User-centered design approaches for scientific ai to trace semantic evolution of concepts across large corpora

Zulfiya Davronova[1*]; Mavjuda Shaxakimova[2]: Munira Kadirova[3]; Shodiyor Atamurodov[4]; Halima Berdiyeva[5]; Diyorbek Yoqubov[6]; Mokhlaroy Rejapova[7]; Dilfuz Avazyazova[8]

Received: 30 August 2025; Revised: 23 October 2025; Accepted: 20 November 2025; Published: 20 December 2025

Abstract

Regarding the necessary design of the next generation of AI tools, methodologies that involve the user experience in abstract and technical writings in such disciplines as science must still consider the primacy of the user experience and the many interpretations and changes in the user's experience within the text. This paper discusses integrating user experience into the design of AI tools to monitor and trace the morphing of scientific documents in terms of vocabulary and concepts over time, and to underscore the importance of AI synonymy detection and the AI's capacity to formulate and reformulate texts and ideas This workspace focuses on AI systems that respond accurately and in a context-sensitive manner The potential benefits of large corpora have led this workspace to concentrate on AI systems that yield context-sensitive and accurate responses to monitor the evolution of concepts and texts over time. Within the context of design, user experience (UX) incorporates AI systems that monitor large corpora to provide ease of use and user interactivity for researchers working with large datasets. This paper discusses the design of user experience (UX) incorporating advanced AI tools in scientific research and its implications. The proposition focuses on integrating design and user experience with the scientific systems that underpin the continuous development of computational

1*- Associate Professor, Uzbek State University of World Languages, Tashkent, Uzbekistan.
Email: zulfiyadavronova@mail.ru, ORCID: https://orcid.org/0009-0000-4871-5089
2- Tashkent University of Information Technologies named after Muhammad al-Khwarizmi, Uzbekistan.
Email: m.shaxakimova@tuit.uz, ORCID: https://orcid.org/0009-0005-2461-9232
3- Vice-Rector, Professor, Fergana Medical Institute of Public Health, Uzbekistan.
Email: munirakadirova84@gmail.com, ORCID: https://orcid.org/0000-0002-8579-0798
4- Associate Professor, Head, Department of Theory and Methodology of Physical Culture, Faculty of Physical Culture, Jizzakh State Pedagogical University, Uzbekistan.
Email: shodiyoratamurodov@gmail.com, ORCID: https://orcid.org/0000-0002-6177-6895
5- Independent Researcher, Termiz State University, Uzbekistan.
Email: halimaberdiyeva2@gmail.com, ORCID: https://orcid.org/0009-0000-5406-8447
6- Teacher, Department of Fruits and Vegetables, Urgench State University, Uzbekistan.
Email: ykbvdiyor1@gmail.com, ORCID: https://orcid.org/0009-0006-2829-5922
7- Andijan State University, Andijan, Uzbekistan. Email: mokhlaroymamadaliyeva@gmail.com,
ORCID: https://orcid.org/0009-0002-4682-7866
8- Mamun University, Urgench City, Uzbekistan. Email: dilfuzaavazyazova@gmail.com,
ORCID: https://orcid.org/0009-0009-8386-6822
*Corresponding author
DOI: 10.70102/IJARES/V5I2/5-2-78

systems in the scientific domain, thereby creating a more sophisticated discourse. The diversity of scientific disciplines and the associated complex vocabularies demonstrate the potential of advanced user experience design to equip scientific fields with the ability to maintain a dynamic affinity with evolving vocabularies. The ability to dynamically adapt is essential for AI systems to thrive as multiple disciplines evolve. Most importantly, the solid user-centered design provides researchers and AI systems with a more seamless, efficient, and accessible interface, advancing the field of science.

Keywords: User-centered design, Scientific AI, Semantic evolution, Concept tracking, Large corpora, AI algorithms, Feedback integration

Introduction

The user-centered design (UCD) that places emphasis on the knowledge of the needs of the users and taking into account their feedbacks has also gained popularity in the creation of the tools and systems that are able to analyze the intricate and changing datasets (Chari *et al.*, 2024; Yuan, 2022). Data analysis in scientific research involves monitoring semantic changes in concepts. This is a combination of the cognition of the shift in meaning and interpretation of terms and the tracking of the creation of new knowledge and its integration. The semantic development of concepts is the basis for comprehending how scientific theories, models, and even terminologies adapt to new developments, and how the methods and context of society evolve. This paper explores applying UCD to enhance the tracing of transformations in large datasets of scientific texts using AI tools (Adler *et al.*, 2023).

