International Journal of Technology

http://ijtech.eng.ui.ac.id

Research Article

Functional Modeling of Distributions of Substantive-Content Message Properties in the Information Background

Evgenii Konnikov[1], **Dmitriy Rodionov**[1], **Darya Kryzhko**[1,*]

[1]Peter the Great St. Petersburg Polytechnic University, 29 Politechnicheskaya Ulitsa, St. Petersburg, 195251 Russia

*Corresponding author: darya.kryz@yandex.ru, Tel.: +7 (812) 775-05-30

Abstract: This paper presents a novel methodology for modeling the distribution of substantive-content message properties in the information background. This study develops a toolkit to analyze and predict information dynamics by identifying key themes, evaluating their importance, and understanding their connections. The proposed approach is based on the concept of multimodality, where properties are characterized by peaks of varying intensity and frequency. Intensity and frequency components are modeled separately and combined into a unified probabilistic framework; model parameters (shape and internal-covariance coefficients) are searched within the range (0,1). The genetic search uses mutation of ±20% with probability 50% and normalization of the intensity scale ($\theta = 1$). Model quality is assessed by the Mean Absolute Error between ranked histogram bins (discretization coefficient DC defines the number of bins). Intensity, reflecting the depth and saturation of the information signal, is modeled using the Gamma distribution, while frequency, reflecting the number of occurrences, is represented by the multivariate normal distribution. A genetic algorithm is employed to identify the optimal parameters for these distributions. The methodology offers a more comprehensive understanding of information dynamics by considering both intensity and frequency, and effectively handles complex interdependencies between properties. It can be applied to various domains, including social media analysis, political communication, and marketing, providing valuable insights for decision-making.

Keywords: Information background analysis; Intensity of presence; Substantive-Content message; Symbolic data constructs; Thematic analysis

1. Introduction[1]

The rapid evolution of socio-economic relations over recent decades has given rise to new trends, rules, and laws governing enterprise operations. Currently, during the emergence of a post-industrial economy, the significance of material factors for an enterprise corresponds to a smaller share of value in terms of the resource structure of activities than in the previous stages of economic development. Meanwhile, the weight of intangible elements, such as information, has a steadily positive trend towards growth (Berawi et al., 2022; Eremina et al., 2022; Nugraha et al., 2022; Zaytsev et al., 2019; Hitt et al., 2019; Damanpour, 2018).

The stability of these trends is explained by the formation of an empirically and scientifically confirmed understanding of the comprehensive essence of information. Information is an intangible resource that breaks barriers and drives globalization (Rodionov et al., 2021; Mueller and Grindal, 2019; Bharadwaj, 2000).

[1]This work was supported by ≪Ministry of Education and Science of the Russian Federation≫ funded by ≪Development of a methodology for instrumental base formation for analysis and modeling of the spatial socio-economic development of systems based on internal reserves in the context of digitalization" (FSEG-2023-0008)≫

https://doi.org/10.14716/ijtech.v16i6.7411
Received September 2024; Revised November 2024; Accepted December 2024

Information, presented as applied knowledge, determines the effectiveness of other classical factors of production (P. Li et al., 2024). Information has a huge impact on classical production resources such as labor, capital, and land (Anisiforov et al., 2017). It plays a role in resource management, production processes, and decision-making, which can lead to increased efficiency in resource utilization and improved productivity (Öberg, 2023; Qiu et al., 2022; Rodionov, Kryzhko, et al., 2022).

Practical information obtained through the analysis of large data sets contributes to an increase in the quality of decision-making, both for individual actors within an enterprise and for the enterprise as a whole (Kushwaha et al., 2021). With the rapid increase in computing power available to analysts, the use of big data analytics for efficiency in small areas of everyday management is becoming a competitive necessity for organizations (Anisiforov et al., 2017). The costs of obtaining actionable information through the volume, variety, and veracity of data, enabling individuals or organizations to make informed, qualitatively balanced decisions underpinned by the fact of reducing the risk of uncertain outcomes, correlate with the income generated by the socio-economic benefits of such an approach (Rodionov, Gracheva, et al., 2022; Cawsey and Rowley, 2016). Therefore, studying the mechanism of information background formation among a group of subjects and developing a methodology for modeling the properties of the substantive-content message that arises in such an environment is becoming a pressing issue in contemporary research.

Karl Deutsch, an American political scientist and sociologist who studied the specifics of communication in the context of the state of political systems and management processes, posited the thesis that information has various properties that determine its impact on the system consuming it. Karl Deutsch was the first to begin to separately highlight what he called the ”tone” or ”temperature” of information, characterizing its emotional coloring. Deutsch also highlighted the profound effects of information on society, particularly its capacity to shape the status quo, stimulate collective action or hinder activity, and influence the allocation of resources and power (Z. Li et al., 2023). As we can see, Deutsch assumed that information is not solely a factual or neutral set of data, but rather a substance carrying a significant impact and driving changes in social dynamics.

Fred Dretske, a pioneer in the philosophy of cognitive science, in turn, focused his research on fundamental questions concerning the nature of information, perception, and cognition. Dretske developed a detailed theory linking information and knowledge, known as the ”semantic conception of information.” In terms of the properties of information, Dretske primarily highlighted quantity and semantic aspects. Dretske posited that the information content of a message can be quantified based on its ability to diminish uncertainty in decision-making for the information recipient, aligning closely with the core principles of Shannon’s information theory.

However, what was unique about Dretske’s research was his postulation of the significance of truth and essentiality in information. For a data construct to be recognized as information in Dretske’s understanding, it must not only be semantically significant but also true. False or misleading messages cannot be considered as true information, as they do not add to existing knowledge.

Furthermore, one of Dretske’s key ideas was that a message or signal must embody a ”claim to mattering” in the context of its interpretation. This means that information must be relevant and appropriate in a given context for it to be meaningful to an individual (Vigo, 2013).

The work of Karl Deutsch and Fred Dretske is the basis for this study. Deutsch emphasized the systemic impact of information on societal and political dynamics, focusing on its ability to mobilize or suppress activity and redistribute power. His identification of properties such as the ”tone” or ”temperature” of information as measures of its emotional impact aligns with the study’s focus on the qualitative dimensions of substantive-content messages. Similarly, Dretske’s semantic theory of information, which links its value to the reduction of uncertainty and its truthfulness, underpins the quantitative aspects of this research. Dretske’s focus on relevant and contextually appropriate information guides the methodological design, specifically in modeling

how intensity and frequency influence the substantive content message's presence. By integrating these perspectives, this study advances a probabilistic framework that not only captures the complexity of information dynamics but also bridges the qualitative and quantitative dimensions identified in Deutsch and Dretske's seminal works.

A significant development of this idea can be seen in the research of Karl Pribram. Karl Pribram was a renowned neuropsychologist and a pioneer in the field of neuroscience, particularly in the study of memory, perception, and other cognitive processes. His contribution to the understanding of information can be viewed through the prism of his "holographic theory of memory", according to which information is encoded in the individual's brain in a distributed manner, similar to the nature of holography. In the context of his research, one can highlight such properties of information as its distributed nature, fractal complexity, interaction with other information, and its connection to context.

