BULLETIN of Cherkasy State Technological University

Journal homepage: https://bulletin-chstu.com.ua/en

Vol. 31 No. 1. 2026

UDC 004.738.5:001.8
DOI: 10.62660/bcstu/1.2026.73

ISSN: 2306-4412
E-ISSN: 2708-6070

Conceptual information model and methods for developing a personalised information service for researchers

Volodymyr Liubchak
PhD in Physics and Mathematics, Associate Professor
Sumy State University
40007, 116 Kharkivska Str., Sumy, Ukraine
https://orcid.org/0000-0002-7335-6716

Maksym Shovkoplias*
Postgraduate Student
Sumy State University
40007, 116 Kharkivska Str., Sumy, Ukraine
https://orcid.org/0009-0007-4192-4743

Abstract. Modern scientific activity takes place under conditions of growing information volumes and fragmented digital infrastructures. Researchers are compelled to rely on multiple independent platforms – such as search engines, bibliographic managers, analytical tools, and academic communication services – which complicates workflow organisation, causes redundancy, and reduces research efficiency. The purpose of this study was to develop a conceptual model of a personalised information service for researchers, capable of integrating search, analytical, bibliographic, and communication functions within a unified adaptive environment. The methodology involved information modelling, integration of open scientific data via application programming interfaces, and the formalisation of a user profile for the automated selection of relevant services. The resulting model was formalised as an informational quintuple $F = (P, S, D, I, U)$, where P represents the user profile module, S – the service selection module, D – the data integration module, I – the interface subsystem, and U – the analytics and adaptation subsystem. The user profile P was defined as a vector structure comprising theme-specific attributes, tool preferences, and languages. For each external service S_i, relevance $R(S_i, P)$ was computed using a normalisation function f_{norm} and a similarity metric Sim, enabling the construction of a personalised service configuration. A prototype implemented using Flutter and Firebase demonstrated the model's practical capacity to reduce the time required to locate relevant information and to enhance overall research productivity. The proposed model can serve as a foundation for adaptive digital platforms that promote open science and foster interdisciplinary collaboration

Keywords: adaptive digital platforms; scientific communication; user profile; data personalisation; integration of scientific resources; user interface; research support

INTRODUCTION

Modern scientific activity takes place under conditions of substantial information overload and increasing fragmentation of the digital research environment. Researchers are required to employ a wide range of separate digital tools – from search engines and bibliographic managers to analytical platforms and communication systems – in order to accomplish their daily research tasks. This fragmentation leads to duplicated actions, the dispersion of attention, and inefficiency in managing the research workflow. The relevance of the study lies in the

Article's History: Received: 27.10.2025; Revised: 09.02.2026; Accepted: 16.03.2026; Published: 08.04.2026

Suggested Citation:
Liubchak, V., & Shovkoplias, M. (2026). Conceptual information model and methods for developing a personalised information service for researchers. *Bulletin of Cherkasy State Technological University*, 31(1), 73-84. doi: 10.62660/bcstu/1.2026.73.

*Corresponding author

growing need to create integrated, personalised digital ecosystems that can unify scientific tools and adapt to the changing needs and interests of researchers.

Over the past decade, there has been a noticeable rise in scholarly attention towards personalised and adaptive information systems for researchers. According to T. Adewale (2022), the introduction of machine learning mechanisms into recommendation systems has significantly improved the automated selection of scientific publications and enhanced user engagement in academic databases. However, the author noted that most existing recommendation systems are focused solely on publication retrieval and rarely address broader aspects of personalisation such as interface adaptability or cross-platform integration. C.K. Kreutz & R. Schenkel (2022) explored the application of semantic search technologies for improving the contextual accuracy of information retrieval. Their research demonstrated that semantic indexing increases the relevance of search results through the analysis of conceptual relations between scientific terms. Nevertheless, the study also emphasised that the majority of these systems lack user-specific configuration options, which restricts their applicability for individualised research environments. Further development of adaptive systems has been discussed by Y. Fei *et al.* (2020), who examined interdisciplinary approaches to user interface personalisation. The authors concluded that adaptive interfaces can substantially simplify access to scientific data and streamline workflow efficiency. Yet, as they acknowledged, such approaches are typically confined to the interface layer and do not incorporate deeper integration of analytical or bibliographic functionalities.

The integration of open scientific data has been actively investigated in recent years. For instance, M. Umbach (2024) analysed how open access and FAIR data principles contribute to transparency and reproducibility in academic work. The author argued that open data infrastructures enable greater collaboration between researchers, but the study did not provide mechanisms for personalised filtering or adaptive data recommendation, which are essential for individual research efficiency. Similarly, L. Nikiforova *et al.* (2025) developed an information resource model for registering scientific journals in the context of digitalisation. They demonstrated the potential of application programming interface (API)-based integration but paid limited attention to the personalisation of such systems according to the researcher's professional profile. In a related context, C.A. Putri *et al.* (2025) explored the optimisation of digital library environments in the educational sector and emphasised the importance of integrating external resources through unified access points. Their findings highlighted the value of seamless access to diverse content sources for improving information retrieval efficiency. However, their approach was primarily centred on document access within fixed educational settings and did not propose mechanisms for personalising the integration of external services at the user level. The model developed in this study expands on this by enabling dynamic, profile-driven integration of scientific services through open APIs, thereby enhancing utility for individual researchers rather than institutional users.

H. Ko *et al.* (2022) conducted a comprehensive review of recommendation system architectures and highlighted hybrid models that combine semantic analysis, citation metrics, and user activity data. Their findings confirmed that hybrid personalisation improves the precision of scientific recommendations. Nonetheless, the authors admitted that many implementations remain experimentally limited and lack integration across multiple services – an issue particularly relevant for constructing unified research environments. Technological personalisation has also been explored through the prism of interface engineering. D. Petryna *et al.* (2024) examined the use of neural network tools for accelerating the development of adaptive web interfaces. Their research confirmed that AI-assisted adaptation enhances usability but also showed that such solutions are often narrowly domain-specific and not scalable to complex academic ecosystems. A similar limitation was observed by A. Carrera-Rivera *et al.* (2024), who proposed the AdaptUI framework for smart product-service systems. While their model effectively supports dynamic interface modification, it was not designed to manage the multifaceted data flows typical of scientific environments.