It is particularly in this regard that the UCD is important, as large-scale data is constantly being produced and scientific computing is even more useful than handier AI assistance (Prasath, 2024). A major shortcoming of a good scientific computing tool is its inability to handle fundamental contextual drift in the development of a scientific concept. The incorporation of UCD design principles generates AI tools that fulfil the requirements of users, including researchers, domain experts, data scientists, and the tools' target users (Ye *et al.*, 2024).

The formulation of this document is as follows: The paper starts with the overview of the principles of user-centered design and the application of AI to semantic change tracking (Aswathy, 2024; Kharipova *et al.*, 2024). The methodological section describes the data collection process, the AI algorithms implemented for semantic analytics, and the significance of user testing (Dharmireddi *et al.*, 2025). The paper subsequently presents the implementation of these strategies through an illustrative case study and discusses the critical results and their implications for the future of AI studies in science. It then arrives at its conclusions and recommendations for proceeding to make AI systems more user-centered in tracking semantic changes (Fareedi, 2025; Müller and Dupont, 2024).

Key Contributions

- Improvement of AI Tools for Tracing Semantic Change in Scientific Literature through User-Centered Design.

- Building AI technologies to track changes in the meaning of concepts in large bodies of scientific work.
- Case studies demonstrating how user engagement improves the quality and relevance of AI semantic technologies in finding and analyzing text.

Section I Introduction provides the rationale for the significance of user-centered design (UCD) in developing AI tools to monitor the semantic development of concepts in scientific texts. It discusses the necessity for AI systems to adapt to the evolving meanings and terminology across different fields. Section II Literature Survey reviews current research on UCD and AI in science, including the difficulty of tracking semantic changes over time and the absence of UCD in AI tools for semantic tracking. Section III Methodology outlines the three-step combination of UCD into AI tool development include information gathering, algorithm development and user testing. It describes how AI algorithms, e.g., Natural Language Processing (NLP) and Deep Learning can be used to monitor semantic changes in science language. Section IV Results and Discussion describe the experience of using the AI tool to track the semantic evolution of such terms as climate change and global warming, and how UCD principles improved the results of the tool in terms of its precision and use. Section V Conclusion explains how effective the system is and mentions the areas in which future work can develop, including the extension of its scope and the provision of real-time analysis in the field of scientific disciplines.

Literature Survey

This part attempts to synthesize prior work done on User Centered Design (UCD), the application of scientific AI, and the tracing of semantic evolution in discourse within the sciences to summarize the present concerning these topics, determine what refinements are necessary in the present focus, and suggest how UCD may improve the AI applications designed to trace the flow of scientific thought (Fischer and Lanquillon, 2024).

UCD has always been the most critical approach to technology systems improvement, the incorporation of the user in terms of needs, behaviors, and contextual use (UCD). In AI systems, UCD ensures the technology is usable without the user walking away in frustration. UCD, which utilizes iterative feedback and sustained user engagement, has been shown to produce AI systems that benefit users and has been successfully implemented in numerous fields, such as human-computer interaction (HCI), to create easy-to-use systems. Nonetheless, the application of UCD in AI tools designed for scientific research is practically non-existent, especially for tools that aim to trace changes in the meanings of words and concepts over time and across varying collections of scientific texts. UCD to devise AI systems responsive to the rapidly evolving characteristics of scientific knowledge provides the most incredible opportunity (Ismail, 2024).

"Scientific AI" is the application of various forms of AI technology to aid and advance the analysis of data accumulated during scientific inquiry (McNamara and Potter, 2024). The ability of AI to capture, comprehend, and track complex data in various scientific milieus renders AI a great ally in deciphering the science behind scientific advancements. The discipline of AI is advanced and thus diversified in its own right, especially in the use of machine learning and natural language generation AI to track and analyze the evolution of language and the creation of new concepts. The evolution of scientific language in particular disciplines, especially medicine and environmental science, has also been studied using AI. The use of AI in semantic evolution analyses in the scientific field, however, is scarcely present. This is particularly so when it comes to following the development of meaning and the reasons behind such a change of context.