According to Pribram's holographic theory, information is not localized in a specific area of the brain but rather distributed throughout the entire volume of the brain, while also having a fractal nature, manifested in the fact that patterns of repetition and nesting are found at different levels of thought processing - from neural networks to higher cognitive functions. At the same time, Pribram postulated that information does not exist in a vacuum and it always interacts with other information. This interaction changes and redefines the significance of information depending on the context, similar to how an image in a hologram can change depending on the perspective from which it is illuminated. Information in Pribram's research is invariably connected to context. This property emphasizes that the brain's information processes do not merely process raw data but actively construct a substantive-content message, influenced by accumulated experience and ongoing expectations (Nishiyama et al., 2024).

Therefore, it can be said that the nature of the human brain determines the substantive content message of the data construct based on the information basis formed in the consumer's consciousness and in combination with the continuous context. Consequently, the very process of perception and the previously formed basis of this perception transform the substantive-content message embedded by the source of information.

The scientific novelty of this study lies in the development of a probability distribution function for the presence of substantive-content message properties in the information background. This function integrates the intensity parameter, which reflects the depth and saturation of the substantive-content message, and the internal covariance parameter, which characterizes the relationships between the manifestations of these properties. Together, these parameters enable a detailed analytical analysis of the information background, capturing both the degree of importance of the topics and their frequency of occurrence. Current research on modeling information flows primarily focuses on either intensity or frequency, often treating these parameters separately and overlooking their interdependence. Moreover, existing methodologies tend to emphasize quantitative aspects while neglecting qualitative dimensions, such as thematic richness and the interrelations among components. These limitations reduce the ability of current models to explain the complex dependencies and dynamics inherent in multidimensional information environments (Cappella and Li, 2023; Lim and Schmälzle, 2023; Bashir et al., 2022; Bruce et al., 2017; Clark et al., 2007).

This study addresses these gaps by introducing an integrative approach that combines intensity and frequency parameters using gamma and multivariate normal distributions. This approach enables a more comprehensive analysis, taking into account the depth, richness, and interrelationships of messages, thereby providing a nuanced and accurate depiction of the information background. Additionally, the use of genetic algorithms to optimize distribution parameters enhances the model's flexibility and adaptability, opening new opportunities for the analysis and forecasting of information dynamics.

2. Methodology

The information background, presented as a collection of substantive-content messages of data constructs, can be represented in many forms at the physical level. The fundamental types in this case are symbolic and natural information. Natural information, as a category, unites data formed by natural phenomena and directly perceived by the individual's sensory organs (Goldberg, 2022; Nadkarni et al., 2011). This type of information encompasses both visual and auditory images, as well as complex structures that emerge in natural languages developed through cultural and social evolution.

One of the key researchers in the analytical specificity of natural information is Leonard Euler, who proposed using graph models to describe systems presented in natural form (Stapleton et al., 2010). Also, a significant contribution to the study of the analytical specificity of natural information was made by Gregor Mendel in the framework of heredity research. Mendel's research in the field of genetics enabled the description of the mechanisms underlying how information is encoded and transmitted at the biological level (Smýkal and Eric, 2023).

Symbolic information, in turn, is characterized by the use of signs and symbols of an artificial nature (created by individuals to represent and convey knowledge, ideas, and concepts). Language, mathematical notation, and computer codes are examples of symbolic systems.

A key specificity of symbolic information is the a priori need for interpretation, according to which each symbol or sequence of symbols is assigned a specific meaning by the regulations established within the cultural or disciplinary context. Symbolic systems are characterized by a high degree of abstraction and formalization, which allows for the use of logical and algorithmic tools for their analytical formalization (Onykiy et al., 2020; Beth, 2012).

The most significant research contribution in the field of symbolic information analysis was also made by Claude Shannon. The concept of entropy, formulated by Shannon, became a key tool for evaluating the amount of symbolic information and its transmission in digital form (Omar and Peter, 2020). It is also necessary to recall the work of Alan Turing, primarily in the context of abstract computing devices capable of modeling algorithmic processes. Conditional "Turing Machines" most clearly demonstrate the concept of symbolic information, as they represent algorithms and data in a symbolic form (Copeland and Fan, 2023).

Within the described methodology, the primary focus is on the universality of encoding the input array for comparison purposes. This leads to the postulation of the thesis that the final form of data constructs containing a substantive-content message requires a transition to a comparable symbolic form. Natural textual form is recommended as this form.

Natural textual form of information represents one way of encoding the substantive-content message using natural language. This language develops organically within social groups and cultures and differs from formal and artificial languages in its polysemy, its non-reducibility to a strict formal structure, and its flexibility in expressing individual elements and contexts of human experience. Natural textual form of information encompasses a wide range of variations in manifestation, including literary works, scientific documents, news articles, social media correspondence, legal documents, and much more. Textual constructs generated by individuals are often imbued with a rich variety of semantic nuances, metaphors, allusions, and other syntactic figures, thereby shaping unique characteristics. Furthermore, natural language possesses a complex hierarchical structure, operates on various levels of abstraction, and intricately interacts with numerous facets of human cognition and social interaction patterns (Y. Li et al., 2022; Boerman et al., 2021; Hsieh and Chen, 2011).

The natural textual form of information has contextual, sociocultural, and cognitive features that determine how the text is perceived and interpreted by the individual. Aspects such as intonation, emotional coloring, and stylistic design largely shape the subjective component of the substantive-content message of the data construct. Thus, it is precisely the natural textual form of information that allows for a balance between the universality of analysis and the richness of substantive content (Wang et al., 2023; Amur et al., 2023; Dehaene and Cohen, 2011).

Methodologically, the primary stage in modeling the properties of the substantive-content

message encoded in the form of symbolic data constructs is the collection and systematization of an array, which in this case is presented in natural textual form. The collection process can be differentiated into two fundamentally different approaches: automated and non-automated (Biggers et al., 2023).

An automated approach implies interaction exclusively within the digital environment, which in turn leads to a dualism of the final data array. On the one hand, the information aggregated within the approach is secondary, which in turn increases the complexity of specifying the target properties of the substantive-content message. However, the dissociation of the subject of generation from the subject of aggregation allows for the primacy of the substantive-content message in relation to the impact that inevitably occurs during communication between these subjects (Weinlich and Semerádová, 2022).

Technologies for automated information gathering in the digital environment are described in the direction of parsing. Parsing is the process of automated extraction of information aggregated by web resources. The concept involves multiple operations, such as requesting web content, extracting, and structuring data to transform information presented in various formats, especially natural text (Lytras et al., 2020).

Parsing and analyzing symbolic textual data involves a multi-stage process: collecting data (automated or manual), preprocessing it through tokenization, normalization, tagging, and lemmatization, then vectorizing it using methods like One-Hot Encoding, TF-IDF, or Word Embedding. These vectors feed into supervised models (classification or regression) or unsupervised ones (like clustering with LDA or K-Means) to uncover patterns and thematic groupings. Model quality is evaluated using appropriate metrics accuracy, precision, recall, ROC-AUC for classification; R^2, MAE, MAPE, MSE for regression; and Silhouette Score, Calinski-Harabasz, WCSS for clustering to ensure effective detection of meaningful content and providing actionable insights into the structure and semantics of text data (Weaver, 2017).

