UDC 004.415.2
DOI https://doi.org/10.32782/2663-5941/2026.1.2/55

Shupyliuk M.V.
https://orcid.org/0009-0005-3457-9895
Kharkiv National University of Radio Electronics

Martovytskyi V.O.
https://orcid.org/0000-0003-2349-0578
Kharkiv National University of Radio Electronics

MODEL OF INFORMATION TECHNOLOGY FOR AUTOMATED HANDWRITING ANALYSIS

Today, handwriting analysis remains an important tool for verifying the authenticity of documents and understanding a person's personality, but this process requires modern digital tools to increase the speed and accuracy of data processing. Such automation plays a special role in such areas as forensics, psychology and personnel management, as well as in the processing of historical archives. Traditional methods of expert assessment often suffer from subjectivity and low data processing speed, which creates a need to create highly efficient intelligent systems. The object of the study is the process of automated analysis of handwritten text. The subject of the study is the model of information technology for automated writer identification and psychological profiling based on images of handwritten text. The paper presents an information technology model based on the integration of computer vision methods, in particular the Vision Transformers architecture for writer identification and psychological profiling of a personality according to the Big Five model. The structure of the system is mathematically formalized in the form of a tuple of sets and transformation functions. The entire process was covered: from personality verification and assessment of character traits to report generation. To protect the integrity of images and prevent unauthorized manipulation, steganographic protection methods, namely digital watermarks, are implemented. The proposed approach minimizes the subjectivity of expert opinions, integrates psychological profiling, and ensures high accuracy through the use of advanced computer vision methods. In the future, the technology can be expanded by integrating multimodal biometric data processing modules, such as voice and dynamic facial parameters, to create a universal identity verification system.

Keywords: *information technology, model, handwriting, digital watermarking, image, mathematical model.*

Formulation of the problem. Development of handwritten text analysis information systems can become a tool for solving urgent problems in various spheres of human activity.

In the realm of forensics, traditional forensic examination of handwriting largely depends on human perception, which leads to subjectivity of conclusions [1]. Priority direction of development of the industry is the introduction of automated handwriting identification system [2]. Integration of specialized software allows to minimize the influence of the human factor and provides a transition to mathematically based verification of the authenticity of the details. Such systems provide high accuracy of writer identification in real time, allow for operational screening of large amounts of documentation and effectively detect forgeries made with the help of modern technical means, such as plotters or autopen.

Another critical application lies in the processing of historical archives, where manual manuscript analysis remains a significant bottleneck. Manuscripts can be very different types of texts: letters from famous personalities, relatives or ancestors, novels or poems, diaries or legal reports [3]. Working with historical documents includes solving problems of transcribing the text as well as attribution of anonymous manuscripts, identification of writers of letters, etc.

Furthermore, the application of handwriting analysis extends to the behavioral sciences. The psychology and personnel management sector is interested in methods for quickly assessing the psychological portrait of candidates. Analysis is considered a tool for identifying "soft skills", assessing stress resistance, predisposition to leadership or even destructive behavior. Integration of automated handwriting analysis in this area can enable real-time assessment and feedback, expanding its accessibility and usability in diverse settings [4].

Therefore, development of new technologies for automated handwriting analysis is relevant.

Analysis of recent research and publications. The field of handwriting analysis software is clearly segmented into two groups: professional forensic systems (Forensic Grade) and commercial/scientific tools for psychological profiling.