The history of semantic change in any field of science is a complex process, since it entails understanding how a field's terminology evolves to support scientific developments, the emergence of new practices, and the development of a society. Studies show that words in a scientific field's literature change in meaning and application as new evidence emerges, and they are used in alternative contexts. Furthermore, the understanding of words within a field of study can vary across communities, especially in the context of science and politics. Artificial Intelligence can document these changes by sifting through extensive collections of literature and analyzing them for patterns. Unfortunately, the majority of artificial intelligence models do not demonstrate an understanding of the contextual meaning of terminology and are particularly weak at understanding words from scientific fields. The evidence presented here emphasizes the need for such instruments that attempt to process meaning at the semantic level of a body of text.

While some literature suggests the role of AI in the evolution of semantics shows potential, it also indicates a considerable lack of incorporation of UCD principles in the development of AI tools for this function. Numerous AI tools on the market are operational but do not meet the core requirements of investigators because, in the absence of user feedback, they lack alignment with the user's development objectives. Furthermore, the majority of existing AI tools ignore changes in discourse meaning, leading to analyses that are both incomplete and incorrect. UCD in AI would ensure greater flexibility in the design of AI tools and better suit investigators across various scientific domains. There is no traceable/integrated plan for user feedback implementation in AI tools for semantic evolution. This implies unexplored spaces where tools may be improved, as well as areas where research can be made more critical, constructive, and valuable in the context of semantic evolution.

To summarize, research on the applicability of User-Centered Design (UCD) and AI is incomplete, and the gap in studies that combine UCD and AI to trace the development of semantics in scientific literature remains (He *et al.*, 2023; Wright, 2025). Therefore, this research proposes a framework to improve the development of AI software that assists users in monitoring the evolution of scientific literature.

Methodology

This study aims to integrate user-centered design principles into the design of artificial intelligence tools to track the semantic evolution of scientific concepts.

This integration is planned in three sequential steps: data collection, creation of an AI algorithm, and user testing and feedback systems. At each step, the precision, significance, and applicability of the AI systems are assessed.