The developed algorithm for multimodal analysis integrates intensity modeling (using the Gamma distribution) and frequency modeling (using the multivariate normal distribution) to capture the depth and recurrence of substantive-content message properties. These two components are combined into a unified probabilistic framework, optimized via a genetic algorithm to identify the most accurate parameters. The resulting optimized model provides a foundation for generating analytical insights and practical applications, enabling the exploration of complex thematic relationships within the information background. Figure 1 presents a visual representation of the algorithm, illustrating the structured operational flow and integration of key mathematical and analytical components.

3. Results and Discussion

The distribution of substantive content message properties in the information background is multimodal, and the specifics of the correlation of peaks are the most content-significant from the point of view of analysis. To describe this distribution, it is proposed to represent it as a product of two components: intensity and frequency of presence of substantive-content message properties in the information background.

Intensity reflects how strong or detailed an information signal is. In this case, we are not referring to the mere presence of a particular topic or narrative in the information background, but rather to the measure of its influence, significance, or emotional saturation. Intensity can manifest through the detail of discussion, the degree of emotional involvement of the audience, or the depth of analytical coverage of the topic.

The frequency of presence of substantive-content message properties, on the other hand, refers to the number of times that specific information or a topic manifests in the information background in a limited time frame. It measures how often the entity manifests in messages, publications, or discussions, regardless of the depth or context of its presence. In terms of probability and statistics, frequency can be associated with a series of events in the Poisson process, where the interest lies in the appearance of specific topics in the information background.

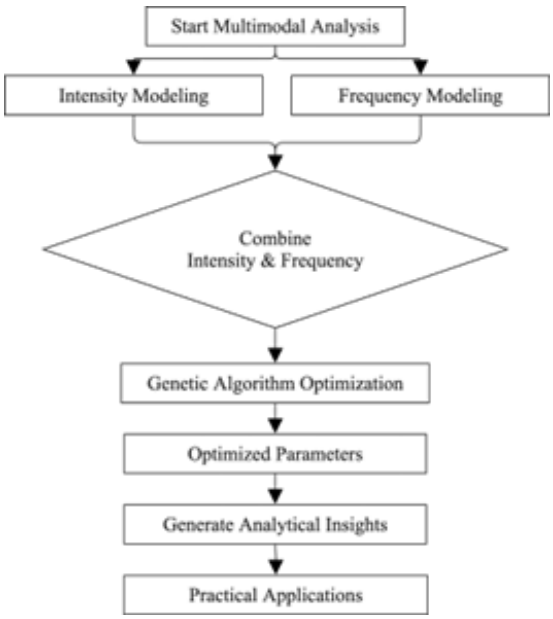

Figure 1 Multimodal Probabilistic Analysis Algorithm for Substantive-Content Message Properties

The key difference between intensity and frequency lies in quality versus quantity: intensity reflects the richness or depth of topic presentation, whereas frequency measures how often a topic appears, regardless of context or depth.

For the purposes of approximating the described properties, it is proposed to use separate types of distributions. To approximate the intensity of presence of substantive-content message properties in the information background, it is proposed to use the Gamma distribution.

The Gamma distribution, denoted as G (k, θ), where k $>$ 0 is the shape parameter and θ $>$ 0 is the scale parameter, which is a continuous distribution defined on the positive semi-axis. The probability density of the Gamma distribution is given by the function (see equation 1):

$$f(x;k,\theta) = \frac{x^{k-1}e^{-\frac{x}{\theta}}}{\theta^k \Gamma(k)}, x > 0 \tag{1}$$

where G(k) is the Gamma function, defined as (see equation 2):

$$G(k) = \int_0^\infty x^{(k-1)} e^{(-x)dx} \tag{2}$$

In the context of analyzing the information background, the Gamma distribution stands out for its adaptability and ability to model diverse phenomena associated with the accumulation of events, making it particularly relevant for the quantitative study of the intensity of the substantive content message. When analyzing the information background, the Gamma distribution can be used as a mathematical model to estimate the parameters of the event distribution over time. In this context, k is interpreted as the expected number of "events" (in this case, the appearance of specific themes or ideas), and θ is a measure of time or space during which these events can occur. This, in turn, forms the basis for probabilistic modeling and statistical analysis of the intensity of presence of thematic narratives in the information background, allowing us to assess both the probability of a specific intensity of discussion and calculate the expected indicators of the appearance of informational entities.

To approximate the frequency of the presence of substantive-content message properties in the information background, it is proposed to use the multivariate (or multidimensional) normal distribution.

The multivariate (or multidimensional) normal distribution is a generalization of the univariate normal distribution for a vector of n variables. The multivariate normal distribution is defined by a vector of means (μ) and a covariance matrix Σ, where each element of the vector (μ) represents the mean of the corresponding variable, and each element of Σ represents the covariance between the variables.

The expression for the probability density of the multivariate normal distribution is given in equation 3.

$$f(x;\mu,\Sigma) = 1/\left((2\pi)^{\frac{n}{2}}|\Sigma|^{\frac{1}{2}}\right) exp\left(-\frac{1}{2(x-\mu)^T\Sigma^{-1}(x-\mu)}\right) \quad (3)$$

where $\mathbf{x}$ — is the vector of variables, n — is the dimension of the variable space, $|\Sigma|$ — is the determinant of the covariance matrix, and $\Sigma-1$ is the inverse covariance matrix.

The multivariate normal distribution is effective for modeling and analyzing the frequency of the presence of the substantive-content message in the information background, as it enables the capture and modeling of correlations between various entities in the information space. This allows for analyzing how the appearance of one topic in the information background is related to the appearance of other topics. Additionally, the multivariate distribution leverages the covariance matrix to accommodate various relationships between variables, including uncorrelated, positively correlated, and negatively correlated data, enabling accurate representation of complex relationships among topic appearance frequencies.

In the modern world, information often has a multidimensional nature (e.g., text data, metadata, context-dependent properties). The multivariate normal distribution can effectively model and analyze this multidimensionality, providing valuable insights into the structure of information interactions. Therefore, using the multivariate normal distribution to describe the frequency of the presence of the substantive-content message in the information background allows for a deeper understanding and quantification of the relationships and structure of information flows. At the same time, the covariance matrix can be represented by a uniform parameter, which meaningfully reflects the probability of dependence of nearby values in the array on each other. Thus, the distribution of the presence of substantive-content message properties in the information background can be described by the following parameters:

k – The shape of the intensity of the presence of substantive-content message properties in the information background. In the context of the Gamma distribution, the shape parameter allows characterizing the following specifics of the analyzed properties:

- Level of concentration of the distribution. Low values of the shape parameter indicate a more concentrated distribution, while higher values are characteristic of more distributed data. This specificity may indicate the significance of the diversity of the intensity of the presence of the analyzed substantive-content message property in the information background.
- Level of homogeneity of the distribution. When analyzing the information background, the shape parameter enables the formulation of comparative conclusions about the diversity or homogeneity of events related to the appearance of the analyzed substantive content message properties. A higher level of this parameter indicates greater diversity, where events are more evenly distributed over time or information space.
- Level of variability of events. The variability of the data relative to the average value increases with an increase in the shape parameter, which in turn indicates a more predictive presence of the substantive-content message properties with higher values of k.