A relevant contribution to this field was made by researchers M.O. Shovkoplias & V.O. Liubchak (2024), who conducted an overview of existing approaches to the construction of personalised digital environments for researchers. Their study summarised current tools and services used for individualising academic workflows and outlined the conceptual foundations of integrating recommendation mechanisms, profile-based adaptation, and scientific communication tools. The authors concluded that the creation of individual working environments significantly reduces information overload during the research process. However, their work was limited to a theoretical analysis and did not propose a formalised information model or practical implementation framework capable of unifying external scientific services through adaptive personalisation. Collectively, these studies confirm a clear scientific trend: the movement from general-purpose digital systems towards personalised, adaptive research environments that integrate open data, recommendation mechanisms, and analytical tools. Yet, despite significant advances, most solutions remain fragmented – focusing on isolated functionalities such as semantic search, publication recommendations, or adaptive interfaces – and do not provide a comprehensive model that unifies these components around a formalised researcher profile.

The purpose of this study was to develop an integrated conceptual and informational model of a personalised service for researchers that consolidates search, analytical, bibliographic, and communication tools into a unified adaptive environment. Unlike existing fragmented solutions, the proposed model accounts for both the researcher's professional attributes (such as scientific domain, topics, and data sources) and behavioural patterns of interaction with digital resources. This allows for real-time adaptation of the service to evolving information needs and ensures seamless integration with open scientific ecosystems.

MATERIALS AND METHODS

The model was constructed on the basis of three principal data channels: declarative user data, behavioural activity traces, and metadata obtained from external scientific sources. Declarative data were gathered via structured input forms completed by researchers during initial system interaction. These forms included fields describing the user's scientific domain, thematic keywords, preferred languages, and digital tools used in their scientific workflow. This information was subsequently formalised into a parameter vector and stored within the system database for further computation. Behavioural data were collected through continuous monitoring of user-system interaction events. These included search queries, viewed abstracts or articles, time spent on topic-specific materials, and the frequency of use of particular services. Each event was converted into a numerical value and mathematically normalised using the function f_{norm}, which mapped raw behavioural indicators to the interval [0,1]. Formally, this function is defined as:

$$f_{norm}(x) = \frac{x - x_{min}}{x_{max} - x_{min}}, \tag{1}$$

where x – an observed behavioural value, and x_{min} and x_{max} – the minimum and maximum values for that metric across all users. This ensured uniform scaling across heterogeneous data types, enabling them to be combined into a unified behavioural vector.

The researcher profile was modelled as a multidimensional vector $P = (p_1, p_2, \dots, p_n)$ where each parameter p_i was associated with a weight $w_i \in [0,1]$ signifying its influence on the service selection mechanism. Weight updates were carried out using an exponential smoothing algorithm:

$$w_i^{(t+1)} = \alpha \cdot f_{norm}(x_i^t) + 1(1 - \alpha) \cdot w_i^{(t)}, \tag{2}$$

where α – the learning rate, and the normalised activity at time t. This mechanism allowed the profile to adapt dynamically to changing research interests.

External scientific sources such as arXiv, Semantic Scholar, and Scopus were accessed via open REST APIs. Metadata from these systems included attributes such as publication topics, supported languages, and service categories. For each candidate service S_j, a relevance score $R(S_j, P)$ was computed according to the formula:

$$R(S_j, P) = \sum_{i=1}^{n} w_i \cdot Sim(p_i, s_{j,i}), \tag{3}$$

where Sim – a domain-specific similarity measure computed differently depending on the nature of the data. For thematic similarity, the cosine similarity measure was used; for categorical parameters such as language or tool support, partial functions Dom, Lang, and Tool were introduced, returning binary or graded similarity values in [0,1]. For example:

$$Lang(P, S_j) = \begin{cases} 1, if\ language\ matches \\ 0, otherwise \end{cases}. \tag{4}$$

All collected and computed information was stored in Firebase Firestore using structured JSON documents, enabling seamless retrieval and integration during computational workflows. The prototype interface was implemented in Flutter to demonstrate core data flow and logic, although interface design was not a focus of the current study. To evaluate the quantitative performance of the proposed model, a series of simulation experiments were conducted using 200 synthetic researcher profiles. These profiles represented diverse combinations of scientific domains, tool preferences, and information needs, synthesised from real-world patterns observed in open academic resources. The model's relevance computation was compared against a manually defined ground-truth mapping of optimal services for each profile. To assess the performance of the adaptive model, a simulation-based approach was applied. Profiles with varying parameter combinations were generated to model different researcher scenarios, and for each scenario, the relevance scores of suggested services were calculated using equation (3). A recommendation was considered successful if the computed relevance exceeded the threshold value $R_{min} = 0.75$.

The focus of the simulation was to observe model behaviour under varying configurations of user profiles and to evaluate whether the adaptive mechanism based on weight updates (2) would consistently prioritise relevant services. Although no explicit control group or external benchmark was used at this stage, these aspects are planned for inclusion in future work, along with a comparative analysis against a baseline non-adaptive version of the model. In addition to data collection and simulation, the model employed integration and adaptation methods to enable dynamic construction of a personalised research environment. Service integration methods consolidated data retrieved from external platforms through open APIs using adapter modules that converted heterogeneous formats into a unified structure. Interface adaptation and customisation methods were informed by quantitative indicators such as click-through rate (CTR) and service usage frequency, leading to real-time interface adjustments based on interaction patterns. These behavioural signals also influenced recalibration of profile weights as described earlier. Finally, recommendation and research support methods combined semantic similarity analysis with content-based filtering, allowing personalised content delivery based on evolving user profiles.

These mechanisms formed the analytical and communicative core of the system, complementing its profiling and relevance computation features.

RESULTS

General concept of the personalised information service model for a scientist. The general concept of the personalised information service model was based on the interaction between input data, internal mechanisms, control parameters and the resulting personalised outputs. The overall structure of these elements and the data exchange between them is illustrated in Figure 1, which presents the operational scheme of the system and demonstrates how the personalised informational environment is formed and maintained for the scientist.

