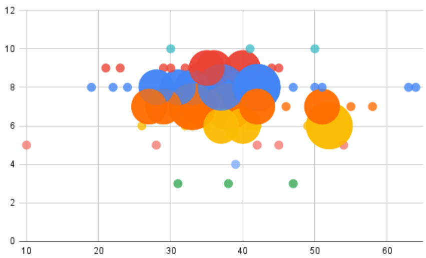

Figure 26. Diagram of GPT4 summary text length for per image score ratio for MidJourney images

V. CONCLUSIONS

AI services that return a short description of the text have better results. However, this is not enough for image generation services. The results of these services are quite poor even if an accurate description of what is happening is provided. Image generation results for other algorithms are even worse or do not match the text at all.

In general, when analyzing the data, we could see that Stable Diffusion in many cases added more persons in the generated image than described in the input data. There was also a correlation when the generated descriptions or keywords and the generated images did not match the input text well. In some cases, Stable Diffusion could not generate anything at all, and as a result, a black image was obtained, or an error was received.

At the same time, Dall-e and MidJourney provided an overall similar level of image quality and text relevance, while at the same time making guesses by adding key details to the images when they were absent in the provided abstraction.

After analyzing the results, we can conclude that the described method and the commands used are hardly up to the task. The generated abstractions contain too general abstraction of the main essence of the text on the page and it is difficult for AI services to focus on the overall picture. Nevertheless, these results showed that this approach is still better than using traditional approaches with algorithms that analyze natural language texts. Results for generated keywords are comparable to the results of the best tested algorithm, while generated abstraction has even better evaluation.

To improve the described approach, it is worth trying in future studies:

– use this approach for individual paragraphs to obtain more precise images that focus on fewer events and objects described;
– use AI services with a different number of keywords or limit the number of words in the generated description;
– change a given command so that the AI tries to separate scenes from a page or separate events from a page of text, and generate images for each of them;
– check and test API for the other AI services with same functionality (if there are such with open API), that won't be limited by general restrictions (i.e. political correctness).

Also, the assessment of the relevance or similarity of the selected words, generated abstractions and images is quite subjective, because any reader can imagine and perceive the text in their own subjective way, but the task of this study is not to guess absolutely accurately and convey the formed idea of a given page before the reader imagines it, but to help imagine, suggest or form a certain direction of images and objects to form a better perception of the work.

References

[1] D. H. Park, S. Azadi, X. Liu, T. Darrell, and A. Rohrbach, "Proceedings of the neural information processing systems track on datasets and benchmarks," vol. 1, 2021. [Online]. Available: https://datasets-benchmarks-proceedings.neurips.cc/paper_files/paper/2021/file/0a09c8844ba8f0936c20bd791130d6b6-Paper-round1.pdf

[2] S. Brade, B. Wang, M. Sousa, S. Oore, and T. Grossman, "Promptify: Text-to-image generation through interactive prompt exploration with large language models." Association for Computing Machinery, 10 2023, pp. 1–14. [Online]. Available: http://dx.doi.org/10.1145/3586183.3606725

[3] M. Elman, P. Emilio, L. B. Jimmy, and S. Ruslan, "Generating images from captions with attention," 2016. [Online]. Available: https://doi.org/10.48550/arXiv.1511.02793

[4] S. Xinyue, Q. Yiting, B. Michael, and Z. Yang, "Prompt stealing attacks against text-to-image generation models," 2024. [Online]. Available: https://arxiv.org/abs/2302.09923

[5] V. Yakymiv, L. Piskozub, and Y. Piskozub, "Using artificial intelligence to generate real-time augmented reality content in book publishing," International Journal of Control Systems and Robotics, vol. 8, pp. 1–5, 2023. [Online]. Available: https://iaras.org/home/caijcsr/using-artificial-intelligence-to-generate-real-time-augmented-reality-content-in-book

[6] R. Scott, A. Zeynep, Y. Xinchen, L. Lajanugen, S. Bernt, and L. Honglak, "Generative adversarial text to image synthesis," in International Conference on Machine Learning, 2016. [Online]. Available: https://arxiv.org/abs/1605.05396

[7] X. Zhu, A. Goldberg, M. Eldawy, C. Dyer, and B. Strock, "A text-to-picture synthesis system for augmenting communication," in Proceedings of the Twenty-Second AAAI Conference on Artificial Intelligence, 2007, pp. 1590–1595.

