Effective Semantic Markers of Using Socio-Communication Technologies in Educational Activities

Halyna Lukash, Oleksii Prihunov and Olha Anisimova

Vasyl’ Stus Donetsk National University, str. 600-richchia, 21, Vinnytsia, 21021, Ukraine

Abstract
The study aims to investigate the role of semantic markers in educational texts within the context of employing socio-communication technologies. The primary goal of the research is to identify optimal strategies for content optimization and adaptation of the semantic core of texts to maintain a consistent flow of traffic to educational resources. The scientific approach to shaping the semantic core has proven to be a tool that enhances the understanding of stakeholders’ needs and facilitates effective semantic analysis of web content. The research confirmed that content optimization and adaptation of the semantic core are crucial tasks that require continuous monitoring and improvement. It is important to note that these principles can be successfully applied not only in education but also in various fields, making them subjects of further research.

Keywords
Socio-communication technologies, keywords, search terms, semantic core, content of the information resource

1. Introduction

Socio-communication technologies are a key modern trend, which in the era of digital information world is expressed by the ratio of theoretical and practical aspects of the concepts of "communication technologies", "social technologies", "information technologies". Technologies of communication processes cover the issues of effective interaction of enterprises, institutions, organizations with the external and internal environment: public relations (PR, image building, thrash image, spin doctor), market research (market research), marketing (including advertising), management (including brand management), electronic networks, mobile communication, etc. [11].

Modern linguistics interprets language as a social phenomenon, because social interaction becomes mostly linguistic interaction and is carried out through texts that are generated and perceived by the individual. It is also worth noting that the most important and basic feature of human communication is the transmission of thoughts, i.e. not words, but the concepts they express; not sentences, but judgments, opinions and considerations that they contain [10].

Advertising has become one of the main communication elements today. To effectively promote their services, advertisers seek to use a variety of communication technologies, combining them into a well-constructed and strategically designed system of action. Every communication technology - advertising, marketing publications or public relations - has its own characteristics of impact on the consumer.

The rapid development and popularity of social networks indicate the great potential of using these platforms to promote products and services of companies, or the formation or promotion of the brand of different firms. Modern youth draws the necessary information from the World Wide Web. There are many online communities on the Internet that bring together different types of users (by age, social,
and gender). About 80% of adolescents over the age of 16 have personal accounts on various social networks and Web communities, and almost 80% of them use them on a daily basis [13]. Due to the wide audience of users, as well as their age diversity and easy access to communication with them, the concept of SMM (social media marketing) emerged. Social media marketing (SMM) is an internet marketing tool that promotes various services, companies or brands through the involvement and use of social networks. Often, SMMs create viral marketing, which not only increases the spread of information about companies, but also increases the traffic and ranking of the company's site in search engines. It is impossible to maintain high activity in social networks without constant creation of quality content.

Until recently, the idea of advertising was a novelty for higher education institutions. Although it is known that educators already used advertising more than 2000 years ago. Greek sophists publicly "demonstrated their knowledge and oratory skills and the ability to argue their point of view to show the value of their services" [4; 17]. The analysis of texts of modern advertising will allow to reveal language "markers" which, according to the plan of advertisers, should exert the directed influence on the recipient. The purpose of the article - involving in the analysis of the concept of keywords, to analyze ways to use socio-communication technologies for a strategically designed system of actions to build the semantic core of the text. The term "keywords" in this case refers to a set of words, phrases or phrases that best reflect the essence of the material.

Semantic text analysis encompasses various methods, such as frequency, content, and statistical analyses. Through the frequency analysis method, the most commonly used words and phrases in the text have been identified, allowing us to discern their significance within the research context. The content analysis method enabled a deep examination of the structure and content of the text to highlight key themes and concepts, aiding in the comprehension of the essence of the subject under investigation. By employing the statistical analysis method, statistical connections and patterns within the text were revealed, unveiling existing data patterns. This analysis facilitated an understanding of the semantic meanings of effective semantic markers when utilizing socio-communication technologies in educational activities.