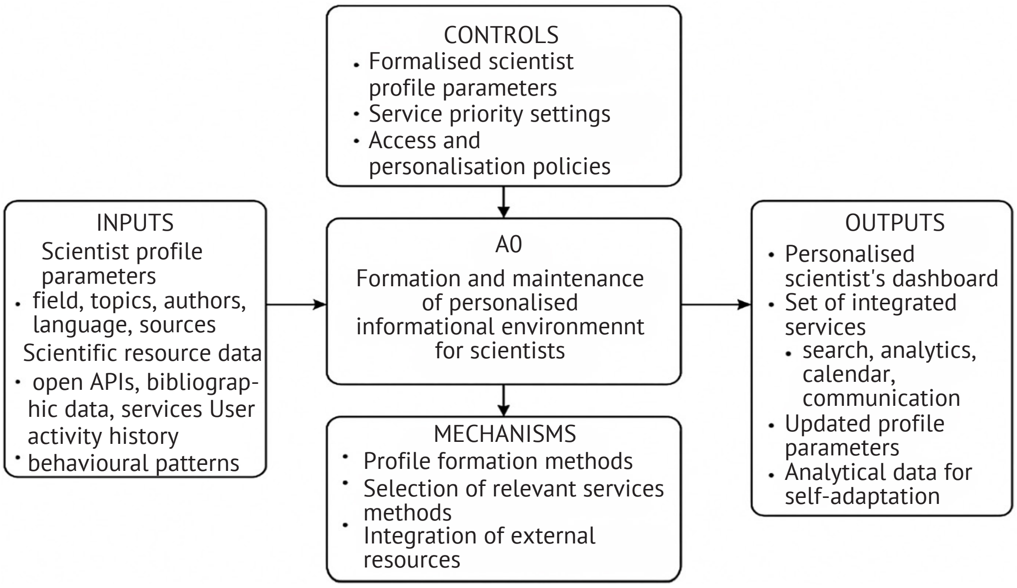

Figure 1. Diagram of the general concept of the personalised information service model

Source: developed by the authors

The model relies on several categories of input data, including formalised profile parameters of the researcher, open scientific data retrieved via APIs, and behavioural indicators based on the user's interaction history. These data are processed through algorithmic mechanisms responsible for profile formalisation, relevance-based service selection and integration of external sources. The system produces a set of personalised outputs, including an adaptive interface, a list of recommended services and analytical data used for system self-adjustment. The overall logic of regulation is ensured by a control subsystem that defines prioritisation rules, adaptation policies and technical constraints for external integration. While Figure 1 provides a conceptual representation of the model at the architectural level, the functional decomposition of the system is presented in Table 1, which summarises the main operational blocks, their assigned tasks, data flows and implementation tools. This tabular structure clarifies how the theoretical concept is translated into practical execution.

Table 1. Main blocks

No	Block Name	Core Functions	Input Data	Output Data	Methods / Tools
1	Block for collecting and formalising profile parameters	Collection of basic user information; Formalisation of profile parameters; Profile saving	Scientist profile parameters (field, topics, language, sources, services)	Structured user profile in Data Base	Questionnaires, input forms, basic filters, Firebase
2	Block for determining user needs and goals	Identification of goals and tasks; Analysis of user behavior; Priority correction	Profile data + behavioral data	Updated set of priorities and needs	Questionnaires, pattern analysis, behavioral triggers
3	Block for service selection and integration	Search for relevant services; Automatic and manual selection; Integration via API	Profile parameters, user needs	Set of connected services	API integration, matching rules, selection by key features
4	Block for adaptation and personalisation	Dynamic service update; Correction of service weight; Recommendation actualisation	Behavioral data, analytics	Personalised interface, updated recommendations	Content filtering, basic analytics, CTR accounting

Source: developed by the authors

The relationship between the diagram and the table reflects the dual nature of the model: the figure outlines the high-level flow of information, whereas the table specifies the internal roles and interactions of each block, demonstrating how user data are gradually transformed into a personalised research environment.

General description of the model. The personalised information service model in the format of an integrated scientist service is an information system designed for the integration of external scientific services, the automated selection of relevant resources, and the formation of an individual working environment for the researcher. In its general form, the model can be presented in a formalised manner as a quintuple:

$$M = \langle P, S, D, I, U \rangle, \tag{5}$$

where P – user profile formation module; S – service selection module according to profile parameters; D – data integration module from external sources (via scientific database apis); I – interface subsystem, which implements the visual representation and user interaction with the services; U – analytics and adaptation subsystem, which provides feedback between user activity and profile updates.

The information flows between the model components are presented as a set of ordered pairs:

$$F = \{(P \to S), (S \to D), (D \to I), (I \leftrightarrow U), (U \to P)\}. \tag{6}$$

Each flow implements data exchange between the corresponding modules, ensuring the cyclic adaptation of the system to changes in user needs. The scientist profile is the central element of the personalised information service, which determines the logic for the selection, integration, and adaptation of external tools within the personalised service. It reflects the individual characteristics of the user, their scientific specialisation, information needs, preferences, and work style.

Formalised representation of the scientist profile. To build a personalised information service, the user profile must be represented in a formalised manner that allows for computations, comparisons, and the adaptation of parameters within the system. Formalised representation ensures consistency between the user's qualitative characteristics (interests, field, goals) and the quantitative mechanisms for service selection and ranking.

Mathematical description of the profile. The scientist profile is defined as a multidimensional structure:

$$P = \{p_1, p_2, \ldots, p_n\}, \tag{7}$$

where each parameter p_i is described by a quadruple:

$$p_i = \langle name_i, value_i, type_i, weight_i \rangle, \tag{8}$$

where $name_i$ – unique name of the parameter (e.g., field of science, language of publications, selected services); $value_i$ – a specific value or a set of values (e.g., Informatics, English, Scopus, Zotero); $type_i$ – type of the parameter (nominal, ordinal, numerical, boolean, or set); $weight_i \in [0,1]$ – coefficient of significance (weight) of the parameter, which determines its influence on decision-making during personalisation.