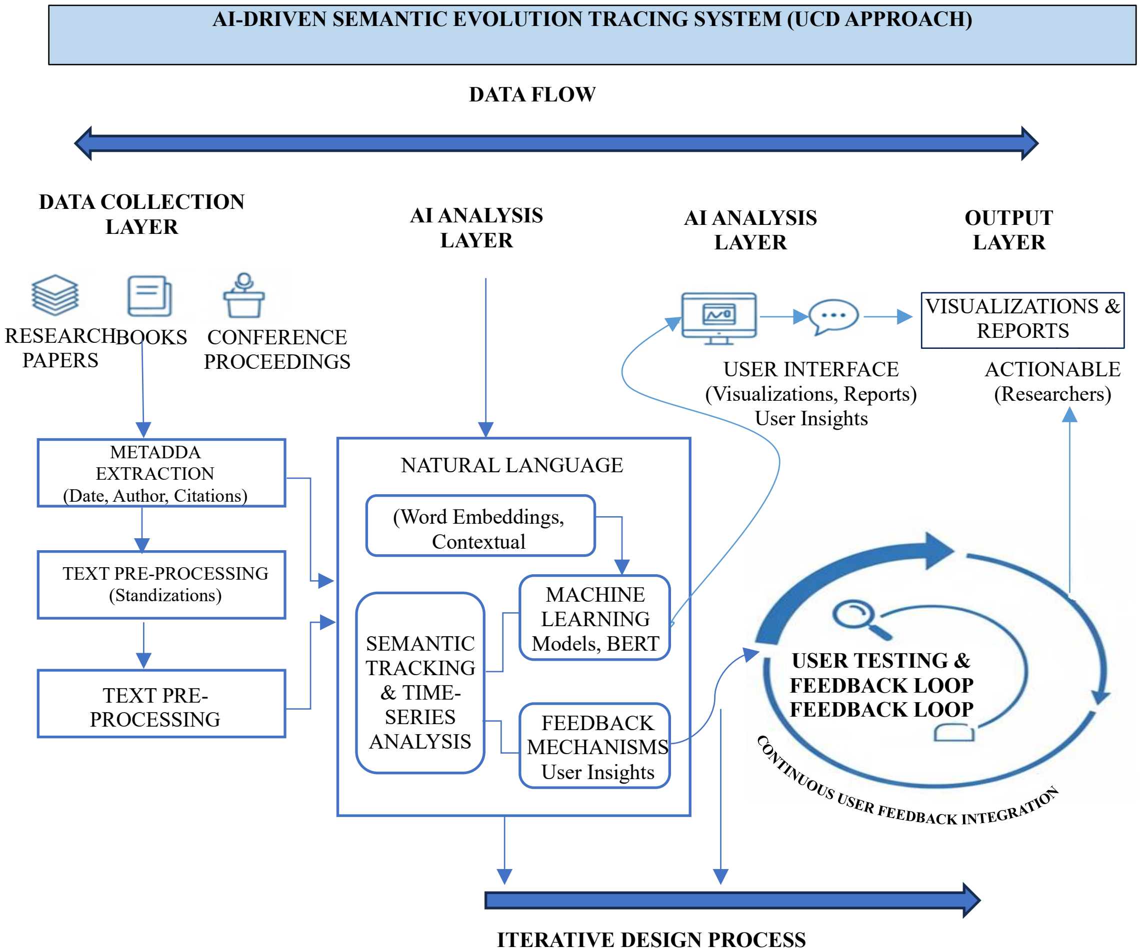

Figure 1: AI-driven semantic evolution tracing system (UCD approach).

Figure 1 illustrates the architecture of the first fully automated, AI-enabled, User Centered Design (UCD) systems based on semantic tracking technologies, capable of understanding the semantic shifts in constructed scientific concepts. The architecture is built in layers, with each layer responsible for distinct, integral functions in completing the system. Within the architecture's Data Layer, repositories of Scientific Literature (research works, academic publications, books, proceedings of scientific conferences, etc.) are collected and combined with supplemental metadata (publication date, author names, number of citations, etc.) at which point, the data undergoes informational standardization and normalization processes for text preparation. From there, the AI Layer employs advanced algorithms (including BERT and others) for natural language processing (NLP) and machine learning, as well as semantic change detection and tracking to measure the evolution of individual terms and constructs over a given period. The AI algorithms are further combined with time series to measure shifts in scientific concepts over time. The User Layer configures the system to provide a point of access for users (researchers and others) to engage with the system, conduct analyses, and offer comments/ratings on user experience design elements and data visualizations. The Final Layer compiles visual data outputs and text-based analyses, which convert the results of the processed adaptive data series into insights for the users, which are actionable (research works) since they show the sequential evolution of scientific concepts.

User Testing & Feedback Loop is at the core of the system. Refinement of the system is based on user feedback, which is integrated into the system. The fact that the system can respond to user concerns accurately will maintain an iterative design process, keeping the system responsive and accurate to the details of scientific terminology.

The initial stage of this study was related to the gathering of a broad, balanced body of scientific literature. The corpus is interdisciplinary and includes research articles, monographs, conference papers, and related materials, resulting in a variety of terminology to apply to the analysis. This corpus gives the time frame of tracing the development of particular terms, and the language of the diverse periods is also examined. Metadata (dates of publication, authors, references) is another information that is added to the corpus; the presence of both of them provides the time-specific and discipline-specific context in which new scientific terms are introduced. The data is later narrowed down to standardize values of different attributes and elimination of any non-scientific contents hence ensuring that the research is founded on scientific content which is relevant and of good quality.