2. Literature Review

The advertising text was studied mostly within the linguistic paradigm. In particular, domestic and foreign scholars have studied the text of advertising in the following areas: a general description of the features of advertising works; structure of advertising text; advertising as a discursive practice; manipulative techniques in advertising; linguistic analysis of advertising texts at different language levels, etc. At the present stage, scholars are interested in certain aspects of pragmalinguistics - an interdisciplinary science, in the forefront of pragmatic research which is the successful interaction of the initiator of speech with the audience as an object of influence. Methods of building an information society, social networks in the WWW, gender and age differences of language in Internet communication explored in their works A. M. Peleshchynshyn and S. S. Fedushko [1; 16].

The key words in the research publication were first used in 1975 in the scientific journal The Journal of Applied Behavior Analysis [14]. Their value for information retrieval was substantiated by M. Kemman, who emphasized that for keywords should be chosen not individual tokens, but phrases that allow to better understand the context and limit the number of search results [14]. E. Gbur and J. Trumbo suggested ways to define effective keywords and phrases, in particular: to avoid too general terms; do not repeat words from the title; avoid prepositions and abbreviations; use people's names as keywords only when they are part of established terminology; use variants of terms. Jacqueline Gomez specializes in semantic analysis within online communications. Her work is dedicated to understanding how the semantics of text impact communication in social media and other online platforms. Mark Holden focuses on the analysis of semantics and expressions within texts, with the goal of identifying conceptual connections and thematic patterns in linguistic data. At the same time, J. Bill described numerous cases when a keyword search may fail, for example, if a term has many synonyms or spellings, is outdated or used in various sciences, and not all concepts can be conveyed by keywords.
3. Basic points statement

According to the "Classifier of Professions" - the state standard for the classification of professions, there are 8,250 names of professions [9]. Higher education institutions train specialists in 340 professions, colleges in 450 professions, and schools in 800 professions. More than 500 new professions appear every year and old ones die out [6]. The Internet helps to understand all this variety of information. The word Google has already been adopted by our language, linking the search process not only to Google itself. People turn to the search engine to find information, answers to questions, advice. The main types of web communities include forums, social networks, personal and collective blogs, multimedia web communities, market web communities, promotional web communities, wiki communities, and educational web communities. The analysis of specialized educational web communities on the basis of web forums was most fully conducted by A. Yu. Schilling [3]. She singled out the following types of them: 1) forums for freelance entrants; 2) forums for students and teachers of free economic zones; 3) other specialized educational Web forums.

Search engines have changed the perception of information. Search engine algorithms are changing, but the requirement for content remains: it must be useful, interesting, convincing and optimized for search engines. Search engine developers compete with each other in the battle for users, constantly improving search mechanisms to provide more relevant information in response to user queries. Therefore, the formation of query results can be quite complex, but the basic stages remain unchanged (Fig. 1).

![Figure 1: Main stages of search engine operation](image)

The search engine continually scans internet resources, gathering information about the content of web pages, indexing and populating the search engine's databases. This process also involves semantic analysis of text and the identification of keywords that best match the meaning of the information published on the internet.

When the search engine receives a query from a user, it breaks it down into individual words or phrases, known as search terms. For example, if a user enters the phrase "semantic markers in education," the search engine breaks it down into separate terms: "semantic," "markers," "education." It then consults its index and searches for pages whose content matches the user-defined terms. The next step in the system's operation is evaluating the relevance of the pages found to the user's query. To improve the accuracy of results, search engines constantly update their search algorithms, adapting to various factors. Understanding natural language, information relevance, source authority, user behavior, external linking, interaction time, user location, search history, and the presence of key terms in page titles or meta tags – these are an inexhaustible list of factors that can influence how the system recognizes, processes semantic markers, and forms a final response to the user's query. Based on
relevance, the search engine sorts and displays the found internet resources. Therefore, the user primarily sees pages that most closely match their query. Thus, search terms play a crucial role in this process, as they help the search engine understand the user's specific intent.