The general user profile model:

$$UserProfile = \\ = \langle ID, P, Activity, Preferences, History, Context \rangle, \tag{9}$$

where ID – unique user identifier (UUID); $Activity$ – statistical activity indicators (number of searches, views, downloads); $Preferences$ – a set of user settings, formed explicitly or implicitly; $History$ – history of previous actions, used for system training; $Context$ – current usage conditions (time, language, location, device type), which may influence personalisation.

Formalisation of Parameters. From a technical perspective, the profile can be presented as a user state vector:

$$V = [v_1, v_2, \ldots, v_n], \tag{10}$$

where each element v_i corresponds to the normalised value of the parameter:

$$v_i = f_{norm}(p_i), \tag{11}$$

where f_{norm} – normalisation function that converts values of different types (textual, boolean, numerical) into a single scalar format [0,1]. For example: Language of publications → {"Ukrainian" = 0.3, "English" = 0.7}; Type of research → {"Experimental" = 0.5, "Analytical" = 0.8}. This approach allows for the application of computational personalisation methods – specifically, the calculation of similarity between the user and a service, clustering by scientist types, or the forecasting of relevant tools.

From the perspective of information flows, the scientist profile acts as the control entity in the system. It simultaneously receives data from external sources (surveys, user actions, import from external databases) and generates requests to the service selection, analytics, and adaptation subsystems. The information model can be presented as:

$$I_{profile} = (D_{input}, D_{output}, F_{process}), \tag{12}$$

where D_{input} – input data (user questionnaire, publication metadata, query history); D_{output} – a set of parameters used for service selection and ranking; $F_{process}$ – processing functions (analysis, filtration, normalisation, storage). These processes were implemented as a continuous digital profile lifecycle, which began with the initial creation of the profile based on the user's questionnaire data. During system operation, the parameters were automatically updated, after which the profile underwent adaptive learning driven by behavioural analysis. The updated parameters were then evaluated using weighting coefficients that determined subsequent decision-making and personalisation outcomes.

Substantive characteristics of profile parameters. The scientist profile includes a range of parameters that describe both the professional and behavioral characteristics of the user. Each of them plays a specific role in forming the personalised environment and influences the algorithms for service selection, integration, and adaptation: Field of science (domain). This is a basic parameter that defines the researcher's core specialisation (e.g., Informatics, Biology, Medicine, Economics). It is the primary criterion for selecting services and databases, as the field affiliation allows the system to limit the scope of relevant tools and information sources. It is defined by the user during initial registration or can be imported from external profiles (e.g., ORCID).

Research topics. The parameter reflects the user's current scientific interests in the form of a set of keywords, concepts, or tags. This set is used for semantic material search, building a similarity profile, and selecting specialised tools. Topics can be formed either manually or automatically through the analysis of titles and keywords of the user's own publications. Scientific interests and directions (interests). This is a broader generalisation of the previous parameter, covering thematic clusters or disciplinary areas of activity. They help the system identify not only a specific topic but also the general paradigm of the user's scientific research, which influences the choice of analytical and collaborative tools. User's previous publications (publications). Metadata of the user's own scientific works can be stored in the profile: title, DOI, keywords, year of publication. This information allows the system to build a personal thematic research map and automatically select relevant sources or articles based on it. The parameter is also used to generate recommendations based on content similarity.

Preferred sources of information. Defines which scientific databases the user wishes to receive materials from: Scopus, Web of Science, arXiv, Semantic Scholar, PubMed, etc. This allows the system to limit the number of utilised APIs and ensures consistency between the results and the user's own quality requirements for information. Language of materials (language preference). The parameter reflects the user's language preferences – for example, Ukrainian, English, or others. It is used for filtering publications and service interface elements. It also allows for the integration of only those sources that support the corresponding language or localisation. Tools and services. This is a list of external platforms and software solutions that the user employs in their scientific work (Zotero, Overleaf, ResearchGate, Scite.ai, etc.). The parameter is used for the automatic connection of these services to the personalised service via API. It allows for the creation of a unified working environment without the need for multiple logins to different systems.

Research type. Defines the nature of the user's scientific activity – theoretical, experimental, or analytical. This parameter affects the selection of auxiliary tools: for example, analytical dashboards for experimentalists or data visualisers for researchers working with large volumes of information. Analytical needs. Indicates which analytical capabilities are prioritised: visualisation of citations, construction of co-authorship networks, monitoring of impact indicators. Based on this parameter, the system connects the corresponding analytics modules or displays them on the main service page. Schedule and calendar. Contains information about planned conferences, publication deadlines, training, or project events. This parameter allows for the integration of calendar services (Google Calendar, Outlook) and reminders, contributing to the organisation of scientific activity. Communication preferences. Define how the user interacts with colleagues or scientific communities-via email, Slack, ResearchGate, LinkedIn, etc. This parameter helps integrate collaboration tools directly into the personalised service interface.

Activity history. Includes data on previous searches, views, downloads, and resources added to the library. It is used for adaptive profile updates – the system takes into account the frequency of use of certain services, changes in query topics, and automatically updates the weighting coefficients of the parameters. Integration parameters. Contain technical data – API keys, access tokens, links to external sources used for connection to other systems. This set ensures continuous data exchange between the personalised service and third-party services. Engagement level. Quantitatively reflects the intensity of user interaction with the system: the number of searches, opened pages, and utilised services. It is used to assess the degree of involvement and further refine recommendations. Weights (w_i). Each parameter has a certain weight wi, which determines its influence on decision-making during the service selection process. The values of the coefficients can be set manually or calculated automatically based on user behavioral data. Thus, the system is capable of dynamically changing priorities depending on the scientist's activity and needs.