The Gamma distribution was selected to model the intensity of substantive-content message properties because of its capacity to represent positively skewed phenomena, which are characteristic of the depth and saturation of information signals. Its shape parameter (k) captures the

concentration and variability of intensity, while the scale parameter (θ) accounts for the range of distribution, making the model adaptable to various contexts. Meanwhile, the multivariate normal distribution was chosen for modeling the frequency of these properties due to its ability to describe correlations and co-occurrences among multiple variables. By using a mean vector (μ) and covariance matrix (Σ), it provides a detailed depiction of thematic relationships and multidimensional data structures. Together, these distributions complement each other, allowing for a comprehensive representation of both the quality (intensity) and quantity (frequency) aspects of the information background. This integration forms a robust probabilistic framework that accurately models multimodal data, enabling advanced analysis and predictive insights.

General specificity of the information background slice or the analyzed information flow. The probabilistic model, including the shape parameter, can reflect the specificity of the information flow, particularly in terms of how often and with what intensity specific properties appear. In particular, in the media space, characterized by intense thematic renewal, the shape parameter will allow us to conclude the level of relevance or saturation of the information background.

θ – The scale of the intensity of the presence of substantive-content message properties in the information background. In the context of the Gamma distribution, the scale parameter allows us to characterize the following specifics of the analyzed properties:

- Intensity of targeted information flows. A high value of θ indicates a broader spread of data, which may indicate a high intensity of targeted information flows with a more diverse frequency of occurrence. This fact indicates that the substantive-content message may appear frequently in some periods and significantly less frequently in others, providing a non-uniformly distributed pattern.
- Range of appearance. The parameter θ also allows us to assess the variability of the appearance of informational entities, i.e., how often and in what volumes the information message is found in the information background.

μ - The vector of mean values, reflecting the central tendency of the presence of substantive-content message properties in the information background. In an analytical sense, the parameter μ indicates the dominant position of substantive-content message properties within the considered multidimensional space. Interpretation of the μ value allows identifying the general orientation of the information field and assessing around which axes of essential or content parameters the data is grouped, thereby reflecting the average statistical characteristic of the prevailing trend in the extensive information flow. Thus, the value of μ in this analytical context is an indicator of the midpoint of the attributes of the substantive-content message, from which one can identify both general trends in the information space and transitions in the dominance of specific properties.

Σ – The covariance matrix, reflecting the degree of interrelationship between substantive-content message properties in the information background. The elements of this matrix indicate the covariance between each pair of variables included in the analysis. In analyzing the distribution of substantive-content message properties within the information background, this parameter reflects the structure and dynamics of their interaction. When analyzing the information background based on the multivariate normal distribution, Σ allows us to assess how synchronously the discussed topics, emotional coloring, and other substantive-content aspects change.

Each of the identified parameters has significant analytical specificity. However, parameters k and μ meaningfully reflect unified characteristics of the presence of substantive-content message properties in the information background, which allows us to equate them to each other. Parameter θ, in turn, is less analytically significant, as the comparison of distributions in this case requires scaling, which in turn implies normalization. Therefore, parameter θ can be equated to 1.

The covariance matrix Σ, on the other hand, has a more specific nature. Since the distribution of each substantive-content message property is considered separately from the others,

interrelationships can be considered solely within the framework of the elements of a single array. Thus, the covariance matrix can be replaced with a homogeneous scalar-augmented matrix, where all diagonal elements equal 1 and all off-diagonal elements are identical but differ from the diagonal values. Off-diagonal elements, in turn, reflect the level of dependence of the manifestation of properties of adjacent elements of the array. As it increases, the probability of the property manifesting in elements adjacent to the dominant one increases. This parameter can be denoted as the coefficient of internal covariance, and the matrix formed based on it will take the form (see equation 4):

$$\sum_{ij} = \{1 \ if \ i = j \ Corr_i \ if \ i \neq j \tag{4}$$

Thus, the distribution of the presence of substantive-content message properties in the information background can be described by two conceptual parameters:

- k: The intensity parameter, reflecting the depth and saturation of the substantive-content message.
- Corr: The internal covariance parameter, characterizing the relationship between the manifestations of the properties of the substantive-content message.

To identify these parameters, it is necessary to approximate the distribution of the data array describing a particular property of the substantive-content message. However, the simultaneous use of continuous and discrete distributions does not allow the use of traditional methods such as the method of moments or the method of maximum likelihood. For the purposes of iterative data parameter selection, genetic algorithms can be used, as:

- Genetic algorithms are helpful in exploring many parameter combinations in complex models.
- Genetic algorithms are able to identify the global optimum of a function, even if the optimization landscape contains many local minima and maxima.
- Thanks to their adaptability and flexibility in tuning, genetic algorithms can be effectively adapted to the task of finding the best parameters for functions describing the distribution of data with complex interdependencies between variables.
- Unlike traditional optimization methods, which require a formal definition of gradients or other function derivatives, genetic algorithms can optimize parameters without the need to define the exact form of the distribution.

These properties indicate the potential effectiveness of using genetic algorithms in the context of the task at hand.

In the first stage of the algorithm, an initial population is formed. The initial population in this case is represented by arrays of values of the internal covariance coefficient of the manifestation of substantive-content message properties in the information background (Corr) and values of the intensity coefficient of the presence of substantive-content message properties in the information background (k) within specified ranges. As ranges of coefficients, values from 0 to 1 (exclusive) can be specified. The size of the initial population is also variable and can be denoted as P.S. (see Equations 5 and 6).

$$Corr = \{Corr_i \ simU(0.01, \ 0.99), i = 1, 2, ..., P.S.\} \tag{5}$$

$$k = \{k_i \ simU(0.01, \ 0.99), i = 1, 2, ..., P.S.\} \tag{6}$$

where:

$Corr_i$	:	The i-th variation of the internal covariance coefficient of the manifestation of substantive-content message properties in the information background.
k_i	:	The i-th variation of the intensity coefficient of the presence of substantive-content message properties in the information background.
$P.S.$	:	The size of the initial population.
$Corr$	:	The generated array of values of the internal covariance coefficient of the manifestation of substantive-content message properties in the information background.
k	:	The generated array of values of the intensity coefficient of the presence of substantive-content message properties in the information background.

A simulation model is constructed based on generated value pairs, incorporating both multivariate normal and Gamma distributions. The modeling process begins with the multivariate normal distribution.