Internationally, there are systems that include both automatic pattern recognition functions and interactive tools that help forensic experts. FISH is considered the ancestor of automated handwriting analysis systems. Developed by the German Federal Criminal Police Office (BKA) in the 1990s, it was adapted and widely used by the US Secret Service to analyze threatening letters to presidents and high-ranking officials [5]. Since it is a decision support system it does not make a final conclusion, but only narrows the search for the expert. Lately, the modernization of FISH was announced by the US Secret Service [6]. The idea behind modernization lies in the belief that using the latest AI algorithms will speed up the search [6]. Even though modernization takes place, FISH is still considered a decision support tool. Besides that, underlying original architecture does not account for the psychological aspect of the writing and focuses mainly on pattern matching. As the evolution of FISH, a new handwriting analysis and writer identification system called WANDA [7] was developed. The WANDA client-server platform is an open system for forensic handwriting and signature examination, which supports the integration of third-party modules and various writing systems through a plugin architecture. The program allows you to automate sample collection, allograph recognition, feature extraction and personal identification based on digitized data, using the WandaML XML standard for document description. Using image import filters (in particular, via the IBIS software) and a convenient graphical interface for expert notes, the platform combines traditional forensic methods with innovative developments of independent researchers into a single working environment. The Center for Excellence in Document Analysis (CEDAR) at the University at Buffalo (USA) is developing a forensic document analysis system known as CEDAR-FOX [8]. CEDAR-FOX analysis process goes from background removal and until Likelihood Ratio determination. It also supports a wide range of text entries including full pages of text and short entries. This document analysis system analyzes both macro-features such as entropy and threshold and micro-features e.g. character structure and supports tools for signature verification. Although it works great at micro-feature analysis and Likelihood Ratio determinations, CEDAR-FOX is a purely deterministic tool that treats handwriting as a physical trace without consideration of psychological projection. The Center for Statistics and Applications in Forensic Evidence (CSAFE) is developing an open source package for the R language. It uses Rainbow Triangle Graph Decomposition and stochastic clustering (k-means) to create groups of glyphs. Since writer identification is cluster-based, the chance to identify correctly depends on how often handwriting elements occur in these clusters. Cluster architecture has its own downsides, for example results may vary depending on sample size or writers ill intents like handwriting disguise.

In Ukraine, forensic handwriting examination is regulated by the Criminal Procedure Code and instructions of the Ministry of Justice. The main subjects are state specialized institutions, such as the Kyiv Scientific Research Institute of Forensic Expertise (KSRIFE). Although the basis of the examination conducted by Ukrainian experts remains traditional (comparative analysis of general and individual features by an expert), tools are actively being introduced. In particular, the use of video spectral comparators and Regula Forensic Studio software is a standard for technical examination of documents and analysis of signatures for technical forgery. On the other hand the reliance on comparative analysis of "general and individual features" is labor-intensive and lacks predictive power.

As for commercial/scientific tools for psychological profiling there is the AvgMISC system that uses a hybrid of SVM and CNN to analyze offline handwriting in different languages (English, Arabic, Chinese) [9]. The authors claim correlation of the results with psychological tests, but these studies often have limited samples. Also there is a Nyckel platform that offers an API for image classification, including the Handwriting Personality Quiz Identifier model. It is positioned as a tool for HR, marketing, and education, providing an assessment of the "confidence" of the model in the presence of a particular trait.

Taking into account the current state it becomes clear that the current landscape is split between deterministic forensic tools that view script purely as a physical trace and flexible but unreliable psychological models and this work proposes information technology for automated handwriting analysis that combines forensic examination with the behavioral insights of psychological profiling into a single framework.

Task statement. The work aims to develop a model of information technology for automated handwriting analysis. It should perform identification of the writer and his psychological profiling based on

provided image samples. To achieve the aim, the following tasks were set:

1. Build a functional model of information technology for automated handwriting analysis.

2. Perform formalization of the work through the construction of a mathematical model.

Outline of the main material of the study. The following general scheme of information technology for automated handwriting analysis is proposed, Fig. 1. At the center of the system is a processing unit that interacts with several data streams: input parameters include user identification data and an image of handwritten text; management and methodological support is carried out through the use of writer identification methods, detection of personal characteristics and the introduction of digital watermarks to protect information. The resource base of the system is based on a secure file storage and user data server. The result of the technology is the output products, namely: a detailed analysis report, a list of identified handwriting characteristics and an array of selected features for further processing.

The process of automated handwriting analysis is proposed to be carried out in the stages presented in the decomposition scheme of information technology Fig. 2. The process includes such steps as user authentication, storage and watermarking, document authorship determination, human character traits identification and report generation.

It is practical to look in detail at the stages shown in Fig. 2 and underlying the functioning of the information technology:

1) User authentication: The initial stage of verifying the subject's access rights to the system, which is carried out using data from the user server.

2) Storage and watermarking: The received images of handwritten text are processed using steganographic protection methods (digital watermarks) and are stored in a secure file storage.