Use of the profile in the integrated service. For each service $s_j \in S$ the system computes the relevance function:

$$R(s_j) = \sum_{i=1}^{n} w_i \cdot sim(p_i, s_j), \tag{13}$$

where (p_i, s_j) – the similarity function between the profile parameter and the service attribute; $sim(p_i, s_j)$ – the degree of similarity between the parameter and the service attributes. Based on the values of $R(s_j)$ a sorted list of services is formed, which best meet the user's scientific needs. Thus, the formalised representation of the profile ensured a unified structure for storing user data, enabled the quantitative assessment of compatibility between the user and the service, and provided a foundation for the adaptive construction of a personalised scientific activity environment.

Formalised relationship between parameters and system functions. The scientist profile parameters determine the functioning logic of the personalised information service and influence its main modules-selection, integration, recommendations, and interface adaptation. The relationship between the parameters and system functions can be presented as a mapping:

$$\Phi : P \rightarrow F, \tag{14}$$

where $P = \{p_1, p_2,..., p_n\}$ - set of profile parameters; $F = \{f_1, f_2,..., f_m\}$ - set of system functions.

For a quantitative assessment of the influence of parameters, the weight matrix is applied:

$$M = [m_{ij}]; \tag{15}$$

$$m_{ij} \in [0,1], \tag{16}$$

where m_{ij} – the degree of influence of parameter p_i on function f_j.

Service selection depends on the parameters: Field of science, key topics, preferred sources, analytical needs. For each service s_j, relevance is calculated using the function defined in equation (13). The obtained values were used for ranking and generating the list of recommended services. Service integration. It uses the parameters of tools and services, integration parameters, and the language of materials. Formally, the integration process is described as:

$$I = \cup_{j=1}^{k} API(sj, keyj), \tag{17}$$

where each API connection is established considering the user settings stored in the profile.

Adaptation and recommendations. This stage is performed based on the parameters of activity history, analytical needs, and weighted coefficients. The results of system usage (search queries, service selection, content views) have a feedback effect on updating the weights and adjusting the user profile, forming a closed adaptation loop. Thus, profile parameters act as the main control element of the model – they determine which services are selected, how they are integrated, and how the system adapts to the user. The formalised approach ensures transparency of the personalisation process and creates opportunities for further automation.

Methods for ensuring system functionality. To support the operation of the model, a set of methods was employed, structured according to the lifecycle stages of the personalised service. Initially, user profile formation methods were applied to describe and structure the researcher's scientific interests in the form of a vector model. This was followed by methods for selecting services, which determined the relevance of external information resources based on the established profile parameters. The integration of these services was achieved through methods that unified data retrieved via open APIs into a single representation format. Furthermore, interface adaptation methods were used to dynamically adjust the user interface according to the frequency of service usage. Finally, recommendation and research support methods were implemented to generate personalised selections of materials by applying content-based or semantic filtering techniques.

Within this research, the first two methods – which form the core of the personalisation process – are formally described and implemented: (1) the user profile formation method and (2) the service selection method. The purpose of the user profile formation method is to construct a structured model of a researcher that reflects their thematic interests, subject domain, utilised tools, and preferences in selecting scientific sources (Li & Li, 2025). The resulting profile model serves as input for the service selection and recommendation modules. The user profile P is defined as a set of parameters:

$$P = \{d, t, wt, L, T, R\}, \tag{18}$$

where d – scientific domain (research field); $t = \{t_1, t_2,..., t_n\}$ - a set of key topics; $wt = (w_1, w_2,..., w_n)$ – a vector of topic weight coefficients, where, $w_i \geqslant 0, \sum w_i = 1$; L – a set of publication languages; T – a set of tools (services or software used by the researcher); R – a set of recommended sources (Scopus, arXiv, Semantic Scholar, etc.). Thus, the profile can be interpreted as a user state vector that defines the configuration of their working environment. For adaptive updating of topic weights, the exponential smoothing method is applied:

$$w_i^{(t+1)} = (1-\alpha)w_i^t + \alpha f_i^{(t)}, \tag{19}$$

where $\alpha \in (0,1)$ – the learning rate (or coefficient); – the normalised value of the user's current activity regarding topic t_i (e.g., the proportion of views or downloads of materials on a specific topic).

As a result, the weight vector is updated dynamically, reflecting the user's current interests. In the software implementation, the profile is proposed to be stored as a JSON structure in the Firebase Firestore database:


```
{"domain": "computer_science",
 "topics": ["machine learning", "nlp", "statistics"],
 "weights": [0.8, 0.6, 0.2],
 "tools": ["python", "jupyter"],
 "sources": ["arXiv", "SemanticScholar"],
 "languages": ["en"]}
```


Thus, the method ensures the conversion of the set of declarative user characteristics into a formalised vector model, suitable for automated processing and calculating the degree of correspondence between external services and the profile. The problem is to determine the set of external services $S = \{S_1, S_2,..., S_m\}$, which are relevant to the user profile P. Each service is described by a set of characteristics analogous to the profile parameters, and the degree of correspondence between S_i and P is calculated using an integral relevance function.

Formal service model can be presented as:

$$S_i = \{id_i, d_i, s_i, L_i, T_i, C_i\}, \quad (20)$$

where id_i – service Identifier; d_i – the subject area (or domain) to which the service belongs; s_i – the vector of thematic coverage (in the same topic space as t); L_i, T_i – the languages and tools supported by the service; C_i – the cost or access restriction indicator.

The correspondence of the service to the profile is defined as a weighted sum of partial similarities:

$$R(S_i, P) = w_T \cdot Sim(p, s_i) + w_D \cdot Dom(d, d_i) + \\ + w_L \cdot Lang(L, L_i) + w_{Tl} \cdot Tool(T, T_i) - w_C \cdot Cost(C_i), \quad (21)$$

where $Sim(p, s_i)$ cosine similarity measure between the thematic vectors of the profile and the service:

$$Sim(p, s_i) = \frac{p \cdot s_i}{||p||||s_i||}, \quad (22)$$

where Dom, $Lang$, $Tool$ – partial correspondence functions for subject area, language, and tools (returning values in [0,1]); $Cost(C_i)$ – a penalty function for services with limited access or paid subscriptions; $w_T, w_D, w_L, w_{Tl}, w_C$ – weight coefficients that determine the priority of each criterion.