It's also important to consider that search engines are constantly evolving and employ numerous other factors to enhance the relevance of results. In 2013 Google announced the launch of a new algorithm - "Hummingbird", which changed the classic methods of SEO-promotion. If before the website was indexed by keywords, then with the advent of "Hummingbird" this is not enough. The robot seeks not only the query, but also the meaning of the query. Hummingbird also appreciates the unique and useful content optimized for the search engine, but already highlights not only the keys, but also associative on the topic of words and phrases. The launch of the algorithm led to the use of texts based on LSA - latent-semantic analysis. Latent-semantic analysis is a method of information processing that reveals the latent connections of the studied object or phenomenon. In general, it is a method that allows you to compile a thematic dictionary and expand the semantic core, which makes it more saturated and valuable. LSA keys can be synonyms, hyponyms, associations, hyperonyms and other mixed terms. For example, when the category "education" is promoted, the search tools will include not only the main key, but also LSA concepts: students, teacher, gymnasium, first grade, six-year-old, fraternal schools, graduates, entrant, bachelor, master, graduate, level knowledge, gain, learn, abilities and other thematic words: teaching, preparation, teaching, reading, writing, mathematics, history, education, development, disciplines, skill, craft, education, as well as definitions: primary, secondary, higher, inclusive, dual, professional, linguistic, classical, university. The latest discourse is distance education with the thematic words laptop, gadgets, education in Europe. In the English-speaking world, LSA copywriting is beginning to displace the classical approach, and in our reality, latent-semantic analysis has become an additional tool for promotion. At the same time, only a well-thought-out thematic dictionary that helps to reveal the topic will work in favor of the site. The search engine algorithm is designed so that it highlights the most important places in the text and looks for relevant keys in them. Therefore, in the process of composing the text should take into account the zoning of keywords.

Search engines have changed the perception of information. Search engine algorithms are changing, but the requirement for content remains: it must be useful, interesting, convincing and optimized for search engines. In 2013, Google announced the launch of a new algorithm - "Hummingbird", which changed the classic methods of SEO-promotion. If before the website was indexed by keywords, then with the advent of "Hummingbird" this is not enough. The robot seeks not only the query, but also the meaning of the query. Hummingbird also appreciates the unique and useful content optimized for the search engine, but already highlights not only the keys, but also associative on the topic of words and phrases. The launch of the algorithm led to the use of texts based on LSA - latent-semantic analysis. Latent-semantic analysis is a method of information processing that reveals the latent connections of the studied object or phenomenon. In general, it is a method that allows you to compile a thematic dictionary and expand the semantic core, which makes it more saturated and valuable. LSA keys can be synonyms, hyponyms, associations, hyperonyms and other mixed terms. For example, when the category "education" is promoted, the search tools will include not only the main key, but also LSA concepts: students, teacher, gymnasium, first grade, six-year-old, fraternal schools, graduates, entrant, bachelor, master, graduate, level knowledge, gain, learn, abilities and other thematic words: teaching, preparation, teaching, reading, writing, mathematics, history, education, development, disciplines, skill, craft, education, as well as definitions: primary, secondary, higher, inclusive, dual, professional, linguistic, classical, university. The latest discourse is distance education with the thematic words laptop, gadgets, education in Europe. In the English-speaking world, LSA copywriting is beginning to displace the classical approach, and in our reality, latent-semantic analysis has become an additional tool for promotion. At the same time, only a well-thought-out thematic dictionary that helps to reveal the topic will work in favor of the site. The search engine algorithm is designed so that it highlights the most important places in the text and looks for relevant keys in them. Therefore, in the process of composing the text should take into account the zoning of keywords.
lacunarity is a universal category, which is characterized by systemic, multivector, polyterminological text [18]. The gaps were differently classified by researchers. There are intracultural (intra-lingual) and intercultural (inter-lingual) gaps. They can relate to intra-lingual and inter-lingual gaps and "be perceived by the recipient as something strange that requires interpretation (explicit gaps), or remain for him in the" zone of insensitivity "(implicit gaps)" [18].