[8] R. Mihalcea and P. Tarau, "Textrank: Bringing order into text." in Proceedings of the 2004 Conference on Empirical Methods in Natural Language Processing. Association for Computational Linguistics, 07 2004, pp. 404–411. [Online]. Available: https://aclanthology.org/W04-3252/

[9] P. D. Turney, "Learning algorithms for keyphrase extraction," Information Retrieval, vol. 2, no. 4, pp. 303–336, 2000. [Online]. Available: https://doi.org/10.1023/A:1009976227802

[10] S. Na, M. Do, K. Yu, and J. Kim, "Realistic image generation from text by using bert-based embedding," Electronics, vol. 11, no. 5, 2022. [Online]. Available: https://www.mdpi.com/2079-9292/11/5/764

[11] S. Mathesul, G. Bhutkar, and A. Rambhad, "Attngan: Realistic text-to-image synthesis with attentional generative adversarial networks," in Sense, Feel, Design, vol. 13198. Springer International Publishing, 2022, pp. 397–403. [Online]. Available: https://doi.org/10.1007/978-3-030-98388-8_35

[12] Y. Xie, Z. Pan, J. Ma, L. Jie, and Q. Mei, "A prompt log analysis of text-to-image generation systems," in Proceedings of the ACM Web Conference 2023. Association for Computing Machinery, 2023, p. 3892–3902. [Online]. Available: https://doi.org/10.1145/3543507.3587430

[13] L. Fan, H. J. Wang, K. Zhang, Z. Pei, and A. Li, "Towards an automatic prompt optimization framework for ai image generation," in HCI International 2023 Posters, vol. 1836. Springer Nature Switzerland, 10 2023, pp. 405–410. [Online]. Available: https://doi.org/10.1007/978-3-031-36004-6_55

[14] V. Liu and L. B. Chilton, "Design guidelines for prompt engineering text-to-image generative models," in Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems, vol. 384, no. 23. Association for Computing Machinery, 2022, pp. 1–23. [Online]. Available: https://doi.org/10.1145/3491102.3501825

[15] F. Stanislav, H. Tobias, R. Federico, H. Jörn, and D. Andreas, "Adversarial text-to-image synthesis: A review," Neural Networks, vol. 144, pp. 187–209, 2021. [Online]. Available: https://doi.org/10.1016/j.neunet.2021.07.019

[16] H. Yaru, C. Zewen, D. Li, and W. Furu, "Optimizing prompts for text-to-image generation," 12 2022. [Online]. Available: http://dx.doi.org/10.48550/arXiv.2212.09611

[17] X. Xu, J. Guo, Z. Wang, G. Huang, I. Essa, and H. Shi, "Prompt-free diffusion: Taking "text" out of text-to-image diffusion models," 05 2023. [Online]. Available: http://dx.doi.org/10.48550/arXiv.2305.16223

[18] X. You, "Automatic summarization and keyword extraction from web page or text file," in 2019 IEEE 2nd International Conference on Computer and Communication Engineering Technology (CCET), 2019, pp. 154–158. [Online]. Available: https://doi.org/10.1109/CCET48361.2019.8989315

[19] X. Mao, S. Huang, R. Li, and L. Shen, "Automatic keywords extraction based on co-occurrence and semantic relationships between words," IEEE Access, vol. 8, pp. 117 528–117 538, 2020. [Online]. Available: https://doi.org/10.1109/ACCESS.2020.3004628

[20] A. Enes, R. C. N. Jason, X. Yang, G. Jie, and L. Shujun, "Improving performance of automatic keyword extraction (ake) methods using pos-tagging and enhanced semantic-awareness," 2022. [Online]. Available: https://doi.org/10.48550/arXiv.2211.05031

[21] L. Lingfei, W. Yueshan, and Z. Xiaoyi, "A study of keyword extraction for short news documents," in Sixth International Conference on Computer Information Science and Application Technology (CISAT 2023), vol. 12800, International Society for Optics and Photonics. SPIE, 2023, p. 128004H. [Online]. Available: https://doi.org/10.1117/12.3003835

[22] "How many words in one page? answers to your questions." [Online]. Available: https://www.anycount.com/word-count-news/how-many-words-in-one-page

[23] R. Mihalcea and T. Paul, "Textrank: Bringing order into text," in Proc. Conf. on Empirical Methods Natural Lang. Process. Association for Computational Linguistics, 07 2004, pp. 404–411. [Online]. Available: https://aclanthology.org/W04-3252/

[24] "Nltk toolkit." [Online]. Available: https://www.nltk.org/

[25] "Pagerank algorithm." [Online]. Available: https://networkx.org/documentation/stable/reference/algorithms/generated/networkx.algorithms.link_analysis.pagerank_alg.pagerank.html

[26] "Cosine distance." [Online]. Available: https://docs.scipy.org/doc/scipy/reference/generated/scipy.spatial.distance.cosine.html

VOLODYMYR YAKYMIV a PhD student, graduated Ukrainian Academy of Printing in 2011. Now he works as an IT engineer in Lviv Polytechnic National University, Institute of Applied Mathematics and Fundamental Sciences, Department of Applied Mathematics. Research areas: Artificial Intelligence, IT Projects.

YOSYF PISKOZUB is a Professor, Doctor of physical and mathematical sciences in Lviv Polytechnic National University, Institute of Applied Mathematics and Fundamental Sciences, Department of Applied Mathematics and Professor of the Department of Applied Mathematics, Faculty of Computer Science and Telecommunications, Cracow University of Technology, Warszawska Str., 24, 31- 155 Krakow, Poland. Research areas: Mathematical modelling of contact phenomena, modelling of fine defects of an arbitrary physical nature for further solving the relevant problems of thermal conductivity, thermoelasticity, viscoelasticity and plasticity. Study of perturbations of relevant physical fields due to the presence of fine defects, structure of nanocomposites.