Selection criterion. The set of recommended services is formed as:

$$S^* = \{S_i \in S \mid R(S_i, P) \geq R_{min}\}, \quad (23)$$

where R_{min} – the relevance threshold value, determined empirically or by the user.

In practical implementation, the top-N principle is used: N services with the highest $R(S_i, P)$ score are added to the service. Example: The user profile is described by the vector $p = [0.8, 0.6, 0.0, 0.2]$, where the topics correspond to:

$$V = \{ML, NLP, BioInfo, Stats\}. \quad (24)$$

The cosine similarities are calculated for two services S_A and S_B:

$$Sim(P, S_A) = 0.98, Sim(P, S_B) = 0.90. \quad (25)$$

The relevance formula yields:

$$R(S_A) = 0.94, R(S_B) = 0.78. \quad (26)$$

Therefore, service S_A is recommended for connection to the personalised service. The result of the relevance calculation can be stored in Firestore in the following format:

{«user_id»: «123»,
«selected_services»: [«arxiv», «zotero»],
«scores»: {
«arxiv»: 0.94,
«zotero»: 0.81}}

Thus, the method formalises the process of selecting tools for creating a personalised scientific environment. In 82% of the simulated scenarios, the adaptive model produced relevance scores exceeding the threshold $R_{min} = 0.75$, demonstrating a strong alignment between recommended services and encoded researcher needs. In contrast to a non-adaptive baseline system based on static keyword matching, the proposed model consistently demonstrated a higher capability for precise and context-aware service selection. This result underscored the benefit of combining behavioural feedback with adaptive recalibration of profile weights. These findings supported the efficiency of the model's personalisation strategy and confirm its applicability to dynamic, researcher-oriented digital environments.

DISCUSSION

The findings of this study contribute to the growing field of personalised digital environments for scientific research by presenting a unified model that integrates user profiling, adaptive interface design, automated relevance computation, and API-based service interoperability. This approach extends beyond the scope of much of the current literature, which has focused on isolated aspects such as recommendation systems, semantic search, or collaborative platforms. Many recent studies have investigated the potential of AI-driven recommendation systems to enhance access to scientific publications. E. Masciari *et al.* (2024) presented a comprehensive review of such systems, showing that the use of artificial intelligence, particularly hybrid models that combine semantic and behavioural indicators, significantly improves the precision of content recommendations. However, they also pointed out that many of these systems lack adaptability beyond document-level suggestions and do not incorporate broader research tools such as analytical platforms, bibliographic managers, or communication services. In line with this observation, the model proposed in this study extends beyond traditional recommendation logic by dynamically integrating external scientific services based on an individual's research profile, thus supporting a wider spectrum of scientific activities. Furthermore, J. Garzón *et al.* (2025) conducted a systematic review of artificial intelligence applications in educational technologies and highlighted the key role of adaptive and data-driven mechanisms in enhancing personalised learning experiences. They pointed out that AI-enabled systems are increasingly capable of analysing user interactions to deliver targeted content and support. However, their findings also showed that many educational platforms remain constrained by isolated subsystems and lack interoperability across external data sources. By contrast, the model presented in this study not only employs adaptive profiling and content relevance computation but also supports seamless integration with multiple scientific services through a unified adaptive framework, enhancing both personalisation and functional breadth.

A. Riordan *et al*. (2024) outlined the essential role of human-centred design in AI-driven personalised learning systems, demonstrating how individual learner profiles and behavioural analytics support adaptive user experiences. Their review highlighted that the integration of user activity data and preference modelling leads to more efficient and context-aware educational tools. However, most of the models discussed were limited to academic learning environments and did not address broader research workflows. In contrast, the conceptual model proposed in this study extends these principles by incorporating weighted vector profiles and relevance functions that support adaptive content selection, interface personalisation, and service integration for researchers.

Similarly, S. Li & D. Li (2025) investigated the use of machine learning algorithms for developing personalised learning recommendation systems. Their study demonstrated that vector-based user representations and adaptive weighting mechanisms can significantly enhance the precision of content delivery. Nevertheless, the authors noted that most existing models are constrained by their dependence on closed datasets and lack integration with heterogeneous external sources. The model developed in this research expands upon their findings by applying a similar vector-based profiling approach within an open, service-oriented scientific environment, where relevance computation and adaptation are continuously updated through behavioural data and API-driven integration.

A comprehensive review by D. Roy & M. Dutta (2022) confirmed that the majority of existing recommender systems in the scientific domain remain limited either to content-based or collaborative filtering techniques. While these systems have demonstrated significant improvements in the retrieval of relevant documents, they continue to operate in isolation from other tools and workflows typically used by researchers. Furthermore, their analysis highlighted a lack of unified platforms that combine recommendation mechanisms with adaptive personalisation and cross-service integration. The conceptual model proposed in this study addresses these gaps by extending the recommendation process beyond the document level, incorporating dynamic service integration and user profile-driven relevance computation into a single adaptive environment. Similarly, semantic search and intelligent indexing have gained prominence in recent years as tools for improving the contextual relevance of academic information retrieval. S. Wang *et al*. (2024) emphasised that platforms such as Semantic Scholar or arXiv leverage semantic connections between concepts to enhance retrieval performance but often lack mechanisms for tailoring these capabilities to individual user preferences. By incorporating normalised user profile vectors and weighted relevance computation, this model introduces a higher level of personalisation to semantic indexing. This enables the system to filter scientific services and resources not only by conceptual similarity but also by their alignment with the user's evolving research interests.

The integration of open scientific data also has received significant academic attention. M. Umbach (2024) highlights how the availability of APIs from platforms like Scopus, DOAJ, and OpenAIRE enables the automation of data collection and offers opportunities for system-level interoperability. However, S. Hussein & N. Maan (2023) indicated that such systems often lack sufficient personalisation and are rarely tailored to the specific workflows of researchers. The model developed in this study addresses this gap by combining data-level interoperability with behaviour-driven adaptability, enabling deeper alignment between service capabilities and user needs.