3) Document authorship determination: Using writer identification methods to establish or confirm the identity of the writer of the submitted manuscript. Underlying method used for writer identification in details described in work [10] and uses Vision Transformers.

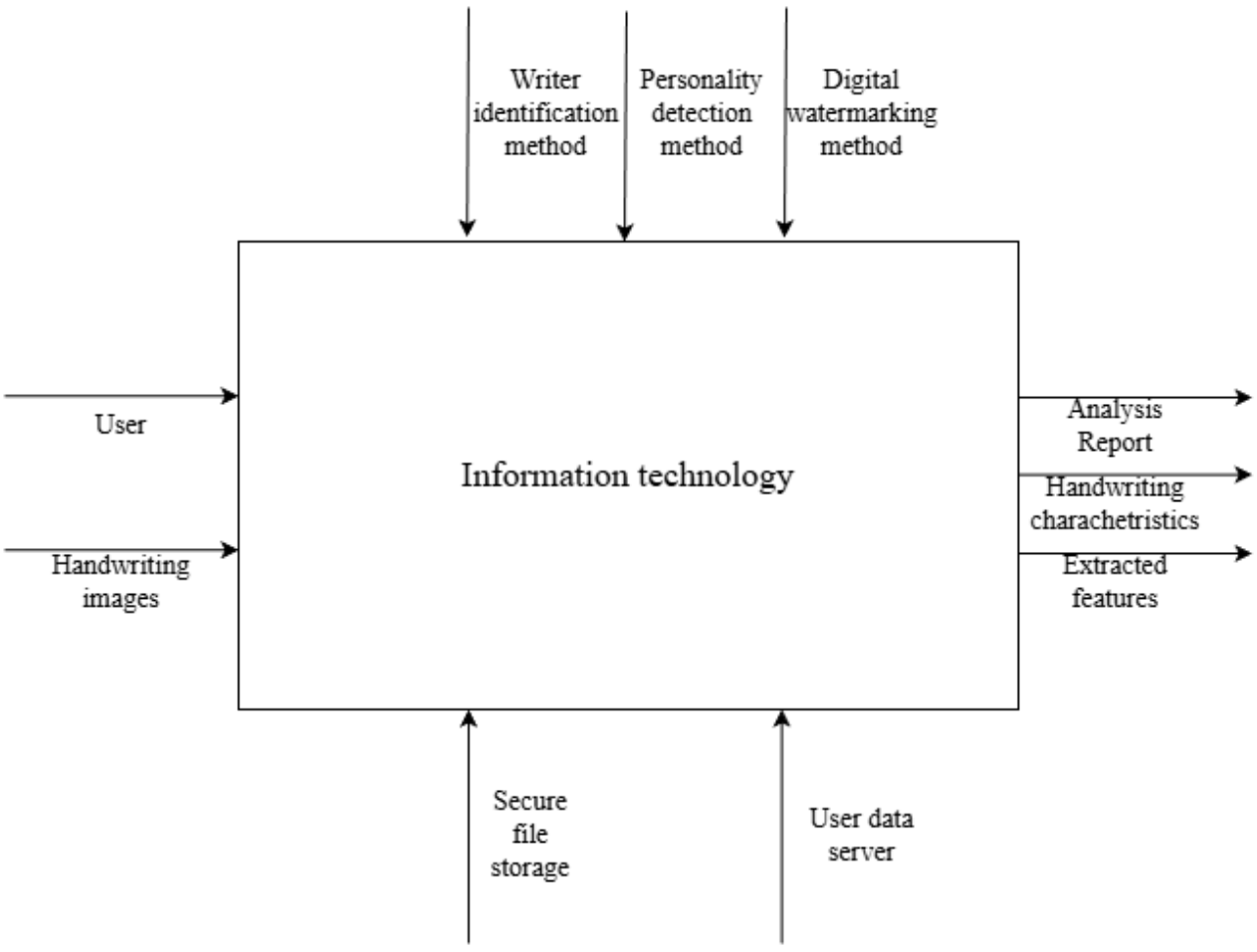

Fig. 1. General scheme of information technology

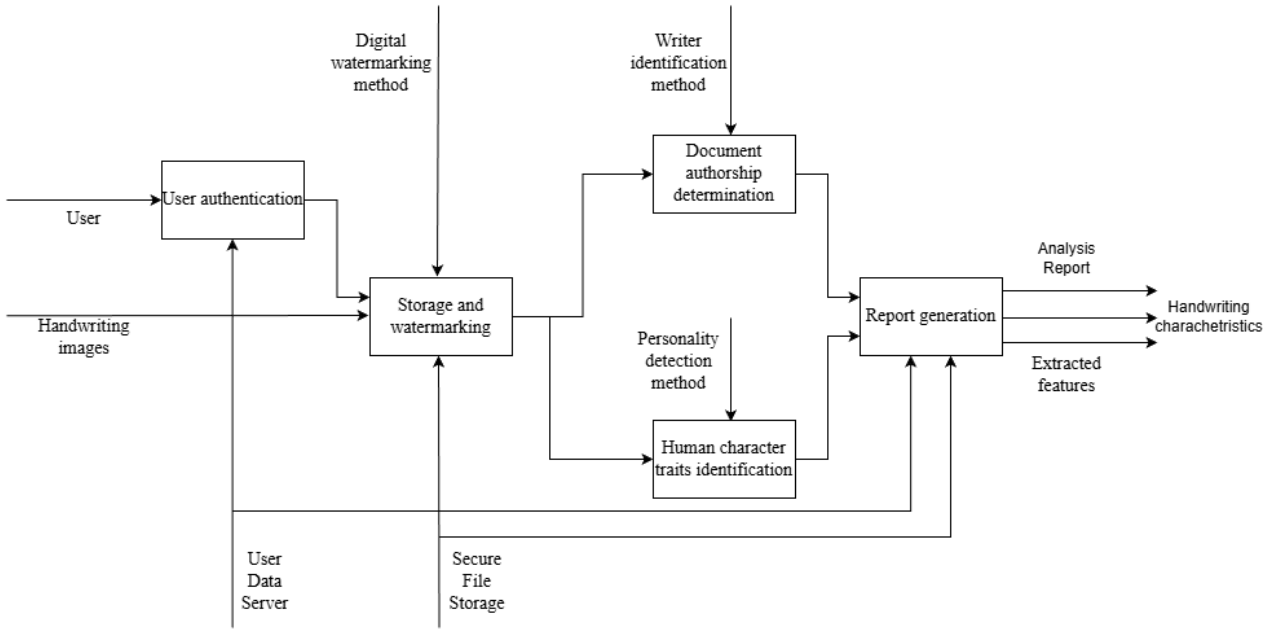

Fig. 2. Decomposition scheme of information technology

4) Human character traits identification: Psychological and graphological analysis of the text using personality detection methods to form a psychological portrait. Underlying method used for personality detection in detail described in work [11] and uses Big Five personality traits model.

5) Report generation: The final stage, at which data from all modules is synthesized and the final results are formed: an analytical report, a list of handwriting characteristics and an array of extracted features.

The formalization of the work of the proposed information technology is carried out through the construction of a mathematical model that describes the entire process – from collecting and storing input images to generating reports based on handwriting analysis algorithms.

System S can be represented as a tuple of sets of input data, output data, control mechanisms, and resources:

$$S=\{X,Y,M,R,F\} \quad (1)$$

where: X – set of input data;
Y – set of output results;
M – set of methods (controlling influences);
R – set of resources;
F – set of transformation functions.

Input data (X) can be represented as:

$$X=\{x_{user},x_{img}\} \quad (2)$$

where: x_{user} – user identification data;
x_{img} – digital images of handwritten text.

Methods and algorithms (M) can be represented as:

$$M=\{m_{wm},m_{id},m_{pi}\} \quad (3)$$

where: m_{wm} – Digital watermarking method;
m_{id} – Writer identification method;
m_{pi} – Personality detection method.

Resources (R)) can be represented as:

$$R=\{r_{srv},r_{store}\} \quad (4)$$

where: r_{srv} – user data server;
r_{store} – secure file storage.

Output results (Y) can be represented as:

$$Y=\{y_{rep},y_{char},y_{feat}\} \quad (5)$$

where: y_{rep} – Analysis Report;
y_{char} – Handwriting characteristics;
y_{feat} – Extracted features.