The first stage of this process involves generating a sample from the multidimensional normal distribution (see equation 7):

$$Y = [y1, ..., y_{NSim}], where y_i \sim \frac{1}{\left((2\pi)^{\frac{n}{2}}|\Sigma|^{\frac{1}{2}}\right) exp\left(-\frac{1}{2(x-k)^T\Sigma^{-1}(x-k)}\right)}, \Sigma_{ij} = \{1\ if\ i = j Corr_i\ if\ i \neq j\} \quad (7)$$

where:

Y	:	A matrix of random vectors of size NSim × n.
y_i	:	The i-th row vector in the matrix Y, represents a single simulation from the distribution.
$NSim$	:	The number of simulations.
Σ_{ij}	:	A homogeneous scalar-augmented matrix, all elements on the main diagonal of which are equal to 1, and all off-diagonal elements are equal to Corr_i

Next, the quantile of the standard normal distribution is determined, and each element of the sample is compared to the quantile, and the result is converted to 0 or 1 (see equation 8):

$$Q = \left[\left(y_1 > F^{-1}(1 - K_i)\right)\right] \quad (8)$$

where F^{-1} denotes the inverse distribution function (quantile function) of the standard normal distribution, and the intensity coefficient of the presence of substantive-content message properties in the information background (k_i) indicates the specified probability level. Each element y_i of the matrix Y is compared to the quantile, and the result of the comparison is transformed: if y_i is greater than the value of the quantile, then 1 is placed in Q, and 0 otherwise. This transformation is applied element-wise for each value in Y.

Finally, the matrix of results is transposed and its type is converted (see equation 9):

$$Q' = transpose(Q) \quad (9)$$

where Q' is the transposed matrix of Q. The matrix formed as a result describes the frequency of manifestation of substantive-content message properties in the information background. To account for intensity, a corresponding simulation of the Gamma distribution is necessary. Generation of the matrix of random intensity values is determined by Equation 10.

$$M = [M_{ij}],\ where\ M_{ij} \sim \frac{x^{k-1}e^{-\frac{x}{\theta}}}{\theta^k \Gamma(k)},\ \theta = 1 \tag{10}$$

Each element Mij of the matrix M is a random number generated from a Gamma distribution with parameters k_i and $\theta = 1$, thus forming a random matrix of dimension NSim × NSim. The resulting arrays are then multiplied and the results normalized (see equations 11,12):

$$S = \sum_{i=1}^{n} Q_i'.M_i \tag{11}$$

$$Z = \left\{ \frac{s_1 - q_S}{\sigma_S}, \frac{s_2 - q_S}{\sigma_S}, ..., \frac{s_N - q_S}{\sigma_S} \right\}, where\ q_S = \frac{1}{n}\sum_{i=1}^{n} S_i,\ \sigma_S = \sqrt{\sum_{i=1}^{n}(S_i - q_S)^2} \tag{12}$$

where:

Z : The resulting simulation array.

s_i : The i-th value of the degree of presence of substantive-content message properties in the information background.

q_S : The average value of the degree of presence of substantive-content message properties in the information background.

Each resulting array Z represents a variation in the distribution of the degree of presence of substantive-content message properties within the information background. The properties of the distribution of the generated arrays Z are correlated with the initial array R, describing the presence of substantive-content message properties in the information background of the enterprise. The comparison is made based on the Mean Absolute Error (MAE) between two rank vectors obtained from the distribution histograms. The number of bins in the histogram determines the level of abstraction of the form and is expressed by the discretization coefficient DC. Thus, the histogram for the data set Z with the number of bins DC, where each bin is denoted as B_p for k = 1, 2, ..., DC. The frequency of data entering each bin B_p, denoted as $C_Z(B_p)$, and defined as (see equation 13):

$$C_Z(B_p) = \sum_{i=1}^{n} 1(Z_i \epsilon B_p), C_R(B_p) = \sum_{i=1}^{n} 1(r_i \epsilon B_p), p = 1, 2, ..., DC \tag{13}$$

Where $1(Z_i \epsilon B_p)$ — is the indicator function, which is equal to 1, if Z_i is within the range of bin B_p, and 0 otherwise. After the frequencies for all bins $C_Z(B_p)$ are calculated, the ranks of each bin can be determined. Let C_Z represent the sequence of frequencies for the bins. First, a sequence is created that results from sorting C_Z in ascending order. Denote this action as $S(C_Z)$, where each element $C_Z(B_p)$ is assigned its index in the sorted sequence. Then, the rank of each bin $R_Z(B_p)$ is defined as the index of its frequency $C_Z(B_p)$ in $S(C_Z)$. Thus, the bin with the lowest frequency is assigned the lowest rank, and so on sequentially for all other frequencies as they increase (see equation 14).

$$R_Z(B_p) = index(C_Z(B_p)\ in\ S(C_Z)), R_R(B_p) = index(C_R(B_p)\ in\ S(C_R)) \tag{14}$$

This process of calculating the rank of each bin in the histogram allows for the formation of comparable sequences, based on which the MAE is then calculated (see equation 15).

$$MAE = \frac{1}{DC} \sum_{k=1}^{DC} |R_Z(B_p) - R_R(B_p)| \tag{15}$$

Thus:

$C_Z(B_p)$	:	The frequency of data entering each bin B_p for the generative set Z.
$C_R(B_p)$	:	The frequency of data entering each bin B_p for the actual set R.
$I(z_i \in B_p)$	:	The indicator function, equal to 1 if z_i is within the range of bin B_p, and 0 otherwise.
B_p	:	A specific bin of the histogram.
DC	:	The discretization coefficient, determining the number of bins in the histogram.
$S(C_Z)$	:	The sequence C_Z sorted in ascending order.
$S(C_R)$	:	The sequence C_R sorted in ascending order.
$R_R(B_p)$	:	The rank of each bin in the histogram of the actual set R.
$R_Z(B_p)$	:	The rank of each bin in the histogram of the generative set Z.
MAE	:	The Mean Absolute Error between the two rank vectors obtained from the histograms of distributions Z and R.

The array of MAE values that has been formed allows for the ranking of the generated pairs of Corr and k according to the level of approximation quality. The subset of the most efficacious pairs constitutes the initial segment of the novel population. The second part of the new population is formulated based on the mutation of the first, implying a change in each of the values of the first part within a range of 20% with a probability of 50% (see equations 16).

$$F.C_i = \left\{ MAE_i < \left(\frac{1}{P.S.} \sum_{j=1}^{P.S.} MAE_j \right) \right\}, \quad \forall i \in \{1, 2, \ldots, P.S.\}$$

$$(Corr'_i, k'_i) = \begin{cases} (Corr_i, k_i), & Prob = 0.5 \\ (Corr_i \cdot (1 \pm rand[0, 0.2]), k_i \cdot (1 \pm rand[0, 0.2])), & Prob = 0.5 \end{cases}$$

$$Corr = [Corr_i, \ldots, Corr_m, Corr'_i, \ldots, Corr'_m]$$

$$k = [k, ..., k_m, k'_i, ..., k'_m] \tag{16}$$

The updated population is redirected to the simulation modeling stage. This process is carried out for the required number of generations (K.T.), as a result of which the pair Corr and k with the lowest MAE value is identified (see equation 17):

$$(\text{Corr}^*, k^*) = \min_{(Corr_i, k_i) \in F.C.} MAE(Corr_i, k_i) \tag{17}$$

The values of Corr* and k* are potentially the most effective in reflecting the characteristics of the distribution of the presence of substantive-content message properties in the information background. The developed algorithm is presented in Figure 2.