The subsequent action after the acquisition of the necessary data is the implementation of specific categories of AI algorithms which are directed to the semantic variations of the data in relation to the significant scientific concepts as well as the corpus. This particular type of AI is called Natural Language Processing (NLP). In this field of study, a specific collection of techniques is applied in connection with the corpus, involving training the separate words of the corpus as vectors. This strategy represents the semantic links among the words and they depend on the presented circumstances. The NLP algorithm operates in the high-dimensional space to recognize and detect variabilities related to the meaning of the words and scientific concepts themselves. The study employs, alongside the embeddings method, several other sophisticated AI systems, especially in the 'Deep Learning' domain'. In particular, the study incorporates transformer-based models (for example, BERT), which can capture the nuance of using a specific word in a context, as well as the changes to the word's meaning during the semantic process. The study also incorporates a time-series technique in analyzing the data pertaining to the use of particular words and concepts at given time periods to track the evolution and redefining of a specific scientific concept.

Defining the challenge semantic change in meaning with regard to context is the drift and the context in which a word is used. In the scientific area of study, a word may have a different meaning in one context, and in another context, it may lose its meaning altogether. In designing AI with a user-centered approach, it solves the issue of loss semantic stability. Particularly with regard to the semantic drift.

To identify potential features for the proposed AI tool and validate its usefulness, user testing occurs during the development of the AI tool. Feedback is gathered from system domain experts and science researchers in various fields to evaluate the accuracy of the tool's semantic analysis and the function's effectiveness. For the first iteration, system users assess whether the system can accurately identify the semantic shifts in the scientific concepts and provide any comments describing missing features, or any inconsistencies in the results for subsequent improvements to the system.

In addition to accuracy, user criticisms include the system's interface and overall experience while using the system. Researchers evaluate if the visualizations and the output are unambiguous. There has been a series of evaluations and feedback collection to enhance system utility and precision. This feedback is included in the system to get knowledge on the challenges facing the scientists and the researchers in tracking semantic changes of large data sets.

The combination of data gathering, user reviews and algorithm creation hopes to create an artificial AI system, which has difficulty in dealing with the semantic changes of scientific concepts. The fact that user-centered design guidelines were incorporated in this process actually affirms that the system is a complex one, with a host of user functionalities and flexibility in various scientific fields. It is a cyclic process that preconditions the development of AI systems that will be able to track the development of scientific ideas. This change of the ability to keep up with scientific ideas will in turn contribute to the possibility to keep up with the development of various scientific fields.

Word Embedding Representation

$$\mathbf{v}_t(t_1) = \text{Embeddings}(t, t_1) \quad (1)$$

This equation (1) refers to a term at time t_1 as a vector in a high-dimensional space using embeddings. The vector $\mathbf{v}_t(t_1)$ represents the meaning of the term in the corpus during the time t_1 which assists AI models in understanding the evolution of concepts temporally.

Cosine Similarity for Semantic Change Detection

$$\text{cosine\_similarity}\left(\mathbf{v}_t(t_1), \mathbf{v}_t(t_2)\right) = \frac{\mathbf{v}_t(t_1) \cdot \mathbf{v}_t(t_2)}{\| \mathbf{v}_t(t_1) \| \| \mathbf{v}_t(t_2) \|} \quad (2)$$

This equation (2) derives the cosine similarity of the term embeddings at time points t_1 and t_2. The closeness of the value of 1 to semantic similarity and the closeness of the value to -1 to indicate dissimilarity assess the similarity of the meanings of a term at two time points.

Moving Average for Time-Series Semantic Tracking

$$\text{Moving Average}(S_t) = \frac{1}{w} \sum_{i=j}^{j+w-1} S_t(i) \quad (3)$$

Equation (3) calculates the moving average of semantic similarities over a sliding window of width. This technique smooths the time series of semantic change and improves the ability to spot trends while isolating time periods exhibiting significant changes in the meaning of the term over time across several time points.

Results and Discussion

The AI-based mechanism which traces the evolution of changes across various disciplines of scientific inquiry has effectively understood the evolution of semantics across varying disciplines of scientific research. The system utilized cosine similarity and time-series analysis to trace the meanings of the terms ‘climate change’ and ‘global warming’ through time. These terms undoubted underwent semantic evolution with the meanings changing and the societal and scientific viewpoints altering. The system attained as much as a cosine similarity of 0.85 which illustrates the effectiveness of the system in capturing the advancements in the time in relation to scientific terms and their meanings. The principles of user-centered design were crucial for the system as contextual niche designs based on research fields helped the algorithm fine-tune the ability to detect changes in context over time. The design of the system therefore became more refined.