According to decomposition scheme Fig. 2 process of automated handwriting analysis can be presented as composition of functions $f_i \in F$.

Authentication function (f_1) – verifies the user using server data can be represented as:

$$a=f_1(x_{user},r_{srv}) \quad (6)$$

where: a – access token.

Storage and Marking Function (f_2) – processes images using watermarking and stores them can be represented as:

$$d_{sec}=f_2(x_{imgr},a,m_{wm},r_{store}) \quad (7)$$

where: d_{sec} – secure data.

Authorship determination function (f_3) – analyzes secure data to determine the writer can be represented as:

$$res_{id} = f_3\left(d_{sec}, m_{id}\right) \quad (8)$$

where: res_{id} – identification result.

Character traits identification function (f_4) – analyzes secure data to determine psychotype can be represented as:

$$res_{traits} = f_4\left(d_{sec}, m_{pi}\right) \quad (9)$$

where: res_{traits} – revealed psychological traits.

Report generation function (f_5) – Aggregates analysis results and accesses resources to generate final output data can be represented as:

$$Y = f_5\left(res_{id}, res_{traits}, r_{srv}, r_{store}\right) \quad (10)$$

By substituting intermediate functions, the overall operation of information technology can be described by the operator Φ:

$$Y = \Phi\left(X, M, R\right) = f_5\left(f_3\left(\Omega, m_{id}\right),\ f_4\left(\Omega, m_{pi}\right), r_{srv}, r_{store}\right) \quad (11)$$

where Ω – securing data result:

$$\Omega = f_2\left(x_{imgr}, f_1\left(x_{user}, r_{srv}\right), m_{wm}, r_{store}\right) \quad (12)$$

Based on the proposed conceptual and mathematical model, an information technology for automated handwriting analysis prototype was developed. The prototype implements the full cycle of data processing and system operation: from collecting and storing input images to generating reports based on handwriting analysis algorithms.

Conclusions. In this work information technology for automated handwriting analysis was developed and presented. The paper proposes an information technology that integrates forensic writer identification methods with psychological personality profiling. The proposed model is based on a combination of well-known personality traits model and modern machine vision methods, in particular the Vision Transformers architecture.

Mathematical formalization of the process in the form of a tuple of sets and transformation functions was done. It covers the whole system from secure storage of graphic information using digital watermarks to the generation of complex analytical reports, which ensures minimization of subjectivity and increased accuracy of analysis in real time.

Reliability of the system is achieved by implementation of steganographic protection methods that allows to confirm the integrity of documents and prevent unauthorized manipulation of the examination results. This approach makes the technology suitable for use in conditions of strict confidentiality and high requirements for data reliability.

In the future, the technology can be expanded by processing and using media files to verify identity using biometric data, such as voice or dynamic facial image. This will allow creating a universal ecosystem for identity verification that can adapt to changing requests in the areas of cybersecurity, forensics, and human resource management.

Bibliography:

1. Galekovic J., Tkalac Z., Ledic A. Forensic Examination of Handwriting and Documents. Forensic Science and Molecular Anthropology – Topics Selected from 12th ISABS Conference on Forensic and Anthropological Genetics. *IntechOpen,* Nov. 27, 2024. DOI: https://doi.org/10.5772/intechopen.115118

2. Hicklin R.A., Eisenhart L., Richetelli N., Miller M.D. et al. Accuracy and reliability of forensic handwriting comparisons. *Proc. Natl. Acad. Sci. U.S.A.* 2022. 119 (32). e2119944119. DOI: https://doi.org/10.1073/pnas.2119944119

3. Tikhonov A., Rabus A. Handwritten Text Recognition of Ukrainian Manuscripts in the 21st Century: Possibilities, Challenges, and the Future of the First Generic AI-based Model. *Kyiv-Mohyla Humanities Journal.* 2024. (11). P. 226–247. DOI: https://doi.org/10.18523/2313-4895.11.2024.226-247

4. Deore S., Kalokhe P., Patil S., Nagarkar S. et al. Identifying the personality traits using handwriting recognition in a real-time environment. *Ingénierie des Systèmes d'Information.* 2025. Vol. 30, No. 1. P. 203–211. DOI: https://doi.org/10.18280/isi.300117