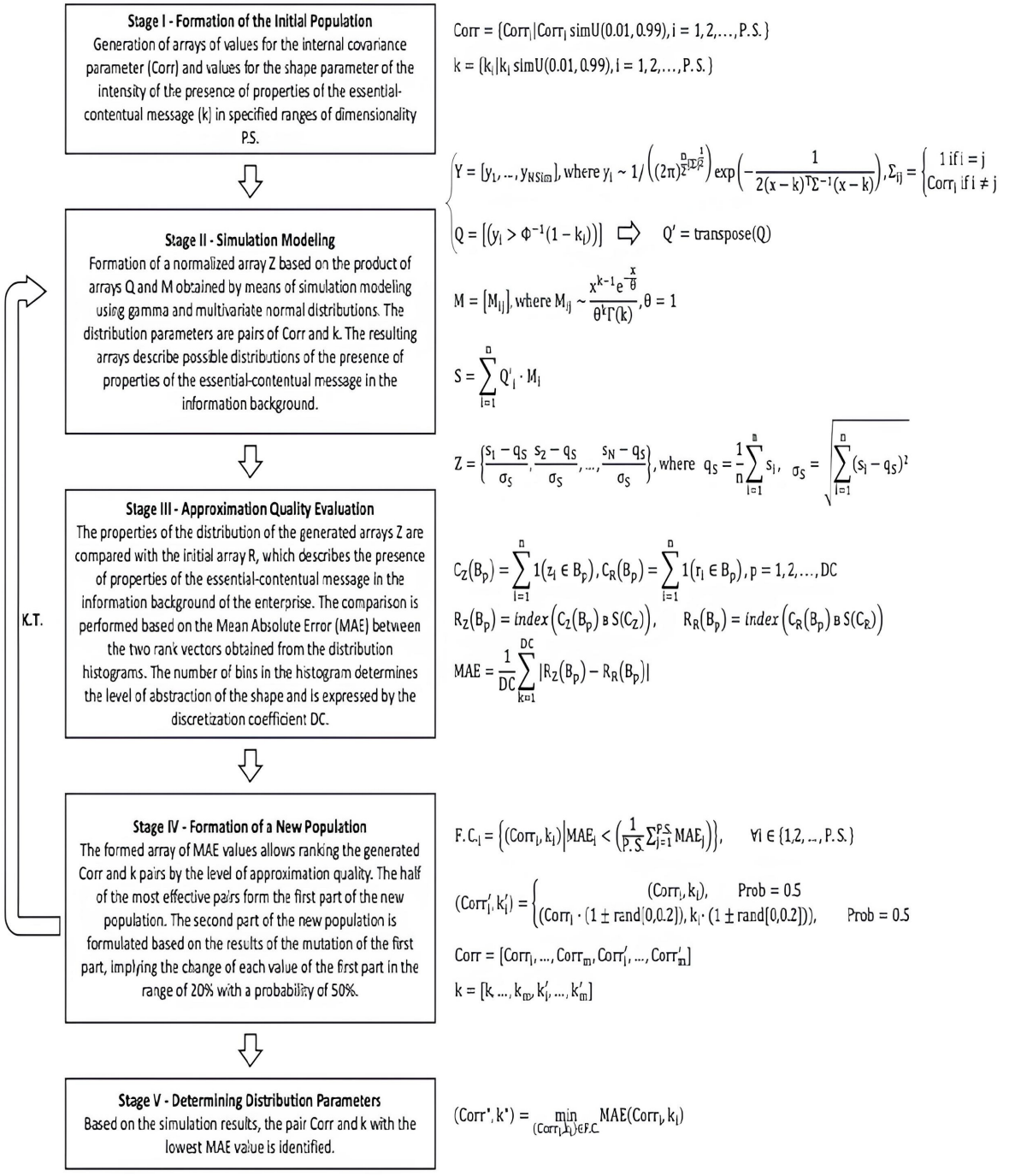

Figure 2 Flowchart of the research methodology an Electric Vehicle Acceptance Model in Indonesia

The analytical data obtained from the reconciliation of this algorithm will enable conclusions regarding the intensity and specificity of particular substantive-content message properties within the information background.

The model can be validated using real-world datasets from domains such as social media or political communication, where multimodal information patterns are prominent (Biggers et al., 2023; Lytras et al., 2020). For instance, analyzing the distribution of substantive-content message properties in datasets derived from online interactions or advertising campaigns could

showcase the model's capability in real-world scenarios (Boerman et al., 2021). Comparing model results with real data demonstrates the model's accuracy.

A comparative review of similar approaches, such as those utilizing Latent Dirichlet Allocation (LDA) (Constantiou and Kallinikos, 2015) or clustering algorithms like K-Means and DBSCAN (Lytras et al., 2020; Cawsey and Rowley, 2016), will highlight the improvements introduced by integrating Gamma and multivariate normal distributions optimized via genetic algorithms. These enhancements, including better handling of multimodal distributions and complex thematic interdependencies, can position the proposed method as a significant advancement over existing methodologies (Goldberg, 2022).

The proposed model is particularly suited for analyzing datasets from social media and political news, where the information background often exhibits multimodal characteristics. Social media platforms generate vast amounts of user-generated content, including posts, comments, and interactions, which can be used to detect dominant themes, measure emotional intensity, and identify patterns in user engagement (Biggers et al., 2023). Similarly, datasets from political news, which often reflect highly polarized and event-driven dynamics, provide an opportunity to validate the model's ability to capture the intensity and frequency of recurring themes (Boerman et al., 2021).

The methodology proposed in this study offers several distinct advantages compared to existing alternatives for analyzing the presence of substantive-content message properties in information backgrounds. Unlike traditional tools such as Latent Dirichlet Allocation (LDA) or K-Means clustering, which often treat information properties in isolation or assume linear separability in data, the developed model integrates both the Gamma distribution for intensity and the multivariate normal distribution for frequency. This dual-probabilistic framework enables a more nuanced representation of multimodal distributions, capturing both the depth (intensity) and recurrence (frequency) of information properties.

Moreover, existing approaches, such as LDA, are limited in their ability to handle the interdependencies between properties and often fail to account for contextual relevance. In contrast, the proposed model explicitly incorporates interrelationships through the covariance matrix in the multivariate normal distribution, allowing it to analyze correlations and co-occurrence patterns more effectively. Additionally, while methods like DBSCAN and Mean-Shift clustering excel in identifying data density and clusters, they do not account for the semantic or thematic content of the information. The developed methodology bridges this gap by integrating genetic algorithms for parameter optimization, ensuring adaptability to complex, multidimensional data structures and improving model accuracy.