We were interested in the theory of lacunarity, proposed half a century ago by Taylor and used in the study of the lexical and semantic aspect of verbal communication. First - to identify the level of language proficiency: respondents were asked to restore the text with the missing words - gaps. Then this technique was used to establish the general content of the message, which gave the second name to the theory of lacunarity - the theory of contextual predictability of the text. Based on it, A. Stern proposed a method of typing keywords [2]. To index the text, she suggested a graph of a set of keywords: a small set - 1-4 words; average - 7-12; large - 15-20. The small set covers the general topic-concept to which the text belongs; middle - characterizes the main characters, the main subject of the image, the climax, i.e., reflects the overall integrity of the text. A large set of keywords covers the subtopics of the text, words that detail its content. In essence, this graph of keywords coincides with the mentioned latent-semantic analysis. But from this perspective, we can look at keywords from a different angle: if the term "keywords" refers to words that are part of the HTML text of a document and in a prospective search engine query that searches for information on the subject of the document, we are dealing with the process of collapsing information. If we turn to LSA - we intensify the process of information deployment. Keywords in this case play the role of a broad acquaintance with the content of information and better to remember it. LSA copywriting is considered a more effective way to promote than SEO content. These texts are nice to read, they do not have unnatural keywords, and search engines quickly recognize them. The only downside of LSA materials is the delicate process of their creation. The author needs to harmoniously enter 1-2 dozen keys, create a readable article. Be that as it may, LSA texts are in demand. And the reason for that is one: they benefit sites.

On the example of educational advertising, consider the means of creating effective advertising texts on social networks. Social networks are designed to communicate with users through the creation or posting of various information in any format (verbal text, video, sound, photo), for example, personal sites for posting personal information of an individual; chats for collective written communication in real time; wikis - websites for self-input of written messages, such as articles for "Wikipedia"; blogs - websites for regular addition of records, images, multimedia. Given their purpose (communication, dating, professional contacts), social networks connect people into different social, professional groups, provide communication between them regardless of geographical location, age, gender, ethnicity, religion.

In 2020, almost half a billion new users appeared on social networks - that is, about 15 people registered every second. Some social networks have become hyperpopular amid the pandemic: for example, the average time of using TikTok per day is a record 95 minutes, more than 40 million people visit this app every week, it is opened by more than 1 billion visitors every month. Tik Tok is one of the most popular applications in the world. In addition, social media has become an even more important source of news. According to the Social Media Flagship Report, 47% of Internet users recognized them from social networks [13]. As a result, the concept of doomscrolling has become relevant - "continuous absorption of a huge amount of negative news online".

Innovations in education are introduced in order to meet the needs of entrants and students in learning and acquiring such necessary knowledge.

Branding in education creates the image of an educational institution that has its own individual style. This style is promoted on social networks, on billboards, in printed materials. After determining the educational institution, the next step is to choose a training course that suits your interests. Most students are under the age of thirty. With this in mind, it is extremely important to understand that most of them are active users of social networks. Therefore, to increase the popularity of educational services, training centers should advertise their programs through social networks such as Facebook, Twitter, LinkedIn. Search engines calculate the relevance of the text by constructing a frequency range of words. The more often the word is used in the text, the more relevant it is to the user's request. Depending on the level of competition and the task, the number of keywords can range from two to three to ten. The fewer keywords a document is set up for, the more often those keywords can appear in the text, and the more relevant they are.
Highly relevant documents focus on two or three keywords, each of which occurs in the text in a ratio close to 5%. The optimal ratio is 1 keyword per 500 characters without a space.

There are a huge number of opportunities for promotion, which are used by educational institutions. These include search engine optimization (SEO) and pay per click; social media marketing (SMM). The complete list of phrases that are placed on the pages of the site is called the semantic core. This is a library of phrases and word forms that most fully and comprehensively describe your activities, products or services presented on the site. To create an effective semantic core, you need to conduct a little research, carefully analyze the statistics of your resource and search engines, all real and potential competitors and, of course, the goods and services offered. Most likely, the list of keywords can consist of several thousand options, from which you will have to choose only the most relevant and accurate.

Note that the use of correct lexical and grammatical markers in the construction of search keywords may be ineffective [14], in the process of semantic analysis of a text document distinguish between "keywords for the search network" and "search terms" (Fig. 2). The first reflect the subject (essence of the content) of the document, its structure; the second - express the views of seekers on its subject or certain expectations. The aim of the researcher is to find the best combination of canonical structures with subjective (commonplace) thought into a single semantic core.