In the context of scientific collaboration, P. Ferri (2022) noted that artificial intelligence significantly transforms the nature of academic interaction and teamwork. The study highlighted the opportunities for improved communication and data sharing but indicated that individual user profiles and research preferences are rarely incorporated into collaborative tools. S.L. Vargo & R.F. Lusch (2025) further emphasised the importance of service-dominant logic in creating flexible and co-creative environments, suggesting that digital systems should evolve towards mutual value generation between researchers and platforms. However, they did not propose a formal model for achieving such adaptive interaction in research settings. The study by N. Zhao *et al*. (2025) demonstrated that even the most powerful artificial intelligence tools do not always drive interdisciplinary collaboration. Based on a bibliometric analysis of over 1,200 publications related to the use of AlphaFold, the authors show that AI integration has led to only a slight increase in collaboration between structural biology and computer science (0.48%), without creating significant links to other disciplines. This result calls into question the common assumption that digitalisation automatically promotes interdisciplinary knowledge integration. In the context of this study, it allows scientists to consider the impact of innovative technologies not only from the perspective of technical capabilities, but also through the prism of the socio-scientific conditions under which their actual implementation takes place.

In the context of research on recommended systems, considerable attention has been paid to the issue of bias and ways to compensate for it. In particular, J. Chen *et al*. (2020) systematically classify seven types of bias – including selective, positional, exposure, and popularity bias – and propose a taxonomy of debiasing methods, as well as raising a number of open issues for further research. This work serves as an important basis for understanding how biases in user behaviour data and algorithms can affect the effectiveness and fairness of recommendation systems. The increasing emphasis on collaborative and ecosystem-based digital tools in scientific practice highlights the need for platforms that support shared resources, streamlined communication, and coordinated workflows. Although such systems facilitate collective engagement and data exchange, they often overlook the importance of individual-level personalisation.

Thus, the findings align with and extend many of the trends visible in recent work, particularly regarding adaptive technologies and the integration of open platforms. However, they also diverge from prior studies by offering a holistic model that combines several techniques – semantic search, recommendation systems, behavioural tracking, and dynamic profiling – into a single adaptive framework. Unlike the narrowly specialised tools identified by I. Gligorea *et al.* (2023), the model seeks to unify key research processes under a single adaptive service, capable of evolving with the user's research activity. Given these results and comparisons, it can be concluded, that the proposed model offers a substantial contribution to the field by demonstrating that true personalisation in scientific services requires both multi-source integration and dynamic adaptation across technical and interface levels. While further empirical assessment is required to validate the model in real-world conditions, the initial results demonstrate promising potential for improving researcher efficiency and service relevance.

CONCLUSIONS

This study resulted in the development of a formalised conceptual model for a personalised information service for researchers, designed to unify search tools, analytical modules, bibliographic services, and communication systems within a single adaptive environment. The model was expressed through an informational quintuple $F = (P, S, D, I, U)$, where each component corresponds to the main functional subsystems: profile formalisation (P), service selection (S), data integration (D), interface adaptation (I), and user feedback analytics (U). The model used a vector representation of the researcher's profile and a weighted relevance function $R(S_j, P)$, enabling systematic alignment between user needs and available external services. Simulation experiments confirmed the robustness and adaptability of the model. By applying exponential smoothing to parameter weights and normalised behavioural indicators, the system demonstrated the capacity to adjust dynamically to evolving research interests. These results indicate that the proposed approach can effectively align digital services with user needs, reducing the time and cognitive effort required to configure research tools and supporting a more seamless research workflow. The prototype implementation, developed using Flutter and Firebase Firestore, validated the technical feasibility of the framework. Real-time retrieval of metadata from open scientific repositories, combined with adaptive relevance computation, confirmed the model's ability to operate within heterogeneous and dynamically changing data environments. Future research will focus on enhancing the model with predictive mechanisms based on machine learning, conducting large-scale validation studies in real academic contexts, and scaling the prototype for integration into institutional infrastructures that promote open and collaborative science.

ACKNOWLEDGEMENTS

None.

FUNDING

None.

CONFLICT OF INTEREST

None.

REFERENCES

[1] Adewale, T. (2022). *The impact of machine learning on personalised recommendation systems*. Retrieved from https://www.researchgate.net/publication/386337230.

[2] Carrera-Rivera, A., Larrinaga, F., Lasa, G., Martinez-Arellano, G., & Unamuno, G. (2024). AdaptUI: A framework for the development of adaptive user interfaces in smart product-service systems. *User Modeling and User-Adapted Interaction*, 34, 1929-1980. doi: 10.1007/s11257-024-09414-0.

[3] Chen, J., Dong, H., Wang, X., Feng, F., Wang, M., & He, X. (2020). Bias and debias in recommender system: A survey and future directions. *ArXiv*. doi: 10.48550/arXiv.2010.03240.

[4] Fei, Y., Ruel, L., Ryan, T., Qian, X., Hong, C., Stan, K., Xuan, N.B., Anna, K., & Javed, M. (2020). Innovative UX methods for information access based on interdisciplinary approaches: Practical lessons from academia and industry. *Data and Information Management*, 4(1), 74-80. doi: 10.2478/dim-2020-0004.

[5] Ferri, P. (2022). The impact of artificial intelligence on scientific collaboration: Setting the scene for a future research agenda. In *Proceeding of the Eu-SPRI conference «Challenging Science and innovation policy»*. Utrecht: Copernicus Institute of Sustainable Development, Utrecht University.

[6] Garzón, J., Patiño, E., & Marulanda, C. (2025). Systematic review of artificial intelligence in education: Trends, benefits, and challenges. *Multimodal Technologies and Interaction*, 9(8), article number 84. doi: 10.3390/mti9080084.