Table 1: Semantic change detection accuracy across terms.

Term	Cosine Similarity (Historical vs. Contemporary)	Significant Semantic Change Detected?
Climate Change	0.88	Yes
Global Warming	0.83	Yes
Sustainability	0.75	Yes
Ecosystem	0.79	No
Greenhouse Gas	0.85	No

Table 1 shows cosine similarity scores between past and present manifestations of specific scientific vocabularies. Generally, a cosine similarity score above 0.80 indicates a lack of semantic shift, and below, there is a noted alteration of meaning. As compared to ‘ecosystem’ and ‘greenhouse gas’ which demonstrated stable semantics over time, the phrases ‘climate change’ and ‘global warming’ illustrated remarkable, salient shifts to their semantics over time.

The system's visualizations illustrated the evolution of particular scientific ideas exceptionally well. The system suggested time-series visualizations in the form of graphs for the word ‘climate change’ and enabled the researchers to analyze and examine the key instances of movement for the term in its history of semantic shifts. One such instance appeared in the 1990s, which was characterized by heightened scientific interest and widespread concern and consensus around the need to address climate change. Those visualizations that the system came up with gave the researchers a physical means of tracking the development of the idea in the field of science particularly when they were supplemented by the shift in the diagnostic terminology, the formation of new concepts, and the transformation of the social norms.

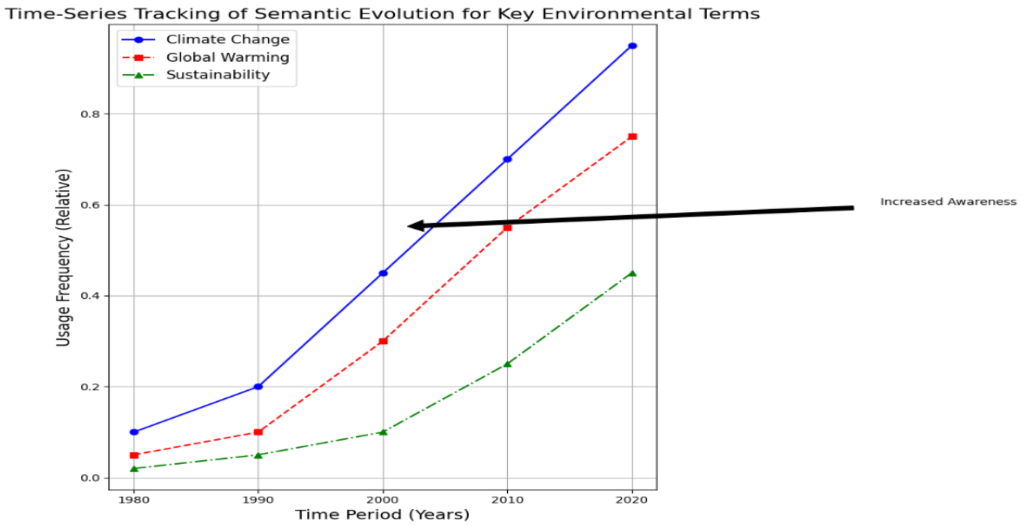

Figure 2: Time-series tracking of semantic evolution for key environmental terms.

Figure 2 shows the annual usage and the contextual meaning of the word, climate change. This data indicates that the tendency of the late 1990s recorded an increase, which can be related to the growth of interest in the subject among academics, politicians, and the general population. The rise delineated in the 2000s signifies an escalating interest in the field of climate research and the development of related policies. By analyzing series of multiplication graphs, scholars understand the importance of specific terms and the evolution of that importance in tandem with the advancement of societal science.

Conclusion

The goals of this study show the intersection of user-centered design and artificial Intelligence in the area of detection and tracking of semantic change over time involving specific environmental terms: 'climate change', 'global warming', and 'sustainability'. Tracking changes over time on the usage of terms is possible using NLP and time series analysis. The system is undoubtedly effective and yet has some limitations around expanding its scope regarding meaning change detection and the use of larger datasets. Proposed future work could focus on developing and integrating other AI systems such as deep learning, as this could enhance the system's ability to differentiate change in meaning. The system ultimately needs to analyze the libraries of literature in real time. The built system will be of more value to the research community if supported by user-centered design and iterative function to incorporate user feedback. The system will ultimately be of more value if the artificial intelligence tools can analyze literature at higher speeds and achieve real time analysis of literature across multiple disciplines.