Furthermore, the model's robustness to noise, outliers, and incomplete data provides a significant advantage over alternative tools that often rely on clean and structured datasets. This feature ensures reliable performance in real-world scenarios, such as social media analysis or political communication, where data is inherently noisy and incomplete.

For instance, the model can be applied to social media data to identify key topics and assess their saturation and recurrence in different communities or demographics. In political news analysis, the methodology can track shifts in narrative focus or sentiment over time, highlighting the correlation between specific topics and audience reactions. By leveraging the Gamma distribution for intensity and the multivariate normal distribution for frequency, the model can effectively represent the diverse properties of information dynamics in these contexts.

The proposed methodology is designed to address common challenges in information environments, including noise, outliers, and incomplete data, ensuring robustness and reliability. By employing the Gamma distribution for intensity modeling, the approach inherently smooths over random fluctuations in data, minimizing the impact of noise through its shape and scale parameters. Outliers are effectively managed through the multivariate normal distribution, where the covariance matrix identifies and reduces the influence of anomalous data points by accounting for established correlations. Similarly, the shape parameter (k) in the Gamma distribution adjusts for dispersion, mitigating the effect of extreme intensity values. To handle incomplete data,

the methodology incorporates imputation techniques during preprocessing, maintaining the integrity of the data array for analysis. Additionally, the genetic algorithm's iterative optimization process is resilient to missing values, as it seeks global optima across the parameter space, compensating for data gaps. These features collectively enable the model to deliver accurate and meaningful insights even in noisy, outlier-prone, or incomplete data environments.

4. Conclusions

This study has presented a novel methodology for modeling the distribution of substantive-content message properties in the information background. The methodology leverages the concept of multimodality in the distribution of these properties, which is characterized by peaks of varying intensity and frequency. The intensity of the presence of a property is defined as the depth or saturation of the information signal associated with it, while the frequency represents the number of times it appears in the information background. To model these aspects, we propose a combined approach using the Gamma and multivariate normal distributions. The Gamma distribution effectively captures the intensity of presence, allowing us to estimate the expected number of occurrences and the time scale of these events. Meanwhile, the multivariate normal distribution accounts for the frequency of presence, capturing the complex relationships between different substantive-content message properties. A key advantage of this approach is its ability to model the interplay between intensity and frequency, thus providing a more comprehensive understanding of the information background. To effectively identify the parameters of these distributions, we employed a genetic algorithm approach, enabling us to explore the multidimensional parameter space and identify global optima. Future research can further enhance this methodology by exploring additional distributions to model different types of substantive-content message properties, developing more sophisticated approaches for handling complex interdependencies between properties, and investigating the potential applications of this methodology in various domains, such as social media analysis, political communication, and marketing. Overall, the proposed methodology represents a valuable tool for understanding the complex interplay of substantive-content message properties within the information background. It offers a robust framework for analyzing and quantifying information dynamics, paving the way for deeper insights and more informed decision-making in various domains.

Acknowledgements

The research is financed as part of the project "Development of a methodology for instrumental base formation for analysis and modeling of the spatial socio-economic development of systems based on internal reserves in the context of digitalization" (FSEG-2023-0008).

Author Contributions

Methodology, Evgenii Konnikov; conceptualization, Dmitriy Rodionov; validation, Darya Kryzhko; data curation, Darya Kryzhko; writing—original draft preparation, Evgenii Konnikov; writing—review and editing, Darya Kryzhko; supervision, Dmitriy Rodionov; project administration, Dmitriy Rodionov; funding acquisition, Dmitriy Rodionov.

Conflict of Interest

The authors declare no conflicts of interest.

References

Amur, Z., Hooi, Y., Bhanbhro, H., Dahri, K., & Soomro, G. (2023). Short-text semantic similarity (stss): Techniques, challenges and future perspectives. *Applied Sciences*, *13*(6), 3911. https://doi.org/10.3390/app13063911

Anisiforov, A., Dubgorn, A., & Lepekhin, A. (2017). Organization of enterprise architecture information monitoring [Conference proceeding: E3S Web of Conferences, Vol. 110, 02051 (2019)]. *Proceedings of the 29th International Business Information Management Association Conference*, 2920–2930. https://doi.org/10.1051/e3sconf/201911002051

Bashir, H., Ojiako, U., Marshall, A., Chipulu, M., & Yousif, A. (2022). The analysis of information flow interdependencies within projects. *Production Planning & Control*, *33*(1), 20–36. https://doi.org/10.1080/09537287.2020.1821115

Berawi, M., Sari, M., Salsabila, A., Susantono, B., & Woodhead, R. (2022). Utilizing building information modelling in the tax assessment process of apartment buildings. *International Journal of Technology*, *13*(7), 1515–1526. https://doi.org/10.14716/ijtech.v13i7.6188

Beth, E. (2012). *Formal methods: An introduction to symbolic logic and to the study of effective operations in arithmetic and logic*. Springer Science & Business Media.

Bharadwaj, A. (2000). A resource-based perspective on information technology capability and firm performance: An empirical investigation. *MIS Quarterly*, *24*(1), 169–196. https://doi.org/10.2307/3250983

Biggers, F., Mohanty, S., & Manda, P. (2023). A deep semantic matching approach for identifying relevant messages for social media analysis. *Scientific Reports*, *13*(1), 12005. https://doi.org/10.1038/s41598-023-38761-y

Boerman, S., Kruikemeier, S., & Bol, N. (2021). When is personalized advertising crossing personal boundaries? how type of information, data sharing, and personalized pricing influence consumer perceptions of personalized advertising. *Computers in Human Behavior Reports*, *4*, 100144. https://doi.org/10.1016/j.chbr.2021.100144

Bruce, N., Murthi, B., & Rao, R. (2017). A dynamic model for digital advertising: The effects of creative format, message content, and targeting on engagement. *Journal of Marketing Research*, *54*(2), 202–218. https://doi.org/10.1509/jmr.14.0117

Cappella, J., & Li, Y. (2023). Principles of effective message design: A review and model of content and format features. *Asian Communication Research*, *20*(3), 147–174. https://doi.org/10.20879/acr.2023.20.023

Cawsey, T., & Rowley, J. (2016). Social media brand building strategies in b2b companies. *Marketing Intelligence & Planning*, *34*(6), 754–776. https://doi.org/10.1108/MIP-04-2015-0079

Clark, D., Hunt, S., & Malacaria, P. (2007). A static analysis for quantifying information flow in a simple imperative language. *Journal of Computer Security*, *15*(3), 321–371. https://doi.org/10.3233/JCS-2007-15302

Constantiou, I., & Kallinikos, J. (2015). New games, new rules: Big data and the changing context of strategy. *Journal of Information Technology*, *30*(1), 44–57. https://doi.org/10.1057/jit.2014.17

Copeland, B., & Fan, Z. (2023). Turing and von neumann: From logic to the computer. *Philosophies*, *8*(2), 22. https://doi.org/10.3390/philosophies8020022

Damanpour, F. (2018). Organizational innovation: A meta-analysis of effects of determinants and moderators. *Organizational Innovation*, *34*(3), 127–162. https://doi.org/10.5465/256406

Dehaene, S., & Cohen, L. (2011). The unique role of the visual word form area in reading. *Trends in Cognitive Sciences*, *15*(6), 254–262. https://doi.org/10.1016/j.tics.2011.04.003

Eremina, I., Yudin, A., Tarabukina, T., & Oblizov, A. (2022). The use of digital technologies to improve the information support of agricultural enterprises. *International Journal of Technology*, *13*(7), 1393–1402. https://doi.org/10.14716/ijtech.v13i7.6184

Goldberg, Y. (2022). *Neural network methods for natural language processing*. Springer. https://doi.org/10.1007/978-3-031-02165-7

Hitt, M., Hoskisson, R., & Ireland, R. (2019). *Strategic management: Concepts and cases: Competitiveness and globalization*. Cengage Learning.