Electronic networks have now become an integral part of social life and necessary means of production, and the main object of Internet communication has become content [12]. Every day, users use different search engines to search for information. TOP-5 most visited sites in Ukraine in August 2023 (Fig. 3) [5; 19]:

1. Google.com;
2. Youtube.com;
3. Facebook.com;
4. Telegram
5. Ukr.net;
6. OLX.

The main task of search engines is to meet the information needs of users. And the more searches there are in the electronic network, the more often the system will be able to influence a potential subject of social interaction. Competing with each other for communication to the largest possible audience as a market for information services, they offer legal entities or individuals powerful tools for socio-communicative influence on it. Their effective practical use depends on understanding the relevant semantic markers of user behavior, each of which may have its own ideas about the semantic content of the search object. In view of the above, we will consider one of the approaches to the definition of semantic markers, which essentially means the formation of the semantic core of the national educational space in the most popular Google search network in Ukraine.

Documents that reflect the activities of higher education institutions can often be characterized by the following basic keywords: university, institute, higher education, specialty, external independent evaluation, student, graduate student, doctoral student, bachelor's or master's degree. However, there is no guarantee that the user will use the base keywords to search. Table 1 shows a snippet of a potential
Analysis of actions and behavior of users on the Internet allows you to: set the frequency of relevant requests for a certain period (Table 1, column 2); divide them into low-frequency, medium-frequency and high-frequency; establish the geodependence of requests, ie relevance in a particular area; the level of competition and the cost of advertising in a particular position (table 1, columns 3-5).

Analysis of the data in table 1 shows the following:

- terms can consist of the same words in different sequences (гете інститут, інститут гете);
- deadlines may contain errors (the term «інститу» has an average frequency of 5,000 requests per month);
- terms can be used in the singular and plural when searching for documents on similar topics (маркетинг університети, університет маркетингу);
- search queries can be formed in different languages (університет, университет);
- different search terms may have the same semantic meaning (дистанційне навчання, навчання, дистанційне навчання бакалаврат);
- search terms can determine the user's intentions to find, buy or learn something (університет шевченка вартість, бакалавр за 2 роки);
- search terms can be indefinite or ambiguous (студент хелп, навчання за обміном).

### Table 1

Fragment of the list of search terms of users for the period from 1/05/2022 to 30/04/2023

<table>
<thead>
<tr>
<th>Keywords</th>
<th>Avg. monthly searches</th>
<th>Competition</th>
<th>Top view rate</th>
<th>Min</th>
<th>Max</th>
</tr>
</thead>
<tbody>
<tr>
<td>гете інститут</td>
<td>500</td>
<td>Low</td>
<td>2,92</td>
<td>84,14</td>
<td></td>
</tr>
<tr>
<td>інститут гете</td>
<td>500</td>
<td>Low</td>
<td>2,93</td>
<td>58,02</td>
<td></td>
</tr>
<tr>
<td>маркетинг університети</td>
<td>500</td>
<td>Low</td>
<td>3</td>
<td>34,45</td>
<td></td>
</tr>
<tr>
<td>магістр</td>
<td>5000</td>
<td>Low</td>
<td>8,25</td>
<td>15,5</td>
<td></td>
</tr>
<tr>
<td>інститут</td>
<td>5000</td>
<td>Low</td>
<td>2,26</td>
<td>14,18</td>
<td></td>
</tr>
<tr>
<td>інститут гринченка</td>
<td>500</td>
<td>Low</td>
<td>0,99</td>
<td>13,63</td>
<td></td>
</tr>
<tr>
<td>менеджмент</td>
<td>50</td>
<td>Low</td>
<td>1,64</td>
<td>9,92</td>
<td></td>
</tr>
<tr>
<td>університеті</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>університет маркетингу</td>
<td>50</td>
<td>Low</td>
<td>2,54</td>
<td>9,91</td>
<td></td>
</tr>
<tr>
<td>магістр юа</td>
<td>500</td>
<td>Low</td>
<td>5,27</td>
<td>9,37</td>
<td></td>
</tr>
<tr>
<td>студентческие работы</td>
<td>50</td>
<td>Medium</td>
<td>4,92</td>
<td>8,3</td>
<td></td>
</tr>
</tbody>
</table>
Thus, one of the key indicators used by the search engine to meet the information needs of the user is the optimization of texts and metadata on the information resource for search terms [8]. This means that the high-frequency search terms of the semantic core should be used in the formation of the content of the information resource, which should provide stable traffic (Table 2).