5. Dikevych K.G. Application of the method of cybernetic modeling in foreign practice during the writing of forensic handwriting examinations. *Kriminalistyka i sudova ekspertyza*. 2021. Vol. 66. P. 766–771. DOI: https://doi.org/10.33994/kndise.2021.66.72

6. You've Been Gone: FSD Forensics Advancements. Behind the Shades. 2025. URL: https://www.secretservice.gov/newsroom/behind-the-shades/2025/02/youve-been-gone-fsd-forensics-advancements (date of access: 10.01.2026)

7. Erp M., Vuurpijl L., Franke K., Schomaker L. The WANDA Measurement Tool for Forensic Document Examination. J Forensic Doc Exam. 2003. Vol. 16

8. Srihari S., Huang C., Srinivasan H. Search engine for handwritten documents. Proc. SPIE 5676, Document Recognition and Retrieval XII. 2005. P. 66–75. DOI: https://doi.org/10.1117/12.585883

9. Garoot A., Suen C.Y. Measuring the Big Five Factors from Handwriting Using Ensemble Learning Model AvgMlSC. Intertwining Graphonomics with Human Movements: materials IGS 2022. Cham: Springer, 2022. Vol. 13424. DOI: https://doi.org/10.1007/978-3-031-19745-1_12

10. Shupyliuk M., Martovytskyi V., Romanenkov Y. Enhancing writer identification and writer retrieval with CenSurE and Vision Transformers. *Technology Audit and Production Reserves.* 2025. 6(2(86)). P. 6–14. DOI: https://doi.org/10.15587/2706-5448.2025.343943

11. Shupyliuk M., Martovytskyi V., Bolohova N., Romanenkov Y. et al. Devising an approach to personality identification based on handwritten text using a vision transformer. *Eastern-European Journal of Enterprise Technologies*. 2025. 1(2 (133)). P. 53–65. DOI: https://doi.org/10.15587/1729-4061.2025.322726

Шупилюк М.В., Мартовицький В.О. МОДЕЛЬ ІНФОРМАЦІЙНОЇ ТЕХНОЛОГІЇ АВТОМАТИЗОВАНОГО АНАЛІЗУ РУКОПИСНОГО ТЕКСТУ

Сьогодні аналіз почерку залишається важливим інструментом для перевірки справжності документів та розуміння особистості людини, проте цей процес потребує сучасних цифрових інструментів для підвищення швидкості та точності обробки даних. Така автоматизація відіграє особливу роль в таких сферах як криміналістика, психологія та управління персоналом, а також при обробці історичних архівів. Традиційні методи експертного оцінювання часто страждають від суб'єктивізму та низької швидкості обробки даних, що створює потребу у створенні високоефективних інтелектуальних систем. Об'єктом дослідження є процес автоматизованого аналізу рукописного тексту. Предметом дослідження є модель інформаційної технології автоматизованої ідентифікації автора та психологічного профілювання за рукописним текстом. У роботі представлено модель інформаційної технології, що базується на інтеграції методів комп'ютерного зору, зокрема архітектури Vision Transformers для ідентифікації автора та психологічного профілювання особистості за моделлю Big Five. Математично формалізовано структуру системи у вигляді кортежу множин та функцій перетворення. Було покрито весь процес: від верифікації особистості та оцінки рис характеру до формування звіту. Для захисту цілісності зображень і запобігання несанкціонованим маніпуляціям впроваджено використання стеганографічних методів захисту, а саме цифрових водяних знаків. Запропонований підхід мінімізує суб'єктивізм експертних висновків, інтегрує психологічне профілювання та забезпечує високу точність за рахунок використання передових методів комп'ютерного зору. У майбутньому технологію можна розширити шляхом інтеграції модулів обробки мультимодальних біометричних даних, таких як голос та динамічні параметри обличчя, для створення універсальної системи верифікації особи.

Ключові слова: *інформаційна технологія, модель, рукописний текст, цифровий водяний знак, зображення, математична модель.*

Дата першого надходження статті до видання: 22.01.2026
Дата прийняття статті до друку після рецензування: 23.02.2026
Дата публікації (оприлюднення) статті: 08.04.2026