Hsieh, Y.-C., & Chen, K.-H. (2011). How different information types affect viewer's attention on internet advertising. *Computers in Human Behavior*, *27*(2), 935–945. https://doi.org/10.1016/j.chb.2010.11.019

Kushwaha, A., Kar, A., & Dwivedi, Y. (2021). Applications of big data in emerging management disciplines: A literature review using text mining. *International Journal of Information Management Data Insights*, *1*(2), 100017. https://doi.org/10.1016/j.jjimei.2021.100017

Li, P., Jinyi, L., Xiangyi, L., Yao, X., & Ziguo, W. (2024). Digitalization as a factor of production in china and the impact on total factor productivity (tfp). *Systems*, *12*(5), 164. https://doi.org/10.3390/systems12050164

Li, Y., Chang, Y., & Liang, Z. (2022). Attracting more meaningful interactions: The impact of question and product types on comments on social media advertisings. *Journal of Business Research*, *150*, 89–101. https://doi.org/10.1016/j.jbusres.2022.05.085

Li, Z., Wenquan, W., & Qingguo, M. (2023). The adjustment of pressure perception in e-government response: A perspective of the political system theory. *Systems*, *11*(3), 158. https://doi.org/10.3390/systems11030158

Lim, S., & Schmälzle, R. (2023). Artificial intelligence for health message generation: An empirical study using a large language model (llm) and prompt engineering. *Frontiers in Communication*, *8*, 1129082. https://doi.org/10.3389/fcomm.2023.1129082

Lytras, M., Visvizi, A., Zhang, X., & Aljohani, N. (2020). Cognitive computing, big data analytics and data driven industrial marketing. *Industrial Marketing Management*, *90*, 663–666. https://doi.org/10.1016/j.indmarman.2020.03.024

Mueller, M., & Grindal, K. (2019). Data flows and the digital economy: Information as a mobile factor of production. *Digital Policy, Regulation and Governance*, *21*(1), 71–87. https://doi.org/10.1108/DPRG-08-2018-0044

Nadkarni, P., Ohno-Machado, L., & Chapman, W. (2011). Natural language processing: An introduction. *Journal of the American Medical Informatics Association*, *18*(5), 544–551. https://doi.org/10.1136/amiajnl-2011-000464

Nishiyama, A., Shigenori, T., Jack, A. T., & Roumiana, T. (2024). Holographic brain theory: Super-radiance, memory capacity and control theory. *International Journal of Molecular Sciences*, *25*(4), 2399. https://doi.org/10.3390/ijms25042399

Nugraha, E., Sari, R., Sutarman, Yunan, A., & Kurniawan, A. (2022). The effect of information technology, competence, and commitment to service quality and implication on customer satisfaction. *International Journal of Technology*, *13*(4), 827–836. https://doi.org/10.14716/ijtech.v13i4.3809

Öberg, C. (2023). Neuroscience in business-to-business marketing research: A literature review, co-citation analysis and research agenda. *Industrial Marketing Management*, *113*, 168–179. https://doi.org/10.1016/j.indmarman.2023.06.004

Omar, Y., & Peter, P. (2020). A survey of information entropy metrics for complex networks. *Entropy*, *22*(12), 1417. https://doi.org/10.3390/e22121417

Onykiy, B., Antonov, E., Artamonov, A., & Tretyakov, E. (2020). Information analysis support for decision-making in scientific and technological development. *International Journal of Technology*, *11*(6), 1125–1135. https://doi.org/10.14716/ijtech.v11i6.4465

Qiu, Q., Hao, Z., & Jiang, L. (2022). Strategic information flow under the influence of industry structure. *European Journal of Operational Research*, *298*(3), 1175–1191. https://doi.org/10.1016/j.ejor.2021.08.041

Rodionov, D., Gracheva, A., Konnikov, E., Konnikova, O., & Kryzhko, D. (2022). Analyzing the systemic impact of information technology development dynamics on labor market transformation. *International Journal of Technology*, *13*(7), 1548–1557. https://doi.org/10.14716/ijtech.v13i7.6204

Rodionov, D., Kryzhko, D., Tenishev, T., Uimanov, V., Abdulmanova, A., Kvikviniia, A., & Konnikov, E. (2022). Methodology for assessing the digital image of an enterprise with its industry specifics. *Algorithms*, *15*(6), 177. https://doi.org/10.3390/a15060177

Rodionov, D., Zaytsev, A., Konnikov, E., Dmitriev, N., & Dubolazova, Y. (2021). Modeling changes in the enterprise information capital in the digital economy. *Journal of Open Innovation: Technology, Market, and Complexity*, *7*(3), 166. https://doi.org/10.3390/joitmc7030166

Smýkal, P., & Eric, W. (2023). A commemorative issue in honor of 200th anniversary of the birth of gregor johann mendel: The genius of genetics. *International Journal of Molecular Sciences*, *24*(14), 11718. https://doi.org/10.3390/ijms241411718

Stapleton, G., Howse, J., & Rodgers, P. (2010). A graph theoretic approach to general euler diagram drawing. *Theoretical Computer Science*, *411*(1), 91–112. https://doi.org/10.1016/j.tcs.2009.09.005

Vigo, R. (2013). Complexity over uncertainty in generalized representational information theory (grit): A structure-sensitive general theory of information. *Information*, *4*(1), 1–30. https://doi.org/10.3390/info4010001

Wang, S., Zhang, Y., Shi, W., Zhang, G., Zhang, J., Lin, N., & Zong, C. (2023). A large dataset of semantic ratings and its computational extension. *Scientific Data*, *10*(1), 106. https://doi.org/10.1038/s41597-023-01995-6

Weaver, W. (2017). The mathematics of communication. In *Communication theory* (pp. 27–38). Routledge.

Weinlich, P., & Semerádová, T. (2022). Emotional, cognitive and conative response to influencer marketing. *New Techno Humanities*, *2*(1), 59–69. https://doi.org/10.1016/j.techum.2022.07.004

Zaytsev, A., Rodionov, D., Dmitriev, N., & Kichigin, O. (2019). Comparative analysis of results on application of methods of intellectual capital valuation. *International Scientific Conference: Digital Transformation on Manufacturing, Infrastructure and Service (DTMIS 2019)*.