Table 2 draws attention to the search term «мультипредметне тестування» with an average number of queries of 500,000 per month with low competition for a high advertising position in the search network.

**Table 2**

A fragment of the list of high-frequency search terms of the educational environment

<table>
<thead>
<tr>
<th>Keywords</th>
<th>Avg. monthly searches</th>
<th>Competition</th>
<th>Top view rate Min</th>
<th>Top view rate Max</th>
</tr>
</thead>
<tbody>
<tr>
<td>магістр</td>
<td>5000</td>
<td>Low</td>
<td>8,25</td>
<td>15,5</td>
</tr>
<tr>
<td>мультипредметне тестування</td>
<td>500000</td>
<td>Low</td>
<td>1,44</td>
<td>5,58</td>
</tr>
<tr>
<td>університет</td>
<td>5000</td>
<td>Low</td>
<td>0,59</td>
<td>4,1</td>
</tr>
<tr>
<td>вища освіта</td>
<td>5000</td>
<td>Low</td>
<td>1,65</td>
<td>2,45</td>
</tr>
<tr>
<td>спеціальність</td>
<td>5000</td>
<td>Low</td>
<td></td>
<td></td>
</tr>
<tr>
<td>студент</td>
<td>5000</td>
<td>Low</td>
<td></td>
<td></td>
</tr>
<tr>
<td>стипендія</td>
<td>5000</td>
<td>Low</td>
<td></td>
<td></td>
</tr>
<tr>
<td>диплом</td>
<td>5000</td>
<td>Low</td>
<td></td>
<td></td>
</tr>
<tr>
<td>бакалавра</td>
<td>5000</td>
<td>Low</td>
<td></td>
<td></td>
</tr>
<tr>
<td>з днем студента</td>
<td>5000</td>
<td>Low</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

It can also be noted that one of the conditions for constant traffic of the target audience is a common semantic core (Fig. 4).
Advertising of educational services has certain functions, among which the main ones are: informational; value-oriented; communicative [7]. In total, humanity has spent more than 1.8 billion years on the Internet in the past year alone. This is the data of the annual report Global Digital 2022. Most of this time – in social networks. Today, 4.9 billion of the world's nearly eight billion people use social networks – more than half. The number of social media supporters has tripled in the last decade and has grown by about 16% in the last year, partly due to the pandemic and wartime.

At present moment the use of social networks became more diverse than before. Previously, the priority was interpersonal communication, and now they are increasingly used for private business and view the latest news.

7. Conclusions

There are many factors that affect the ranking positions of representative information resources on the Internet. They change over time, but for a long time one of the most important is the optimization of content (texts and metadata) for targeted user requests. The identified problem is always relevant, including in the educational space, as it requires constant updating of official resources. The problem is exacerbated by the fact that users’ views on the semantic content of resources are constantly changing and may also depend on geographical location, age, gender, ethnicity, religion. A scientific approach to forming the semantic core of the educational space can be the solution, with a key focus on understanding the needs and demands of stakeholders (prospective students and their parents, teachers, employers). It’s also important to understand that the semantic core is not static and requires continuous improvement based on new trends, changes in the educational landscape, and user queries.

The approaches to creating a semantic core demonstrated in the article can be successfully applied to other semantic directions, not limited to the educational space. The semantic core is a tool that can be applied to any type of business, project, or website. The fundamental principles remain similar: studying the target audience to understand its needs, identifying key words and phrases that best reflect the content or services directly (even competitor websites can serve as sources of ideas), optimizing content in accordance with the chosen key words, and creating high-quality content (informative, engaging, and valuable for the specific audience).

The general principles and approaches to creating a semantic core remain consistent, but the key is adaptation to a specific industry or field of activity, which may be the subject of further scientific research.
8. References


[7] Keywords: concepts and types. URL: https://aptxt.com/klyuchevye-slova-ponyatie-i-tipy.html/


[17] Strategic marketing for educational institutions. URL: https://koloro.ua/ua/blog/brending-i-marketing/marketing-i-brending-obrazovatelnych-uslug.html.