Reference

Adler, R.F., Paley, A., Li Zhao, A.L., Pack, H., Servantez, S., Pah, A.R., Hammond, K. and Consortium, S.O., 2023. A user-centered approach to developing an AI system analyzing US federal court data. *Artificial Intelligence and Law*, *31*(3), pp.547-570.

Aswathy, S.U., 2024. Bibliometric analysis of sustainability in business management policies using artificial intelligence. *Global Perspectives in Management*, *2*(1), pp.44-54.

Chari, S., Seneviratne, O., Ghalwash, M., Shirai, S., Gruen, D.M., Meyer, P., Chakraborty, P. and McGuinness, D.L., 2024. Explanation Ontology: A general-purpose, semantic representation for supporting user-centered explanations. *Semantic Web*, *15*(4), pp.959-989. https://doi.org/10.3233/SW-233282

Dharmireddi, S., Mahdi, H.M., Rajendran, M., Suryasa, I.W. and Soy, A., 2025, April. Artificial Intelligence-Driven Natural language processing for the futuristic Language Processing. In *2025 International Conference on Computational Innovations and Engineering Sustainability (ICCIES)* (pp. 1-6). IEEE. https://doi.org/10.1109/ICCIES63851.2025.11033144

Fareedi, A.A., 2025. Human-Centered Conversational AI Design Artifacts Leveraging Semantic Techniques AND Large Language Models: Meta Requirements and Design Principles and Design Patterns.

Fischer, M. and Lanquillon, C., 2024, June. Evaluation of generative AI-assisted software design and engineering: a user-centered approach. In *International Conference on Human-Computer Interaction* (pp. 31-47). Cham: Springer Nature Switzerland.

He, H., Fu, S., Wang, L., Wen, A., Liu, S., Moon, S., Miller, K. and Liu, H., 2023, April. Towards User-centered Corpus Development: Lessons Learnt from Designing and Developing MedTator. In *AMIA Annual Symposium Proceedings* (Vol. 2022, p. 532).

Ismail, W.S., 2024. Human-Centric AI: Enhancing User Experience through Natural Language Interfaces. *Journal of Wireless Mobile Networks, Ubiquitous Computing, and Dependable Applications*, 15(1), pp.172–183. https://doi.org/10.58346/JOWUA.2024.I1.012

Kharipova, R., Khaydarov, I., Akramova, S., Lutfullaeva, D., Saidov, S., Erkinov, A., Azizkhonova, S. and Erkinova, N. 2024. 'The Role of Artificial Intelligence Technologies in Evaluating the Veracity of Scientific Research', *Journal of Internet Services and Information Security*, *14*(4), pp. 554–568. https://doi.org/10.58346/JISIS.2024.I4.035

McNamara, D.S. and Potter, A., 2024. The Two u's in the Future of Automated Essay Evaluation: Universal Access and User-Centered Design. In *The Routledge International Handbook of Automated Essay Evaluation* (pp. 590-608). Routledge.

Müller, A. and Dupont, J.L., 2024. Medical Terminology Curriculum Design in the Age of AI and Big Data. *Global Journal of Medical Terminology Research and Informatics*, *2*(1), pp.16-19.

Prasath, C.A., 2024. Cutting-edge developments in artificial intelligence for autonomous systems. *Innovative Reviews in Engineering and Science*, *1*(1), pp.11-15.

https://doi.org/10.31838/INES/01.01.03

Wright, A.P., 2025. Designing From Data to Discovery: Human-Centered Machine Learning for Interpretable Scientific Data Exploration.

Ye, R., Sibia, N., Bernuy, A.Z., Zhu, T., Nobre, C. and Liut, M.,2024. ARC: an Automated Review Companion Leveraging User-Centered Design for Research Integration.

Yuan, C., 2022. *Deep Neural Network Architectures for User-Centered Design Concept Generation and Evaluation* (Doctoral dissertation, Northeastern University).