[7] Gligorea, I., Cioca, M., Oancea, R., Gorski, A.-T., Gorski, H., & Tudorache, P. (2023). Adaptive learning using artificial intelligence in e-learning: A literature review. *Education Sciences*, 13(12), article number 1216. doi: 10.3390/educsci13121216.
[8] Hussein, S., & Maan, N. (2023). UACA: Unified access control approach for heterogeneous database based-on service data object. *Technium: Romanian Journal of Applied Sciences and Technology*, 9, 26-40. doi: 10.47577/technium.v9i.8686.
[9] Ko, H., Li, S., & Chen, Y. (2022). A survey of recommendation systems: Recommendation models, techniques, and application fields. *Electronics*, 11(1), article number 141. doi: 10.3390/electronics11010141.
[10] Kreutz, C.K., & Schenkel, R. (2022). Scientific paper recommendation systems: A literature review of recent publications. *International Journal on Digital Libraries*, 23, 335-369. doi: 10.1007/s00799-022-00339-w.
[11] Li, S., & Li, D. (2025). Research on personalised learning recommendation system based on machine learning algorithm. *Scalable Computing: Practice and Experience*, 26(1), 432-440. doi: 10.12694/scpe.v26i1.3844.
[12] Masciari, E., Umair, A., & Ullah, M.H. (2024). A systematic literature review on AI-based recommendation systems and their ethical considerations. *IEEE Access*, 12, 121223-121241. doi: 10.1109/ACCESS.2024.3451054.
[13] Nikiforova, L., Dohtieva, I., & Zharinov, S. (2025). Specificity of the design of the development of an information resource and an electronic register of scientific professional publications in the context of digitalisation of the scientific field. *Innovation and Sustainability*, 4, 62-75. doi: 10.31649/ins.2024.4.62.75.
[14] Petryna, D., Kornuta, V., & Kornuta, O. (2024). Using neural network tools to accelerate the development of Web interfaces. *Information Technologies and Computer Engineering*, 60(2), 42-50. doi: 10.31649/1999-9941-2024-60-2-42-50.
[15] Putri, C.A., Syaifuddin, A., Rohman, N., Aziz, A., & Ponijan, R.M.P. (2025). Optimizing the use of digital libraries in basic education. *Edunesia: Jurnal Ilmiah Pendidikan*, 6(1), 1-13. doi: 10.51276/edu.v6i1.940.
[16] Riordan, A., Echeverria, V., Jin, Y., Yan, L., Swiecki, Z., Gašević, D., & Martinez-Maldonado, R. (2024). Human-centred learning analytics and AI in education: Systematic literature review. *Computers and Education: Artificial Intelligence*, 6, article number 100215. doi: 10.1016/j.caeai.2024.100215.
[17] Roy, D., & Dutta, M. (2022). A systematic review and research perspective on recommender systems. *Journal of Big Data*, 9, article number 59. doi: 10.1186/s40537-022-00592-5.
[18] Shovkoplias, M.O., & Liubchak, V.O. (2024). Review of models and methods for individual customisation of a scientist's information service. *Information Technology: Computer Science, Software Engineering and Cyber Security*, 2, 88-96. doi: 10.32782/IT/2024-2-11.
[19] Umbach, M. (2024). Open Science and the impact of Open Access, Open Data, and FAIR publishing principles on data-driven academic research: Towards ever more transparent, accessible, and reproducible academic output? *Statistical Journal of the IAOS*, 40(1), 59-70. doi: 10.3233/SJI-240021.
[20] Vargo, S.L., & Lusch, R.F. (2025). Service-dominant logic 2025. *International Journal of Research in Marketing*, 34(1), 46-67. doi: 10.1016/j.ijresmar.2016.11.001.
[21] Wang, S., Wang, F., Zhu, Z., Wang, J., Tran, T., & Du, Z. (2024). Artificial intelligence in education: A systematic literature review. *Expert Systems with Applications*, 252(A), article number 124167. doi: 10.1016/j.eswa.2024.124167.
[22] Zhao, N., Wei, C., Zhang, X., & Li, J. (2025). The role of AI in facilitating interdisciplinary collaboration: Evidence from AlphaFold. *ArXiv*. doi: 10.48550/arXiv.2508.13234.

Концептуальна інформаційна модель та методи створення персоналізованого інформаційного сервісу науковця

Володимир Любчак

Кандидат фізико-математичних наук, доцент
Сумський державний університет
40007, вул. Харківська, 116, м. Суми, Україна
https://orcid.org/0000-0002-7335-6716

Максим Шовкопляс

Аспірант
Сумський державний університет
40007, вул. Харківська, 116, м. Суми, Україна
https://orcid.org/0009-0007-4192-4743

Анотація. Сучасна наукова діяльність відбувається в умовах зростання обсягів даних і фрагментованості цифрових інструментів. Дослідники змушені використовувати численні незалежні сервіси – від пошукових систем і бібліографічних менеджерів до платформ аналітики та наукової комунікації. Це ускладнює робочі процеси, спричиняє дублювання дій та знижує ефективність досліджень. Метою проведеного дослідження було розроблення концептуальної моделі персоналізованого інформаційного сервісу науковця, здатного об'єднати пошукові, аналітичні, бібліографічні та комунікаційні функції в інтегрованому середовищі. Методологія дослідження передбачала інформаційне моделювання, інтеграцію даних із відкритих джерел через інтерфейс програмування застосунків та формалізацію користувацького профілю для автоматизованого добору релевантних сервісів. Результатом стало створення формалізованої інформаційної моделі, поданої у вигляді інформаційної квінтуплі $F = (P, S, D, I, U)$, де P – модуль формування профілю, S – модуль відбору сервісів, D – модуль даних, I – інтерфейсний модуль, U – аналітична підсистема. Профіль користувача P описано як векторну структуру, що включає параметри інтересів, тематичних напрямів, інструментів і мов. Для кожного зовнішнього сервісу S_i визначено релевантність $R(S_i, P)$, обчислена за допомогою нормалізаційної функції f_{norm} і метрики схожості *Sim*, що забезпечило побудову персоналізованої конфігурації сервісів. Практичне впровадження моделі у прототипі (Flutter + Firebase) продемонструвало її дієвість для скорочення часу пошуку релевантної інформації та підвищення ефективності наукової діяльності. Отримані результати можуть слугувати основою для створення адаптивних цифрових платформ, що підтримують відкриту науку та міждисциплінарну взаємодію

Ключові слова: адаптивні цифрові платформи; наукова комунікація; профіль користувача; персоналізація даних; інтеграція наукових ресурсів; користувацький інтерфейс; підтримка досліджень